一种基于QC-LDPC码的数字签名算法

杨雪菲 郑 东 任 方
(西安邮电大学通信与信息工程学院 西安 710121)
(西安邮电大学无线网络安全技术国家工程实验室 西安 710121)

摘 要 基于编码的公钥密码技术能够抵抗量子算法的攻击,针对经典的CFS签名方案密钥量大的缺陷,文中提出了一种基于QC-LDPC码的CFS签名方案。该方案基于QC-LDPC码改进了传统的CFS签名方案,签名过程中使用了QC-LDPC码的BP快速译码算法。分析表明,新方案在不降低安全性的同时,能够有效抵抗现有量子算法的攻击,减小了CFS签名方案的密钥存储空间,提高了方案的签名效率。

关键词 公钥密码,QC-LDPC码,CFS签名方案,BP译码算法
中图法分类号 TP309 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2019.06.024

Digital Signature Algorithm Based on QC-LDPC Code

YANG Xue-fei ZHENG Dong REN Fang
(School of Telecommunication and Information Engineering, Xi'an University of Posts and Telecommunications, Xi'an 710121, China)
(National Engineering Laboratory for Wireless Security, Xi'an University of Posts and Telecommunications, Xi'an 710121, China)

Abstract Code-based public key cryptography can resist the attack of quantum algorithms. Aiming at the large amount of key in classical CFS signature scheme, this paper proposed a kind of CFS signature scheme based on QC-LDPC codes. This scheme improves the traditional CFS signature scheme based on QC-LDPC codes. The BP fast decoding algorithm of QC-LDPC codes is used in the signature process. The analysis shows that the new scheme can reduce the key storage space of CFS, improve the efficiency of signature, and effectively resist the attack of quantum algorithm without reducing the security.

Keywords Public key cryptography, QC-LDPC codes, CFS signature scheme, BP decoding algorithm

目前,公钥密码已成为现代通信及信息安全领域不可或缺的重要技术之一,数字签名是公钥密码学的重要组成部分,用于鉴别用户身份和保证数据的完整性。现在广泛使用的公钥密码技术的安全性都基于数学困难问题,如RSA算法基于大整数分解问题,DSA算法、ElGamal签名算法基于离散对数困难问题。但是,Shor等[1]于1997年提出的量子攻击算法能够对上述困难问题进行有效的攻击,使得基于数学困难问题的公钥密码体制受到严重的威胁。2006年,第一届后量子密码会议(International Conference on Post-Quantum Cryptography, PQCrypto)总结了4种目前能够有效抵抗量子攻击的公钥密码体制[2],其中基于编码的公钥密码体制受到了众多研究者的广泛关注。

1978年,McEliece[3]首次提出了基于编码的公钥密码体制,即McEliece公钥密码体制。1986年,Niederreiter[4]也提出了一种基于编码的公钥密码体制,其与McEliece公钥密码体制的安全性等价,被称为Niederreiter公钥密码体制。2001年,Courtois等[5]利用校验子译码问题提出了第一个严格意义上的安全的基于编码的数字签名方案,被称为CFS签名方案。

1962年,Gallager[6]提出了LDPC码(Low Density Check-Parity Code),它是一类校验矩阵为稀疏矩阵的线性分组码。刚提出LDPC码时,Gallager虽然证明了它是具有渐进特性的好码,性能逼近Shannon极限,但是由于受到当时计算能力的限制,LDPC码一度被认为是一种不实用的码,在很长时间内被人们忽视。直到1996年,MacKay等[7]证明了LDPC码是一种好码,并推广了Gallager的概率迭代译码算法,论述了置信传播(Belief Propagation, BP)算法,极大地推动了LDPC码的发展。2007年,Baldi等[8]用QC-LDPC码代替Goppa码,构造了一个新的McEliece公钥密码体制。QC-LDPC码的准循环结构弥补了基于Goppa码的McEliece密码方案的密钥开销大的缺点,有效地降低了密钥存储空间。

目前,CFS算法有较高的安全性,但是密钥开销大,签名效率很低,签名成功的概率为$1/t!$(t为Goppa码的纠错能力)。本文首次提出了一种基于QC-LDPC码的CFS签名方

到稿日期:2018-04-08 返修日期:2018-07-21 本文受国家自然科学基金(61472472),陕西省自然科学基础研究计划项目(2015JQ6262,2017JQ6010)资助。

杨雪菲(1991-),女,硕士生,主要研究方向为信息安全;郑 东(1964-),男,博士,教授,主要研究方向为密码学、云存储安全;任 方(1981-),博士,副教授,主要研究方向为密码学与网络安全,E-mail:renfang_81@163.com(通信作者)。

案。与 Goppa 码相比，LDPC 码的校验矩阵为稀疏矩阵，可以大大节省密钥存储空间。QC-LDPC 码的校验矩阵具有准循环结构，这种结构使得 LDPC 码的译码复杂度比 Goppa 码的译码复杂度低。LDPC 码的 BP 译码算法在硬件中能够并行实现，极大地提高了译码速度，从而能够提高 CFS 算法的签名效率。

目前，基于 CFS 签名算法构造的盲签名[9]、环签名[10]、群签名[11]等都基于 Goppa 码，本文提出的新方案也能够使得这些算法得到改进。

1　基础知识

1.1　线性分组码基础

定义 1(线性分组码)　有限域 F_2 中的一个(n,k)线性分组码 $\boldsymbol{C}$ 是 n 维线性空间 F_2^n 的一个 k 维子空间，F_2^n中的向量被称为字，$\boldsymbol{C}$ 中的向量被称为码字，n 为码长，k 为维数。

定义 2(生成矩阵，generating matrix)　(n,k)线性分组码的生成矩阵是一个阶为 $k\times n$ 的矩阵 $\boldsymbol{G}$，生成矩阵 $\boldsymbol{G}$ 不是唯一的。

定义 3(校验矩阵，parity check matrix)　(n,k)线性分组码的校验矩阵是一个阶为$(n-k)\times n$ 的矩阵 $\boldsymbol{H}$，校验矩阵 $\boldsymbol{H}$ 不是唯一的，且校验矩阵 $\boldsymbol{H}$ 和生成矩阵 $\boldsymbol{G}$ 满足 $\boldsymbol{HG}^{\mathrm{T}}=0$。

长度为 n 的向量 $\boldsymbol{c}$ 是线性分组码 $\boldsymbol{C}$ 的一个码字的充要条件是 $\boldsymbol{Hc}^{\mathrm{T}}=0$。任意的字 $\boldsymbol{c}$ 的校验子为 $\boldsymbol{Hc}^{\mathrm{T}}$。

1.2　困难问题

1)校验子译码(Syndrome Decoding，SD)问题[12]

对于一个$(n-k)\times n$ 阶矩阵 $\boldsymbol{H}\in F_q^n$，向量 $\boldsymbol{s}\in F_q^{n-k}$ 及整数 $\omega>0$，是否存在字 $\boldsymbol{x}\in F_q^n$ 满足 $\boldsymbol{Hx}^{\mathrm{T}}=\boldsymbol{s}$ 且 $w(x)\leqslant\omega$。

2)*Goppa* 码的区分(*Goppa Code Distinguishing*，*GD*)问题

已知一个(n－k)×n 阶矩阵 $\boldsymbol{H}\in F_q^n$，判断 $\boldsymbol{H}$ 是一个(n,k)Goppa 码的校验矩阵还是一个随机(n,k)码的校验矩阵。

1.3　CFS 签名算法

CFS 签名算法是第一个可证明安全性的基于编码的数字签名方案，它是基于 Niederreiter 加密方案构造的。传统的基于 Goppa 码的 CFS 签名方案的安全性可以归结为校验子译码问题和 Goppa 码的区分问题。CFS 签名算法的具体过程如下：

1)密钥生成算法

给定一个(n,k,t)不可约 Goppa 码 $\boldsymbol{C}\in F_2$，$(n-k)\times n$ 阶校验矩阵 $\boldsymbol{H}_1$，Goppa 码的有效校验子译码算法 γ，$(n-k)\times(n-k)$阶的非奇异矩阵 $\boldsymbol{S}\in F_2$，$n\times n$ 阶置换矩阵 $\boldsymbol{P}\in F_2$，公开的安全 Hash 函数 h。

公钥为 $\boldsymbol{H}=\boldsymbol{S}\times\boldsymbol{H}_1\times\boldsymbol{P}$，私钥为$\langle\boldsymbol{S},\boldsymbol{H}_1,\boldsymbol{P},\gamma\rangle$。

2)签名算法

$\boldsymbol{z}=\gamma(S^{-1}h(h(m)\parallel i))$，$i\in N$，输出消息 m 的签名 $\sigma=[z|i]$。

3)验证算法

$\boldsymbol{s}_1=\boldsymbol{Hz}^{\mathrm{T}}$，$\boldsymbol{s}_2=h[h(m)|i]$；

如果 $\boldsymbol{s}_1=\boldsymbol{s}_2$ 相等，则签名 σ 有效；反之，则签名 σ 无效。

1.4　LDPC 码

F_2 上的(n,k)LDPC 码 $\boldsymbol{C}$ 是一种特殊的线性分组码，可以由$(n-k)\times n$ 阶奇偶校验矩阵 $\boldsymbol{H}$ 唯一定义。$\boldsymbol{H}$ 中非零元素的个数特别少，且尽量随机排列。也就是说 LDPC 码是一种校验矩阵密度非常低的分组码，它的核心思想是用一个稀疏的向量空间把信息分散到整个码字中。其中 n 表示码长，k 表示信息位长度，r 表示校验位长度，并且满足 $r=n-k$。

LDPC 码可以用 Tanner 图直观地表示。Tanner 图由 n 个变量节点和 r 个校验节点组成。图 1 给出了 $n=7$，$r=3$ 的 Tanner 图，其中 $v_j(j=0,1,\cdots,6)$表示变量节点，$c_i(i=0,1,2)$表示校验节点。

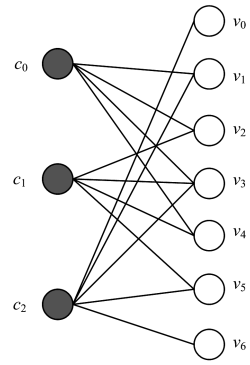

图 1　(7,4)LDPC 码的 Tanner 图

Fig. 1　Tanner graph of (7,4) LDPC codes

当且仅当校验矩阵 $\boldsymbol{H}$ 中的 $h_{ij}=1$ 时，v_j 和 c_i 之间存在一条边，这意味着位置 j 处的码字参与第 i 个奇偶校验方程。图 1 中 Tanner 图对应的奇偶校验方程 $\boldsymbol{H}$ 如下：

$$\boldsymbol{H}=\begin{bmatrix}0&1&1&1&1&0&0\\0&0&1&1&1&1&0\\1&1&0&1&0&1&1\end{bmatrix}$$

QC-LDPC 码[13]是一种准循环低密度奇偶校验码，其校验矩阵具有分块循环特性，设 F_2 上的二元(n,k) QC-LDPC 码的码长为 $n=n_0\times p$，维数为 $k=k_0\times p$，校验矩阵如下所示：

$$\boldsymbol{H}_{QC}=\begin{bmatrix}\boldsymbol{A}_{0,0}&\boldsymbol{A}_{0,1}&\cdots&\boldsymbol{A}_{0,n_0-1}\\\boldsymbol{A}_{1,0}&\boldsymbol{A}_{1,1}&\cdots&\boldsymbol{A}_{2,n_0-1}\\\vdots&\vdots&&\vdots\\\boldsymbol{A}_{k_0-1,0}&\boldsymbol{A}_{k_0-1,1}&\cdots&\boldsymbol{A}_{k_0-1,n_0-1}\end{bmatrix}\tag{1}$$

校验矩阵 $\boldsymbol{H}_{QC}$ 由 $k_0\times n_0$ 个子矩阵排列而成，$\boldsymbol{A}_{ij}$ $(1\leqslant i\leqslant m,1\leqslant j\leqslant n)$是 $p\times p$ 阶的全零矩阵或者单位循环矩阵。由于 QC-LDPC 码校验矩阵的分块循环特性，存储时只需要存储每一个非零循环子矩阵的位置和循环移位位数，这样明显缩小了签名过程中需要的密钥存储空间。

1.5　置信(Belief Propagation，BP)译码算法

LDPC 码的 BP 译码算法[14]也被称为和积译码算法(Sum-Product Algorithm，SPA)，是现在已知的性能最优的基于迭代译码的 LDPC 码译码算法。其主要思想是每次迭代过程中，利用收到的信息，在变量节点和校验节点之间不断进行信息的传递和迭代运算，从而进行译码。该算法是完全并行实现的，极大地提高了译码速度。迭代过程中，如果译码成功，则立刻结束迭代过程，而不是进行固定次数的迭代；并且 BP 算法的运算量不会随着码长的增加而快速增加，复杂度较低。BP 算法的译码流程如图 2 所示。

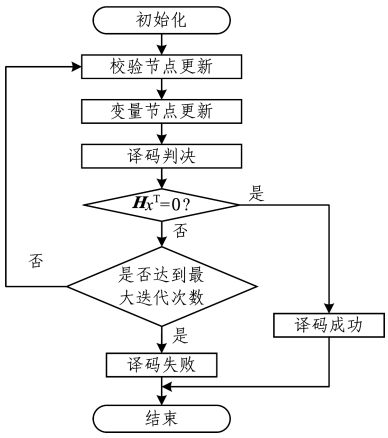

图 2 BP 算法的译码流程图

Fig. 2 Flow diagram of BP algorithm

BP 译码算法分为概率域 BP 算法和对数似然比(Log-Likelihood Ratio,LLR)BP 算法。不同之处在于:概率域 BP 算法的消息用概率形式表示;LLR-BP 算法[15]的消息用对数似然比表示。与概率域 BP 算法相比,LLR-BP 算法把大量的乘法运算转换成了加法运算,降低了译码复杂度和硬件资源的损耗。本文将使用 LLR-BP 算法。LLR-BP 算法的具体步骤如下:

首先对 LLR-BP 算法中用到的符号含义进行解释说明。

对于(n,k)LDPC 码,冗余长度为 r,Tanner 图的校验节点为$\{c_0,\cdots,c_{r-1}\}$,变量节点为$\{v_0,\cdots,v_{n-1}\}$。$A(k)$表示与校验节点 c_k 相连的所有变量节点集;$A(k)\backslash i$ 表示除变量节点 v_i 外,与 c_k 相连的所有变量节点集;$B(i)$表示与变量节点 v_i 相连的所有校验节点集;$B(i)\backslash k$ 表示除 c_k 外,与 v_i 相连的所有校验节点集。

$q_{i\to k}(x)$,$x\in\{0,1\}$表示根据 $B(i)\backslash k$,变量节点 v_i 传递给 c_k 的概率信息;$r_{k\to i}(x)$,$x\in(0,1)$表示根据 $A(k)\backslash i$,校验节点 c_k 传递给 v_i 的概率信息。

LLR-BP 算法中,消息从变量节点发送到校验节点的概率信息的对数形式如下:

$$\Gamma_{i\to k}(x_i)=\ln\left[\frac{q_{i\to k}(0)}{q_{i\to k}(1)}\right] \tag{2}$$

消息从校验节点发送到变量节点的概率信息的对数形式如下:

$$\Lambda_{k\to i}(x_i)=\ln\left[\frac{r_{k\to i}(0)}{r_{k\to i}(1)}\right] \tag{3}$$

1)初始化

在二元对称信道(BSC)中,$\forall i,k\in R$,Tanner 图的边连接节点为 c_k 和 v_i。

$$\Gamma_{i\to k}(x_i)=LLR(x_i) \tag{4}$$

$$\Lambda_{k\to i}(x_i)=0 \tag{5}$$

式(4)中的 $LLR(x_i)$表示在给定接收信号的情况下,变量节点的初始化概率信息,即:

$$LLR(x_i)=\ln\left[\frac{P(x_i=0\mid y_i=y)}{P(x_i=1\mid y_i=y)}\right] \tag{6}$$

由 BSC 信道概率分布及贝叶斯定理可得:

$$LLR(x_i\mid y_i=0)=\ln\left(\frac{1-p}{p}\right)=\ln\left(\frac{n-t}{t}\right) \tag{7}$$

$$LLR(x_i\mid y_i=1)=\ln\left(\frac{p}{1-p}\right)=\ln\left(\frac{t}{n-t}\right) \tag{8}$$

2)校验节点更新

通过式(9)计算从校验节点发送到变量节点的信息:

$$\Lambda_{k\to i}(x_i)=2\cdot\tanh^{-1}\left\{\prod_{j\in A(k)\backslash i}\tanh\left[\frac{1}{2}\Gamma_{j\to k}(x_i)\right]\right\} \tag{9}$$

3)变量节点更新

通过式(10)计算从变量节点发送到校验节点的信息:

$$\Gamma_{i\to k}(x_i)=LLR(x_i)+\sum_{j\in B(i)\backslash k}\Lambda_{j\to i}(x_i) \tag{10}$$

$$\Gamma_i(x_i)=LLR(x_i)+\sum_{j\in B(i)}\Lambda_{j\to i}(x_i) \tag{11}$$

4)尝试判决

利用式(11)得到的可靠值进行尝试判决,得到码字估计值$\hat{\boldsymbol{x}}=\{\hat{x}_0,\cdots,\hat{x}_n\}$的判决准则如下:

$$\hat{\boldsymbol{x}}_i=\begin{cases}0, & \text{if } \Gamma_i(x_i)\geqslant 0\\ 1, & \text{if } \Gamma_i(x_i)<0\end{cases} \tag{12}$$

然后,通过 $\boldsymbol{H}$ 计算估计值$\hat{\boldsymbol{x}}$的校验子,如果 $\boldsymbol{H}\hat{\boldsymbol{x}}^{\mathrm{T}}=0$,译码成功;否则,转向步骤 2),当达到最大迭代次数,并且不满足 $\boldsymbol{H}\hat{\boldsymbol{x}}^{\mathrm{T}}=0$,则说明译码失败,停止译码。

2 一种基于 QC-LDPC 码的 CFS 签名方案

2.1 经典 CFS 签名算法的分析

尽管基于$(n=2^m,k=n-mt)$Goppa 码的 CFS 算法的安全性较高,但由于在签名时需要进行多次译码尝试,因此签名速率特别低。

对于一个给定的(n,k)Goppa 码,满足 $n=2^m$,$k=n-mt$,其可译码的校验子个数为:

$$N_d=\sum_{i=1}^{t}\binom{n}{t}\approx\binom{n}{t}\approx\frac{n^t}{t!} \tag{13}$$

总的校验子个数为:

$$N_t=2^{n-k}=2^{mt}=n^t \tag{14}$$

随机获得校验子可译码的概率为:

$$p=\frac{N_d}{N_t}=\frac{1}{t!} \tag{15}$$

因此,其校验子译码成功的概率为 $1/t!$,即 CFS 签名算法签名成功的概率为 $1/t!$。当 $t=9$ 时,平均需要尝试 $9!=362\,880$ 次才能得到一个签名[16]。但是,文献[17]中已证明,该参数不再安全,重新建议参数可以取 $m=15$,$t=12$ 或者 $m=16$,$t=10$。随着攻击方法的不断更新,t 的取值会不断增大,这使得尝试译码的次数呈指数增长,签名速度越来越低,CFS 算法的签名效率也随之越来越低。

经典的 CFS 签名算法的公钥量为$(n-k)\times n$ bit,其需要的密钥存储空间较大,这使得该算法在实际中很少得到应用。LDPC 码的校验矩阵 $\boldsymbol{H}$ 的稀疏性以及 QC-LDPC 码的校验矩阵 $\boldsymbol{H}$ 的分块循环特性,可以大幅降低密钥存储空间,改善经典 CFS 算法的缺陷。因此本文提出了基于 QC-LDPC 码的 CFS 签名算法。

2.2 基于 QC-LDPC 码的 CFS 签名算法

经典的 CFS 签名算法利用的是 Goppa 码,本节提出一种基于 LDPC 码的 CFS 签名方案,即用 QC-LDPC 码替换 CFS

算法中的 Goppa 码，用 LDPC 码的 BP 译码算法替换 Goppa 码的快速译码算法。这里用可逆变换矩阵 $\boldsymbol{Q}$ 代替置换矩阵 $\boldsymbol{P}$[18]，以抵抗密度降低攻击。

由于经典的 CFS 算法中，消息经过哈希函数处理后得到的消息摘要 $\boldsymbol{s}$ 的长度为 $r(r<n)$，但是在签名过程中，LDPC 码的 BP 译码算法的输入序列长度为 n，因此本文做出了以下改进：先将消息摘要 $\boldsymbol{s}$ 转换为长度为 n 的序列，然后再进行尝试译码。具体过程见算法 1，算法 1 中的矩阵 $\boldsymbol{H}$ 即为 LDPC 码的校验矩阵 $\boldsymbol{H}$。

已知 $r\times n$ 阶矩阵 $\boldsymbol{H}$，r 维向量 $\boldsymbol{v}$，可以使 $\boldsymbol{H}\boldsymbol{v}^{\mathrm{T}}=\boldsymbol{s}$（$\boldsymbol{v}$ 不限重量）。如果 $\boldsymbol{v}$ 的重量限定为 $t(t\leqslant r)$，则属于 SD 问题。

算法 1

输入：长度为 r 的消息摘要 **s**，校验矩阵 **H**

输出：长度为 n 的序列 **v**

1. 利用行变换将矩阵 **H** 转换成行最简形 $\mathbf{H}'$，即存在可逆矩阵 **M**，使得 $\mathbf{M}\cdot\mathbf{H}=\mathbf{H}'$，可以得到 $\mathbf{H}=\mathbf{M}^{-1}\cdot\mathbf{H}'$；
2. 根据 $\mathbf{H}=\mathbf{M}^{-1}\cdot\mathbf{H}'$ 以及 $\mathbf{H}\cdot\mathbf{v}^{\mathrm{T}}=\mathbf{s}$，可以得到 $\mathbf{M}^{-1}\cdot\mathbf{H}'\cdot\mathbf{v}^{\mathrm{T}}=\mathbf{s}$，然后两边同时左乘可逆矩阵 **M**，即 $\mathbf{M}\cdot\mathbf{M}^{-1}\cdot\mathbf{H}'\cdot\mathbf{v}^{\mathrm{T}}=\mathbf{M}\cdot\mathbf{s}$，得到 $\mathbf{H}'\cdot\mathbf{v}^{\mathrm{T}}=\mathbf{M}\cdot\mathbf{s}$；
3. 因为 $\mathbf{H}'$，**M**，**s** 都已知，因此可以得到满足条件的 **v**。

算法 2 为基于 QC-LDPC 码的 CFS 签名算法的具体过程。

算法 2　基于 QC-LDPC 码的 CFS 签名算法

1. 初始化

设二元(n,k)线性码 **C** 是 F_2 上的 QC-LDPC 码，其校验矩阵 $\mathbf{H}_1$ 为 $(n-k)\times n$ 阶矩阵，纠错能力为 t。**S** 为随机选择的 $(n-k)\times(n-k)$ 阶可逆矩阵，**Q** 是随机选择的 $n\times n$ 阶的可逆变换矩阵，且 **Q** 为对角分块矩阵，行重和列重 w 均大于 1。计算 $\mathbf{H}=\mathbf{S}\times\mathbf{H}_1\times\mathbf{Q}$。α 为 LDPC 码对应的 BP 译码算法，β 表示算法 1，即将长度为 r 的消息转为长度为 n 的序列的算法。

选择公开的安全 Hash 函数 $h:\{0,1\}^{*}\rightarrow F_2^{n-k}$；**H** 为公钥，$\langle\mathbf{S},\mathbf{H}_1,\mathbf{Q},\alpha\rangle$ 为私钥。

2. 签名过程

设待签名消息为 m。

2.1. 签名者利用 Hash 函数 h 对 m 进行哈希运算，得到消息摘要 **s**：$\mathbf{s}=\mathbf{h}(\mathbf{m})$。

2.2. 签名者利用 Hash 函数 h 计算 $\mathbf{s}_i=h(\mathbf{s}|i)$，$i=0,1,2,\cdots$。

2.3. 利用 β 算法将 $\mathbf{s}_i$ 转换为长度为 n 的序列 $\mathbf{v}_i$。

2.4. 利用 α 译码算法对 $\mathbf{v}_i$ 进行尝试译码，找到使得 $\alpha(\mathbf{v}_i)$ 存在的最小的 i，记作 i_0，并用 $\mathbf{s}_{i_0}$ 表示 i_0 对应的 $\mathbf{s}_i$，用 **z** 表示 i_0 对应的译出的字，满足 $\mathbf{H}\mathbf{z}^{\mathrm{T}}=\mathbf{s}_{i_0}$，$w(\mathbf{z})=t$。

消息 m 的签名记作 $\sigma=[\mathbf{z}|i_0]$，用 (m,σ) 表示消息-签名对。

3. 验证算法

设验证者收到的消息-签名对为 (m,σ)。

3.1. 利用 **z** 和公钥 **H** 计算 $\mathbf{s}_1=\mathbf{H}\mathbf{z}^{\mathrm{T}}$。

3.2. 根据消息摘要 h(m) 和 i_0，计算 $\mathbf{s}_2=h[h(m)|i_0]$。

3.3. 如果 $\mathbf{s}_1=\mathbf{s}_2$ 相等，则签名 σ 有效；反之，则签名 σ 无效。

3　算法分析

3.1　安全性分析

3.1.1　理论安全性

该方案中，哈希函数 h 的单向性依赖于 SD 问题，消息 m 的哈希值 $\boldsymbol{s}_i=h(h(m))$ 是 LDPC 码的校验子。通过尝试译码，得到的可译码 s_{i_0} 与 $\boldsymbol{z}$ 的关系相当于校验子与错误向量的关系，通过公钥 $\boldsymbol{H}$ 和 $\boldsymbol{s}_{i_0}$ 直接求解方程 $\boldsymbol{H}\boldsymbol{z}^{\mathrm{T}}=\boldsymbol{s}_{i_0}$，其中 $w(\boldsymbol{z})\leqslant t$。由纠错码理论可知，这是一个 NPC 问题，能够抵抗现有的量子攻击，如 Shor 算法和 Grover 算法，因此本文提出的方案能够抵抗现有的量子攻击算法。

3.1.2　Stern 攻击

针对 LDPC 码的低密度特点，攻击者可以对该算法进行攻击。由于 LDPC 码是低重量码字，攻击者可以直接根据 Stern 算法[19]获取消息。根据文献[20]提出的 (n,k)LDPC 码的 Stern 算法可知，经过一次迭代，寻找码重为 w 的码字的概率为：

$$\pi_w=\frac{\binom{w}{p}\binom{n-w}{\frac{k}{2}-p}\binom{w-p}{p}\binom{n-w-\frac{k}{2}+p}{\frac{k}{2}-p}}{\binom{n}{\frac{k}{2}}\binom{n-\frac{k}{2}}{\frac{k}{2}}}\cdot\frac{\binom{n-k-w+2p}{l}}{\binom{n-k}{l}}\tag{16}$$

参考文献[19]选择参数 p 和 l，使性能达到最优，则寻找一个低重量码字所需的平均迭代次数为 π_w^{-1}。

寻找最小码距，使用 Stern 算法代替随机码字的选择。执行 Stern 算法，进行迭代，直到错误概率小于 ε(ε 为一个非常小的实数)，每次迭代之后得到的码字相互独立，因此，r 次迭代之后，未能发现重量小于 w 的码字的概率为 $(1-\pi_w)^r$。

每次迭代的位操作平均数约为：

$$N\approx\frac{\frac{(n-k)^3}{2}+k(n-k)^2+2pl\binom{\frac{k}{2}}{p}+2p(n-k)\binom{\frac{k}{2}}{p}^2}{2^l}\tag{17}$$

一般用工作因子 W 来衡量算法抵抗攻击的能力，普遍认为 $W\geqslant 2^{80}$ 时，方案是安全的。寻找一个低重量码字的工作因子为 $W=\pi_w^{-1}\cdot N$。当取经典值 $n=4\,096$，$k=2048$，$d=82$，$p=3$，$l=36$ 时，$W=2^{98.39}$，则该方案可以抵抗 Stern 攻击。

3.1.3　OTD 攻击

攻击者还可以利用校验矩阵的稀疏性对方案进行攻击。矩阵 $\boldsymbol{S}$ 和 $\boldsymbol{Q}$ 都由大小为 $p\times p$ 的循环块组成，并且都为稀疏矩阵，这样可以减少译码的复杂性。设它们的生成多项式分别为 $s_{i,j}(x)$ 和 $q_{i,j}(x)$，$\boldsymbol{Q}$ 为对角形式，则有 $q_{i,j}(x)=0(i\neq j)$，$\boldsymbol{Q}$ 的块对角形式可以表示为：

$$\boldsymbol{Q}=\begin{bmatrix}\boldsymbol{Q}_0 & & & \\ & \boldsymbol{Q}_1 & & \\ & & \ddots & \\ & & & \boldsymbol{Q}_{n_0-1}\end{bmatrix}\tag{18}$$

根据校验矩阵 H_1 的后 $n-k$ 列为单位阵的形式及 $\boldsymbol{H}=\boldsymbol{S}\times\boldsymbol{H}_1\times\boldsymbol{Q}$ 可以得到：

$$\boldsymbol{H}_{n-k}=\boldsymbol{S}^{-1}\cdot\begin{bmatrix}\boldsymbol{Q}_0^{-1} & & & \\ & \boldsymbol{Q}_1^{-1} & & \\ & & \ddots & \\ & & & \boldsymbol{Q}_{n_0-1}^{-1}\end{bmatrix} \tag{19}$$

$\boldsymbol{H}_{n-k}^{-1}$在$(i,j)$处的循环块为$\boldsymbol{Q}_i\boldsymbol{S}_{i,j}$,$\boldsymbol{S}_{i,j}$表示矩阵$\boldsymbol{S}$在$(i,j)$处的循环块,因此,$\boldsymbol{H}_{n-k}^{-1}$依然是循环矩阵,其多项式表达为$h_{i,j}(x)=q_i(x)\cdot s_{i,j}(x) \bmod (x^p-1)$。由于矩阵$\boldsymbol{S}$和$\boldsymbol{Q}$是稀疏矩阵,它们的行重量和列重量$w\ll n/n_0$。$h_{i,j}(x)$的最高次为$w^2$,攻击者可以枚举$h_{i,j}(x)$所有的多项式子集,最终得到密钥。

OTD 攻击算法的具体描述如下:

1)$\boldsymbol{R}_i$表示$\boldsymbol{H}_{n-k}^{-1}$的第$i$行:$\boldsymbol{R}_i=[\boldsymbol{Q}_i\boldsymbol{S}_{i,k}\mid\boldsymbol{Q}_i\boldsymbol{S}_{i,k+1}\mid\cdots\mid\boldsymbol{Q}_i\boldsymbol{S}_{i,n-1}]$;

2)线性码的校验矩阵为$\boldsymbol{H}_{OTD3}=(\boldsymbol{Q}_i\boldsymbol{S}_{i,k})^{-1}\cdot\boldsymbol{R}_i=[\boldsymbol{I}\mid\boldsymbol{S}_{i,k}^{-1}\boldsymbol{S}_{i,k+1}\mid\cdots\mid\boldsymbol{S}_{i,k}^{-1}\boldsymbol{S}_{i,n-1}]$;

3)将校验矩阵$\boldsymbol{H}_{OTD3}$变换成$\boldsymbol{H}'_{OTD3}=\boldsymbol{S}_{i,k}\cdot H_{OTD3}=[\boldsymbol{S}_{i,k}\mid\boldsymbol{S}_{i,k+1}\mid\cdots\mid\boldsymbol{S}_{i,n-1}]$;

$\boldsymbol{S}$为稀疏矩阵,说明存在低重量码字。$\boldsymbol{H}'_{OTD3}$的行重为$w\cdot(n-k)$,该值相对于码长非常小。可以使用 Stern 算法找到低重量码字,然后恢复出$[\boldsymbol{S}_{i,k}\mid\boldsymbol{S}_{i,k+1}\mid\cdots\mid\boldsymbol{S}_{i,n-1}]$;但是如果矩阵$\boldsymbol{S}$是稠密矩阵,则该攻击的工作量会变得非常大,攻击者将很难得到密钥信息,因此该方案是安全的。

3.2 密钥开销分析

公钥密码体制存在密钥存储量大和信息传输速率低的缺点,使得其很少得到实际应用。基于 Goppa 码的经典 CFS 签名算法的密钥开销为$(n-k)\times n$ bits。例如,Goppa 码的典型参数[18]取值为:码长n为 1024。当纠错能力$t=50$时,密钥开销约为512×10^3 bits;当$t=40$时,密钥开销约为410×10^3 bits。虽然经典的 CFS 算法的安全性较高,但是密钥存储空间较大,算法的效率会随之降低。

本文是基于 QC-LDPC 码的 CFS 签名方案,利用 LDPC 码校验矩阵的稀疏性,可以大大减少密钥的存储空间,只需存储$\boldsymbol{H}$矩阵中非零元素的个数以及每一个非零元素的位置。QC-LDPC 码校验矩阵的循环特性不仅能降低密钥存储量,而且能够使得码字的信息位变大,增强纠错能力[12],有效地提高信息传输速率。本文方案中的公钥为准循环矩阵,只需存储每个循环块的第一行,取 QC-LDPC 码的典型参数为$n_0=4$,$k_0=3$,$q=4096$,$n=n_0\times q=16384$,$k=k_0\times q=12288$,则该方案的公钥量$k_0\times n_0\times q=49152$ bits$=6144$ Bytes。因此,QC-LDPC 码可以很好地改善经典 CFS 签名算法的密钥开销,降低密钥存储空间。表 1 列出了基于不同纠错码的 CFS 签名的性能比较。

表 1 数据分析

Table 1 Data analysis

纠错码	公钥量/Bytes	传信率
Goppa 码(1024,524)	32750	0.57
Goppa 码(2048,1036)	131054	0.57
Goppa 码(16384,12288)	6291456	0.75
LDPC 码(16384,12282)	20481	0.75
QC-LDPC 码(16384,12288)	6144	0.75

由表 1 可以看出,使用 QC-LDPC 码可以大幅降低 CFS 算法的密钥量,并且由于 QC-LDPC 码的结构特点,码字的信息位较大,能有效提高传信率。

同样,分析可知,随着码长的变化,基于 Goppa 码的 CFS 签名方案的密钥开销呈指数增长,密钥开销非常大。与其相比,基于 LDPC 码的 CFS 签名方案的密钥开销增长的速度小很多,由于 QC-LDPC 码的准循环特性,其密钥开销更小。因此,QC-LDPC 码可以大大节省密钥开销,增加算法的签名效率。

结束语 经典的 CFS 算法属于基于纠错码的数字签名算法,由于具有较高的安全性,自提出以来一直受到广泛的关注及研究,但是该算法存在密钥开销大、随着 Goppa 码的纠错能力t的增加签名速度急速下降的缺点。因此本文提出了一种基于 QC-LDPC 码的 CFS 签名方案。该方案利用 QC-LDPC 码的校验矩阵的稀疏性,有效地改善了经典 CFS 签名算法的密钥空间大的缺点,使用 LDPC 码的 LLR-BP 译码算法,可以在不降低安全性的前提下,提高签名效率。该方案也属于基于编码的公钥密码方案,可以抵抗现有的量子算法攻击。

对于抵抗量子攻击的基于编码的数字签名体系的发展,还需要不断地进行研究。数字签名的种类很多,下一步还可以对其他数字签名方案进行改进,使它们能够基于编码问题实现,成功抵抗现有的量子攻击,从而不断地扩大数字签名方案的应用范围。

参 考 文 献

[1] SHOR P W. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer[J]. Siam Review,1997,41(2):1484-1509.

[2] BERNSTEIN D J. Introduction to post-quantum cryptography[J]. Post Quantum Cryptography,2009,85(1-2):1-14.

[3] MCELIECE R J. A Public-Key Cryptosystem Based on Algebraic Coding Theory[J]. Deep Space Network Progress Report,1978,42(44):114-116.

[4] NIEDERREITER H. Knapsack-type cry ptosystems and algebraic coding theory[J]. Problems Control Inform Theory,1986,15(2):159-166.

[5] COURTOIS N,FINIASZ M,SENDRIER N. How to Achieve a McEliece-Based Digital Signature Scheme[C]// Advances in Cryptology- ASIACRYPT 2001,International Conference on the Theory and Application of Cryptology and Information Security. Australia:DBLP,2006:157-174.

[6] GALLAGER R G. Low-density parity-check codes[J]. Information Theory Ire Transactions on,1960,8(1):21-28.

[7] MACKAY D J C,NEAL R M. Near Shannon limit performance of low density parity check codes[J]. Electronics Letters,1996,33(6):457-458.

[8] BALDI M,CHIARALUCE F,GARELLO R,et al. Quasi-Cyclic Low-Density Parity-Check Codes in the McEliece Cryptosystem[C]// IEEE International Conference on Communications. IEEE,2007:951-956.

[9] BLAZY O,GABORIT P,SCHREK J,et al. A code-based blind signature[C]// IEEE International Symposium on Information Theory. IEEE,2017:2718-2722.

[10] CHEN S,ZENG P,CHOO K K R,et al. Efficient Ring Signature and Group Signature Schemes Based on q-ary Identification Protocols[J]. Computer Journal,2018,61(4):545-560.

[11] LING S,NGUYEN K,ROUX-LANGLOIS A,et al. A lattice-based group signature scheme with verifier-local revocation [J]. Theoretical Computer Science,2018,730(19):1-20.

[12] REN F,ZHENG D,FAN J L. Survey of Digital Signature Technology based on Error Correcting Codes[J]. Chinese Journal of Network and Information Security,2016,2(11):1-10. (in Chinese)
任方，郑东，范九伦. 基于纠错码的数字签名技术综述[J]. 网络与信息安全学报，2016，2(11)：1-10.

[13] DRAGOI V,KALACHI H T. Cryptanalysis of a public key encryption scheme based on QC-LDPC and QC-MDPC codes[J]. IEEE Communications Letters,2017,PP(99):264-267.

[14] BALDI M. QC-LDPC Code-Based Cryptosystems [M]// QC-LDPC Code-Based Cryptography. Springer International Publishing,2014:91-117.

[15] ZHANG X R,LI J P,CAI C S. A Novel LLR-BP Algorithm for LDPC Codes Based on Taylor Series and Least Squares[J]. Applied Mechanics & Materials,2014,462-463:193-197.

[16] REN F,ZHENG D,WANG W J. An Efficient Code Based Digital Signature Algorithm[J]. IJ Network Security,2017,19(6):1072-1079.

[17] FINIASZ M,SENDRIER N. SECUrity Bounds for the Design of Code-Based Cryptosystems[C]// Advances in Cryptology- ASIACRYPT 2009, International Conference on the Theory and Application of Cryptology and Information Security. Tokyo: DBLP,2009:88-105.

[18] VAMBOL A,KHARCHENKO V,POTII O,et al. McEliece and Niederreiter Cryptosystems Analysis in the Context of Post-Quantum Network Security[C]// International Conference on Mathematics & Computers in Sciences & in Industry. IEEE Computer Society,2017:134-137.

[19] STERN J. A method for finding codewords of small weight [C]// International Colloquium on Coding Theory and Applications. New York:Springer-Verlag,1989:106-113.

[20] HIROTOMO M,MOHRI M,MORII M. A probabilistic computation method for the weight distribution of low-density parity-check codes[C]// International Symposium on Information Theory. IEEE,2005:2166-2170.