基于OPTICS聚类与集成学习的数据库入侵检测研究

摘要

本文提出一种基于OPTICS聚类以及集成学习实现的对数据库入侵检测的方法。对用户事务使用OPTICS聚类，通过对请求的中各属性的分析，计算异常事务，对异常事务继续由集成学习的进一步分析，即使用Bagging、Boosting、Stacking进行分类，最终创建用户行为特征库。特征库分为有效行为与危险行为，对未来产生的新的用户事务进行快速响应，对危险行为进行主动拦截。本文使用随机模型对该系统进行测试，已验证了该系统的有效性。

关键词

数据库安全；入侵检测；OPTICS；集成学习

0引言

在网络信息安全体系中，数据库的安全保护一直被人们所关注，数据库威胁主要分为外部攻击与内部攻击。外部攻击由外部未经授权的用户，利用系统漏洞进行非法的数据库访问，尽管传统的数据库安全服务能够提供一定的防御措施，有效抵御外部攻击，但是其却不易应对来自内部人员滥用权限造成的安全问题，此类攻击则被称为内部攻击，因此对于防范内部攻击目前变得尤为重要。

DIDS的引入

举例说明其他人的做法

针对此类问题，本文提出一种……使用了xxx说明本文系统的有效性

1相关理论

* 1. OPTICS聚类

OPTICS是一种不需要用户提供特定密度阈值的密度聚类算法，是基于具有噪声应用的基于密度的空间聚类（Density-Based Spatial Clustering of Applications with Noise）的改进算法，其并不显式地产生数据集聚类，通过输出簇排序，表达数据的基于密度的聚类结构。核心思想为：对于簇Ci中任意对象p，在其ε邻域Nε(p)中，至少存在MinPts-1个其他对象，其中ε代表欧里几何距离半径，MinPts表示一个簇中至少应含有的对象数量。另外，OPTICS计算中，存在核心距离(core-distance,cd)和可达距离(reachability-distance,rd)两个重要信息。

对象p的核心距离cdε,MinPts(p)定义为ε'，使得p的ε'-邻域刚好包含MinPts个对象，若p不是关于ε和MinPts的核心对象，则p的核心距离没有定义，表达式为

<公式>

对象p到另一个对象q的可达距离rdε,MinPts(p,q)定义为使得p从q密度可达的最小半径值，其中q必须是核心对象，并且p必须在q的领域内，

* 1. 集成学习

预备知识