异常检测方法综述[1]

张剑，龚俭
（东南大学 计算机科学与工程系，南京 210096）

Summary of Anomaly-Detection Approaches

Zhang Jian, Gong Jian
(University of SouthEast Deparment of Computer Science and Engineering, Nan Jing 210096)

【Abstract】

The approach of anomaly detection is a vigorously adaptive technique because it can detect unknown intrusions. The paper summarizes the advantage and the shortcoming of known anomaly-detection approaches in the past, which is based on the model of intrusion detection proposed by Dorothy Denning. Moreover, the development current of anomaly-detection is proposed on the above.

【Key】Intrusion Detection, Anomaly Detection, Misuse Detection, Data Mining, Neural Network, Bayesian Statistics

【摘要】:

异常检测方法是适应性较强的入侵检测技术，因为它能检测到未知的入侵方式。本文在 Dorothy Denning 提出的入侵检测模型的基础上，总结了以往主要的异常检测方法的优点和缺点，并据此提出其发展趋势。

【关键字】入侵检测，异常检测，滥用检测，数据挖掘，神经网络，贝叶斯统计

一．引言

计算机联网技术的发展改变了以单机为主的计算模式，但是，网络入侵的风险性和机会也相应地急剧增多。设计安全措施来防范未经授权访问系统的资源和数据，是当前网络安全领域的一个十分重要而迫切的问题。目前，要想完全避免安全事件的发生并不太现实，网络安全人员所能做到的只能是尽力发现和察觉入侵及入侵企图，以便采取有效的措施来堵塞漏洞和修复系统，这样的研究称为入侵检测。为此目的所研制的系统就称为入侵检测系统（intrusion detection system，简称 IDS）

入侵检测技术根据检测方法的不同分为两大类，一种是滥用检测方法，该方法通过对采集的信息按已知的知识进行分析，发现正在发生和已经发生的入侵行为；另一种叫做异常检测方法，它通过采集和统计发现网络或系统中的异常行为，然后按照某种决策算子来判断它是否入侵。滥用检测方法的误报率较低，但它不能检测出未知的攻击方式；而异常检测方法能检测出未知的攻击方式，但其误报率较高。通常将两种方法结合起来能够扬长避短，从而取得比较好的检测效果。

通常衡量一种检测方法的检测效果有两个指标，一个是漏检率（false negative rate），它是不能检测到的入侵行为的比率，该指标越低，表明检测效果越好；另一个是误检率（false positive rate），它是误将正常行为判为入侵行为的比率，该指标越低，表明检测效果越好。

二．Dorothy Denning 的入侵检测模型

入侵检测技术模型最早由 Dorothy Denning[1]提出，如图 1 所示。目前，检测技术及其体系均是在此基础上的扩展和细化。

[1] 本文受国家自然科学基金项目 90104031 资助

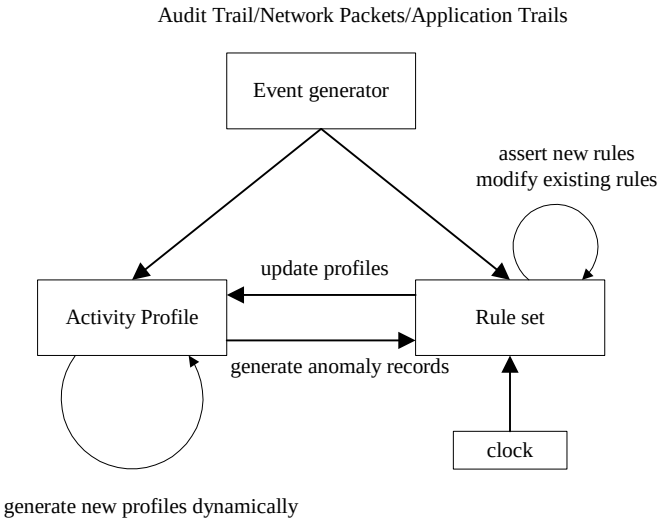

图1 通用入侵检测模型

该模型的第一个部件是事件发生器（Event generator），它负责从不同的数据源采集数据，数据源可以是主机的审计日志、应用程序的日志或网络报文。在目前实现的大多数入侵检测系统中，这个部件一般被细化为采集器和数据预处理模块。采集器只负责从数据源采集数据，而数据预处理模块将原始数据转化为检测模型所能识别的数据格式。行为映像（activity profile）相当于一个异常检测模块，它也可以细化为很复杂的结构，一般包括数据源、训练模块、检测模块和决策引擎。行为映像能根据新数据来更新自己，以保持检测模型对用户行为或网络行为的变化的同步；如果发现入侵事件，则将该事件转移给规则集（Rule set）。规则集实质是滥用检测模块，它根据其内部对已知攻击的知识集对事件发生器传递下来的事件进行模式或规则匹配，一旦发现有与知识集匹配的行为，就判定其为入侵行为并产生报警。如果当前事件与知识集不匹配，它将被传递给行为映像来更新其正常行为映像，知识集不一定是一组规则，也可以是状态迁移图、神经网络等滥用检测模型。

下图是 Aurobindo Sundurum[2]提出的一个典型的异常检测系统结构：

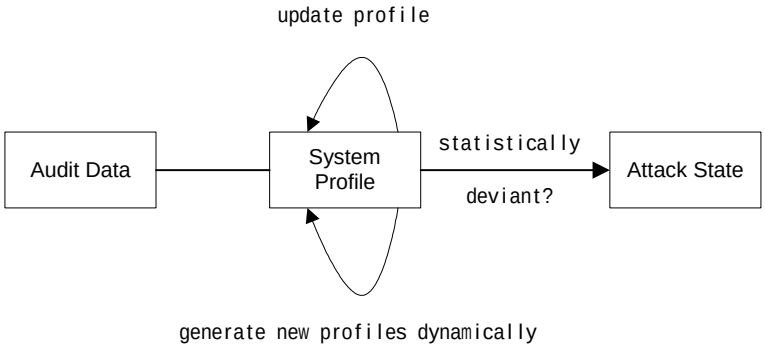

图2 A typical anomaly detection system

该结构包含了异常检测系统所必须具备的部件，它是 Dorothy Denning 模型中行为映像的细化。其中系统映像(System Profile)就是检测模型，它是系统或用户等正常行为的总结，在建立以后还要根据新数据动态地更新，以减少误检率和漏检率。异常检测方法的特点一般都集中在检测模型，下面将按其不同来对以往的异常检测方法进行分类。

三．异常检测模型分类

3.1 统计性模型

统计性模型首先要选择描述主体行为的测度集，然后在采集到的安全事件集合中建立基于该测度集的检测模型，该模型有可能是用户的正常行为映像，也可能是正常网络流量测度的概率分布，取决于异常检测系统的检测目标。取决于实际的检测模型，某种度量算法被用来计算当前的主体行为与检测模型的背离程度，然后根据某种决策方法来决定是否入侵。统计性模型的主要优点是能自适应地学习主体的行为，因此对异常行为比人更加敏感。另外统计性模型不要求训练数据全部是纯粹的正常行为，只要是真实环境的数据就行，这一点很重要，因为从数据源获取的数据一般是没有标记的。采用统计性模型最有代表性的要数 SRI 的 IDES[3]和 NIDES[4]，这两个入侵检测系统包含了面向用户的异常检测模块，通过建立描述用户行为的各测度的概率分布函数作为其检测模型，并采用某种距离算法评价用户当前行为与其模型的差异并作出响应。统计性模型的缺点是：（1）它容易被入侵者所训练，最后使得异常行为也变成正常的。（2）主观确定的入侵阀值决定了误检率和漏检率的高低（3）它对于依赖于事件之间关系的入侵不敏感，因为这种模型忽略了事件之间的关系。（4）需要假设测度的概率分布，目前一般采用正态分布或泊松分布，而这有可能与实际不符合。

3.2 预测模型

预测模型的检测对象是事件的时间序列，其目的是发现在构成入侵的安全事件[1]集合在时间上的相关性，从而预测未来发生的事件，如果实际发生的事件与预测结果有较大的差异，则表明有异常现象发生。Teng 和 Chen[5][6]提出基于时序的推导性归纳方法，产生时序性规则来建立用户的正常行为映像。这些规则在训练阶段会动态地调整，只有较高的预测准确性的规则被保留。例如从大量的正常事件序列中发现了以下一条规则：

$$E_1 - E_2 \rightarrow (E_3 = 80\%, E_4 = 15\%, E_5 = 5\%)$$

E_1、E_2、E_3、E_4 和 E_5 都是安全事件，该规则的语义是指在 E_2 继 E_1 发生后，发生 E_3 的概率是 80%，而发生 E_4 的概率是 15%，而发生 E_5 的概率是 5%。这种模型有以下的优点：（1）它能检测到传统的模型（例如统计性模型）不能检测的入侵（2）它对主体行为的变化有高度的适应性，因为低质量的时间序列模式会不断地被排除，而留下高质量的模式。（3）它对在训练阶段企图训练它的行为比较容易察觉。（4）它能在入侵完成前检测到并发出报警。

3.3 基于机器学习的异常检测模型

基于机器学习的异常检测模型是用机器学习的方法来建立系统映像。它的最大特点是根据正常来分辨异常，因为其训练数据大多是代表清一色的正常行为。这种方法的优点是检测速度快，而且误检率低。但该方法在用户动态行为变化以及单独异常检测方面还有待改善。复杂的相似度量和先验知识加入到检测中可能会提高系统的准确性，但需要做进一步的工作。机器学习方法在异常检测中的运用非常广泛，下面是一些有代表性的方法：

（1） 基于数据挖掘的异常检测模型

数据挖掘能从审计记录或数据流中提取出感兴趣的知识，这些知识是隐含的、事先未知的、潜在的有用信息，提取的知识表示为概念、规则、规律、模式等形式，并可用这些知识去检测异常入侵和已知的入侵。Wenke Lee 和 Salvatore J. Stolfo[7][8][9]等在 1998 年和 1999 年提出通过对正常数据建立决策树的预测模型来作为检测模型，然后用该模型来检测实际发生的网络报文是否异常。他们在 2001 年提出将基于数据挖掘的检测模型应用到实时环境中，着重解决三个关键问题：检测的准确性、效率和可用性。数据挖掘的优点是能自动、快速地产生异常检测模型，这在海量的历史数据中提取知识是非常重要的，通过人工建立的方法很难实现。数据挖掘方法的缺点在于：①误报率较高；②由于在训练和评价时计算的复杂度较高，难以应用到实时环境中；③需要大量的训练数据，而且对数据的纯洁性要求较高。

（2） 基于神经网络的异常检测模型

---

[1] 安全事件是指一条审计记录，或一个检测实体单元

神经网络可以应用到各种异常检测模型中，人工神经网络力图模拟生物神经系统，通过接受外部输入的刺激，不断获得并积累知识，进而具有一定的判断预测能力。有的基于神经网络的异常检测模型是分类器，它通过训练和学习，记忆了系统的正常行为或入侵行为，并能根据系统现状进行自我调节，有效地发现并阻止各种入侵行为，这种神经网络与基于数据挖掘的决策树的作用是类似的，例如李鸿培和王新梅[10]设计的基于神经网络的入侵检测系统模型。另一种神经网络可用来建立预测模型，例如通过在神经网络的输入端输入用户所用的命令序列可以预测下一个命令，如果不符合就可判定为异常，这种神经网络的例子是Kevin[11]等设计的面向入侵检测的神经网络模型。神经网络的优点是：①它的实现不依赖对潜在数据的统计假设；②能较好地处理噪声数据；③能自动调节影响输出的各测度的权重，而这在传统的异常检测方法中通常是人为确定的。神经网络的缺点在于：①神经网络的拓扑结构和各元素的权重只有在训练后才能确定；②输入窗口的大小是该方法的一个主观因素，如果设得太低，该模型的检测能力就会下降，如果太高，就会碰到许多不相关的输入。

（3） 基于免疫学原理的异常检测模型

基于免疫学原理的异常检测模型受生物免疫系统的启发，试图为要保护的对象建立一个“免疫系统”，该方法的关键是如何有效地定义“自我”（self）和识别自我，并据此来排斥“异类”。Stephanie Forrest[12]设计了一个保护 Unix 进程的“免疫系统”，他将进程的正常系统调用序列作为“自我”，然后据此建立后缀树（Suffix tree）和有限状态机（FSM），在检测阶段 FSM 被用来分析该进程当前的系统调用序列，一旦发现背离程度大于某个阀值，就发出报警。

（4） 基于 IBL 的异常检测模型

IBL（Instance-based Learning）是基于实例的学习方法。该方法将入侵检测问题形式化地表述成根据离散数据的时间序列来代表个人、系统或网络的特征，并采用某种相似度测量方法将离散数据的时间序列转化为可度量比较的空间，从而量化正常序列和异常序列的差异并据此作出决策。Terran Lane 和 Carla E. Brodley[13]曾用用户命令数据来测试该方法的有效性，获得了比较理想的效果。

四．检测算法分类

检测算法用来比较系统当前行为与检测模型，将多个测度的值合成，从而获得一个量化的检测评价。比较有代表性的决策算法有：

（1） 基于贝叶斯推理[14]的检测算法

基于贝叶斯推理的异常检测方法是通过在任意给定的时刻，测量 A_1，A_2，…，A_nn 种测度值，推理判断系统是否有入侵事件发生。贝叶斯推理的原理在于已知异常情况下各测度值出现的概率，从而推出在 A_1，A_2，…，A_n 的情况下，入侵发生的概率。贝叶斯推理的公式是：

$$\frac{P(I \mid A_1, A_2, \mathbf{L}, A_n)}{P(\neg I \mid A_1, A_2, \mathbf{L}, A_n)} = \frac{P(I)}{P(\neg I)} \frac{\prod_{i=1}^{n} P(A_i \mid I)}{\prod_{i=1}^{n} P(A_i \mid \neg I)}$$

等式左边如果大于 1，则判定为入侵。

（2） 贝叶斯网络[14][15]

上述的贝叶斯推理公式是基于各测度相互独立的假设，而实际许多测度之间是相互联系的，贝叶斯网络是一种信任网络，当它应用于异常检测时，网络中每个节点代表一种测度及其概率分布，对于根节点，概率分布是不依赖其他测度的，即是客观概率；对于子节点，它是以根节点为条件的条件概率，输入各测度的当前值后，该网络就能输出一个考虑到测度之间关系的综合异常评价意见。

（3）协方差测量[3]

该方法用协方差矩阵来描述各测度之间的关系度。如果 A_1，A_2，…，A_n 等测度用向量 A 来表示，那么复合异常测度，即综合异常评价意见就表示为：$A^T C^{-1} A$，其中 C 为协方差矩阵。

（4）模糊逻辑（*Fuzzy Logic*）[16][17]

模糊逻辑是模糊数学的基础理论，它把“真假”作为语言变量来处理，使得对一个命题的判断可以采用“比较真”，“比较假”等词。模糊逻辑推理采用的是假言推理的近似式而不是精确形式。假言推理的规则可以写为：

$$\begin{cases} A \to B & \text{大前提（蕴含）} \\ A^* & \text{小前提} \\ B^* = A^* \circ (A \to B) & \text{结论} \end{cases}$$

在应用模糊逻辑推理的异常检测系统中，根据当前的各测度值和模糊规则，得到一个异常的模糊度，如果模糊度大于某个阀值，则可判定为入侵。

四．总结

检测算法是依赖于检测模型的，有些检测模型不需要复杂的检测算法，例如决策树和神经网络，只需将各测度值作为它们的输入，就能获得综合评价意见。而对于统计性模型，通常需要专门的检测算法来获得综合评价意见。

异常检测方法假设所有的入侵行为都是异常行为，因此如何判断异常行为是异常检测技术的核心，而这通常是带来误检和漏检的主要环节。提高判断异常行为的准确性关键要建立强壮的检测模型和采用与模型相适应的检测算法。随着网络入侵技术的不断发展，入侵的行为表现出不确定性、复杂性、多样性等特点，入侵检测面临许多有待解决的关键问题，如高效率的检测算法、入侵模式确认、入侵实时监测、入侵描述语言、检测数据标准化、高速网络中的入侵检测、IDS 评估、IDS 与其他系统的协同工作等一系列问题都有待进一步研究和实现。

【参考文献】

1．[Dorothy E. Denning,1987]Dorothy E. Denning;*An intrusion detection model*;In IEEE Transactions on software engineering

2．[Aurobindo Sundurum]*An introduction to intrusion detection*

3．[T.F. Lunt,1992]T. F. Lunt,A. Tamaru, F. Gilham; *A real-time intrusion detection expert system(IDES)-Final technical report*; Computer Science Laboratory,SRI

4．[Debra Anderson,1995]Debra Anderson,Teresa F. Lunt,Harold Javitz; *Detecting unusual program behavior using the statistical component of the next-generation intrusion detection expert system(NIDES)*;Computer Science Laboratory,SRI

5．[K.chen,1988]K.chen; *An inductive engine for the acquisition of temporal knowledge*; PH.D Thesis; Department of computer science,University of I11inois at Urbana-champaign.

6．[Henry,1990]Henry S.Teng,Kaihu Chen,and Stephen C Lu; *Security audit trail analysis using inductively generated predictive rules*; In proceedings of the sixth conference on Artificial Intelligence Application

7．[Lee and Stolfo,1998] W. Lee and S. J. Stolfo. Data mining approaches for intrusion detection. In *Proceedings of the 7th USENIX Security Symposium*, San Antonio, TX, January 1998.

8．[Lee et al., 1998] W. Lee, S. J. Stolfo, and K. W. Mok. Mining audit data to build intrusion detection models. In *Proceedings of the 4th International Conference on Knowledge Discovery and*

Data Mining, New York, NY, August 1998. AAAI Press

9．[Lee et al., 1999a] W. Lee, S. J. Stolfo, and K. W. Mok. *A data mining framework for building intrusion detection models.* In Proceedings of the 1999 IEEE Symposium on Security and Privacy, May 1999.

10．[李鸿培，1999]李鸿培，王新梅；基于神经网络的入侵检测系统模型；西安电子科技大学学报

11．[Kevin L.Fox,1990]Kevin L.Fox,Ronda R. Henning. Jonathan H.Reed, and Richard Simonian; *A neural network approach towards intrusion detection*; In proceedings of the 13th national computer security conference

12．[Stephanie Forrest,1996]S. Forrest,S. A. Hofmeyr,A. Somajayi; *A sense of self for unix process.* IEEE symposium on Computer Security and Privacy

13．[Terran Lane,1999]Terran Lane, Carla E. Brodley; *Temporal sequence learning and data reduction for anomaly detection*; ACM

14．[慕春棣，2000]慕春棣，戴剑彬，叶俊；用于数据挖掘的贝叶斯网络；软件学报；

15．[Judea Pearl,1988]Judea Pearl; *Probabilistic reasoning in expert systems*

16．[John E. Dickerson,2000]John E. Dickerson; *Fuzzy network profiling for intrusion detection*; IEEE 2000

17．[陈鸣钊，1992]陈鸣钊，张志烈，樊宝康;《模糊数学及其实用》；河海大学出版社。

18．[Sandeep Kumar,1995]Sandeep Kumar; *Classification and detection of computer intrusions*; Doctor thesis

19．[蒋建春，2000]蒋建春，马恒太，任党恩，卿斯汉；网络安全入侵检测：研究综述；软件学报

20．[刘美兰，1999]刘美兰，姚京松；神经网络在入侵检测系统中的应用；计算机工程与应用

21．[汪立东，1999]汪立东，李亚平，方滨兴等；一个基于神经网络的入侵检测系统；计算机工程

22．[郭翠英，2001]郭翠英，余雪丽；基于神经网络的入侵检测模型；太原理工大学学报

23．[李之棠，2000]李之棠，杨红云；模糊入侵检测模型；计算机工程与科学

24．[Kumar,1995]Kumar；*Classification and Detection of Computer Intrusions*；Ph.D degree thesis,Purdue University