陈洪军,景清武,李丹妮,等. 基于 DNA 和对称密码的一次一密加密算法[J]. 智能计算机与应用,2024,14(9):150-154.
DOI:10.20169/j.issn.2095-2163.240924

基于 DNA 和对称密码的一次一密加密算法

陈洪军, 景清武, 李丹妮, 石 升
(东北大学 计算机科学与工程学院 国家级计算机实验教学示范中心, 沈阳 110819)

摘 要: 针对区块链中现有的密码学算法无法对抗量子计算机攻击的问题,本文提出了一种新的基于 DNA 密码和对称密码的一次一密加密算法。通过建立一次一密 DNA 密码本、利用 PCR(Polymerase Chain Reaction)技术实现密钥分配;根据对称密码学中的 DES(Data Encryption Standard)算法形成密文来完成加密过程,给出加密算法的实现算例、加密算法的应用过程和安全分析。由于 DNA 密码和对称密码的特性,该算法具有良好的抗量子性能,利用该算法对区块链中现有的算法加以改进,可以实现更高效和更安全的加密。

关键词: 区块链; 对称密码学; 一次一密; 加密算法
中图分类号: TP311 **文献标志码**: A **文章编号**: 2095-2163(2024)09-0150-05

One-time-pad cryptography algorithm based on DNA and symmetric codes

CHEN Hongjun, JING Qingwu, LI Danni, SHI Sheng
(National Computer Experimental Teaching Demonstration Center, School of Computer Science and Engineering, Northeastern University, Shenyang 110819, China)

Abstract: In response to the problem that existing cryptographic algorithms in blockchain cannot resist quantum computer attacks, this paper proposes a new one-time encryption algorithm based on DNA cryptography and symmetric cryptography. By establishing a one-time encrypted DNA codebook and utilizing PCR (Polymerase Chain Reaction) technology to achieve key distribution; According to the DES (Data Encryption Standard) algorithm in symmetric cryptography, ciphertext is formed to complete the encryption process. An implementation example of the encryption algorithm, its application process, and security analysis are provided. Due to the characteristics of DNA cryptography and symmetric cryptography, this algorithm has good quantum resistance. By using this algorithm to improve existing algorithms in blockchain, more efficient and secure encryption can be achieved.

Key words: blockchain; symmetric cryptography; one-time-pad cryptography; encryption algorithm

0 引 言

区块链作为一种新兴的分布式计算系统,拥有去中心化、信息不可篡改、集体维护等显著优点,区块链中有四项不可或缺的关键技术——分布式存储、共识机制、智能合约和密码学原理[1]。密码学是区块链底层的核心技术之一,在保证交易的完整性、真实性和私密性方面都发挥着重要作用。

在区块链中,密码学算法主要由哈希函数和公钥密码学组成[2]。哈希函数主要用于数据完整性、数据加密、共识计算的工作量证明、区块之间链接等;公钥密码学,又叫非对称密码学,包括公钥加密算法和数字签名算法,主要用于用户标识、操作权限校验、数字资产地址的生成、资产所有权的标识和数字资产的流转。

公钥密码系统的安全性随着 Shor 算法的提出和量子计算的发展受到了威胁[3]。以公钥密码学中的 RSA 算法为例,其数学原理概括来说是大数素因子分解的困难性,然而量子计算机不同于传统计算的基于 0 和 1 的二维计算,其可实现 N 维并行运算,且计算速度相当快,可以在极短时间内完成大数分解[4]。一旦这样的量子计算机开始被大规模使用,那么基于对固定数学问题的计算复杂性猜想的公钥密码体系就丧失了防护能力,主要通过公钥密

基金项目: 中央高校基本科研业务专项资金资助项目(N182410001)。
作者简介: 陈洪军(1971-),男,学士,高级实验师,主要研究方向:计算机网络技术,Email:chenhj@cc.neu.edu.cn; 景清武(1969-),男,硕士,高级实验师,主要研究方向:网络与信息安全。
收稿日期: 2024-04-28

码体系来实现加密的区块链的安全性也就受到了很大的威胁。

为了对抗量子计算机对于密码学界的攻击，2016 年美国国家标准与技术研究院正式发起了对公钥加密、密钥封装和数字签名这 3 类基本公钥密码算法相关的抗量子密码算法标准的征集工作[5]。经过多年研究，基于格的密码取得了丰硕成果，格密码可以抵抗量子计算机的攻击，并且几乎所有经典密码概念都可以在格密码中实现，但是格密码的密钥长度和通信代价比较大，涉及与安全强度相关的参数较多，选取特定安全强度的参数和评估给定参数的具体安全强度的困难性较高[6]。

对抗量子计算机攻击的另一种有效方式是利用 DNA 密码学建立密码系统。DNA 密码学是在 1994 年 Adleman[7] 提出了 DNA 计算之后才衍生出现的。2002 年，Shimanovsky 等[8] 将信息隐藏在 mRNA 序列中，给出了一种冗余密码子参与的加密算法；2004 年，Gehani [9] 用 DNA 加密技术设计了两种一次一密加密方法：映射替代法和异或法。2006 年，卢明欣等[10] 提出了利用 DNA 芯片加密技术实现的 DNASC（DNA Symmetric Cryptography）算法；2010 年，来学嘉等[11] 在 DNA 探针对称加密算法的基础上，提出非对称加密和签名技术；2013 年，Legoff 等[12] 将 DNA 微粒子技术与热缩片结合，构建了一种微粒子阵列加密模型；2020 年，Grass 等[13] 通过读取人类基因组和 DNA 合成技术设计了一种 DNA 加密存储系统，使得消息序列能够受到个性化密钥的保护。

DNA 密码学的独特优势，使其在密码学界具有重要的地位，对于 DNA 密码学的深入研究具有重要的理论意义。

本文提出了一种基于 DNA 密码学和对称密码学的一次一密加密算法，通过建立 DNA 密码本，利用 PCR 技术进行密钥分配，利用对称密码学算法进行加密运算，确保一次一密的实现。将该算法应用于区块链加密系统中，可以保证区块链技术良好的抗量子攻击的性能。

1　一次一密加密算法

1.1　一次一密的实现方式

1.1.1　一次一密的无条件安全性

一次一密的加密方式是指密钥数量和明文数量一样多，且是真随机的，每条密钥在使用一次后就被删除，下次使用时再采用一条新的临时随机密钥。一次一密加密方式中，密文给出关于明文的信息量：

$$I(M;C) = 0 \tag{1}$$

其中，M 为明文空间，C 为密文空间。

一次一密是无条件安全的[14]。

1.1.2　一次一密 DNA 密码本的制作

DNA 密码是指将 DNA 作为信息的载体，通过一系列生化反应来实现加密、认证和签名等密码学功能[15]。DNA 具有很小的体积，可以储存大量信息，一个碱基只有几个原子大小，1 克 DNA 可储存 4 500 亿 GB 字节数据。

实现一次一密加密方式的关键是拥有一个一次一密密码本，该密码本必须拥有真随机和高储存量的特点，而由于 DNA 密码的特性，利用 DNA 密码是一个很好的选择。可以利用人工合成 DNA 单链来建立一个一次一密 DNA 密码本，由于人工合成 DNA 单链具有真随机性，因此该 DNA 密码本产生的密钥是真随机的；另外，通过 PCR 技术可以有效实现密钥的分配[16]。因此，建立一个 DNA 密码本可以很好地解决一次一密加密方式中密钥的生成、管理和分配的问题。

1.2　一次一密加密算法的实现过程

1.2.1　明文编码与碱基编码

为了建立明文消息序列与 DNA 碱基序列之间的对应关系，将二进制编码作为中介来实现这一过程。DNA 由 A（腺嘌呤脱氧核苷酸）、T（胸腺嘧啶脱氧核苷酸）、C（胞嘧啶脱氧核苷酸）和 G（鸟嘌呤脱氧核苷酸）4 个脱氧核苷酸组成，通过 4 种碱基编码，其中 A、T 和 C、G 分别两两互补，定义如下编码方式：00 表示 A、01 表示 C、10 表示 G、11 表示 T，可以将碱基序列转化为二进制编码序列[17]；而明文消息序列转化为二进制编码序列，通过 ASCII 码表实现。

1.2.2　明文分组

通过密钥对明文进行加密的过程采用对称密码学中的 DES（Data Encryption Standard）算法[18]。由于在 DES 算法中第 8、16、24、32、40、48、65、64 位是奇偶校验位，因此这 8 位不能作为 DES 加密的有效数据位，为了确保一次一密的实现，本文将每 56 bit 长度的明文分为一组，不足 56 bit 的采用补零法。

设明文长度为 L_0，则明文的二进制编码长度 L_1 为：

$$L_1 = 8 \times L_0 \tag{2}$$

明文二进制编码被分为 n 组：

$$n = \lceil \frac{L_1}{56} \rceil \quad (3)$$

每组明文二进制编码的长度为 56。

1.2.3 密钥分配

人工合成若干条长度为 N_0 的 DNA 单链,其中前 20 个碱基为前引物序列,后 20 个碱基为后引物序列,其余碱基为密钥序列,保证任意两条 DNA 单链之间的引物序列和密钥序列都不相同,这就构成了一次一密的 DNA 密码本[19]。

加密时,利用 PCR 技术从 DNA 密码本中提取每对引物序列所对应的 DNA 单链。为了便于密钥分配,每条 DNA 单链的密钥序列恰好用于加密一组明文消息,因此每条 DNA 单链的长度 N_0:

$$N_0 = 20 + 64/2 + 20 = 72 \quad (4)$$

每条密钥序列的长度为 64 bit,即除去奇偶校验位后的实际密钥长度为 56 位。

1.2.4 密文生成

从 DNA 密码本中提取到密钥序列后,按照传统 DES 算法的加密流程进行加密。首先,通过初始置换 IP 对明文序列重新排列;其次,根据密钥序列生成 16 个长度为 48 的子密钥,进行 16 轮 feistel 结构迭代,包括扩展置换 E、S 盒迭代和 P 盒置换;最后,经过逆初始置换后,得到密文序列[20]。由于明文与密钥的长度不同,因此对每条明文需要进行后 8 位的补零操作,由于这 8 位是无效信息位,因此密文给出关于明文的信息量依旧是为 0 的。为了便于传输,将二进制的密钥序列转换为十六进制编码。

综上所述,基于 DNA 密码学和对称密码学的一次一密加密算法过程如图 1 所示。

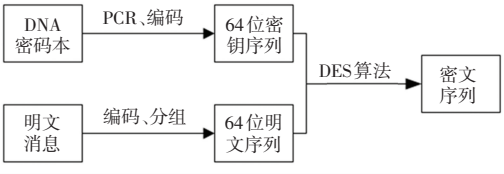

图 1 加密算法过程

Fig. 1 Encryption algorithm procedure

解密时,解密方拥有一个与加密方相同的 DNA 密码本,当解密方接受到密文消息后,首先将十六进制的密文消息转化为二进制序列;其次,根据接收到的 DNA 引物序列,利用 RCR 技术从 DNA 密码本中提取出密钥;最后,利用 DES 算法进行解密,去掉每条序列的后 8 位无效信息位,便可得到原始消息的编码序列,一次一密解密算法过程如图 2 所示。

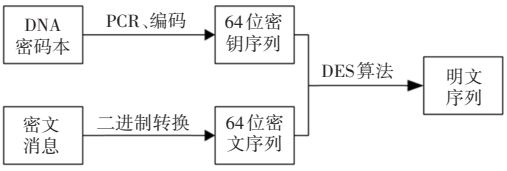

图 2 解密算法过程

Fig. 2 Decryption algorithm procedure

2 算例分析

以加密明文“Hello blockchain!”为例,利用 C++ 语言模拟加密流程。

首先,对明文进行编码与分组。表 1 给出了示例中各符号对应的二进制编码。

表 1 部分符号的 ASCII 码表

Table 1 ASCII encoding of partial symbols

符号	二进制编码
a	01100001
b	01100010
c	01100011
e	01100101
h	01101000
i	01101001
k	01101011
l	01101100
n	01101110
o	01101111
H	01001000
!	00100001
空格	001 0000

明文“Hello blockchain!”转化为二进制编码后为:

X = 01001000011001010110110001101100011011
11001000000110001001101100011011110110
00110110101101100011011010000110000101
10100101101110001000 01

每 64 位明文分为一组,明文分组后的结果为:

X_1 = 0100100001100101011011000110110001101 1
110010000001100010

X_2 = 0110110001101111011000110110101011011000
110110100001100001

X_3 = 0110100101101110000000000000000000000000
0000000000000100001

随机生成一次一密 DNA 密码本,每条 DNA 单链含有 72 个碱基,其中 32 个碱基为密钥序列,见表 2。

表 2　一次一密 DNA 密码本

Table 2　One-time-pad DNA code book

DNA 单链序列	引物序列
TGCATACCCATTTGTGGCGAGACTACTACGGCGCTAGGGCGGCTTATGTAGACTTGGTTCAGCCACACAACT	ACGTATGGGTAAACACCGCTGAACCAAGTCGGTGTGTTGA
CTCCCTTATGACAACAGAAGCGTCCCTTTGGCACTCTCTTAACTTTTACCTAGGGCGACCTACCGGGTTGTG	GAGGGAATACTGTTGTCTTCCCCGCTGGATGGCCCAACAC
TTTCTACCCCTAAGCGCGATAGGTCCGACAATAGAATCTGCCAAGGATAGGTCTTCTAGGATGACGAAATTC	AAAGATGGGGATTCGCGCTAGAAGATCCTACTGCTTTAAG
GTTGGGCGAACCAACTTCTTACCACGCTCGGAAGCCAGACCCGTTCAATCCGTGTAATGAACGTTATTGGCT	CAACCCGCTTGGTTGAAGAAACATTACTTGCAATAACCGA
CTACTTAATCTTATAATGCTTATACTGGGAGTTTCTTGTACTACGCCTAGCGTGTTAAATTATCCCTAACTT	GATGAATTAGAATATTACGAACAATTTAATAGGGATTGAA
GCTCAGAGTACCTTCTCCACGTCGCTTCCACCAGTCGTCTTCGAATCATGCGGTGAGATTGATTTGACCCAA	CGAGTCTCATGGAAGAGGTGCACTCTAACTAAACTGGGTT
GATCCCGTGAAGTCACGAGTACCCGCCTGACAGAAGACCTCGAGGCATCCTGTCTGTAGACCGATTGTAACA	CTAGGGCACTTCAGTGCTCAAGACATCTGGCTAACATTGT
CCGGGGTACAAGAGCTATCCTGTACGGATATATGTCAATATAGAGAACGTAACTCTTGTTACTATATTGTTT	GGCCCCATGTTCTCGATAGGGAGAACAATGATATAACAAA

接下来，提取 3 条 DNA 单链的密钥序列作为加密密钥，每条密钥只使用一次。将提取的 DNA 密钥序列按照规定的方式转化为二进制编码，转换后的密钥编码分别为：

Y_1 = 10000111000111000110100110011100101010011010011111001110110011001000

Y_2 = 01101101010111111111010010001110111011111000001111111110001011100

Y_3 = 00101011010110000100001100100000110111100101000010100011001010011

由于密钥编码为 64 位，而明文编码为 56 位，因此对明文编码的后 8 位补 0，将新的明文编码记为 X_1'，X_2'，X_3'，分别通过 3 条密钥 Y_1，Y_2，Y_3 按照 DES 算法加密 3 组明文 X_1'，X_2'，X_3'，得到的密文结果为：

Z_1 = 00110011101111010011011000010011111100111111010100000101010110000

Z_2 = 00010010100100011001001111110101001011010111010111111101011010

Z_3 = 10010011101110011100101101111111100110110001000000010001010001

将其转化为十六进制编码后为：

M_1 = 33bd3613e7f50a00

M_2 = 129193f52d75bf00

M_3 = 93b9cb7f36340800

最后，将密文 M_1，M_2，M_3 以及相应的引物序列传输给接收方，接收方按照解密算法解密后，再按照 ASCII 码将二进制编码还原为明文，得到明文消息。

3　一次一密加密算法的应用

在区块链中，每个区块的内部以及区块与区块之间的连接，都需要密码学来确保其安全性。依托上述加密过程，对区块链现有的加密算法进行改进，以提高加密算法的安全性。本文以实现数字签名为例来说明该算法的应用。由于该算法的加密密钥与解密密钥是相同的，因此本质上是一种对称加密的方式，两方实现通信的前提是拥有两个相同的一次一密 DNA 密码本。假设 Tom 是一个可信的并且拥有所有用户的一次一密 DNA 密码本，现在 Alice 要将明文消息发送给 Bob，就需要通过 Tom 这个媒介来实现。假设 Alice 拥有的 DNA 密码本为 D_1，Bob 拥有的 DNA 密码本为 D_2，传递消息的具体过程如下：

Step 1　Alice 从 D_1 中提取密钥，通过加密算法将明文消息转化为密文消息，传输给 Tom；

Step 2　Tom 从 D_1 中提取相同的密钥进行解密，得到解密后的消息；

Step 3　Tom 从 D_2 中提取密钥，把解密后的消息再次进行加密，并将发送者是 Alice 的声明也进行加密，然后传输给 Bob；

Step 4 Bob 从 D_2 中提取相同的密钥进行解密,读取 Alice 所发的消息和 Tom 的证明消息。

这个签名是可信的、不可伪造的和不能抵赖的,因为除了 Tom,只有 Alice 拥有密码本 D_1,而 Tom 是每个人都信任的,因此 Tom 的证书对 Alice 起着证明作用。

因为 Tom 是每个人都信任的,而密码本 D_1、D_2 是只有 Tom 和 Alice、Bob 才拥有的,因此整个消息的传递过程是可信的。

消息的传递在区块链交易中是经常发生的,只有保证了底层密码学技术的安全性,才能使整个区块链的交易是安全可靠的。

4 安全性分析

本文构建的一次一密密码系统中,结合了 DNA 密码学与对称密码学。利用 DNA 密码构造的一次一密的密码本,很好地解决了对称密码中密钥的储存、管理与分配的困难性问题;利用 DES 算法,加密速度快且加密效率高,也可以将 DES 算法改进为 3DES 算法来进一步提高加密算法的安全性。事实上,将 DNA 密码与 DES 加密算法结合使用,可以形成算法的优势互补,DES 算法保证了攻击者在破解每条密钥时尚且是不易的,而 DNA 密码本产生密钥的方式保证了单条密钥的破解并不能对其他组密钥产生影响,因此整个系统是安全的。

量子计算机破解对称密码体系可以归结为找到密钥,而量子计算机上的 Gover 算法可以在 $O(n^{0.5})$ 的时间内在 n 个加密字符中找到符合条件的加密字符,这就相当于将密钥长度减半,弥补的方法只要将密钥长度加倍即可,再加上利用生物学技术实现的一次一密的 DNA 密码本本身就具有不可破解性,使得本文提出的加密系统在面对量子计算机的攻击时具有了双重安全性。

5 结束语

本文提出的基于 DNA 密码学和对称密码学的一次一密加密算法,通过建立一次一密 DNA 密码本、利用 PCR 技术实现密钥分配、根据对称密码学中的 DES 算法形成密文来完成加密过程,具有良好的加密性能和安全性能。在区块链中应用该改进算法,能有效地抵抗量子计算机的攻击,保证量子计算机普及后区块链仍然是安全的。但该算法的实现存在费用较高等问题,后续研究将开展低成本的区块链加密算法。

参考文献

[1] 张亮,刘百祥,张如意,等. 区块链技术综述[J]. 计算机工程,2019,45(5):1-12.

[2] 王化群,吴涛. 区块链中的密码学技术[J]. 南京邮电大学学报(自然科学版),2017,37(6):61-67.

[3] 王潮,姚皓南,王宝楠,等. 量子计算密码攻击进展[J]. 计算机学报,2020,43(9):1691-1707.

[4] 胡云. RSA 算法研究与实现[D]. 北京:北京邮电大学, 2010.

[5] 王洋,沈诗羽,赵运磊,等. 基于模格的密钥封装方案的比较分析与优化[J]. 计算机研究与发展,2020,57(10):2086-2103.

[6] 倪亮,王念平,谷威力,等. 基于格的抗量子认证密钥协商协议研究综述[J]. 计算机科学,2020,47(9):293-303.

[7] ADLEMAN L M. Molecular computation of solutions to combinatorial problems[J]. Science, 1994, 266(5187): 1021-1024.

[8] SHIMANOVSKY B, FENG J, POTKONJAK M. Hiding data in DNA[C]// Proceedings of International Workshop on Information Hiding. Cham: Springer, 2002: 373-386.

[9] GEHANI A,LABEAN T H, REIF J H. DNA-based cryptography [C]//Proceedings of Lecture Notes in Computer Science. IEEE, 2950:34-50.

[10] 卢明欣,来学嘉,肖国镇,等. 基于 DNA 技术的对称加密方法[J]. 中国科学(E 辑:信息科学),2007(2):175-182.

[11] 来学嘉,卢明欣,秦磊,等. 基于 DNA 技术的非对称加密与签名方法[J]. 中国科学:信息科学,2010,40(2):240-248.

[12] LEGOFF G C, BLUM L J, MARQUETTE C A. Shrinking hydrogel - DNA spots generates 3D microdots arrays [J]. Macromolecular Bioscience, 2013, 13(2): 227-233.

[13] GRASS R N, HECKEL R, DESSIMOZ C, et al. Genomic encryption of digital data stored in synthetic DNA [J]. Angewandte Chemie, 2020, 132(22): 8554-8558.

[14] 李恕海. 量子密码若干问题研究[D]. 西安:西安电子科技大学,2008.

[15] 肖国镇,卢明欣,秦磊,等. 密码学的新领域——DNA 密码[J]. 科学通报,2006(10):1139-1144.

[16] 黄留玉,王恒樑. PCR 最新技术原理、方法及应用[M]. 北京:化学工业出版社,2005.

[17] 赵晓航. 基于 DNA 计算的加密方法研究[D]. 郑州:郑州轻工业学院,2013.

[18] 潘建生,孔苏鹏,程实. 实现 DES 加密算法安全性的分析与研究[J]. 网络空间安全,2020,11(4):104-107.

[19] 王子成,赵晓航,王宏,等. 基于 DNA 密码的一次一密加密算法[J]. 计算机工程与应用,2014,50(15):97-100.

[20] 管莹,敬茂华. DES 算法原理及实现[J]. 电脑编程技巧与维护,2009(4):5-7,13.