学校编码：10384　　分类号＿＿＿密级＿＿＿

学号：X2013231092　　UDC＿＿＿＿

厦门大学

工 程 硕 士 学 位 论 文

基于改进 Apriori 算法的云平台安全审计系统的设计与实现

Design and Implementation of the Cloud Platform Security Audit System Based on Improved Apriori Algorithm

彭怡

指 导 教 师： 王鸿吉副教授

专 业 名 称： 软 件 工 程

论文提交日期： 2015 年 10 月

论文答辩日期： 2015 年 11 月

学位授予日期： 2015 年 12 月

指 导 教 师：＿＿＿＿＿＿

答辩委员会主席：＿＿＿＿＿＿

2015 年 10 月

厦门大学学位论文原创性声明

本人呈交的学位论文是本人在导师指导下，独立完成的研究成果。本人在论文写作中参考其他个人或集体已经发表的研究成果，均在文中以适当方式明确标明，并符合法律规范和《厦门大学研究生学术活动规范（试行）》。

另外，该学位论文为（　　）课题（组）的研究成果，获得（　　）课题（组）经费或实验室的资助，在（　　）实验室完成。（请在以上括号内填写课题或课题组负责人或实验室名称，未有此项声明内容的，可以不作特别声明。）

声明人（签名）：

年　月　日

厦门大学学位论文著作权使用声明

本人同意厦门大学根据《中华人民共和国学位条例暂行实施办法》等规定保留和使用此学位论文，并向主管部门或其指定机构送交学位论文（包括纸质版和电子版），允许学位论文进入厦门大学图书馆及其数据库被查阅、借阅。本人同意厦门大学将学位论文加入全国博士、硕士学位论文共建单位数据库进行检索，将学位论文的标题和摘要汇编出版，采用影印、缩印或者其它方式合理复制学位论文。

本学位论文属于：

（　　）1.经厦门大学保密委员会审查核定的保密学位论文，于　　年　月　日解密，解密后适用上述授权。

（　√　）2.不保密，适用上述授权。

（请在以上相应括号内打“√”或填上相应内容。保密学位论文应是已经厦门大学保密委员会审定过的学位论文，未经厦门大学保密委员会审定的学位论文均为公开学位论文。此声明栏不填写的，默认为公开学位论文，均适用上述授权。）

声明人（签名）：

年　月　日

摘 要

随着信息技术的不断进步，人们对于计算和信息存储的要求越来越高，而个人电脑的计算和存储能力都是有限的，云计算平台正是在这样的背景下应运而生，云计算平台可以显著降低企业和个人客户在计算和存储上的成本，但是云计算平台的安全性问题转而成为人们最为担心的问题。对于云平台服务提供商来说，在服务的过程中发现各种异常行为，并对各种异常行为进行审计是尤为重要的。

本文实现了基于改进 Apriori 算法的云平台安全审计系统，主要研究内容如下：

1、设计云平台 Agent 用于收集审计信息，云平台 Agent 主要用于采集云环境下的审计信息，在审计信息采集完成之后，需要将审计信息转化为标准格式并送入数据库存储。云平台 Agent 的分布式部署，提高云平台 Agent 收集审计信息的能力。

2、将审计信息以对称密钥加密方式进行存储，审计信息在云平台数据库服务器中存储，为了防止云平台服务提供商内部窃取资料或者篡改资料，需要对审计信息以加密方式存储并应支持密文检索，采用这种方式的目的是为了在半可信的云环境中确保审计信息安全。

3、实时审计和事后审计相结合，实时审计由本地规则库完成，事后审计由审计分析模块中的规则库完成。在事后审计中，传统 Apriori 算法的无效连接和比较次数较多，增加了系统负荷，本文提出了改进的 Apriori 算法，其核心思想是按照既定的最小支持度和最小信任值，得到频繁集，将其作为用户正常的行为模式，提高了算法的效率，降低了系统的负荷。

经过测试，本文实现的基于改进 Apriori 算法的云平台安全审计系统能够满足设计要求，具有友好的操作界面，易用性较强，可以在现实环境中使用。本文的研究期望能够对云环境下的安全审计理论能够起到一定的推动作用，对当前如火如荼迅猛发展的云计算领域在对自身和顾客的信息安全防护方面能够起到积极的作用。

关键词：云平台；安全审计；Apriori 算法

Abstract

Along with the advance of information technology, people more and more high to the requirement of computing and information storage, and personal computers, computing and storage capacity is limited, the cloud computing platform is precisely in this context came into being, cloud computing platform can significantly reduce the enterprises and individual customers in the cost of computing and storage, but the security issues of cloud computing platform to become the most concerned about. For cloud service providers, found in the service process of all kinds of abnormal behavior, and audit all kinds of abnormal behavior is particularly important.

This thesis implements the cloud platform security audit system based on improved Apriori algorithm, the main research content is as follows:

1, design a cloud platform Agent to collect audit information, cloud platform Agent is mainly used for collecting a cloud environment audit information, in the audit information acquisition is completed, will need to audit information is converted into a standard format and stored into the database. Cloud platform Agent distributed deployment, improve the ability of cloud platform Agent to collect audit information.

2, the audit information stored on the basis of symmetric key encryption, audit information stored in the cloud platform database server, in order to prevent internal cloud platform service providers to steal or tamper with the data, the need to audit information stored in encrypted way and should support the cipher text retrieval, in this way is for the purpose of the partially trusted cloud environment to ensure that the audit information security.

3, combining the real-time audit and post audit, real-time audit completed by local rule base, post audit by audit analysis module to complete the rules in the library. In post audit, the number of invalid connection and comparison of the traditional Apriori algorithm is more, increases the system load, this thesis puts forward the improved Apriori algorithm, its core idea is in accordance with the established minimum support and minimum confidence value, get frequent sets, as a normal user

behavior patterns, to improve the efficiency of the algorithm and reduce the load of the system.

After the test, this thesis implemented the cloud platform security audit system based on improved Apriori algorithm can meet the design requirements, has friendly interface, ease of use is stronger, can be used in the real environment. Hope this research to a cloud environment security auditing theory can play a role, for the current rapid development in the cloud computing in information security protection of themselves and the customer can play a positive role.

Key Words：Cloud platform; Security audit; Apriori algorithm

目录

第一章 绪论 1

1.1 研究背景和意义 1

1.2 国内外研究现状 2

1.3 主要研究内容 4

1.4 论文的组织结构 4

第二章 关键技术研究 6

2.1 云平台安全审计概述 6

2.2 云平台安全审计关键技术 9

2.2.1 Agent 技术 9

2.2.2 对称加密 11

2.2.3 审计分析方法 11

2.3 关联分析技术 12

2.3.1 关联分析技术概述 12

2.3.2 关联规则算法的分类 13

2.3.3 经典关联分析 Apriori 算法 13

2.4 本章小结 15

第三章 系统需求分析 17

3.1 信息收集功能 17

3.2 数据存储功能 17

3.3 数据分析功能 17

3.4 报警功能 17

3.5 响应功能 18

3.6 查询功能 18

3.7 报表功能 18

3.8 打印输出功能 18

3.9 本章小结 18

第四章 系统设计 19

4.1 系统架构设计 19

4.1.1 云平台典型架构 19

4.1.2 云平台安全审计模型 20

4.1.3 系统总体架构 25

4.2 系统详细设计 28

4.2.1 审计信息收集子系统详细设计 28

4.2.2 审计信息存储子系统详细设计 35

4.2.3 审计信息分析子系统详细设计 36

4.3 本章小结 40

第五章 系统实现 42

5.1 改进 Apriori 算法的实现 42

5.1.1 传统 Apriori 算法的不足 42

5.1.2 当前对 Apriori 算法进行改进的几种技术 42

5.1.3 本文对 Apriori 算法的改进 43

5.2 审计信息收集子系统的实现 46

5.3 审计信息存储子系统的实现 51

5.3.1 加解密及密钥管理的实现 51

5.3.2 密文检索的实现 53

5.4 审计信息分析子系统的实现 55

5.5 本章小结 59

第六章 系统测试 61

6.1 测试准备 61

6.2 功能测试 62

6.3 性能测试 65

6.4 本章小结 67

第七章 总结与展望 67

7.1 总结.....67

7.2 展望.....68

参考文献 69

致谢.....72

Contents

Chapter 1 Introduction 1

1.1 Research Background and Significance.....1

1.2 Research Statues at Home and Abroad 2

1.3 Main Research Contents.....4

1.4 Structures.....4

Chapter 2 Key Technologies.....6

2.1 Introduction of Security Audit of Cloud Platform 6

2.2Key Technologies of Security Audit of Cloud Platform 9

2.2.1 Agent Technology 9

2.2.2 Symmetrical Encryption.....11

2.2.3 Audit Analysis Method.....11

2.3Relational Analysis Technology 12

2.3.1 Introduction of Relational Analysis Technology.....12

2.3.2 Types of Relational Rules Arithemttic 13

2.3.3 Typical Apriori Arithemetic for Relational Analysis 13

2.4 Summary.....15

Chapter 3 Demand analysis of System……..17

3.1 Information gathering function 17

3.2 Data storage function.....17

3.3Data analysis function.....17

3.4 Alarm function.....17

3.5 Response function.....18

3.6 Query function.....18

3.7 Report function.....18

3.8 Print function.....18

3.9 Summary 18

Chapter 4 Design of Cloud Platform of Security Audit System 19

4.1 Design of System Framwork 19

4.1.1Typical Framwork of Cloud Platform 19

4.1.2 Security Audit of Cloud Model 20

4.1.3 General Framwork of System 25

4.2 Detailed Design of System 28

4.2.1Design of Sub-System for Audit Information Collection 28

4.2.2 Design of Sub-System for Audit Information Storage 35

4.2.3Design of Sub-System for Audit Information Analysis 36

4.3Summary 40

Chapter 5 Implementation of System 42

5.1 Achievement of Advanced Apriori Arithmetic 42

5.1.1 Defects of Traditional Apriori Arithemttic 42

5.1.2 Technologies to Develop Present Apriori Arithemttic 42

5.1.3 Development for Apriori Arithemttic in This Work 43

5.2 Achievement of Sub-system for Audit Information Collection 46

5.3 Achievement of Audit Information stored in sub-system 51

5.3.1 Achievement of Encryption&Decryption and Key Management 51

5.3.2Achievement of Ciphertext retrieval 53

5.4Achievement of Sub-system for Audit Information Analysis 55

5.5 Summary 59

Chapter 6 System Testing 61

6.1 Praparation for Testing 61

6.2 Function Testing 62

6.3 Performance Testing 65

6.4Summary 67

Chapter 7 Conclusions and Future Work 67

7.1Conclusions 67

7.2Future work 68

References 69

Acknowledgments 72

第一章 绪论

1.1 研究背景和意义

随着信息技术的快速发展，无论是企业还是个人的日常生活都充斥着各种各样的信息系统，这些信息系统为我们的工作和生活带来了极大的便利，但人们对电脑计算能力的要求和对海量信息的存储要求也是越来越高，正是在这样的背景下，云计算平台很快从概念进入到实际运用领域[1]。云计算通过把用户的计算和存储搬到云端来实现，极大地降低了企业用户和个人用户的成本投入，极大地提高了运算效率，并且云计算平台还具有按需服务、弹性伸缩的特点，对中小企业和个人用户来说，可以有效的平衡成本和收益，以在激烈的市场竞争中赢得一席之地[2]。

云计算平台的巨大应用价值吸引着各大厂商相继投入到这个领域之中，对于用户来说，他们在享受极大便利和低成本的同时，最为关注的还是信息安全问题，一个不安全的云平台很难获得用户的青睐，也不具有推广的价值，因此，各个云平台服务厂商无一不大力关注自身云平台的安全问题[3]。

CSA（cloud security alliance，云安全联盟）在其发布的《云安全指南》中对云平台的安全问题进行了导向性论述，主要包括以下几个主要内容：

1、在云计算平台中，对于未经授权的行为，应进行监督，云服务提供商应监控公共黑名单上的用户行为以及客户的网络活动以堵塞安全漏洞。

2、在云服务商内部，也存在内部人员对云平台恶意破坏的可能性，因此云平台服务商应完善自身的内部控制流程。

3、云服务商应完善自身基础设施建设，特别是在传统安全、容错、容灾、业务连续等方面。

总之，对云平台的安全保护来说，应加强技术防护，但也不能忽略管理的重要性[4]。本文正是在网络安全和风险评估的传统理论上结合云环境的特殊性提出基于改进 Apriori 算法的云平台安全审计系统。

实质上，云计算就是通过大规模的分布式计算为消费者提供相关服务，云平台由集群的虚拟化计算机组成，通过统一的接口以面向服务的形式为用户提供服

务，在现阶段，云计算平台提出了三种服务模式，即基础设施即服务、平台即服务和软件即服务，基础设施即服务提供包括数据处理、存储等服务，平台即服务提供给用户按自身需求在云服务提供商的平台上构建主机应用，软件即服务即通过云平台使用软件服务提供商提供的服务[5]。传统的网络安全和风险评估理论在云平台上同样适用，但云平台有其特殊性，本文提出的基于改进 Apriori 算法的云平台安全审计系统期望能够对云环境下的安全审计理论起到一定的推动作用。

云平台日志信息在云平台的安全运行，责任追究和事后查证等方面扮演着极为重要的角色，云平台安全审计机制应做到监控用户网络活动，记录、分析用户行为，及时发现潜在威胁的行为等，云平台安全审计机制是对防火墙和入侵检测的有效补充，有利于云平台的安全以保障用户的利益不受损失。

1.2 国内外研究现状

就目前来看，云计算平台已经成为各大公司的研究重点，首当其冲的是亚马逊公司，微软、谷歌也投入了巨大资源用于云计算平台的构建，云计算平台有可能成为下一代的网络入口，这也是各大主流厂商争相进入的主要原因。

除了商业用途以外，维基百科还构建了非商业用途的云平台，用户可以把信息在云平台上共享，也可以免费获得别人分享的各种资源，这种云平台一般被称为公共云平台。

云平台的未来发展前景不可限量，我国的很多企业也大力投入资源进行研究，中国移动投入了大量资金构建了先进的实验室和硬件设施，并提出了“大云”计划[6]。中国联通提出了和中国移动“大云”计划类似的“互联云”计划，在基础设施即服务、平台即服务和软件即服务三个领域展开了深入研究[7]。但和国外同类技术相比，我国的企业同亚马逊、微软等企业还有着相当大的差距。云计算平台的原理有时看起来并不复杂，但真正构建的时候却绝非易事，还有很多未知的领域有待研究和探索，信息在未来的重要性无需多言，我国企业还需快马加鞭迎头赶上。

网络安全一直都是随着网络的发展而同步发展的，Anderson（1980）首次提出了安全审计的思想，他认为网络安全审计就是按照一定的安全策略发现系统漏洞、入侵行为进而改善系统安全性能的过程。网络安全审计是记录与审查用户网

络活动的过程，其作用主要包含以下 5 个主要内容：1、警示潜在的攻击者；2、保证安全策略的运行；3、责任追究；4、评价和反馈；5、及时发现潜在的系统安全漏洞[8]。

自 Anderson 以后，网络安全审计的思想在不断发展和完善之中，至今已经形成了较为完整的理论体系，也有了一定数量的实际应用系统。

在 Unix 环境下实现的 syslog 机制提出了较为完备的安全审计方案，主要包括审计对策、安全要求、审计级别的区分、判断策略等内容。

微软公司设计了在 Windows 环境下的安全审计机制 SCE（Security Configuration Editor），其机制和 syslog 机制是异曲同工的。

NetlQ 公司开发的 WebTrends Firewall Suite 能够获取防火墙所产生的日志信息，并对这些日志信息进行全面的分析，最后为用户提供安全审计报表[9]。

GFI 公司开发的 LANguardSecurity Event Log Monitor 产品提供了微软 SCE 安全审计机制以外的安全审计功能，它能够提供对于入侵 Windows 系统的违规行为的审计，并为用户提供安全审计报表。

NFR 公司开发的 SLR（Secure Log Repository）可以实现 Windows、Unix、Linux 的跨系统安全审计，对不同系统的记录日志进行统一保存和审计。

思福迪公司研发的 LogBase 系统针对云平台，可以提供网络安全预警，设备故障检测、用户非法行为侦查等功能[10]。

在安全审计方面，我国企业研发的系统也具有较好的可靠性和易用性。汉邦科技集团研制的“信息安全综合强审计监控系统”能够满足基本的网络安全审计要求，还可以审查内部人员的操作[11]。

天融信公司研发的 TOPSEC Auditor 系统可以采集包括防火墙、入侵检测系统产生的日志，还可以采集路由器、交换机、操作系统产生的日志信息，在日志信息采集方面较为完善。

复旦大学开发的 S-Audit 系统将入侵检测和安全审计合为一体。

西安交通大学开发的 Jump 系统在日志信息采集、实时审计和事后审计等方面都能够发挥较大的作用[12]。

综上，在云平台的构建方面，我国还落后较多，但在网络的安全审计方面，我国在理论和实践上都已有大量的研究成果，但针对于云平台，由于云平台的特

殊性，在云平台收集审计信息的能力、半可信的云环境中审计信息安全、实时审计和事后审计相结合、算法效率等方面的研究都显得非常不足，本文提出的基于改进 Apriori 算法的云平台安全审计系统期望能够对云环境下的安全审计理论起到一定的推动作用。

1.3 主要研究内容

本文实现了基于改进 Apriori 算法的云平台安全审计系统，主要研究内容如下：

1、设计云平台 Agent 用于收集审计信息，云平台 Agent 主要用于采集云环境下的审计信息，在审计信息采集完成之后，需要将审计信息转化为标准格式并送入数据库存储。实现了云平台 Agent 的分布式部署，提高云平台 Agent 收集审计信息的能力。

2、将审计信息以对称密钥加密方式进行存储，审计信息在云平台数据库服务器中存储，为了防止云平台服务提供商内部窃取资料或者篡改资料，需要对审计信息以加密方式存储并应支持密文检索，采用这种方式的目的是为了在半可信的云环境中确保审计信息安全。

3、实时审计和事后审计相结合，实时审计由本地规则库完成，事后审计由审计分析模块中的规则库完成。在事后审计中，传统 Apriori 算法的无效连接和比较次数较多，增加了系统负荷，本文提出了改进的 Apriori 算法，对传统的 Apriori 算法的连接步进行了改进，提高了算法的效率，降低了系统的负荷。

1.4 论文的组织结构

本文共分七部分：

第一部分是绪论，主要包括研究的背景和意义，国内外研究现状，主要研究内容和论文的组织结构；

第二部分是相关理论综述，主要包括传统安全审计的标准和模型，云环境下安全审计和传统安全审计的区别等；

第三部分是云平台设计系统的需求分析，主要包括云平台信息收集功能；数据分析功能；报警功能；审计功能；响应功能；查询功能；打印输出功能；报表