FatSeal：一种基于格的高效签名算法

谢天元[①②] 李昊宇[①②] 朱熠铭[①②] 潘彦斌*[①] 刘 珍[①②] 杨照民[①②]

[①](中国科学院数学与系统科学研究院数学机械化重点实验室 北京 100190)

[②](中国科学院大学数学科学学院 北京 100049)

摘 要：当前基于格设计的能够抵抗量子计算机攻击的签名方案是基于数论难题的传统签名方案的热门候选替代。通过Fiat-Shamir变换以及拒绝采样技术构造格签名是一种重要方法，共有5个格签名方案提交到美国国家标准与技术局的后量子算法项目中，基于Fiat-Shamir变换进行设计的有两个方案。其中Dilithium是基于模错误学习(MLWE)问题构造的Fiat Shamir签名，它的一个特性是在签名算法中使用了高效简洁的均匀采样。Dilithium签名方案构造在一般格上，为了获得更紧凑的公钥尺寸，Dilithium对公钥进行了压缩。另一方面，NTRU格上的密码方案比一般格上的密码方案在效率和参数尺寸上有更大的优势，该文给出了Dilithium签名在NTRU格上的一个高效变种方案，在继承Dilithium简洁设计的基础上，综合了NTRU和拒绝采样的技术优势而无需额外的压缩处理，进一步提升了基于格的Fiat-Shamir签名的效率。

关键词：数字签名；格；Fiat-Shamir签名；后量子；拒绝采样

中图分类号：TN918.1 **文献标识码**：A **文章编号**：1009-5896(2020)02-0333-08

DOI: 10.11999/JEIT190678

FatSeal: An Efficient Lattice-based Signature Algorithm

XIE Tianyuan[①②] LI Haoyu[①②] ZHU Yiming[①②] PAN Yanbin[①]
LIU Zhen[①②] YANG Zhaomin[①②]

[①](*KLMM, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing* 100190, *China*)

[②](*School of Mathematical Sciences, University of Chinese Academy of Sciences, Beijing* 100049, *China*)

Abstract: The lattice-based signature schemes are promising quantum-resistant replacements for classical signature schemes based on number theoretical hard problems. An important approach to construct lattice-based signature is utilizing the Fiat-Shamir transform and rejection sampling techniques. There are two Fiat-Shamir signatures among five lattice signature schemes submitted to the post-quantum project initiated by National Institute of Standards and Technology. One of them is called Dilithium, which is based on Module-Learning-With-Errors (MLWE) problem, it features on its simple design in the signing algorithm by using uniform sampling. The Dilithium is built on the generic lattices, to make the size of public key more compact, Dilithium adopts compression technique. On the other hand, schemes using NTRU lattices outperform schemes using generic lattices in efficiency and parameter sizes. This paper devotes to designing an efficient NTRU variant of Dilithium, by combining the advantage of NTRU and uniform rejection sampling, this scheme enjoys a concise structure and gains performance improvement over other lattice-based Fiat-Shamir signature without using extra compression techniques.

Key words: Digital signature; Lattice; Fiat-Shamir transform; Post-quantum; Rejection sampling

1 引言

数字签名是公钥密码体制的基础原件之一，其在通讯中有着非常广泛的应用。由于整数分解和离散对数问题被证明在量子计算机下可以高效求解，因此基于上述困难问题的传统公钥密码体制在量子计算机下是不安全的，构造能够抵抗量子攻击的密码体制即后量子密码体制是公钥密码一个非常重要的研究方向。

基于格的密码体制是后量子密码体制的一类重要候选方案，就格签名而言主要有两条发展路线：

收稿日期：2019-09-04；改回日期：2019-12-11；网络出版：2019-12-19
*通信作者： 潘彦斌 panyanbin@amss.ac.cn
基金项目：国家自然科学基金(61572490)
Foundation Item: The National Natural Science Foundation of China (61572490)

一类格签名根据Hash-and-Sign范式进行构造，另一类签名通过Fiat-Shamir变换及拒绝采样技术进行构造。

典型的Hash-and-Sign签名，公钥是一个单向陷门函数$f(x)$，私钥是利用陷门可以高效计算的逆函数$f^{-1}(x)$。签名一个消息m，首先将消息m哈希到函数f的值域上，即$y=H(m)$，然后输出签名$\sigma=f^{-1}(y)$。验签的时候需要检查条件$f(\sigma)=H(m)$是否成立。最早利用Hash-and-Sign范式构造的签名是(Goldreich-Goldwasser-Halevi, GGH)方案[1]，公钥选用格的一组“坏基”$\boldsymbol{H}$，私钥选取“好基”$\boldsymbol{B}$，签名的时候将消息哈希到实空间上的一点$\boldsymbol{v}$，利用好基$\boldsymbol{B}$通过Babai算法[2]找到离$\boldsymbol{v}$最近的格向量，验签的时候利用坏基$\boldsymbol{H}$验证签名是否是一个格点，当格点离$\boldsymbol{v}$足够近，则判定签名合法。GGH签名的安全性建立在敌手无法由坏基多项式时间内恢复出好基以及敌手无法利用坏基高效求解最近向量问题(Closest Vector Problem, CVP)。将GGH所采用的一般格限制到NTRU格上衍生出了NTRUSign方案[3]，这种签名方案有紧凑的密钥和签名长度。然而GGH型的签名易于遭受基本区域学习(learning a parallelepiped)攻击[4]，其缺陷是签名会泄露私钥的统计信息，通过收集足够多的消息签名对可以勾勒出好基的形状。第1个具有可证安全的Hash-and-Sign签名是2008年提出的(Gentry-Peikert-Vaikuntanathan, GPV)签名方案[5]，相较于GGH类签名直接在格中找离消息的哈希值$H(\boldsymbol{m})$最近的格点，GPV签名是在垂直格$\Lambda^{\perp}$中找离$\boldsymbol{c}_0$近的格点$\boldsymbol{s}$，其中$\boldsymbol{c}_0$满足$\boldsymbol{A}\boldsymbol{c}_0=H(\boldsymbol{m})$，这个格点服从中心在$\boldsymbol{c}_0$的离散高斯分布$\mathcal{D}_{\boldsymbol{c}_0+\Lambda^{\perp},\sigma}$，基本区域学习攻击不再能得到关于私钥的有效信息。GPV签名方案虽然在安全性上有了保障，但陷门采样过程效率较低且签名尺寸较大(达到了10^6字节级别)。提交到美国国家标准与技术局(National Institute of Standards and Technology, NIST)后量子算法竞赛的基于格的NTRU上的快速傅里叶紧签名(FAst-fourier Lattice-based Compact signatures Over NTRU, FALCON)算法[6]在陷门采样上做了改进，通过将GPV的一般格改为NTRU格得到了紧凑的公钥与签名尺寸。

Lyubashevsky[7]在2009年以ring-SIS为底层困难问题给出了第1个Fiat Shamir格签名。这类签名从一个身份鉴别方案出发进行Fiat Shamir变换，将身份鉴别方案中的脚本(承诺commitment，挑战challenge，回应response)应用到签名的生成过程中，将身份鉴别方案中验证者最后一轮对回应的检查应用到签名方案的验签算法中，由此得到一个非交互的方案。身份鉴别方案中的回应作为签名的一个部分，其分布应独立于私钥才能保证方案的安全性。Lyubashevsky[8]在2012年提出的Fiat Shamir签名方案，拒绝采样前的回应服从平移的离散高斯分布，平移量与私钥相关，他使用拒绝采样技术使得输出分布服从中心在原点的离散高斯分布，从而保证了签名方案是零知识的。双峰格签名方案(Bimodal LattIce Signature Scheme, BLISS)[9]对回应的分布进行了改进，将平移的离散高斯分布变成了双峰高斯分布，这样处理能够减少签名算法中的采样迭代次数并且使得输出的签名尺寸更加紧凑，BLISS方案用NTRU格进行了实例化。

提交到NIST后量子算法竞赛的Dilithium算法[10]也采用了Fiat-Shamir变换及拒绝采样技术，但避免了使用离散高斯采样，仅需均匀采样从而使签名方案更加易于实现。Dilithium方案的安全性基于模错误学习(Module Learning With Errors, MLWE)问题，其签名在选择消息攻击下具有存在不可伪造性。Dilithium使用的是一般格，要在保证Dilithium方案优点的基础上，进一步构造更加紧凑的签名方案，本文给出的一个解决思路是构造NTRU版的Dilithium签名。基于NTRU假设构造的后量子密码方案在工程上有很大的优势，其算法运行速度快，私钥尺寸非常小。尺寸较小的签名如文献[9,11]中的方案等都是基于NTRU假设，这些方案在签名过程中使用了需要高精度运算的离散高斯采样来生成秘密向量和错误向量，但生成服从离散高斯分布的样本在实现中比较复杂且容易遭到侧信道攻击[12–14]。Dilithium在一般格上，采用文献[4,15,16]等方案的策略，选择均匀采样设计。本文给出了一个基于“Fiat-Shamir with Aborts”方法[7]的高效格签名方案FatSeal(Fiat-shamir-transformation-based Signature algorithm with lattice), FatSeal签名的形式受到Dilithium签名的启发，但二者基于的问题不同。本文的方案基于NTRU假设进行设计，继承了Dilithium的简洁设计，签名过程简单，仅需均匀采样。由于使用了NTRU结构，FatSeal的公钥完全压缩，相较而言Dilithium需要进行额外压缩，FatSeal在密钥生成和验签算法上效率很高，方案的实用性很强。FatSeal的乘法运算在多项式环$\mathbb{Z}_q[x]/(x^n+1)$中进行，参数的选取保证运算能用数论变换(Number-Theoretic Transform, NTT)技术进行加速。FatSeal方案参数的设定由实际攻击决定，FatSeal方案私钥的恢复可以归结为NTRU格上CVP问题的求解，也可以转换为LWE实例进行

分析，签名的伪造可以归结为短整数解(Short Integer Solution, SIS)问题的求解。本文考虑了混合攻击、Primal攻击以及Dual攻击等具体攻击，由基本求解工具格基约化算法的计算复杂度确定FatSeal相应参数的安全级别，本文给出的两组参数能够使FatSeal签名方案在经典计算机下分别达到128 bit和256 bit安全性。综合来看，FatSeal签名方案在Fiat Shamir with Abort框架下创新地应用了高效的NTRU结构，同时保持了采样模块为均匀分布的简洁设计，能够抵抗已知的格攻击以达到安全目标，是一种非常高效简洁的格签名方案。

2　数学准备

2.1　记号描述

(1) 环与环运算：令q为一个正整数，对于环R，定义R_q为商环R/qR。本文选用商环$R_q=\mathbb{Z}_q[x]/(x^n+1)$，令$n$为2的次幂且$q$为奇素数满足$q=1\,(\mathrm{mod}2n)$。对于任意的$f=f_0+f_1x+\cdots+f_{n-1}x^{n-1}\in R_q$，将多项式$f$与向量$(f_0,f_1,\cdots,f_{n-1})\in\mathbb{Z}_q^n$等同，成立$x\cdot f=(-f_{n-1},f_0,\cdots,f_{n-2})$，$\mathrm{rot}(f):=\begin{pmatrix} f\\ xf\\ \vdots\\ x^{n-1}f\end{pmatrix}$则对任意两个多项式$a,b\in R_q$进行乘法运算有$ab=a\cdot\mathrm{rot}(b)\bmod q$。

(2) 剩余系：令p为2的某个小幂次，记正整数$\alpha=(q-1)/p$，并令$\mathbb{Z}_\alpha=\left\{-\frac{\alpha}{2},-\frac{\alpha}{2}+1,\cdots,\frac{\alpha}{2}-1\right\}$。对任一整数$y$做$\mathrm{mod}^\alpha q$运算时，即$x=y\bmod^\alpha q$，令$x$为剩余系$\mathbb{Z}_q=\left\{-\frac{\alpha}{2},-\frac{\alpha}{2}+1,\cdots,q-1-\frac{\alpha}{2}\right\}$中与$y$相对应的元素，对$x$关于$\alpha$做带余除法，定义

$$x=\mathrm{quo}(x)\alpha+\mathrm{rem}(x),\ \mathrm{rem}(x)\in\mathbb{Z}_\alpha \tag{1}$$

记$k=\mathrm{quo}(x)$，则$x\in I_k=k\alpha+\mathbb{Z}_\alpha$，这时有$\mathbb{Z}_q=\left\{q-1-\frac{\alpha}{2}\right\}\cup\cup_{k=0}^{p-1}I_k$。记号$x=y\bmod^\pm q$表示$x$是剩余系$\mathbb{Z}_q=\left\{-\frac{q-1}{2},-\frac{q+1}{2},\cdots,\frac{q-1}{2}\right\}$中与$y$相对应的元素；记号$x=y\bmod^+q$表示$x$是$y$在剩余系$\mathbb{Z}_q=\{0,1,\cdots,q-1\}$里对应的唯一元素；模运算$y\bmod q$表示元素所在的剩余系选择不影响运算结果，可以使用任一剩余系。

(3) 小系数多项式集合：集合B_t^n是$\mathbb{Z}_q[x]/(x^n+1)$的一个子集，集合中任一多项式恰有t个系数为1，其余系数为0。集合$T(d+1,d)$也是$\mathbb{Z}_q[x]/(x^n+1)$的一个子集，集合的任一多项式恰有$d+1$个系数为1，d个系数为–1，其余系数等于0。

(4) 元素的尺寸：对于任意$y\in\mathbb{Z}_q$，令$\|y\|_\infty:=\left|y\bmod^\pm q\right|$。对于任意的$f=f_0+f_1x+\cdots+f_{n-1}x^{n-1}\in R_q$，定义$\|f\|_\infty:=\max_i\|f_i\|_\infty$。

2.2　格与格基约化算法

在数学中，欧式空间上的一个离散加法子群成为一个格。特别地，设行向量$\boldsymbol{b}_1,\boldsymbol{b}_2,\cdots,\boldsymbol{b}_m\in\mathbb{R}^n$为$\mathbb{R}^n$中的线性无关向量，记其对应的矩阵为$\boldsymbol{B}$，$\mathcal{L}(\boldsymbol{B}):=\left\{\sum_{i=1}^m x_i\boldsymbol{b}_i|x_i\in\mathbb{Z},i=1,2,\cdots,m\right\}$构成一个$\mathbb{R}^n$上的格。称$\boldsymbol{B}$为对应格的一组格基。以下是两种常用的格：

定义1 q-ary格：对于任意的正整数m,n,q和矩阵$\boldsymbol{A}\in\mathbb{Z}_q^{m\times n}(m\ge n)$，定义$m$维$q$-ary格为

$$\Lambda^\perp(\boldsymbol{A})=\{\boldsymbol{x}\in\mathbb{Z}^m|\boldsymbol{x}\boldsymbol{A}=0\ \bmod\ q\} \tag{2}$$

定义2 NTRU格：对于任意$f,g\in T(d+1,d)$，其中f在$\mathbb{R}_q$中可逆，则关于$h=f^{-1}g$的NTRU格定义为

$$\mathcal{L}_h=\{(v,u)\in\mathcal{R}_q^2|uh=v\} \tag{3}$$

格$\mathcal{L}_h$也可以由矩阵$\begin{pmatrix} q\boldsymbol{I}_n & 0\\ \mathrm{rot(h)} & \boldsymbol{I}_n\end{pmatrix}$的行向量张成。

格基约化算法是分析格密码方案的实际困难性的重要工具。格基约化算法，包括(Lenstra–Lenstra–Lovász, LLL)算法[17], (Block Korkine-Zolotareff, BKZ)算法[18]等，其主要想法是将原有格基做垂直投影转化为低维格，在低维的投影格上利用最短向量问题(Shortest Vector Problem, SVP)求解算法找低维格的最短向量，再扩张到高维格上。例如，以块大小b为参数的BKZ算法首先调用b维SVP求解器去找$\pi_{i+1}(\boldsymbol{b}_i),\pi_{i+1}(\boldsymbol{b}_{i+1}),\cdots,\pi_{i+1}(\boldsymbol{b}_j)\ (i\le n,j=\min(n,i+b))$上的最短格向量(其中$\pi_i:\mathbb{R}^n\to\mathrm{span}(\boldsymbol{b}_1,\boldsymbol{b}_2,\cdots,\boldsymbol{b}_{i-1})^\perp$为投影映射)，之后将这组短向量扩展为高维格的一组格基，多次重复此过程进而求得一组较短的格基。关于BKZ约化基，有等比数列假设(Geometric Series Assumption, GSA)，BKZ约化基正交化后长度近似满足关系$\|\boldsymbol{b}_i^*\|^2/\|\boldsymbol{b}_1^*\|^2=r^{i-1}$，这里$r$是[3/4, 1)中的某个实数，称为GSA常数。

块大小b会影响输出格基的质量和算法总的运行时间，不同模型下的SVP求解算法有不同运行时间估计。一般而言运行时间约为$2^{c\cdot b}$，关于常数c的估计有如下结果：

(1) Classical：目前已知最好的经典SVP求解算法[19]取$\mathrm{c}=\log_2\sqrt{3/2}\approx0.292$;

(2) Quantum：目前已知最好的量子SVP求解算法[20]取$\mathrm{c}=\log_2\sqrt{13/9}\approx0.265$;

(3) Plausible：密码学界猜测未来的算力可能会达到$\mathrm{c}=\log_2\sqrt{4/3}\approx0.2075$[21]。

2.3 困难问题

FatSeal签名方案与以下格上的困难问题密切相关：

定义3 γ-SVP和uSVP：给定一个格$\mathcal{L}$和一个近似因子$\gamma \geq 1$，在格中找一个不长于$\gamma \cdot \lambda_1(\mathcal{L})$的非0向量叫做近似最短向量问题($\gamma$-SVP)，这里$\lambda_1(\mathcal{L})$表示格中最短非0向量的长度。对于特定的存在唯一最短非0向量的格$\mathcal{L}'$，寻找该向量即为唯一最短向量问题(uSVP)。

定义4 γ-CVP：给定一个n维格$\mathcal{L}$，一个目标向量$\boldsymbol{t} \in \mathbb{R}^n$和一个近似因子$\gamma \geq 1$，在格中寻找向量$\boldsymbol{y}$满足$\|\boldsymbol{t}-\boldsymbol{y}\| \leq \gamma \cdot \mathrm{d}(\boldsymbol{t}, \mathcal{L})$的问题叫做近似最近向量问题($\gamma$-CVP)，这里$\mathrm{d}(t, \mathcal{L})$表示目标向量到格的距离。

定义5 $\mathrm{SIS}_{q,n,m,\beta}$问题：对于从$\mathbb{Z}_q^{m\times n}(m \geq n)$中随机均匀选取的矩阵$\boldsymbol{A}$，求解短向量$\boldsymbol{x} \in \mathbb{Z}^m/\{0\}$使得$\boldsymbol{x}\boldsymbol{A} = 0 \bmod q$且$\|\boldsymbol{x}\|_\infty \leq \beta$。

这个问题也可以看成是$\Lambda^{\perp}(\boldsymbol{A})$格上找非0短向量的问题。

3 FatSeal签名方案

3.1 方案描述与正确性

在这个部分给出本文的新签名方案如表1、表2和表3所示及正确性证明。

FatSeal的私钥是从$T(d+1, d)$中均匀选取的短多项式f, g，其中要求f在R_q中可逆。公钥$h \in R_q$由$(g+\alpha)f^{-1}$计算得到。

表 1 FatSeal密钥生成算法

算法1 FatSeal.KeyGen()
输入：n, q
输出：公钥h，私钥(f, g)
(1) $f, g \leftarrow T(d+1, d)$
(2) 如果f在R_q中不可逆，跳转(1)
(3) 计算$h = (g+\alpha)f^{-1}$
(4) 输出$pk = h$, $sk = (f, g)$

表 2 FatSeal签名算法

算法2 FatSeal.Sign()
输入：消息M，公钥h，私钥(f, g)
输出：消息M的签名(z, c)
(1) $r \leftarrow \mathbb{Z}_\alpha[x]/(x^n+1)$
(2) 计算$w = hr \bmod^{\alpha} q$
(3) 若w的某分量等于$q-1-\alpha/2$，跳转(1)
(4) 计算$c = H(M, \mathrm{quo}(w))$
(5) 计算$z = r + cf$
(6) 若$\|cg\|_\infty \leq \gamma$, $\|cf\|_\infty \leq \gamma$, $\|cg + \mathrm{rem}(w)\|_\infty \leq \alpha/2-\gamma$, $\|z\|_\infty < \alpha/2-\gamma$，输出$(z, c)$
(7) 否则，跳转步骤(1)

表 3 FatSeal验签算法

算法3 FatSeal.Verify()
输入：消息M，公钥h，签名(z, c)
输出：验证成功输出1，否则输出0
(1) 如果$\|z\|_\infty \geq \alpha/2-\gamma$或$w' = hz - \alpha c \bmod^{\alpha} q$的某个分量等于$q-1-\alpha/2$，输出0
(2) 否则
(3) 计算$c' = H(M, \mathrm{quo}(w'))$
(4) 如果$c' = c$，输出1
(5) 否则，输出0

签名者要对消息M进行签名，首先从$R_\alpha = \mathbb{Z}_\alpha[x]/(x^n+1)$中随机选取多项式$r$，即$r$的每个系数在$\{-\alpha/2, -\alpha/2+1, \cdots, \alpha/2-1\}$中均匀选取。然后计算$w = hr \bmod^{\alpha} q$，如果$w$的某个分量为$q-1-\alpha/2$，则重新选取$r$。计算$c = H(M, \mathrm{quo}(w)) \in B_t^n$，哈希值$c$是$R_q$中的多项式，其中恰有$t$个系数为1，其余系数为0。接着计算$z = r + cf$，这里$cf$的系数绝对值的最大可能值是$t$，此时若直接输出$z$，则有可能泄露私钥的信息，FatSeal利用拒绝采样来保证签名算法是零知识的。对某个选定的正数$\gamma(\leq t)$，如果能同时满足以下4个条件：$\|cg\|_\infty \leq \gamma, \|cf\|_\infty \leq \gamma$, $\|cg + \mathrm{rem}(w)\|_\infty \leq \alpha/2-\gamma$, $\|z\|_\infty < \alpha/2-\gamma$，则输出签名$(z, c)$。输出的$z$在$(\gamma-\alpha/2,\ \alpha/2-\gamma) \cap \mathbb{Z}$上均匀分布，对于任意$x \in (\gamma-\alpha/2,\ \alpha/2-\gamma) \cap \mathbb{Z}$, $j \in \{-\gamma, -\gamma+1, \cdots, \gamma\}$，有$-\alpha/2+1 \leq x-j \leq \alpha/2-1$，可以计算得

$$\begin{aligned}\Pr[z_i = x] &= \Pr[(cf)_i = \gamma]\Pr[r_i = x-\gamma] + \cdots \\ &\quad + \Pr[(cf)_i = -\gamma]\Pr[r_i = x+\gamma] \\ &= \frac{1}{\alpha}\sum_{j=-\gamma}^{\gamma}\Pr[(cf)_i = j] \end{aligned} \tag{4}$$

即z_i取到$(\gamma-\alpha/2,\ \alpha/2-\gamma) \cap \mathbb{Z}$中各值的概率是均等的。

对于输入的签名(z, c)，验签算法检查了3个条件：

(1) $\|z\|_\infty < \alpha/2-\gamma$;

(2) 计算$w' = hz - \alpha c \bmod^{\alpha} q$, w'的每一个分量都不等于$q-1-\alpha/2$;

(3) $c = H(M, \mathrm{quo}(w'))$。

如果以上3个条件都能满足，则签名(z, c)通过验证。

对于合法签名(z, c)，条件(1)显然满足。计算$hz - \alpha c = hr + cg$，有

$$\begin{aligned} w' &= w + cg \bmod^{\alpha} q = \mathrm{quo}(w)\alpha + \mathrm{rem}(w) \\ &\quad + cg \bmod^{\alpha} q \end{aligned} \tag{5}$$

因为合法签名满足$\|cg+\mathrm{rem}(w)\|_\infty \le \alpha/2-\gamma$，故成立$\mathrm{quo}(w')=\mathrm{quo}(w)$，进而条件(2)和条件(3)可满足。

3.2 签名迭代次数

对于计算$w=hr \bmod^{\alpha} q$，如果w的某个分量为$q-1-\alpha/2$，则算法重启，假设w各分量独立且在剩余系内均匀分布，则签名算法在这一步不重启的概率为$(1-1/q)^n$。要输出一个签名，还需要满足$\|cg\|_\infty \le \gamma$，$\|cf\|_\infty \le \gamma$，$\|cg+\mathrm{rem}(w)\|_\infty \le \alpha/2-\gamma$，$\|z\|_\infty < \alpha/2-\gamma$这4个条件。注意到

$$\begin{aligned}&\Pr[|z_i|<\alpha/2-\gamma \,|\, |(cf)_i|\le\gamma]\\&=\frac{1}{\alpha}\sum_{k=-\gamma}^{\gamma}\Pr[|r_i+k\le\alpha/2-\gamma-1]\cdot\Pr[(cf)_i=k]\\&=\frac{\alpha-2\gamma-1}{\alpha}\end{aligned}\tag{6}$$

假设各分量独立，于是有$\Pr[\|z\|_\infty<\alpha/2-\gamma \,|\, \|cf\|_\infty\le\gamma]=\left(\frac{\alpha-2\gamma-1}{\alpha}\right)^n \approx \mathrm{e}^{-\frac{2\gamma+1}{\alpha}n}$。类似地，有$\Pr[\|cg+\mathrm{rem}(w)\|_\infty\le\alpha/2-\gamma \,|\, \|cg\|_\infty\le\gamma] \approx \mathrm{e}^{-\frac{2\gamma-1}{\alpha}n}$。

下面估计$\|cf\|_\infty\le\gamma$的概率。由$c\in B_t^n$是一个公开的值，有t个分量为1，其余为0，而$cf=c\cdot\mathrm{rot}(f)$，则cf的各分量是t个f的系数相加(减)得到。其中f的系数取1的概率为$(d+1)/n$，取−1的概率为d/n，取0的概率为$(n-2d-1)/n$。为了便于分析，这里考虑f的系数取1和−1的概率均为d/n，取0的概率为$(n-2d)/n$。由于f的系数取1和取−1的概率相同，此时cf的各分量即可视为是t个f的系数相加。参数选取时若使t的值大于30，那么应用中心极限定理可知cf的各系数服从期望为0，方差为$2td/n$的正态分布，从而$\|cf\|_\infty\le\gamma$的概率为$\mathrm{erf}\left(\frac{\gamma}{2\sqrt{td/n}}\right)$。

综合上述分析，可以估计出同时满足$\|cg\|_\infty\le\gamma$, $\|cf\|_\infty\le\gamma$, $\|cg+\mathrm{rem}(w)\|_\infty\le\alpha/2-\gamma$, $\|z\|_\infty<\alpha/2-\gamma$的概率为$\mathrm{e}^{-\frac{4\gamma}{\alpha}n}\cdot\mathrm{erf}\left(\frac{\gamma}{2\sqrt{td/n}}\right)^{2n}$。

4 FatSeal的参数选取

本节给出FatSeal方案的实例化。对于NTRU系统，其私钥一般从小系数集合中选取，本文将小系数集合选为$T(d,d+1)$。由于使用的多项式环为$\mathbb{Z}_q[x]/(x^n+1)$, FatSeal方案中的多项式乘法运算可以利用NTT技术进行高效实现。为使用NTT，本文选择的模数q为素数，满足$q\equiv 1 \bmod 2n$，群$\mathbb{Z}_q^*$中的$2n$阶元记为w。参数α取为$(q-1)/8$，此时整数关于α的商需要3 bit来表示。表4列出了两组参数使得FatSeal在经典计算机下分别具有128 bit和256 bit安全性，同时也给出了相应的迭代概率p，期望迭代$1/p$次FatSeal.Sign()算法能返回一个合法签名。在这两组参数下，哈希函数值域规模分别大于2^{256}和2^{512}。

表 4　FatSeal签名方案的参数选取及迭代概率估计

n	$T(d+1,d)$	q	ω	t	α	γ	迭代概率(%)
1024	T(257, 256)	286712	106	44	35840	20	10.2
2048	T(413, 412)	724993	287	87	90624	24	11.4

签名的尺寸由z主导，需要$n\lceil\log_2(\alpha/2-\gamma)\rceil/8$ Byte存储，公钥尺寸为$n\lceil\log_2(q)\rceil/8$Byte。表5给出FatSeal两套参数对应的规模，长度数据来源于FatSeal的参考实现。提交到NIST的基于格的Fiat-Shamir类型签名只有Dilithium和qTESLA[22]，Dilithuim方案在签名尺寸上比qTESLA的大，表5和表6的数据表明FatSeal作为Dilithium的NTRU变种在参数规模上优于同安全等级下的qTESLA方案。

5 具体安全性分析

FatSeal签名方案的密钥恢复攻击与伪造签名攻击依赖于NTRU问题与无穷范数下的SIS问题的困难性；这两个问题均是格理论中的经典困难问题。

5.1 私钥恢复

FatSeal体制私钥的安全性与NTRU格上求解近似CVP问题紧密相关。若能求解短多项式对$(f,g)\in R_q$使得$hf=g+\alpha$，则可以恢复体制的私钥。矩阵$\begin{pmatrix} q\boldsymbol{I}_n & 0 \\ \mathrm{rot}\,(h) & \boldsymbol{I}_n\end{pmatrix}$行向量张成的NTRU格记为$L_h$，则$(g+\alpha,f)\in L_h$。如果能求解$L_h$中离$(\alpha,0)\in R_q^2$足够近的格向量，则可以恢复$(f,g)$。

混合攻击：考虑矩阵$\boldsymbol{B}=\begin{pmatrix} q\boldsymbol{I}_n & 0 & 0 \\ \mathrm{rot}\,(h) & \boldsymbol{I}_n & 0 \\ \alpha & 0 & 1\end{pmatrix}=\begin{pmatrix} q\boldsymbol{I}_{r_1} & 0 & 0 \\ * & \boldsymbol{L}_1 & 0 \\ * & * & \boldsymbol{I}_{r_2}\end{pmatrix}\in\mathbb{Z}^{2n+1}$，其中$\boldsymbol{L}_1\in\mathbb{Z}^{N\times N}$。对

表 5　FatSeal签名128 bit和256 bit安全强度下的参数大小(Byte)

体制	公钥长度	私钥长度	签名长度
FatSeal-1024	2321	385	2048
FatSeal-2048	4984	719	4352

表 6　qTESLA签名128 bit和256 bit安全强度下的参数大小(Byte)

体制	公钥长度	私钥长度	签名长度
qTESLA-Ⅱ	2336	1600	2144
qTESLA-Ⅴ	5024	3520	4640

$\boldsymbol{L}_1$进行格基约化，再对列向量进行Gram-Schmidt正交化，最后进行旋转变换得到下三角矩阵$\boldsymbol{T}'$，即$\boldsymbol{T}'=\boldsymbol{T}'\boldsymbol{L}_1\boldsymbol{Y}'$，其中$\boldsymbol{U}'$是幺模矩阵，$\boldsymbol{Y}'$是正交阵。对$\boldsymbol{B}$进行变换使得子矩阵$\boldsymbol{L}_1$变为$\boldsymbol{T}'$，相应的矩阵$\boldsymbol{B}$变为矩阵$\boldsymbol{T}=\boldsymbol{U}\boldsymbol{B}\boldsymbol{Y}$，即

$$\boldsymbol{T}=\begin{pmatrix}\boldsymbol{I}_{r_1}&0&0\\0&\boldsymbol{U}'&0\\0&0&\boldsymbol{I}_{r_2}\end{pmatrix}\cdot\begin{pmatrix}q\boldsymbol{I}_{r_1}&0&0\\*&\boldsymbol{L}_1&0\\*&*&\boldsymbol{I}_{r_2}\end{pmatrix}\cdot\begin{pmatrix}\boldsymbol{I}_{r_1}&0&0\\0&\boldsymbol{Y}'&0\\0&0&\boldsymbol{I}_{r_2}\end{pmatrix}=\begin{pmatrix}q\boldsymbol{I}_{r_1}&0&0\\*&\boldsymbol{T}'&0\\*&*&\boldsymbol{I}_{r_2}\end{pmatrix}\tag{7}$$

其中，$(g,f,-1)\boldsymbol{Y}$是格$\mathcal{L}(\boldsymbol{T})$中的一个向量。设$\boldsymbol{T}$对角线上的元素分别为$q^{\alpha_1},q^{\alpha_2},\cdots,q^{\alpha_{2n+1}}$，有$\alpha_1+\alpha_2+\cdots+\alpha_{2n+1}=n$。当$1\le i\le r_1$时，$\alpha_i=1$；当$2n+1-r_2<i\le 2n+1$时，$\alpha_i=0$；当$r_1<i\le 2n+1-r_2$，根据GSA假设，$\alpha_i$线性递减，满足

$$\left.\begin{aligned}\alpha_{r_1}&=\frac{n-r_1}{N}+N\log_q(\delta)\\&\vdots\\\alpha_{2n+1-r_2}&=\frac{n-r_1}{N}-N\log_q(\delta)\end{aligned}\right\}\tag{8}$$

其中，δ是根Hermite因子。由文献[23]中的引理1，当$\boldsymbol{y}=\boldsymbol{u}\boldsymbol{T}+\boldsymbol{x}$，$\boldsymbol{u},\boldsymbol{x}\in\mathbb{Z}^{2n+1}$，若$\boldsymbol{x}$的各分量满足$-T_{i,i}/2<x_i\le T_{i,i}/2$，$1\le i\le 2n+1$，针对$\boldsymbol{T}$应用Babai算法约化$\boldsymbol{y}$可以恢复$\boldsymbol{x}$。对于$\mathcal{L}(\mathbf{B})$中的最短向量$\boldsymbol{v}$，若能保证$\alpha_{r_2}>\log_q(2||\boldsymbol{v}||_\infty)$，枚举$\boldsymbol{v}$的最后$r_2$个分量，就能找到$\boldsymbol{v}$。攻击运行时间为格基约化时间加上枚举时间，利用Grover算法加速搜索，运行时间估计为$2^{c\cdot b}+3^{0.5\cdot r_2}$。通过平衡约化时间和搜索时间可以得到该攻击下的比特安全性，如表7所示。

Primal攻击：取$\mathrm{rot}(h)$的任意m列得到矩阵$\boldsymbol{A}\in\mathbb{Z}^{n\times m}$，多项式$-g$和$f$对应的列向量分别记作$\boldsymbol{e}\in\mathbb{Z}^m$和$\boldsymbol{s}\in\mathbb{Z}^n$，令$\boldsymbol{b}=(\alpha,0,\cdots,0)\in\mathbb{Z}^m$。考虑LWE实例$(\boldsymbol{A},\boldsymbol{b}=\boldsymbol{s}\boldsymbol{A}+\boldsymbol{e})$，构造维数为$d=m+n+1$的格

$$\Lambda=\left\{\boldsymbol{x}\in\mathbb{Z}^{n+m+1}\middle|x\begin{pmatrix}\boldsymbol{A}\\-\boldsymbol{I}_n\\-\boldsymbol{b}\end{pmatrix}=0\mod q\right\}\tag{9}$$

其中$(\boldsymbol{s},\boldsymbol{e},1)$是格$\Lambda$的一个短向量，长度近似为$\sigma\sqrt{m+n}$，这里$\sigma$为$\boldsymbol{s},\boldsymbol{e}$的标准差。利用BKZ算法求解短向量，攻击能够生效当且仅当块大小b满足$\sigma\sqrt{b}\le\delta^{2b-d-1}q^{\frac{m}{d}}$，其中$\delta=\left((\pi b)^{\frac{1}{b}}\dfrac{b}{2\pi\mathrm{e}}\right)^{\frac{1}{2(b-1)}}$为根Hermite因子，该攻击运行时间约为$2^{c\cdot b}$。调整$m$和$b$的大小使得运行时间最少，以此给出Primal攻击的具体复杂度估计。

Dual攻击：在对偶攻击的模型下，考虑LWE实例$(\boldsymbol{A},\boldsymbol{b}=\boldsymbol{s}\boldsymbol{A}+\boldsymbol{e})$，可以构造$d=m+n$维格$\Lambda=\left\{(\boldsymbol{x},\boldsymbol{y})\in\mathbb{Z}^{m+n}|\boldsymbol{A}\boldsymbol{x}^{\mathrm{T}}=\boldsymbol{y}^{\mathrm{T}}\bmod q\right\}$。由文献[24]中的结论，BKZ算法可以找到格$\Lambda$中的一个长度为$l=\delta^{d-1}q^{n/d}$的短向量$\boldsymbol{v}=(\boldsymbol{x},\boldsymbol{y})$，$\boldsymbol{v}^{\mathrm{T}}\boldsymbol{b}$的分布和均匀分布的统计距离不超过$\epsilon=4\exp\left(-2\pi^2\tau^2\right)$，这里$\tau=l\sigma/q$。用这个攻击即可以优势$\varepsilon$打破判定性LWE问题。攻击的计算复杂度由BKZ算法给出，块大小取为b，重复运行筛法$R=\max\left(1,1/\left(\gamma\varepsilon^2\right)\right)$次可以找到$1/\epsilon^2$个短向量，这里$\gamma=\sqrt{4/3}^b$是筛法$b$维投影格里最短向量个数的估计。

NewHope模型下的Primal攻击和Dual攻击复杂度如表8所示。

5.2 签名伪造

对于FatSeal体制的签名伪造，可以考虑求解SIS问题。给定公钥h，如果能对$v\in\{0,1,\cdots,p-1\}$和消息M找到找到足够短的$(\boldsymbol{z},\boldsymbol{r})\in\mathbb{Z}_q\times\mathbb{Z}_\alpha$，使得$h\boldsymbol{z}-\boldsymbol{r}=(v+c)\alpha$成立，这里$c=H(M,v)$，则可以得到关于消息M的一个合法签名$(\boldsymbol{z},\boldsymbol{r})$。

对$h\boldsymbol{z}-\boldsymbol{r}=(v+c)\alpha$的求解可以转化为SIS问题的求解，记$\boldsymbol{A}=\begin{pmatrix}\mathrm{rot}(h)\\\boldsymbol{I}_n\\(v+c)\,\alpha\end{pmatrix}$，考虑解齐次方程$\boldsymbol{x}\boldsymbol{A}=0\bmod q$使得$\|\boldsymbol{x}\|_\infty<\alpha/2-\gamma$。从矩阵$\boldsymbol{A}$中任选$w(>n)$行，余下的行对应的解分量取为0。对于这$w$行张成的格，求解其正交格的一组格基，经过BKZ算法约化后格基形为$\begin{pmatrix}q\boldsymbol{I}_{r_1}&0&0*&*&0*&*&\boldsymbol{I}_{r_2}\end{pmatrix}$，其中$0\le r_1,r_2<w$。由GSA假设$\log_2\|\boldsymbol{b}_i^*\|\,(r_1<i<w-r_2)$满足斜率为$\dfrac{1}{b-1}\log_2\left(\dfrac{b}{2\pi\mathrm{e}}(\pi b)^{\frac{1}{b}}\right)$的线性关系。利用SVP求解算法可以得到$\sqrt{4/3}^b$个$\pi_{r_1}(\mathcal{L})$上的向量，长度约为$\left\|\mathbf{b}_{r_1+1}^*\right\|$。可假设这$\sqrt{4/3}^b$个向量的前

表 7 混合攻击下的FatSeal比特安全性

模型	体制	b	N	r_1	比特安全性
classical	FatSeal-1024	510	1770	91	150
quantum	FatSeal-1024	522	1799	75	139
plausible	FatSeal-1024	549	1865	40	115
classical	FatSeal-2048	1130	3440	240	331
quantum	FatSeal-2048	1157	3503	207	307
plausible	FatSeal-2048	1219	3646	132	253

表 8 NewHope攻击模型下的比特安全性

体制	攻击模型	m	b	classical	quantum	plausible
FatSeal-1024	primal	874	503	147	133	104
FatSeal-1024	dual	862	502	146	133	104
FatSeal-2048	primal	1671	1076	314	285	223
FatSeal-2048	dual	1697	1072	313	284	222

r_1个分量在$\mathbb{Z}_q$上均匀分布，余下各分量满足中心等于0，方差等于$\left\|\boldsymbol{b}^*_{r_1+1}\right\|^2/(w-r_1-r_2)$的高斯分布。设向量前$w-r_2$个分量绝对值不超过$\alpha/2$的概率为$p$，那么总的运行时间等于$2^{c\cdot b}/p$，利用此分析可以得到表9。

表 9　SIS攻击下的比特安全性

模型	体制	b	w	比特安全性
classical	FatSeal-1024	577	2048	181
quantum	FatSeal-1024	577	2048	166
plausible	FatSeal-1024	577	2048	132
classical	FatSeal-2048	1276	4039	379
quantum	FatSeal-2048	1278	4041	344
plausible	FatSeal-2048	1284	4049	270

6　结束语

FatSeal签名方案在Dilithium签名方案的基础上融合了NTRU的优势，是一个新的基于NTRU格的Fiat-Shamir签名构造，相对于已有的Fiat-Shamir格签名得到了效率上的提升和参数规模的减小。FatSeal签名方案两套参数的设置能够抵抗现有已知最好的攻击，在经典计算机下分别能够达到128 bit安全性和256 bit安全性。未来的工作是继续研究FatSeal在量子RO模型下的可证安全性以及改进FatSeal的具体安全性评估模型。对于FatSeal的参考实现，在算法优化这一块尚有许多工作可以开展，如考虑基于AVX的多项式乘法加速，从而进一步提升FatSeal签名的实用性。

参 考 文 献

[1] GOLDREICH O, GOLDWASSER S, and HALEVI S. Public-key cryptosystems from lattice reduction problems[C]. The 17th Annual International Cryptology Conference, Santa Barbara, USA, 1997: 112–131. doi: 10.1007/BFb0052231.

[2] BABAI L. On Lovász' lattice reduction and the nearest lattice point problem[J]. *Combinatorica*, 1986, 6(1): 1–13. doi: 10.1007/BF02579403.

[3] HOFFSTEIN J, PIPHER J, and SILVERMAN J H. NSS: An NTRU lattice-based signature scheme[C]. International Conference on the Theory and Application of Cryptographic Techniques, Innsbruck, Austria, 2001: 211–228. doi: 10.1007/3-540-44987-6.

[4] NGUYEN P Q and REGEV O. Learning a parallelepiped: Cryptanalysis of GGH and NTRU signatures[C]. The 24th Annual International Conference on the Theory and Applications of Cryptographic Techniques, St. Petersburg, 2006: 271–288. doi: 10.1007/11761679_17.

[5] GENTRY C, PEIKERT C, and VAIKUNTANATHAN V. Trapdoors for hard lattices and new cryptographic constructions[C]. The 40th Annual ACM Symposium on Theory of Computing, Victoria, 2008: 197–206. doi: 10.1145/1374376.1374407.

[6] FOUQUE P A, HOFFSTEIN J, KIRCHNER P, *et al*. Fast-fourier lattice-based compact signatures over NTRU[EB/OL]. https://falcon-sign.info/, 2019.

[7] LYUBASHEVSKY V. Fiat-shamir with aborts: Applications to lattice and factoring-based signatures[C]. The 15th International Conference on the Theory and Application of Cryptology and Information Security, Tokyo, 2009: 598–616. doi: 10.1007/978-3-642-10366-7_35.

[8] LYUBASHEVSKY V. Lattice signatures without trapdoors[C]. The 31st Annual International Conference on the Theory and Applications of Cryptographic Techniques, Cambridge, 2012: 738–755. doi: 10.1007/978-3-642-29011-4_43.

[9] DUCAS L, DURMUS A, LEPOINT T, *et al*. Lattice signatures and bimodal gaussians[C]. The 33rd Annual Cryptology Conference, Santa Barbara, 2013: 40–56. doi: 10.1007/978-3-642-40041-4_3.

[10] AVANZI R, BOS J, DUCAS L, *et al*. Cryptographic suite for algebraic lattices[EB/OL]. https://pq-crystals.org/, 2019.

[11] DUCAS L, LYUBASHEVSKY V, and PREST T. Efficient identity-based encryption over NTRU lattices[C]. The 20th International Conference on the Theory and Application of Cryptology and Information Security, Kaoshiung, 2014: 22–41. doi: 10.1007/978-3-662-45608-8_2.

[12] BRUINDERINK L G, HÜLSING A, LANGE T, *et al*. Flush, gauss, and reload - a cache attack on the BLISS lattice-based signature scheme[C]. The 18th International Conference on Cryptographic Hardware and Embedded Systems, Santa Barbara, 2016: 323–345. doi: 10.1007/978-3-662-53140-2_16.

[13] ESPITAU T, FOUQUE P, GÉRARD B, *et al*. Side-channel attacks on bliss lattice-based signatures: Exploiting branch tracing against strongswan and electromagnetic emanations in microcontrollers[C]. The 2017 ACM SIGSAC Conference on Computer and Communications Security, Dallas, 2017: 1857–1874. doi: 10.1145/3133956.3134028.

[14] PESSL P, BRUINDERINK L G, and YAROM Y. To BLISS-B or not to be: Attacking strongSwan's implementation of post-quantum signatures[C]. The 2017 ACM SIGSAC Conference on Computer and Communications Security, Dallas, 2017: 1843–1855. doi: 10.1145/3133956.3134023.

[15] GÜNEYSU T, LYUBASHEVSKY V, and PÖPPELMANN T. Practical lattice-based cryptography: A signature scheme

for embedded systems[C]. The 14th International Workshop on Cryptographic Hardware and Embedded Systems, Leuven, 2012: 530–547. doi: 10.1007/978-3-642-33027-8_31.

[16] BAI Shi and GALBRAITH S D. An improved compression technique for signatures based on learning with errors[C]. Cryptographers' Track at the RSA Conference, San Francisco, 2014: 28–47. doi: 10.1007/978-3-319-04852-9_2.

[17] LENSTRA A K, LENSTRA H W Jr, and LOVÁSZ L. Factoring polynomials with rational coefficients[J]. *Mathematische Annalen*, 1982, 261(4): 515–534. doi: 10.1007/BF01457454.

[18] SCHNORR C P and EUCHNER M. Lattice basis reduction: Improved practical algorithms and solving subset sum problems[J]. *Mathematical Programming*, 1994, 66(1/3): 181–199. doi: 10.1007/BF01581144.

[19] LAARHOVEN T. Search problems in cryptography: From fingerprinting to lattice sieving[D]. [Ph.D. dissertation], Eindhoven University of Technology, 2015.

[20] BECKER A, DUCAS L, GAMA N, *et al*. New directions in nearest neighbor searching with applications to lattice sieving[C]. The 27th Annual ACM-SIAM Symposium on Discrete Algorithms, Arlington, 2016: 10–24. doi: 10.1137/1.9781611974331.

[21] LAARHOVEN T, MOSCA M, and VAN DE POL J. Finding shortest lattice vectors faster using quantum search[J]. *Designs, Codes and Cryptography*, 2015, 77(2/3): 375–400. doi: 10.1007/s10623-015-0067-5.

[22] AKLEYLEK S, ALKIM E, BARRETO P S L M, *et al*. qTesla[EB/OL]. https://qtesla.Org, 2019.

[23] HOWGRAVE-GRAHAM N. A hybrid lattice-reduction and meet-in-the-middle attack against NTRU[C]. The 27th Annual International Cryptology Conference, Santa Barbara, 2007: 150–169. doi: 10.1007/978-3-540-74143-5_9.

[24] ERDEM A, DUCAS L, PÖPPELMAN T, *et al*. Post-quantum key exchange-a new hope[C]. The 25th USENIX Security Symposium, Vancouver, 2016: 327–343.

谢天元：女，1992年生，博士，研究方向为格密码算法设计与分析.
李昊宇：男，1990年生，博士，研究方向为格密码算法设计与分析.
朱熠铭：男，1994年生，博士，研究方向为格算法、编码.
潘彦斌：男，1982年生，副研究员，研究方向为格算法及格密码算法分析.
刘 珍：女，1994年生，博士，研究方向为格算法及格密码算法分析.
杨照民：男，1995年生，硕士，研究方向为格算法、后斯诺登密码.