入侵检测研究综述*)

A Survey of Intrusion Detection

刘 勇 茅 兵 谢俊元

(南京大学计算机科学与技术系 南京210093)

Abstract Intrusion Detection(ID)is one of the important technologies of computer system security and has been widely studied recently because the traditional security technologies such as AC, Audit and Encryption cant' meet with the security demands of contemporary computer system. In this paper the studies on Intrusion Detection are introduced broadly including the definitions, classifications and important techniques.

Keywords Computer system security, Intrusion detection, Anomaly detection, Misuse detection

1 引言

随着计算机和网络技术在社会生活各方面应用的普及和深入发展,计算机系统安全已成为计算机科学当前的研究热点之一。传统的计算机系统安全研究包括设计一定的安全策略、建立支持该策略的形式化安全模型,然后通过身份认证、访问控制、审计等技术使安全模型在计算机系统中得以实现,使之成为针对各种系统入侵的防御屏障。然而近年来随着系统入侵行为的程度和规模加大,安全模型理论自身的局限以及实现中存在的漏洞被逐渐暴露出来。由形式化的安全模型和传统的安全技术所构成的被动防御体系,已不足以保证计算机系统的安全。

入侵检测作为一种主动防御技术,弥补了传统安全技术的不足。当今的入侵检测系统可以对计算机主机和网络进行实时监控,分析发现可疑事件。一旦入侵行为被检测出来,系统就会采取相应的措施(如通知管理员,切断网络连接等),从而及时消除即将对系统安全产生的危害。入侵检测作为系统安全技术的重要组成部分,日益受到各国政府和学者的重视,而我国这方面的研究工作还相对较少,特别是对国外已有的研究成果没有一个系统的考察。本文对国外近年来入侵检测技术的研究状况及其成果进行总结与分析,并展望了未来入侵检测的发展趋势。

2 入侵检测概述

入侵的定义有多种,一种流行的定义是由 R.-Heady 等人提出的[1-2]:入侵,是任何企图破坏资源的完整性、保密性和可用性的行为集合。对于入侵检测,一种定义认为它是一种试图通过观察行为、安全日志或审计数据来检测针对计算机或网络入侵的技术,这种检测通过手工或专家系统软件对日志或其它网络信息进行分析来完成[2]。而更广义的说法是:识别企图侵入系统非法获得访问权限行为的过程[3]。

入侵检测的依据是系统中的异常行为,入侵检测技术要发挥作用存在一个前提:假定任何一种入侵都会引起系统运行的异常。但由于系统异常不一定是由于入侵行为引起的,就会使检测系统产生误报。同时,也有一些入侵行为并未被检测系统察觉,也会发生漏报。为了衡量入侵检测避免上述两种情况的能力,人们提出了两个重要指标,它们分别是正确率和覆盖率。

- 正确率=检测出的真正入侵行为/检测出的所有入侵行为
- 覆盖率=检测出的真正入侵行为/系统中发生的所有入侵行为

同时具有较高正确率和覆盖率的入侵检测技术是研究者们追求的目标,但实际上很难做到。为了提高检测的覆盖率,检测系统要变得更加警觉,将更多的可疑行为判断为入侵行为,而这样做又降低了检测的正确率,显然,单纯改变报警阈值无法真正获得可靠、高效的入侵检测系统。

3 入侵检测的研究成果

入侵检测的相关研究工作已进行了近二十年,早

*)本课题获国家863基金资助(863-301-6-4)。刘 勇 硕士研究生,研究方向:信息安全,安全操作系统。茅 兵 副教授,研究方向:信息安全。谢俊元 教授,研究方向:系统安全、智能操作系统。

期的研究工作由 J. Anderson 等人展开，但并未形成规模。1985年，D. Denning 和 Oakland 提出第一个实时入侵检测专家系统模型[4]，以及实时的、基于统计量分析和用户行为轮廓(Profile)的入侵检测技术。该模型是入侵检测研究领域的一块里程碑，此后大量的入侵检测系统模型开始出现，其中很多都是基于 Denning 的统计量分析理论。进入90年代以来，随着 P. Porras 和 R. Kemmerer 基于状态转换分析的入侵检测技术[5]的提出和完善，根据已知攻击模型进行入侵检测的方法成为该领域研究的另一热点。与此同时，大型系统中原用于评价系统性能和明确责任的大量审计数据开始被作为入侵检测的数据来源，这在很大程度上提高了检测力度，使入侵检测技术真正具有了实用价值。

对于目前已有入侵检测的分类有多种方法，一种比较公认的方法是根据检测所基于的原则不同，将入侵检测划分为异常检测和误用检测。

3.1 异常检测

异常检测又称为基于行为的入侵检测，这类检测方法的原则是：任何与已知正常行为(包括用户和系统)不符合的行为都是入侵行为。异常检测系统在准备阶段通过一定时间的学习为用户正常情况下的行为建立行为轮廓(Profile)，在使用阶段系统一方面通过比较用户当前行为与原先行为轮廓的偏差来检测入侵，一方面继续根据用户的正常行为来修正行为轮廓。同样，异常检测系统也可为整个计算机系统建立正常行为轮廓，如可根据整个系统单位时间内的资源消耗情况来检测 DoS 攻击。这种检测方法的优点在于可以发现未出现过的入侵行为，另外由于能为单个用户建立行为轮廓，系统可以进行有针对性的检测。

早期的异常检测系统将用户登录时间、登录失败次数、资源访问频度等一些特征量作为随机变量，并通过统计模型计算出这些随机变量的新观察值落在一定区间内的概率。常用的统计模型有：均值和标准差模型(Mean and Standard Deviation Model)、多元模型(Multivariate Model)、状态转换矩阵模型(Markov Process Model)、时间序列模型(Time Series Model)等[6]。也有系统直接根据管理员的经验指定一个阈值，一段时间内某特征量的观察值超出该阈值即认为有入侵发生。基于统计量分析的入侵检测技术具有高可移植性而且容易实现，但由于检测的正确率极低，极少被单独使用。

随着入侵检测研究的深入，人们发现基于单个用户行为轮廓的入侵检测存在一些缺点，如：1)当用户合法地改变行为时(如使用新的应用程序)系统会误认为入侵发生；2)入侵者可通过对正常行为轮廓缓慢的偏离使系统逐渐适应；3)对于新用户，系统的学习阶段何时结束不易确定，同时在该阶段难以对用户进行正常的检测。这些问题仅靠为单个用户建立行为轮廓是无法解决的，因此后来的异常检测研究重点逐渐移到了如何观察和描述整个系统的正常行为轮廓上。使用有限的特征量来描述整个系统的正常行为是困难的，而通过大量测试来生成系统的正常行为轮廓就更加困难。由于已知大部分的入侵都是通过利用系统程序缺陷来完成的，因此比较现实的办法是建立系统中重要服务程序和安全相关程序的行为轮廓，通过这些程序的当前行为来考察系统的安全状态。C. Ko 等人在分布式环境中使用规范语言来描述安全相关程序的行为轮廓，通过对 UNIX 系统中15个 suid 超级用户程序进行安全规范分析来监控整个系统的安全状态[7]。A. Kosoresow 和 S. Hofmeyr 通过对操作系统产生的系统调用记录进行分析，归纳出系统中重要服务程序(如 ftp、sendmail 等)的正常系统调用序列，作为系统的正常行为轮廓[8]，随后 H. Debar 和 A. Wespi 等人进一步分析了系统调用序列的定长模式抽取与变长模式抽取两种方法，并给出抽取算法[9~11]。

基于系统调用序列模式抽取的异常检测实现简单，但它的缺点也很突出：1)由于大型服务程序执行路线复杂，抽取的模式库无法准确概括程序所有的正常系统调用序列；2)由于丢弃了没有构成模型的系统调用序列信息，检测结果具有较大偶然性。该技术的一个改进是基于状态转换图的入侵检测模型 STGIDM，该模型保留了不能形成模式的系统调用序列以及模式是否来源于同一次抽样及是否相邻的信息，并引入状态转换图来表示正常状态下的程序执行路线。

3.2 误用检测

误用检测(Misuse Detection)又称为基于知识的入侵检测，这类检测方法的原则是：任何与已知入侵模型符合的行为都是入侵行为。它要求首先对已知的各种入侵行为建立签名，然后将当前的用户行为和系统状态与数据库中的签名进行匹配。这种检测方法的特点是检测正确率高而覆盖率偏低，另外它还有一个最大的弱点，即只能发现已知入侵行为。但是理论上的局限性并未影响误用检测的实际应用价值，由于实际情况中大部分入侵者使用的都是已知攻击方法，该技术还是可以有效抵御大部分攻击行为。同时要特别指出的是，误用检测的正确率要明显高于异常检测，这些原因使误用检测仍然引起了很多研究者的兴趣。

早期的误用检测系统是一个专家系统，构成入侵威胁的审计记录会触发相应规则。这些规则可以识别出危及系统安全的单个审计事件，也可分析出构成一个入侵过程的简单审计事件序列。IDES[12]、NADIR[13]和 W&S[14]系统中都使用了这种技术。基于规则的误

用检测缺点在于:1)入侵检测建立在对系统审计纪录的逐个匹配上,效率太低。2)规则库的创建和更新需要富有经验的专家手工进行。针对这些问题,Porras 和 Kemmerer 设计了一个误用检测工具 STAT[15]。该工具使用状态转换分析技术,不再是通过规则逐个匹配审计纪录,而是根据成功入侵行为所必需的关键步骤(如创建具有特定访问权限的文件、发特定邮件给超级用户等)进行推理。后来 Ilgun 在 UNIX 上实现了一个原型系统 USTAT[15],该原型系统使用了 SunOS4.1.1的 C2-BSM 产生的审计数据。这种检测方法最主要的优点在于匹配的不是单个审计记录,从而获得了审计记录格式独立性,同时建立新的入侵签名也较容易。

如何获取和形式化地表示知识是误用检测研究的重点之一,也是基于规则的专家系统面临的最大困难。为了解决这一问题人们提出了一些新的入侵签名表示方法。在 Garvey 和 Lunt 提出的基于模型的误用检测[16]中,系统根据入侵过程的审计记录建立抽象的高层模型,管理员通过建立入侵模型来描述入侵过程,入侵模型到审计纪录序列的转换由系统完成。这在很大程度上减轻了管理员的负担,方便了入侵过程的描述。J. Kolodner 等人提出的基于案例的误用检测[17]适合无法将知识分解为规则的场合,它只使用原先的案例记录来完成对入侵行为的检测。由于不需要根据一类问题总结出通用的模型以适合所有该类问题,使用该技术的入侵检测系统可自动获取新知识而无需专家手工干预。

将神经网络技术应用于误用检测是近年来的研究热点之一,J. Bonifacio 等人通过使用 MLP 神经网络进一步分析网络数据包与入侵签名匹配的结果,得到了较高的正确率。R. Lippmann 和 R. Cunningham 通过将关键字匹配和神经网络结合到同一入侵检测系统中,使原系统性能有了明显改进[18]。由于神经网络具有自组织、自适应、自学习的能力,可以很容易将新的系统入侵知识融合到原有系统中,使系统具有较好的适应性。神经网络应用于入侵检测的缺点在于:1)网络拓扑结构需反复大量的训练才能确定;2)阈值的原则对检测的结果有很大影响;3)入侵者可在系统学习期间逐渐改变行为特征,使入侵行为变得合法。

除了上述一些主要的入侵检测技术以外,很多入侵检测系统或系统原型还采用了其他领域的一些成熟技术,如 S. Kumar 和 E. Spafford 在误用检测系统进行模式匹配时使用了 CPA (Colored Petri Automation)[19],IBM 公司在 RTID 系统中使用数据挖掘技术处理分析大量的报警信息[22]。随着这些技术在入侵检测领域中应用的逐步成熟,入侵检测系统的性能将有很大提高。

4 入侵检测的发展趋势

从目前国外入侵检测领域的发展趋势来看,今后入侵检测的研究工作将侧重在以下几个方面:

·大规模分布式的入侵检测系统以及异构系统之间的协作和数据共享。随着分布式技术和网络技术的发展,分布式或网络环境下的入侵检测将成为未来研究的热点,其中由于异构系统间有相互配合与数据共享的要求,建立与具体审计数据结构无关的入侵描述成为当前重要的研究课题。

·入侵检测系统的自身保护。目前入侵检测面临的最大挑战就是自身的安全性,一旦系统中的入侵检测部分被入侵者控制,整个系统的安全防线将面临崩溃的危险。如何防止入侵者对入侵检测系统功能的削弱乃至破坏的研究将在很长时间内持续下去。

·入侵检测与其它安全技术的结合。目前,信息安全受到前所未有的挑战,单一的安全技术很难保证系统的真正安全。与其它安全技术的结合也将成为入侵检测技术的趋势之一,如具有人工智能特性的自适应访问控制技术、多重身份认证技术等。

参 考 文 献

1 Heady R, et al. The Architecture of a Network Level Intrusion Detection System: [Technical Report]. University of New Mexico, Department of Computer Science, August 1990

2 NSA Glossary of Terms Used in Security and Intrusion Detection. SANS Institute, 1999. Available at: http://www.sentinelsys.com/glossary.html

3 An Introduction to Computer Security: The NIST Handbook. National Institute of Standards and Technology, Technology Administration, U. S. Department of Commerce. Available at: http://seachpdf.adobe.com/proxies/1155/80/29.html

4 Denning D, et al. Requirements and model for IDES: A real-time intrusion detection expert system: [Technical Report]. CSL, SRI Int., August1985

5 Porras P, Kemmerer R. Penetrations State Transition Analysis A Rule-Based Intrusion Detection Approach. In: Proc of the 8th Annual Computer Security Applications Conf. San Antonio, Texas, Dec. 1992

6 Denngn D. An Intrusion Detectoin Model. IEEE Transactions on Software Engineering, 1987, 13(2): 222~232

7 Ko C, et al. Executoin Monitoring of Security-Critical Programs in Distributed Systems: A Specification-based Approach. In: Proc. of the IEEE Symposium on Security and Privacy, 1997

8 Kosoresow A, Hofmeyr S. Intrusion Detection via System

Call Traces. IEEE Software, 1997, 14(5): 35～41

9 Forrest S, et al. A Sense of Self for Unix Processes. In: IEEE Symposium on Security and Privacy, Oakland, 1996

10 Debar H, et al. Fixed vs. variable-Length Patterns for Detecting Suspicious Process Behavior. In: 5th European Symposium on Research in Computer Security (ESORICS'98), 1998. 1～15

11 Wespi A, et al. Audit Trail Pattern Analysis for Detecting Suspicious Process Behavior. In: Proc of the 1st Workshop on the Recent Advances in Intrusion Detection (RAID'98), Louvainla-Neuve, Belgium, Sept. 1998. 14～16

12 Lunt T. DES: An Intelligent System for Detecting Intruders. In Proc of the Symposium on Computer Security, Threat and Countermeasures, Rome, Italy, Nov. 1990

13 Hubbards B, et al Computer System Intrusion Detection. [Tech. Rep RADC-TR-90-413, Final Technical Report]. Trusted Information Systems, Inc., Dec. 1990

14 Vaccaro H, Liepins G. Detection of Anomalous Computer Session Activity. In Proc of IEEE Computer Society Symposium on Security and Privacy, Oakland, California, May, 1989. 280～289

15 Ilgun K. USTAT: A Real-time Intrusion Detection System for UNEX. In Proc of the IEEE Symposium on Research in Security and Privacy, Oakland, CA, May, 1993 16～28

16 Garvey T, Lunt T. Model-based intrusion detection In: Proc. of the 14th National Computer Security Conf. Washington DC, Oct. 1991

17 Kolodner J. Case-Based Reasoning. San Mateo, CA, USA: Morgan Kaufmann Publishers, Inc., 1993

18 Lippmann R, Cunningham R Improving Intrusion Detection Performance Using Keyword Selection and Neural Network. In: Proc. of the 2nd Workshop on the Recent Advances in Intrusion Detection (RAID'99), 1999

19 Kumar S, Spafford E. A Pattern Matching Model for Misuse Imtrusion Detection. In: Proc. of the National Computer Security Conf. Oct. 1994. 11～22

20 Manganaris S. et al A Data Mining Analysis of RTID Alarms. In: Proc of the 2nd Workshop on the Recent Advances in Intrusion Detection (RAID'99), 1999

（上接第59页）

取相应的易抽取特征作为探测依据，来降低算法的计算复杂性。我们的实验表明，该算法是完全有效可行的，虽然跳过大量的帧，却保持有较高的查全率。另外，使用不同编码类型帧中含有的不同特点的原始特征信息，同样获得了较高的正确率。具有约为其它压缩域上探测算法2.5～5.2倍的探测速度是本算法的重要特色。

算法中采用宏块类型作为P帧、B帧的特征。实验表明对于P帧、B帧，该特征对于探测渐变镜头过渡类型的效果不好，但对于切变这种最常用的镜头过渡类型具有很好的探测效果。新闻类视频素材中的镜头过渡方式主要是切变类型，因此该算法适合于象新闻这样以切变镜头过渡类型为主的视频节目。我们对CCTV新闻联播的统计数据表明仍存在3%-5%渐变类型的镜头过渡方式。寻找对渐变类型镜头过渡方式的快速高效探测算法是我们未来对系统优化所要解决的一个新的挑战性课题。

参 考 文 献

1 Zhang H J, Kankanhalli A, Smoliar S W Automatic partitioning of full-motion video. Multimedia Systems, 1993, 1: 10～28

2 Shrahraray B. Scene change detection and content-based sampling of video. In: Proc. of the SPIE, vol 2419, 1995

3 Zabih R, Miller J, Mai K. A Feature-based algorithm for detecting and classifying production effects. Multimeida Systems, 1999, 7: 119～128

4 Meng J, Juan Y, Chang S F. Scene change detection in a MPEG compressed video sequence. In: Proc. of the SPIE, vol, 2419, 1995

5 Shen K, Delp J. A fast algorithm for video parsing using MPEG compressed sequences. In: Proc. of the IEEE Intl Conf. on Image Processing. 1995

6 Yeo B L. Liu B. Rapid scene analysis on compressed video. IEEE transaction on circuits and systems for video technology, 1995, 5(6) (Transactions Best Paper Award)

7 Kobla V, DeMenthon D, Doermann D. Special effect edit detection using VideoTrails: a comparison with existing techniques. In: Proc of SPIE conf. on Storage and Retrieval for Image and Video Databases VII, Jan. 1999

8 王伟强，高文. 一种MPEG-2流的索引模型及其应用. 投软件学报，已录用