

一种基于免疫的网络安全态势感知方法

刘念^{1,2} 刘孙俊³ 刘勇⁴ 赵辉¹

(四川大学计算机学院 成都 610065)¹ (四川大学电气信息学院 成都 610065)²
(成都信息工程学院软件工程学院 成都 610225)³ (中国科学院成都计算机应用研究所 成都 610041)⁴

摘要 网络安全态势感知技术作为积极主动的防御技术,目前已成为网络安全领域新的研究方向。在分析与总结国内外网络安全态势感知技术的基础上,提出了一种基于免疫的网络安全态势感知系统。该方法采用基于免疫的入侵检测模型作为态势感知的基础,实现对网络中已知和未知入侵行为的检测;依据生物免疫系统抗体浓度的变化与病原体入侵强度的对应关系,建立网络安全态势定量评估模型,并采用灰色马尔可夫模型对网络安全态势进行预测。实验结果表明,该方法有助于及时有效地调整网络安全策略,为系统提供更全面的安全保障,是网络安全主动防御的一个较好的解决方案。

关键词 人工免疫,网络安全,网络安全态势

中图分类号 TP391 **文献标识码** A

Method of Network Security Situation Awareness Based on Artificial Immunity System

LIU Nian^{1,2} LIU Sun-jun³ LIU Yong⁴ ZHAO Hui¹

(College of Computer Science, Sichuan University, Chengdu 610065, China)¹

(School of Electrical Engineering and Information, Sichuan University, Chengdu 610065, China)²

(College of Software Engineering, Chengdu University of Information Technology, Chengdu 610225, China)³

(Chengdu Institute of Computer Application, Chinese Academy of Sciences, Chengdu 610041, China)⁴

Abstract As a positive defense technology, Network Security Situational Awareness has become the orientation of research in the field of network security. Based on the analysis of the papers from domestic and foreign on technologies for network security situational awareness, a method of network security situational awareness based on the profound research of AIS was designed and built. The method uses network intrusion detection based on the theory of immunity as the base of situational awareness, to detect known and unknown intrusions with the help of biological technology. According to correspondence relations of density change of antibody in the artificial immune systems and pathogen invasion intensity, a novel network security situational evaluation model was also established. In the tendency prediction for network security situational, this paper used Grey Markov Model to make quantitative prediction. Experiment results show that this model is also helpful to resemble network security tragedy effectively, therefore, it is a better solution for network security initiatives defense.

Keywords Artificial immunity, Network security, Network security situation

1 引言

近年来随着网络的普及,其面临的威胁也越来越大,计算机病毒、木马程序、DOS/DDOS 攻击日益猖獗。为保证网络安全运行,目前采用的入侵检测、防火墙、病毒检测^[1]等技术属于被动防御手段,只能对系统局部进行检测,获取的信息之间缺乏关联。而现有的网络态势感知技术仅能粗略地就网络过去所处的风险进行定性的描述,对系统正在遭受的攻击的风险缺乏网络安全态势实时定量的感知能力,无法依据网络环境威胁的变化主动调整整个系统的防御策略,同时缺乏自适应性,无法识别未知攻击,不能防范日益严重的网络安全威胁。

近年来,由生物引发的信息处理方法的研究引起了人们高度的重视^[2]。Burner^[3]率先提出克隆选择原理而获得诺贝尔奖;丹麦学者 Jerne^[4]提出免疫系统的第一个数学模型,奠定了免疫计算基础;Forrest^[5]等人提出否定选择算法并提出了计算机免疫概念。人工免疫理论不断发展,已成为一种动态的计算模型^[6],突破了传统的方法论和思维方式,通过免疫系统能够自己学习和判断,反映了“细胞自我学习”过程^[7]。免疫机体根据自身所面临的危险,能及时有效地作出响应,这些优秀的特性决定了它在网络安全领域具有广阔的应用前景^[8]。

本文将人工免疫技术应用于网络安全态势感知^[9]领域,

到稿日期:2009-05-03 返修日期:2009-06-09 本文受国家自然科学基金项目(60373110,60573130),国家 863 计划项目(2006AA01Z435)资助。

刘念(1973—),男,讲师,主要研究领域为网络安全、人工免疫等,E-mail:liunianis@gmail.com;刘孙俊(1975—),男,博士,主要研究领域为智能计算;刘勇(1970—),男,博士,主要研究领域为数据挖掘;赵辉(1977—),男,博士,主要研究领域为网络安全。

设计并构建了一种基于免疫的网络安全态势感知系统 NS-SASI (Network Security Situation Awareness System based Immune)^[10],旨在对网络安全态势状况进行实时监控,对恶意的网络行为变得无法控制之前,实现对网络安全态势的实时、定量感知,这有助于及时有效地调整网络安全策略,为系统提供全面的安全保障。

2 现有网络安全态势感知方法简介

目前国外提出的态势感知模型框架主要有 Edward Waltz 提出的使用数据融合技术和数据挖掘提取安全态势的方法。美国国防部提出的 JDL^[10] (Joint Director of Laboratories) 模型,通过对来自传感器、信息源的数据进行关联分析、数据组合,从而获得对战场情况和威胁程度的完整评价。从网络安全角度出发,建立的网络安全态势感知系统架构主要有 Tim Bass^[10] 提出的基于入侵检测融合的网络安全态势感知框架,以及 Jason Shifflet 采用本体论 (Ontology) 对网络安全态势感知相关概念进行了分析比较后提出的基于模块化的技术无关框架结构。Christopher 等人也采用了类似的方法,但这些研究多止步于框架结构的介绍,未见其精确数学模型。国内管晓宏教授利用入侵检测系统的日志库,结合服务、主机自身的重要性及网络系统的组织结构,提出了自下而上、先局部后整体的层次化网络安全态势定量评估模型,该方法在一定程度上完善了网络安全态势评估的内容。但提出的模型是基于统计模型来完成的,在参数的选择上没有进行标准化,而是采用专家经验选择,因而不能很好地体现网络安全态势的真实情况。

3 基于免疫的网络安全态势感知模型架构

态势感知 (Situation Awareness) 的定义为在一定的时空条件下,对环境因素的获取、理解以及对未来状态的预测。它最早出现于军事领域,是指了解己方的优势和敌方的弱点从而进行战场决策的过程。它的研究重点在于各种复杂信息的高效组织和可视化技术,目的就是缩短得到信息到做出决策之间的时间。此后在军事战场、核反应控制以及医疗应急调度等领域,态势感知被广泛地研究。由于在动态复杂的环境中,决策者需要借助态势感知工具显示当前环境的连续变化状况,才能准确地做出决策,因此态势感知越来越成为一项热门的研究课题。

网络安全态势是对网络运行状况的宏观反映,它反映了一个网络过去和当前的状况,并预测下一个阶段可能的网络状态。它首先对网络原始事件进行预处理,把具有一定相关性、反映某些网络安全事件的特征的信息提取出来,通过一系列数学方法处理,将网络安全特征信息归并融合成有意义的数值。这些数值具有表现网络运行状况的特性,随着网络安全事件发生的频率、数量以及网络受到威胁程度的不同,该数值的大小也会随之产生特征性的变化。

网络安全态势感知是指在大规模网络环境中,通过对网络攻击行为、网络服务状况及正常行为进行特征提取,获取态势信息,综合这些来自不同网络设备在时间及空间上的特征进行态势评估,准确了解网络的当前安全状况,提高对网络安全状况的认知,根据评估结果进行决策,采取及时有效的响应措施抑制当前网络所遭受的攻击,保障系统安全,缓解决策者

的认知压力。

网络安全态势感知是对动态变化的网络安全态势元素进行觉察、认识、理解和预测的过程。据此,本文提出一种基于人工免疫原理的网络安全态势感知模型,该模型主要分为 3 层,依次为网络安全态势觉察、态势理解以及态势预测,如图 1 所示。

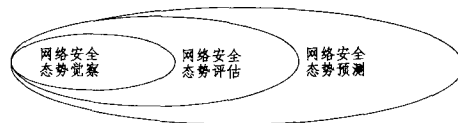


图 1 网络安全态势感知模型体系结构

NSSASI 将 IP 包抽象为抗原,将免疫细胞抽象为入侵检测系统中的检测器,将抗体抽象为相应的匹配器,当检测器发现网络入侵时(对应抗体匹配抗原),模拟人体免疫系统的克隆原理,提高相应抗体的浓度;当网络攻击减弱时,模拟免疫反馈机制,降低相应抗体的浓度,并使之恢复到正常的水平。由于当前系统中抗体的浓度与系统正在遭遇的攻击的强度、风险等具有一一对应的关系,因此,通过测量当前系统中抗体的浓度,就可以准确获知当前系统遭受网络攻击时的风险,从而解决了网络攻击风险或者说网络安全风险实时、定量检测这一难题。同时根据实时定量评估出的网络风险对其变化趋势进行定量预测。

4 基于免疫的网络安全态势感知模型实现

4.1 态势觉察

态势觉察是网络安全态势感知的基础,为态势感知提供实时的原始数据,即提取态势元素,并对态势元素进行分类。态势要素一般是通过入侵检测系统提供,但传统的入侵检测系统提供的海量攻击信息中包括大量的系统信息和底层数据,冗余度太高,关联性不足,不能给网络管理员提供有效的态势信息。系统必须将攻击数据提炼成攻击信息,进而提炼成攻击知识后才能为网络管理员所理解、运用。

生物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,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)来实现的。

人工免疫系统以生物免疫原理为基础,突破了传统的方法论和思维方式,不再是单纯地分析已知规则,而是让免疫系统能够自己学习和判断,反映了“细胞自我学习”过程,提高了传统 IDS 发现攻击的能力。特别是动态克隆选择算法的提出,增强了基于免疫的网络攻击检测器的适应性,为检测变异攻击和未知攻击提供了理论依据。

基于免疫的入侵检测模型借鉴生物免疫原理,对输入的抗原集合 Ag 通过模拟免疫细胞 B 的入侵检测器将其区分为自体和非自体,模型体系结构如图 2 所示。检测器根据自身的演变过程分为:未成熟检测器、成熟检测器以及记忆检测器 3 类。新生成的未成熟检测器经过自体耐受后进化为成熟检测器。成熟检测器在生命周期内匹配到一定数目的非自体抗原原则时被激活进化为记忆检测器,否则会由于年龄过大而死亡。记忆检测器有无限的生命周期,其工作流程由两大循环组成:一类是免疫检测器识别入侵的过程,另一类是免疫检测

在GM(1,1)预测过程中,首先对含有 n 个已知网络安全态势值 $r(t_k)(k=1,2,\dots,n)$ 的时间序列 $R(t)=\{r(t_1),r(t_2),\dots,r(t_n)\}$ 进行累加变换得到递增时间序列 $\hat{R}(t)$,对其建立微分方程,求解后得到反映态势递增时间序列 $\hat{R}(t)$ 变化规律的预测函数 $F(t)$,得到与已知安全态势值 $r(t_k)(k=1,2,\dots,n)$ 对应的预测值 $\hat{r}(t_k)(k=1,2,\dots,n)$,计算出这两个值之间的残差 $e(t_k)=r(t_k)-\hat{r}(t_k)(k=1,2,\dots,n)$,将其按时间顺序排列构成网络安全态势残差序列 $\hat{E}(t)=\{e(t_1),e(t_2),\dots,e(t_n)\}$ 。

在基于马尔可夫状态概率矩阵的误差修正过程中,首先确定网络安全态势残差序列 $\hat{E}(t)=\{e(t_1),e(t_2),\dots,e(t_n)\}$ 中的最大值 e_{max} 和最小值 e_{min} ,以此为基准划分若干个状态区间。根据 $e(t_k)(k=1,2,\dots,n)$ 所处的状态,通过对相邻时刻 $e(t_k)$ 所处的状态变化的统计,计算出 $e(t_k)$ 在各个状态之间的转移概率,建立状态转移概率矩阵。并根据当前时刻 $e(t_n)$ 所处状态,结合状态转移概率矩阵判断出 $n+1$ 时刻残差 $e(t_{n+1})$ 最有可能存在的状态,以该状态区间的中间值为 $e(t_{n+1})$ 的取值,对GM(1,1)模型预测的 $n+1$ 时刻的风险值 $\hat{r}(t_{n+1})$ 进行误差修正,得到灰色马尔可夫模型对 $n+1$ 时刻的网络安全态势预测值 $\hat{r}'(t_{n+1})=\hat{r}(t_{n+1})-e(t_{n+1})$ 。

5 仿真试验

5.1 实验环境及参数设定

实验环境由配置相同的 20 余台主机、服务器构成,选取 MIT LINCOLN 实验室的 KDDCUP 99^[7] 的部分数据作为试验数据,采用 guess_passwd, buffer_overflow, land, spy, perl 等 10 余种攻击对处于网络内的 ftp 服务器、打印服务器、数据库服务器等进行攻击。ftp 服务器、打印服务器、数据库服务器等服务器重要性分别设为 0.5, 0.2, 0.8 等。Synflood, land, smurf 等攻击的直接危险性设为 0.8, 0.5, 0.9。

5.2 实验结果分析

在完成系统参数配置后,根据基于免疫的网络安全态势感知系统 NSSASI 的安全态势定量评估结果绘制网络安全态势曲线,并与相应的网络攻击强度曲线(每秒发送的攻击数据包数目)相比较,试验结果如图 3 所示。

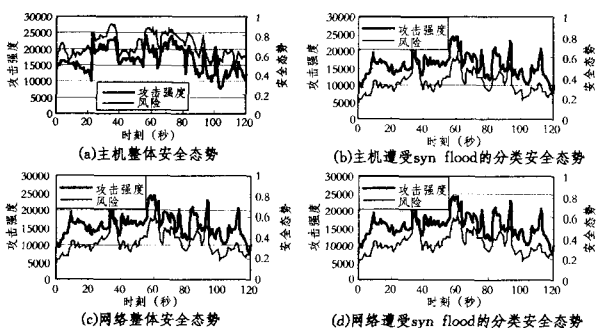


图3 网络安全态势曲线及相应的攻击强度曲线

从图3可以看出,随着攻击强度的增加,其相应的安全态势也迅速上升,这是由于记忆抗体匹配到抗原后就迅速克隆,导致抗体浓度上升。当攻击强度下降时,其相应的安全态势指标也降低,但下降的斜率相对攻击强度下降的斜率要小。

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。

以定量评估的网络实时安全态势时间序列中波动性较大指标值作为灰色马尔可夫模型的原始数据序列,对其未来的变化趋势进行预测,并将预测结果与实际风险进行对比,其结果如图4所示。

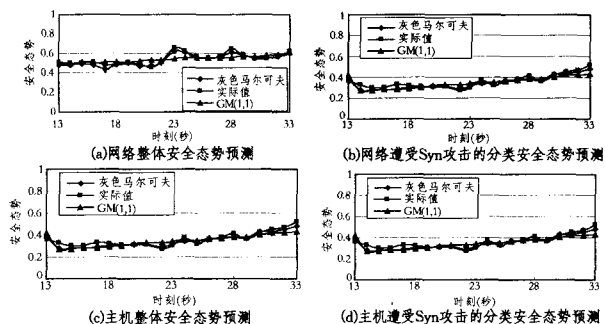


图4 灰色马尔可夫模型实时风险预测、GM(1,1)模型安全态势预测与实际态势值对比

从两种模型预测出的网络安全态势与实际攻击强度曲线的对比可以发现,GM(1,1)模型对网络安全态势预测结果的几何图形是一条较为平滑的曲线,能够把握网络安全态势的总体变化趋势,但对安全态势的随机波动性处理能力较差,当实际安全态势波动程度上升时,GM(1,1)模型预测出的安全态势与实际状况误差较大,随着时间的推移,误差急剧上升。

灰色马尔可夫模型利用马尔可夫理论中的状态转移概率反映风险变化过程中各种随机因素的影响程度和各状态之间的内在规律性,根据系统状态之间的转移概率确定未来时刻安全态势所处的状态,以状态区间的残差中间值来对GM(1,1)模型预测结果进行误差修正,降低随机因素对安全态势变化的影响,预测出的网络安全态势与实际状况较为接近,具有较高的精度,表明基于灰色马尔可夫模型的网络安全态势预测具有较强的可用性和较高的准确率。

结束语 本文将人工免疫技术应用于网络安全态势感知领域,提出了一种基于免疫的网络安全态势感知模型。该模型将免疫系统的基本原理应用于网络安全态势感知领域,从网络安全态势觉察、网络安全态势评估、网络安全态势预测等方面建立了立体的网络安全态势感知体系架构,通过识别恶意攻击行为,对网络信息系统当前安全状况及未来变化趋势进行实时、定量的分析和预测,使网络信息系统和生物免疫系统同样具有自学习性和自适应性,增强免疫力和生存能力,缓解网络攻击造成的危害,为制定合理准确的响应决策提供依据,从而提高网络信息系统的应急响应能力。

参考文献

- [1] Pilz A, Swoboda J. Network management information models [J]. Aeu-International Journal of Electronics and Communications, 2004, 58: 165-171
- [2] Esponda F, Forrest S, Helman P. A formal framework for positive and negative detection schemes [J]. IEEE Transactions On Systems Man and Cybernetics Part B-Cybernetics, 2004, 34(1): 357-373
- [3] 李涛. 计算机免疫学[M]. 北京: 电子工业出版社, 2004: 1-232

号:

- 元组角色,用 s_0, s_1, \dots 表示;
- 属性值(AV)-角色,用 q_0, q_1, \dots 表示。

定义 5 (i) 每一元组概念名称都是元组概念;

如果 C, D 是元组概念,那么 $\neg C, C \cap D, C \sqcup D$ 也是元组概念;

如果 C 是元组概念并且 s 是元组角色,那么 $\exists s. C$ 和 $\forall s. C$ 是元组概念;

如果 V 是 AV-概念并且 a 是角色,那么 $\exists a. V$ 和 $\forall a. V$ 是元组概念;

(ii) 每一 AV-概念名称都是 AV-概念;

如果 V, W 是 AV-概念,那么 $\neg V, V \cap W, V \sqcup W$ 也是 AV-概念;

如果 V 是一个 AV-概念并且 q 是 AV-角色,那么 $\forall q. V$ 和 $\exists q. V$ 是 AV-概念;

如果 C 是概念并且 a 是角色,那么 $\exists a^- . C$ 和 $\forall a^- . C$ 是 AV-概念;

(iii) 如果 a 是角色,那么 a^- 也是角色。

一个模型 M 是一个三元组 (Δ, Σ, I) ,使得对每一元组角色 s ,有:

$$I(s) \in \Delta^2$$

并对 AV-角色 q ,有:

$$I(q) \subseteq \Sigma^2$$

归纳地定义概念、AV-概念和角色的解释如下:

$$I(\exists s. C) = \{x \in \Delta: \exists y \in \Sigma((x, y) \in I(s) \& y \in I(C))\}$$

$$I(\forall s. C) = \{x \in \Delta: \forall y \in \Sigma((x, y) \in I(s) \Rightarrow y \in I(C))\}$$

$$I(\exists q. V) = \{v \in \Sigma: \exists w \in \Delta((v, w) \in I(q) \& w \in I(V))\}$$

$$I(\forall q. V) = \{v \in \Sigma: \forall w \in \Delta((v, w) \in I(q) \Rightarrow w \in I(V))\}$$

在这种描述逻辑中,可以定义概念高于 $2m$ 的学生为 $\exists a. (\exists gt. \{2\})$,其中 a 是角色属性 $height$, gt 是 AV-角色高于, $\{2\}$ 是只有实例 2 的概念(假定对于每一常量 v , $\{v\}$ 都是 AV-概念)。

结束语 本文给出了两种双层描述逻辑:一种中的角色联系的是元组和属性值,另一种中的角色分为 3 类:联系元组之间关系的角色、元组和属性值之间关系的角色以及属性值之间关系的角色。

本文试图表明:为了用形式逻辑(像描述逻辑)来形式化计算机科学与技术中的系统,应存在多种表示形式。通常,给出一个实际系统的完全的逻辑描述表示是很难的。可以看出,即使是数据库关系这种简单系统,用描述逻辑表示时也存在几种可选择的形式表示。

根据知识表示的粒度,为了表示数据库中的查询、关系代数运算和关系,需要不同的动态描述逻辑,这与文献[14-16]中给出的是非常不同的。我们将在未来的论文中讨论这些不

同的动态描述逻辑。

参 考 文 献

- [1] Baader F, Calvanese D, McGuinness D L, et al. The Description Logic Handbook[M]. Cambridge University Press, 2002
- [2] Baader F, Laux A. Terminological Logics with Modal Operators [C]//Proceedings of the International Workshop on Description Logics-DL-95. Roma, Italy, 1995, 6-12
- [3] Baader F, Laux A. Terminological Logics with Modal Operators [C]//Proceedings of the 14th International Joint Conference on Artificial Intelligence. Morgan Kaufman, Montreal, Canada, 1995, 808-814
- [4] Beneventano D, Bergamaschi S, Lodi S. Terminological logics for schema design and query processing in OODBs[C]//Proceedings of 1st Workshop KRDB'94. Saarbrücken, Germany, 1994
- [5] Beneventano D, Bergamaschi S, Sartori C. Subsumption for Semantic Query Optimization in OODB[C]//Proceedings of the International Workshop on Description Logics-DL-94. Bonn, Germany, 1994
- [6] Borgida A. Description Logics for Querying Databases [C] // Proceedings of the International Workshop on Description Logics-DL-94. Bonn, Germany, 1994
- [7] Borgida A. Description Logics in Data Management [J]. IEEE Transactions on Knowledge and Data Engineering, 1995, 7: 671-682
- [8] Borgida F, Lenzerini M, Rosati R. Description logics for data bases[M]//Baader F, Calvanese D, McGuinness D L, et al., eds. The Description Logic Handbook. Cambridge University Press, 2002, 472-494
- [9] Calvanese D, De Giacomo G, Lenzerini M, et al. Source Integration in Data Warehousing[C]//Proceedings of the Ninth International Workshop on Database and Expert Systems Applications. Vienna, 1998, 192-197
- [10] Calvanese D, Lenzerini M, Nardi D. Description logics for conceptual data modeling [M] // Chomicki J, Saake G, eds. Logics for Databases and Information Systems. Kluwer Academic Publisher, 1998, 229-263
- [11] Lenzerini M. Description logics and their relationships with databases[C]//ICDT'99. LNCS 1540. 1998, 32-38
- [12] Lutz C. Reasoning with concrete domains[C]//Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence. Stockholm, Sweden, 1999, 90-95
- [13] Simovici D A, Tenney R L. Relational Database Systems[M]. Academic Press, 1995
- [14] Wolter F, Zakharyashev M. Modal description logics: Modalizing roles[J]. Fundamenta Informaticae, 1999, 39: 411-438
- [15] Wolter F, Zakharyashev M. Temporalizing description logics [C]//Proceedings of Fro-CoS'98. Amsterdam, 1998
- [16] Wolter F, Zakharyashev M. Dynamic description logics[C]//Advances in Modal Logic II. Stanford: CSLI Publications, 2000, 449-463
- [7] Hofmeyr S A, Forrest S. Architecture for an artificial immune system[J]. Evolutionary Computation, 2000, 8(4): 443-473
- [8] 李涛. 网络安全概论[M]. 北京: 电子工业出版社, 2004: 1-458
- [9] Bass T. Intrusion Detection Systems and Multisensor Data Fusion: Creation Cyberspace Situation Awareness[J]. Communications of the ACM, 2000(43): 99-105
- [10] Steinburg A N. Revision to the JDL Data Fusion Model[C]//Joint NA TO/IRIS Conference. Quebec, October 1998
- [11] Bass T. Intrusion Detection Systems and Multi-sensor Data Fusion: Creating Cyberspace Situational Awareness[J]. Communications of the ACM, 2000, 43(4): 99-105

(上接第 129 页)

- [4] Jerne N K. Towards a Network Theory of the Immune System [J]. Annual Immunology, 1974, 24(3): 125-134
- [5] Forrest S, Perelson A S, Allen L, et al. Self-Nonsel Self Discrimination in a Computer[C]//Proceedings of IEEE Symposium on Research in Security and Privacy. Oakland, 1994, 54-64
- [6] Hofmeyr S A, Forrest S. Immunity by design: an artificial immune system[A]//Proc. of the Genetic and Evolutionary Computation Conference [C]. Morgan-Kaufmann, San Francisco, 1999, 1289-1296