基于危险理论的网络风险评估模型

彭凌西[1,2]，陈月峰[2]，刘才铭[1]，曾金全[1]，刘孙俊[1]，赵 辉[1]

(1. 四川大学计算机学院 成都 610065; 2. 广东海洋大学信息学院 广东 湛江 524088)

【摘要】为有效评估网络信息系统的网络风险，提出了一种新的基于危险理论的风险评估模型(DTREM)。在给出网络活动中自体、非自体、免疫细胞的定义，建立由记忆检测器、成熟检测器、未成熟检测器集合构成的入侵检测子模型后，进一步给出了基于危险理论的网络风险定量计算子模型。利用该模型，可以实时定量地计算出主机和网络当前所面临攻击的各类攻击和总体网络风险强度，理论分析和实验结果均表明，该模型为实时网络安全风险评估提供了一种有效的新途径。

关 键 词 人工免疫系统; 危险理论; 网络信息系统; 网络风险评估
中图分类号 T P389.1 文献标识码 A

Danger Theory Based Network Risk Evaluation Model

PENG Ling-xi[1,2]，CHEN Yue-Feng[2]，LIU Cai-ming[1]，ZENG Jin-quan[1]，LIU Sun-jun[1]，ZHAO Hui[1]

(1. School of Computer Science, Sichuan University Chengdu 610054;
2. Information School, Guangdong Ocean University Zhanjiang Guangdong 524025)

Abstract To effectively evaluate the network risk of network information system, a Danger Theory based Network Risk Evaluation Model (DTREM) is proposed. With definitions of self, non-self, and immunocyte, the intrusion detection sub-model is given. DTERM is composed of memory detectors, mature detectors, and immature detectors. Furthermore, the danger theory based network risk evaluation sub-model is given. In the proposed model, the risk of each network attack, including holistic risk of the host and network, can be calculated in real time and quantificationally. Both the theory analysis and experimental results prove that DTERM provides an effective and novel approach for network risk evaluation.

Key words artificial immune system; danger theory; network information system; network risk evaluation

实时网络安全风险评估对网络安全防御技术的研究具有重要的意义，从风险评估的自动化程度来看，风险评估方法主要包括人工评估和自动评估[1]，人工评估通常以问卷调查的方式开展，依赖专家经验，尽管评估比较全面，但容易引入主观因素，而且复杂的评估过程会导致用户面对庞大的开销；而自动评估通常采用自动识别弱点或攻击的方法对目标网络进行评估，由于该方法有自动性、高效性和易管理性等优点，针对自动评估技术的相关研究工作较多。

1 自动评估技术相关研究

文献[1]提出了用于评估网络信息系统的风险传播模型；文献[2]基于COPS提供的数据，采用权限图理论建模系统漏洞，使用马尔科夫模型计算攻击者攻破系统安全目标可能付出的平均代价，以定量度量系统安全；文献[3]提出了一种网络弱点的模型检测方法；文献[4]设计了一种拓扑弱点分析工具TVA，可以检测网络系统弱点，并以网络攻击图的形式分析网络系统的安全风险；文献[5]提出了一种基于免疫的网络安全风险检测模型，用于对网络系统面临攻击时的实时风险评估，但因计算量较大，在实际应用中受到了一定的限制；文献[6]提出了一种根据网络系统的组织结构，基于服务、主机本身的重要性的层次化网络安全威胁态势量化评估方法。这些方法都介于静态评估和实时检测之间，有效性及实时性不能满足真实网络环境的需要。总体来讲，目前有关网络安全风险评估的系统性研究比较少见，所提出的一些检测模型或方法，大多缺乏严格、全面、定量的数学模型，因而在具体应用中存在很大的局限性。

计算机安全问题与生物免疫系统(Biological

收稿日期：2007 – 09 – 06
作者简介：彭凌西(1978 –)，男，博士生，讲师，主要从事网络安全技术和人工免疫方面的研究.

Immune System，BIS)所遇到的问题具有惊人的相似性，两者都要在不断变化的环境中维持系统的稳定性。文献[7]的克隆学说，建立了“自体-非自体”模型，该模型认为各种淋巴细胞带有的受体具有特异性，特异性受体和具有高亲和力的外来抗原发生相互作用，带有该受体的淋巴细胞被激活并发生克隆扩增，从而能识别自身抗原的淋巴细胞在发育早期被清除。据此，文献[8]提出了否定选择算法(Negative Selection Algorithm，NSA)，并被广泛用于入侵检测系统。基于人工免疫(Artificial Immune System，AIS)的网络安全技术具有多样性、自适应、鲁棒性等特点，被认为是一个非常重要且有意义的研究方向[5]。针对“自体-非自体”模型无法解释的自我免疫疾病现象(如多发性硬化症疾病中，免疫系统会对某些属于“自体”的细胞产生免疫应答)，文献[9]提出了危险理论[9]。危险理论指出，免疫系统中不只是进行“自体-非自体”的识别，而是当受侵组织的“危险”积累到某种程度时，则发出危险信号，意味着免疫系统检测到了外来的入侵抗原。

为有效评估实时信息系统的网络风险，本文提出了一种基于危险理论的网络安全风险评估模型(A Danger Theory Based Risk Evaluation Model for Network Security，DTREM)，在建立网络安全环境免疫系统的入侵检测子模型后，进一步给出了基于危险理论的网络风险计算子模型。利用DTREM，可实时并定量地计算出主机和网络当前所面临攻击的类别、数量、强度及风险数值等。理论分析和实验结果表明该模型为网络安全风险评估提供了一种有效的新途径。

2　模型理论

DTREM由入侵检测子模型和风险评估子模型两部分组成，入侵检测子模型检测入侵或攻击行为；而风险评估子模型对检测到的入侵或攻击行为进行评估，并计算出子机和网络以及各类攻击的网络风险，下面分别对这两个子模型进行介绍。

2.1　入侵检测子模型

入侵检测子模型由记忆检测器检测模块，成熟检测器检测模块和未成熟检测器的耐受模块组成，它们分别对应记忆检测器集合，成熟检测器集合和未成熟检测器集合三个重要的集合。这些集合的具体定义如下。

定义论域 $D=\{0,1\}^l$，抗原集合 $\mathrm{Ag}\subset D$，自体集合 $\mathrm{Self}\subset \mathrm{Ag}$，非自体集合 $\mathrm{Nonself}\subset \mathrm{Ag}$。

$$\mathrm{Self}\cup \mathrm{Nonself}=\mathrm{Ag}$$
$$\mathrm{Self}\cap \mathrm{Nonself}=\phi$$

式中　Ag为通过从网络IP数据包中提取的IP地址、端口号、协议类型等网络事务特征的二进制表示；Self 集为正常网络服务事务，Nonself 集为来自网络的攻击。

定义集合 $s\mathrm{Ag}\subset \mathrm{Ag}$，且 $|s\mathrm{Ag}|=|\mathrm{Ag}|\eta$，其中 $\eta(0<\eta<1)$ 为检测系数，sAg 的元素为随机从 Ag 中选取而来。

定义免疫检测器集合为 $B=\{<d,\mathrm{age},\mathrm{count},s>|\ d\in D,\mathrm{age},\mathrm{count}\in N,s\in R\}$，其中 d 为抗体基因；age 为抗体年龄；count 为匹配数；s 为检测器受到攻击后的危险性；N 为自然数集合；R 为实数集合。免疫检测器由记忆检测器和成熟检测器组成，即 $B=M_\mathrm{b}\cup T_\mathrm{b}$。对于记忆检测器集有：

$$M_\mathrm{b}=\{x\mid x\in B,x\mathrm{count}\geq\beta,\ \forall y\in \mathrm{Self}<xd,yd>\notin \mathrm{Match}\}$$

式中　β 为激活数阈值。

定义成熟免疫检测器集合 T_b，它由对自体耐受且还未被激活的免疫细胞组成，有：

$$T_\mathrm{b}=\{x\mid x\in B,x\mathrm{count}<\beta,\forall y\in \mathrm{Self}<x.d,y>\notin \mathrm{Match}\}$$

未成熟检测器通过抗体基因库产生或者随机方法生成，其形式化描述如下：

$$I_\mathrm{b}=\{<d,\mathrm{age}>|\ d\in D,\mathrm{age}\in N\}$$

为对一个输入的抗原集合 Ag 分 δ 代(δ 为常数)进行训练，每代选出一定数量的抗原组成 sAg 抗原集合，通过B集合的检测把抗原集合分类为自体和非自体，整个过程分为三个阶段。(1) 第一阶段为从0时刻到一个耐受期 α 结束的时刻，需要定义初始的自体集合Self(0)和未成熟检测器集合$I_\mathrm{b}(0)$，后者经前者耐受后成为成熟检测器。(2) 第二阶段为从 α +1时刻到记忆检测器产生的时刻，为自学习阶段，成熟检测器通过克隆选择产生能识别大量不同非自体抗原的记忆检测器，通过检测被分类为自体的抗原最后送给未成熟检测器进行耐受。(3) 第三阶段为从记忆检测器产生到系统终止，免疫系统各部件产生完毕，进行实际环境中的检测；首先是记忆检测器对抗原进行检测，然后是成熟检测器对剩下抗原的检测，最后未成熟检测器以剩余抗原为自体进行耐受。耐受采用否定选择算法，具体的匹配可采用海明距离、欧氏距离以及r连续位匹配规则，即如果：

$$\exists i,\ j(x.d_i=y_i,x.d_{i+1}=y_{i+1},\cdots,x.d_j=y_j)$$
$$j-i\geq r,0<i<j\leq l$$

则x与y就匹配成功。

当记忆检测器和成熟检测器与一个抗原匹配成

功后，即检查该抗原是否属于当前 Self(t)集合。(1) 如果属于，则进行协同刺激，把该抗原交给外部系统进一步判断是自体还是非自体，并完成相应处理。若是自体，则删除这个能检测出当前自体的免疫检测器，否则删除该抗原，且删除已不是当前自体的自体元素。(2) 如果不属于，则把该抗原视为非自体。

运行过程中，外部系统可以向Self(t)添加当前认为是合法的自体，新添加的自体就会使系统产生对该自体耐受的免疫检测器。但是当更新抗原后，如果该抗原没有在一定时间内出现，则系统新产生的免疫检测器就对它不再耐受，因此系统具有很好的自适应性。如果系统中同时存在对一种抗原耐受和不耐受的两种不同的免疫检测器，它们的竞争就交给外部系统仲裁(协同刺激)。

2.2 风险评估子模型

在入侵检测子模型完成对入侵或攻击抗原进行检测的同时，主机中的记忆检测器则对检测到的入侵或者攻击抗原进行风险评估。在从时刻 $t-1$ 到 t 的单位时间内，主机中的第 j ($0\le j\le J$)个记忆检测器检测到一个入侵或攻击抗原，其网络风险按式(1)增加，同时将年龄赋值为 0。如果检测到多个抗原，则按式(1)对网络风险进行累计计算，表明攻击威胁在持续增加。式(1)中，η_1 (>0 的常数)为起始的损失程度值，η_2 (>0 的常数)模拟奖励因子(监视遭到连续的类似的网络攻击)。

$$M_{\mathrm{b},j}s(t)=\eta_1+\eta_2 M_{\mathrm{b},j}s(t-1) \tag{1}$$

反之，如果该记忆检测器在该时间间隔没有检测到入侵或攻击抗原，则危险性按式(2)衰减 $1/\lambda$，同时将其年龄增加 1。

$$M_{\mathrm{b},j}s(t)=\begin{cases} M_{\mathrm{b},j}s(t-1)\left(1-\dfrac{1}{\lambda-M_{\mathrm{b},j}\mathrm{age}(t-1)}\right) & M_{\mathrm{b},j}\mathrm{age}(t-1)<\lambda \\ 0 & M_{\mathrm{b},j}\mathrm{age}(t-1)\ge\lambda \end{cases} \tag{2}$$

可以看出，当记忆检测器的年龄增大到 λ 时，记忆检测器的危险性衰减为 0，这与实际的入侵情况基本一致，即在一定时间如果没有检测到该类攻击，则认为该类攻击已经停止。

采用著名的林肯实验室从攻击中抽取关键属性的方法对入侵检测子系统检测到的网络攻击进行分类，该基于属性的攻击分类的方法在普适性、全面性、准确性、可扩展性等方面具有较好的表现，并已经用于实际的工作中[10]。

设风险指标 $0\le r_k(t)\le 1$ 为主机 k 在 t 时刻所面临的风险，$r_k(t)=1$，表明当前系统极度危险；$r_k(t)=0$，表明当前系统没有危险。$r_k(t)$值越大，表明当前系统面临的风险越高，考虑到各种主机的资产权重以及各类攻击的危险性不一样，设定μ_i 表示该类攻击的危险性，ϖ_k 表示该主机的资产权重。

对于主机 k，t 时刻面临第 i($1\le i\le I$)类 $A_i(t)$攻击的安全风险按式(3)进行计算：

$$r_{k,i}(t)=1-\frac{1}{1+\ln\left(\mu_i\left(\sum\limits_{M_{\mathrm{b},j}\in A_i(t)}M_{\mathrm{b},j}s(t)\right)+1\right)} \tag{3}$$

式中 $1\le k\le K$，下同。

对于主机k，t时刻主机的整体安全危险按式(4)进行计算：

$$r_k(t)=1-\frac{1}{1+\ln\left(\sum\limits_{i=1}^{I}\mu_i\left(\sum\limits_{M_{\mathrm{b},j}\in A_i(t)}M_{\mathrm{b},j}s(t)\right)+1\right)} \tag{4}$$

网络面临第i($1\le i\le I$)类$A_i(t)$攻击的安全风险为 R_i，按式(5)进行计算。与文献[5]不同的是，DTREM首先通过计算各主机的风险，然后再通过主机资产权重加权的方式层层向上计算上层网络风险，因此减小了风险计算过程中的因攻击分类而产生的计算量。具体计算如下：

$$R_i(t)=1-\frac{1}{1+\ln\left(\mu_i\left(\sum\limits_{k=1}^{K}\varpi_k\left(\sum\limits_{M_{\mathrm{b},j}\in A_{k,i}(t)}M_{\mathrm{b},j}s(t)\right)\right)+1\right)} \tag{5}$$

整个网络整体安全风险按式(6)进行计算：

$$R(t)=1-\frac{1}{1+\ln\left(\sum\limits_{k=1}^{K}r_k(t)\varpi_k+1\right)} \tag{6}$$

3 实 验

为验证模型对主机和网络的风险评估程度，对模型进行了仿真实验。实验环境为100 M的局域网，通过一个C类IP地址222.18.2.*连接到Internet；服务器的操作系统为Red Hat 9.0，服务器提供WWW、Email以及FTP服务，抗原定义为定长为从IP包中提取的包含IP地址、端口号、协议类型、数据包内容等网络事务特征的128位二进制字符串。采用海明距离计算抗原和免疫检测器之间的亲和力阈值，取值为80。

考虑到一般情况下网络的正常行为变化不大，实验设定未成熟检测器耐受期 α 为1，记忆检测器危险性衰减周期设置为1。而成熟检测器的激活阈值 β 以及生命周期 λ 的取值标准为确保入侵检测子模型

取得较高的检测率(True Positive, TP)和较低误报率(False Positive，FP)[5]。实验时取β=5和λ=40，实验得到满意的结果[5]，本文对记忆检测器初始风险和奖励因子η_1和η_2，分别取值0.001 0和0.999 8。

为验证DTREM对网络风险评估的效果，对网络进行udpstorm网络攻击，并将DTREM和文献[5]中的Insre模型进行实验结果对比，Insre被证实是一种有效的网络风险评估方法，攻击强度曲线和网络风险曲线如图1所示。

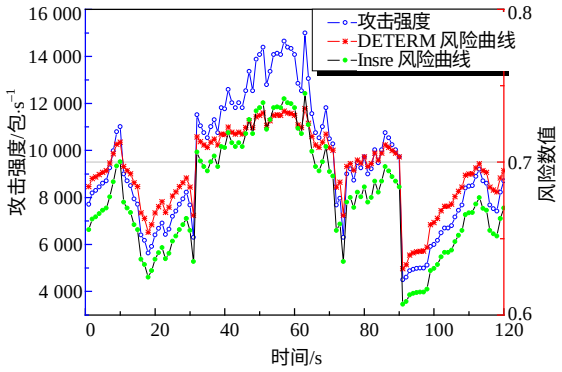

图1　udpstorm网络攻击强度曲线和网络风险曲线

从图1可看出，在时刻0～10 s之间，主机攻击强度有所增加，对应的风险曲线也呈上升的趋势，在时刻11～19 s之间，攻击强度下降，对应的风险指标呈下降的趋势。实验结果与真实网络环境的情形较为一致：当系统遭到持续高强度的网络攻击时，系统所面临的风险较高；反之，当网络攻击强度降低时，系统所面临的风险降低。

实验结果显示，DTREM和文献[5]中的Insre模型均能实时准确地计算当前网络以及网络中任意主机面临的整体综合风险，计算得出的风险值与当前网络所面临的实际攻击强度具有很好的一致性，表明模型能够很好地实时反映当前网络风险的实际变化情况。但与Insre模型相比，DTREM模型更能反映真实的网络风险情况。

4　结　论

本文提出了一种新的基于危险理论的网络风险评估模型，该模型成功地将生物免疫系统中的相关原理运用到网络安全技术中，与已有的风险评估模型比较，该模型能够实时准确地计算当前网络以及网络中主机面临的整体综合风险，以及每一种网络攻击的风险，使管理员能够对系统面临的安全威胁状况有宏观的了解，并可从网络风险态势中发现系统的安全趋势和规律，根据需求制定安全建议，避免危险事件的发生。

参　考　文　献

[1] 张永铮, 方滨兴, 迟　悦. 用于评估网络信息系统的风险传播模型[J]. 软件学报, 2007, 18(1): 137-145.

[2] ORTALO R, DESWARTE Y, KAANICHE M. Experimenting with quantitative evaluation tools for monitoring operational security[J]. IEEE Trans on Software Engineering, 1999, 25(5): 633-650.

[3] RITCHEY R, AMMANN P. Using model checking to analyze network vulnerabilities[C]//Proceedings of the IEEE Symp on Security and Privacy. Berkeley: IEEE Computer Society Press, 2000.

[4] JAJODIA S, NOEL S, OBERRY B. Topological analysis of network attack vulnerability[C]//Managing Cyber Threats: Issues, Approaches and Challenges. [S.l.]: Springer-Verlag, 2005: 248-266.

[5] LI Tao. An immunity based network security risk estimation[J]. Science in China Ser F Information Sciences, 2005, 48(5): 557-578.

[6] 陈秀真, 郑庆华, 管晓宏, 等. 层次化网络安全威胁态势量化评估方法[J]. 软件学报, 2006, 17(4): 885-89.

[7] BURNET F M. The Clonal selection theory of acquired immunity[M]. London: Cambridge University Press,1959.

[8] FORREST S, PERELSON A, CHERUKURI R. Self-nonself discrimination in a computer[C]// Proceedings of 1994 IEEE Computer Society Symposium on Research in Security and Privacy. Los Almitos: IEEE Computer Society, 1994.

[9] MATZINGER P. The danger model: a renewed sense of self[J]. Science, 2002, 296(5566): 301-305.

[10] HAINS W, LIPPMANN R, DAVID J F, et al. 1999 DARPA intrusion detection evaluation: design and procedures[R]. MIT Lincoln Laboratory, 2001.

编　辑　熊思亮