第30卷 第22期 计 算 机 工 程 2004年11月
Vol.30 №22 Computer Engineering November 2004
· 安全技术 · 文章编号：1000—3428(2004)22—0111—03 文献标识码：A 中图分类号：TP 393.08

基于贝叶斯网络的网络安全评估方法研究

刘 勃，周荷琴
（中国科学技术大学信息网络安全研究中心自动化系，合肥 230027）

摘 要：通过分析贝叶斯网络在计算机网络安全评估中的适用性，提出了一种新的网络安全评估方法。将贝叶斯网络应用于网络安全评估，建立了不依赖于安全漏洞的安全评估模型，引入虚拟主机概念解决多层网络结构的安全评估问题。结果表明，基于贝叶斯网络的网络安全评估方法能够综合考虑网络的特性（即先验信息）和环境（样本信息），减少主观偏见和噪声影响，能够缩短网络安全评估时间，而且易于实现。

关键词：贝叶斯网络；虚拟主机；网络安全；安全评估

Network Security Evaluation Method Research Based on Bayesian Net

LIU Bo, ZHOU Heqin
(Department of Automation, Research Center of Information Network Security, University of Science and Technology of China, Hefei 230027)

【Abstract】 Through the analysis of Bayesian net's applicability in computer network security evaluation, this paper presents a novel method to evaluate network security. It applies Bayesian net to the network security evaluation and builds a security evaluation model independent of security vulnerability. An idea of virtual host is put forward to resolve the problem of multiple-layers structure network security evaluation. The novel method takes network's characteristics and environment into account. Examples show it reduces the subjective bias and noise effect. The novel method reduces the time of network security evaluation, and it is also facile to achieve.

【Key words】 Bayesian net; Virtual host; Network security; Security evaluation

目前大部分网络安全评估方法从本质上来看，都是从安全漏洞的角度进行网络安全评估，通过扫描网络中是否存在某些已知漏洞，然后给出相应的评估结果和解决方案。这类方法的缺点是耗时长，占用大量带宽，干扰网络的正常运行。为了寻找更好的网络安全评估方法，不少研究者做了许多有益的工作。文献[1,2]从图论的角度对网络安全进行了量化分析，但主要不足是没有针对网络安全的实际情况对模型的粒度进行深入分析，其模型对网络安全的分析过于理想化，在实际网络特别是大型网络中难以使用。

本文引入贝叶斯网络(Bayesian net)对计算机网络进行建模，采用贝叶斯推理方法对计算机网络进行安全评估。用贝叶斯网络描述计算机网络模型时无须考虑具体安全漏洞，也无须对整个网络扫描，不会干扰网络的正常运行，只需要对模型进行训练，就能够快速有效地进行网络安全评估。贝叶斯推理方法比较成熟，进行网络安全评估的可操作性强。

1 贝叶斯网络在安全评估中的适用性分析

贝叶斯网络是概率分析和图论相结合的产物，它是一种有向图模型，用于不确定性知识的表达和推理。简单来说，贝叶斯网络表现为一个赋值的因果关系网络图。在贝叶斯网络中，原因和结果变量都用节点表示，每个节点都有自己的概率分布，它们之间用有向弧连接。贝叶斯网络的例子如图1所示。

图1中的4个节点分别代表现实中的4个事件，由于每个节点对应的随机变量是二值的(真或者假)，它们是离散的节点，因此可以将每个节点的条件概率分布(Conditional Probability Distribution, CPD)用表格的形式表示，这些表格称为CPT(Conditional Probability Table)。图1中的事件“草是湿的”(W=true)有两种可能的原因：洒水车(S=true)或者下雨(R=true)。它们之间的关系用表格表示，例如P(W=true | S=true, R=true)= 0.99，P(W=false | S=true, R=false)= 0.01。

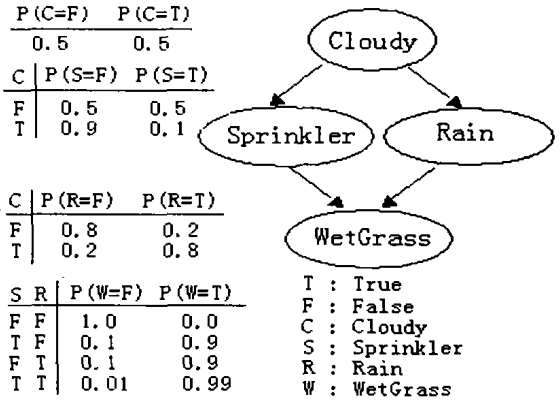

图1 4个变量的贝叶斯网络

在贝叶斯网络中，如果某个节点的概率分布独立于它的父节点，就称该节点是条件独立的。所谓父子关系是按某种固定的拓扑顺序来定义的，例如在图1中，节点S和R是节点W的父节点，节点C是节点S和R的父节点，节点C没有父节点，称为根节点。

根据概率的链式规则，图1中所有事件(节点)的联合分布可表示为：

$$P(C, S, R, W) = P(C) * P(S|C) * P(R|C, S) * P(W|C, S, R)$$

应用条件独立关系，在上式等号右边第3项中R是独立于S的，第4项中W是独立于C的，所以可重写上式为：

$$P(C, S, R, W) = P(C) * P(S|C) * P(R|C) * P(W|S, R)$$

应用条件独立关系，可以将联合概率密度表达得更简洁，节省贝叶斯网络的存储空间和训练时间。

贝叶斯网络主要功能就是进行概率推理，给定贝叶斯网络的所有事件(节点)联合概率分布，理论上能够回答所有的

作者简介：刘 勃(1977－)，男，博士生，主研方向：信息安全，数字信号处理；周荷琴，教授、博导

收稿日期：2003-09-26 E-mail: boliu@ustc.edu

推理问题。但是Cooper在文献[3]中指出，即使在条件独立的假定下，贝叶斯网络的概率推理仍是一个NP问题，不过针对一些特殊的网络(如多树型网络)仍然存在可行的算法[4]，所以在网络安全评估中引入贝叶斯网络时，必须合理构造贝叶斯网络，适当简化网络结构，以降低计算的难度。

贝叶斯网络推理主要有3种形式：(1)因果推理(causal or top-down inference)，由原因推知结论；(2)诊断推理(diagnostic or bottom-up inference)，由结论推知原因；(3)支持推理(explaining away)，提供解释以支持现象。

在建立网络安全评估的贝叶斯网络模型时，主要考虑第(1)、(2)两种概率推理。第(1)种推理用来评估网络的安全等级。在网络不安全的时候，用第(2)种推理给出最大的安全隐患。

贝叶斯网络模型实际上作了两个假定：无环假定(即假定网络图中不存在环)和静态假定(不考虑原因节点影响结果节点的滞后时间)，所以在建立网络安全评估模型的时候，必须避免环结构，这一点通过合理选取贝叶斯网络的节点，完全可以做到。至于静态假定，本文在定义贝叶斯节点的时候选取的节点都是时间无关的，所以不存在滞后问题。

2 单机安全评估模型的建立

用贝叶斯网络建立一台计算机的安全评估模型，可以分为3个步骤[5]：

(1)确定建模目标和模型中的各个要素

单机安全评估模型的目标是根据定义的指标评估单机的安全等级。按照黑客攻击的一般步骤为单机安全评估模型定义了5个指标，按照黑客攻击的一般结果为单机定义了4个安全等级，这样的定义充分考虑了影响网络安全的各个方面。

黑客攻击的步骤[6]：(a)远程信息收集、(b)对收集到的数据进行分析、(c)远程攻击、(d)本地攻击、(e)本地数据收集。如图2所示，定义的5个指标为：

(a)单机与外部网络的连接性，定义为节点A。

(b)单机可被外部收集到的信息量，定义为节点B。

(c)单机的污染度，指单机被“污染”的程度，定义为节点C。

(d)单机本地安全性，指单机的本地安全机制的可靠程度，定义为节点D。

(e)单机与内部网络的连接性，指单机与网络内部其他机器的连接性，定义为节点E。

定义单机4个安全等级为高、中、低、危险，对应的节点定义为Z。

(2)建立有向无环图并使其满足条件独立性

(3)给出所有节点的条件概率分布(CPD)

使用Bayes Net Toolbox BNT[7]辅助建模，BNT所支持的条件概率分布类型(节点类型)如图3所示[7]。

节点A、B、C、D、E是连续节点(对应连续随机变量)，可以看作是根节点(root)[8]，它是用来定义没有父节点、没有参数、能够被观察的外部输入变量。节点Z是离散节点(对应离散随机变量)，节点Z的父节点都是连续节点，所以可以用softmax函数[9]定义它的CPDCPD。借助BNT，通过训练数据可以计算出A、B、C、E、D、Z的CPD。

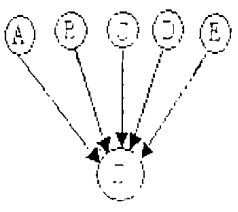

图2 单机的贝叶斯网络安全评估模型

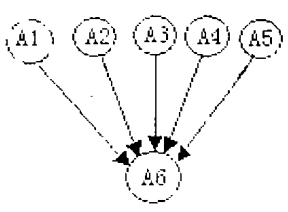

图3 网络的贝叶斯网络安全评估模型

贝叶斯网络的拓扑结构和CPD定义好以后，接着训练每一个CPD的参数(根节点不用训练)，确定了参数就确定了各个节点的CPD，也就可以对网络开始进行安全评估。本文中的贝叶斯网络的结构已知，数据完全可见，可以采用最大似然估计(Maximum Likelihood Estimation，MLE.)来计算(训练)节点的CPD参数。

3 网络安全评估模型的建立

虽然黑客攻击一般是从单机开始的，但是在评估一个网络安全的等级时，必须考虑整个网络的结构。网络中一台主机的安全性并不能代表整个网络的安全性，因为如果只是一台外围的不重要的机器被入侵，并不会直接对网络造成破坏，如果入侵的是一台重要的服务器，如Web服务器，才会给网络造成影响。和单机安全评估模型一样，分3步来建立网络安全评估模型。

(1)确定建模目标和模型中的各个要素

网络安全评估模型的目标是根据定义的指标评估网络的安全等级。如图3，我们为网络安全评估模型定义了5个指标：

(a)网络网络中被评估为“危险”等级的单机所占的比例，定义为节点A1。

(b)网络的洁净污染度，指网络中被入侵过的单机所占的比例，定义为节点A2。

(c)网络提供的服务分散程度，指提供服务最多的服务器所提供的服务(假设为a)，在整个网络对内、对外所提供的全部服务(假设为b)中的比例(a/b)，定义为节点A3。

(d)网络中操作系统类型的分散程度，指在网络中使用最多的操作系统在整个网络中所有操作系统中所占的比例，定义为节点A4。

(e)网络的内部连接性，指网络中内部连接最多的服务器所连接的内部机器占网络中所以有机器的比例，定义为节点A5。

定义网络4个安全等级为高、中、低、危险，对应的节点定义为A6。

(2)建立一个有向无环图并使其满足条件独立性

图3中A1节点用虚线表示，因为在进行网络安全评估时，对单机的安全评估可能还没有完成，这时A1节点是未知的，可以去掉节点A1，用剩下的节点完成对网络的安全评估。等单机的安全评估完成了，把A1节点加上，再对整个网络进行安全评估，将两次评估的结果进行比较，可以得到一个单机安全性对网络的安全性的影响因子，因子越小，说明网络的整体安全性对单机的安全性依赖越小。

(3)给出所有节点的条件概率分布(CPD)

将A1、A2、A3、A4、A5看作是root节点，将A6看作是softmax节点，那么图3中各个节点的CPD的训练方法和单机安全评估模型是一样的。

对于一个大型的网络，它可能由多个子网络组成，或者多个子网和多台单机混合组成，把这种网络称为具有多层网络结构。在评估这样的网络时，传统安全评估方法只是把它当成相互独立的多个网络，或者干脆就无视这种多层结构的存在。传统方法在对这类网络进行评估时，不仅耗时特别长，而且完全割裂了网络结构的关联性，此时它们的评估必然是不准确的。

为了解决多层网络的安全评估问题，引入虚拟主机的概念。首先独立地评估子网的安全性，然后将整个子网看作是一个虚拟的主机，放到整个网络中去考虑，采用相同的模型去评估整个网络，即对多层网络系统进行嵌套式的安全评估。理论上来说，这种嵌套是可以一直进行下去的。但是因

为每层评估都会存在误差，这种误差会被累加到更高层的评估中。所以在作两层以内的网络评估效果比较好，更多层次的网络评估中的误差分析将是下一步研究的方向。

4 实验分析结果

下面介绍使用贝叶斯网络进行网络安全评估的实例，其为单层结构，网络内部有3台服务器，分别是Web服务器、数据库服务器、FTP(文件传输)服务器，采用本文第2、3节给出的模型对该网络建模。

为单机安全评估模型和网络安全评估模型分别定义了一个训练集，每个训练集都有100个训练样本，如表1和表2所示，表中的所有指标都进行了归一化处理。

表1 单机安全评估模型的训练集

训练样本	P1 (指标1)	P2 (指标2)	P3 (指标3)	P4 (指标4)	P5 (指标5)	单机安全等级 (P)
1	0.9	0.5	0.1	0.5	0.4	中
2	1	0.3	0.3	0.5	0.5	低
…		…	…	…	…	…
99	0.01	0.5	0.1	0.5	0.9	高
100	0.5	0.7	0.1	0.5	0.9	中

表2 网络安全评估模型的训练集

训练样本	Q1 (指标1)	Q2 (指标2)	Q3 (指标3)	Q4 (指标4)	Q5 (指标5)	网络安全等级 (Q)
1	0.1	0.4	0.4	0.4	0.1	中
2	0.2	0.8	0.8	0.2	0.6	危险
…	…	…	…	…	…	…
99	0.05	0.3	0.2	0.4	0.2	高
100	0.2	0.45	0.5	0.5	0.3	低

表3 缺少指标1的网络安全评估模型的训练集

训练样本	Q2 (指标2)	Q3 (指标3)	Q4 (指标4)	Q5 (指标5)	网络安全等级 (Q)
1	0.4	0.4	0.4	0.1	0.69
2	0.8	0.8	0.2	0.6	0.27
…	…	…	…	…	…
99	0.3	0.2	0.4	0.2	0.88
100	0.45	0.5	0.5	0.3	0.58

如果要计算单机安全性对网络的安全性的影响因子，在设计训练集的时候就必须把安全量化，比如让0~0.249 9对应“危险”等级，0.25~0.499 9对应“低”等级，0.5~0.749 9对应“中”等级，0.75~0.999 9对应“高”等级，如表3所示。作了以上修改后，就可以分别计算出考虑指标1(Q1)和不考虑指标1(Q1)时的网络安全系数(量化值)full_Q和lack_Q，对于本例的结果是full_Q=0.79，lack_Q=0.89。所以本例中的网络单机安全性对网络的安全性的影响因子为：

$$\lambda = lack\_Q / full\_Q = 0.89 / 0.79 = 1.126$$

如果在某个现实网络环境中影响因子λ在一个局部范围内只做很小的波动，比如假定在校园网内λ在1.12~1.13之间波动，就可以把λ看作一个常数1.12，那么对于一个校园网内的子网(如本例中的网络)的安全评估可以不用对网络子网中的单机进行安全评估，就可以便能算出它的安全系数：

$$full\_Q = lack\_Q * \lambda$$

这将大大简化网络安全评估的过程，比起传统方法将节省大量时间，而且准确性有保证。

5 结束语

将贝叶斯网络应用到网络安全评估中，可以综合先验知识和样本知识，减少使用先验知识带来的主观偏见和使用样本知识带来的噪声影响，提高了网络安全评估的准确性。通过建立不依赖于安全漏洞的安全评估模型，使得贝叶斯网络安全评估模型只需使用不同的训练样本就能适用于不同的网络结构，提高了建模的速度，缩短了评估时间。在网络结构大致相同的情况下，对多个网络可以只用一个训练样本集，只训练一次就评估多个网络，提高了网络评估效率。

参考文献

1 Ortalo R, Deswarte Y, Kaaiche M. *Experimenting with Quantitative* Evaluation Tools for Monitoring Operational Security [R]. LAAS Report 96369, 1997-01

2 Dacier M, Deswarte Y, Kaaiche M. Quantitative Assessment of Operational Security: Models and Tools [R]. LAAS Research Report 96493, 1996-05

3 Cooper G. Computational Complexity of Probabilistic Inference Using Bayesian Belief Networks (Research Note) [J]. Artificial Intelligence, 1990, 42 (2/3) : 393-405

4 Russe L S,Niorvig P.Artificial Intelligence:A Modem Approach [[M]. Englewood Cliffs, NI : Prentice Hall, 1995

5 Heckerman D. *A Tutorial on Learning with Bayesian Networks* [R]. Microsoft Research Report MSR-TR-95-06, 1995-03

6 Spencer W. Network Security Assessment White Paper [R]. Network System Architects, Inc., http://www.nsai.net, 2000-04-20

7 Murphy K P. The Bayes Net Toolbox for Matlab [R]. Technical Report,U.C. Berkeley, Dept. Comp. Sci, CA, 94720-1776, 2001-10-09

8 Murphy K P. How to Use the Bayes Net Toolbox [R]. http://www. csberkeley. edu/~murphyk/Bayes/usage. html#root

☆☆

(上接第73页)

4 总结和下一步工作

本文对开放式运行平台ORP的代码进行了修改，使之能够获得SUN的CLDC库的支持，并能够运行简单的J2ME程序。实验证明ORP经过适当的修改，完全具备运行J2ME程序的能力。但是因为ORP没有针对内存和计算能力进行优化，所以还不能作为正式的J2ME运行平台。下一步的工作将对ORP进行进一步的修改，使之可以获得SUN的MIDP库的支持。

参考文献

1 卢 军.J2ME手机.PDA程序设计.北京:中国铁道出版社,2002-09

2 http://www.gnu.org/software/classpath/

3 Topley K. J2ME in a Nutshell. OReilly，2002-03

4 Liang Sheng .The Java TM Native Interface: Programmers Guide and Specidfication .ADDISON WESLEY，1999

5 Ciemiak M,Eng M.The Open Runtime Platform:A Flexible High-performance Managed Runtime Environment. Intel Technology Journal,2003-02