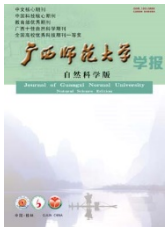

广西师范大学学报(自然科学版)
*Journal of Guangxi Normal University(Natural Science Edition)*
ISSN 1001-6600,CN 45-1067/N

# 《广西师范大学学报(自然科学版)》网络首发论文

题目：几类纠错码及其相关 McEliece 密码体制
作者：李志豪，吴严生，张钰芃
DOI：10.16088/j.issn.1001-6600.2022122001
收稿日期：2022-12-20
网络首发日期：2023-03-16
引用格式：李志豪，吴严生，张钰芃．几类纠错码及其相关 McEliece 密码体制[J/OL]．广西师范大学学报(自然科学版).
https://doi.org/10.16088/j.issn.1001-6600.2022122001

DOI: 10.16088/j.issn.1001-6600.2022122001 http://xuebao.gxnu.edu.cn

# 几类纠错码及其相关 McEliece 密码体制

李志豪[1]，吴严生[1*]，张钰芃[2]

（1.南京邮电大学 计算机学院、软件学院、网络空间安全学院，江苏 南京 210023；
2.南京航空航天大学 计算机科学与技术学院/人工智能学院/软件学院，江苏南京 210016）

**摘要：**随着量子计算机的迅速发展，抗量子攻击密码算法俨然成为密码学的研究热点。基于纠错码的 McEliece 公钥加密体制由 Robert McEliece 于 1978 年提出，是一种非对称的抗量子攻击加密算法。本文首先介绍几种常见的纠错码——Reed-Solomon 码、Goppa 码、Twisted Reed-Solomon 码的定义与基本性质，系统分析线性码的纠错译码算法；其次，详细阐述基于纠错码的 McEliece 公钥加密体制的加解密过程，并从 NP-Hard 问题对该加密体制进行安全性分析；最后，在前人研究成果的基础上，结合最近研究热点，提出几点思考和一些公开问题，为未来的研究提供参考。

**关键词：**纠错码；Goppa 码；Reed-Solomon 码；Twisted Reed-Solomon 码；McEliece 加密体制；纠错译码算法。
**中图分类号：**TN918.3 **文献标志码：**A

现代密码学的历史可以追溯到 20 世纪中叶。1949 年，香农发表《保密系统的通信理论》[1]，这一开创性工作标志着现代密码学的真正开始。1976 年前所有密码系统均属于对称密码学范畴。1976 年，Diffie 等[2]发表著名论文 *New directions in cryptography*，在论文中提出 1 个崭新的思想，他们认为加密算法本身可公开，甚至于加密过程使用的密钥也可以公开，这就是著名的公钥密码体制原创思想的肇始[3]。

在当今时代，小到日常支付、智能家居，大到卫星通信、国家安全，密码学都发挥着无可替代的作用。密码学技术是信息安全的核心技术。2015 年，网络空间安全在我国被正式确立为一级学科，随着网络空间安全问题成为全世界关注的热点，网络安全已成为国家安全的重要部分。2019 年，《中华人民共和国密码法》成功立法，以法律制度形式把信息安全的重要性升格为国家意志[4]。

量子计算机是以量子态直接进行信息处理的一种新型计算机，其在无序数据库搜索和大数分解等困难问题上显示出优于经典计算机的良好能力。随着科学技术的迅猛发展，量子计算机渐渐走进民众的视野。量子计算机的强大计算能力在诸多方面将给人们带来便利，但同时也会对基于数论等密码体制带来严重威胁。RSA、ElGamal、椭圆曲线等经典公钥加密体制均依赖于数学困难问题，例如大数分解与离散对数问题，经典计算机在多项式时间内难以有效解决上述 2 个问题，即对应密码体制的安全性在量子计算机时代受到严峻挑战，传统密码学将无法承担起保护信息安全的重任，网络空间安全变得岌岌可危。

为了应对挑战，许多学者开始构造抗量子密码体制(Post-Quantum Cryptography)，主要类型有 4 种[5]：基于格(Lattice-based)、基于多变量(Multivariate-based)、基于编码(Code-based)以及基于哈希函数(Hash-based)的公钥密码体制。

基于纠错码的密码体制最初由 McEliece 于 1978 年提出[6]，是一种非对称加密算法。本文从纠错码的角度出发，分析线性码的纠错译码算法，并简单介绍几种常见的纠错码——Reed-Solomon 码、Goppa 码、Twisted Reed-Solomon 码；接着详细阐述基于纠错码的 McEliece 公钥加密体制的加解密过程，并从 NP 类问题与相关攻击的角度对其进行安全性分析；最后，对已有研究成果作出总结，提出一些公开问题，为后续研究提供参考。

## 1 纠错码

### 1.1 纠错码

纠错码是信息论中的重要研究内容。给定信息序列$M$，纠错码[7-8]就是按照一定规则对该信息序

收稿日期：2022-12-20 修回日期：2023-03-06
基金项目：国家自然科学基金青年基金（12101326）；江苏省自然科学基金青年基金（BK20210575）；江苏省高等学校基础科学（自然科学）面上项目（21KJB110005）；南京市留学人员科技创新项目择优资助、南京邮电大学华礼人才计划、综合业务网理论及关键技术国家重点实验室开放课题（ISN23-22）
通信作者：吴严生(1989-)，男，安徽安庆人，南京邮电大学副教授，博士。E-mail: yanshengwu@njupt.edu.cn

列增加一些冗余，使没有规律性的信息序列$M$变换为具有某种规律性的数字序列$c$，称为码序列。纠错编码的基本思想是在信息接收端利用译码器将接收到的信息准确译码为原本的信息序列。

分组码是一种重要的纠错码，是把信息序列以$k$个码分组，通过编码器将每组的$k$位信息按一定规律产生$r$个校验位，输出长度为$n=k+r$的一个码字。每个码组的$r$个校验位仅与本组的信息位有关[8]。分组码用$[n,k]$表示，$n$表示码长，$k$表示信息位个数，分组码的码率为

$$R=\frac{k}{n}.$$

按照校验位与信息位之间的关系可把纠错码分为线性码[8-9]和非线性码。线性码的所有码字在线性运算下是封闭的，而非线性码则不封闭。把向量空间$F_q^n$的一个$F_q$上的线性子空间$C$叫作$q$元线性码[7,10]。换言之，$F_q^n$的一个非空子集合$C$叫作$q$元线性码，是指若$\boldsymbol{c},\hat{\boldsymbol{c}}\in C$，则对于任意$a,b\in F_q$，均有

$$a\boldsymbol{c}+b\hat{\boldsymbol{c}}\in C.$$

## 1.2 纠错译码算法

### 1.2.1 生成阵和校验阵[7]

对于线性码$C$，可利用线性代数工具，取一组$F_q$-基$\{\boldsymbol{v}_1,\boldsymbol{v}_2,\ldots,\boldsymbol{v}_k\}$，式中

$$\boldsymbol{v}_i=(a_{i1},a_{i2},\ldots,a_{in})\ (1\leqslant i\leqslant k).$$

$a_{ij}\in F_q(1\leqslant j\leqslant n,1\leqslant i\leqslant k)$. 则每个码字可惟一表示成

$$\boldsymbol{c}=b_1\boldsymbol{v}_1+b_2\boldsymbol{v}_2+\ldots+b_k\boldsymbol{v}_k=(b_1,b_2,\ldots,b_k)G.$$

式中$\boldsymbol{G}$为$F_q$上秩为$k$的$k$行$n$列矩阵，

$$G=(\boldsymbol{v}_1,\boldsymbol{v}_2,\ldots,\boldsymbol{v}_k)^T=\begin{bmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\vdots&\vdots&\ddots&\vdots\\a_{k1}&a_{k2}&\cdots&a_{kn}\end{bmatrix},$$

$\boldsymbol{G}$称为线性码$C$的一个生成阵。

把$K=q^k$个信息编为$F_q^k$中向量$(b_1,b_2,\ldots,b_k)G$，即是$F_q$-线性的单射

$$\Phi:\ F_q^k\to C\subseteq F_q^k\mid(b_1,b_2,\ldots,b_k)\mapsto(b_1,b_2,\ldots,b_k)G.$$

另一方面，$F_q^n$的一个$k$维向量子空间$C$必是某个齐次线性方程组

$$\begin{cases}b_{11}x_1+b_{12}x_2+\ldots+b_{1n}x_n=0\\b_{21}x_1+b_{22}x_2+\ldots+b_{2n}x_n=0\\\qquad\qquad\ldots\ldots\\b_{n-k,1}x_1+b_{n-k,2}x_2+\ldots+b_{n-k,n}x_n=0\end{cases}$$

的全部解，其中

$$H=(b_{ij})_{1\leqslant i\leqslant n-k,1\leqslant j\leqslant n}$$

是$F_q$上$n-k$行$n$列矩阵，并且秩为$n-k$。$H$叫作线性码$C$的一个校验阵。由定义可知，对每个$\boldsymbol{v}\in F_q^n$，

$$\boldsymbol{v}\in C\Longleftrightarrow\boldsymbol{v}H^T=0,$$

式中$0$是$n-k$维零向量。

### 1.2.2 线性码的纠错译码算法[7]

设$C$为参数$[n,k,d]$的$q$元线性码，$L=\left[\frac{d-1}{2}\right]$，$C$有校验阵$H=(\boldsymbol{u}_1,\boldsymbol{u}_2,\ldots,\boldsymbol{u}_n)$，其中$\boldsymbol{u}_i(1\leqslant i\leqslant n)$是$F_q$上长为$n-k$的列向量。若码字$\boldsymbol{c}\in C$在传送时发生的错位个数$\leqslant L$，即收到向量$\boldsymbol{y}=\boldsymbol{c}+\boldsymbol{\varepsilon}$，其中$w(\boldsymbol{\varepsilon})\leqslant L$，则用下列算法可纠错：

① 计算$F_q$上长为$n-k$的列向量，叫作$\boldsymbol{y}$的校验子

$$\boldsymbol{v}=H\boldsymbol{y}^T.$$

② 如果$\boldsymbol{v}=0$(零向量)，则$\boldsymbol{\varepsilon}=0$且$\boldsymbol{y}=\boldsymbol{c}$(无错)。

③ 如果$\boldsymbol{v}\neq 0$，则$\boldsymbol{v}$必可表示成$\boldsymbol{u}_1,\boldsymbol{u}_2,\ldots,\boldsymbol{u}_n$中不超过$L$个列向量的线性组合

$$\boldsymbol{v}=a_{i_1}\boldsymbol{u}_{i_1}+a_{i_2}\boldsymbol{u}_{i_2}+\ldots+a_{i_t}\boldsymbol{u}_{i_t}(1\leqslant i_1<i_2<\ldots<i_t\leqslant n),$$

式中：$1\leqslant t\leqslant L$；$a_{i_1},a_{i_2},\ldots,a_{i_t}$均为$F_q$中非零元素，此时$\boldsymbol{\varepsilon}=(\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_n)$, 其中

$$\varepsilon_{i_1}=a_{i_1},\varepsilon_{i_2}=a_{i_2},\ldots,\varepsilon_{i_t}=a_{i_t},$$

当$i\neq i_1,i_2,\ldots,i_t$时，$\varepsilon_i=0$，于是

$$\boldsymbol{c}=\boldsymbol{y}-\boldsymbol{\varepsilon}.$$

### 1.3 常见纠错码简介

#### 1.3.1 Reed-Solomon 码

经典代数编码中最广为使用的是 Reed-Solomon 码。Reed-Solomon 码自从 1960 年被 Reed 和 Solomon 提出后，因其优美的代数结构被广泛应用在通信和存储领域。

**定义 1** 设向量$\boldsymbol{\alpha}=(\alpha_0,\alpha_1,\ldots,\alpha_n)\in F_q^n$，选取固定值$k(1\leqslant k\leqslant n)$，则长度为$n$，维数为$k$的 Reed-Solomon 码(RS 码)可定义为

$$RS_{k,n}[\boldsymbol{\alpha}]:=\{ev_{\boldsymbol{\alpha}}(f):f\in F_q p[X],\ deg\ f\leqslant k-1\}\subseteq F_q^n,$$

其中，向量$\boldsymbol{\alpha}$的元素称为$RS_{k,n}[\boldsymbol{\alpha}]$的位置子(locators)。

RS 码是 MDS 码，即达到了 Singleton 界$(d\leqslant n-k+1)$，同时具有快速译码算法来纠错，其权重可达惟一解码半径$r=\left[\frac{n-k}{2}\right]$.

#### 1.3.2 Goppa 码

20 世纪 70 年代初，俄国学者 Goppa 在认真研究了 RS 码后，作为推广，系统性地构造出了一类有理分式码—Goppa 码。Goppa 码是一类特殊的代数几何码，它的某些子类能达到香农信息论中的信道编码定理[11]所给出的性能，且具有快速译码算法。

由于 Goppa 码不等价的码类数目非常大，McEliece 选用了二元 Goppa 码来构造一类公钥密码体制[6]，由此开始了使用纠错码构造密码体制及各种认证码的研究。Goppa 码也具有与 BCH 码相类似的好的译码算法，并且 Goppa 码族可以渐近地达到 G-V 界。因此，Goppa 码的研究在理论和实际应用中均具有极其重要的意义。

**定义 2**[12]　设$g(z)$是$F_{q^m}[z]$中$t$次首一多项式，$L=\{\gamma_0,\gamma_1,\ldots,\gamma_{n-1}\}$为$F_{q^m}$中$n$个不同元素组成的集合，且$g(\gamma_i)\neq 0(0\leqslant i\leqslant n-1)$，则码长为$n$的$q$元 Goppa 码可定义为

$$G(L,g)=\left\{\boldsymbol{C}=(c_0,c_1,\ldots,c_{n-1})\in F_q^n\left|\sum_{i=0}^{n-1}\frac{c_i}{z-\gamma_i}\equiv 0(mod\ g(z))\right.\right\}.$$

Goppa 码是$q$元线性码，接下来可以估计它的参数$k$和$d$[8]。上述求和式可写成

$$\left[\frac{1}{z-\gamma_0}\frac{1}{z-\gamma_1}\cdots\frac{1}{z-\gamma_{n-1}}\right][c_0,c_1,\ldots,c_{n-1}]^T=H_1\boldsymbol{C}^T\equiv 0(mod\ g(z)),$$

式中

$$H_1=\left[\frac{1}{z-\gamma_0}\frac{1}{z-\gamma_1}\cdots\frac{1}{z-\gamma_{n-1}}\right]$$

称为 Goppa 码的校验矩阵。把$H_1$化成其他形式, 为此把$H_1$中的每个元素进行化简。因为

$$\frac{g(z)}{z-\gamma_0}\equiv 0(mod\ g(z)),$$

所以

$$\frac{-1}{z-\gamma_0} \equiv \frac{g(z)-g(\gamma_i)}{z-\gamma_0} g^{-1}(\gamma_i)\,(mod\ g(z)).$$

把$H_1$中的每个元素都用上式代入，得到

$$H_2 = \left[\frac{g(z)-g(\gamma_0)}{z-\gamma_0} g^{-1}(\gamma_0),\ldots,\frac{g(z)-g(\gamma_{n-1})}{z-\gamma_{n-1}} g^{-1}(\gamma_{n-1})\right]。$$

若$\boldsymbol{C}(z) = c_{n-1}z^{n-1} + c_{n-2}z^{n-2} + \ldots + c_0$是由$g(z)$生成的 Goppa 码的一个码字，由$H_2\boldsymbol{C}^T = 0$可知

$$\sum_{i=0}^{n-1} c_i \frac{g(z)-g(\gamma_i)}{z-\gamma_0} g^{-1}(\gamma_i) = 0.$$

化简后，得到矩阵

$$H_2 = A\alpha B = \begin{bmatrix} g_t & 0 & 0 & \cdots & 0 \\ g_{t-1} & g_t & 0 & \cdots & 0 \\ g_{t-2} & g_{t-1} & g_t & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ g_1 & g_2 & g_3 & \cdots & g_t \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ \gamma_0 & \gamma_1 & \gamma_2 & \cdots & 0 \\ \gamma_0^2 & \gamma_1^2 & \gamma_2^2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \gamma_0^{t-1} & \gamma_1^{t-1} & \gamma_2^{t-1} & \cdots & \gamma_{n-1}^{t-1} \end{bmatrix} \begin{bmatrix} g^{-1}(\gamma_0) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & g^{-1}(\gamma_{n-1}) \end{bmatrix}.$$

该式中的矩阵$A$最后可化为主对角线以外元素均为 0 的矩阵, 因此它的存在与否不会影响码的纠错能力，所以矩阵$H_2$最后可化简为

$$H_3 = \alpha B = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ \gamma_0 & \gamma_1 & \gamma_2 & \cdots & 0 \\ \gamma_0^2 & \gamma_1^2 & \gamma_2^2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \gamma_0^{t-1} & \gamma_1^{t-1} & \gamma_2^{t-1} & \cdots & \gamma_{n-1}^{t-1} \end{bmatrix} \begin{bmatrix} g^{-1}(\gamma_0) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & g^{-1}(\gamma_{n-1}) \end{bmatrix}$$

$$= \begin{bmatrix} g^{-1}(\gamma_0) & g^{-1}(\gamma_1) & g^{-1}(\gamma_2) & \cdots & g^{-1}(\gamma_{n-1}) \\ \gamma_0 g^{-1}(\gamma_0) & \gamma_1 g^{-1}(\gamma_1) & \gamma_2 g^{-1}(\gamma_2) & \cdots & \gamma_{n-1} g^{-1}(\gamma_{n-1}) \\ \gamma_0^2 g^{-1}(\gamma_0) & \gamma_1^2 g^{-1}(\gamma_1) & \gamma_2^2 g^{-1}(\gamma_2) & \cdots & \gamma_{n-1}^2 g^{-1}(\gamma_{n-1}) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \gamma_0^{t-1} g^{-1}(\gamma_0) & \gamma_1^{t-1} g^{-1}(\gamma_1) & \gamma_2^{t-1} g^{-1}(\gamma_2) & \cdots & \gamma_{n-1}^{t-1} g^{-1}(\gamma_{n-1}) \end{bmatrix}.$$

式中$g^{-1}(\gamma_0) \in F_{q^m}$，对于一组固定的$F_q$-基，这些元素可以看成长度为$m$的列向量，从而$H_3$可以看成是$F_q$上$mt$行$n$列的矩阵，它在$F_q$上的秩$r(H_3) \leqslant mt$, 于是$k = dim_{F_q} G(L,g) \geqslant n - mt$. 进而，$H_3$中任意$t$列构成的 Vandermonde 行列式均不为零，所以$d \geqslant t+1$.

当$q=2$时，可以给出 Goppa 码的最小距离的更强结果。

**定理 1**[7] 对于**定义 2** 中的 Goppa 码$G(L,g)$，若$q=2$，并且$g(z)$在$F_2$的扩域中没有重根，则$d \geqslant 2t+1$.

**定理 2**[7] 设$g(x)$为$F_q[x]$中$t$次首 1 多项式$0 \leqslant t \leqslant n-1$, $L = \{\gamma_0, \gamma_1, \ldots, \gamma_{n-1}\}$为$F_q$中$n$个不同元素$(n \leqslant q)$，且$g(\gamma_i) \neq 0 (0 \leqslant i \leqslant n-1)$，则$q$元 Goppa 码

$$G(L,g) = \left\{C = (c_0, c_1, \ldots, c_{n-1}) \in F_q^n \middle| \sum_{i=0}^{n-1} \frac{c_i}{z-\gamma_i} \equiv 0 (mod\ g(x))\right\}$$

的参数为$[n,k,d]$，其中$k = n-t$，$d = t+1$，从而为 MDS 码。

1.3.3 Twisted Reed-Solomon 码

Twisted Reed-Solomon 码是 RS 码的一种推广。设$L\geqslant 1$，$n\geqslant k\geqslant 1$。向量$\boldsymbol{h}\in\{0,1,\ldots,k-1\}^l$取自成对不同的递增钩(pairwise distinct increasing hooks)，$\boldsymbol{t}\in\{0,1,\ldots,n-k\}^l$取自成对不同的扭曲(twists)，$\boldsymbol{\eta}\in(F_q\backslash\{0\})^l$，则由$[\boldsymbol{t},\boldsymbol{h},\boldsymbol{\eta}]$决定的扭曲多项式[13]为

$$P_{k,n}[\boldsymbol{t},\boldsymbol{h},\boldsymbol{\eta}]:=\left\{\sum_{i=0}^{k-1}f_iX^i+\sum_{j=1}^{L}\eta_jf_{h_j}X^{k-1+t_j}:\ f_i\in F_q\right\}\subseteq F_q[X].$$

**定义 3**[13]　若向量$\boldsymbol{\alpha}\in(\alpha_1,\alpha_2,\ldots,\alpha_n)\in F_q^n$，$k(1\leqslant k\leqslant n)$，$[\boldsymbol{t},\boldsymbol{h},\boldsymbol{\eta}]$的定义如上，则特征参数为$[\boldsymbol{t},\boldsymbol{h},\boldsymbol{\eta}]$，长度为$n$，维数为$k$，位置子为$\boldsymbol{\alpha}$的 Twisted Reed-Solomon 码(TRS codes)可定义为

$$TRS_{k,n}[\boldsymbol{\alpha},\boldsymbol{t},\boldsymbol{h},\boldsymbol{\eta}]:=\{ev_{\boldsymbol{\alpha}}(f):\ f\in P_{k,n}[\boldsymbol{t},\boldsymbol{h},\boldsymbol{\eta}]\}.$$

根据定义$TRS_{k,n}[\boldsymbol{\alpha},\boldsymbol{t},\boldsymbol{h},\boldsymbol{\eta}]$的生成矩阵为：

$$G_{\boldsymbol{\alpha},\boldsymbol{t},\boldsymbol{h},\boldsymbol{\eta}}=\begin{pmatrix}1\\ \boldsymbol{\alpha}^1\\ \vdots\\ \boldsymbol{\alpha}^{h_1-1}\\ \boldsymbol{\alpha}^{h1}+\eta_1\boldsymbol{\alpha}^{k-1+t_1}\\ \boldsymbol{\alpha}^{h_1+1}\\ \vdots\\ \boldsymbol{\alpha}^{h_L+1}\\ \boldsymbol{\alpha}^{h_L}+\eta_L\boldsymbol{\alpha}^{k-1+t_L}\\ \boldsymbol{\alpha}^{h_L+1}\\ \vdots\\ \boldsymbol{\alpha}^{k-1}\end{pmatrix},$$

式中

$$\boldsymbol{\alpha}^i:=(\alpha_1^i,\alpha_2^i,\ldots,\alpha_n^i)\,(1\leqslant i\leqslant k-1).$$

### 1.4　NP-Hard 类问题

若一个算法的复杂度小于某个多项式，则该算法是多项式时间算法。例如，二分查找的时间复杂度为$\log(n)\leqslant n$，故为多项式时间算法[14-15]。能够在多项式时间内用算法求解的问题为 P 问题，即计算机能够在有限时间内完成计算。整数排序就是 P 问题。

NP 问题即非确定性多项式时间问题，指不确定是否存在多项式时间内的求解算法，但可以在多项式时间内验证一个猜测解的正确性的问题。由定义易知，P 问题属于 NP 问题，但 NP 问题不一定属于 P 类问题。定义问题 P 可约化到另一问题 Q，当解决 Q 的方法也可用来解决 P 时。由约化的定义可知，一个问题约化为另一个问题，时间复杂度增加了，问题的应用范围也扩大了。

NPC 问题即这样一个 NP 问题，所有的 NP 问题都可以约化成它。NP-Hard 问题[14-15]即可由 NPC 问题约化到它。如果它本身也是 NP 问题的话，那么它就是 NPC 问题。简单来说 NPC 问题是 NP-Hard 问题的一个子集。

一般线性码的纠错译码算法可以用多项式时间求解，故一般线性码的译码算法是 NPC 问题[16]，属于 NP-Hard 问题，也是纠错码构造某些密码体制的理论基础。但对某些具有特殊代数结构的$[n,k,d]$线性分组码，例如 Goppa 码，存在快速译码算法，因此可以利用某些置换把 Goppa 码转化成一般线性码，再用其构造某些公钥密码体制。

## 2　McEliece 加密体制

### 2.1　McEliece 加密体制简介

McEliece 公钥加密体制的基本思想是先选取一个特殊的编码，其解码相对容易，然后将其伪装成一般的编码。原先的编码作为私钥，变换成的一般线性码作为公钥。McEliece 没有有效的攻击算法，但因为它的公钥太长，故很少用于实际。下面展示 McEliece 加密体制的详细算法过程[17]。

#### 2.1.1 密钥生成算法

① 确定$k,n,t$作为系统参数；

② 选取一个能纠$t$个错误的二元线性码的一个$k\times n$阶生成矩阵$G$，且知道该线性码的有效解码算法；

③ 随机选取一个$k\times k$阶二元非奇异矩阵$S$；

④ 随机选取一个$n\times n$阶二元置换矩阵$P$；

⑤ 计算$k\times n$阶矩阵$\widehat{G}=SGP$；

⑥ 输出公钥$\left(\widehat{G},t\right)$和私钥$(S,G,P)$.

2.1.2 加密算法

① 将信息$\boldsymbol{m}$分割为几个长度为$k$的二进制串；

② 选取一个随机长为$n$的错误向量$\boldsymbol{z}$，至多有$t$个1；

③ 计算$\boldsymbol{c}=\boldsymbol{m}\widehat{G}+\boldsymbol{z}$，

④ 输出密文$\boldsymbol{c}$.

2.1.3 解密算法

① 计算$\widehat{\boldsymbol{c}}=\boldsymbol{c}P^{-1}$，其中$P^{-1}$是$P$的逆；

② 用$G$生成的编码译码算法解密$\widehat{\boldsymbol{c}}$得$\widehat{\boldsymbol{m}}$；

③ 计算$\boldsymbol{m}=\widehat{\boldsymbol{m}}S^{-1}$.

流程图见图1~3.

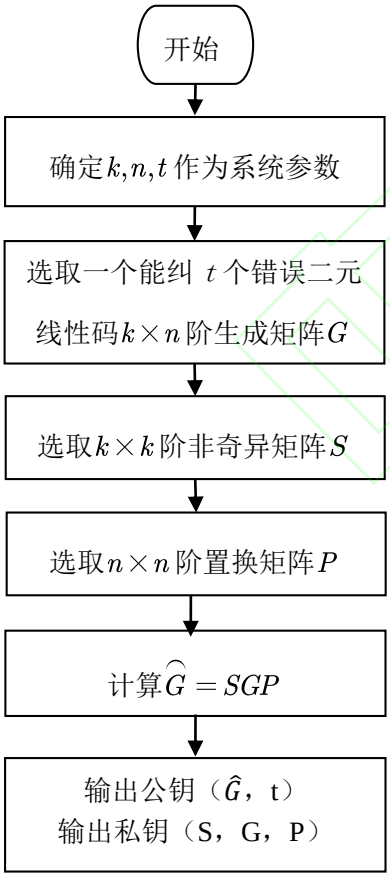


图1 密钥生成

Fig. 1 Key generation

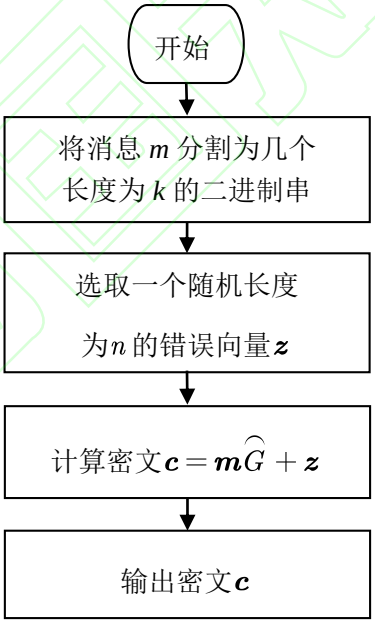


图2 加密

Fig. 2 Encryption

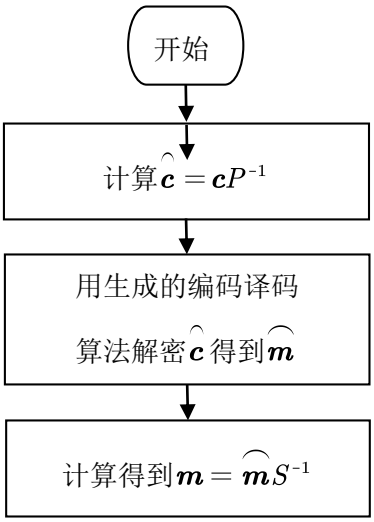


图3 解密

Fig. 3 Decryption

2.1.4 正确性验证

若所有算法都按步骤执行，则解密者可以正确恢复明文。因为

$$\widehat{\boldsymbol{c}}=\boldsymbol{c}P^{-1}=\left(\boldsymbol{m}\widehat{G}+\boldsymbol{z}\right)P^{-1}=(\boldsymbol{m}SGP+\boldsymbol{z})P^{-1}=(\boldsymbol{m}S)G+\boldsymbol{z}P^{-1},$$

又因为

$$W_H(\boldsymbol{z}P^{-1})=W_H(\boldsymbol{Z})\leqslant t,$$

所以通过译码，去掉纠错部分可得

$$\widehat{\boldsymbol{m}}=\boldsymbol{m}S.$$

2.1.5 示例

举例展现 McEliece 密钥生成和加解密算法过程，为方便起见，参数选取较短，设已知的编码生成矩阵为

$$G=\begin{pmatrix}1&0&0&0&1&1&0\\0&1&0&0&1&0&1\\0&0&1&0&0&1&1\\0&0&0&1&1&1&1\end{pmatrix}.$$

其编码和译码可以通过查表 1 完成。

**表 1　消息和码字对应表**

Tab. 1 Correspondence betweenmessages and codewords

| 消息 | 码字 |
|---|---|
| (0,0,0,0) | (0,0,0,0,0,0,0) |
| (0,0,0,1) | (0,0,0,1,1,1,1) |
| (0,0,1,0) | (0,0,1,0,0,1,1) |
| (0,0,1,1) | (0,0,1,1,1,0,0) |
| (0,1,0,0) | (0,1,0,0,1,0,1) |
| (0,1,0,1) | (0,1,0,1,0,1,0) |
| (0,1,1,0) | (0,1,1,0,1,1,0) |
| (0,1,1,1) | (0,1,1,1,0,1,0) |
| (1,0,0,0) | (1,0,0,0,1,1,0) |
| (1,0,0,1) | (1,0,0,1,0,0,1) |
| (1,0,1,0) | (1,0,1,0,1,0,1) |
| (1,0,1,1) | (1,0,1,1,0,1,0) |
| (1,1,0,0) | (1,1,0,0,0,1,1) |
| (1,1,0,1) | (1,1,0,1,1,0,0) |
| (1,1,1,0) | (1,1,1,0,0,0,0) |
| (1,1,1,1) | (1,1,1,1,1,1,1) |

1) 密钥生成：选取矩阵

$$S=\begin{pmatrix}1&0&0&1\\1&1&0&1\\0&1&0&1\\1&1&1&0\end{pmatrix},\ P=\begin{pmatrix}0&0&1&0&0&0&0\\1&0&0&0&0&0&0\\0&0&0&0&1&0&0\\0&0&0&0&0&1&0\\0&0&0&0&0&0&1\\0&1&0&0&0&0&0\\0&0&0&1&0&0&0\end{pmatrix}.$$

计算

$$\widehat{G}=SGP=\begin{pmatrix}0&0&1&1&0&1&0\\1&0&1&0&0&1&1\\1&1&0&0&0&1&0\\1&0&1&0&1&0&0\end{pmatrix},$$

则 $\widehat{G}$ 为公钥，$G,S,P$为私钥。

2)加密过程：设加密消息为$\boldsymbol{m}=(1,0,1,1)$，选取随机错误向量$\boldsymbol{z}=(0,1,0,0,0,0,0)$，计算密文

$$\boldsymbol{c}=\boldsymbol{m}\widehat{G}+\boldsymbol{z}=(0,0,0,1,1,0,0).$$

3)解密过程：根据密文计算

$$\widehat{\boldsymbol{c}}=\boldsymbol{c}P^{-1}=(0,0,1,0,0,0,1),$$

纠错得码字$\boldsymbol{x}=(0,0,1,0,0,1,1)$，解码得$\widehat{\boldsymbol{m}}=(0,0,1,0)$，最后得真实消息

$$\boldsymbol{m}=\widehat{\boldsymbol{m}}S^{-1}=(1,0,1,1).$$

### 2.2 McEliece 加密体制安全性分析

McEliece 公钥加密体制中使用的纠错码是既约二元 Goppa 码,该码的不等价数量随着参数的增加而快速增加。对于 McEliece 公钥加密体制的攻击可以归纳总结为下列问题:给定一个加密矩阵$\widehat{G}$，如果存在$k\times k$阶可逆矩阵$S$，$n\times n$阶置换矩阵$P$，密码分析者利用码$C$的快速译码算法和生成矩阵$G$，计算得到$\widehat{G}=SGP$。若如此，McEliece 公钥加密体制就可以被攻破，但从公钥信息推测出原始纠错码生成矩阵$G$属于一般线性码的解码问题，是 NP-Hard 的，因此这种情况发生的概率是很小的。

若有两个消息$\boldsymbol{m}_1$和$\boldsymbol{m}_2$被加密，假定密码分析者知道$\boldsymbol{m}_1+\boldsymbol{m}_2$，可以得到

$$\boldsymbol{c}_1=\boldsymbol{m}_1\widehat{G}+\boldsymbol{z}_1,$$

$$\boldsymbol{c}_2=\boldsymbol{m}_2\widehat{G}+\boldsymbol{z}_2\ \ (\boldsymbol{m}_1\neq\boldsymbol{m}_2),$$

因此

$$\boldsymbol{c}_1+\boldsymbol{c}_2=\boldsymbol{m}_1\widehat{G}+\boldsymbol{z}_1+\boldsymbol{m}_2\widehat{G}+\boldsymbol{z}_2=(\boldsymbol{m}_1+\boldsymbol{m}_2)\widehat{G}+\boldsymbol{z}_1+\boldsymbol{z}_2.$$

由于预先知道$\boldsymbol{m}_1+\boldsymbol{m}_2$，可以计算$(\boldsymbol{m}_1+\boldsymbol{m}_2)\widehat{G}$，因此该种攻击方式仅需要进行比较少的猜测就能攻击成功[18]。

### 2.3 研究现状与改进方向

McEliece 公钥密码体制具有加解密速度快、计算复杂度低的优点。对于码长为$n$的编码，McEliece 加密算法的时间复杂度为$O(n^2)$，而密钥长度为$n$的 RSA 等算法的时间复杂度为$O(n^3)$。但是存在公钥过长(约为$2^{19}$比特)、密钥存储空间太大、码率低等缺陷[17]。

目前 McEliece 加密体制的改进方向可分为 2 类[19]。

一是进一步改进 McEliece 原始方案。最早 McEliece 加密体制使用的是既约二元 Goppa 码。深入了解原始方案的细节和底层编码，选择性能更好的纠错码来替代原始方案的 Goppa 码，从而达到减小密钥长度之目的，使其能更好地投入到实际运用中。

二是进一步分析 McEliece 原始方案和改进版密码体制的安全性，综合多种攻击方法分析它们存在安全隐患的根本原因，从而更好地改善密码方案。

# 3 思考与总结

## 3.1 Goppa 码的编码性质

Goppa 码族是线性码的一个最有趣的子类, McEliece 密码体制中使用了这组纠错码[6]，该密码体制可以抵抗量子攻击。

### 3.1.1 计数

了解具有固定参数的不等价二元 Goppa 码的数量,有助于评价这种密码体制的安全性。设$n$为大于 3 的素数，Ryan[20]发现了长度为$2^n+1$的 4 次扩展不可约二元 Goppa 码数目的一个上界。Huang 等[21]给出了长度为$2^n+1$的 6 次扩展不可约二元 Goppa 码Gamma(L,g)数目的一个上界。Chen 等[22]得到了长度为$2^n+1$，次数为$r$的等价扩展不可约二元 Goppa 码数目的一个上界，其中$r$>3, 且满足$gcd(r,n)=1$, $gcd(r,2^n(2^{2n}-1))=1$。

**公开问题 1** 不同于$2^n+1$，对于更多的码长，确定扩展不可约二元 Goppa 码数目的更好上界甚至准确值。

### 3.1.2 自同构

McEliece 公钥加密方案的主要实际限制是其密钥的大小，克服这个问题的一个趋势是关注具有非平凡自同构群的 Goppa 码。

Thiong-Ly[23]给出了 2 类扩展非二进制 Goppa 码的自同构子群。Li 等[24]通过自同构群的非平凡子集来推广自同构诱导 Goppa 码的思想，构造出具有良好块结构的 Goppa 码。

**公开问题 2** 通过相关 Goppa 非平凡自同构群，考虑通过采用 QC 或 QD-Goppa 码来实现公密钥缩减。

### 3.1.3 代数译码算法

Berlekamp[25]证明 Goppa 码代数译码算法的 4 个最重要的性质]。Yasuo 等[26]证明了用欧几里得算法求解 Goppa 码解码的关键方程是可行的,其复杂度是 Burton 改进 BCH 码的 Berlekamp-Massey 算法的几倍。但该方法较为简单，易于求解任何 Goppa 多项式的关键方程。Patterson[27]提出 Goppa 码的代数译码算法，算法只比 Berlekamp-Massey 的 BCH 代码算法稍复杂。

McEliece 密码体制是使用 Goppa 编码来定义的,解码 Goppa 码是其解密的关键步骤,而 Patterson 算法是最为著名的 Goppa 码译码算法。Patterson 算法最有效的实现是使用预计算且在原始译码算法中，需要在有限域上计算多项式环商域的平方根，经过 Lim 等[28]的不懈努力，Patterson 译码算法得

到了改进，使其在保持最佳效率的前提下，去除了预计算部分，且在他们的修正下，计算只涉及有限域上多项式环的计算，不涉及商域，从而效率得到较大提高。

此外，Shen 等[29]在 Porter 等[30]提出的算法的基础上，给出了一种新的几何 Goppa 码的译码算法，其主要步骤是通过求解仿射环上的一个关键方程来完成。该算法适用于大量的 Goppa 码，为 Feng-Rao 译码算法提供了一种可行的替代方案，可将几何 Goppa 码解码到设计的极小距离。Agnes[31]开发了一种获取函数空间零增基的方法，从而实现了基于 Sudan 改进算法的译码方案。Liu 等[32]提出二次时间译码算法，其译码误差可达极小距离的一半。

**公开问题 3** 灵活运用 Goppa 码的译码算法，构造特殊的校验子，使得二元不可约 Goppa 码的覆盖半径有更优的下界。

### 3.2 Twisted Reed-Solomon 码的编码性质

尽管 RS 码是 MDS 码，但是文献[13]中已经证明根据定义构造的 TRS 码不一定是 MDS 码。近 MDS 码(NMDS codes)是通过弱化 MDS 码的定义而引入的，在秘密共享方案中有广泛应用。Sui 等[33]给出了 TRS 码是 MDS 或近 MDS 码的充要条件。

$F_q^m$ 上的 Goppa 码是代数纠错码的一个著名子类，如果$m=1$，则为 GRS 码。Huang 等[34]给出了 TGRS 码是 MDS 或近 MDS 码的充要条件，并尝试利用 2 类 TGRS 码构造了几类 MDS 或近 MDS 码。

线性对偶互补(LCD)码是一种线性码，可用于通信系统、电子消费、密码学等领域。因此 LCD MDS 编码的构造是编码理论中有趣的热点问题。Carlet 等[35]证明存在$q$元$[n,k]$ LCD MDS 码。Wu 等[36]构造出一类新的 Euclidean LCD MDS 码和 Hermitian LCD MDS 码，该码与 RS 码不是单项式等价(monomially equivalent)。

**公开问题 4** 构造更多新参数的性能良好的 Twisted Reed-Solomon 码。

### 3.3 基于上述编码的密码体制研究

#### 3.3.1 基于 Goppa 码的编码体制研究

McEliece 公钥加密方案的主要实际限制是其密钥的大小，克服此问题的一个趋势是关注具有非平凡自同构群的 Goppa 码。Li 等[24]提出一种新的通用结构，以此来减小自同构诱导 Goppa 码构造的 McEliece 密码体制的公钥大小。Faugère 等[37]在自同构群的作用下，添加属于同一轨道码字坐标，得到结果能用来增强 McEliece 体制上任何键恢复攻击，特别是代数攻击。

McEliece 方案的另一个缺点是它在语义上不安全，由 Barreto 和 Misoczki 提出的准二元 McEliece 变体解决了这 2 个问题。Heyse[38]给出了此公钥密码系统的实现方式，使其在语义上是安全的，并使用了较短的密钥。

**公开问题 5** 进一步研究 Goppa 码的编码性质，修改 Patterson 的解码算法，提高解码效率以及抵抗攻击，有效改进其所对应密码体制的性能。

#### 3.3.2 基于 Twisted Reed-Solomon 码的编码体制研究

TRS 码是一类包含大量最大距离可分码的码，最近被提议作为基于 McEliece 密钥体制的 Goppa 码的替代方案，以潜在减少密钥尺寸。

Lavauzelle 等[13]针对 McEliece 密码体系中使用的 TRS 变体，提出一种有效的密钥恢复攻击方法，该攻击使用了一种新方法，即基于公共代码中精选子域子字段的结构恢复，且攻击效果较好。可见，基于 TRS 码的密码体制还有待修正和完善。

**公开问题 6** 在深入研究 TRS 码编码性质的基础上，修改相应的解码算法，提高解码效率以及抵抗攻击，提出更为高效的密码体制。

## 4 结语

量子计算机时代的到来，传统密码学将无法承担起保护信息安全的重任。如何构造抗量子计算的密码体制是当今亟待解决的问题。理论上已经证明，基于纠错码的公钥密码体制是一种能够抵御量子攻击的新型密码体制，故此密码体制有非常好的研究前景和迫切的研究需求。本文对已有研究进行整理总结，尤其是 Goppa 码、Twisted Reed-Solomon 码的编码性质以及相关的公钥密码体制，并根据目前研究现状提出了几点思考和一些公开问题，为后续研究提供参考。

**参 考 文 献：**


[1] Shannon C E. Communication theory of secrecy systems[J]. The Bell System technical Journal, 1948, 28(4): 656-715.
[2] Diffie W, Hellman M E. New directions in cryptography[J]. Democratizing Cryptography,2022,365-390.
[3] 陈仲津, 张先俊. 有纠错能力的 PKC 密码体制—CS 码及其在电报通信中的应用[J]. 南京邮电学院学报, 1989, 9(1): 30-37.
[4] 吕佳璐. 十三届全国人大常委会第十四次会议表决通过密码法[EB/OL].(2019-10-26)[2022-12-16]. http://news.youth.cn/gn/201910/t2019102612103743.htm.
[5] 巫光福, 江林伟. 抗量子密码学综述[J]. 长江信息通信, 2021, 34(7): 55-60.
[6] Mceliece R J. A public-key cryptosystem based on algebraic coding theory[J]. DSN Progress Report, 1978, 114-116.
[7] 冯克勤. 纠错码的代数理论[M]. 北京: 清华大学出版社, 2005.
[8] 朱陆费. 基于纠错码的公钥密码体制研究[D]. 南京: 南京邮电大学, 2009.
[9] 杨淑娣, 岳勤. 一类线性码的完全重量分布[J]. 计算机工程与科学, 2019, 41(2): 281-285.
[10] 陈恭亮. 信息安全数学基础[M]. 北京: 清华大学出版社, 2004.
[11] 杨庚. 计算机通信与网络[M]. 北京: 清华大学出版社, 2009.
[12] López H H, Matthews G L. Multi- variate Goppa codes[J]. IEEE Transactions on Information Theory, 2023, 69(1): 126 – 137.
[13] Lavauzelle J, Renner J. Cryptanalysis of a system based on twisted Reed-Solomon codes[J]. Designs, Codes and Cryptography, 2020, 88(7): 1285-1300.
[14] Blateyang. NP 问题、NP 难问题(NPH)和 NP 完全问题(NPC)理解[EB/OL].(2017-10-28)[2022-12-16]. https://blog.csdn.net/Blateyang/article/details/78375885.
[15] Databatman. 算法中的 P 问题、NP 问题、NP 完全问题和 NP 难问题[EB/OL]. (2015-10-21)[2022-12-16]. https://blog.csdn.net/databatman/article/details/49304295.
[16] 曹东, 赵生妹, 宋耀良. 一种基于量子准循环 LDPC 码的 McEliece 公钥密码算法[J]. 南京邮电大学学报(自然科学版), 2011, 31(2): 64-68.
[17] 谷利泽, 郑世慧, 杨义先. 现代密码学教程[M]. 北京: 北京邮电大学出版社, 2009.
[18] Berson T A. Failure of the McEliece public-key cryptosystem under message-resend and related-message attack[J]. Lecture Notes in Computer Science, 1997, 1294: 213-220.
[19] 李喆, 韩益亮, 李鱼, 等. 基于编码的加密体制综述[J]. 国防科技大学学报, 2020, 42(4): 134-142.
[20] Ryan J. A. Counting extended irreducible binary quartic Goppa codes[J]. IEEE Trans- actions on Information Theory, 2015, 61(3): 1174-1178.
[21] Huang D T, Yue Q. Extended irreducible binary sextic Goppa codes[J]. IEEE Transactions on Information Theory, 2021, 68(1): 230-237.
[22] Chen B C, Zhang G H. The number of extended irreducible binary Goppa codes[J]. 2022, arXiv: 2204.02083.
[23] Thiong-Ly J A. Automorphisms of two families of extended non binary cyclic Goppa codes[J]. Lecture Notes in Computer Science, 1984, 229:112-121.
[24] Li Z, Xing C P, Yeo S L. Reducing the key size of McEliece cryptosystem from automorphism-induced Goppa codes via permutations[J]. Lecture Notes in Computer Science, 2019, 11443,:599-617.
[25] Berlekamp E R. Goppa codes[J]. IEEE Transactions on Information Theory, 1973, 19(5): 590-592.
[26] Sugiyama Y, Kasahara M, Hirasawa S. et al, A method for solving key equation for decoding Goppa codes[J]. Information and Control, 1975, 27(1): 87-99.
[27] Patterson N J. The algebraic decoding of Goppa codes[J]. IEEE Transactions on Information Theory,1975,21(2):203-207.
[28] Lim S, Lee H-S, Choi M. An efficient decoding of Goppa codes for the McEliece cryptosystem[J]. Fundamenta Informaticae, 2014, 133(4): 387-397.

[29] Shen B Z, Tseng K K. Decoding geometric Goppa codes up to designed minimum distance by solving a key equation in a ring[J]. IEEE Transactions on Information Theory, 1995, 41(6): 1694-1702.
[30] Porter S C, Shen B-Z, Pellikaan R. Decoding geometnc Goppa codes using an extra place[J]. IEEE Transactions on Information Theory, 1992, 38(11): 1663-1676.
[31] Heydtmann A E. Sudan-decoding generalized geometric Goppa codes[J]. Finite Fields and Their Applications, 2003, 9(3): 267- 285.
[32] Liu H D, Pircher S, Zeh A. et al, Decoding of (interleaved) generalized Goppa codes[C]. // 2021 IEEE International Symposium on Information Theory (ISIT), Melbourne, Australia , 2021: 664-669.
[33] Sui J Z, Zhu X M, Shi X Y. MDS and near-MDS codes via twisted Reed-Solomon codes[J]. Designs Codes and Cryptography, 2022, 90(8): 1937-1958.
[34] Huang D T, Yue Q, Niu Y F. MDS or NMDS LCD codes from twisted Reed-Solomon codes[J]. Cryptography and Communications, 2023, 15: 221-237.
[35] Carlet C, Mesnager S, Tang C. et al. Euclidean and Hermitian LCD MDS codes[J]. Designs Codes and Cryptography, 2018, 86(11): 2605-2618.
[36] Wu Y S, Hyun J. Y., Lee Y. New LCD MDS codes of non-Reed- Solomon type[J]. IEEE Transactions on Information Theory, 2021, 67(8): 5069-5078.
[37] Faugère J-C, Otmani A, Perret L. et al. Folding alternant and Goppa codes with non-trivial automorphism groups[J]. IEEE Transactions on Information Theory, 2014, 62(1): 184-198.
[38] Heyse S. Implementation of McEliece based on quasi-dyadic Goppa codes for embedded devices[J]. Lecture Notes in Computer Science, 2011, 7071: 143-162.

# Several Kinds of Error-Correcting Codes and Related McEliece Cryptosystem

LI Zhihao[1], WU Yansheng[1*], ZHANG Yupeng[2]

(1. School of Computer Science, Nanjing University of Posts and Telecommunications, Nanjing Jiangsu 210023, China; 2. Colloge of Computer Science and Technology/Colloge of Artificial Intelligence/Colloge of Software, Nanjing University of Aeronautics and Astronautics, Nanjing Jiangsu 210016, China)

**Abstract:** With the rapid development of quantum computers, cryptographic algorithms against quantum attacks have become a research hotspot in cryptography. McEliece encryption system based on error-correcting codes, proposed by Robert McEliece in 1978, is an asymmetric encryption algorithm against quantum attacks. In this paper, we first introduce the definitions and basic properties of several common error-correcting codes, such as Reed-Solomon codes, Goppa codes, and Twisted Reed-Solomon codes. Secondly, the encryption and decryption process of McEliece public-key encryption system based on error-correcting codes is described in detail, and the security of the encryption system is analyzed from the perspective of NP-Hard problems. Finally, on the basis of previous research results, combined with recent research hotspots, several considerations and open problems are proposed to point out the direction for future research.

**Keywords:** error-correcting codes；Goppa codes；Reed-Solomon codes；Twisted Reed-Solomon codes；McEliece encryption system；error-correcting decoding algorithm