模式识别环境下智能入侵检测系统的研究

文/王安娜

摘要

本文主要分析入侵检测的基本理论与实践技术，根据模式识别的基础原理和入侵检测的相似性特点，把模式识别的具体方法运用在入侵检测技术的可行性环节，同时通过近邻法、K 近邻法与 K 均值法等各种关键技术提出模式识别环境下智能入侵检测系统的理论模型，理论上可表明系统具备较高性能的检测率与较低的误报率。

【关键词】入侵检测系统 人工智能 模式识别

伴随着互联网技术的广泛应用，基于计算机网络的业务应用领域已经逐步深入到社会各行各业范围中，计算机网络的安全性能显得十分关键。计算机网络的安全定义主要包括保密性、完整性、可用性以及认证等四个重要环节。因为计算机网络在理念设计、实践部署以及实际应用过程中存在较大的缺陷，使得计算机网络安全服务无法得到满意的结果，所以研发安全可靠的信息安全互联网产品已经发展成为学术界领域努力的前进方向。入侵检测技术作为扩充计算机系统安全确保能力、提高信息安全基础架构完整性的关键性领域。因为入侵检测的操作过程需要面对复杂的网络环境与变化多端的攻击方式，这就需要入侵检测系统具备灵活性、主动性以及自适应性等优秀性能。模式识别环境下入侵检测技术已经逐步社会重点关注的方向，尤其是模式识别的实际运用，更加是提高入侵检测系统性能的重要方法。

1 入侵检测系统的概述

入侵检测系统的理论定义主要是指在入侵检测过程中所需要具备的各种基本软件与硬件的配置组合，其通过对计算机网络信息系统的实际工作状态进行实时性的有效监测，发现各种类型的攻击意图、攻击行为或者攻击后果同时作出相应的响应，从而可以确保计算机系统资源的安全性、运行性与可靠性。其主要功能分别表现在：监控行为、分析系统用户与执行活动；检测计算机系统的技术配置与操作漏洞；评估系统取决于计算机资源与数据信息的完整性；模式识别已具备的攻击行为、统计分析异常行为；对于操作系统进行日志的操作管理；模式识别违反安全策略的系统用户活动；系统响应入侵行为的事件等。

2 智能入侵检测技术

现阶段大部分入侵检测系统可以符合大部分系统用户的实际需求，然而在重点技术领域 (金融、商务以及军事等) 的实际应用仍然存在各方面问题，通常表现在：误报率比较高、报警信息比较多；缺少检测未知入侵行为的有效技术；自适应与自学习能力比较低；互操作性比较差，无法形成协同防御的完善体系等。人工智能技术的实际应用，为能够解决上述各种问题积累坚实的基础。模式识别技术的基本原理是：把一个输入模式和储存在计算机系统中的多个参考模式相互对比，寻找出最接近的参考模式，把这种参考模式所代表的类名作为输入模式的类名输出。模式识别技术能够分成学习与识别这两个具体过程。学习是为了构造识别系统而进行的一种行为，参考模式是通过学习之后确定的。在应用识别系统的过程中，必须实时更新参考模式以增强系统的自适应性，这需要对识别结果集进行学习。本质上，模式识别是对未知样本进行类归属判定的过程；而入侵检测也是将一个新的实例与原有的规则集进行比较归类的过程。两者工作机理非常相似。模式识别的应用对于改善入侵检测系统的识别精度、识别能力以及智能特性有着重要的影响。

3 智能入侵检测系统

智能入侵检测系统主要采取模块化思想进行设计，其中包含数据采集模块，特征提取模块，规则处理模块，分析检测模块和异常响应模块等。

系统各个模块的功能如下：

数据采集模块：实时采集计算机网络系统的原始数据信息，同时根据各自不同的网络协议进行解码操作，然后对解码处理之后的数据信息进行分片重组、流重组以及代码转换等种技术处理，还原数据包的原始数据含义与数据包相关之间的实际关系。

特征提取模块：对于数据采集模块直接采集得到的数据信息进行特征化选取，然后对信息数据进行向量化处理，最后生成待检测的数据样本。

规则处理模块：进行规则集的向量化与聚类处理工作。首先根据条读入的处理规则，对于各条规则进行向量化处理，获得一个规则向量集，然后对规则向量集进行聚类分析处理，在向量集规模较小的情况不需要进行聚类入生成精简的参考规则集。

分析检测模块：这是计算机系统的核心控制模块。把待检测的数据样本和参考规则集进行比较分析处理，从而确定是否出现入侵状况。具体的处理过程为：

（1）采取近邻法分析待检测的数据样本和参考规则集。

（2）当欧氏距离 d=0 的时候，即待检测的数据样本和参考规则集中某部分规则进行匹配处理，从而得到分析结果。

（3）当 $d \neq 0$ 的时候则采取 k- 近邻法进行二次检测处理，从而得到相应的分析结果。

（4）根据具体的分析结果从而判断分析待检测数据样是否出现异常行为。

（5）假如是异常行为，则会马上启动异常响应的处理措施，同时对原规则数据库进行更新操作；假如是正常行为，则直接退出。

异常响应模块：对于入侵行为作出响应 (报警、日志记录等)。

4 结语

入侵检测理论是防火墙技术、数据加密技术以及访问控制等各种传统安全技术的重要基础，作为网络信息安全防护体系的关键构成环节。入侵检测系统能够对计算机网络入侵行为作出相应的识别与响应，其不但能够检测来自计算机网络的实际攻击行为，也能够监督系统内部用户未经授权的访问活动。模式识别是处于不断提升发展的新型学科技术，其理论基础与应用范围也处于不断发展的阶段。本文提出将模式识别方法具体运用在入侵检测的技术领域中，把入侵检测的相关问题转变成模式识别问题来进行处理，这实际上是一种富有价值的技术解决方案。基于模式识别的入侵检测系统自适应 / 学习能力强、成本低和健壮性好，能有效提高系统的安全性。但是，本系统仍存在缺陷：为保证参考规则集的有效性和实时性，需要提取海量的对象行为特征；在高带宽的网络环境下，为缩短检测响应时间，对检测算法的时空效率提出更高的要求。这两点对入侵检测系统的效能来说具有决定性意义，如何快速构建入侵参考模式知识库、进一步提高检测算法的智能性和效率，将是进一步研究的方向。

参考文献

[1] 沟口理一郎，石田亨 . 人工智能 [Ml. 北京：科学出版社，2005.

[2] 蔡自兴，徐光祐 . 人工智能及其应用 [M]. 北京：清华大学出版社，2004.

[3] 简清明，曾黄麟，叶晓彤 . 粗糙集特征选择和支持向量机在入侵检测系统中的应用 [J]. 四川理工学院学报：自然科学版，2009,22(5).

[4] 赵丽萍 . 基于模式识别的入侵检测模型 [J]. 电脑开发与应用，2008,21(6).

[5] 胡煜 . 主分量分析法和 K 近邻法应用于基因芯片数据分析 [J]. 北华大学学报：自然科学版，2008,9(1).

作者单位

吉林化工学院 信息与控制工程学院 吉林省吉林市 132022