DOI:10.19304/j.cnki.issn1000-7180.2014.11.008

网络安全态势量化模型研究

周 健[1],章 倩[2]
(1 合肥工业大学 信息与网络中心,安徽 合肥 230009;2 合肥工业大学 计算机与信息学院,安徽 合肥 230009)

摘 要:采用改进的D-S证据理论方法,将多源告警事件信息进行融合,综合网络中结点的存在的脆弱性、威胁的危害程度以及资产价值,绘制量化的网络安全态势曲线,对网络安全状态进行量化评估.通过相应的实验,表明该方法能够有效地评估网络安全态势.

关键词:网络安全;态势评估;D-S证据理论;数据融合
中图分类号:TP393 **文献标识码**:A **文章编号**:1000-7180(2014)11-0032-04

Network Security Situation Quantitative Model Study

ZHOU Jian[1], ZHANG Qian[2]
(1 Information and Network Center, Hefei University of Technology, Hefei 230009, China;
2 School of Computer and Information, Hefei University of Technology, Hefei 230009, China)

Abstract: This paper uses the improved D-S evidence theory method to fuse multi-source alarm event information, combines with vulnerability, threat and assessment of the nodes in the network, to draw quantification curve of network security situation, to quantify the network security status assessment. Through the corresponding experiment, this method can effectively evaluate the network security situation.

Key words: network security; situation assessment; D-S evidence theory; data fusion

1 引言

随着网络技术的不断发展,网络所提供的信息越来越庞大,因此,它所面临的安全问题也越来越突出.传统的安全管理手段,如入侵检测系统[1]、防火墙技术,以及网络扫描等,他们以独立的方式进行工作,之间不能相互联系,从而导致这些方法不能有效地评估当前的网路安全状态,同时存在着较大的误报率和漏报率.

态势感知(SA)最初定义是"在一定的时空范围内,认知、理解环境因素,并且对未来的发展趋势进行预测"[2].用网络态势评估网络安全状态最初是由Bass[3]等人提出来的,他提出了采用智能算法对网络态势进行评估,为后来的研究者们提供一定的研究方法,但并没有实现具体的模型.

王春雷等人提出了基于知识发现的网络安全态势感知系统[4],设计了网络安全态势建模与网络安全态势建模,该系统对于网络态势进行实时量化评估;陈秀真[5]等人提出了层次化网络安全威胁态势量化评估方法,主要从服务、主机、网络三个方面入手,结合IDS海量信息以及网络性能指标,对网络态势进行量化,得到网络态势图;文献[6]采用隐马尔科夫模型对实时网络存在的风险进行量化,但是并没有提出对多源信息的处理方法.

上述的一些方法都给网络安全态势做出了一定的评估,但是也存在了一些问题,如考虑的因素比较少,不全面;采用的数据单一,只是通过某一个IDS告警信息就做出了评估,可靠性不高;对于庞大源数

收稿日期:2014-01-13;修回日期:2014-03-07
基金项目:国家自然科学基金(60873194)

据没有进行处理而直接采用等等. 针对上述存在的问题，本文提出了网络安全态势量化模型.

2　网络安全态势建模

2.1　算法分析

在网络中，为了有一个安全的运行环境，网络管理人员部署了很多的安全设备，这就导致了获取到的数据存在着多样性. 一方面这些数据使得在评估网络安全状态时更全面；另一方面却造成了使用数据出现了不确定性、复杂性. 因此，本文算法首先对收集到的数据进行一个预处理；其次，对已处理后的数据，针对不同的设备得到的数据，采用数据融合技术得到攻击发生的可能性；再次，综合结点的脆弱性计算出攻击成功的概率；最后判断发生的威胁的威胁程度以及资产价值综合得到网络节点的态势. 算法的具体流程如图 1 所示.

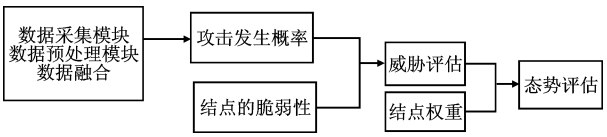

图 1　网络结点态势建立流程图

2.2　数据预处理

对于网络中产生的各种攻击行为，不同的安全设备接收到的数据各异，对此，这里采用统一格式，用一个多元数组来表示：T＝{time，type，srcIP，desIP，srcPort，Protocal，valus}，time 表示警告发生的时间，type 表示警告类型，srcIP、desIP 分别表示源 IP 地址与目的 IP 地址，srcPort、desPort 分别表示源端口号与目的端口号，Protocol 表示协议类型，valus 表示攻击所依赖的漏洞信息. 这些属性均可以从获取的数据包中提取，这样将会减少所需要处理的数据的容量.

对于上述所处理后的数据再进一步做如下处理：(1)对于任意的事件 T，如果不属于事先定义的告警事件，则删除；(2)对于某个告警事件 T，如果判断其为告警事件的关键属性空缺，则删除；(3)对于任意的某些告警事件，如果他们上述的属性值完全相同，规定在一定时间阀值内的事件进行合并，这样处理的好处是，假设发生 DDOS 攻击事件，则可以很大程度上减少要处理的告警事件. 经过上述几个步骤，可以使得管理人员避免处理海量的告警信息.

2.3　攻击事件融合

由于检测到的海量告警信息误报率、漏报率高，同时存在着不确定性. 所以采用信息融合技术进行融合，可以提高告警信息的可信度. 对于每个结点定义一个多元祖 N＝{Id，St，Service，Valus，Asset}，Id 表示结点的编号，St 表示攻击事件发生的可信度，Service 表示结点提供的服务，Valus 表示结点存在的脆弱性，Asset 表示结点的资产价值，主要由结点提供的 Service 以及结点在网络中所处的位置来确定. 这里 St 的获取需要融合多源告警事件加以确认，这里采用 D-S 证据理论加以融合，证据理论是定义在一个识别框架 Θ 上，下面给出 D-S 证据理论的相关定义.

定义 1　（信任分配函数 m）设 Θ 为一个识别框架且 Θ 由有限的互斥元素组成，在幂集 2^{Θ} 上定义信任分配函数：$m:2^{\Theta}\to[0,1]$，且满足条件 $m(\varnothing)=0$，$\sum_{e\subseteq\Theta}m(e)=1$

其中信任分配函数根据设备对检测到所得到的的数据来确定的，不同的设备各自根据检测数据确定有一定的差异. 为了提高事件检测的可信度，降低误报率，需要融合不同的设备提供的证据，下面给出多个证据合成的规则.

定义 2　（多个证据的组合规则）对于信任函数 $m_1(e_1)$，$m_2(e_2)\cdots m_n(e_n)$，则 $m_1(e_1)$，$m_2(e_2)\cdots m_n(e_n)$ 的组合的信任分配函数：

$$m(e)=\begin{cases}0,\text{若 } e=\varnothing\\ K\sum\limits_{e_1\cap e_2\cap\cdots\cap e_n=e}m_1(e_1)m_2(e_2)\cdots m(e_n),\text{若 } e\neq\varnothing\end{cases}.$$

式中，K 为规范数且 K 的值为 $K=(1-\sum\limits_{e_1\cap e_2\cap\cdots e_n=\varnothing}m_1(e_1)m(e_2)\cdots m(e_n))^{-1}$.

对于定义 2 的多个证据的组合规则，容易产生一些悖论，从而得到一些与实际情况不符合的结论，对此，本文引用文献[7]提供的改进的多个证据的合成规则，根据本文的需要加以一定的处理，定义如下：

定义 3　（改进的多个证据合成规则）

$$m(e)=\begin{cases}0,\text{若 } e=\varnothing\\ \prod\limits_{i=1}^{n}m_i(e)+f(e),\text{若 } e\neq\varnothing\end{cases},$$

其中，

$$f(e)=K\times q(e),q(e)=\frac{1}{n}\sum_{i=1}^{n}m_i(e)$$

$$K=1-\sum_{\cap e_i=e}\prod_{i=1}^{n}m_i(e).$$

对于某个结点上的报警信息，无论是入侵检测

告警信息还是防火墙日志等，检测到的结果归结起来只有两种：攻击成功或者攻击不成功，因此设识别框架 $U=\{Y,N\}$，其中 Y 表示攻击发生，N 表示攻击没有发生，则 $2^{\Theta}=\{\varnothing,Y,N,H\}$，$\varnothing$ 表示不可能发生的事件 $m(\varnothing)=0$；H 表示攻击可能发生也有可能没有发生，由它的定义容易得到 $H=Y\cap N=\varnothing$，则 $m(H)=0$；$m(Y)=p(Y)$ 表示攻击发生的概率；$m(N)=p(N)$ 表示攻击没有发生的概率. 至此得到单个源数据的攻击是否发生的概率，然后根据定义 3 融合成不同数据源的攻击事件的概率 St. 通过数据融合，一方面可以很大程度上减少原来告警事件的数量，使得网络管理人员摆脱处理 G 数量级的告警事件，另一方面，可以解决告警事件中存在的虚报警以及漏报警问题，这样可以有效的提高告警事件的可靠性，降低误报率、漏报率.

2.4 安全态势融合

综合各种检测结果得到的 St 表明了攻击发生的可信度，但是计算安全事件发生的可能性还需要结合结点存在的脆弱行，这里脆弱性主要是指容易被攻击利用的结点存在的漏洞信息，包括有协议漏洞、系统漏洞以及服务漏洞. 即：(安全事件发生的可能性) $=L$(攻击发生的可信度，脆弱性) $=L(St,valus)$，此外，安全态势的计算还要考虑其他方面的因素，包括有资产信息，威胁造成危害程度等. 综上所述，对于任意结点某个时间段 t 得到安全态势的算法如下：

(1) 利用 D-S 证据理论融合技术得到攻击发生的可信度 $P(St)$；

(2) 利用漏洞扫描技术和告警中的提供的漏洞信息，判断结点是否存在被攻击利用的漏洞信息，如果有，则 $P(valus)=0.9$，否则 $P(valus)=0.1$，这里取 $P(valus)=0.9$，$P(valus)=0.1$ 而不是 $P(valus)=1$，$P(valus)=0$ 主要考虑攻击发生但没有攻击所需的漏洞信息时，大量的无效攻击也会对结点造成严重影响，这一点可以从后面的实验结果中得出. 至此得到安全事件发生的可能性

$$L(St,Valus)=P(St)\ P(Valus).$$

(3) 计算安全事件对结点造成的损失 R，R 的取值与威胁值 M 相关，不同的威胁造成的影响不同：

$$R(\text{安全事件发生的可能性，威胁值})=R(L(L(St,Valus),M)=P(St)\ P(Valus)M$$

(4) 对于网络态势评估，需要考虑结点资产价值 A 的计算，网络中的结点主要考虑的有主机、服务器、路由器等. 不同的结点，由于其位置和提供的服务不同，它的重要程度也不同，这里以资产价值 A 表示，所以 t 时段内单个结点 N 安全态势：

$$S_N=LMA=P(St)P(Valus)MA.$$

若结点 N 在 t 时段内遭受 n 个攻击，则结点 N 的安全态势为 n 个 S_N 的累加，即

$$S=\sum_{i=1}^{n}S_{N_i},$$

n 为攻击类型个数，S_{N_i} 为单个攻击造成的影响.

在得到各个结点安全态势的情况下，可通过节点安全态势融合得到整个网络的 t 时段内安全态势 SA，SA 的计算方法如下：

$$SA=\sum_{i=1}^{n}S_i,$$

n 为结点的个数，S_i 为单个节点的安全态势.

至此，整个过程主要是经过攻击事件融合得到攻击发生的可能性，融合各种要素得到结点安全态势，融合各个节点的安全态势得到网络安全态势 SA，其中 SA 反映了整个网络的安全状态，可以根据不同的时段 t 绘制网络的安全态势走向，从而对网络进行监控.

3 实验仿真与结果

本文采用 Lincoln 实验室提供的实际 IDS 数据集 DARPA 2000[8] 作为实验数据集，这个数据集是通过监控四个 C 类子网得到的，包括两个攻击场景 LLDOS 1.0 和 LLDOS 2.0.2，均含有 5 个攻击步骤，第一个场景包括 IPsweep 寻找活跃主机、寻找存在 sadmind 漏洞的主机、利用漏洞入侵主机、安装 DDOS 木马软件、利用被控制的主机发动 DDOS 攻击；第二个场景的五个阶段包括 DNS HINFO queries、通过 sadmind 漏洞信息入侵主机、利用 DDOS 软件入侵更多主机、安装 DDOS 攻击程序、发动 DDOS 攻击.

对于上述两个数据集，利用 tcpreplay 在实验室组成的局域网中回放，利用原有的数据集提供的信息，构造相应的配置以及安装 snort 等多种插件. 对每个数据集的 5 个步骤，分开进行回放. 根据 snort 抓获的告警事件以及原有数据提供的边界网络告警事件和内部网络告警事件，按照本文第二部分介绍的方法进行数据融合以及态势计算. 由于主机数目比较多，10 个步骤都是进行相同的操作，下面具体给出第一阶段关键主机的分析结果. 由第二部分介绍的方法，需要知道结点的漏洞信息，而第一部分主要进行的是 IPsweep 寻找活跃主机，主要由地址为 202.77.162.213 进行扫描，发送大量的 ICMP 探测报文，利用的主要漏洞信息为 ICMP 配置不当. 具体

分析结果如表 1、2 所示所示.

表 1　关键主机漏洞信息

主机地址	ICMP 配置漏洞	RPC 配置漏洞	Sadmind 漏洞	SYN flood
172.16.115.20	Y	Y	Y	
172.16.113.105	Y			
172.16.112.10	Y	Y	Y	
172.16.112.50	Y	Y	Y	
172.16.112.100	Y			
202.77.116.213				
131.84.1.31				Y

表 2　部分结点的态势评估结果

主机地址	$P(St)$	$P(Valus)$	M	A	S
172.16.115.20	0.917	0.9	2	7	11.55
172.16.113.105	0.333	0.9	1	1	0.30
172.16.112.10	0.917	0.9	2	7	11.55
172.16.112.50	0.917	0.9	2	7	11.55
172.16.112.100	0.667	0.9	2	4	4.80
202.77.116.213	0.917	0.1	3	3	11.00

表 1 显示了关键主机的漏洞信息，其中 Y 表示存在相关方面的漏洞信息，空白表示不存在相关方面的漏洞. 表 2 所示为部分重要结点的态势融合，其中 P(St)通过 D-S 证据理论融合多源告警信息得到的，对于结点 172.16.115.20 的安全态势计算方法如下：

$$S = \sum_{i=1}^{n} S_{N_i} = \sum_{i=1}^{n} LMA = \sum_{i=1}^{n} P(St)P(Valus)MA = 11.55.$$

通过表格可以看出 172.116.115.20，172.16.112.10，172.16.112.50 以及 202.77.116.213 这四个主机在第一步骤中的态势比较高，其中前三个结点在多个告警信息中提示告警，又存在攻击必备的漏洞信息，结合结点在网络中的作用确定资产的值，主要根据文献[5]的判断方法来确定，从而使得综合后的态势值比较大，提醒网络管理人员要注意；而最后一个结点，虽然发生了攻击，但是没有漏洞信息，造成最终态势比较高的原因是因为发送了大量的无效攻击，造成整体态势比较高，同样引起管理人员注意.

由于该实验主要是攻击 172.16.115.20，172.16.112.10，172.16.112.50 以及最后对 131.84.1.31 发动 DDOS 攻击，下面给出这四个结点在这 10 个阶段的安全态势走向，如图 2 所示.

图 2 中描述了结点在 10 个步骤中的态势，描述了结点的态势走向，结点的态势值越大，则说此结点的安全态势值越大，存在的风险越大. 该图反映了网络中结点的安全态势状态，能够对网络管理人员采取的措施提供很好的建议，采取积极的措施.

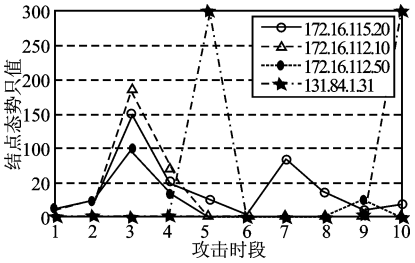

图 2　部分结点安全态势曲线图

4　结束语

本文提出了网络安全态势量化评估模型. 实验结果表明，该模型能够很好的量化网络安全状态，

下一步工作主要是对本文提出的系统进一步完善；研究统一的指标体系进行量化；根据已得到的网络安全态势状态对后期的态势进行预测，以达到提前预防的目的.

参考文献：

[1] 章倩，周健. 基于贝叶斯的入侵检测模型与仿真研究[J]. 微电子学与计算机，2013，30(11)：119-122.

[2] Endsley M R. Design and evaluation for situation awareness enhancement [C]// Proceeding of the 32nd Human Factors Society Annual Meeting. Santa Monica: Human Factors and Ergonomics Society, 1988: 97-101.

[3] Bass T. Intrusion detection systems & multisensor data fusion: Creating cyberspace situational awareness [J]. Communications of t he ACM, 2000,43 (4):992105.

[4] 王春雷，方兰. 基于知识发现的网络安全态势感知系统[J]. 计算机科学，2012，39(7)：16-18.

[5] 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)：885-897.

[6] 李明伟，雷杰. 一种优化的实时网络安全风险量化方法[J]. 计算机学报，2009，32(4)：793-804.

[7] 许红波，丁建江，胡伟稿. D-S 规则推广及其在飞机目标识别中应用研究[1]. 雷达与对抗，2006(1)：34-38.

[8] 2000 DARPA intrusion detection scenario specific data sets [EB/OL]. [2008-01-24]. http: // www. ll. mit. Edu/IST/ideval/data/2000/2000_data_index. html.

作者简介：

章　倩　女，(1987-)，硕士研究生. 研究方向为复杂网络安全.

周　健　男，(1960-)，博士，副教授. 研究方向为复杂网络安全