一种基于人工免疫的网络安全实时风险检测方法

王益丰, 李 涛, 胡晓勤, 宋 程
(四川大学计算机系, 四川成都 610065)

摘 要: 依据人体免疫系统中抗体浓度的变化与人体生病的严重程度的对应关系, 本文提出了一种基于人工免疫的网络安全实时风险检测的方法, 给出了自体、非自体、抗体和抗原实现方式, 建立了抗体的克隆选择、学习机制以及生命周期模型, 提出了基于抗体浓度的网络安全风险的定量计算模型. 实验表明该方法具有定量计算、实时性等优点, 是网络安全实时风险检测的一个较好的解决方案.

关键词: 人工免疫; 网络入侵检测; 风险评估
中图分类号: TP393. 08 文献标识码: A 文章编号: 0372-2112 (2005) 05-0945-05

A Real-Time Method of Risk Evaluation Based on Artificial Immune System for Network Security

WANG Yi-feng, LI Tao, HU Xiao-qin, SONG Chen
(*Dept. of Computer Science, Sichuan University, Chengdu, Sichuan 610065, china*)

Abstract: With the relationship between the antibody concentration and the degree of illness in the human immune system, a real-time method of risk evaluation for network security based on artificial immune system is proposed. Self, nonself, antibody and antigen are defined. Thus clone selection, learning scheme and lifecycle of antibody are built. A quantitative computation model based on the concentration of the antibodies for computer network security has thus been presented. The experiment shows that this method has the features of quantitative calculation and real-time process ability, and is a good solution for real-time risk evaluation for network security.

Key words: artificial immune; network intrusion detection; risk evaluation

1 引言

目前网络安全风险检测的方法主要有两类: 静态评估和实时检测. 大部分静态评估方法利用一些评估标准来评估, 例如 TESEC[1]、ITSEC[2]、CC[3]等; 也可以利用一些网络安全专家 (例如 Tiger Team[4]) 对目标网络的风险进行评估, 这些人对系统潜在的安全漏洞具有很强的敏感性. 由于静态评估方法主要通过系统的静态因素来评估, 所以静态网络安全风险评估方法缺乏实时性, 只能粗略地估计网络长期所处的安全风险状态, 对网络正在遭受的攻击缺乏实时的网络安全风险检测, 从而导致现阶段网络安全处于被动防御的局面[5].

目前在国内外, 实时网络安全风险检测的研究尚处于探索阶段. E Jonsson 和 T Olovsson[6] 基于入侵检测中的一些实验数据对攻击者行为进行了分析, 可对系统的可靠性进行一定的评估. R Ortalo 等人[7]提出了一种将已知的 UNIX 安全漏洞映射为"权利图"的方法. 该模型根据估计攻击者攻破系统的已知漏洞需要花费的平均时间来评估系统的安全风险, 然而该模型只能根据系统的已知漏洞来评估, 因此具有很大的局限性. B B Madan 等人[8]提出了一种基于状态机的模型, 该模型利用 9 个安全状态来描述在已知攻击和未知攻击下系统的动态行为, 该模型强调攻击的结果, 对系统面临的整体安全缺乏实时评估的能力. 李辉等人[9]提出了一种基于贝叶斯网络和目标树的网络安全风险评估的方法, 但该方法只是根据目标系统正在遭受的攻击的种类来粗略的评估系统面临的安全风险. C K Chu 等人[10]综合静态和实时的检测方法, 提出了一种网络安全的风险评估框架. 在实时检测方面, 该方法只能以简单的突发事件 (运行条件的突然改变、组件的丢失等) 为依据, 就某些网络攻击进行简单的实时风险评估.

实时的网络安全风险检测在积极的网络安全技术中具有十分重要的意义. 然而目前提出的一些方法或模型, 由于缺乏定量的描述和可理解的物理解释, 因而在具体应用中存在很大的局限性.

计算机安全问题和人体免疫系统有许多相似之处, 近几年来, Forrest 等人[11, 12]、Dasgupta 等人[13, 14]和 Kim 等人[15~17]将人体免疫系统的一些免疫机制应用到网络安全研究中, 从而开辟了网络安全研究的一个新局面[18].

收稿日期: 2004-02-09; 修回日期: 2004-12-23
基金项目: 国家自然科学基金 (No. 60373110); 教育部博士点基金项目 (No. 20030610003)

人体免疫系统是一个由免疫活性分子、免疫细胞、免疫组织和器官组成的非常复杂的系统. 人体免疫系统的主要功能就是区分自体(对人体无害的)和非自体(对人体有害的),并且消灭非自体,这主要是通过分布在全身的不同种类的淋巴细胞(*B* 细胞、*T* 细胞等)来实现. 一旦 *B* 细胞与抗原匹配数达到一定数目并且受到 *T* 细胞的刺激,就会释放出大量的抗体来匹配抗原,这时该类抗体的浓度就会大量增加,当抗原被杀死之后抗体的释放就会受到抑制从而该类抗体的浓度会降低,使免疫系统趋于稳定. 正常情况下,人体各种抗体的浓度基本保持不变,所以可以通过测量各种类型抗体的浓度来判断人体是否生病以及生病的严重程度.

根据人体免疫系统中抗体浓度的变化与人体生病严重程度之间的对应关系,本文提出了一种基于人工免疫的网络安全实时风险检测方法,给出了网络安全环境下免疫系统中抗体和抗原的表达方式,就免疫系统的基本原理如自体耐受、克隆选择等进行了仿真,并且给出了网络安全风险的定量计算模型. 实验结果表明该方法是网络安全实时风险检测的一个较好解决方法.

2 基于免疫的风险检测方法

本文提出的基于免疫的风险检测方法模拟了人体免疫系统,人体免疫系统到该方法的映射关系如表 1 所示. 整个网络

表 1 人体免疫系统与基于免疫的风险检测方法间映射关系

人体免疫系统机理/功能	基于免疫的风险检测方法
抗原	IP 包进行特征提取后得到的二进制字符串
B 细胞, *T* 细胞和抗体	用二进制字符串表示的抗体
抗体和抗原的绑定	*r* 连续位匹配算法
自体耐受	否定选择算法
B 细胞结合的抗原数目超过亲和力阈值	超过激活阈值的匹配
细胞克隆	抗体的复制
抗体浓度的增加	系统面临的安全风险的增加
人体	网络
淋巴结	网络中的主机

相当于人体,网络中的主机模拟了人体免疫系统中的淋巴结,但它又与人体免疫系统不同:每台主机都可以独立的产生和训练抗体,而在人体免疫系统中的未成熟免疫细胞的成长统一在胸腺中完成. 网络中的每台主机完成的免疫功能如图 1 所示. 未成熟抗体通过随机生成,成熟抗体由通过自体耐受的

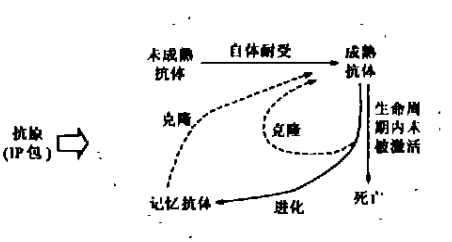

图 1 主机的免疫功能

未成熟抗体进化而来,而记忆抗体由被激活的成熟抗体进化而来. 抗原由成熟抗体和记忆抗体来检测,成熟抗体在一定的生命周期内匹配到一定的抗原就会被激活而进化为记忆抗体,同时克隆自己,否则就会死亡,而记忆抗体有无限的生命周期,并且一匹配到抗原就会克隆自己.

2.1 自体与非自体

整个系统的网络活动可以抽象为一个由长度为 *l* 的二进制字符串组成的集合 U[5], 即 $U=\{0,1\}^l$. 集合 U 分为两个子集:自体集合(正常网络活动)S,非自体集合(非法网络活动或网络攻击)T. S 和 T 满足下面两个关系: $U=S\cup T$ 和 $S\cap T=\Phi$.

2.2 抗原与抗体

定义抗原为对网络上传输的 IP 包进行特征提取(类似免疫系统中的抗原提呈)后得到的长度为 *l* 的二进制字符串,定义 $Ag(Ag\subseteq U)$ 为抗原集合.

定义抗体集合 $D=\{d\mid d=\langle s, age, count, ag\rangle, s, ag\in U, age, count\in N\}$,每个抗体为一个四元组,其中 s 为抗体的基因(长度为 *l* 的二进制字符串), age 为抗体的年龄, $count$ 为抗体的累计亲和力,即该抗体匹配到的抗原数目, ag 为检测到的抗原(被该抗体捕获到的抗原), N 为自然数集合. 抗体与抗原的匹配采用 *r* 连续位匹配算法,如公式(1)所示,其中 1 表示匹配, 0 表示不匹配, $x\in U, y\in U, r\in N$ 并且 $0\leqslant r\leqslant l$, x_k 为二进制字符串 x 的第 k 个字符, y_k 为二进制字符串 y 的第 k 个字符.

$$f_m(x,y)=\begin{cases}1, \text{ iff } \exists\, i,j, x_k=y_k\wedge k=1,\cdots, j\wedge j-1\geqslant r\wedge i,j,r,k\in N\\ 0, \text{otherwise}\end{cases} \tag{1}$$

抗体分为三类:成熟抗体、记忆抗体和未成熟抗体. 成熟抗体是指对自体耐受,但未被抗原激活(在一定的生命周期 δ 内匹配到一定数目的抗原)的抗体,定义 $M(M\subseteq D)$ 为成熟抗体集合. 记忆抗体由被激活的成熟抗体进化而来,定义 $R(R\subseteq D)$ 为记忆抗体集合. 未成熟抗体为尚未进行自体耐受的抗体,定义 $I(I\subseteq D)$ 为未成熟抗体集合.

2.3 自体耐受

为了防止抗体匹配到自体,新生成的未成熟抗体必须通过自体耐受才能与抗原匹配,对新生成的未成熟抗体 ab 可以用公式(2)进行自体耐受,其中 1 表示通过自体耐受, 0 表示没通过自体耐受, $ab\in I$. 通过自体耐受的未成熟抗体将进化为成熟抗体,而未通过自体耐受的未成熟抗体将死亡.

$$f_t(ab)=\begin{cases}0, \text{ iff } \exists\, y\in S\wedge f_m(ab.s,y)=1\\ 1, \text{otherwise}\end{cases} \tag{2}$$

2.4 克隆选择过程

成熟抗体的累计亲和力达到一定的激活阈值 θ,就会被激活成为记忆抗体,并且克隆自己,这相当于生物免疫系统的初次应答过程,这过程需经过一个亲和力累计的过程,即需要一个较长的学习期. 记忆抗体只要一匹配到抗原就会被激活而克隆自己,这相当于生物免疫系统中的二次应答过程,这过程无需学习过程,系统能对以前识别过的抗原做出迅速的反

应. 克隆选择的策略为: 与抗原具有高亲和力的将会被克隆, 即优势抗体将会被优先使用来抵抗网络攻击. 整个克隆选择过程如图2所示. $f_{num}(ab)$ 为抗体 ab 激活后被克隆的数目, 如公式3所示, 其中 σ 为集合 M 和 R 中与抗体 ab 有相同抗体基因的抗体的数目, ζ 为一常数. 集合 M 和 R 中与被激活的抗体 ab 有相同抗体基因的抗体数目越多, 则抗体 ab 的克隆将受到抑止, 克隆的数目越少. 克隆出来的抗体的累计亲和力和年龄全置为0. 抗体克隆的结果将导致系统在遭受某一攻击后, 与之对应抗体的浓度升高, 也就是说, 随着攻击强度的增加, 单位时间内抗体的浓度也随着相应提高, 从而为风险评估提供了直接依据.

$$f_{num}(ab) = \lceil \zeta(1 - \sigma/(|M| + |R|)) \rceil \tag{3}$$

```
Procedure Clone( )
Begin
  For(抗原集合 Ag 中的每个 ag) Do
    If(M 和 R 中有一抗体 ab 与 ag 匹配)Then
      Begin
        增加抗体 ab 的亲和力;
        If(ab ∈ R)Then /* 二次应答 */
          Begin
            克隆 f_num(ab) 个 ab 加入到 M 中;
            在 ab 和克隆的抗体中记录 ag;
          End;
        Else If(ab ∈ M)Then /* 初次应答 */
        If(ab.count > θ)Then
          Begin
            将 ab 进化为记忆抗体;
            克隆 f_num(ab) 个 ab 加入到 M 中;
            在 ab 和克隆的抗体中记录 ag;
          End;
      End;
End.
```

图2　克隆选择过程

2.5　免疫监视

对网络活动的检测过程是基于主机的, 其过程如图3所示. 初始过程时, 集合 M 为空. 集合 R 中的记忆抗体为保存下来的记忆抗体, 第一次运行时, 集合 R 为空集. 更新抗原集合时, 抗原集合 Ag 中的抗原由上一次迭代过程中收集到的抗原替代. 系统通过克隆选择, 利用成熟抗体的初次应答以及记忆细胞的二次应答实现对抗原集合 Ag 中的抗原的分类, 从而实现免疫监视的目的. 被记忆抗体匹配和使成熟抗体累计到足够的亲和力而被激活的抗原分类为非自体抗原(非法网络活动或网络攻击), 其余分类为自体抗原(正常网络活动). 被分类为自体的抗原中有一部分是被成熟抗体匹配但未使该抗体累计到足够的亲和力而被激活的抗原, 因此该免疫监视过程采用了容忍入侵的网络安全积极防御策略(在不能确定网络活动是否为攻击时, 先容忍它). 成熟抗体有固定的生命周期, 在生命周期内未被激活的成熟抗体就会走向死亡, 并由新生的通过自体耐受的成熟抗体取代. 而记忆抗体有无限的生命周期.

```
Procedure Surveillance( )
Begin
  初始化集合 M 和 R;
  While(1) Do
    Begin
      更新抗原集合 Ag;
      增加所有成熟抗体的年龄; /* 模拟进化一代 */
      克隆选择过程; /* 通过克隆选择实现抗原的分类, 见图 2 */
      /* 删除那些在生命周期内未被激活的抗体 */
      /* 模拟抗体的死亡机制 */
      M := M - {x | x ∈ M ∧ x.age > δ};
      /* 将通过自体耐受的未成熟抗体进化为成熟抗体 */
      M := M ∪ {x | x ∈ I ∧ f_t(1) = 1};
      /* 随机生成 ρ 个未成熟抗体并加入 I 中进行自体耐受 */
      I := I ∪ {x_1, x_2, ..., x_ρ}; /* x_i (i = 1, 2, ..., ρ) 为随机生成的抗体 */
    End
End
```

图3　免疫监视

2.6　网络安全风险评估

在免疫监视过程中, 当主机没遭受攻击时, 主机上的抗体的数目保持相对不变; 当主机遭受攻击时, 由于记忆抗体和被激活的成熟抗体的克隆, 主机上的抗体数目就会急剧增加; 当攻击退去后, 由于成熟抗体的生命周期, 主机上的抗体数目就会减少并趋于稳定状态(这些抗体因不能在其生命周期中捕获到足够的抗原而死亡). 因此, 主机以及网络的安全风险可以根据成熟抗体集合和记忆抗体集合来计算.

由于每台主机在网络中的重要性和不同类型攻击的危害性不同, 所以在计算网络或主机面临的安全风险时, 应考虑每台主机的重要性以及每类攻击的危害性. 设 β_i ($0 \leqslant \beta_i \leqslant 1$) 为主机 i 在网络中的重要性, α_j ($0 \leqslant \alpha_j \leqslant 1$) 为 j 类型攻击的危险性. 另外由于攻击的不同, 产生检测到该类型攻击的抗体的数目也将不同, 例如当遇到 DOS 攻击时, 则将产生大量检测到该攻击的抗体. 设 $\tilde{\omega}_j$ 为能检测到 j 类型攻击的抗体数目的比例系数. 定义 $n_{ij}(t)$ 表示 t 时刻主机 i 上检测到 j 类型攻击的抗体数目. 设风险指标 $0 \leqslant R(t) \leqslant 1$ 为整个网络在 t 时刻面临的风险, $0 \leqslant R_j(t) \leqslant 1$ 为整个网络在 t 时刻面临 j 类型攻击的风险, $0 \leqslant r_i(t) \leqslant 1$ 为主机 i 在 t 时刻面临的风险, $0 \leqslant r_{ij}(t) \leqslant 1$ 为主机 i 在 t 时刻面临 j 类型攻击的风险: 为0时, 表明没有风险; 为1时, 表明系统极度危险. 风险指标值越大, 表明当前系统的面临的风险也越大. $R(t)$、$R_j(t)$、$r_i(t)$ 和 $r_{ij}(t)$ 的计算方法如下:

$$R(t) = \frac{2}{1 + e^{-\sum_i \beta_i \sum_j \alpha_j \tilde{\omega}_j n_{ij}}} - 1 \tag{4}$$

$$R_j(t) = \frac{2}{1 + e^{-\alpha_j \tilde{\omega}_j \sum_i \beta_i n_{ij}(t)}} - 1 \tag{5}$$

$$r_i(t) = \frac{2}{1 + e^{-\sum_j \alpha_j \tilde{\omega}_j n_{ij}}} - 1 \tag{6}$$

$$r_{ij}(t) = \frac{2}{1 + e^{-\alpha_j \tilde{\omega}_j n_{ij}}} - 1 \tag{7}$$

主机安全风险值的计算开销主要为对抗体的分类. 抗体根据其捕获到的抗原特征进行分类. 如果抗体不属于已知类别, 则表明检测到了未知攻击, 将其归类为未知攻击类别, 等待网络管理员的确定. 如果确定检测到的不是攻击, 则将所有与该抗体有相同抗体基因的抗体从抗体集合中删除. 按最坏的情况考虑(每个抗体都将重新分类, 并检测到了未知攻击), 则主机的安全风险值(r_i 和 r_{ij})的时间复杂性为 $O(c \cdot num \cdot (|M| + |R|))$, 其中 c 为一抗体捕获到的抗原和一类攻击的特征进行比较的时间, num 为攻击的种类. 同理, 设网络中有 K 台主机, 按最坏的情况考虑, 网络的安全风险值(R 和 R_j)的时间复杂性为 $O(\max_{i=1}^{K})$, 其中 O_i 为主机 i 的安全风险值的计算复杂性. 由于网络和主机的安全风险值的计算与抗体的数目成线性关系, 其时间复杂度较小. 仿真实验同时表明(图 4), 计算出来的安全风险变化曲线和实际系统面临的攻击强度基本同步.

3 实验结果

在实验中, 抗原由源/目的 IP 地址、端口、协议类型以及 TCP/UDP/ICMP 等域组成. 成熟抗体的激活阈值 θ 取 10. 生命周期 δ 取迭代次数 100 次. 克隆时 ζ 取 5. 每一代加入的新非成熟抗体数目 ρ 为 5. 实验中对网络中的 20 台主机, 如 A、B、C 等进行监控, 主机 A、B、C 等的重要性分别设为: 0. 8、0. 4、0. 5 等. 攻击过程中对网络中主机进行 *synflood*、*land*、*teardrop* 等 20 种攻击, 设 *synflood*、*land*、*teardrop* 等攻击的危险性为 0. 8、0. 5、0. 6 等.

图 4 为主机 A 遭受 *synflood* 攻击的实际强度(每秒发送的攻击数据包的数目)变化曲线及相应的检测到的风险变化曲线. 图 5 为检测到的网络的整体安全风险变化曲线、网络遭受 *synflood* 攻击的安全风险变化曲线和网络遭受 *land* 攻击的安全风险变化曲线. 从图 4 中看出, 随着攻击强度的增加, 其相应的风险也迅速跟着上升(这是由于记忆抗体一匹配到抗原, 就迅速克隆自己). 当攻击强度下降时, 其相应的风险也降低, 但下降的斜率相对攻击强度下降的斜率要小(斜率的大小由成熟抗体的生命周期来决定的). 这在真实的网络环境下具有重要的意义: 当某一攻击在短时间内再次发生时, 网络仍可保持较高的警戒度.

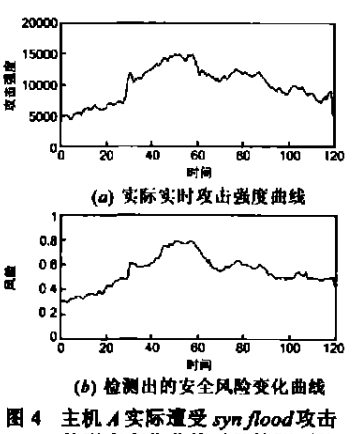

图 4 主机 A 实际遭受 *synflood* 攻击的强度变化曲线以及检测到的相应的安全风险变化曲线

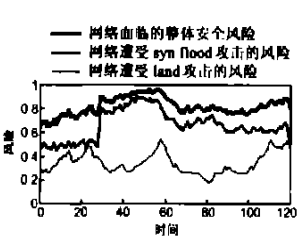

图 5 检测到的网络安全风险变化曲线

在实验中发现, 主机或网络的性能随着面临风险的增加而降低. 当主机或网络面临的风险超过 0. 8 时, 系统基本上处于瘫痪状态, 但当风险降到 0. 3 以下时, 系统工作就很正常.

目前大多数实时风险评估方法基本上从概率上来估计[6~9]; 实时评估安全风险的能力有限[6,8]; 评估的依据过于简单[9,10]; 缺乏自适应性, 只能根据已知攻击或已知漏洞来评估[7,9]; 不能有效地评估某种攻击的安全风险[6,8~10]. 与上述方法对比, 本文提出的基于免疫的网络安全风险检测方法能够实时地和定量地计算主机和网络面临的安全风险; 并同时能够有效地评估主机和网络面临某种攻击的安全风险; 而且对未知攻击具有一定的学习能力.

4 结论

本文提出的网络安全风险评估的方法能实时、定量地计算系统的整体风险以及遭受某种攻击的风险的大小, 从而给系统实时调整自己的防御措施提供了依据. 与传统的网络安全风险评估方法不同, 本文提出的风险检测方法具有分布式实时计算、定量描述、自学习、自适应等特点, 是网络安全风险检测的一种较好解决方案.

参考文献:

[1] National Computer Security Center, Dept. of Defense, no. DoD 5200. 28. STD, Trusted computer system evaluation criteria[S].

[2] ISBN 92-826-7024-4, Information technology security evaluation criteria: provisional harmonized criteria[S].

[3] ISO/IEC 15408, Common criteria for information technology security evaluation (CC v2. 1)[S].

[4] P D Goldis. Questions and answers about tiger teams[J]. EDPACS, The EDP Audit, Control and Security Newsletter, 1989, 27(4): 1– 10.

[5] 李涛. 网络安全概论[M]. 北京: 电子工业出版社, 2004.

[6] E Jonsson, T Olovsson. A quantitative model of the security intrusion process based on attacker behavior[J]. IEEE Trans on Software Engineering, 1997, 23(4): 235– 245.

[7] R Ortalo, Y Deswarte, M Kaâniche. Experimenting with quantitative evaluation tools for monitoring operational security[J]. IEEE Trans on Software Engineering, 1999, 25(5): 633– 650.

[8] B B Madan, K Gõseva-Popstojanova, K Vaidyanathan, et al. Modeling and quantification of security attributes of software systems[A]. In Proc Int Conf on Dependable Systems and Networks (IPDS stream)[C]. Washington, DC, 2002, 2: 505– 514.

[9] 李辉, 蔡忠闽, 韩崇昭, 管晓宏. 基于信息融合的入侵检测模型与方法[J]. 小型微型计算机系统, 2003, 27(9): 1602– 1606.

[10] C K Chu, M Chu. An integrated framework for the assessment of network operations, reliability, and security[J]. Bell Labs Technical Journal, 2004, 8(4): 133– 152.

[11] S A Hofmeyr and S Forrest. Immunity by design: an artificial immune system[A]. Proc of the Genetic and Evolutionary Computation Conference (GECCO)[C]. Morgarr-Kaufmann, San Francisco, 1999. 1289– 1296.

[12] S A Hofmeyr, S Forrest. Architecture for an artificial immune system[J]. Evolutionary Computation, 2000, 8(4): 443– 473.

[13] D Dasgupta. Immunity-based intrusion detection systems: a general framework[A]. Proc of the 22nd National Information Systems Security Conference (NISSC)[C], Arlington, Virginia, 1999. 147- 160.

[14] D Dasgupta, F Gonzalez. An immunity-based technique to characterize intrusions in computer networks[J]. IEEE Trans on Evolutionary Computation, 2002, 6(3): 281- 291.

[15] Jungwon Kim, Peter J Bentley. The artificial immune model for network intrusion detection[A]. the 7th European Congress on Intelligent Techniques and Soft Computing (EUFIT'99)[C]. Aachen, Germany, 1999.

[16] Jungwon Kim, Peter J Bentley. Evaluating negative selection in an artificial immune system for network intrusion detection[A]. Genetic and Evolutionary Computation Conference 2001 (GECCO-2001)[C]. San Francisco, 2001. 1330- 1337.

[17] Jungwon Kim, Peter J Bentley. Towards an artificial immune system for network intrusion detection: an investigation of dynamic clonal selection [A]. the Congress on Evolutionary Computation (CEC-2002)[C]. Honolulu, 2002. 1015- 1020.

[17] 李涛. 计算机免疫学[M]. 北京: 电子工业出版社, 2004.

作者简介:

王益丰　男, 1981 年生于江苏, 硕士研究生, 主要研究方向为网络安全、人工免疫. E-mail: Lonwind@163.net.

李　涛　男, 1965 年年生于四川, 教授, 博士生导师, 主要研究方向为网络安全, 人工智能.