计算机数据安全系统的研制

刘惠珍 张民 周昕

（北京工业大学计算机科学系，100022）

摘 要 概述了计算机数据安全系统涉及的基本概念，基本理论和主要技术方法．建立了一个数据文件加密/解密工具箱．该工具箱具有良好的人机界面和多个功能模块，它为数据文件的安全系统继续研究提供了有效的工具．

关键词 密码，加密，解密

分类号 TP309

目前，以计算机为中心的数据传输、存贮、处理和网络迅速发展，人们的各种活动都通过信息系统紧密地联系起来，信息系统中存贮和处理大量的重要数据，若无适当的安全措施，这些数据很容易被截取，篡改和删除，造成巨大损失．因此，在信息社会中，信息系统的安全和保密是十分重要的，是整个社会安全与稳定的重要因素，这给计算机数据安全保密学研究带来巨大的推动力．

1 计算机数据保密系统的研究

1.1 密码学与数据安全

保密学(cryptology)是研究密码系统或通信安全的科学．它包含两个分支，即密码学(cryptography)和密码分析学(cryptanalytics)．密码学是对信息进行编码实现隐蔽信息的一门学问，密码分析学是研究分析破译密码的学问．两者相互对立，又互相促进．在计算机系统的安全性研究中，密码学在数据存贮和真实性鉴别方面具有增加保护措施的功能．

在用户向计算机系统注册时，计算机要对用户进行真实性鉴别．鉴别方法是要求用户给出口令，这是密码学在计算机安全问题上应用最广的地方．

密码学的另一个的用途是对文件进行加密保护．在没有加密保护措施时，用户的任何文件对系统管理人员都是一目了然的，而加密后的文件，一般很难看懂，只有用技术手段将它转化为明码文才能窥见其内容．密码学不能取代一般的文件保护措施，但只要适当地与系统的其他保护功能密切配合，就可以大大增加系统的整体安全性．

密码学在计算机安全中的另一个作用是对外存的保护，例如：对磁盘和磁带的保护．如果对这些可携带的存贮设备进行加密，使得只有当初给它们写入数据的驱动器能够读它们，别的驱动器无法读取．那么，它们的安全性就处在系统的控制下，哪怕盗走磁盘、磁带，

收稿日期：1996—11—01

也无法得到正确信息.

传统密码体制 (cryptographic system)所用的加密密钥和解密密钥相同[1], 或实质上等同, 称其为单钥或对称密码体制 (one-key or symmetric cryptosystem). 若加密密钥和解密密钥不相同, 从一个难于推出另一个, 则称为双钥或非对称密码体制 (two-key or asymmetric cryptosystem). 在信息传输和处理系统中, 除了意定的接收者外, 还有非授权者, 他们通过各种办法 (如搭线窃听、电磁窃听、声音窃听等)来窃取机密信息. 他们虽然不知道系统所用的密钥, 但通过分析可能从截获的密文推断出明文.

研制一种强的密码算法要把住两道难关: 设计和鉴定. 算法设计是制定准则和提出满足这些准则的待定算法. 算法鉴定则是对待定算法进行详尽的全面的严格的分析.

为了保护信息的保密性, 抗击密码分析, 保密系统应当符合下述要求: ①系统即使达不到理论上是不可破的, 也应该为实际上不可破的, 也就是说, 从截取的密文或某些已知明文密文对, 要决定密钥或任意明文在计算上是不可行的. ②系统的保密性不依赖于对加密体制和算法的保密, 而依赖于密钥. ③加密和解密算法适用于所有密钥空间中的元素. ④系统便于实现和使用方便.

1.2 信息加密基本方法及设计

密码学是对文件进行加密保护的主要理论. 数据文件可通过在文件名、属性、使用次数、位置、内容等方面进行加密. 对文件内容的加密有多种方法, 经典加密技术主要有单表代替、多表代替、换位技术等方法, 近代加密技术有序列密码体制和分组密码体制. 现代加密技术主要有 DES 算法和 RSA 体制.

代替密码, 它利用字母间对应的代换, 实现对明文信息的加密[1].

本课题当中的明文减密钥、明文加密钥、明文异或密钥及弗南姆 (Vernam)密码都属于多表代替法.

多字母代换体制(矩阵系数密码), 是对两个以上字母进行代换的密码体制, 优点是容易将字母的自然频度隐蔽或均匀化而有利于抗击统计分析.

分组密码把明文消息 M分析成相连的分组 M_1, M_2, ……并用同一个密钥 K对分组 M_i 进行加密, 即 $E_k(M) = E_k(M_i)\cdots$, 而序列密码则把信息 M拆成了相连的字符式比特 m_1, m_2, …, 并用序列密钥 $K=K_1, K_2, \cdots$, 的第 i 个成份对信息序列的第 i 个成份进行加密.

换位密码, 又称转置密码, 它只是对明文消息所含全部字母在文中的位置加以重排列. 例如, 在美国数据加密标准 (DES)对信息的加密过程中, 就多次利用字组的重排 (换位)来提高非线性加密变换的保密程度.

代码密码是对明文中较长的语言单元, 如词或短语进行代换, 从信息论的观点出发, 用较短的字符来代换那些经常出现的语言单元, 而用较长的字符代表那些很少出现的词和短语. 实现了明文的压缩. 数据压缩技术本身起到了数据保密的作用.

DES 是美国数据加密标准[2], 它是当前世界最广泛应用的一种分组加密体制. 数据分组长度为 64bit, 密文分组长度也是 64 bit, 没有数据扩展. 密钥长度为 64 bit, 其中有 8 bit 奇偶校验, 有效密钥长度是 56 bit. DES 的整个体制是公开的, 系统的安全性全靠密钥的保密. 算法主要包括: 初始置换 IP; 在密钥控制下的 16 轮迭代的乘法变换; 逆初始置换 IP^{-1}. 自 DES 正式成为美国标准以来, 已有许多公司设计并推出实现 DES 算法的产品.

1.3' 双钥体制的 RSA 算法

为了解决传统密码体制所面临的困难，美国学者 Diffie 和 Hellman 于 1976 年提出了公开密钥的新型密码体制[3]. 公钥密码体制因其工作基础是利用了单向函数的单向性，所以加解密的计算过程较复杂. 该体制的理论基础是数论中的下述论断：要求得两大素数（如大到 100 位)的乘积，在计算机上很容易实现. 但要分解两个大素数的乘积（即以乘积求它的两个素因子)，在计算上几乎不可能实现.

密码学还可以对模拟消息进行加密，如对语音、传真和电视图像等的保密，这是通过对这些消息的模拟信号进行变换来实现.

1.4 密码分析

密码分析可以按密码分析员掌握的关于密码系统的知识分级①密码分析员除密钥外，掌握密码系统的加密和解密算法. ②“仅知道密文的攻击”——密码分析员能够搜集到密文信息. ③“已知明文的攻击”——密码分析员搜集到某些明文和与之对应的密文信息. ④“选择明文攻击”——密码分析员可以有选择地搜集到某些明文和与之对应的密文信息. ⑤密码分析员可以（象合法用户那样)发送加密信息. ⑥密码分析员可以截取或重新发送信息. 密码学的方法实际上属于数学方法. 它所产生的结果是恒定的. 如以密钥 B 加密明文 E 得出密文 F，这在逻辑上肯定为真.

密码分析必须凭经验，靠测试来获取有关的材料. 因此，密码分析学的方法不依赖数学逻辑的恒定不变的真理，而依赖从客观世界觉察得出的事实

密码分析学的经验特性表现在它的作业方法上，作业方法包括通常称为“科学方面”的 4 个步骤. 它们是：分析（例如统计字母)；假设（A 可能是 B)；推断（如果 A 是 B，则应出现可能有的明文)；证实（它们正确)或否定（它们错误，因此 A 可能不是 B)，这两种情况都会导出一连串新的推断.

一般，密码分析学用两种方法进行作业：演译法和归纳法. 演译破译法是破译任一密码制的一般方法，它以频率分析为基础. 归纳破译法是以可能字，或以额外条件（如相同明文的两份密报)为基础的方法，是较特殊的破译法.

基于频率分析的破译法，是应用已知的字母频率分布特性分析已获取的密文. 如对英文单表代替系统的破译，频率分析的一种典型演译法是，以“密文中频率最高的字母大概是代替 e 的字母（频率分析的结果是字母 e 出现频率最高)”作为它的大前提，以“x 是密文中频率最高的字母”作为它的小前提，而以“x 可能是代替 e 的字母”作为它的结论，由于一般语言都有清晰的字母频率特征，所以这种演译方法非常适用于简单的移位和单表代替等密码体制.

另外，归纳法又只有在某些条件得到满足时才有效. 因为密码分析人员在已经获得密文并了解它的其它信息之前，不能确认这些条件是否已确实满足. 在用归纳法破译密文时，由于可能字和特殊情况能使密码分析人员取得有关的额外情报和信息，所以这种破译法显示了巨大的效力和收益，而且往往是在新体制中首先获得成功的方法.

1.5 密钥在保密系统中的作用

密钥的重要意义已日益为人们所重视，当前实用的电子密码机多为对称性密码. 它主

要由硬件和软件两部分组成，硬件提供编码的基本逻辑或算法，是一个相对固定因素，也是保密的基本因素，可称之为静态密度；软件，即密钥[4]，是随密码技术的发展而逐渐成为编码领域中的一个独立成份的，它提供编码逻辑的变化参数，它与密码机的使用和变化相联系，是保证密度的可变因素，可称之为动态密度.

现在加密设备的保密度取决于密钥以及密钥管理，而不取决于算法的保密及硬件的加密过程的保密. 其理由有以下 3 点：

1)一种密码，无论理论上的密度多么强，如果实际使用不当，也可能丧失保密. 而密钥(动态)保密则是实用保密的重要环节.

2)对于没有独立设计和生产能力的用户，只能使用商品密码. 其硬件已无密度可言，就必须依靠密钥保密.

3)专用密码机也有泄密的可能，而一部现代密码机从设计到生产耗资几十万，需时多年，不可能一泄密就更换. 因此，也要准备靠密钥的变化来实现保密.

1.6 计算机数据保密系统的研究

计算机加密/解密涉及理论深，范围广，每一种加密技术都是由各类基本方法综合发展形成，因此从基本理论、基本算法入手，结合当前常用的加密解密手段，逐步深入研究解决，是我们解密研究的重要途径.

我们从传统加密方法，近代加密方法和现代加密方法 3 方面进行信息加密. 根据以上编码方法，实现了单表代换，多表代换的多种运算及弗南姆算法，多字母代换的普雷发算法. 在现代密码方法中，我们研究了快速面向 32 位微处理器的数据加密算法；这种算法采用 128 位的分组密钥进行加密，其中运用了异或、可变位旋转运算，从而具有较高的安全强度和较快的加解密速度. 我们还研究了基于"滑动窗"的数据压缩方法，实现代码变换.

对于解密技术，我们应用了目前国际上成熟的概率论 密码分析论和贝叶思定理等理论， 通过对两种以上的加密算法进行研究，利用统计规律和穷尽原则来测试、分析、跟踪、计算，以实现解密，进而探索现代密码技术研究途径.

为充分满足用户要求，系统把全部程序集合成窗口界面，对英文文本文件 8 种类型和汉字文本文件 4 种类型方法实现加密/解密. 对于给定明文，输入密钥，生成密文，完成加密操作. 对于解密操作，若给定密文，在和本系统相同算法下，有的能自动进行解密，恢复明文，有的会给出有关参考信息，如密钥长度和可能的密钥. 除此以外系统还提供了试探解密法的窗口，供用户选择.

数据文件加密也可以通过磁盘结构加密技术进行. 我们深入研究了特殊的常驻程序调用和中断服务程序及反跟踪技术，完成了主要针对非标准格式化软磁盘在结构上的 3 种变化(非法磁道和附加磁道，非法扇区，错误的 CRC 校验)进行测试、分析和恢复破解工具. 我们还研究了复杂的无缝锁技术和伪随机数加密技术，制作钥匙盘和对可执行文件加密/解密的工具.

我们研究的数据文件加密解密系统是一个功能完备，使用方便的计算机加密、解密工具箱，利用此工具箱，用户可以采用多种方法(传统密码技术和现代密码技术)对数据文件进行加密和破译分析，同时，也对目前常用的软磁盘结构加密解密系统提供了一个较好的磁盘结构测试分析工具. 整个系统是在 PC 机上 DOS 环境下，采用 C 语言和汇编语言实现.

该系统模块性强，便于功能扩充，具有良好的人机交互能力.

2 结束语

计算机加密解密理论涉及概率论、统计学、数论，等数学分支还涉及到集合论、信息论、计算复杂性理论等，是一门综合的学科，在技术上涉及到计算机技术、硬件保护技术、无线电技术等多方面，此项研究有很深远的理论意义和学术价值. 我们的研究虽然取得了一些成果，但是还有很多工作需要做，无论是数据文件的加密解密技术还是磁盘结构的加密解密，都需要进一步扩充方法，完善功能，提高水平.

参 考 文 献

1 王化文. 计算机安全保密原理与技术. 北京：科学技术出版社，1993. 29～50
2 魏仲山. 计算机信息保护. 天津：天津大学出版社，1989. 73～80
3 曹珍富. 公钥密码学. 哈尔滨：黑龙江教育出版社，1993. 19～25
4 陈爱民. 计算机安全与保密. 北京：电子工业出版社，1992. 112～122

Development of the Computer Data Security System

Liu Huizhen Zhang Min Zhou Xin
(Department of Computer, Beijing Polytechnic University, 100022)

Abstract This article outlines the basic concept, theory and main technical methods on the computer data security system. A toolbox for encrypting and decrypting of data file has is established. The toolbox has friendly user interface and clear modules. It provides effective tools for the study on computer data security system in future.

Keywords security code, encryption, decryption