入侵检测技术研究综述

卿斯汉，蒋建春，马恒太，文伟平，刘雪飞

（1. 中国科学院 软件研究所，北京 100080；
2. 中国科学院 信息安全技术工程研究中心，北京 100080；3. 中国科学院 研究生院，北京 100080）

摘 要：入侵检测是信息安全保障的关键技术之一。本文综述了入侵检测系统的最新研究进展，包括基本概念、模型、方法等，讨论了该领域尚存在的问题及今后的发展趋势。

关键词：入侵检测；误用检测；异常检测；评估；标准化

中图分类号：TP393 **文献标识码：**A **文章编号：**1000-436X(2004)07-0019-11

Research on intrusion detection techniques: a survey

QING Si-han, JIANG Jian-chun, MA Heng-tai, WEN Wei-ping, LIU Xue-fei

(1. Institute of Software of Chinese Academy of Sciences, Beijing 100080, China;
2. Engineering Research Center for Information Security Technology, Chinese Academy of Sciences, Beijing 100080, China; 3. Graduate School of Chinese Academy of Sciences, Beijing 100080, China)

Abstract: Intrusion detection is one of the critical techniques in information assurance. In this paper we give a survey of the state of the art in the development of intrusion detection systems including basic notions，models, methods, etc. We also document some remaining problems and emerging trends in this area.

Key words: intrusion detection; misuse detection; anomaly detection; evaluation; standardization

1 引言

信息系统安全保障是一种防御体系，包括防护（protect）、检测（detect）、反应（react）和恢复（recovery）4 个层面[1]。入侵检测系统是其中一个重要的组成部分，扮演着数字空间“预警机”的角色。

本文综述了入侵检测系统的最新研究进展，包括：基本概念与模型；各类入侵检测技术；入侵检测系统的评估与标准化等。本文也讨论了该领域尚存在的问题及今后的研究方向。

2 入侵检测的基本概念与模型

早在 20 世纪 80 年代初期，Anderson 将入侵定义为：未经授权蓄意尝试访问信息、篡改

收稿日期：2004-02-10
基金项目：国家自然科学基金资助项目（60083007）；国家“973”重点基础研究发展规划基金资助项目（G1999035810）

信息、使系统不可靠或不能使用[2]。Heady 认为入侵是指试图破坏资源的完整性、机密性及可用性的行为集合[3]。Smaha 从分类角度指出[4]，入侵包括尝试性闯入、伪装攻击、安全控制系统渗透、泄漏、拒绝服务、恶意使用六种类型。卡内基－梅隆大学的研究人员将入侵定义为非法进入信息系统，包括违反信息系统的安全策略或法律保护条例的动作[5]。我们认为，入侵的定义应与受害目标相关联，该受害目标可以是一个大的系统或单个对象。判断与目标相关的操作是入侵的依据是：对目标的操作超出了目标的安全策略范围。因此，入侵系指违背访问目标的安全策略的行为。入侵检测通过收集操作系统、系统程序、应用程序、网络包等信息，发现系统中违背安全策略或危及系统安全的行为。具有入侵检测功能的系统称为入侵检测系统，简称 IDS。

最早的入侵检测模型是由 Denning[6]给出的，该模型主要根据主机系统审计记录数据，生成有关系统的若干轮廓，并监测轮廓的变化差异发现系统的入侵行为，如图 1 所示。

入侵行为的种类不断增多，涉及的范围不断扩大，而且许多攻击是经过长时期准备，通过网上协作进行的。面对这种情况，入侵检测系统的不同功能组件之间、不同 IDS 之间共享这类攻击信息是十分重要的。为此，Chen 等提出一种通用的入侵检测框架模型，简称 CIDF[7]。该模型认为入侵检测系统由事件产生器（event generators）、事件分析器（event analyzers）、响应单元（response units）和事件数据库（event databases）组成，如图 2 所示。

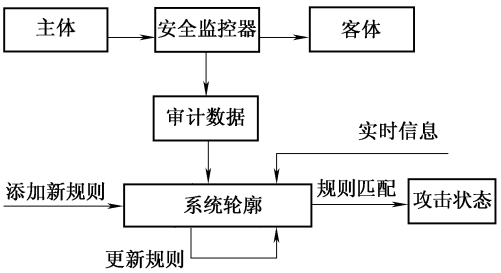

图 1 IDES 入侵检测模型

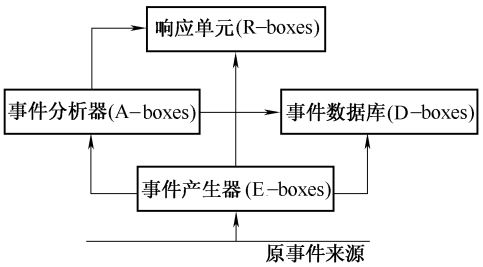

图 2 CIDF 各组件之间的关系图

CIDF 将入侵检测系统需要分析的数据统称为事件，它可以是网络中的数据包，也可以是从系统日志等其它途径得到的信息。事件产生器是从整个计算环境中获得事件，并向系统的其它部分提供事件。事件分析器分析所得到的数据，并产生分析结果。响应单元对分析结果做出反应，如切断网络连接、改变文件属性、简单报警等应急响应。事件数据库存放各种中间和最终数据，数据存放的形式既可以是复杂的数据库，也可以是简单的文本文件。CIDF 模型具有很强的扩展性，目前已经得到广泛认同。

3 入侵检测技术

入侵检测技术传统上分为两大类型：异常入侵检测（anomaly detection）和误用入侵检测（misuse detection）[8]。异常入侵检测系指建立系统的正常模式轮廓，若实时获得的系统或用户的轮廓值与正常值的差异超出指定的阈值，就进行入侵报警。异常入侵检测方法的优点是不依赖于攻击特征，立足于受检测的目标发现入侵行为。但是，如何对检测建立异常指标，如何定义正常模式轮廓，降低误报率，都是难以解决的课题。误用入侵检测系指根据已知的攻击特征检测入侵，可以直接检测出入侵行为。误用检测方法的优点是误报率低，可以发现已知的攻击行为。但是，这种方法检测的效果取决于检测知识库的完备性。为此，特征库必

须及时更新。此外，这种方法无法发现未知的入侵行为。混合型检测方法试图综合上述两种方法的优点，我们将在 3.3 节进行介绍。

3.1 异常入侵检测

异常检测的前提是异常行为包括入侵行为。理想情况下，异常行为集合等同于入侵行为集合，此时，如果 IDS 能够检测所有的异常行为，就表明能够检测所有的入侵行为。但是在现实中，入侵行为集合通常不等同于异常行为集合。事实上，行为有以下 4 种状况：(1) 行为是入侵行为，但不表现异常；(2) 行为不是入侵行为，却表现异常；(3) 行为既不是入侵行为，也不表现异常；(4) 行为是入侵行为，且表现异常。异常检测方法的基本思路是构造异常行为集合，从中发现入侵行为。异常检测依赖于异常模型的建立，不同模型构成不同的检测方法。异常检测需要获得入侵的先验概率，如何获得这些入侵先验概率就成为异常检测方法是否成功的关键问题。下面对不同的异常入侵检测方法进行论述。

3.1.1 基于特征选择的异常检测方法

基于特征选择的异常检测方法，系指从一组度量中选择能够检测出入侵的度量，构成子集，从而预测或分类入侵行为。异常入侵检测方法的关键是，在异常行为和入侵行为之间作出正确判断。选择合适的度量是困难的，因为选择度量子集依赖于所检测的入侵类型，一个度量集并不能适应所有的入侵类型。预先确定特定的度量，可能会漏报入侵行为。理想的入侵检测度量集，必须能够动态地进行判断和决策。假设与入侵潜在相关的度量有 n 个，则 n 个度量构成 2^n 个子集。由于搜索空间同度量数之间是指数关系，所以穷尽搜索理想的度量子集，其开销是无法容忍的。Maccabe[9]提出应用遗传方法搜索整个度量子空间，以寻找正确的度量子集。其方法是通过学习分类器方案，生成遗传交叉算子和基因突变算子，允许搜索的空间大小比其它启发式搜索技术更加有效。其它基于特征选择的异常检测方法，可以参考文献[10,11]。

3.1.2 基于贝叶斯推理的异常检测方法

基于贝叶斯推理的异常检测方法，系指在任意给定的时刻，测量 A_1，A_2，…A_n 变量值，推理判断系统是否发生入侵行为。其中，每个变量 A_i 表示系统某一方面的特征，例如磁盘 I/O 的活动数量、系统中页面出错的数目等。假定变量 A_i 可以取两个值：1 表示异常，0 表示正常。令 I 表示系统当前遭受入侵攻击。每个异常变量 A_i 的异常可靠性和敏感性分别用 $P(A_i=1|I)$ 和 $P(A_i=1|\neg I)$ 表示。于是，在给定每个 A_i 值的条件下，由贝叶斯定理得出 I 的可信度为

$$P(I|A_1,A_2,\cdots A_n)=P(A_1,A_2,\cdots A_n|I)\frac{P(I)}{P(A_1,A_2,\cdots A_n)} \tag{1}$$

其中，要求给出 I 和 $\neg I$ 的联合概率分布。假定每个测量 A_i 仅与 I 相关，与其它的测量条件 $A_j(i\neq j)$ 无关，则有

$$P(A_1,A_2,\cdots A_n|I)=\prod_{i=1}^{n}P(A_i|I) \tag{2}$$

$$P(A_1,A_2,\cdots A_n|\neg I)=\prod_{i=1}^{n}P(A_i|\neg I) \tag{3}$$

从而得到

$$\frac{P(I|A_1,A_2,\cdots A_n)}{P(\neg I|A_1,A_2,\cdots A_n)}=\frac{P(I)}{P(\neg I)}\frac{\prod_{i=1}^{n}P(A_i|I)}{\prod_{i=1}^{n}P(A_i|\neg I)} \tag{4}$$

因此，根据各种异常测量的值、入侵的先验概率、入侵发生时每种测量得到的异常概率，能够判断系统入侵的概率。但是为了保证检测的准确性，还需要考查各测量 A_i 之间的独立性。一种方法是通过相关性分析，确定各异常变量与入侵的关系[12]。

3.1.3 基于贝叶斯网络的异常检测方法

贝叶斯网络实现了贝叶斯定理揭示的学习功能，用于发现大量变量之间的关系，是进行预测和数据分类的有力工具。基于贝叶斯网络的异常检测方法，系指建立异常入侵检测的贝叶斯网络，通过它分析异常测量结果。贝叶斯网络允许以图形方式表示随机变量之间的相关关系，并通过指定的一个小的与邻接结点相关的概率集计算随机变量的联接概率分布。按给定全部结点组合，所有根结点的先验概率和非根结点概率构成这个集。贝叶斯网络是一个有向图 DAG，在 DAG 中，弧表示父结点与子结点之间的依赖关系。这样，当随机变量的值变为已知时，就允许将它吸收为证据，为其它的剩余随机变量条件值判断提供计算框架。需要解决的关键课题是，判断根结点的先验概率值与确定每个有向弧的连接矩阵。Valdes 和 Skinner 提出了一个基于贝叶斯网络的异常检测模型 eBayes TCP，用于发现网络中针对 TCP 协议的入侵行为[13]。

3.1.4 基于模式预测的异常检测方法

基于模式预测的异常检测方法的前提条件是，事件序列不是随机发生的而是服从某种可辨别的模式，其特点是考虑了事件序列之间的相互联系。Teng 和 Chen 给出一种基于时间的推理方法，利用时间规则识别用户正常行为模式的特征[14]。通过归纳学习产生这些规则集，并能动态地修改系统中的这些规则，使之具有较高的预测性、准确性和可信度。如果规则大部分时间是正确的，并能够成功地用于预测所观察到的数据，那么规则就具有较高的可信度。例如，TIM（time-based inductive machine）给出下述产生规则

$$(E1!E2!E3)(E4 = 95\%, E5 = 5\%)$$

其中 $E1 \sim E5$ 表示安全事件。上述规则说明，事件发生的顺序是 $E1$，$E2$，$E3$，$E4$，$E5$。事件 $E4$ 发生的概率是 95%，事件 $E5$ 发生的概率是 5%。通过事件中的临时关系，TIM 能够产生更多的通用规则。根据观察到的用户行为，归纳产生出一套规则集，构成用户的行为轮廓框架。如果观测到的事件序列匹配规则的左边，而后续的事件显著地背离根据规则预测到的事件，那么系统就可以检测出这种偏离，表明用户操作异常。这种方法的主要优点有：(1) 能较好地处理变化多样的用户行为，并具有很强的时序模式；(2) 能够集中考察少数几个相关的安全事件，而不是关注可疑的整个登录会话过程；(3) 容易发现针对检测系统的攻击。

3.1.5 基于贝叶斯聚类的异常检测方法

基于贝叶斯聚类的异常检测方法，系指在数据中发现不同数据类集合。这些类反映了基本的类属关系，同类成员比其它成员更相似，以此可以区分异常用户类，进而推断入侵事件发生。Cheeseman 和 Stutz 在 1995 年提出的自动分类程序（autoclass program），是一种无监督数据分类技术[8]。Autoclass 应用贝叶斯统计技术，对给定的数据进行搜索分类。其优点是：根据给定的数据，自动判断并确定类型数目； 不要求特别的相似测量、停顿规则和聚类准则； 可以混合连续属性与离散属性。

基于统计的异常检测方法对所观测到的行为分类处理，到目前为止，所使用的技术主要是监督式的分类，即根据观测到的用户行为建立用户行为轮廓。而贝叶斯分类方法允许理想化的分类数、具有相似轮廓的用户群组以及遵从符合用户特征集的自然分类。但是，该方法目前只限于理论讨论，还没有实际应用。自动分类程序怎样处理固有的次序性数据，在分类

中如何考虑统计分布特性等问题，还没有很好地解决。由于统计方法的固有特性，自动分类程序还存在异常阈值的选择和防止攻击者干扰类型分布等问题。

3.1.6 基于机器学习的异常检测方法

基于机器学习的异常检测方法，系指通过机器学习实现入侵检测，其主要方法有死记硬背、监督学习、归纳学习、类比学习等[15]。Carla 和 Brodley 将异常检测问题归结为，根据离散数据临时序列特征学习获得个体、系统和网络的行为特征；并提出了一个基于相似度的实例学习方法 IBL（instance based learning），该方法通过新的序列相似度计算，将原始数据（如离散事件流和无序的记录）转化成可度量的空间。然后，应用 IBL 学习技术和一种新的基于序列的分类方法，发现异常类型事件，从而检测入侵行为。其中，阈值的选取由成员分类的概率决定 [16,17]。新的序列相似度定义如下：

设 l 表示长度，序列 $X=(x_0,x_1,\cdots x_{l-1})$ 和 $Y=(y_0,y_1,\cdots y_{l-1})$

$$w(X,Y,i)=\begin{cases}0 & if \quad i<0 \quad or \quad x_i\neq y_i\\ 1+w(X,Y,i-1) & if \quad x_i=y_i\end{cases} \tag{5}$$

$$Sim(X,Y)=\sum_{i=0}^{l-1}w(X,Y,i) \tag{6}$$

$$Dist(X,Y)=Sim_{\max}-Sim(X,Y) \tag{7}$$

令 D 表示用户的模式库，由一系列的序列构成，X 表示最新观测到的用户序列，则

$$Sim_D(X)=\max_{Y\in D}\{Sim(Y,X)\} \tag{8}$$

上式用于分类识别，检测异常序列。实验结果表明这种方法检测迅速，而且误报率低。然而，这种方法对于用户动态行为变化以及单独异常检测还有待改善。总之，机器学习中许多模式识别技术对于入侵检测都有参考价值，特别是用于发现新的攻击行为。

3.1.7 基于数据挖掘的异常检测方法

计算机联网导致大量审计记录，而且审计记录大多数以文件形式存放（如 UNIX 系统中的 Sulog）。因此，单纯依靠人工方法发现记录中的异常现象是困难的，难以发现审计记录之间的相互关系。Lee 和 Stolfo 将数据挖掘技术引入入侵检测领域，从审计数据或数据流中提取感兴趣的知识。这些知识是隐含的、事先未知的潜在有用信息。提取的知识表示为概念、规则、规律、模式等形式[18~20]，并用这些知识检测异常入侵和已知的入侵。基于数据挖掘的异常检测方法，目前已有 KDD 算法可以应用。这种方法的优点是，适于处理大量数据。但是，对于实时入侵检测，这种方法还需要加以改进，需要开发出有效的数据挖掘算法和相应的体系[21,22] 。数据挖掘的优点在于处理大量数据的能力与进行数据关联分析的能力。因此，基于数据挖掘的检测算法将会在入侵预警方面发挥优势。

3.1.8 基于应用模式的异常检测方法

一般来说，入侵行为与应用联系密切。因此，对特定应用行为建模，发现异常入侵行为是一种可行的方法。Krugel 等人[23]提出一种基于服务相关的网络异常检测算法，用服务请求类型（type of request）、服务请求长度（length of request）、服务请求包大小分布（payload

distribution）计算网络服务的异常值。异常值的计算公式如下

$$AS=0.3\cdot AS_{type}+0.3AS_{len}+0.4\cdot AS_{pd} \tag{9}$$

其中，AS_{type}，AS_{len} 和 AS_{pd} 分别表示服务请求类型、服务请求长度和服务请求包的异常值。该方法利用已知的攻击方法训练异常阈值，在实际检测中，通过实时计算出的异常值和所训练出的阈值作比较，分析判断是否有针对某种网络服务的攻击发生。

3.1.9 基于文本分类异常检测方法

基于文本分类的异常检测方法[24]由 Liao 和 Vemuri 提出，其基本原理是将程序的系统调用视为某个文档中的“字”，而进程运行所产生的系统调用集合就产生一个“文档”。对于每个进程所产生“文档”，利用 K 最近邻聚类（K-nearest neighbor）文本分类算法，分析文档的相似性，发现异常的系统调用，从而检测入侵行为。

3.1.10 其它

由于篇幅所限，基于神经网络、统计的异常检测方法请参看文献[12,25]。

3.2 误用入侵检测

误用入侵检测的前提是，入侵行为能按某种方式进行特征编码。入侵检测的过程，主要是模式匹配的过程。入侵特征描述了安全事件或其它误用事件的特征、条件、排列和关系。特征构造方式有多种，因此误用检测方法也多种多样。下面列举主要的误用检测方法。

3.2.1 基于条件概率的误用检测方法

基于条件概率的误用检测方法，系指将入侵方式对应一个事件序列，然后观测事件发生序列，应用贝叶斯定理进行推理，推测入侵行为[8]。令 ES 表示事件序列，先验概率为 $P(\text{intrusion})$，后验概率为 $P(ES\mid\text{intrusion})$，事件出现概率为 $P(ES)$，则

$$P(Intrusion\mid ES)=P(ES\mid Intrusion)\frac{P(Intrusion)}{P(ES)} \tag{10}$$

通常网络安全员可以给出先验概率 P（intrusion），对入侵报告数据统计处理得出 $P(ES\mid Intrusion)$ 和 $P(ES\mid\neg Intrusion)$，于是可以计算出

$$P(ES)=((P(ES\mid Intrusion)-P(ES\mid\neg Intrusion))\cdot P(Intrusion)+P(ES\mid\neg Intrusion) \tag{11}$$

因此，可以通过事件序列的观测推算出 $P(Intrusion\mid ES)$。基于条件概率的误用检测方法，是基于概率论的一种通用方法。它是对贝叶斯方法的改进，其缺点是先验概率难以给出，而且事件的独立性难以满足。

3.2.2 基于状态迁移分析的误用检测方法

状态迁移分析方法以状态图表示攻击特征，不同状态刻画了系统某一时刻的特征。初始状态对应于入侵开始前的系统状态，危害状态对应于已成功入侵时刻的系统状态。初始状态与危害状态之间的迁移可能有一个或多个中间状态。攻击者执行一系列操作，使状态发生迁移，可能使系统从初始状态迁移到危害状态。因此，通过检查系统的状态就能够发现系统中的入侵行为。采用该方法的 IDS 有 STAT（state transition analysis technique）[26]和 USTAT（state transition analysis tool for UNIX）[27]。

3.2.3 基于键盘监控的误用检测方法

基于键盘监控的误用检测方法，假设入侵行为对应特定的击键序列模式，然后监测用户击键模式，并将这一模式与入侵模式匹配发现入侵行为[8]。这种方法的缺点是，在没有操作系统支持的情况下，缺少捕获用户击键的可靠方法。此外，也可能存在多种击键方式表示同一种攻击。而且，如果没有击键语义分析，用户提供别名（例如 Korn shell）很容易欺骗这种

检测技术。最后，该方法不能够检测恶意程序的自动攻击。

3.2.4　基于规则的误用检测方法

基于规则的误用检测方法（rule-based misuse detection），系指将攻击行为或入侵模式表示成一种规则，只要符合规则就认定它是一种入侵行为。Snort 入侵检测系统就采用了基于规则的误用检测方法[28]。基于规则的误用检测按规则组成方式分为以下两类：

(1) 向前推理规则。根据收集到的数据，规则按预定结果进行推理，直到推出结果时为止。这种方法的优点是，能够比较准确地检测入侵行为，误报率低；其缺点是，无法检测未知的入侵行为。目前，大部分 IDS 采用这种方法。

(2) 向后推理规则。由结果推测可能发生的原因，然后再根据收集到的信息判断真正发生的原因。因此，这种方法的优点是，可以检测未知的入侵行为，但缺点是，误报率高。

3.2.5　其它方法

由于篇幅所限，基于专家系统、模型误用推理及 Petri 网状态转换等的误用检测方法，请参见文献[25]。

3.3　混合型入侵检测

3.3.1　基于规范的检测方法

Ko 等[29]提出了一种介于异常检测和误用检测之间的入侵检测方法，称之为基于规范的入侵检测方法（specification-based intrusion detection），用于发现对系统特权程序的入侵行为。其基本原理是，用一种策略描述语言 PE-grammars，定义系统特权程序的有关安全的操作执行序列。每个特权程序都有一组安全操作序列，这些操作序列构成特权程序的安全跟踪策略（trace policy）。若特权程序的操作序列不符合已定义的操作序列，就进行入侵报警。这种方法的优点是，不仅能够发现已知的攻击，而且也能发现未知的攻击。

3.3.2　基于生物免疫的检测方法

基于生物免疫的检测方法，系指模仿生物有机体的免疫系统工作机制，使受保护的系统能够将“非自我”（non-self）的攻击行为与“自我”（self）的合法行为区分开来[30]。该方法综合了异常检测和误用检测两种方法，其关键技术在于构造系统“自我”标志以及标志演变方法。

3.3.3　基于伪装的检测方法

基于伪装的检测方法，系指将一些虚假的信息提供给入侵者，如果入侵者应用这些信息攻击系统，就可以推断系统正在遭受入侵；并且还可以诱惑入侵者，进一步跟踪入侵的来源。读者可以进一步参考文献[31]。

3.3.4　基于入侵报警的关联检测方法

目前，入侵检测系统的检测方法，基本上都是从检测可疑事件入手，这样就无法防止误报警和重复报警。对于各种报警信息没有进行关联分析，只能起到记录可疑事件的作用。如果报警信息量过大，会使安全管理人员无所适从，导致 IDS 的作用受到限制。并且，网络攻击者也会在攻击之前故意制造大量的可疑事件，以降低入侵检测系统的警觉，使真实入侵事件淹没在大量的可疑事件之中。因此，对报警信息的分析处理成为当前的研究热点。其研究方法可以分为三类[32,33]。第一类基于报警数据的相似性进行报警关联分析；第二类通过人为设置参数，或通过机器学习的方法进行报警关联分析；第三类根据某种攻击的条件与结果（preconditions and consequences）进行报警关联分析。基于入侵报警的关联检测方法，有助于在大量报警数据中挖掘出潜在的关联安全事件，消除冗余安全事件，找出报警事件的相关

度及关联关系，从而提高入侵判定的准确性。

4 入侵检测的发展趋势

4.1 体系结构演变

入侵检测系统的结构大致可以分为主机型、网络型和分布型三种。主机型和网络型入侵检测系统是一种集中式系统，但是，随着网络系统的复杂化和大型化以及入侵行为所具有的协作性[34]，入侵检测系统的体系结构由集中向分布式发展。不同 IDS 之间通过共享信息，协同检测复杂的入侵行为，如攻击策略识别[35]。除此之外，现代网络技术的发展带来的新问题是，IDS 需要进行海量计算，因而高性能检测算法及新的入侵检测体系也成为研究热点，高性能并行计算技术将用于入侵检测领域[36,37]。

4.2 安全技术综合集成

IDS 尽管能够识别并记录攻击，但不能及时阻止攻击，而且 IDS 的误报警造成与之联动的防火墙无从下手。要解决当前的实际网络安全需求,入侵检测系统将与弱点检查系统、防火墙系统、应急响应系统等逐渐融合，形成一个综合的信息安全保障系统[38]。例如，Securedecisions 公司研究开发了一个安全决策系统产品，集成 IDS、扫描器、防火墙等功能，并将报警数据进行可视化处理[39]。

4.3 标准化

标准化有利于不同类型 IDS 之间的数据融合及 IDS 与其他安全产品之间的互动。IETF（Internet engineering task force）的入侵检测工作组（IDWG）已制定了入侵检测消息交换格式(IDMEF)、入侵检测交换协议(IDXP)、入侵报警(IAP)等标准，以适应入侵检测系统之间安全数据交换的需要[40]。目前，这些标准协议得到 silicon defense、defcom、UCSB 等不同组织的支持，而且按照标准的规定进行实现[41]。开放源代码的网络入侵检测系统 Snort 也已经支持 IDMEF 的插件。因此，具有标准化接口的功能将是下一代 IDS 的发展方向。

4.4 安全性评估

近几年来，攻击者不仅攻击网络服务的主机系统，而且采取各种手段逃避 IDS 的检测，攻击网络入侵检测系统。有鉴于此，国外特别重视网络入侵检测系统的评估问题，期望提高网络入侵检测系统的顽健性。从 1998 年起，麻省理工学院的 Lincoln 实验室在美国国防部支持下，研究了离线环境下 IDS 性能评估工具，并开发了 IDS 评估基准数据集，该数据集得到广大研究人员的认可[42]。此外，IBM 公司的 Zurich 实验室也研究了一套 IDS 测评工具[43]。Cohen 博士从理论上探讨了攻击入侵检测系统的若干方法[44]。另外，一些黑客组织也研究入侵检测系统攻击技术和评估方法，包括著名的反 IDS 工具 Snot、Stick、Fragrouter、Whisker、nidsbench 等[45]。目前，IDS 评估还没有工业标准可以参考，评估指标一般包括： 可靠性。系统具有容错能力，可以不间断地运行。 可用性。系统开销小，不会严重降低网络系统性能。 适应性。系统具有模块化结构，易于添加新的功能，能够随时适应系统环境的改变。 实时性。系统能够实时地发现入侵企图、报警并采取相应响应措施。 准确性。系统具有较低的误报率和漏报率。 抗攻击性。IDS 产品本身难于被攻击者欺骗，能够很好地保护自身的安全[46~48]。未来，网络入侵检测系统的攻击技术与评估方法研究是一个热点[49,50]。

4.5 面向 IPv6 的入侵检测

目前绝大多数入侵检测系统是面向 IPv4 的。IPv6 是针对 IPv4 地址空间有限和安全性不够而提出的。随着 IPv6 应用范围的扩展，入侵检测系统支持 IPv6 将是一大发展趋势，如

Snort2.0 就增加了对 IPv6 协议的分析。

IPv6 扩展了地址空间，协议本身提供加密和认证功能，因此，面向 IPv6 的入侵检测系统主要解决如下问题： 大规模网络环境下的入侵检测：由于 IPv6 支持超大规模的网络环境，面向 IPv6 的入侵检测系统要解决大数据量的问题，需要融合分布式体系结构和高性能计算技术。 认证和加密情况下的网络监听：IPv6 协议本身支持加密和认证的特点，极大地增加了面向 IPv6 的入侵检测系统监听网络数据包内容的难度，极端情况下，甚至需要首先获得通信双方的会话密钥。

面向 IPv6 的入侵检测技术是未来几年该领域研究的主流。

5 结束语

本文讨论了入侵检测的基本概念与模型，研究分析异常、误用和混合型入侵检测方法，同时，就入侵检测的发展趋势进行了概括总结。随着攻击技术不断发展及网络环境的变化，入侵检测还有许多问题有待研究，如协同入侵检测体系、攻击意图识别算法、攻击模式自动获取、入侵实时响应、报警数据关联分析与可视化处理、IDS 自身安全和新型网络环境的入侵检测等。

参考文献：

[1] 卿斯汉. 密码学与计算机网络安全[M]. 北京: 清华大学出版社,2001.

[2] ANDERSON J P. Computer Security Threat Monitoring and Surveillance[R]. James P Anderson Co, Fort Washington, Pennsylvania, 1980.

[3] SPAFFORD E. Crisis and aftermath[J]. Communications of the ACM, 1989,32(6):678-687.

[4] STEVEN E,SMAHA. Haystack: an intrusion detection system[A]. Proceedings of the Fourth Aerospace Computer Security Applications Conference[C]. Washington: IEEE Computer Society Press, 1988.37-44.

[5] ELLIS J, *et al*. State of the practice of intrusion detection technologies[EB/OL]. http://www.sei.cmu.edu/publications/documents/99.reports/99tr028/99tr028abstract.html.

[6] DOROTHY E. Denning, an intrusion-detection model[J]. IEEE Transactions on Software Engineering, 1987, 13(2):222-232.

[7] CHEN S, TUNG B, SCHNACKENBERG D. The Common Intrusion Detection Framework-data Formats[R]. Internet draft draft-ietf-cidf-data-formats-00.txt, 1998.

[8] KUMAR S. Classification and Detection of Computer Intrusions[D]. Dissertation, Purdue University, 1995.

[9] HEADY R, LUGER G, MACCABE A, *et al*. The Architecture of a Network Level Intrusion Detection System[R]. Department of Computer Science, University of New Mexico, 1990.

[10] DOAK J. Intrusion Detection: The Application of Feature Selection-A Comparison of Algorithms, and the Application of a Wide Area Network Analyzer[D]. Department of Computer Science, University of California, Davis, 1992.

[11] 边肇祺等. 模式识别[M]. 北京: 清华大学出版社, 1988.

[12] LUNT T F, TAMARU A, GILHAM F, *et al*. A Real-Time Intrusion Detection Expert System (IDES)-Final Technical Report[R]. Computer Science Laboratory, SRI International, Menlo Park, California, 1992.

[13] VALDES A, SKINNER K. Adaptive model-based monitoring for cyber attack detection[EB/OL]. http://www.sdl.sri.com/projects/emerald/adaptbn-paper/adaptbn.html.

[14] TENG H S, CHEN K, LU S C. Adaptive real-time anomaly detection using inductively generated sequential patterns[A]. Proceedings of the IEEE Symposium on Research in Security and Privacy[C]. Oakland CA, 1990,12(4):278-284.

[15] 何华灿. 人工智能导论[M]. 西安: 西北工业大学出版社，1988.

[16] CARLA T L, BRODLEY E. Temporal sequence learning and data reduction for anomaly detection[A]. Proceedings of the 5th

Conference on Computer &Communications Security[C]. New York: ACM Press, 1998.150-158.

[17] CARLA T L, BROALEY E. Detecting the Abnormal: Machine Learning In Computer Security[R]. Technical Report TR-ECE 97-1, Purdue University, West Lafayette, 1997.

[18] LEE W,STOLFO S.Data mining approaches for intrusion detection[EB/OL]. http://www.usenix.org/publications/library/ proceedings/sec98/full_papers/lee/lee_html/lee.html.

[19] **胡侃, 夏绍玮.基于大型数据仓库的数据采掘:研究综述**[J].**软件学报**, 1998,9(1):53-63.

[20] LEE W, STOLFO S, MOK K. Mining in a data-flow environment: experience in network intrusion detection[EB/OL]. http://www.cs.columbia.edu/~sal/hpapers/kdd99-id.ps.gz.

[21] LEE W,STOLFO S,MOK K. A data mining framework for adaptive intrusion detection[EB/OL]. http://www.cs.columbia.edu/~sal/hpapers/framework.ps.gz.

[22] LEE W, STOLFO S J, MOK K. Algorithms for mining system audit data[EB/OL]. http://citeseer.ist.psu.edu/lee99algorithms.html. 1999.

[23] KRUEGEL C, TOTH T, KIRDA E.Service specific anomaly detection for network intrusion detection[A]. Proceedings of the 2002 ACM Symposium on Applied Computing[C]. Madrid, Spain, 2002. 201-208.

[24] LIAO Y, VEMURI V R. Use of text categorization techniques for intrusion detection[A]. 11th USENIX Security Symposium[C]. San Francisco, CA, 2002.

[25] **阮耀平, 易江波, 赵战生. 计算机系统入侵检测模型与方法**[J]. **计算机工程**, 1999,25(9):63-65.

[26] An extensible stateful intrusion detection system[EB/OL]. http://www.cs.ucsb.edu/~kemm/NetSTAT/doc/index.html.

[27] ILGUN K. USTAT: A Real-Time Intrusion Detection System for UNIX[D]. Computer Science Dep University of California Santa Barbara, 1992.

[28] The open source network intrusion detection system [EB/OL]. http://www.snort.org/.

[29] KO C, FINK G, LEVITT K. Automated detection of vulnerabilities in privileged programs by execution monitoring[A]. Proceedings of the 10th Annual Computer Security Applications Conference [C]. Orlando, FL: IEEE Computer Society Press, 1994. 134-144.

[30] Computer security & other applications of immunology[EB/OL]. http://www.cs.unm.edu/~forrest/isa_papers.htm.

[31] GRUNDSCHOBER S. Sniffer Detector Report[R]. IBM Research Division Zurich Research Laboratory Global Security Analysis Lab, 1998.

[32] NING P, CUI Y, REEVES D S. Constructing attack scenarios through correlation of intrusion alerts[A]. Proceedings of the 9th ACM Conference on Computer & Communications Security[C]. Washington, USA: ACM Press, 2002. 245-254.

[33] VALDES A, SKINNER K. Probabilistic alert correlation[A]. Proceedings of the 4th International Symposium on Recent Advances in Intrusion Detection[C]. Springer-Verlag, 2001. 54-68.

[34] A note on distributed coordinated attacks[EB/OL]. http://www.all.net/books/dca/background.html.

[35] NING P, XU D. Learning attack strategies from intrusion alerts[EB/OL]. http://discovery.csc.ncsu.edu/~pning/pubs/ccs03-ids.pdf.

[36] KRUEGEL C, VALEUR F, VIGNA G, *et al*. Stateful intrusion detection for high-speed networks[A]. Proceedings of the IEEE Symposium on Security and Privacy[C]. Berkeley, California, USA: IEEE Computer Society Press, 2002. 285-294.

[37] FISK M, VARGHESE G. An Analysis of Fast String Matching Applied to Content-Based Forwarding And Intrusion Detection[R]. University of California - San Diego, 2002.

[38] Intrusion prevention systems: the next step in the evolution of IDS [EB/OL]. http://www.securityfocus.com/infocus/1670.

[39] DARPA summary power point[EB/OL]. http://www.securedecisions.com/darpa.htm.

[40] Intrusion detection exchange format[EB/OL]. http://www.ietf.org/html.charters/idwg-charter.html.

[41] Software[EB/OL]. http://www.silicondefense.com/idwg/.

[42] DARPA intrusion detection evaluation [EB/OL]. http://www.ll.mit.edu/IST/ideval/index.html.

[43] DEBAR H, DACIER M, *et al*. An Experimentation Workbench for Intrusion Detection Systems[R]. IBM Zurich Research Laboratory, 1998.

[44] COHEN F. 50 ways to defeat your intrusion detection system[EB/OL]. http://all.net/.

[45] Anti-IDS tools and tactics[EB/OL]. http://www.sans.org/rr/intrusion/anti-ids.php.

[46] 刘美兰，姚京松．入侵检测预警系统及其性能设计[A]．第一届中国信息和通信安全学术会议论文集[C]．北京：科学出版社，1999.105-111.

[47] AXELSSON S. The base-rate fallacy and its implications for the difficulty of intrusion detection[A].Proceedings of the 6th Conference on Computer And Communication Security[C]. New York: ACM Press, 1999. 1-7.

[48] FAN W, LEE W, *et al*. A multiple model cost-sensitive approach for intrusion detection[A]. Proceedings of the Eleventh European Conference on Machine Learning[C]. Barcelona, Spain,2000. 142-153.

[49] PTACEK T, NEWSHAM T. Secure networks insertion, evasion, and denial of service: eluding network intrusion detection[EB/OL]. http://citeseer.nj.nec.com/ptacek98insertion.html.

[50] 卿斯汉，蒋建春．网络攻防技术原理与实战[M]．北京:科学出版社,2004.

作者简介：

卿斯汉（1939-），男，湖南隆回人，中国科学院软件研究所研究员，博士生导师，主要研究方向为信息系统安全理论和技术。

蒋建春（1971-），男，广西桂林人，工程师，中国科学院软件研究所博士生，主要研究方向为信息安全对抗、网格计算。

马恒太（1970-），男，山东临朐人，工程师，中国科学院软件研究所博士，主要研究方向为网络信息安全和分布式计算。

文伟平（1976-），男，湖南桃江人，中国科学院软件研究所博士生，主要研究方向为信息安全对抗、恶意代码。

刘雪飞（1975-），女，湖南宁乡人，中国科学院软件研究所博士研究生，主要研究方向为计算机网络、信息安全、数据挖掘。