计算机科学
COMPUTER SCIENCE

# 基于I-SM4和SM2的混合加密算法

孙敏, 陕童, 续森炜

引用本文

孙敏, 陕童, 续森炜. 基于I-SM4和SM2的混合加密算法[J]. 计算机科学, 2023, 50(11A): 221100116-4.

SUN Min, SHAN Tong, XU Senwei. Hybrid Encryption Algorithm Based on I-SM4 and SM2[J]. Computer Science, 2023, 50(11A): 221100116-4.

---

计算机科学 COMPUTER SCIENCE
http://www.jsjkx.com
DOI:10.11896/jsjkx.221100116

# 基于 I-SM4 和 SM2 的混合加密算法

孙 敏 陕 童 续森炜
山西大学计算机与信息技术学院 太原 030006

**摘 要** 近年来,数据泄露事件频发,信息安全问题日益突出。由于单一的加密算法无法满足信息在传输过程中的安全需求,因此一般采用混合加密算法进行数据加密。现有的混合加密算法主要基于国外设计的加密算法,不符合网络空间安全自主可控的要求。针对这一问题,结合改进的 SM4 算法(I-SM4)与 SM2 算法,设计了一种新的混合加密算法。该算法改进了 SM4 加密算法的密钥扩展部分,采用线性同余序列代替原有的密钥扩展方式对轮密钥进行扩展,降低了轮密钥之间的相关性,提高了密钥的安全性。此外,采用将 I-SM4 与 SM2 相结合的方法,一方面可以加强对 I-SM4 密钥的管理,提高安全性;另一方面可以缩短单独使用 SM2 加密算法所需的时间。通过实验与分析证明,文中提出的混合加密算法能够有效提高网络传输过程中信息的保密性、完整性和不可否认性。

**关键词**:混合加密;SM2;SM3;SM4;线性同余
**中图法分类号** TP309

## Hybrid Encryption Algorithm Based on I-SM4 and SM2

SUN Min,SHAN Tong and XU Senwei
School of Computer & Information Technology,Shanxi University,Taiyuan 030006,China

**Abstract** In recent years,data leakage incidents have occurred frequently,and information security issues have become increasingly prominent. Since a single encryption algorithm cannot meet the security requirements of information in the transmission process,data encryption is generally performed through a hybrid encryption algorithm. The existing hybrid encryption algorithms are mainly based on encryption algorithms designed abroad,which do not meet the autonomous and controllable requirements of cyberspace security. Aiming at this problem,a new hybrid encryption algorithm is designed by combining the improved SM4 algorithm(I-SM4) and SM2 algorithm. It improves the key expansion part of the SM4 encryption algorithm,and uses the linear congruence sequence instead of the original key expansion method to expand the round key,which reduces the correlation between the round keys and improves the security of the key. In addition,the combination of I-SM4 and SM2 can strengthen the management of I-SM4 keys and improve security on the one hand. On the other hand,it can reduce the time required to use the SM2 encryption algorithm alone. Through experiments and analysis,it is proved that the hybrid encryption algorithm proposed in this paper can effectively improve the confidentiality,integrity and non-repudiation of information during network transmission.

**Keywords** Mixed encryption,SM2,SM3,SM4,Linear congruence

## 1 引言

当今社会高速信息化、网络化,导致数据数量呈爆发式增长,且海量的数据中包含大量用户的敏感信息[1]。数据在传输过程中面临着诸多安全风险,具有数据安全与隐私保护的需求[2]。混合加密技术是结合了多种加密体制和算法的新型加密技术,可以跟据实际的需求选择不同的加密体制进行加密方案的构建[3]。Bian 提出将 SM4[4] 中系统的固定参数由固定值改为动态选择生成,并采用 ECC[5] 完成 SM4 的密钥管理,以有效提高加密的安全性[6]。Li 提出基于国密算法 SM2[7],SM3[8],SM4 的高速混合加密系统硬件设计,通过对 SM2 和 SM4 算法的底层硬件结构进行优化,缩短了加密所需要的时间[9]。

本文提出了一种结合 SM4 与 SM2 的混合加密算法,对在网络传输中的信息进行加密传输。I-SM4 降低了轮密钥间的相关性,提高了信息的保密性。利用 SM2 对 I-SM4 密钥进行加密,一方面使密钥便于管理,降低 I-SM4 密钥泄露的危险,另一方面使信息具有不可否认性。同时,在数据传输过程中使用 SM3 算法进行数据的完整性检验,避免数据在传输过程中被恶意篡改。

## 2 相关工作

### 2.1 SM2 算法

SM2 是我国国家密码管理局发布的数字签名标准,用于保证身份的真实性、数据的完整性和行为的不可否认性。SM2 作为一种非对称加密算法,其安全性高,但是因为计算

基金项目:山西省基础研究计划项目(20210302123455);山西省基础研究计划项目(201701D121052)
This work was supported by the Shanxi Province Basic Research Program,China(20210302123455) and Shanxi Province Basic Research Program,China(201701D121052).
通信作者:孙敏(minsun@sxu.edu.cn)

量大导致其加解密的时间明显长于对称加密算法。

### 2.2 SM3 算法

SM3 算法适用于商用密码应用中的数字签名和验证、消息验证码的生成与验证及随机数的生成。SM3 算法还可以用于数据的完整性验证。

### 2.3 SM4 算法

SM4 是国家采用的一种分组密码标准，用于无线局域网产品。SM4 算法是一种对称加密算法，其加密效率高，节省资源及时间，但是 SM4 加解密使用同一个密钥导致其安全性低。在 SM4 的密钥扩展部分，其轮密钥具有相关性，导致破解出一轮轮密钥后密钥可以被破解。

在商用密码体系中，SM4 主要用于数据加密，其算法公开，分组长度与密钥长度均为 128 位，加密算法与密钥扩展算法都采用 32 轮非线性迭代结构，非线性变换中的基本运算单元为 $S$ 盒，$S$ 盒为固定的 8 位输入和 8 位输出。SM4 主要的运算以字($Z_2^{32}$)为单位，一次运算为一轮变换加解密结构一致，子密钥顺序相反。

### 2.4 线性同余发生器

线性同余发生器利用同余来产生随机数，其递推公式如下：

$$x_n \equiv (a \times x_{n-1} + c) (\mathrm{mod}\ m) \tag{1}$$

其中，$m$ 为模，$a$ 为乘数，$c$ 为增量，$x_0$ 为初始值，用线性同余发生器产生的随机数的特点是非常容易实现的，生成速度快[10]。

## 3 混合加密方案

### 3.1 SM4 算法密钥扩展

SM4 算法容易受到唯密文故障分析的影响，可以利用错误密文和密钥候选值推出子密钥，随后利用其扩展密钥间的相关性，可以依次推导出所有轮密钥的值，从而恢复主密钥[11]。SM4 密钥扩展过程如图 1 所示。

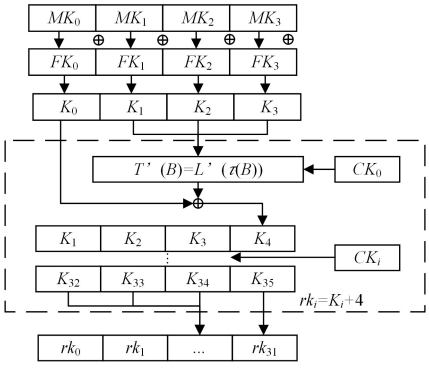


图 1 SM4 密钥扩展过程

Fig. 1 SM4 key expansion process

$Z_2^e$ 表示 $e$ 比特的二进制向量集，$\oplus$ 表示按位异或，$\tau$ 表示非线性变换，$L$ 表示线性变换，$T=L(\tau(\cdot))$ 表示 $\tau$ 和 $L$ 的组合，$MK \in (Z_2^{32})^4$ 表示 128 比特主密钥，$(X_i, X_{i+1}, X_{i+2}, X_{i+3}) \in (Z_2^{32})^4$ 表示加密算法第 $i+1$ 轮的输入 $i \in [0, 31]$，$CK_i \in (Z_2^8)^4$ 表示固定参数，$(K_i, K_{i+1}, K_{i+2}, K_{i+3}) \in (Z_2^{32})^4$ 表示密钥编排方案第 $i+1$ 轮的输入，$rk_i \in Z_2^{32}$ 表示第 $i+1$ 轮子密钥，$X \in (Z_2^{32})^4$ 表示明文 $X$，$Y \in (Z_2^{32})^4$ 表示密文 $Y$。

图 1 中，初始密钥和系统参数进行异或得到 $K_0, K_1 K_2 K_3$。每一轮的运算依赖于上一轮，依次下推可得到所需的任意轮子密钥。这种密钥扩展方法具有高效性和即时性的优点，但是如果知道其一轮的密钥就可以推出剩下的子密钥，从而获得原密钥。

### 3.2 I-SM4 算法密钥扩展

针对上述安全隐患，本文从抗攻击强度并兼顾程序执行时间方面考虑，利用线性同余发生器生成随机序列实现密钥扩展。

首先对初始密钥和系统参数进行异或。随后将异或后得到的值 $K_0, K_1 K_2 K_3$ 作为线性同余发生器的参数，传入公式后计算出随机值，随后传给子密钥 $rk_i$。

改进密钥扩展的具体流程如图 2 所示。这样改进后原始密钥不变，但子密钥是通过线性同余发生器产生出来的随机值。这使得攻击者无法通过某一轮密钥来推出全部子密钥。

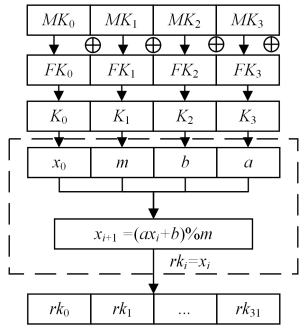


图 2 改进密钥扩展

Fig. 2 Improved key expansion

### 3.3 混合加密算法

混合加密算法使用 I-SM4 算法对隐私数据进行加密，使用 SM2 算法对其密钥进行加密，方便了对密钥的管理，也保证了信息的安全性。此外，使用 SM3 算法对明文进行哈希值的计算，并在解密完成后对明文进行完整性验证，防止数据在传输过程中遭到恶意篡改，保证传输过程中数据的完整性。本文的混合加密算法的步骤如下：

1)数据请求方产生非对称公私钥对 $P$ 和 $R$，并将公钥 $P$ 发送给数据拥有者。

2)数据拥有者产生密钥 $MK$，对明文 $M$ 进行哈希值计算得到 $Z$。使用密钥 $MK$ 并通过 I-SM4 加密算法加密数据得到密文 $Y$。使用 SM2 公钥 $P$ 加密密钥 $MK$ 得到密钥 $MK'$。

3)将密文 $Y$ 和密钥 $MK'$ 发送给数据请求方。

4)数据请求方使用私钥 $R$ 通过 *SM2* 解密算法解密 $MK'$ 得到密钥 $MK$。使用 $MK$ 通过 *I-SM4* 解密算法进行解密得到明文 $M'$。对 $M'$ 进行哈希值计算得到 $Z'$，并判断 $Z$ 是否等于 $Z'$，若等于则输出解密出来的明文 $M$；若不等于则拒绝接收。

## 4 实验及分析

实验环境为 Windows10 系统，处理器 CPU 为 Intel(R) Core(TM) i7-4790 CPU @ 3.60 GHz，3.60 GHz，运行内存 8 GB，编程语言为 Python 语言，工具为 PyCharm。

为对比本文提出的混合加密算法的有效性，使用 GmSS[1)] 测试算法性能。

1) https://github.com/gongxian-ding/gmssl-python

### 4.1　改进 SM4 算法扩散混淆性测试

扩散和混淆是 Shannon 提出的设计密码体制的两种基本方法，其目的是为了抵抗敌手对密码体制的统计分析。混淆的目的，是使密码的统计信息和加密密钥的值之间的关系尽可能复杂，使得敌手即使获得了关于密文的一些统计特性，也无法推出密钥。扩散的目的是让明文中的每一位影响密文中的许多位，这样可以隐蔽明文的统计特性。

首先测试其扩散性，加密数据为 128 比特。保证密钥不变，记录明文每变化一位，密文变化的位数（由于篇幅原因，这里只给出明文变化三位的测试结果）。当明文改变一位时，密文变化如图 3 所示，SM4 密文改变位数的范围是 65±8，I-SM4 密文改变位数的范围是 63±8。当明文改变两位时，密文变化如图 4 所示，SM4 密文改变位数的范围是 63±8，I-SM4 密文改变位数的范围是 61±7。当明文改变三位时，密文变化如图 5 所示，SM4 密文改变位数的范围是 63±7，I-SM4 密文改变位数的范围是 62±8。

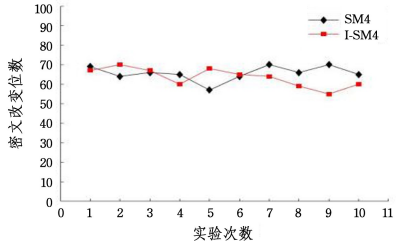


图 3　明文改变一位

Fig. 3　Change one bit in plaintext

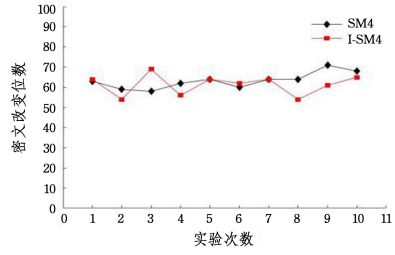


图 4　明文改变两位

Fig. 4　Change two bits in plaintext

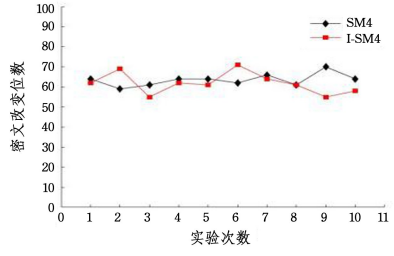


图 5　明文改变三位

Fig. 5　Change three bits in plaintext

其次测试其混淆性，加密密钥为 128 比特。保证加密数据不变，记录密钥每变化一位，密文变化的位数（由于篇幅原因，这里只给出密钥变化三位的测试结果）。当密钥改变一位时，密文变化如图 6 所示，SM4 密文改变位数的范围是 66±5，I-SM4 密文改变位数的范围是 65±7。当密钥改变两位时，密文变化如图 7 所示，SM4 密文改变位数的范围是 64±9，I-SM4 密文改变位数的范围是 66±5。当密钥改变三位时，密文变化如图 8 所示，SM4 密文改变位数的范围是 64±7，I-SM4 密文改变位数的范围是 67±8。本文一共测试了 10 位明文和密钥的变化情况，密文变化的位数都在 64 位左右，说明改进算法并没有影响原算法的扩散性和混淆性。

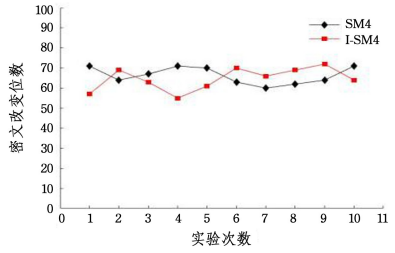


图 6　密钥改变一位

Fig. 6　Change one bit in key

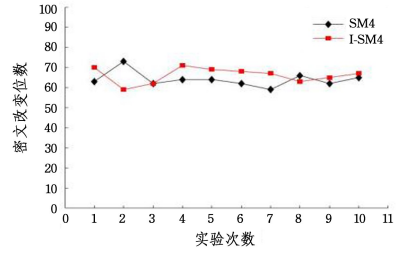


图 7　密钥改变两位

Fig. 7　Change two bits in key

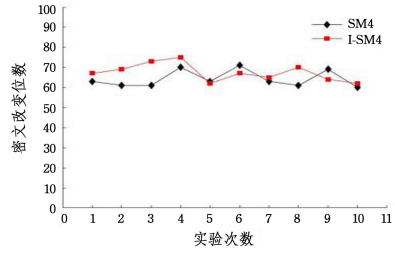


图 8　密钥改变三位

Fig. 8　Change three bits in key

### 4.2　I-SM4 加解密速率测试

本文使用 I-SM4 算法对 500 组以内，每组长度为 128 比特的数据进行加解密时间测试，另外对 SM4 算法也进行了测试。

由表 1 可知，相比原算法，I-SM4 算法并没有增加额外的时间开销。

表 1　加解密耗时对比

Table 1　Time-consuming comparison of encryption and decryption

（单位：ms）

| 数据/组 | SM4 | | I-SM4 | |
|---|---|---|---|---|
| | 加密 | 解密 | 加密 | 解密 |
| 100 | 13.334 | 12.239 | 12.187 | 11.320 |
| 200 | 25.247 | 24.377 | 24.470 | 23.499 |
| 300 | 38.378 | 34.493 | 38.230 | 33.726 |
| 400 | 50.138 | 45.882 | 49.946 | 44.083 |
| 500 | 61.269 | 57.550 | 59.255 | 56.256 |

### 4.3　混合加密算法性能测试

本文使用混合加密算法对 1000 组以内，长度为 128 比特的数据进行加解密时间测试，另外对 SM2 算法也进行了测试。

由表 2 可知，通过混合加密算法可以有效提高 SM2 算法的加解密速率。

表 2 加解密耗时对比

Table 2 Time-consuming comparison of encryption and decryption

(单位:ms)

| 数据/组 | SM2 | | SM2+I-SM4 | |
|---|---|---|---|---|
| | 加密 | 解密 | 加密 | 解密 |
| 10 | 62.03 | 53.53 | 38.00 | 14.99 |
| 100 | 490.41 | 475.32 | 48.17 | 25.31 |
| 500 | 2319.69 | 2383.90 | 95.24 | 70.24 |
| 1000 | 4785.24 | 4608.85 | 154.63 | 124.30 |

对 SM4 加密算法密钥扩展部分进行改进,采用线性同余序列代替原先密钥扩展方式对轮密钥进行扩展,降低轮密钥间的相关性,提高密钥安全性。使用 SM2 加密算法对 I-SM4 算法的密钥进行加密,一方面可以提高对 I-SM4 密钥的管理,提高安全性,另一方面可以缩短 SM2 加密算法所需的时间。

**结束语** 本文提出了一种基于 I-SM4 和 SM2 的混合加密算法,对网络传输过程中的信息进行加密。通过实验分析得出该算法能提高 I-SM4 密钥的安全性,并缩短单独使用 SM2 加密算法所需的时间。本文提出的混合加密算法提高了信息在传输过程中的保密性、完整性、不可否认性,并解决了目前缺乏基于国密标准的混合加密算法的问题,符合网络空间安全自主可控的需求。

## 参 考 文 献


[1] YANG Q, LIU Y, CHEN T, et al. Federated machine learning: Concept and applications[J]. ACM Transactions on Intelligent Systems and Technology(TIST), 2019, 10(2): 1-19.

[2] YAN Y X, MA M, JIANG H. An Efficient Privacy Protection Sifang Machine Learning Scheme Based on Secret Sharing [J]. Computer Research and Development, 2022, 59(10): 2338-2347.

[3] KANG H Y, DENG J. Enhanced hybrid encryption method for secure storage of medical data[J]. Journal of Beijing Institute of Technology, 2021, 41(10): 1058-1068.

[4] Cryptography Administration. SM4 Block Cipher Algorithm: GM/T 0002-2012 [S]. Beijing: China Standard Press, 2012.

[5] WANG J W, ZHANG S H, LI C. A Controllable Identity Management and Authentication Model of Agricultural Product Supply Chain Based on ECC-ZKP[J]. Computer Application Research, 2022, 39(10): 2916-2922, 2928.

[6] BIAN J X, LI Y J, WANG J H. Research on Hybrid Encryption Algorithm Based on SM4 and ECC[J]. Computer Applications and Software, 2016, 33(10): 303-306, 324.

[7] General Administration of Quality Supervision, Inspection and Quarantine of the People's Republic of China, Standardization Administration of China. Information Security Technology SM2 Elliptic Curve Public Key Cryptographic Algorithm Part 1: General Provisions: GB/T 32918.1-2016 [S]. Beijing: China Standards Press, 2017.

[8] Cryptography Administration. SM3 Cryptographic Hash Algorithm: GB/T 32905-2016 [S]. Beijing: China Standard Press, 2012.

[9] LI J L, MO Y N, SU T, et al. Hardware Design of High-speed Hybrid Encryption System Based on National Secret Algorithms SM2, SM3, SM4 [J]. Computer Application Research, 2022, 39(9): 2818-2825, 2831.

[10] HUANG X L, SHI H S, ZHANG C B, et al. Unpredictability of a Class of Combinatorial Linear Congruential Generators[J]. Journal of Tsinghua University(Natural Science Edition), 2016, 56(1): 22-27.

[11] LI W, WANG M L, GU D W, et al. Ciphertext-only Failure Analysis of SM4 Cryptographic Algorithm[J]. Journal of Computers, 2022, 45(8): 1814-1826.


**SUN Min**, born in 1966, master, professor. Her main research interests include computer network and information security.