基于贝叶斯网络的入侵容忍系统*)

秦华旺　戴跃伟　王执铨
（南京理工大学自动化学院　南京 210094）

摘　要　提出一种基于贝叶斯网络的入侵容忍系统，给出系统的运行流程。用进程特性向量来表示一个具体的进程，并对进程特性进行具体的分类。提出利用贝叶斯网络模型来描述进程的运行过程，给出基于贝叶斯网络推理的进程类型概率值的计算公式，构造了用于确定进程危险程度的危险函数，并用实例说明了对入侵进程的具体识别过程。

关键词　入侵容忍，贝叶斯网络，信息安全

Intrusion Tolerant System Based on Bayesian Networks

QIN Hua-wang　DAI Yue-wei　WANG Zhi-quan
(School of Automatization, Nanjing University of Science & Technology, Nanjing 210094, China)

Abstract　A kind of intrusion tolerant system based on Bayesian networks is proposed in this paper. The work flow of the system is also given. The process is denoted by process character vector which is classified concretely. The work flow of process is described by Bayesian networks model. The calculative expressions of probability about process type are given, and which are based on the Bayesian networks illation. The danger function used to confirm the degree of danger about process is built. An example is given to explain the course of distinguishing intruded process.

Keywords　Intrusion tolerance, Bayesian networks, Information security

1　引言

随着网络技术的高速发展，网络安全问题也日益受到人们的关注。传统的安全措施如入侵检测、防火墙等通常只关注一些已知的和定义好的攻击，对于新出现的攻击方法，往往显得无能为力。针对该问题，诞生了一种新的网络安全方法——入侵容忍。与传统安全技术的思路不同，入侵容忍关注的是系统在已经遭到入侵的情况下，如何屏蔽或遏制入侵，从而尽量使系统能够继续安全运行。

近几年来，随着分布式密码学研究，特别是秘密共享和门限密码学方面研究的逐渐成熟与完善，再加上分布式网络系统的大量应用，入侵容忍的理论、方法与应用逐渐成为信息安全业内人士关注的一个焦点。国内外学术界对入侵容忍的相关问题展开了大量研究，并取得了许多丰富的成果，如：基于多状态自动切换的入侵容忍系统模型建立[4]、基于门限密码和秘密共享的入侵容忍系统[14]、基于大数表决或拜占庭一致性的入侵容忍系统[11,13]、具有自动清除或自动恢复的入侵容忍系统[1,8]、基于可信任时间计算基准的入侵容忍系统[1,2]、基于角色访问控制的入侵容忍系统[10]、具有入侵容忍特性的数据库系统[6]等，同时，对入侵容忍系统的各种安全属性也进行了一些定量研究[7,15]。

对于一个入侵容忍系统而言，只有在其识别到入侵行为以后，才会采取相应的安全措施，防止入侵的蔓延，因此，如何尽快并且尽可能准确地发现入侵行为，是入侵容忍系统设计中的关键要素。然而，由于入侵行为的诸多不确定性，任何一个入侵容忍系统都不可能百分之百正确地识别出所有的入侵行为，即必然会存在一定的漏报率和虚警率，所以，对于一次入侵行为的识别，只能从概率上进行判断，贝叶斯网络正是一种应用广泛的基于概率的不确定性推理方法，鉴于这两者相似性的启发，本文提出了一种利用贝叶斯网络从概率上识别入侵行为的入侵容忍系统。

2　基于贝叶斯网络的入侵容忍系统

2.1　系统概述

在网络系统中，入侵者的入侵行为是通过运行其相应的入侵进程来实现的，本文将网络系统中的进程分为四种类型：正常进程、轻微异常进程、严重异常进程、入侵进程。对于一个运行中的进程，利用贝叶斯网络得出该进程对应于上述每一种进程类型的概率值，通过将这些概率值代入危险函数，求出该进程对应的危险函数值，然后入侵容忍系统再根据该危险函数值采取相应的安全措施，如：监视进程运行、严格控制进程运行、清除进程等。系统的运行流程如图1所示。

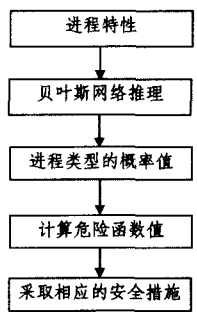

图1　系统运行流程图

设进程 x 对应于正常进程的概率值为 P_n，对应于轻微异常进程的概率值为 P_l，对应于严重异常进程的概率值为 P_s，对应于入侵进程的概率值为 P_i，则进程 x 的危险函数为 $f(x)=K_nP_n+K_lP_l+K_sP_s+K_iP_i$，其中 K_n，K_l，K_s，K_i 为权系数，用于调整不同的进程类型对危险函数值的影响程度，可

*)基金项目：国家自然科学基金资助项目(60374066)。秦华旺　博士研究生，讲师，研究方向为入侵容忍；戴跃伟　教授，研究方向为信息安全、数字水印；王执铨　教授，博导，研究方向为信息安全、复杂系统、容错。

以根据经验数据确定，$K_n>K_l>K_s>K_i>0$。危险函数值越大，该进程为入侵进程的概率就越大，即该进程越危险。

2.2 进程特性描述

为了能够描述一个进程在运行时的具体特性，本文用向量 $x=(F,O,K,M,R)$ 来定义一个具体的进程，该向量中的各个元素解释如下：

(1)F 表示该进程的 IP 地址对系统的访问频率，并定义：F_1=很低，F_2=较低，F_3=一般，F_4=较高，F_5=很高。

(2)O 表示该进程的操作对象类型，并定义：O_1=普通文件，O_2=应用文件，O_3=系统文件。

(3)K 表示该进程的类型，并定义：K_1=正常，K_2=轻微异常，K_3=严重异常，K_4=入侵。

(4)M 表示该进程的操作类型，并定义：M_1=只读，M_2=修改，M_3=删除，M_4=创建。

(5)R 表示该进程所占用的系统资源，并定义：R_1=很少，R_2=较少，R_3=一般，R_4=较多，R_5=很多。

入侵容忍系统对进程的判断过程为：首先确定一个进程的 F,O,M,R 特性，然后利用贝叶斯网络推理，计算出该进程的 K 特性对应于 K_1,K_2,K_3,K_4 的概率值，即该进程对应于每一种进程类型的概率值，最后根据所得的四个概率值计算该进程的危险函数值。

2.3 贝叶斯网络模型

贝叶斯网络是一种模拟人类推理过程中因果关系的不确定性处理模型，其网络拓扑结构是一个有向无环图，该图由节点和弧段构成，节点代表相应的事件或变量，弧段代表节点之间的因果关系或概率关系，弧段是有向的，不构成回路。

本文用图 2 所示的贝叶斯网络模型来描述进程的运行，图中的节点表示进程的具体特性，节点之间的有向弧段表示各个特性之间的条件概率。该贝叶斯网络图的解释如下：

(1)$F=\{F_1,F_2,F_3,F_4,F_5\}$，$O=\{O_1,O_2,O_3\}$，$K=\{K_1,K_2,K_3,K_4\}$，$M=\{M_1,M_2,M_3,M_4\}$，$R=\{R_1,R_2,R_3,R_4,R_5\}$，分别表示进程的访问频率、操作对象类型、进程类型、操作类型、占用系统资源等特性。

(2)进程的操作对象类型 O 和访问频率 F 在概率上是独立的。

(3)进程的操作类型 M 和所占用的系统资源 R 在概率上是独立的。

(4)进程的操作对象类型 O 以及访问频率 F 对进程类型 K 有因果影响。

(5)进程类型 K 对进程的操作类型 M 以及所占用的系统资源 R 有因果影响。

采用贝叶斯网络的目的，就是要根据 F,O,M,R 的值以及它们之间的条件概率，计算出 K 对应于 K_1,K_2,K_3,K_4 的概率值。

根据贝叶斯原理可得：

$$P(F,O,K,M,R)=P(M|K)P(R|K)P(K|F,O)P(F)P(O) \quad (1)$$

由于：

$$P(K_i)=P(F,O,K_i,M,R)/\sum_{j=1}^{4}P(F,O,K_j,M,R) \quad (2)$$

所以进程类型 K 的概率计算公式为：

$$P(K_i)=P(M|K_i)P(R|K_i)P(K_i|F,O)P(F)P(O)/\sum_{j=1}^{4}P(M|K_j)P(R|K_j)P(K_j|F,O)P(F)P(O) \quad (3)$$

即：

$$P(K_i)=P(M|K_i)P(R|K_i)P(K_i|F,O)/\sum_{j=1}^{4}P(M|K_j)P(R|K_j)P(K_j|F,O),i=1,2,3,4 \quad (4)$$

公式中的条件概率 $P(M|K)$，$P(R|K)$，$P(K|F,O)$ 可以根据经验和实验统计数据得到。

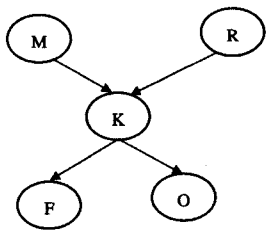

图 2 进程运行的贝叶斯网络图

3 实例分析

设某进程 $x=(F_5,O_2,K,M_3,R_4)$，即该进程的访问频率为很高，操作对象类型为应用文件，操作类型为删除，所占用的系统资源为较多，其中部分概率值设为：

$P(K_1|F_5,O_2)=0.1$，$P(K_2|F_5,O_2)=0.2$，

$P(K_3|F_5,O_2)=0.3$，$P(K_4|F_5,O_2)=0.4$；

$P(R_4|K_1)=0.1$，$P(R_4|K_2)=0.15$，$P(R_4|K_3)=0.2$，

$P(R_4|K_4)=0.25$；

$P(M_3|K_1)=0.15$，$P(M_3|K_2)=0.15$，$P(M_3|K_3)=0.2$，$P(M_3|K_4)=0.2$。

则根据上述的进程类型概率值计算公式可得：

$P(K_1)=P(M_3|K_1)P(R_4|K_1)P(K_1|F_5,O_2)/\sum_{j=1}^{4}P(M_3|K_j)P(R_4|K_j)P(K_j|F_5,O_2)=0.0395$；

$P(K_2)=P(M_3|K_2)P(R_4|K_2)P(K_2|F_5,O_2)/\sum_{j=1}^{4}P(M_3|K_j)P(R_4|K_j)P(K_j|F_5,O_2)=0.1184$；

$P(K_3)=P(M_3|K_3)P(R_4|K_3)P(K_3|F_5,O_2)/\sum_{j=1}^{4}P(M_3|K_j)P(R_4|K_j)P(K_j|F_5,O_2)=0.3158$；

$P(K_4)=P(M_3|K_4)P(R_4|K_4)P(K_4|F_5,O_2)/\sum_{j=1}^{4}P(M_3|K_j)P(R_4|K_j)P(K_j|F_5,O_2)=0.5263$。

则该进程 x 的危险函数值为 $f(x)=K_nP_n+K_lP_l+K_sP_s+K_iP_i$，所以，$f(x)=K_nP(K_1)+K_lP(K_2)+K_sP(K_3)+K_iP(K_4)$，即，$f(x)=0.0395K_n+0.1184K_l+0.3158K_s+0.5263K_i$。

由计算结果可知，该进程 x 的危险函数值较大，即该进程为入侵进程的可能性较大，因此，系统需要采取如严格控制进程运行或清除进程等较为保险的安全措施。

结束语 由于入侵容忍系统中入侵行为的不确定性，因此只能从概率上识别入侵行为，鉴于贝叶斯网络是一种应用广泛的基于概率的不确定性推理方法，本文提出了一种基于贝叶斯网络的入侵容忍系统。本文用详细分类的进程特性来表示一个具体的进程，并利用贝叶斯网络模型来描述进程的运行过程，基于此网络模型，给出了进程类型概率值的计算公式，以及确定进程危险程度的危险函数，并用实例说明了对入侵进程的具体识别过程。

参 考 文 献

[1] Castro M, Liskov B. Proactive recovery in a Byzantine-fault-tolerant system [J]// 4th Symp. on Operating Systems Design and Impl. 2222: 273-288

[2] Neves NF, Veríssimo P. Complete specification of APIs and Protocols for the MAFTIA middleware. Project MAFTIA IST-1999-11583 deliverable D9, July 2002

[3] Kreidl O P, Frazier T M. Feedback control applied to survivability: a host-based automatic defense system. Reliability, IEEE Transactions on, 2004, 53(1): 148-166

[4] Goseva-Popstojanova K, Wang F, Wang R. Characterizing intru-

sion tolerant Systems using a state transition model// DARPA Information Survivability Conference and Exposition. vol. 2, 2001:211-221
[5] Wu T, Malkin M, Boneh D. Building intrusion tolerant applications//DARPA Information Survivability Conference and Exposition. 2000. DISCEX '00. Volume 1, Jan. 2000:74-87
[6] Liu P. Architecture for Intrusion Tolerant Database Systems// Computer Security Applications Conference, 2002. Proceedings. 18th Annual. Dec. 2002:311-320
[7] Stroud R, Welch I, Warne J. A qualitative analysis of the intrusion-tolerance capabilities of the MAFTIA architecture// Dependable Systems and Networks, 2004 International Conference. June-July 2004:453-461
[8] Huang Y, Arsenault D, Sood A. Incorruptible system self-cleansing for intrusion tolerance// Performance, Computing, and Communications Conference, 2006. IPCCC 2006. 25th IEEE International. April 2006:4
[9] 郭渊博,马建峰.入侵容忍的国内外研究现状及所存在的问题分析.信息安全与保密通信,2005(7)
[10] 彭文灵,王丽娜,张焕国,等.基于角色访问控制的入侵容忍机制研究.电子学报,2005,33(1):91-95
[11] 俞艳苹,郭渊博,马建峰.基于自适应大数表决机制的容忍入侵模型.系统工程与电子技术,2005,27(6):1098-1101
[12] 荆继武,冯登国.一种入侵容忍的CA方案[J].软件学报,2002,13(8):1417-1422
[13] 邹立新,丁建立.基于拜占庭协议的入侵容忍系统模型设计.计算机工程,2005,31(增刊):88-90
[14] 季称利,杨晓元,胡予濮,等.二方共享与t_n门限方案相结合的容侵CA方案.计算机工程,2005,31(21):138-142
[15] 殷丽华,方滨兴.入侵容忍系统的安全属性分析.计算机学报,2006,29(8):1505-1512
[16] 李伟生,王宝树.基于贝叶斯网络的态势评估.系统工程与电子技术,2003,25(4):480-483
[17] 黄光球,孙周军,刘兆明.基于贝叶斯置信网的日志服务系统容侵方法研究.微电子学与计算机,2006,23(12):53-60
[18] 陈健,李忠民,王永明,等.基于贝叶斯网络的装备部件战斗损伤评估模型.系统工程与电子技术,2007,29(2):329-332

(上接第77页)

点数目增加时,导致层次的不断扩大,路径长度增长速度在一定比例时要高于NICE,这是为了节省带宽和提高效率而作的牺牲,可能造成了节点接收数据的延迟,但是这也只对组员较多的情况。

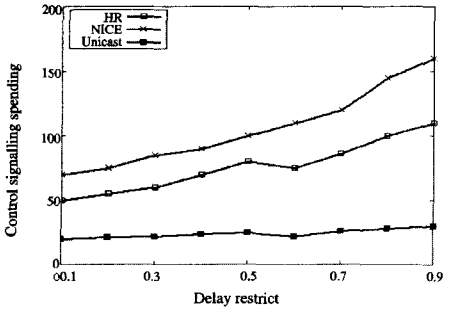

图7 延迟和控制开销

图7表示的是多播组成员数量为128时不同延时约束下平均控制信令开销,对于NICE和HR来说,虽然在约束增加时,平均开销都有所增加,但是从图中我们可以看出HR的性能总要优于NICE,尽管实际的网络中协议的寻径还受网络带宽、延时抖动和包丢失率等的影响。

结束语 多播是一种重要的组通信模式。目前的Internet中,由于IP多播部署的问题,应用层多播(ALM)逐渐得到了广泛应用。文中提出了一种分层环状新型多播模型,实现了分段传输和有序转发,避免了资源在网络中无限制的相互发送和接收,节省了带宽,重点解决了在带宽很小的情况下如何提高多播机制系统性能问题。

模型采用有序泛洪算法,使数据在层间单向传播,在层内双向传输,传输速率要比NICE快,且只占用固定的带宽。同时由于有序传播,使得数据出现差错时,易于恢复,下层节点只要向上一层发送重发请求就可以了。

采用动态规划方法,使系统结构随时间增长不断趋于合理,同时消除了分层结构中的leader节点,将计算工作放在终端执行,使系统更加稳定。

在安全性方面由于双环结构的自愈性而使系统更加健壮,而消除了leader节点也使得系统不再受leader丢失的影响。

通过和NICE的比较以及仿真结果可知HR模型在性能方面优于以往的多播模型。下一步工作的重点在于动态规划过程中的优先权公式参数选择以及在数据传播过程中的内容安全和节点安全问题,另外应用层多播的差错恢复也是需要认真考虑的一个方面。

参 考 文 献

[1] Banerjee S, Bhattacharjee B, Kommareddy C. Scalable Application Layer Multicast// Proceedings of ACM Sigcomm. Aug. 2002:205-217
[2] Chu Y-H, Rao S G, Zhang H. A Case for End System Multicast //Proceedings of ACM SIGMETRICS. 2002, 20(18):1456-1471
[3] Stoica I, Morris R, Karger D, et al. Chord: a scalable peer-to-peer lookup protocol for Internet applications// Proceedings of ACM Sigcomm. 2003, 11(1):17-32
[4] Ratnasamy S, Francis P, Handley M, et al. A scalable content-addressable network. Ph. D. Thesis. University of California, Berkeley, October 2002
[5] 龙白滔,孙立峰,陈文萍,等.具有层间冗余链路的应用层组播及其性能分析[J].电子学报,2004,32(11):1844-1848
[6] 张冰,原冰,刘增基.一种新的固定速率分层组播拥塞控制协议[J].中国科学,2006,33(10):23-28
[7] Pendarakis D, Shi S, Verma D, et al. ALMI: An application level multicast infrastructure// Proc. of the 3rd Usenix Symp. on Internet Technologies and Sys. (USITS 2001). SanFrancisco, CA, Mar. 2001:49-60
[8] Chawathe Y. Scattercast: an architecture for internet broadcast distribution as an infrastructure service [D]. USA: University of California, Berkeley, 2000
[9] Sobeih A, Yurcik W, Hou J C. VRing: a case for building application-layer multicast rings (rather than trees); Modeling, Analysis, and Simulation of Computer and Telecommunications Systems, 2004:437-446
[10] Zhao J, Yang F, Zhang Q, et al. LION: Layered Overlay Multicast With Network Coding Multimedia. IEEE Transactions, 2006, 8(5):1021-1032
[11] Yiu W-P K, Chan S-H G. SOT: secure overlay tree for application layer multicast//Communications, 2004 IEEE International Conference. Volume 3, June 2004:1451-1455
[12] Yiu W-P K, Wong K-F S, Chan S-H G, et al. Lateral error recovery for media streaming in application-level multicast Multimedia. IEEE Transactions, 2006, 8(2):219-232
[13] Tian Ruixiong, Zhang Qian, Xiang Zhe, et al. Robust and efficient path diversity in application-layer multicast for video streamingCircuits and Systems for Video Technology. IEEE Transactions, 2005, 15(8):961-972
[14] Gao L T, Sze T W, Crawford D, et al. A Cross-Functional Service-Oriented Architecture to Support Real-Time Information Exchange in Emergency Medical Response Hauenstein, Engineering in Medicine and Biology Society, 2006// EMBS '06. 28th Annual International Conference of the IEEE Volume Supplement. 2006:6478-6481
[15] Jin X, Wong W-C, Chan S-H G. Serving dynamic groups in application-level multicast High Performance Switching and Routing, 2005// HPSR. 2005 Workshop. May 2005:432-436