一种简易汉字加密与解密算法的设计与实现

刘瑞芳，梅孝安
（湖南理工学院，湖南 岳阳 414000）

摘要：本文通过对常用的加密与解密算法的研究，总结了汉字加密与解密技术意义，介绍了汉字加密与解密的原理，并运用异或运算原理在 VB6 中实现汉字的加密与解密。

关键词：加密解密技术；信息安全；密钥；异或运算；Visual Basic
中图分类号：TP314　　**文献标识码**：A　　**文章编号**：1009-3044(2006)32-0159-01

Design and Implementation of a Simple Algorithm of Chinese Character Encryption and Decryption

LIU Rui-fang, MEI Xiao-an
(Hunan Institute of Science and Technology, Yueyang 414000,China)

Abstract:Through to the commonly used encryption and decryption algorithm research, the paper summarizes the importance of the Chinese character encryption and decryption, introduces the principle of the Chinese character encryption and decryption, and implements the Chinese character encryption and decryption by using XOR operation in VB.

Key words:Encryption(Decryption) technology; information safe; XOR operation; Visual Basic

1 引言

在网络化的今天，网络信息的安全显得越来越重要，计算机犯罪每年使各个领域遭受的损失极其巨大，而且还在发展中。密码是有效可行的保护信息安全的办法。由于在现实生活中，我们要确保一些敏感的数据只能被有相应权限的人看到，要确保信息在传输的过程中不会被篡改，截取，这就需要很多的安全系统大量地应用于政府、大公司以及个人系统。所以，对加密技术、解密技术方面的研究是必要也是必须的，此技术的发展潜力也是巨大的。我们现在、将来都会在各方面利用加密技术、解密技术。

加密技术是保障通信安全的一种重要手段，汉字的加密与解密又是我国信息安全中最基本和最常用的方法。然而，现有的加密方法很少有直接对汉字进行加密的，因此研究汉字文本在传输时的加密就很有必要。汉字的加密与解密又是我国信息安全中首先而且必须研究的课题。

2 异或法汉字加密解密算法原理

汉字不同于一般的数据，因此，在对汉字的加密解密算法上就有特别之处。我们知道一个汉字就是 2 个字节，1 个字节都有自己 ASCII 码，一个汉字为二个 ASCII 码，也可以说汉字本身是用 ASCII 码构造出来的。汉字在电脑中是以 0 和 1 表示的，如果我们对其进行异或的话，A xor B=C，A 为原文件，B 为密钥，C 为加密后的文件，解密时，只要（C xor B=A）就可以了。从加密的主要方法看，换位法过于简单，特别是对于数据量少的情况很容易由密文猜出明文，替换法则不失为一种行之有效的简易算法。首先，Windows 下的字符集采用 Unicode 字符集，它容量大，可置换的范围广；其次，在 Unicode 字符集中，所有字符的内码都占两个字节，不再象 DOS 下西文字符占一个字节，中文字符占两个字节，这样无论西文还是中文都可以互换。因此，将某个字符的高字节和低字节分别加以运算，生成另外一个 0-255 之间的数，然后再将它们合成为另一个字符，从而置换数据达到数据加密的作用，解密时则相反。

从各种位运算的特点看，异或运算最适合用于简易加解密运算，因为当一个数 A 和另一个数 B 进行异或运算会生成另一个数 C，如果再将 C 和 B 进行异或运算则 C 又会还原为 A。如：128 和 253 进行异或运算的结果是 125，125 和 253 再进行异或运算则结果又是 128。其中，128 就是要加密的数据，253 则是密钥。利用这个特性可以将加密和解密用一个函数实现。用同一密钥进行奇数次运算时，是对数据进行加密运算，当进行偶数次运算时，是对数据进行解密运算。如果对 Unicode 字符进行这种运算，需要两个密钥，其取值范围为 0-255，其中一个用于对高字节加密，一个用于对低字节进行加密，这样对同一字符的加密就有 255×255=65025 种可能。如果更换密钥后对密文再进行加密，则会有无穷的可能性。如：密钥 K1=68，K2=134，则字符 A 的低字节为 65，它和 K1 异或后为 5，A 的高字节为 0，它和 K2 异或后还是 134，两者合成的字符为"衢"。再如："密"的低字节为 198，和 K1 异或后为 130，它的高字节为 91，和 K2 异或后为 221，两者合成则为一个不可见的字符。这种加解密数据的方法对任何字符都是有效的，不象有些简易加密算法，只对西文字符有效，对中文加密后再解密无法还原为原来的字符。

3 汉字的加密与解密的具体算法及实现

3.1 算法原理

一般的加解密算法如 DES 和 RSA 等过于复杂，且运算速度慢，特别是它的移位操作，实现起来也比较困难。而一般的汉字的加密与解密对于 VB 这样的高级语言实现起来也比较容易。研究简易汉字加解密解密算法还是有相当的现实意义。在 VB 内部提供的 XOR 函数，它既可对字符，也可对数字、布尔变量进行异或，两次异或的结果即为原值。因此，它是基本加解密函数。例如：Ascw("息")XOR 28 值为 24943，24943 XOR 28 值为 24687，Chrw(24687)值为"息"。在 VB 内部，字符全部作为 Unicode 处理，并且 VB 支持三种类型的字符集，并提供了相应的字符处理的基本函数。经过反复实验证实，在 VB 中使用 Unicode 字符集及 Unicode 字符集函数可较好的加密汉字。另外在对包含有汉字的文本（明文、密文）文件（.txt）进行读、写处理时，要考虑汉字的问题，主要是长度问题，用不同的语句/函数，读出的结果不一样，有的超出实际长度。

3.2 哈希函数构造

散列函数（也叫哈希函数）是密码学和数学中的一个概念。其作用是能够基于给定的输入字串、文件或其它类型二进制数据产生一个独一无二值。此外，该函数采用的算法能够保证人们不能从它反向推得该值的原始信息。对于 Hash 函数，其初始值（种子值）的选择至关重要。为了保证产生和检查散列代码的双方使用相同的初始值，可以选择固定的伪随机初始值。加解密方案中为保证发送方（加密方）和接收方（解密方）得到相同的随机种子值在 HASH 函数中调用 DoXor 之前必须初始化种子值（过程名为 initialize）：

（下转第 167 页）

收稿日期：2006-08-26
作者介绍：刘瑞芳（1972-），女，湖南省岳阳市人，湖南理工学院数学系讲师，研究方面：计算机应用；梅孝安（1973-），男，湖南省岳阳市人，湖南理工学院物电系讲师，研究方面：电子通信。

´ X´ from … … where… …);

通常绝大多数人会使用第一种格式，因为它比较容易编写，而实际上第二种格式要远比第一种格式的效率高。在 SQL 系统中几乎将所有的 IN 操作符子查询可以改写为 EXISTS 的子查询。在第二种格式中，子查询以 select ´ X´ 开始。运用 EXISTS 子句不管子查询从表中抽取什么数据它只查看 where 子句。这样优化器就不必遍历整个表而仅根据索引就可完成工作。使用 EXIST，SQL 系统会首先检查主查询，然后运行子查询直到它找到第一个匹配项，这就节省了时间。那么在执行 IN 子查询时，首先执行子查询，并将获得的结果列表存放在在一个加了索引的临时表中。在执行子查询之前，系统先将主查询挂起，待子查询执行完毕，存放在临时表中以后再执行主查询。这也就是为什么使用 EXISTS 比使用 IN 通常查询速度快的原因。同时应尽可能使用 NOT EXISTS 来代替 NOT IN，尽管二者都使用了 NOT，但 NOT EXISTS 要比 NOT IN 查询效率更高。

3 结束语

总之，有多种技术可用于对 SQL 数据库进行优化查询，提高 SQL 数据库的整体性能。以上技巧仅仅是进行测试实验之后的经验之谈，期望对各层次的数据库爱好者提供借鉴。

参考文献：

[1]王珊，陈红编著. 数据库系统原理教程[M]. 北京：清华大学出版社，1998.

[2]杜兆将，郭鲜凤，刘占文. SQL Server 数据库管理与开发[M]. 北京：北京大学出版社.

[3]李伟红. SQL Server 2000 实用教程[M]. 中国水利水电出版社，2003.

[4]彭林，余艳. SQL Server 2000 经典教程[M]. 人民邮电出版社，2001.

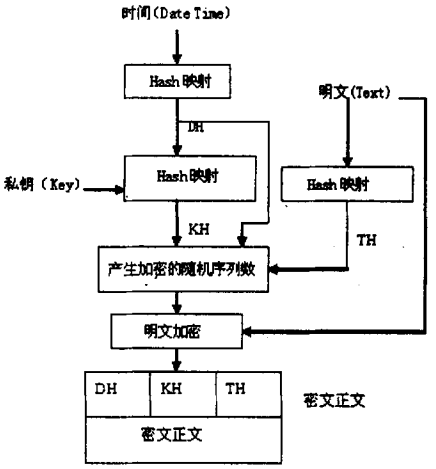
(上接第 159 页)

基于上面的知识，本加密解密方案中 HASH 函数如下：

```
Public Function Hash(ET As String) As String
Dim BitLenString as String，KeyString as String，
FileText  as String
BitLenString="12345678"
KeyString = ET & BitLenString
Call Initialize(KeyString) ´ 根据 KeyString 产生随机数序列
FileText = ET & BitLenString
Call DoXor(FileText) ´ 根据上述随机数序列对 FileText 加密
KeyString = FileText
Call Initialize(KeyString) ´ 根据上述的加密结果产生新的随机数序列
FileText = BitLenString
Call DoXor(FileText) ´ 根据上述随机数序列对 FileText 加密，8 位字符
Hash = FileText ´ 8 位字符送作 HASH 值
End Function
```

3.3 加密过程实现

本加密算法原理流程图如下：



加密数据变化过程为：

明文=Text

密文文本=HASH(DateTime)+HASH(DateTime+Key)+ Encryption(Text,HASH(DateTime)+HASH(DateTime+Key)

密钥=HASH(DateTime)+HASH(DateTime+Key)

加密后的密文文本=密钥+Encryption(Text, 密钥)

3.4 解密过程实现

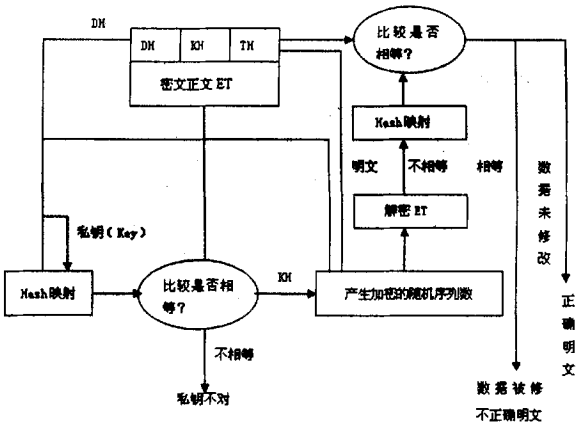
本解密算法原理流程，如下图所示。

解密密数据变化过程为：

密钥=从密文文本获取的密钥

密文正文 Text=从密文文本获取的正文

解密后的明文=密钥+Dncryption(Text, 密钥)



4 总结

本文设计的加解密方案将私钥进行 HASH 的映射（实际上也是加密）后，再与随机数序列等其它信息一起构成特定位数的加密字符串，去加密加密明文，使得加密强度与用户口令密码长度无关，并且加密方法更安全，总之，本加解密方案的特点如下：

(1)从技术要求上看，本算法安全可靠性高，加密后的密文很难被破解，因为加密时使用了时间标志，所以每次加密后的密文都不一样。这样，使用者可以对需要加密的密文放心，该算法达到了密码系统安全性的需求，而且还有比其它现有的密码系统更好的抗攻击性能；

(2)从本加密、解密的算法要求上来看，本算法正确、可读性好、时间复杂度和空间复杂度合适，程序代码精简适用，程序运行时内存占用量不大，与一般的加密程序的 CPU 的占用率相比是比较低的，满足算法中对高效率和低存储量的需求，该算法是一种用 VB 实现私钥加密技术的方法，具有比现有的公钥密码系统更快的计算速度；

(3)本加解密方案健壮性、稳定性好，当输入数据非法时，该方案也能正确地做出反应，并经多次逐一论证与分析，解密后的文字与原文一样。

参考文献：

[1]李克洪，王大玲. 实用密码学与计算机数据安全[M]. 沈阳：东北大学社出版，2001,63-65.

[2]Steve Burnett, Stephen Paine. 密码工程实践指南[M]. 北京：清华大学社出版，2001,137-139.

[3]Ross J. Anderson .信息安全工程[M]. 北京：机械工业出版社，2003,201-203.

[4]Oded Goldreich. 密码学基础(第二卷)[M]. 北京：人民邮电出版社，2005,175-177.

一种简易汉字加密与解密算法的设计与实现

作者： 刘瑞芳， 梅孝安， LIU Rui-fang， MEI Xiao-an
作者单位： 湖南理工学院,湖南,岳阳,414000
刊名： 电脑知识与技术（学术交流）
英文刊名： COMPUTER KNOWLEDGE AND TECHNOLOGY
年，卷(期)： 2006(11)

参考文献(4条)

1. Oded Goldreich 密码学基础 2005

2. Ross J Anderson;蒋佳;刘新喜 信息安全工程 2003

3. Steve Burnett;Stephen Paine;冯登国 密码工程实践指南 2001

4. 李克洪;王大玲;董晓梅 实用密码学与计算机数据安全 2001