基于日志审计与性能修正算法的网络安全态势评估模型

韦 勇[1),2)] 连一峰[2)]

[1)](中国科学技术大学电子工程与信息科学系 合肥 230027)

[2)](中国科学院软件研究所信息安全国家重点实验室 北京 100190)

摘 要 文章分析和比较了目前的安全态势评估方法,提出了一种基于日志审计与性能修正算法的网络安全态势评估模型.首先利用日志审计评估节点理论安全威胁,并通过性能修正算法计算节点安全态势.然后利用节点服务信息计算网络安全态势,并且采用多种预测模型对网络安全态势进行预测,绘制安全态势曲线图.最后构建了一个网络实例,使用网络仿真软件对文中提出的态势评估模型和算法进行了验证.实验证明该方法切实有效,比传统方法更准确地反映了网络的安全态势和发展趋势.

关键词 安全态势评估;日志审计;性能修正;安全态势曲线图;预测

中图法分类号 TP393 **DOI 号**: 10.3724/SP.J.1016.2009.00763

A Network Security Situational Awareness Model Based on Log Audit and Performance Correction

WEI Yong[1),2)] LIAN Yi-Feng[2)]

[1)](*Department of Electronic Engineering and Information Science, University of Science and Technology of China, Hefei* 230027)

[2)](*State Key Laboratory of Information Security, Institute of Software, Chinese Academy of Sciences, Beijing* 100190)

Abstract This paper analyzes and compares the existing situational awareness methods and proposes a network security situational awareness model based on log audit and performance correction algorithm. First, nodes theoretic security threat is got by log audit and the value of nodes security situation is computed by performance correction algorithm. Then the value of network security situation is computed using service information, the future threat is predicted by several prediction models, and the Security Situational Graph(SSG) is drawn. Finally an example is given to validate the network security situational awareness model and algorithm by simulation software. The example proves that the model is more effective and accurate to reflect the network security situational and its trends than traditional methods.

Keywords security situational awareness; log audit; performance correction; security situational graph; predict

1 引 言

随着计算机技术和通信技术的迅速发展以及用户需求的不断增加,计算机网络规模日益庞大,应用系统日益复杂.同时,网络安全事件层出不穷,使得计算机网络面临着严峻的信息安全形势,传统的单一的防御设备或者检测设备已经无法满足安全需

收稿日期:2008-11-24;最终修改稿收到日期:2009-01-16.本课题得到国家"八六三"高技术研究发展计划项目基金(2006AA01Z437,2007AA01Z475,2006AA01Z412,2006AA01Z433)资助.**韦 勇**,男,1981 年生,博士研究生,主要研究方向为网络安全、态势评估.E-mail: weiyong@is.iscas.ac.cn. **连一峰**,男,1974 年生,博士,副研究员,主要研究方向为网络安全、脆弱性评估.

求.网络安全态势评估技术能够综合各方面的安全因素,从整体上动态反映网络安全状况,并对安全状况的发展趋势进行预测和预警,为增强网络安全性提供可靠的参照依据.因此,针对网络的安全态势评估模型及关键技术已经成为目前网络安全领域的研究热点.

安全态势评估(security situational awareness)是指感知和获取一定时间和空间环境中的安全元素,对获取的数据和信息进行整合和分析,并基于分析结果预测其未来的发展趋势.态势评估最初出现在航空和军事领域,后来被广泛应用到社会各个领域,包括生产制造控制、医学研究、人工智能等.近年来,态势评估技术被逐渐地应用于计算机网络,国内外的研究人员从不同角度出发,设计并实现了大量针对计算机网络的安全态势评估方法.

Lau[1]提出了 Spinning Cube,它是一个三维立方体视图,可以将网络连接以点的方式映射到立方体中,并根据连接时间和内容的危险性显示不同的颜色,该方法仅以网络连接作为评估指标,无法综合分析各种安全因素并准确反映出安全态势状况.

SIFT 项目组研制了 NVisionIP[2]和 VisFlowConnect[3]两种可视化工具.NVisionIP 可以显示一个 B 类网络的连接状态,并且提供了 3 种不同精度的视图;VisFlowConnect 能动态显示网络连接状态和网络流量,并且具有数据过滤能力,但是这两种工具仅反映了网络连接状态和网络流量,评估指标较为单一,对管理员经验水平要求也很高.

Bass[4]提出了利用入侵检测系统的分布式传感器进行数据融合的方法,对计算机网络安全态势进行评估,通过数据融合和数据挖掘的方法评估计算机网络的安全性,但没有实现具体的原型系统.

Information Extraction & Transport 开发了 SSARE 系统[5],用于广域网的计算机攻击检测、态势评估和响应评估.该系统实现了入侵检测、态势评估和响应评估的有机结合,但是信息获取方式较为单一.

Yegneswaran 等人[6]提出了利用 Honeynets 进行因特网安全态势评估的方法,该方法使用 Honeynets 提供的大量网络活动信息,利用入侵检测工具 Bro 对这些活动的报警信息来构建安全态势曲线,但该曲线只有在大规模病毒或蠕虫爆发时才能体现出明显效果.

陈秀真等人[7]提出了层次化网络安全威胁态势量化评估方法,利用入侵检测系统报警信息和网络性能指标,并且结合主机的漏洞信息,对服务、主机和网络进行层次化的安全定量评估,得到直观的安全态势图.

以上方法为网络安全态势评估工作提供了可行的解决思路,为评估模型及算法的研究奠定了良好的基础,但也普遍存在着一些技术缺陷.例如,缺乏对网络安全因素的全面考虑,评估数据源单一,使得评估结果不够准确;所采用的量化评估算法存在一定缺陷,导致量化结果与实际结果出现偏差;另外,这些方法没有对安全态势状况的发展趋势进行预测分析.

针对上述问题,本文提出一种基于日志审计与性能修正算法的网络安全态势评估模型,针对网络安全性的各种关键因素进行建模,利用日志审计评估节点理论安全威胁,并通过性能修正算法计算节点安全态势,再利用节点服务信息计算网络安全态势,同时采用多种预测模型对网络安全态势进行预测,绘制安全态势曲线图,直观反映网络安全状况和发展趋势,从而实现网络安全态势的量化分析和趋势预测.

本文第 2 节介绍网络安全态势评估模型;第 3 节介绍基于日志审计与性能修正算法的态势评估量化算法;第 4 节在安全态势量化结果基础上利用多种预测模型进行趋势预测和比较;第 5 节给出上述评估模型及算法的网络实例和实验结果;最后是全文的总结.

2 网络安全态势评估模型

计算机网络中存在大量的主机、服务器和各种网络设备,在网络系统运行过程中,会产生大量的系统日志、安全日志、应用日志和告警日志,这些日志之间存在一定的关联性,包含了安全事件的相关信息.在传统的安全态势评估方法中,通常是对各个日志进行单独分析和处理,以此作为网络安全威胁,却忽略了日志之间的相关性,同时偏重于理论安全威胁值的计算,而不关心网络实际运行的性能状态,使得分析结果无法准确反映网络安全态势状况和趋势.本文提出了基于日志审计与性能修正算法的网络安全态势评估模型,利用日志审计技术对各种日志进行相关性分析,获取安全事件信息,计算出节点理论安全威胁值,然后利用节点性能信息进行修正得到节点安全态势和网络安全态势,最后采用多种预测模型对网络安全态势进行预测,从而有效地弥

补了传统安全态势评估方法的不足.

下面首先给出本评估模型中定义的一些术语.

定义 1.　日志审计. 它是指对各种主机、服务器和网络设备在运行过程中产生的日志和消息进行实时采集，并进行实时关联分析，发现和报告各种安全事件.

定义 2.　理论安全威胁. 它是指利用日志审计得出的安全事件和先验知识计算被攻击的网络节点理论上受到的影响程度.

定义 3.　性能修正. 它是指利用网络节点的实际性能变化对理论安全威胁进行修正，从而得到更加准确的节点安全态势评估结果.

本文提出的网络安全态势评估模型的评估框架如图 1 所示.

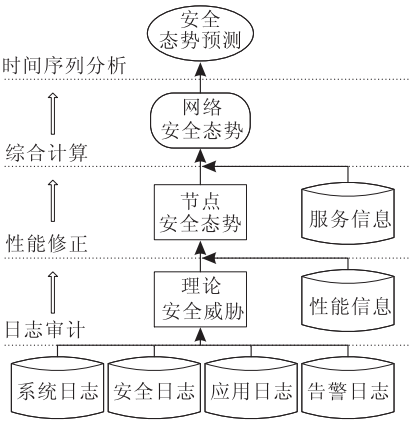

图 1　网络安全态势评估框架

第 1 步，利用多种日志，通过日志审计算法计算网络节点理论安全威胁；

第 2 步，利用性能信息，通过性能修正算法计算网络节点安全态势；

第 3 步，利用服务信息计算节点权重，然后综合计算网络安全态势；

第 4 步，根据网络安全态势评估结果，通过时间序列分析方法对网络安全态势进行趋势预测.

根据上述评估框架，结合计算机网络自身的特点，将安全态势的评估指标归纳为 3 类：日志信息 L、性能信息 P 和服务信息 S，并针对网络安全态势评估进行如下建模：

(1) L：日志信息. 它包含了网络运行时产生的系统日志、安全日志、应用日志和告警日志等信息. 对于网络中的任何一条日志信息 $l \in L$，都用一个六元组 $(id_l, time_l, type, content, id_c, id_h)$ 来表示，其中 id_l 是日志信息的唯一标识符；$time_l$ 是日志信息产生的时间；$type$ 是日志信息的类型，并且 $type \in \{Sys, Sec, App, Alert\}$，其中 Sys 表示系统日志，Sec 表示安全日志，App 表示应用日志，$Alert$ 表示告警日志；$content$ 是日志信息的内容；id_c 是产生日志的节点标识符，id_h 是日志对应安全事件的目标节点标识符.

(2) P：性能信息. 对于网络节点的任何一条性能信息 $p \in P$，都用一个八元组 $(id_p, time_p, \gamma, \mu, \kappa, \rho, \delta, id_h)$ 来表示，其中 id_p 是性能信息的唯一标识符，$time_p$ 是性能信息产生的时间；γ 是节点处理器使用率；μ 是节点内存使用率；κ 是连接数；ρ 是流量；δ 是丢包率；id_h 是该性能信息所属节点标识符.

(3) S：服务信息. 对于网络节点的任何一条服务信息 $s \in S$，都用一个四元组 $(id_s, name, w_s, id_h)$ 来表示，其中 id_s 是服务信息的唯一标识符，$name$ 是服务名称；w_s 是服务权重；id_h 是该服务信息所属节点标识符.

SA 表示网络的安全态势值，由日志信息 L、性能信息 P 和服务信息 S 共同组成，表示为 $SA = (L, P, S)$.

下面首先介绍利用上述模型进行安全态势评估的量化算法.

3　网络安全态势评估量化算法

本节介绍利用评估模型对网络安全态势进行量化分析，主要包含 3 个步骤：日志审计、性能修正和综合计算. 整个量化算法流程如图 2 所示.

```
算法1. 网络安全态势值量化分析算法.
输入: 日志信息, 性能信息, 服务信息
输出: 网络安全态势值
  BEGIN
  1  For each host H_0, H_1, ..., H_n
  2    Let L represent the logs, Eve represent the security events;
  3    Eve = audit(L);                  //日志审计获取安全事件;
  4    Let VoT represent the value of theoretic threat;
  5    VoT = evaluate(Eve);             //评估理论安全威胁;
  6    Let ΔP represent the change of performance;
  7    SA_h = correction(VoT, ΔP);//性能修正计算主机安全态势值;
  8  SA = compute(SA_h, w_h);         //综合计算网络安全态势值;
     where w_h is the weight of network nodes;
  9  return SA.
  END
```

图 2　网络安全态势值量化分析算法

在算法 1 中，首先对评估模型中的日志信息 L 进行日志审计得到安全事件信息，根据已知的安全事件知识评估网络节点的理论安全威胁值；然后利用评估模型中的性能信息 P 得到节点实际性能的

变化值 ΔP，再利用它对节点理论安全威胁值进行修正，得到节点安全态势值；最后利用各节点权重综合计算网络安全态势值，并且利用多个时段的安全态势值绘制安全态势曲线图.

下面详细介绍网络安全态势评估量化算法的 3 个步骤.

3.1 日志审计

在网络运行过程中会产生大量的日志信息，如系统日志、安全日志、应用日志和告警日志等，这些日志信息包含了网络的安全事件及其内在联系，通过日志审计可以挖掘出日志中包含的安全事件. 日志审计通常有基于规则库、基于数理统计、基于有学习能力的数据挖掘等方法. 基于规则库方法主要通过规则匹配，优点是准确率较高，缺点在于很难分析出未知安全事件；基于数理统计方法主要通过对实时数据的统计确定正常值范围，优点是可以发现未知安全事件，缺点是设定统计量阈值较难，安全事件类型也难于区别；基于有学习能力的数据挖掘方法是建立挖掘模型并不断进行迭代和调整，优点是可以减少规则手工编码和经验成分，缺点是算法较复杂，要求较高.

本文主要采用基于规则库的日志审计方法，通过对日志信息进行规则库匹配得到初步的安全事件，然后通过安全事件归并去掉重复的安全事件，再通过安全事件关联分析挖掘出安全事件内在的联系，实现日志审计功能，最后量化评估节点理论安全威胁，算法流程如图 3 所示.

```
算法2. 基于规则库的日志审计算法.
输入: 日志信息, 安全事件规则库
输出: 节点理论安全威胁值
BEGIN
Let Eve represent the security events database;
Set Eve = ∅;
For each log l_0, l_1, ..., l_n of nodes
  Let r represent the rules database,
  Eve+=match(l_i, r);            //匹配规则库得到安全事件;
Let Eve=merge(Eve);              //归并重复的安全事件;
Let Eve=relate(Eve, name, time, target); //安全事件关联分析;
  where name, time and target are security events' parameters;
Let VoT represent the value of theoretic threat;
VoT=evaluate(Eve);               //评估节点理论安全威胁;
return VoT.
END
```

图 3 日志审计算法

在算法 2 中，首先将安全事件库置为空，然后利用规则库依次对每一条日志信息进行匹配，将匹配到的安全事件添加到安全事件库中；再对安全事件进行归并处理，将发生时间相同、类型相同、攻击源相同和攻击目标相同的多条重复的安全事件归并为一条安全事件；然后通过安全事件关联分析，对类型相同、发生时间相同、攻击源相同但攻击目标不同的安全事件，类型相同、发生时间相同、攻击源不同但攻击目标相同的安全事件，以及攻击类型不同、发生时间不同但攻击目标相同或相关的安全事件进行相关性分析，得到抽象层次更高的安全事件信息；最后通过已知的安全事件知识得到节点理论安全威胁 VoT.

3.2 性能修正

节点理论安全威胁只反映了安全事件理论上对网络节点的影响程度，而且无法对未知攻击的影响进行评估. 本文结合实际性能信息，利用性能修正算法计算节点安全态势值，更加准确地反映出节点的安全态势状况，性能修正算法需满足两个假设.

假设 1. 评估过程中未对网络进行人工安全操作，比如升级硬件、清除病毒等. 人工操作会导致网络性能提升，抵消或降低攻击所带来的实际威胁.

假设 2. 评估过程中合法用户的正常服务请求在时间上为均匀分布. 正常服务请求的数量变化也会影响网络性能，进行安全态势评估时需要排除此类影响.

安全态势评估模型中的性能信息 P 用 $(id_p, time_p, \gamma, \mu, \kappa, \rho, \delta, id_h)$ 表示，这些参数用于对节点性能变化量进行计算. 对某 id_h 节点，在 $time_p$ 时刻其性能参数 $(\gamma, \mu, \kappa, \rho, \delta)$ 的最小值都为 0，对应的最大值为 $(1, 1, \kappa_0, \rho_0, 1)$，其中 κ_0 是最大允许连接数；ρ_0 是最大流量. 网络节点性能由当前可利用资源来衡量，采用如下公式计算节点当前的性能值 P：

$$P=1-\frac{1}{5}\left(\gamma+\mu+\frac{\kappa}{\kappa_0}+\frac{\rho}{\rho_0}+\delta\right) \tag{1}$$

设在某时段开始时刻某主机的性能参数为 $(\gamma_1, \mu_1, \kappa_1, \rho_1, \delta_1)$，该时段结束时刻的性能参数为 $(\gamma_2, \mu_2, \kappa_2, \rho_2, \delta_2)$，则利用式(1)计算可得

$$P_1=1-\frac{1}{5}\left(\gamma_1+\mu_1+\frac{\kappa_1}{\kappa_0}+\frac{\rho_1}{\rho_0}+\delta_1\right),$$

$$P_2=1-\frac{1}{5}\left(\gamma_2+\mu_2+\frac{\kappa_2}{\kappa_0}+\frac{\rho_2}{\rho_0}+\delta_2\right),$$

$$\begin{aligned}\Delta P &= P_1-P_2\\ &=\frac{1}{5}\left(\gamma_2-\gamma_1+\mu_2-\mu_1+\frac{\kappa_2-\kappa_1}{\kappa_0}+\frac{\rho_2-\rho_1}{\rho_0}+\delta_2-\delta_1\right).\end{aligned}$$

使用性能变化量 ΔP 对节点理论安全威胁 VoT 进行修正，就可以得到节点安全态势值 SA_h，计算公式为

$$SA_h=(1-\eta)\times VoT+\eta\times\Delta P \tag{2}$$

其中 η 为修正系数，取值为[0,1]，表示性能修正在节点安全态势值计算中所占的比例. 当 $\eta=0$ 时，节点安全态势仅反映理论安全威胁；当 $\eta=1$ 时，节点安全态势仅反映实际性能改变值.

3.3　综合计算

得到各个网络节点的安全态势值后，即可利用各节点权重信息加权计算网络安全态势. 节点服务信息 S 以(id_s,$name$,w_s,id_h)表示，可由以下公式计算节点权重 w_h：

$$w_h=\sum_{i=1}^{m} w_{si} \tag{3}$$

其中，m 为节点提供的服务数，w_{si} 为每个服务所占的权重. 网络中各种服务的权重和为 1，如果同时有几个节点提供同一种服务，那么该服务的权重将被平均分配给这几个节点，因此计算得到的节点权重和也为 1.

最后由各个节点的安全态势值 SA_h 和节点权重 w_h 可以计算得到网络安全态势值 SA，计算公式如下：

$$SA=\sum_{i=1}^{n} w_{hi}\times SA_{hi} \tag{4}$$

其中 n 为网络节点数量，w_{hi} 为式(3)计算得到的每个节点所占的权重.

利用以上量化分析算法可以计算网络不同时段的安全态势值，并绘制安全态势曲线图，安全态势曲线图从整体上反映网络的安全态势变化情况，方便网络管理员掌握整个网络的安全态势状况.

4　网络安全态势预测算法

本节介绍基于时间序列分析的态势预测算法，对利用量化算法计算出的多个时段的网络安全态势值样本进行时间序列分析，从而实现对未来的网络安全趋势进行预测. 时间序列分析是根据系统观测得到的时间序列数据，通过曲线拟合和参数估计来建立数学模型的理论和方法. 时间序列分析常用在国民经济宏观控制、市场潜量预测、气象预报等方面. 在本文中将利用一阶灰色预测 GM(1,1)模型、ARMA 模型和 Holt-Winter 模型 3 种方法对网络安全态势值进行预测分析.

4.1　GM(1,1)模型

灰色系统理论是邓聚龙教授[8]于 1982 年创立的，已经得到了广泛的应用，取得了较好的成果. 灰色系统理论是一种研究一些既含有已知信息又含有未知或未确知信息的系统理论和方法，它从杂乱无章的、有限的、离散的数据中找出数据的规律，然后建立相应的灰色模型进行预测. 灰色理论的实质是对原始随机数列采用生成信息的处理方法来弱化其随机性，使原始数据序列转化为易于建模的新序列.

GM(1,1)模型是一种最常用的灰色预测模型，它是由一个只包含单变量的一阶灰微分方程构成的模型，在本文中这个单变量即为网络安全态势值. 下面介绍 GM(1,1)模型的预测方法和步骤.

首先将原始数据序列做一次累加处理，获得新的数据序列，设安全态势值原始序列为

$$X^{(0)}(t)=\{x^{(0)}(1),x^{(0)}(2),\cdots,x^{(0)}(n)\}.$$

令 $X^{(1)}(t)=\sum_{i=1}^{t} x^{(0)}(i)$，$t=1,2,\cdots,n$，则可得新的数据序列：

$$X^{(1)}(t)=\{x^{(1)}(1),x^{(1)}(2),\cdots,x^{(1)}(n)\}.$$

而 GM(1,1)模型的单变量一阶灰微分方程为

$$\frac{\mathrm{d}x^{(1)}(t)}{\mathrm{d}t}+\alpha x^{(1)}(t)=\mu \tag{5}$$

上式中，α 和 μ 是模型的待定参数，可利用最小二乘法由下式计算：

$$(\alpha,\mu)^{\mathrm{T}}=(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B})^{-1}\boldsymbol{B}^{\mathrm{T}}\boldsymbol{y} \tag{6}$$

上式中，

$$\boldsymbol{B}=\begin{bmatrix} -1/2[x^{(1)}(1)+x^{(1)}(2) & 1 \\ \vdots & \vdots \\ -1/2[x^{(1)}(n-1)+x^{(1)}(n)] & 1 \end{bmatrix},$$

$$\boldsymbol{y}=[x^{(0)}(2),x^{(0)}(3),\cdots,x^{(0)}(n)].$$

所以计算得到预测的累加序列值为

$$\hat{x}^{(1)}(t+1)=\left[x^{(1)}(1)-\frac{\mu}{\alpha}\right]\mathrm{e}^{-at}+\frac{\mu}{\alpha} \tag{7}$$

最后将该累加序列值还原即可得到预测值为

$$\hat{x}^{(0)}(t+1)=\hat{x}^{(1)}(t+1)-\hat{x}^{(1)}(t) \tag{8}$$

一阶灰色预测模型算法简单，易于实现，速度也较快，并且预测过程不需要进行参数设定或其它人工干预，预测结果比较平滑地反映出原序列的发展趋势，缺点在于预测结果无法体现随机性和周期性等因素.

4.2　ARMA 模型

ARMA 模型即自回归移动平均模型，是美国统计学家 George 和英国统计学家 Gwilym 提出的一种时间序列预测方法[9]，它的基本思想是：将预测对象随时间推移而形成的数据序列视为一个随机序列，即除去个别的因偶然因素引起的观测值外，时间序列是一组依赖于时间的随机变量，这组随机变量

所具有的依存关系或自相关性表征了预测对象发展的延续性，而这种自相关性一旦被相应的数学模型描述出来，就可以从时间序列的过去值及现在值预测其未来的值. 下面介绍 ARMA 模型的预测方法和步骤.

首先进行平稳性检验，ARMA 模型要求时间序列为平稳序列，所以平稳性检验的目的就是验证时间序列的平稳性，从而为建模提供依据. 平稳性检验通常包括参数检验法和非参数检验法. 经过检验，如果是平稳时间序列，则继续分析；如果是非平稳序列，则对其差分序列进行平稳性检验，直到满足平稳性条件为止.

然后计算时间序列的样本自相关系数和偏自相关系数，根据样本自相关系数和偏自相关系数的拖尾或截尾性质，确定采用的时间序列模型及其阶数. 设 $x_1, x_2, \cdots, x_n$ 为网络安全态势值时间序列，$\bar{x}$ 为网络安全态势值时间序列的平均值，则样本自相关系数 $\hat{\rho}_k$ 和偏自相关系数 $\hat{\phi}_{kk}$ 的计算公式如下：

$$\hat{\rho}_k = \frac{\sum_{t=1}^{n-k}(x_t-\bar{x})(x_{t+k}-\bar{x})}{\sum_{t=1}^{n}(x_t-\bar{x})^2} \tag{9}$$

$$\hat{\phi}_{kk} = \begin{cases} \hat{\rho}_1, & k=1 \\ \dfrac{\hat{\rho}_k - \sum_{j=1}^{k-1}\hat{\phi}_{k-1,j}\cdot\hat{\rho}_{k-j}}{1-\sum_{j=1}^{k-1}\hat{\phi}_{k-1,j}\cdot\hat{\rho}_j}, & k=2,3,\cdots \end{cases} \tag{10}$$

其中 $\hat{\phi}_{k,j} = \hat{\phi}_{k-1,j} - \hat{\phi}_{kk}\cdot\hat{\phi}_{k-1,k-j}$，$j=1,2,\cdots,k-1$.

计算出样本自相关系数和偏自相关系数后，由如下规则确定时间序列的数学模型和阶数：如果 $\hat{\rho}_k$ 拖尾，而 $\hat{\phi}_{kk}$ 为 p 阶截尾，则采用 $AR(p)$ 模型；如果 $\hat{\rho}_k$ 为 q 阶截尾，而 $\hat{\phi}_{kk}$ 拖尾，则采用 $MA(q)$ 模型；如果 $\hat{\rho}_k$ 和 $\hat{\phi}_{kk}$ 均拖尾，则采用 $ARMA(p,q)$ 模型.

时间序列模型和阶数确定以后，再利用参数估计方法计算模型的各个参数. 常用的参数估计方法有矩估计、极大似然估计和最小二乘估计. 计算出模型参数后，就得到了安全态势值时间序列预测模型的表达式.

最后利用预测模型表达式对网络安全态势值进行计算和预测，并对模型的残差序列进行白噪声检验. ARMA 模型算法反映了时间序列的自相关性，而且预测结果体现了时间序列的随机性和周期性等因素，缺点是预测过程需要进行较多的人工操作，并且由于掌握的样本个数总是有限的，如果预测的时间点越远，则预测结果与实际值的偏差将会越大. 在实际应用中，需要即时掌握最新的安全态势数据，对预测模型进行修正，以使其达到最佳，反映网络安全的最新状况和趋势.

4.3 Holt-Winter 模型

Holt-Winter 模型的基本思想是把具有线性趋势、季节变动和随机波动的时间序列进行分解研究，并与指数平滑法相结合，分别对长期趋势、趋势的增量和季节波动做出估计，然后建立预测模型，外推预测值[10]. 该模型由以下 3 个平滑方程和一个预测公式组成：

$$S_t = \alpha X_t / I_{t-L} + (1-\beta)(S_{t-1}+b_{t-1}) \tag{11}$$

$$b_t = \gamma(S_t - S_{t-1}) + (1-\gamma)b_{t-1} \tag{12}$$

$$I_t = \beta X_t / S_t + (1-\beta) I_{t-L} \tag{13}$$

其中 X_t 为时间序列，L 为季节长度，S 是稳定成分，b 是线性趋势成分，I 为季节成分，α、β 和 γ 为加权系数，而预测公式为

$$f_{t+m} = (S_t + b_t m) I_{t+m-L} \tag{14}$$

其中 m 为从当前时期所要预测时期的数目，f_{t+m} 为第 $t+m$ 时期的预测值.

假设 X_t 为网络安全态势值时间序列，下面详细介绍利用该模型进行预测的步骤.

首先计算时间序列前两个周期中平均每个周期的增量 B，计算公式如下：

$$B = (1/L)(V_2 - V_1) \tag{15}$$

其中 $V_1 = (1/L)\sum_{t=1}^{L} X_t$，$V_2 = (1/L)\sum_{t=L+1}^{2L} X_t$.

然后计算初始增量 b_{2L+1} 及初始指数平滑值 S_{2L+1}，计算公式如下：

$$b_{2L+1} = B \tag{16}$$

$$S_{2L+1} = V_2 + (L-1)B/2 \tag{17}$$

接下来分别计算前两个周期内每个时期的季节指数，计算公式如下：

$$I_{1t} = X_t / \{V_1 - [(L+1)/2 - m]B\} \tag{18}$$

$$I_{2t} = X_{t+L} / \{V_2 - [(L+1)/2 - m]B\} \tag{19}$$

以上两式中 $t=1$ 时，$m=1,\cdots,t=L$ 时，$m=L$.

然后计算前两个周期中平均每个时期的季节指数，计算公式如下：

$$I'_t = (I_{1t} + I_{2t})/2,\ t=1,2,\cdots,L \tag{20}$$

再然后将季节指数正态化，得到 L 个正态化后的季节指数 I_t，计算公式如下：

$$L'=\sum_{t=1}^{L} I'_t,\ \ I_t=(L/L')/I'_t,\ \ t=1,2,\cdots,L \quad (21)$$

最后是确定最优加权系数 α、β 和 γ 的值，确定原则是使预测值与实际值之间误差的均方差最小.

经过以上步骤后即可利用模型参数和预测公式对安全态势值进行预测，Holt-Winter 模型对时间序列的趋势、随机性、周期和季节等因素都进行了考虑，预测结果均方差较小，在短期季节预测中效果不错，但随着预测的时间点越远，预测结果与实际值的偏差也会越来越大.

5　实例分析

为了验证本文模型的适用性，我们构造了一个实验网络. 实验网络的拓扑图如图 4 所示，拓扑中包含 4 个服务器节点和路由器、防火墙、入侵检测系统等网络设备. 服务器节点提供了相应的一些网络服务，正常用户 User 和攻击者 Attacker 都可以通过网络访问服务器节点，对其进行正常使用或者攻击. 网络运行时在各个服务器节点、入侵检测系统、路由器和防火墙中会产生相应的日志信息，同时服务器节点的性能信息也可以实时地进行采集分析.

各个网络服务器节点的服务信息 S 按照本文提出的模型表示如下：

(id1，Web，0.4，Server1)；

(id2，FTP，0.2，Server2)；

(id3，Database，0.2，Server3)；

(id4，Database，0.2，Server4).

由式(3)计算可得 4 个网络服务器节点权重依次为 0.4、0.2、0.2、0.2.

实验采用 NS2[11] 进行模拟，对服务器节点和路由器、防火墙、入侵检测系统等网络设备分别设置模拟节点，并且把正常用户集合和攻击者集合都设置为单独的节点，然后通过从正常用户节点到网络各节点的连接和流量以及攻击者节点到网络各节点的连接和流量模拟整个攻击过程，采用本文第 3 节所描述的量化算法分析网络安全态势，然后利用本文第 4 节所描述的预测算法对网络安全态势进行趋势预测.

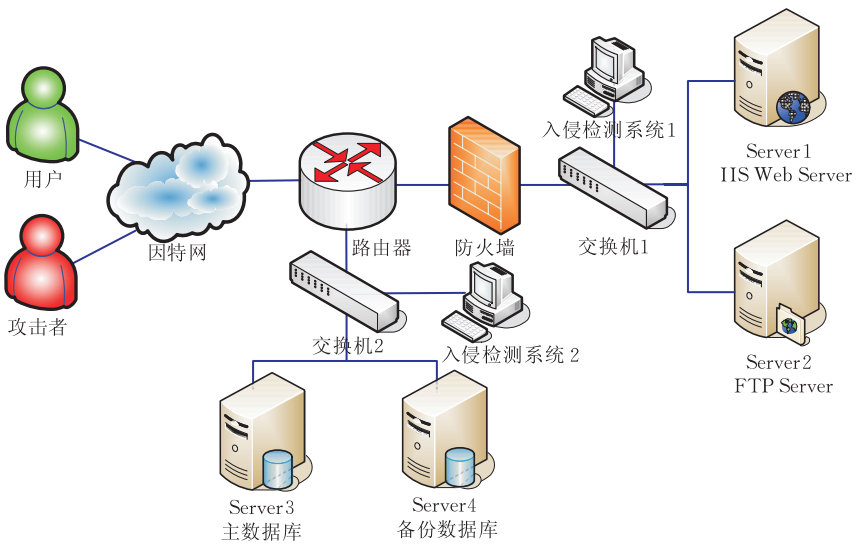

图 4　网络实例拓扑图

在模拟正常使用和攻击过程时，按照如下的步骤进行模拟：

0. 正常用户集合在模拟过程中一直访问 Server 1、Server 2 和 Server 3；

1. Server 3 受到 SQL 注入攻击；

2. Server 4 受到通过 Server 3 进行的 SQL 注入攻击；

3. Server 1 受到 UDP FLOOD 攻击(DoS 攻击)；

4. Server 2 受到 MSBLAST 蠕虫攻击；

5. Server 1、Server 2 和 Server 3 同时受到以上 1、3、4 步骤中的攻击；

6. Server 3 受到 SQL 注入攻击和一种未知攻击；

7. Server 1 受到 Unicode 解码漏洞攻击和 SYN FLOOD 攻击.

NS2 模拟过程中，所有数据包信息记录在 trace 文件中，利用编写的 awk 程序对攻击前后时刻的服务器节点的流量、连接数和丢包率等参数进行计算，并根据这些参数计算服务器节点的 cpu 使用率等性能参数，并将实验数据用于网络安全态势值的量化和预测，从而验证本文提出的安全态势评估模型和方法.

模拟结束后，首先按照第 3.1 节介绍的日志审计算法评估节点理论安全威胁 VoT，我们以分析步骤 1

产生的日志为例，该步骤将在 NIDS 2 和 Server 3 上产生相关的日志，经过预处理后根据本文的模型可表示为

(id1，time1，Alert，SQL injection，NIDS2，Server3)；

(id2，time1，App，SQL injection，Server3，Server3).

根据以上日志分析出安全事件并进行归并后得到 SQL 注入攻击，利用对该攻击的已有知识，评估该时段节点 Server 3 的理论安全威胁为 0.2，用同样的方法可以评估其它时段所有节点的理论安全威胁.

然后利用各时段所有节点的性能变化量 ΔP，对节点理论安全威胁 VoT 进行性能修正得到节点安全态势 SA_h，最后再利用节点权重计算网络安全态势 SA. 模拟实验数据及计算结果如表 1 所示 ($\eta=0.5$).

表 1 模拟实验数据及计算结果表

时间	网络节点及权重												网络安全态势值SA
	Server 1(0.4)			Server 2(0.2)			Server 3(0.2)			Server 4(0.2)			
	VoT	ΔP	SA_h	VoT	ΔP	SA_h	VoT	ΔP	SA_h	VoT	ΔP	SA_h	
时段 1	0	0	0	0	0	0	0.200	0.038	0.119	0	0	0	0.024
时段 2	0	0	0	0	0	0	0	0.020	0.010	0.200	0.041	0.121	0.026
时段 3	0.400	0.457	0.429	0	0	0	0	0	0	0	0	0	0.172
时段 4	0	0	0	0.400	0.472	0.436	0	0	0	0	0	0	0.087
时段 5	0.400	0.396	0.398	0.400	0.395	0.398	0.200	0.038	0.119	0	0	0	0.263
时段 6	0	0	0	0	0	0	0.200	0.345	0.273	0	0	0	0.055
时段 7	0.300	0.345	0.323	0	0	0	0	0	0	0	0	0	0.129

根据以上模拟实验的计算结果进行绘图，得到网络安全态势曲线图如图 5 所示，图中横轴为时间，纵轴为网络安全态势值，态势值越大则表明网络安全状况越严重.

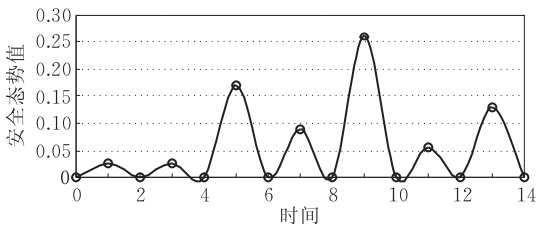

图 5 网络安全态势曲线图

从网络安全态势曲线图可以看出，如果受到攻击的节点权重越大、攻击的数量越多或者攻击的威胁程度越大，那么网络安全态势状况越差. 网络安全态势曲线图与模拟过程的安全态势状况基本一致，并且采用性能修正算法在一定程度上反映出了未知攻击对网络安全态势的影响.

最后，按照第 4 节介绍的 3 种预测模型分别对网络安全态势时间序列进行预测，并利用预测结果进行绘图，得到网络安全态势预测曲线图如图 6 所示. 从网络安全态势预测曲线图可以看出，3 种模型的预测结果均显示网络安全态势总体趋势逐渐上升，网络安全状况有逐渐变差的趋势，需要即时采取相应安全措施.

根据预测结果分析，3 种模型的预测误差和适用范围各有不同：GM(1,1)模型的预测误差较大，

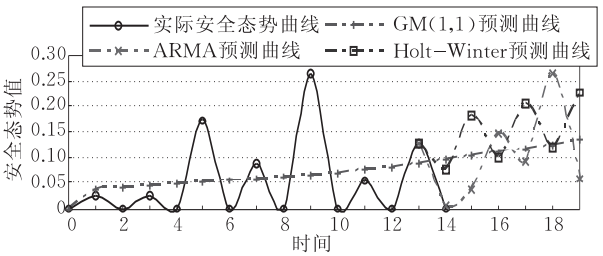

图 6 网络安全态势预测曲线图

主要反映了安全态势的总体平滑趋势，因此只适合于安全趋势平缓的大规模网络安全态势预测；ARMA 模型的预测误差较小，反映了安全态势的自相关性、总体趋势、随机性和周期性因素，因此适合于各种规模和安全趋势的网络安全态势预测；Holt-Winter 模型的预测误差最小，反映了安全态势的总体趋势、随机性、周期性和季节性等因素，因此适合于周期和季节规律较强的网络安全态势预测. 经过综合分析和比较，并结合网络安全事件自相关的特点，ARMA 模型预测误差较小，适用范围更广，反映的安全因素与网络实际情况更加相符，所以 ARMA 模型更加适合于网络安全态势预测.

以上实验结果表明：本文提出的基于日志审计和性能修正算法的网络安全态势评估模型和方法可以有效地将日志理论分析结果和实际性能参数结合在一起，分析攻击对网络的实际影响，并且通过多种预测模型给出安全态势预测结果. 本文方法与传统方法对比结果如表 2 所示，由该表可以看出本文方法安全元素更加全面，比传统方法更准确地反映了

网络的安全态势和发展趋势，方便管理员了解网络整体安全状况和趋势，采取相应防护措施，增强网络的安全性.

表 2　本文方法与传统方法对比表

评估方法	安全元素					量化	预测
	网络连接	安全事件	脆弱性	资产服务	性能信息		
Spinning Cube	√						
NVisionIP	√						
SSARE		√	√			√	√
层次化方法		√	√	√		√	
本文方法	√	√	√	√	√	√	√

6　结　论

本文对已有的网络安全态势评估方法进行了分析和比较，并分析了日志审计、安全威胁、性能改变和安全态势之间的关系，提出了一种基于日志审计与性能修正算法的网络安全态势评估模型，利用日志审计评估节点理论安全威胁，并通过性能修正算法计算节点安全态势，再利用节点服务信息计算网络安全态势，同时采用多种预测模型对网络安全态势进行预测，绘制安全态势曲线图，实现了网络安全态势的量化分析和趋势预测，并结合针对网络实例的分析进一步验证了本文提出的评估模型、量化算法和预测算法的适用性和特点.

今后的研究工作包括进一步完善网络安全态势评估模型及其量化评估方法，研究更全面的安全态势指标及其表示方法，在量化算法上突破传统的利用权重进行加权的方法，研究网络安全态势的多维可视化问题，并改进各类网络安全事件的关联分析算法，从而更加准确地预测网络的安全威胁来源.

参　考　文　献

[1] Lau S. The spinning cube of potential doom. Communications of the ACM, 2004, 47(6): 25-26

[2] Lakkaraju K, Yurcik W, Lee A J. NVisionIP: Netflow visualizations of system state for security situational awareness//Proceedings of the 2004 ACM Workshop on Visualization and Data Mining for Computer Security. Washington DC, 2004: 65-72

[3] Yin X, Yurcik W, Treaster M. VisFlowConnect: Netflow visualizations of link relationships for security situational awareness//Proceedings of the 2004 ACM Workshop on Visualization and Data Mining for Computer Security. Washington DC, 2004: 26-34

[4] Bass T. Intrusion detection systems & multisensor data fusion: Creating cyberspace situational awareness. Communications of the ACM, 2000, 43(4): 99-105

[5] D'Ambrosio B. Security situation assessment and response evaluation (SSARE)//Proceedings of the DARPA Information Survivability Conference & Exposition II. Anaheim, 2001: 387-394

[6] Yegneswaran V, Barford P, Paxson V. Using honeynets for internet situational awareness//Proceedings of the 4th Workshop on Hot Topics in Networks. Maryland, 2005

[7] Chen Xiu-Zhen, Zheng Qing-Hua, Guan Xiao-Hong, Lin Chen-Guang. Quantitative hierarchical threat evaluation model for network security. Journal of Software, 2006, 17(4): 885-897(in Chinese)
(陈秀真，郑庆华，管晓宏，林晨光. 层次化网络安全威胁态势量化评估方法. 软件学报，2006，17(4)：885-897)

[8] Deng Ju-Long. Gray Forecast and Decision-Making. Wuhan: Press of Huazhong University of Science and Technology, 1986(in Chinese)
(邓聚龙. 灰色预测与决策. 武汉：华中科技大学出版社，1986)

[9] George E P B, Gwilym M J. Time Series Analysis: Forecasting and Control. San Francisco: Holden-Day Inc., 1976

[10] Tong Ming-Rong, Xue Heng-Xin, Lin Lin. Study on the forecast of railway freight traffic volume based on Holt-Winter model. Railway Transport and Economy, 2007, 29(1): 79-81(in Chinese)
(童明容，薛恒新，林琳. 基于 Holt-Winter 模型的铁路货运量预测研究. 铁道运输与经济，2007，29(1)：79-81)

[11] Fall K, Varadhan K. The ns manual (formerly ns notes and documentation). California: UC Berkeley, LBL, USC/ISI, and Xerox PARC, 2007

WEI Yong, born in 1981, Ph. D. candidate. His research interests are network security and situational awareness.

LIAN Yi-Feng, born in 1974. Ph. D., assistant researcher. His research interests include network security and vulnerability assessment.

Background

As a result of the wider application of computer network, network security gets more and more attention. Security situational awareness is a hot topic in network security. There are many works on situational awareness modeling and quantitatively analyzing, but they have some shortcomings such as lacking for security factors, inaccurate quantitative analysis, with out forecasting, and so on. In this paper, a new network security situational awareness model based on log and performance correction algorithm is proposed. First, nodes theoretic security threat is got by log audit and the value of nodes security situational is computed by performance correction algorithm. Then the value of network security situational is computed using service information, the future threat by several prediction models is predicted, and the Security Situational Graph is drawn. This model is more effective and accurate to reflect the network security situational and its trends than traditional methods through the example. This research work is a part of research plan on security model and key technology of distributed computing. The plan is to establish security policy framework of distributed computing, research network vulnerability assessment model and security situational awareness model. The security model, vulnerability assessment model and security situational awareness model can help network administrators to comprehend the overall network security and execute some security policies to improve network security. This research work is supported by the National High Technology Research and Development Program (863 Program) of China under grant Nos. 2006AA01Z437, 2007AA01Z475, 2006AA01Z412, 2006AA01Z433.