学校编码：10384　　分类号____密级_____

学号：23220121153049　　UDC _____

厦门大学

硕 士 学 位 论 文

基于 Windows 操作系统下隐私数据保护系统的设计与实现

The Design and Implementation of Data Privacy Protection System Based on Windows Operating System

石 欣 晨

指导教师姓名： 吴顺祥 教授

专 业 名 称： 模式识别与智能系统

论文提交日期： 2015 年 5 月

论文答辩时间： 2015 年 5 月

学位授予日期： 2015 年 月

答辩委员会主席：_____

评 阅 人：_____

2015 年 5 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

本人完全了解厦门大学有关保留、使用学位论文的规定。厦门大学有权保留并向国家主管部门或其指定机构送交论文的纸质版和电子版，有权将学位论文用于非赢利目的的少量复制并允许论文进入学校图书馆被查阅，有权将学位论文的内容编入有关数据库进行检索，有权将学位论文的标题和摘要汇编出版。保密的学位论文在解密后适用本规定。

本学位论文属于

1、保密（　），在　　年解密后适用本授权书。

2、不保密（　）

（请在以上相应括号内打“√”）

作者签名：　　日期：　　年　月　日

导师签名：　　日期：　　年　月　日

摘 要

随着计算机技术的迅猛发展，计算机在人民生产生活中的地位日益提升，成为日常生活工作中不可或缺的一部分。计算机中经常会保存人们生活或者工作中的隐私文件，由于个人保护意识不强，或者被别有用心的人所利用，隐私泄露事件层出不穷。给个人财产、精神、人格权利都造成了不同程度的侵犯，这让我们深刻的意识到了电子数据隐私保护的重要性和必要性。于是，设计一款操作简单、界面友好、适用面广的电子隐私数据保护系统显得尤为必要。本文将隐私发现、隐私销毁、隐私保密三大主题融为一体，设计了一款基于 Windows 平台的电子数据隐私保护系统。

本文首先全面的介绍了信息安全领域近年来的国内外研究现状，阐述了电子数据隐私保护研究的背景、意义及创新。本文深入研究了国内外前沿的隐私数据保护技术，成功解析了浏览器使用痕迹、常用软件使用痕迹、系统历史痕迹、媒体播放痕迹、下载痕迹等，全方位地保护了用户隐私；然后详细地分析了 Windows 磁盘下 FAT 和 NTFS 两大文件系统，设计了文件系统解析引擎。同时，分析了 NTFS 的新特性 USN，并将其与数据库技术相结合设计了一种快速文件搜索引擎。在引擎基础上，设计了嵌入文档搜索、快速文件销毁、彻底清除已删文件、磁盘擦除等功能模块，实现了用户隐私数据的发现与彻底销毁；最后研究了数据加密技术，采用 AES 分组加密算法设计了一种文件加密方法。针对传统加密方式的缺陷和不足，深入研究了基于过滤驱动的虚拟磁盘加密技术，并对文件系统过滤驱动模型和著名的磁盘加密软件 TrueCrypt 的安全机制进行了深度的分析，设计了一种虚拟磁盘加密方法，实现了用户隐私数据的有效保密。

本文所设计的基于 Windows 操作系统下隐私数据保护系统，力争成为一款功能全面、性能稳健、界面友好、易于操作的产品。通过系统集成化、自动化的一键式操作，轻松实现对用户的隐私数据保护。

关键词：隐私保护；痕迹清理；文件搜索；文件销毁；数据加密

Abstract

With the rapid development of computer technology, role of computer in people's production and life in the increasingly, become an integral part of daily life and work. Computer often save people live or work in the privacy of files. Because of personal protection awareness is not strong enough, or attracked by criminals, the event of privacy leaks often happen. This may hurt people's property, spirit, personality rights violations in different degree, which allows us to the necessity and importance of protecting the privacy of the electronic data of deep consciousness.Therefore,design an electronic privacy data protection system which simple operation,friendly interface,wide range of application is particularly necessary.Privacy search,privacy destory,privacy protect is the three main themes.a design based on electronic data privacy protection system on Windows platform.

This paper first introduced the research status in China and abroad in recent years the field of information security, the electronic data privacy protection background, significance and innovation research.In this paper,the author researched the newest technology of the privacy and data protection both in China and abroad,seccessfully resolved the browser using traces,commonly used software,the system traces of history,media play-history,download history and so on.Protect the users' privacy all-around.And then analysed the FAT file system and NTFS file system of Windows disk type.Designed a file analytical engine system.And analyzed the new features of NTFS--USN,with the combination of SQL database technology to design a fast file search engine.Based on the search engine,design the embedded document search system,fast file destruction system,disk wipe system.Realize the user privacy data discovery and destroy.Then researched the technology of data encryption,based on AES encryption algorithm,design of a file encryption method.Aiming at the defects of the traditional encryption methods and problems, in-depth study of the virtual disk encryption technology based on filter driver, security mechanism and the file system filter driver model and the famous disk encryption software TrueCrypt to carry out in-depth analysis, design a virtual disk encryption method, realize the effective security of user privacy data.

Privacy and data protection system based on Windows operating system is designed in this paper, and strive to become a comprehensive functions, stable performance, friendly interface, easy operation of the product. Through system integration, the automation of the one button operation, easy realization of user privacy and data protection.

Keywords: Privacy protection; Privacy clear; File search; Disk wipe; Data encryption

目 录

CONTENTS

第一章 绪论

1.1 选题背景

计算机隐私数据（Computer Private Data）是指个人或单位在使用计算机后，在计算机本地或者网络上所遗留的电子信息隐私数据，包括在磁盘操作所遗留的历史痕迹（包括浏览器使用痕迹、媒体播放记录、下载记录、远程使用痕迹等）、用户个人资料（包括个人文档、影音文件、个人 E-mail 等）、自动表单密码、银行账户密码等。这些隐私数据以二进制的形式保存在个人电脑或服务器中，由于电子数据的特殊性质（信息量丰富、易拷贝、传播速度快等），如果不加以保护，个人隐私、机密很容易被泄露。一旦泄露可能会给个人财产、精神、人格权利造成了不同程度的侵犯。

《2014 年第 34 次中国互联网发展状况统计报告》调查研究表明，到 2014 年 5 月底，我国互联网用户数量达 6.32 亿，与 2013 年同期相比增加 1442 万人。互联网在走入寻常百姓家的同时，也给个人的信息安全带来了极大的隐患。各种计算机病毒、钓鱼网站、木马等成为了不法分子谋取私利的工具，各种个人隐私泄露的事情也层出不穷。在一些非法网站上，网络病毒和木马正在公开叫卖。同时全球大型企业（年收益超过 10 亿美元）检测到的安全事件数量达到 13138 起，较去年增 44%[2]。信息安全事件频发，管理及降低损害的财务支出成本也持续攀升。已检测的信息安全事件带来的财务损失正在迅速增加。仅在中国大陆与香港地区，失窃的知识产权或者商业机密的价值已经远超 10 亿美元。

2014 年 3 月 26 日，携程爆发"安全门"事件，引发一场"换卡潮"。根据报道，携程网的支付功能存在严重系统漏洞，许多用户的银行卡信息被曝光，用户财产安全遭到严重损坏。2014 年 5 月 11 日，优视浏览器被查出登录接口存在漏洞。用户只要通过 QQ、微信、微博等方式登录该浏览器，其提交的用户信息和密码都有可能被不法分子所利用。同时，部分手机病毒可拦截用户短信，终端用户面临网银资金被盗的问题。2014 年 10 月，摩根大通银行"7600 万"与"700 万"信息被泄露。此次事件波及了约 7,600 万户家庭和 700 万中小企业。窃取了包括用户姓名、家庭住址、手机号码和 E-mail 等信息。2014 年 12 月 3 日，智联招聘网站被曝出 86 万用户简历信息泄露的漏洞。通过该漏洞，黑客

可获取包含用户姓名，身份证号码，手机号码，毕业院校等各种信息。这些个人隐私数据泄露事件的发生，对当事人的个人财产、精神、人格都造成了不同程度的损害，保护个人隐私数据的任务既刻不容缓，又任重而道远。

纵观国内市场，现有 360 安全卫士、金山电脑卫士、百度安全卫士等相关的安全产品，其中 360 安全卫士拥有超过 4 亿的用户量，用户规模数最大，但这三者的主要功能基本上都集中与查杀木马、修复漏洞、清理插件、清理用户痕迹、电脑加速等方面，整个产品都侧重于个人电脑的安全防护，并非致力于个人隐私数据的保护。因此，研发国内空缺的电子数据隐私保护产品具有广阔的市场前景。

1.2 国内外研究现状

信息技术的快速发展使人们的生活发生了天翻地覆的改变。人们足不出户，就可以享受到信息化技术所带来的便利。但由于计算机信息具有易共享、易扩散的特性，使得个人隐私数据经常面临着被篡改、破坏、泄露的风险。信息安全作为一门新兴学科，历了一个漫长的历史阶段，从上个世纪 80 年代开始萌芽，90 年代以来得到深化，到了 21 世纪随着信息技术的不断发展，信息安全不再是单纯的计算机技术问题，而成为了计算机科学、管理学、法律等问题相结合的交叉学科。

在信息安全领域，国内外研究的重点主要集中在密码学、可信计算、网络安全、信息隐藏、云计算安全、电子数据取证及隐私保护等方面。

- 密码学的研究

密码学作为信息安全基础研究工作，为信息安全提供了良好的理论基础，是信息安全学科的发展的重要保障。当前国内外的密码学研究主要方向是基于数学的密码理论研究[3]，主流研究方向有：散列函数的研究、对称密钥的研究和非对称密钥的研究。

1）散列函数的主要功能是将任意长度的输入转换成特定长度的输出。国外著名的哈希函数有 SHA-1/256/384/512、MD5、RIPMD 和 HAVAL 等等[3]，Hash 函数在信息安全领域中多用于加密算法，主要体现在文件准确性校验、数字签名、鉴权协议等方面。文献[4] 提出了一种基于散列的 JPEG 2000 图像识别方法；文献[5]提出了一种新的基于非可逆矩阵的单向散列函数。

2）对称密钥是指加密密钥和解密密钥必须相同[3]。国外著名的对称密钥加密算法有 DES 加密算法、IDEA 加密算法、Skipjack 加密算法、Rijndael 加密算法等以及新的数据

加密标准 AES。文献[6]将混沌序列引入流密码的设计；文献[7]采用混合 AES-DES 算法来增强数字运动图像传输的安全性。

3）非对称密钥（公钥密码）是指加密和解密使用两个不同的密钥并且不可能有加密密钥（公钥）推导出对应的解密密钥（私钥）的一种密码体制。国外著名的公钥密码体制有 RSA、HM 背包、Rabin、Elgamal、椭圆曲线密码、MeEliece 等[3]。国外著名的数字签名有 RSA 签名、Rabin 签名、ElGamal 签名、Schnorr 签名和美国国家数字签名标准 DSS，数字签名应用十分广泛，例如代理签名、不可否认签名、在线/离线签名、指定验证者签名、确认者签名等等[3]。文献[9]南相浩教授提出了组合公钥技术（CPK），在椭圆曲线的基础上实现了将标识作为公钥的标识认证算法，解决了密钥托管的问题。

- 可信计算的研究

可信计算技术指的是既要保证计算结果的准确性，又要保证计算过程的安全性。1983 年，美国国防部制定了《可信计算机系统安全评价准则》TCSEC。至今，美国已研制出包括安全操作系统、安全数据库、安全网络不见在内的百余种达到 TCSEC 要求的安全系统[10]。1999 年，多家著名 IT 公司联合发起成立了可信计算平台联盟，标志着可信计算进入迅速发展阶段。在国内，可信计算领域起步也不晚，2004 年第一款拥有自主产权的可信计算平台出现，2005 年联想集团的“恒慧”芯片和可信计算机相继研发成功，由此引发越来越多的企业加入可信计算的研发行列之中，各高校也展开了可信计算的研究。文献[11]由武汉大学、华中科大与 HP 公司合作提出了一种改进的增强网格安全的可信计算；文献[12]提出一种基于可信计算技术的改进的无线网格安全体系结构。

- 网络安全的研究

网络安全是指通过网络管理控制和采取相关技术措施，保证网络环境中的用户数据的完整性、保密性以及可使用性，主要包括物理上的网络自身的安全性，和逻辑上的网络信息的安全性[13]。在网络安全领域，国内外研究主要集中在网络认证授权、网络入侵检测、网络内容安全、防火墙、虚拟专用网、网络安全审计、网络应急、公开密钥基础设施等方面并研发了大量的相关产品[3]，已经形成一定的规模，可以预测，基于网络的安全技术将是未来信息安全技术发展的趋势和研究热点。文献[14]提出了一种具备持久性防御暴力破解能力的 RSA 体系可信网络通信系统；文献[15]研究了多类 SVM（支持向量机），并将其数据融合应用在网络安全态势感知方面；文献[16]将图论和神经网络相结合来分析网络的安全性。

- 信息隐藏的研究

信息隐藏技术主要分为隐蔽信道技术和多媒体信息隐藏技术[17]。隐蔽信道技术又分为潜信道和隐信道，潜信道是建立在公钥密码体制的数字签名和认证上的一种隐藏信道。它的宿主是密码系统。隐信道是在公开信道中建立起来的一种进行隐蔽通信的信道，为公开信道的非法拥有者传输秘密信息[17]。文献[18]提出了一种基于相邻像素分析的改进的 LSB 信息隐藏方法。

多媒体信息隐藏是以多媒体信号作为宿主载体，利用多媒体数据的数据冗余和人们的视觉、听觉冗余来隐藏机密信息的技术[17]，主要的研究工作集中在信息隐藏的数字水印算法、多媒体版权保护、多媒体认证、信息隐藏协议等方面，应用前景广泛，是一类新的信息安全技术，近些年受到研究者们的密切关注。文献[19]实现了一种基于 DCT 域的快速变换的数字水印算法并运用在数字文献的版权保护中，具有较好的鲁棒性。

- 云计算安全的研究

云计算安全是计算安全下的一个分支，与网络安全以及信息安全息息相关。指的是通过一系列的政策支持、安全技术手段和控制部署手段来保护云端数据、应用程序和云计算相关的基础设施。云计算安全问题可分为两大类：一类是云计算数据安全，另一类是云计算服务供应商所面临的安全问题。文献[20]提出一种新的云计算安全管理框架，使得云提供商和云计算使用者得到了安全保障。文献[21]提出在云计算安全中嵌入强密码认证措施。

- 电子数据取证的研究

电子数据取证技术研究的是各种电子信息在存储介质上的存储和运行特性，通过静态取证、动态仿真、网络取证、远程取证等取证分析技术，客观、真实地展现用户的行为痕迹，深入挖掘出电子数据之间的关联关系，为解决司法诉讼及民事纠纷等问题提供依据。当前有关电子数据取证技术的研究成果主要体现在如下四个方面：

1）电子数据取证技术与云计算技术相结合的研究与应用。文献[22]提出一种基于日志模型的云取证系统，通过这种模型可以快速的获取取证信息，降低取证的复杂度；文献[23]提出元数据标记的方法跟踪原始取证信息并借助于云计算来应对取证中出现的数据篡改等情况。

2）移动设备上电子数据取证技术的研究与应用。文献[24]提出了借助于手机传感器信息的取证技术；文献[25]对 Android 系统的即时通讯软件进行研究，提出一种基于 Android 系统的即时通讯记录取证方法。