基于数据挖掘的恶意行为检测方法*)

蔡皖东
(西北工业大学计算机科学与工程系　西安710072)

Detection Approach for Malicious Behaviors Based on Data Mining

CAI Wan-Dong
(Department of Computer Science and Engineering, Northwestern Polytechnical University, Xian 710072)

Abstract Intrusion Detection System (IDS) is an important network security technique. It can dynamic detect the network attack behaviors. At present, great issues for IDS are incapable of detecting the unknown malicious attack behaviors. So that, some new detection techniques are presented, data mining-based detection technique is an effective method in them. In this paper, data mining-based detection method and its key techniques are discussed in detail.

Keywords Network security, Intrusion detection system, Data mining, Detection models

1. 引言

入侵检测系统(Intrusion Detection System, IDS)是一种动态的网络攻击检测技术，能够在网络系统运行过程中发现入侵者的恶意行为和踪迹，并适时地作出反应。它是防火墙之后的第二道安全防线，与防火墙相辅相成，构成了一个完整的网络安全防护系统。

IDS 中的入侵检测方法主要分为两种：异常检测和误用检测。异常检测通过建立典型网络活动的轮廓(Profile)模型来实现，它通过提取审计踪迹(如网络流量和日志文件)中的特征数据来描述用户行为，建立轮廓模型。每当检测到一个新的行为模式，则与轮廓模型相比较，如果两者之差超过一个给定的阈值，则将会引发报警，表示检测到一个异常行为。在异常检测方法中，需要解决的问题是：从审计踪迹中提取特征数据来描述用户行为、正常行为和异常行为的分类方法以及轮廓模型的更新技术等。这种入侵检测方法的检测率较高，但误检率也比较高。

在误用监测系统中，入侵检测是基于已知的攻击模式，检测精确度取决于攻击模式库的完整性。通常，它只能检测攻击模式库中已有的攻击模式，而不能发现未知的攻击模式，甚至不能发现轻微变异的攻击模式，并且添加新的攻击模式是非常困难的。这种入侵检测方法的检测率较低，但误检率也比较低。大多数的商用入侵检测系统属于这类系统。

目前，入侵检测技术的研究重点是针对未知攻击模式的检测方法及其相关技术，并提出了一些新的检测方法，如基于数据挖掘、遗传算法、免疫系统等。其中，基于数据挖掘的检测方法通过分类、连接分析和顺序分析等数据分析方法来建立检测模型，增强了入侵检测系统对未知攻击模式的检测能力。

2. 数据预处理

在基于数据挖掘的入侵检测方法中，首先需要采集大量的审计踪迹数据，其中应当包含代表"正常"行为和"异常"行为的两类数据。然后对数据进行预处理，构造两个样本数据集：训练数据集和测试数据集。也可以先构造一个较大的样本数据集，然后将样本数据集分成训练数据集和测试数据集两部分，两者的比例大致为6:4。

样本数据集主要来自于每个主机上的日志文件或实时采集的网络数据包。为了描述一个程序或用户的行为，需要从样本数据集中提取有关的特征数据。例如，可以采用 TCP 连接数据来描述一个用户连接行为，对于每个 TCP 连接，与用户连接行为有关的特征数据有：

(1)建立 TCP 连接时的信息。在建立 TCP 连接时是否完整经历了三次握手过程，可能的错误信息有：被拒绝的连接、有连接请求但连接没有建立起来(发起主机没有接收到 SYN 应答包)、无连接请求却接收到了 SYN 应答包等；

(2)在 TCP 连接上传送的数据包、应答(ACK)包以及统计数据。统计数据包括数据重发率、错误重发率、两次 ACK 包比率、错误包尺寸比率、双方所发送的数据字节数、数据包尺寸比率和控制包尺寸比率等；

(3)关闭 TCP 连接时的信息。一个 TCP 连接以何种方式被终止的信息，如正常终止(双方都发送和接收了 FIN 包)、异常中断(一方发送了 RST 包，并所有的数据包都被应答)、半关闭(只有一方发送了 FIN 包)和断开连接等。

在样本数据中，每个 TCP 连接形成一个连接记录，并包含有如下属性信息：开始时间、持续时间、参与主机地址、端口号、连接统计值(双方发送的字节数、重发率等)、状态信息(正常的或被终止的连接)和协议号(TCP 或 UDP)等。这些属性信息构成了一个用户连接行为的基本特征。

考虑到网络拓扑对于网络入侵检测的重要性，根据 TCP 连接的发起方和接收方，可以将每个连接分成三种类型：从内部网到外部网；从外部网到内部网和内部网到内部网。一般情况下，网络入侵主要来自于外部，首先在外部网到内部网的连接上显示出某些异常模式(如渗透企图)，随后波及到其它两种连接类型。通过区分连接类型，有助于构造更为精确的检测模式。

3. 基于数据挖掘的入侵检测模型

在审计数据中包含了两类数据：代表正常行为的数据和代表异常行为的数据。在数据挖掘中，通常采用分类方法对这些数据进行分析，建立相应的检测模型，并依据检测模型从当前和今后的审计数据中检测出已知的和未知的入侵行为。为了自动构造检测模型，这里使用了分类、关联规则和频繁事件

*) 本文得到航空科学基金的资助。蔡皖东　教授，博士，主要研究方向为计算机网络、网络信息安全和多媒体通信等。

等方法，其检测模型的精确度依赖于大量的训练数据和正确的特性数据集。关联规则和频繁事件算法主要用于计算审计数据的一致模式，这些模式组成了一个审计追踪的轮廓，可用于指导审计数据的收集、系统特性的选择以及入侵模式的发现等。

3.1 数据分类

分类是数据挖掘中常用的一种数据分析方法，通过分类算法将一个数据项映射到预定义的某种数据类上，并生成相应的模型或分类器输出。分类一般分为两个阶段：

第一阶段是使用一种分类算法建立模型或分类器，描述预定的数据类集合。分类算法首先在一个由样本数据组成的训练数据集上进行学习，然后根据数据特征和描述将一个数据项映射到预定义的某一数据类中，并建立分类器模型。分类算法可以基于分类规则、判定树或数学公式等。

第二阶段是在测试数据集上应用分类器进行数据分类测试，对分类器的精确度和效率进行评估。

将分类方法应用于入侵检测中时，首先需要采集大量的审计数据，其中包含“正常”和“异常”两类数据，经过数据预处理后，构造一个训练数据集和一个测试数据集。然后在训练数据集上应用一种分类算法，建立分类器模型，分类器中的每个模式分别描述了一种系统行为样式。最后将分类器应用于测试数据集，评估分类器的精确度。一个优良的分类器应当具有高检测率和低误检率，检测率是指正确检测到异常行为的几率，误检率是指错误地将正常行为当作异常行为的几率，它也称为假肯定率。一个优良的分类器可以用于今后对未知恶意行为的检测。

为了提高检测精确度，可以采用基于多个检测模型联合的分类模型，将多个分类器输出的不同证据组合成一个联合证据，以便产生一个更为精确的断言。这种联合分类模型可以采用一种层次化检测模型来实现，它定义了两种分类器：基础分类器和中心分类器，并按两层结构来组织这些分类器。底层是多个基础分类器，基础分类器的每个模式对应于一种系统行为样式，其作用是根据训练数据中的特征数据来判断一种系统行为是否符合该模型，然后作为证据提交给中心分类器进行最后的决策；高层是中心分类器，它根据各个基础分类器提交的证据产生最终的断言。这种层次化检测模型的基本学习方法如下：

(1)构造基础分类器：每个模型对应于不同的系统行为样式；

(2)表达学习任务：训练数据中的一个记录可以看作是一个基础分类器所采集的证据，基础分类器将根据一个记录中的每个属性值来判定该系统行为是属于“正常”还是属于“异常”，即它是否符合该模型；

(3)建立中心分类器：使用一种学习算法来建立中心分类器，并输出最终的断言。

基于不同系统行为样式的多个证据进行综合决策，显然可以提高分类模型的精确度。这种层次化检测模型可以映射成一种分布式系统结构，不仅有利于提高检测精确度，并且还有利于分散检测任务负载，提高分类模型的执行效率。

3.2 关联规则

关联规则主要用于从大量数据中发现数据项之间的相关性。数据形式是数据记录集合，每个记录由多个数据项(项目)组成。

一个关联规则可以表示成：X→Y、置信度(confidence)和支持度(support)，其中 X 和 Y 是一个记录中的项目子集，支持度是包含 X+Y 记录的百分比，置信度是 support(X+Y)/support(X)比率。

在入侵检测中，关联规则主要用于分析和发现审计数据之间的相关性，为正确地选择入侵检测系统特性集合提供决策依据。

审计数据被表示成格式化的数据库表，其中每一行是一个审计记录，每一列是一个审计记录的属性字段，以表示系统特性。在这些系统特性中，明显存在着用户行为的频繁相关性。例如，为了检测出一个已知的恶意程序行为，可以将一个特权程序的访问权描述为一种程序策略，它应当与读写某些目录或文件的特定权限相一致，通过关联规则可以捕获到这些行为的一致性。

例如，将一个用户使用 shell 命令的历史记录表示成一个关联规则：trn→rec.log;[0.4,0.15]。其中，置信度为0.4，支持度为0.15，它表示该用户调用 trn 时，40%的时间是在读取 rec.log 中的信息，并且这种行为占该用户命令历史记录中所有行为的15%。

3.3 频繁事件

频繁事件是指频繁发生在一个滑动时间窗口内的事件集，这些事件必须以特定的最小频率同时发生在一个滑动时间窗口内。频繁事件分为顺序频繁事件和并行频繁事件，一个顺序频繁事件必须是按局部时间顺序地发生，而一个并行频繁事件则没有这样的约束。

对于 X 和 Y，X+Y 则是一个频繁事件，而 X→Y、confidence = frequency(X+Y)/frequency(X)和 support = frequency(X+Y)称为一个频繁事件规则。例如，在一个 Web 网站日志文件中，一个顺序频繁事件规则可以表示为：home，research→security;[0.3，0.1],[30s]。它表示当用户访问该网页(home)和研究项目简介(research)时，在30秒时间内随后访问信息安全组(security)网页的情况为30%，并且发生这个访问顺序的置信度为0.3，支持度为0.1。

由于程序执行和用户命令中明显存在着顺序信息，使用频繁事件算法可以发现审计记录中的顺序信息以及它们之间的内在联系。这些信息可用于构造异常行为轮廓。

4. 模式发现和评价

使用关联规则和频繁事件算法可以从审计踪迹中生成一个规则集，它们由关联规则和频繁事件组成，可用于指导审计处理。为了从审计踪迹中发现新的模式(规则)，可以多次以不同的设置来运行一个程序，以便生成新的审计踪迹。对于每次程序运行所发现的新规则，可以通过合并处理加入到现有的规则集中，并使用匹配计数器(match_count)来统计规则集中规则的匹配情况。

当规则集稳定(即无新规则的加入)后，便产生一个基本的审计数据集。然后通过修剪规则集，去除那些 match_count 值低于某一阈值的规则，其中阈值是基于 match_count 值占审计踪迹总量的比率来确定的，通常由用户指定。

从审计数据中发现的模式可以直接用于异常检测。首先使用关联规则和频繁事件算法从一个新的审计踪迹中生成规则集，然后与已建立的轮廓规则集进行比较，通过评分(scoring)功能进行模式评估。通常，它可以识别出未知的新规则、支持度发生改变的规则以及与支持度/置信度相悖的规则等。

为了评估分类器的精确度，通常使用一个测试数据集对分类器进行测试。根据有关的研究实验，基于数据挖掘的入侵检测方法具有较高的检测率和较低的误检率，具体的与所采用挖掘算法、训练数据集以及系统构成等因素有关。

(下转第84页)

t_2,FALSE)增加到HC中,再发送给MH。MH收到数据,因最后一项为FALSE,知道这是所求数据,把它写入缓存,更新时戳为t_2,并回答查询。假设此时MH进入睡眠状态(即可能是主动断连,也可能是网络故障),如图3所示,MA发送的失效项B、A被丢失。当MH被唤醒后,它发送一个探测消息(*,t_2)。MA收到此消息后,删除所有时戳小于等于t_2的三元组最后一项为TRUE的数据项,本例删除(C,t_1,TRUE);然后发送所有时戳大于t_2的失效项。

2.2.2 *移动代理的迁移* 前一小节我们讨论了移动代理在特定MSS上的工作原理,但随着移动主机的移动,移动代理也必须跟着迁移,如何迁移是这一小节需要讨论的问题。

移动代理系统由移动代理(MA)和移动代理环境(Mobile Agent Environment)两个部分组成。MAE是分布在网络各计算机设备上的软件系统,它也称为移动代理服务器和移动代理平台,它建立在操作系统之上,为MA提供运行环境。MA是只能存活在MAE中的软件实体,MA的移动便是从一个MAE移动到另一个MAE。本文中的假设3已说明,要求所有的MSS装载MAE。

移动代理的迁移原理如图4所示,其工作原理说明如下:

(1)假设移动主机MH_i从无线网络N_i漫游到无线网络N_j;

(2)MH_i向网络发送一个迁移消息,该消息包含其当前的缓存时戳t;

(3)当在MSS_j上的MA_i收到此消息后,它先在MSS_j上创建一个新的移动代理MA'_i,再把HC_i中所有大于等于t的失效项传给新代理。

(4)新的代理MA'_i被创建后,与Mh_i通信复制缓存副本。

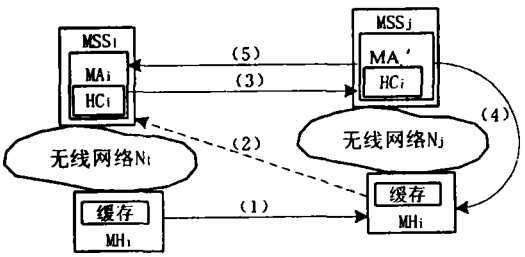

图4 移动代理的迁移

(5)在确认新代理收到所有大于等于t的失效项后,在原位置的MA_i被撤销,在此期间的所有由服务器发送的失效报告由新代理接管。

结束语 本文提出的基于移动代理的缓存失效方案最大的优势在于:不论是什么原因引起的失效报告丢失,例如移动主机主动断连、网络故障引起的断连或单纯的报文丢失,它都能处理,而且方法简单。特别是当MH位于某特定MSS覆盖网络时,工作非常稳定,若MH跨网段移动频繁,其代理迁移方案还有待进一步改进,这将是我们下一步工作的重点。

参考文献

1 Barbara D, Imielinski T. Sleepers and Workaholics: Caching Strategies in Mobile Environments (Extended Version). MOBIDATA: An Interactive J. Mobile Computing, 1994,1(1)

2 Barbara D, Imielinski T. Sleeper and Workaholics: Caching Stragegies in Mobile Enviroments. Very Large Database J. ,Dec. 1995

3 Kahol A, et al. A Strategy to Manage Cache Consistency in a Disconnected Distributed Environment. IEEE transactions on parallel and distributed systems,2001,12(7)

4 Tan K-L, Cai Jun, Ooi B C. A Evaluation of Cache Invalidation Strategy in Wireless Environments. IEEE transactions on parallel and distributed systems, 2001,12(8)

5 Liu G Y,McGuire G Q,Jr. A Mobility-Aware Dynamic Database Caching Scheme for Wirless Mobile Computing and Communication. Distributed and Parallel Database, 1996,4:271～288

6 Sistla A P. Wolfson O,Huang Y. Minimization of Communication Cost Through Caching in Mobile Enviroments. In: Proc. ACM Special Interest Group on Management of Data, May 1994

7 Cai C, Tan K L. Energy-Efficient Selective Cache Invalidaion. Wireless Networks, 1999,5(6):489～502

8 Sistla A P,Wolfson O,Huang Y. Minimization of Communication Cost Through Caching in Mobile Enviroments. IEEE Trans. Parallel and Distributed Systems,1998,9(4):378～389

9 Mobile Agent Facility Specification. OMG TC Document ef/xx-xx-x,1997

10 Wong D, et al. Concordia: An infrastructure for collaborating mobile agents. In: Mobile Agent-First International Workshop, MA'97. Berlin,1997

11 The Agent Society. Agent Product and Research Activities. 2000. http://www.agnent.org

(上接第66页)

结束语 在入侵检测技术中,需要解决的关键问题是未知恶意行为的检测问题。基于数据挖掘的入侵检测方法为检测未知的恶意行为提供了一种可行的方法。这种方法在检测精度、算法效率、实时检测以及系统实现等方面还存在着一些问题,有待于进一步研究和解决。

参考文献

1 Jiawei H, Micheline K 著, 范明等译. 数据挖掘—概念与技术. 机械工业出版社,2001

2 Wenke L, Salvatore J S. Data Mining Approaches for Intrusion Detection. In: Proc. of the 7th USENIX Security Symposium, January, 1998

3 Wenke L, Salvatore J S. Automated Intrusion Detection Using NFR: Methods and Experiences. In: Proc. of the Workshop on Intrusion Detection and Network Monitoring, April, 1999

4 Henry S, Teng K C, Stephen C L. Security Audit Trail Analysis Using Inductively Generated Predictive Rules. In: Proc. of the 11th National Conference on Artificial Intelligence Applications, IEEE, IEEE Service Center, Piscataway, NJ, March 1990. 24～29

5 Debar H, Becker M, Siboni D. A Neural Network Component for an Intrusion Detection System. In:Proc. of the IEEE Symposium on Research in Security and Privacy, Oakland, CA, May 1992

6 Rebecca G B 著, 陈明奇译. 入侵检测. 人民邮电工业出版社, 2001

7 Kohavi R. A study of cross-validation and bootstrap for accuracy estimation and model selection. IJCAI, 1995

8 Matthew G S, Eleazar E, Erez Z, Salvatore J S. Data Mining Methods for Detection of New Malicious Executables. To appear in IEEE Symposium on Security and Privacy, May 2001