4-2020

On the security of LWE cryptosystem against subversion attacks

Zhichao YANG

Rongmao CHEN

Chao LI

Longjiang QU

Guomin YANG
Singapore Management University, gmyang@smu.edu.sg

Advance Access publication on 10 September 2019
doi: 10.1093/comjnl/bxz084

On the Security of LWE Cryptosystem against Subversion Attacks

Zhichao Yang[1], Rongmao Chen[1,*], Chao Li[1,2], Longjiang Qu[2] and Guomin Yang[3]

[1]*410000 College of Computer, National University of Defence Technology, Hunan, Changsha, P. R. China*
[2]*410000 College of Liberal Arts and Sciences, National University of Defence Technology, Hunan, Changsha, P. R. China*
[3]*2500 School of Computing and Information Technology, University of Wollongong, Wollongong, Australia*
**Corresponding author: chromao@nudt.edu.cn*

Subversion of cryptography has received wide attentions especially after the Snowden Revelations in 2013. Most of the currently proposed subversion attacks essentially rely on the freedom of randomness choosing in the cryptographic protocol to hide backdoors embedded in the cryptosystems. Despite the fact that significant progresses in this line of research have been made, most of them mainly considered the classical setting, while the research gap regarding subversion attacks against post-quantum cryptography remains tremendous. Inspired by this observation, we investigate a subversion attack against existing protocol that is proved post-quantum secure. Particularly, we show an efficient way to undetectably subvert the well-known lattice-based encryption scheme proposed by Regev (STOC 2005). Our subversion enables the subverted algorithm to stealthily leak arbitrary messages to the outsider who knows the backdoor. Through theoretical analysis and experimental observations, we demonstrate that the subversion attack against the LWE encryption scheme is feasible and practical.

Keywords: Public Key; Subversion Attack; Post-Quantum; Lattice

Received 18 January 2019; Revised 23 May 2019; Editorial Decision 16 July 2019
Handling editor: Liqun Chen

1. INTRODUCTION

The revelations of Edward Snowden in 2013 indicated [1–3] that, in practice, cryptosystems may be insecure as they could be possibly embedded with backdoors. Precisely, cryptographic backdoors could make the system far less reliable as thought and even completely broken. The subverted cryptosystems may still behave normally while in fact it could undetectably leak secret information via the public channel to the outside world. The attacker who plants the backdoor into the system could recover the secret by simply collecting and analysing all public communication transcripts of the cryptosystems.

Since modern cryptographic implementations are usually of extreme complexities, even cryptographic experts cannot easily detect these backdoors and thus let alone typical users. Even if such code is proved to be safe, the compiler or interpreter may also be subverted which makes that code less "clean". To formalize such strong attacks, in 1996, Young and Yung [4, 5] introduced the concept of `Kleptography` which models the cryptographic subversion in the reality. Since then on, subversion attacks (SAs) against cryptographic systems have received wide attentions and particularly the Snowden revelations reemphasized the need to further explore the power of subverting cryptographic systems and effective countermeasures in practice [6–9].

With the development of quantum technology, most of the current public-key encryption schemes, especially those rely on the hardness of discrete logarithms or big integer factorization problems, are insecure any more when facing quantum computers. Thus, post-quantum cryptography algorithms have received lots of attentions recently. Among them, lattice-based cryptography is regarded as the most promising candidate because of its great performance and strong security guarantee. NTRU and LWE (learning with errors) cryptosystems are two of the most famous types of lattice-based

cryptographies. As one of the most well-known lattice-based schemes, NTRU cryptosystems [10] has been standardized by the IEEE, and its security originally relies on the NTRU problem [11] which is a heuristic argument. Moreover, some other provably secure NTRU variants [12–14] suffer from low efficiency. The LWE cryptography was introduced by Regev in [15]. For some suitable parameters, its security relies on the hardness of the GapSVP (decisional approximate shortest vector problems; SVP) on arbitrary random lattices; it is also believed to be post-quantum secure.

Most recently, National Institute of Standards and Technology (NIST) decided to standardize the post-quantum cryptography and analyse lattice-based cryptosystems for better understanding. By now, people have proposed various attacks to analyse the security of lattice-based cryptosystems. Some are mathematical, e.g. combinatorial attacks [16, 17], while some are based on algorithmic methods, such as lattice-reduction attacks [18, 19]. Though these cryptanalysis results can be used as a somewhat systematical methodology to estimate the security of lattice-based cryptography, few works have been done to explore the SA on post-quantum cryptography. In [20], a class of possible backdoors has been proposed to subvert the NewHope system. However, those backdoors are only applicable to fixed public parameters and can be trivially disabled via changing the fixed parameter to a hash value of a common reference string. Kwant and Thissen [21] designed a backdoor for NTRU such that each ciphertext contains some underlying message which can only be recovered efficiently by the backdoor owner. For NTRU and LWE based signature schemes, the backdoor owner can also get some information about the signing key. Most recently, Xiao and Yu [22] showed how to subvert the classical Ring-LWE scheme. As the security of the proposed backdoor also relies on the Ring-LWE problem, the whole scheme is post-quantum secure.

Our Contributions. In this paper, we show how to perform SA on a classical LWE encryption scheme proposed by Regev [15]. Similarly, the backdoor we design is based on a post-quantum cryptographic scheme, and the whole scheme is considered under the post-quantum setting. We insist that the SA in NTRU [21] relies on an ECC-based backdoor and thus cannot resist quantum-computation analysis. Our technical idea is to select a random vector $\boldsymbol{x}$ carefully in the process of the LWE encryption, such that the ciphertext of LWE scheme will contain some other information which can be recovered by the attacker. Compared with the other lattice-based SA [22], the vector $\boldsymbol{x}$ in our new subverted encryption can be calculated in advance; thus our embedded backdoor will not slow down the modified encryption algorithm, and we demonstrate via experiment that our proposed new SA is practical when the related parameters are not too large.

Organizations. We organize the rest of the paper as follows. Section 2 describes some useful notations and concepts. The description of NTRU cryptosystems and the LWE cryptosystem will be given in Section 3. After some definitions in Section 4, our LWE encryption backdoor will be introduced in Section 5. In Section 6, experiments have been done to evaluate the backdoor. Section 7 presents some countermeasures. Finally, Section 8 concludes this paper.

2. NOTATIONS AND DEFINITION

We denote $\mathbb{Z}$ be the integer ring, and $\mathbb{Z}_q$ is defined by $\mathbb{Z}/q\mathbb{Z}$ where q is a positive parameter. We suppose that a bold letter $\boldsymbol{v}$ represents vector in row notation and $\boldsymbol{v}_i$ is the i-th component of $\boldsymbol{v}$. The capital letter $\boldsymbol{A}$ denotes a matrix.

For any constant c, a function $f(n) = o(n^{-c})$ is said to be `negligible`. Generally, if ϵ is a negligible function then $1 - \epsilon$ is `overwhelming`. The notion $z \leftarrow\!\!\$\ D$ means that the variable z is sampled from the distribution D, and the probability of $z = x$ is denoted by $D(x)$. $\log(\cdot)$ is the base 2 logarithms, respectively.

2.1 Background on Lattices

2.1.1 *Lattice*

A lattice $\mathcal{L}$ is a discrete subgroup in $\mathbb{R}^n$ generated by several linear independent vectors $\boldsymbol{b}_1, \cdots, \boldsymbol{b}_{\mathfrak{m}} \in \mathbb{Z}^n$ over the integer ring, and $m \leq n$,

$$\mathcal{L} = \{a_1\boldsymbol{b}_1 + \cdots + a_{\mathfrak{m}}\boldsymbol{b}_{\mathfrak{m}} : a_1 \in \mathbb{Z}, \cdots, a_{\mathfrak{m}} \in \mathbb{Z}\}.$$

The volume $\mathrm{vol}(\mathcal{L})$ of this lattice is $\sqrt{\det(\boldsymbol{B} \cdot \boldsymbol{B}^T)}$, where $\boldsymbol{B}^T = (\boldsymbol{b}_1^T, \cdots, \boldsymbol{b}_m^T)$ is a basis of this lattice. The dual lattice $\mathcal{L}^*$ is defined as follows:

$$\mathcal{L}^* = \{\boldsymbol{x} \in \mathbb{R}^n : \forall \boldsymbol{v} \in \mathcal{L}, \langle \boldsymbol{x}, \boldsymbol{v} \rangle \in \mathbb{Z}\}.$$

The length of the shortest non-zero vectors in $\mathcal{L}$ is represented by $\lambda_1(\mathcal{L})$. The famous shortest vector problem (SVP) is to find a shortest vector in a random lattice $\mathcal{L}$ [23]. The approximating SVP, $\mathrm{GapSVP}_{f(m)}$ is to find some lattice vectors of length within $f(m)\lambda_1(\mathcal{L})$, where $f(m)$ is a function of m and an approximating factor. A breakthrough result of Ajtai [24] proved that the SVP is NP-hard. Another proof by Micciancio [23] asserts that, under the randomized reduction, approximating SVP within a constant factor is NP-hard. For the latest development we refer to Khot [25], under the same assumption, approximating SVP within a quasi-polynomial factor is also NP-hard.

2.1.2 *Gaussian Measures*

The gaussian distribution over $\mathbb{Z}$ is represented by D_s, and the parameter s is called deviation. Given a function $\rho_s(x) = \exp\left(-\frac{x^2}{s}\right)$, the value of $D_s(x)$ is equal to

$$D_s(x) = \rho_s(x)/\rho_s(\mathbb{Z}),$$

where $\rho_s(\mathbb{Z}) = \sum_{x \in \mathbb{Z}} \rho_s(x)$.

LEMMA 2.1. *Suppose $n > 0$ belongs to $\mathbb{Z}$, and for any parameter $s \geq \omega(\sqrt{\log n})$, the following inequations will hold:*

$$Pr_{\boldsymbol{x} \leftarrow\$ D_{\mathbb{Z}^n,s}}[||\boldsymbol{x}|| > s\sqrt{n}] \leq 2^{-n+1}.$$

The smoothing parameter η_ϵ [26] is defined as the smallest positive parameter s such that $\rho_{1/s}(\mathcal{L}^* \setminus \{\mathbf{0}\}) \leq \epsilon$, for any $\epsilon > 0$.

LEMMA 2.2. *$\mathcal{L}$ denotes an arbitrary n-dimensional lattice, and $\lambda_1(\mathcal{L}^*)$ is the length of the shortest vector in the dual lattice $\mathcal{L}^*$; the smoothing parameter will satisfy*

$$\eta_\epsilon(\mathcal{L}) \leq \sqrt{n}/\lambda_1\left(\mathcal{L}^*\right),$$

where $\epsilon = 2^{-m}$.

We say that a real variable X follows *subgaussian* distribution with a parameter s if

$$\Pr[|X| > t] \leq 2\exp\left(-\pi t^2/s^2\right),$$

for every $t \leq 0$. That means the distribution of X is dominated by a Gaussian.

2.2 Public Key Encryption

Firstly, the syntax of public-key encryption is given below.

DEFINITION 2.1. *There are three probabilistic polynomial-time (PPT) algorithms $(\mathsf{KGen}, \mathsf{Enc}, \mathsf{Dec})$ in the public-key encryption scheme Π which satisfy the following:*

- $\mathsf{KGen}(1^n)$*: It begins with a security parameter 1^n and returns $(\mathsf{pk}, \mathsf{sk})$. We refer to pk as the* `public key` *and sk as the* `private key`*.*
- $\mathsf{Enc}(m, \mathsf{pk})$*: Enc takes a message m and the public key pk as input, output $c \leftarrow\$ \mathsf{Enc}(m, \mathsf{pk})$ as ciphertext.*
- $\mathsf{Dec}(c, \mathsf{sk})$*: The inputs of Dec are cipertext c and private key sk. Dec returns a symbol $\perp$ denoting failure or outputting the message m. In the later case, $m = \mathsf{Dec}(c, \mathsf{sk})$.*

The correctness of a public key encryption scheme requires that for any parameter 1^n, $(\mathsf{pk}, \mathsf{sk}) \leftarrow\$ \mathsf{KGen}(1^n)$ and appropriate message m, we have

$$\mathsf{Dec}\left(\mathsf{Enc}\left(m, \mathsf{pk}\right), \mathsf{sk}\right) = m.$$

In this paper, we mainly consider the security under the indistinguishability from random 'bits/strings' under chosen-plaintext attack (`IND$-CPA`) security [27] which is defined by a game between an adversary $\mathcal{A}$ and a challenger $\mathcal{C}$.

Game $\texttt{IND\$-CPA}^{\mathcal{A}}_{\mathsf{Enc}}(n)$	$\overline{\mathsf{Enc}}(m, \mathsf{pk}, b)$
$b \leftarrow\$ \{0,1\}$	**if** $(b = 1)$ **then**
$(\mathsf{pk}, \mathsf{sk}) \leftarrow\$ \mathsf{KGen}(1^n)$	$c \leftarrow\$ \mathsf{Enc}(\mathsf{pk}, m)$
$c \leftarrow\$ \overline{\mathsf{Enc}}(m, \mathsf{pk}, b)$	**else**
$b' \leftarrow\$ \mathcal{A}(c, \mathsf{pk})$	$c \leftarrow\$ \{0,1\}^{\lvert c\rvert}$
return $b = b'$	**return** c

FIGURE 1. Game used to define `IND$-CPA` security.

As shown in the Fig. 1, $\mathcal{C}$ samples a random coin b and generates a key pair $(\mathsf{pk}, \mathsf{sk})$. $\mathcal{A}$ selects messages and issues queries on the encryption oracle. $\overline{\mathsf{Enc}}$ returns a ciphertext of m when $b = 1$, returns a random string otherwise. Finally, $\mathcal{A}$ takes c and pk as inputs and outputs a gauss b'.

DEFINITION 2.2. *$\Pi = (\mathsf{KGen}, \mathsf{Enc}, \mathsf{Dec})$ is a public-key scheme, and ϵ denotes a negligible function; the scheme is* `IND$-CPA` *secure if for arbitrary PPT adversaries $\mathcal{A}$*

$$Pr[\texttt{IND\$-CPA}^{\mathcal{A}}_{\mathsf{Enc}}(n) = 1] \leq \frac{1}{2} + \epsilon,$$

where the probability depends on both coins used by $\mathcal{A}$ and the experiment.

3. LATTICE-BASED CRYPTOSYSTEMS

3.1 NTRU Cryptosystems

We give a brief introduction about a NTRUEncrypt cryptosystems which is sometimes referred to as SVES-3, and it is also a candidate in NIST-round 1 [28]. The NTRU cryptosystems are determined by some sets of polynomials and three positive integer parameters p, q and N, where N is set to be prime. The parameter q is much bigger than p and satisfies $\gcd(p, q) = 1$. Three sets

$$B_N = \{\text{binary polynomials}\}$$

$$T_N = \{\text{trinary polynomials}\}$$

and $T_N(r, s)$ contain polynomials in T_N with r ones and s minus ones. All of them are subsets of the polynomial ring $R = \mathbb{Z}[X]/(X^N - 1)$.

The multiplication in R is denoted by $*$, and in NTRU cryptosystems it is performed modulo parameter q or p. The hash function H used in NTRU cryptosystems is to map arbitrary messages to a binary string with arbitrary length, and the sampling algorithms are determined by some random seeds. For simplicity, we assume that the message space is R_2 and no padding method is used. Let $\Pi_{\texttt{ntru}} = (\mathsf{KGen}_{\texttt{ntru}}, \mathsf{Enc}_{\texttt{ntru}}, \mathsf{Dec}_{\texttt{ntru}})$ be the NTRU encryption scheme which is constructed as followings

$\mathsf{KGen}_{\mathtt{ntru}}(1^N)$: It begins with two small polynomials $g \in T_N$ and $f \in T_N(d+1, d)$, f is invertible. We generate the public key h by

$$h \equiv g/(pf+1) \mod q,$$

and the private key is (pf, g).

$\mathsf{Enc}_{\mathtt{ntru}}(m, h)$: To encrypt the message m with the public key h, one takes a seed `rseed` as input and samples a polynomial r from T_N, where $\mathtt{rseed} = H(m, h)$ and computes

$$t = r * h.$$

Then, he samples m_{mask} from T_N based on seed $\mathtt{tseed} = H(t)$ and calculates $m' = m - m_{mask} \mod p$. Finally, output ciphertext as

$$e = t + m'$$

$\mathsf{Dec}_{\mathtt{ntru}}(e, f)$: To decrypt e with the private key f, one recovers m' by

$$m' = f * e \mod p,$$

and gets $t = e - m$. One recovers m_{mask} from T_N on $\mathtt{tseed} = H(t)$ and obtains $m = m' + m_{mask} \mod p$; the parameter r can also be sampled with $\mathtt{rseed} = H(m, h)$. Finally, outputs m if the equation

$$pr * h = t$$

holds, otherwise outputs $\perp$.

3.1.1 Security

As a well-known lattice-based public key scheme, redthe NTRU cryptosystem is efficient and standardized but lacks a solid security guarantee. The first provably secure variant of NTRUEncrypt was proposed by Stehlé and Steinfeld [12] in 2011, which is defined over power-of-2 cyclotomic rings and usually denoted by pNE. The pNE scheme is IND-CPA secure if some classical problems over ideal lattices are hard. Recently, a new variant is constructed by Yu, Xu and Wang [13]. They changed the ring into prime cyclotomic rings at the cost of using rather large parameters. Later, they [14] generalized cyclotomic ring by modifying the key generation algorithm.

Our work is mainly based on the pNE scheme and assumes that the scheme has `IND$-CPA` security, which means that the $\boldsymbol{c}$ outputted by algorithm $\mathsf{Enc}_{\mathtt{ntru}}$ is indistinguishable from vector randomly chosen from $\mathbb{Z}_q^N$ and it is a little different from the `IND-CPA` security. Because in the `IND-CPA` secure model, the adversary is asked to judge the output ciphertext comes from which message and no random vector is considered. We believe that our assumption is reasonable; otherwise a distinguisher can be used to attack the NTRU scheme.

3.2 Learning With Errors Cryptosystems

3.2.1 Learning With Errors (LWE)

Regev [15] first proposed the average-case `learning with errors` problem in 2005. n and q are arbitrary positive integers, and χ denotes the error distribution over $\mathbb{Z}$; the LWE distribution is defined as follows.

DEFINITION 3.1. *A vector $\boldsymbol{s} \in \mathbb{Z}_q^n$ is called secret and the LWE distribution $A_{\boldsymbol{s},\chi}$ over $\mathbb{Z}_q^n \times \mathbb{Z}_q$ are pairs with the following form:*

$$(\boldsymbol{a}, b = \langle \boldsymbol{s}, \boldsymbol{a} \rangle + e \mod q),$$

where $e \hookleftarrow \chi$ and $\boldsymbol{a}$ is sampled from $\mathbb{Z}_q^n$ uniformly at random.

In practical applications, χ is usually taken as a discrete Gaussian distribution of width αq, and the parameter $\alpha < 1$ is called "error rate". For m independent samples $(\boldsymbol{a}_1, b_1), \ldots, (\boldsymbol{a}_{\mathtt{m}}, b_{\mathtt{m}})$ from $A_{\boldsymbol{s},\alpha}$, we represent them in matrix form $(\boldsymbol{A}, \boldsymbol{b}^t) \in \mathbb{Z}_q^{m \times n} \times \mathbb{Z}_q^{m \times 1}$, where $\boldsymbol{A}^t = (\boldsymbol{a}_1^t, \ldots, \boldsymbol{a}_{\mathtt{m}}^t)$ and $\boldsymbol{b} = (b_1, \ldots, b_{\mathtt{m}})$. Given matrix $(\boldsymbol{A}, \boldsymbol{b}^t)$, LWE problem is to calculate the secret s, and Regev [15] proved the hardness of this problem as below.

THEOREM 3.1. *Let $q \leq 2^{poly(n)}$ and $m = poly(n)$. The error rate satisfies $\alpha q \geq 2\sqrt{n}$; the $LWE_{n,q,\chi,m}$ problem is at least hard as $GapSVP_\gamma$ problem on arbitrary n-dimension lattices, where $\gamma = \tilde{O}(n/\alpha)$.*

3.2.2 LWE Cryptosystems

Regev also proposed an LWE-based public-key cryptosystems in [15]. With the parameter $\alpha = \tilde{\Omega}\left(1/\sqrt{n}\right)$, the new scheme has semantic security if the $\mathrm{LWE}_{n,q,\chi,m}$ problem is hard. Suppose n is the degree of the related LWE problem, both of the secret keys and ciphertexts will have the same size: $\tilde{O}(n)$ and the size of the public key is $\tilde{O}\left(n^2\right)$. A single message bit will be encrypted each time.

For security and correct decryption, the parameters n, q, the number of samples m and the error distribution χ should satisfy various conditions. The LWE scheme $\Pi = (\mathsf{KGen}, \mathsf{Enc}, \mathsf{Dec})$ is described as follows:

$\mathsf{KGen}(1^n)$: The secret key is the secret vector $\boldsymbol{s} \in \mathbb{Z}_q^n$ in LWE problem. Take $m \approx (n+1)\log q$ samples from the LWE distribution $A_{\boldsymbol{s},\chi}$ and construct the public key $\boldsymbol{A}$ by collecting those samples as the rows of a matrix

$$\boldsymbol{A} = [\bar{\boldsymbol{A}}, \boldsymbol{b}^t] \in \mathbb{Z}_q^{m \times (n+1)}.$$

Since $\boldsymbol{b}^t = \bar{\boldsymbol{A}}\boldsymbol{s}^t + \boldsymbol{e}^t \mod q$, we have

$$\boldsymbol{A} \cdot (-\boldsymbol{s}, 1)^t = \boldsymbol{e}^t \approx \boldsymbol{0} \mod q.$$

$\mathsf{Enc}(\mu, \mathsf{pk})$: Take a bit message $\mu \in \mathbb{Z}_2$ and a public key $\boldsymbol{A}$ as input, one first calculates the sum of an arbitrary subset of the LWE samples and then encodes μ in the last coordinate. Formally, one samples a binary vector $\boldsymbol{x}$ from $\{0,1\}^m$ uniformly and outputs the ciphertext as

$$\boldsymbol{c} = \boldsymbol{x} \cdot \boldsymbol{A} + \left(\boldsymbol{0}, \mu \cdot \lfloor \frac{q}{2} \rceil\right) \in \mathbb{Z}_q^{n+1}.$$

$\mathsf{Dec}(\boldsymbol{c}, \mathsf{sk})$: Given the secret key $\boldsymbol{s}$, we can recover the message μ form $\boldsymbol{c}$ by computing

$$\begin{aligned} \boldsymbol{c} \cdot (-\boldsymbol{s}, 1)^t &\equiv \boldsymbol{x} \cdot \boldsymbol{A} \cdot (-\boldsymbol{s}, 1)^t + \mu \cdot \lfloor \frac{q}{2} \rceil \mod q \\ &= \boldsymbol{x} \cdot \boldsymbol{e} + \mu \cdot \lfloor \frac{q}{2} \rceil \mod q \\ &\approx \mu \cdot \lfloor \frac{q}{2} \rceil \mod q, \end{aligned}$$

and output 1 when the result is closer to $\lfloor \frac{q}{2} \rceil$, otherwise 0. The approximation based on the fact that the norm of vectors $\boldsymbol{e}, \boldsymbol{x} \in \mathbb{Z}^m$ is small.

Correctness: [29] The decryption is always correct when the absolute value of the inner product $\langle \boldsymbol{e}, \boldsymbol{x} \rangle \in \mathbb{Z}$ is less than $q/4$; it will hold easily when the error vector $\boldsymbol{e}$ and the parameter m are sufficiently small compared with the parameter q. Specifically, when χ is set to be a discrete Gaussian distribution $D_{\mathbb{Z},r}$ with parameter r, the inner product $\langle \boldsymbol{e}, \boldsymbol{x} \rangle$ will follow subgaussian distribution with parameter less than $r\sqrt{m}$. Section 2 indicates that the value of $\langle \boldsymbol{e}, \boldsymbol{x} \rangle$ will less than $r\sqrt{m \ln(1/\epsilon)/\pi}$ except a small probability $2\epsilon^2$. Thus, correct decryption comes from small parameters $r = \Theta(\sqrt{n})$, $q = \tilde{O}(n)$, and rate $\alpha = r/q = 1/\tilde{O}(\sqrt{n})$.

Security: Relied on the worst-case hardness of lattice problems, the decision-$\mathsf{LWE}_{n,q,\chi,m}$ is supposed to be hard for appropriate parameters. Based on those results, the LWE-based cryptosystems has semantical security. Moreover, Lindner and Peikert [30] gave a more compact LWE-based scheme in which the random vector $\boldsymbol{x}$ is taken from set $\mathbb{Z}^n$ with coordinates being drawn from the error distribution. Also, they proved that, under the same assumption, the changed scheme is semantically secure against passive eavesdroppers. In this paper, we suppose that the random vector $\boldsymbol{x}$ belongs to set $\mathbb{Z}^m$.

4. MODELING SUBVERSION OF PUBLIC KEY ENCRYPTION

In the next section, we formalize the concept of SA and the subversion model for public key encryption scheme. We first give an overview of the SA and then formally describe its two key properties, one is called `post-quantum secret undetectability` (PQSU) and the other is `arbitrary message recoverability`.

Game $\mathrm{DETECT}^{\mathcal{D}}_{\Pi,\widetilde{\Pi}}(n)$	$\widetilde{\mathsf{Enc}}(\mathsf{pk}, \mathsf{spk}, M, M')$
$(\mathsf{ssk}, \mathsf{spk}) \leftarrow\!\!\$\ \widetilde{\mathsf{KGen}}(1^n)$	$b \leftarrow\!\!\$\ \{0,1\}$
$(\mathsf{sk}, \mathsf{pk}) \leftarrow\!\!\$\ \mathsf{KGen}(1^n)$	**if** $(b = 1)$ **then**
Sample M, M' randomly	$c \leftarrow\!\!\$\ \mathsf{Enc}(M, \mathsf{pk})$
Send $(\mathsf{pk}, \mathsf{spk}, M, M')$ to $\widetilde{\mathsf{Enc}}$	**else**
$b' \leftarrow\!\!\$\ \mathcal{D}(\mathsf{spk}, \mathsf{pk}, \mathsf{sk}, c)$	$c \leftarrow\!\!\$\ \widetilde{\mathsf{Enc}}(M, M', \mathsf{spk}, \mathsf{pk})$
return b'	**return** c to $\mathcal{D}$

FIGURE 2. Game used to define detection security.

4.1 Subversion Attack (SA)

An SA on public key cryptosystems requires a public/private subversion key pair. Particularly, the public subversion key is embedded in the encryption algorithms, and the secret subversion key is hold by the attacker for recovering the underlying message. Formally, let $\Pi = (\mathsf{KGen}, \mathsf{Enc}, \mathsf{Dec})$ be a public key cryptosystems, and three algorithms $\widetilde{\mathsf{KGen}}$, $\widetilde{\mathsf{Enc}}$ and Recv are used to subversion attack it. The attack is performed as followings

Key Generation: The subversion attacker gets a subversion key pair $(\mathsf{spk}, \mathsf{ssk})$ by running the subversion key generation algorithm $\widetilde{\mathsf{KGen}}$.

Subverted Encryption: To embed a backdoor in scheme Π, the attacker replaces its encryption algorithm Enc to an algorithm $\widetilde{\mathsf{Enc}}$. $\widetilde{\mathsf{Enc}}$ takes the public key pk, the public subversion key spk, a message M and an underlying message M' as input and outputs a ciphertext c. This algorithm also ensures that with ciphertext c, the user can obtain the message M through its decryption algorithm Dec but has no idea about the underlying message M'.

Recovery: Given the ciphertext c and the subversion private key ssk, the backdoor owner is able to recover the underlying message M' via running algorithm Recv.

4.2 Post-Quantum Secret Undetectability (PQSU)

In this paper, we assume that the ordinary user knows their secret keys but has no idea about the subversion private key. Moreover, he/she can access to a quantum computer. As for detectability, a detector $\mathcal{D}$ is asked to judge whether the algorithm Enc has been replaced with $\widetilde{\mathsf{Enc}}$ from some given ciphertexts. Suppose that $\Pi = (\mathsf{KGen}, \mathsf{Enc}, \mathsf{Dec})$ is a public key encryption scheme. We consider the game given in the Fig. 2 where $\mathcal{D}$ has the quantum computation ability. In the game, $\mathcal{D}$ samples messages M and M' randomly and sends them to the challenging encryption oracle $\overline{\mathsf{Enc}}$.

Let

$$\mathbf{Adv}^{\mathtt{det}}_{\Pi,\widetilde{\Pi}}(\mathcal{D}) = 2\Pr[b = b'] - 1$$

be the advantage of $\mathcal{D}$ in detecting the subversion.

Game $\mathrm{AMR}^{\mathcal{A}}_{\Pi,\widetilde{\Pi}}(n)$	$\overline{\mathsf{Enc}}(M', \mathsf{pk}, \mathsf{spk})$:
$(\mathsf{sk}, \mathsf{pk}) \leftarrow_{\$} \mathsf{KGen}(1^n)$	$M \leftarrow_{\$} \mathcal{M}$
$(\mathsf{ssk}, \mathsf{spk}) \leftarrow_{\$} \widetilde{\mathsf{KGen}}(1^n)$	$c \leftarrow_{\$} \widetilde{\mathsf{Enc}}(M, M', \mathsf{pk}, \mathsf{spk})$
$M'' \leftarrow_{\$} \mathcal{A}^{\overline{\mathsf{Enc}}}(\mathsf{spk}, \mathsf{ssk}, \mathsf{pk})$	**return** c
return $M'' = M'$	

FIGURE 3. Game used to define arbitrary message recovery security.

DEFINITION 4.1. *(PQSU).*

The SA on public key cryptosystems scheme $\Pi = (\mathsf{KGen}, \mathsf{Enc}, \mathsf{Dec})$ *is of* ϵ*-PQSU under chosen-plaintext attacks if detection adversary with quantum computation ability satisfies the inequality below:*

$$Pr[b = b'] \leq \frac{1}{2} + \frac{\epsilon}{2},$$

that is $\mathbf{Adv}^{det}_{\Pi,\widetilde{\Pi}}(\mathcal{D}) \leq \epsilon$*. In particular, we denote that SA on* Π *is of PQSU if* ϵ *is a negligible function.*

One can note that compared with a general SA, our defined SA captures stronger undetectability. Precisely, the quantum computer would not help the adversary to detect, and nobody can break the subverted public key scheme with the help of quantum computer. Moreover, a reverse analyst manages to get the embedded subversion key spk; he is still unable to detect other subverted system embedded with the same subversion key in a black-box manner.

4.3 Arbitrary Message Recovery

In this work, it is similar to the notion of key recovery considered by Bellare *et al*. [7], we slightly generalize their notion to *arbitrary message recovery* as the strong goal of SAs. To model this notion, we define a game as depicted in the Fig. 3.

We remark that in the game the message M is sampled via running the algorithm $\mathcal{M}$ in the Fig. 3. This reflects the fact that in reality the to-be-encrypted message is independently chosen by the sender. M' is an arbitrary underlying message that the attacker wants to leak to the outside world. Given the ciphertext $c \leftarrow_{\$} \widetilde{\mathsf{Enc}}(\mathsf{pk}, \mathsf{spk}, M, M')$, the attacker wins the game if $\mathcal{A}$ recovers the underlying message M' successfully. The advantage of $\mathcal{A}$ to recover an arbitrary message is measured by

$$\mathbf{Adv}^{mr}_{\Pi,\widetilde{\Pi}}(\mathcal{A}) = \mathtt{Pr}[\mathtt{AMR}^{\mathcal{A}}_{\Pi,\widetilde{\Pi}} = 1].$$

DEFINITION 4.2. *(Arbitrary Message Recoverability).*

A SA on public key scheme $\Pi = (\mathsf{KGen}, \mathsf{Enc}, \mathsf{Dec})$ *is* $(1-\epsilon)$*-recoverable for an arbitrary underlying message if for all PPT subversion attacker adversaries satisfying*

$$\mathbf{Adv}^{mr}_{\Pi,\widetilde{\Pi}}(\mathcal{A}) \geq 1 - \epsilon.$$

In particular, we say that the SA on Π *is arbitrary message recoverable if* ϵ *is negligible.*

5. SUBVERSION ATTACK ON LWE CRYPTOSYSTEMS

5.1 SA on LWE Cryptosystems

The SA on Regev's LWE cryptosystems is presented in the Fig. 4 and it has the following properties: (i) the subverted encryption is efficient and also quantum computation resistant. (ii) It is proven post-quantum undetectable. (iii) The attack here is to recover an underlying message for any sampling algorithm.

5.1.1 *Parameters*

There are some parameters involved in our attack which are defined below:

- N, q', p are positive integers used in NTRU cryptosystems.
- Parameters n, q, m, α are involved in LWE problems.
- Polynomial ring is set to be $R = \mathbb{Z}[x]/(x^{N'} - 1)$.

Those LWE parameters satisfy $q = \tilde{O}(n)$, $m \approx (n+1)\log q$ and $\alpha = 1/\tilde{O}(\sqrt{n})$.

5.1.2 *Preliminaries*

To begin with, two maps will be given which are useful in subverted encryption and underlying message recovering algorithms. We suppose that $k = \log q$, $N = nk$ and q is a power of 2. The first one is defined by

$$\phi_{\mathtt{ext}} : \mathbb{Z}_q^n \to \mathbb{Z}_2^N$$

$$(v_0, v_1, \ldots, v_{n-1}) \to (m_0, m_1, \ldots, m_{N-1}),$$

where $(m_{ik}, m_{ik+1}, \ldots, m_{ik+k-1})$ is the 2-adic expansion of v_i for $i = 0, 1, \ldots, n-1$. The other map is described as follows:

$$\phi_{\mathtt{com}} : \mathbb{Z}_2^N \to \mathbb{Z}_q^n$$

$$(m_0, m_1, \ldots, m_{N-1}) \to (v_0, v_1, \ldots, v_{n-1}).$$

In fact, $\phi_{\mathtt{ext}}$ extends a vector in $\mathbb{Z}_q^n$ to a vector in $\mathbb{Z}_2^N$ and $\phi_{\mathtt{com}}$ compresses a vector in $\mathbb{Z}_2^N$ into a vector in $\mathbb{Z}_q^n$. Furthermore, it is worth noting that $\phi_{\mathtt{ext}}$ and $\phi_{\mathtt{com}}$ are inverse function of each other and both injective.

We propose a modified LWE encryption algorithm Enc'. The key generation algorithm KGen and the decryption algorithm

- $\overline{\mathsf{KGen}}(1^n, 1^N)$ consists of two sub-algorithms:
  - Run LWE key generation algorithm and obtain the LWE public/private key pair
  $$(\boldsymbol{A}, \boldsymbol{s}) \leftarrow_{\$} \mathsf{KGen}(1^n),$$
  where $\boldsymbol{A} = [\bar{\boldsymbol{A}}, \boldsymbol{b}^t] \in \mathbb{Z}_q^{m\times(n+1)}$ and $\boldsymbol{b}^t = \bar{\boldsymbol{A}}\boldsymbol{s}^t + \boldsymbol{e}^t \mod q$.
  - The NTRU key generation algorithm generates the subversion keys as
  $$(h, f) \leftarrow_{\$} \mathsf{KGen}_{\mathtt{ntru}}(1^N).$$
  The subversion public key is embedded in the modified encryption algorithm and f is taken by the attacker.
- $\widetilde{\mathsf{Enc}}(\mathsf{pk}, \mathsf{spk}, M, M')$: Suppose $M' \in R_2$. The subverted encryption algorithm contains three sub-algorithms, i.e., $\mathsf{Enc}_{\mathtt{ntru}}$, $\mathsf{LatticeSolve}$ and Enc' working as follows.
  - $\mathsf{Enc}_{\mathtt{ntru}}(M', \mathsf{spk})$: Given the message M' and NTRU public key h, computes the temper ciphertext $\boldsymbol{c}'$ by
  $$\boldsymbol{c}' \leftarrow_{\$} \mathsf{Enc}_{\mathtt{ntru}}(M', h).$$
  - $\mathsf{LatticeSolve}(\bar{\boldsymbol{A}}, \boldsymbol{c}')$: Input $\boldsymbol{c}'$ and matrix $\bar{\boldsymbol{A}}$, obtain $\boldsymbol{x}'$ satisfies the equation below
  $$\boldsymbol{x}'\bar{\boldsymbol{A}} \equiv \boldsymbol{c}' \mod q',$$
  where q' is a modulo parameter in NTRU cryptosystems and $\boldsymbol{x}' \in \mathbb{Z}^m$.
  - $\mathsf{Enc}'(M, \mathsf{pk}, \boldsymbol{x}')$: With the modified LWE encryption algorithm, we can get a ciphertext of the plaintext $M \in \{0, 1\}$ by
  $$\boldsymbol{c} = \mathsf{Enc}'(M, \boldsymbol{A}, \boldsymbol{x}').$$
- $\mathsf{Recv}(\boldsymbol{c}_1, \mathsf{ssk})$: Let $\boldsymbol{c}_1$ be the first N bits of $\boldsymbol{c}$. The attacker recovers the underlying message $M' \in R_2$ through the NTRU decryption algorithm.
$$M' = \mathsf{Dec}_{\mathtt{ntru}}(\boldsymbol{c}_1, f).$$

FIGURE 4. A SA on LWE Cryptosystems.

Dec are left unchanged. For a fixed vector $\boldsymbol{x}' \in \mathbb{Z}^m$, the new encryption performs as below:

$$\begin{aligned}\boldsymbol{c}' &= \mathsf{Enc}'(M, \mathsf{pk}, \boldsymbol{x}') \\ &= \boldsymbol{x}'\boldsymbol{A} + \left(\boldsymbol{0}, M \cdot \lfloor \frac{q}{2} \rceil\right) \mod q,\end{aligned}$$

where the plaintext $M \in \{0, 1\}$ and the public key $\boldsymbol{A} \in \mathbb{Z}_q^{m\times(n+1)}$. In the original encryption algorithm, vector $\boldsymbol{x}$ is randomly chosen.

5.1.3 *Attack Description*

We propose a SA on LWE encryption scheme $\Pi = (\mathsf{KGen}, \mathsf{Dec}, \mathsf{Enc})$. The procedure is depicted in Fig. 4.

In the process of the subversion encryption $\widetilde{\mathsf{Enc}}$, we can verify that

$$\begin{aligned}\boldsymbol{c} &= \boldsymbol{x}'\boldsymbol{A} + \left(\boldsymbol{0}, M \cdot \lfloor \frac{q}{2} \rceil\right) \mod q \\ &= (\boldsymbol{x}'\bar{\boldsymbol{A}}, \boldsymbol{x}' \cdot \boldsymbol{b}) + \left(\boldsymbol{0}, M \cdot \lfloor \frac{q}{2} \rceil\right) \mod q \\ &= \left(\boldsymbol{x}'\bar{\boldsymbol{A}}, \boldsymbol{x}' \cdot \boldsymbol{b} + M \cdot \lfloor \frac{q}{2} \rceil\right) \mod q.\end{aligned}$$

In fact, $\boldsymbol{x}' \in \mathbb{Z}^m$ is set to be a solution of the modular equations

$$\boldsymbol{x}'\bar{\boldsymbol{A}} \equiv \boldsymbol{c}' \mod q', \tag{1}$$

where q' is the modulo parameter in NTRU cryptosystems and $\boldsymbol{c}'$ is the output of $\mathsf{Enc}_{\mathtt{ntru}}(M', \mathsf{spk})$. It is easy to check that $\boldsymbol{c}$ can be represented as

$$\boldsymbol{c} = \left(\boldsymbol{c}', \boldsymbol{x}' \cdot \boldsymbol{b} + M \cdot \lfloor \frac{q}{2} \rceil\right).$$

Hence, $\boldsymbol{c}'$ is equal to $\boldsymbol{c}_1$ and we can recover M' through the NTRU decryption algorithm

$$M' = \mathsf{Dec}_{\mathtt{ntru}}(\boldsymbol{c}_1, f).$$

Though M is a one bit message, M' can be any message in R_2 or arbitrary vector in $\mathbb{Z}_q^n$. In the later case, the subversion attacker intends to get the vector $\boldsymbol{v}$, that is

$$\phi_{\mathtt{ext}}(\boldsymbol{v}) = \mathsf{Dec}_{\mathtt{ntru}}(\boldsymbol{c}', f).$$

$\widetilde{\mathsf{Enc}}(\mathsf{pk}, \mathsf{spk}, M_i, M')$: Let $M' \in R_2$ be the underlying message. The subverted encryption algorithm contains three sub-algorithms, i.e., $\mathsf{Enc}_{\mathtt{ntru}}$, $\mathsf{LatticeSolve}$ and Enc' working as follows.

- $\mathsf{Enc}_{\mathtt{ntru}}(M', \mathsf{spk})$: Given the message M' and NTRU public key h, calculate the temper ciphertext $\boldsymbol{c}'$ as

$$\boldsymbol{c}' \leftarrow_{\$} \mathsf{Enc}_{\mathtt{ntru}}(M', h).$$

  Suppose $N = kt$ and $k < n$, then $\boldsymbol{c}'$ can be rewrote as $\boldsymbol{c}' = (\boldsymbol{c}'_1, \ldots, \boldsymbol{c}'_t)$, with $\boldsymbol{c}'_i \in \mathbb{Z}_q^k$.
- $\mathsf{LatticeSolve}(\bar{\boldsymbol{A}}, \boldsymbol{c}'_i)$: Obtain $\boldsymbol{x}'_i \leftarrow_{\$} \mathsf{LatticeSolve}(\bar{\boldsymbol{A}}_{m\times k}, \boldsymbol{c}'_i)$ such that the equation below holds

$$\boldsymbol{x}'_i \bar{\boldsymbol{A}}_{m\times k} \equiv \boldsymbol{c}'_i \mod q,$$

  $\bar{\boldsymbol{A}}_{m\times k} \in \mathbb{Z}_q^{m\times k}$ consists by the first k column vectors of $\bar{\boldsymbol{A}}$ and $\boldsymbol{x}'_i \in \mathbb{Z}^m$.
- $\mathsf{Enc}'(M_i, \mathsf{pk}, \boldsymbol{x}'_i)$: With the modified LWE encryption algorithm, we can encrypt the plaintext $M_i \in \{0, 1\}$ by

$$\boldsymbol{c}_i \leftarrow_{\$} \mathsf{Enc}'(M_i, \boldsymbol{A}, \boldsymbol{x}'_i).$$

$\mathsf{Recv}(\boldsymbol{c}', \mathsf{ssk})$: The attacker firstly collects enough ciphertext $\boldsymbol{c}_i$ and constructs the ciphertext $\boldsymbol{c}'$. Then, the message M' can be calculated through the NTRU decryption algorithm

$$M' = \mathsf{Dec}(\boldsymbol{c}', f).$$

FIGURE 5. An improved subverted encryption algorithm

Since $\phi_{\mathtt{ext}}$ and $\phi_{\mathtt{com}}$ are inverse function of each other, the backdoor owner can recover $\boldsymbol{v}$ through $\phi_{\mathtt{com}}(\phi_{\mathtt{ext}}(\boldsymbol{v}))$.

5.1.4 *Inhomogeneous Small Integer Solution problem*

The Inhomogeneous Small Integer Solution (ISIS) problem was proposed by Gentry *et al*. in [31]. Given parameters (n, m, q, β), a random matrix $\boldsymbol{A} \in \mathbb{Z}_q^{m\times n}$ and a fix vector $\boldsymbol{y} \in \mathbb{Z}^n$, that problem is to find a solution of the equation

$$\boldsymbol{x}\boldsymbol{A} = \boldsymbol{y} \mod q,$$

where the norm of $\boldsymbol{x}$ is bounded by β. Gentry *et al*. also proved that the ISIS problem is at least hard as approximating the SIVP_γ problem with $\gamma = \beta \cdot \tilde{O}(\sqrt{n})$ when $\beta = \mathrm{poly}(n)$ and $q \geq \beta \cdot \omega(n \log n)$. In fact, finding a solution to Eq. (1) is similar to solve an ISIS problem except that we require less on the norm of $\boldsymbol{x}$. Moreover, in the following section, we use the ideal of intersection to make that process easier.

5.1.5 *Arbitrary Message Recoverability*

It is obviously that finding a solution of the modular Eq. (1) is the key point of the SA. No matter what the underlying message is, it can be recovered as long as a suitable solution $\boldsymbol{x}$ is found. That makes our new scheme arbitrary message recoverable. We will analyse those equations and propose a lattice algorithm $\mathsf{LatticeSolve}(\boldsymbol{A}, \boldsymbol{c}_1)$ in the following section.

5.2 Solving Linear Modular Equations

Finding a vector $\boldsymbol{x} \in \mathbb{Z}^m$ that satisfies the Eq. (1) will be hard when n is large and in most cases, we only need to transfer a certain part of the underlying message. Let

$$\boldsymbol{A} = \left[\boldsymbol{A}_{m\times k}, \boldsymbol{A}_{m\times(n-k)}\right] \in \mathbb{Z}_q^{m\times n}$$

and $\boldsymbol{c} = \left[\boldsymbol{c}_k, \boldsymbol{c}_{n-k}\right] \in \mathbb{Z}^n$ for an integer $k < n$. Experiment results indicate that solving the modular linear equations

$$\boldsymbol{x}\boldsymbol{A}_{m\times k} \equiv \boldsymbol{c}_k \mod q \tag{2}$$

will be much easier, where $\boldsymbol{A}_{m\times k} \in \mathbb{Z}^{m\times k}$ is a submatrix of $\boldsymbol{A}$ and $\boldsymbol{c}_k \in \mathbb{Z}_q^k$. The improved SA is presented in the Fig. 5.

To solve the linear modula the Eq. (2), we consider a lattice $\mathcal{L}'$ spanned by the $(m + k + 1) \times (m + k + 1)$ basis

$$\boldsymbol{B} = \begin{pmatrix} \boldsymbol{I}_{m\times m} & \boldsymbol{0}_{m\times 1} & \boldsymbol{A}_{m\times k} \\ \boldsymbol{0}_{1\times m} & 1 & \boldsymbol{c}_{1\times k} \\ \boldsymbol{0}_{k\times m} & \boldsymbol{0}_{k\times 1} & q\boldsymbol{I}_{k\times k} \end{pmatrix}_{(m+k+1)\times(m+k+1).}$$

Suppose $\boldsymbol{A}_{m\times k} = [\boldsymbol{a}_1, \ldots, \boldsymbol{a}_k]$, $\boldsymbol{c}_k = [c_1, \ldots, c_k]$ and $\boldsymbol{x} = [x_1, \ldots, x_k]$, the Eq. (2) can be rewritten as the following k equations:

$$\sum_{i=1}^{m} x_i a_{i,j} - c_j \equiv 0 \mod q\,(1 \leq j \leq k), \tag{3}$$

where $\boldsymbol{a}_j = [a_{1,j}, \ldots, a_{k,j}]$. Hence, we can conclude that the lattice $\mathcal{L}'$ contains the relatively short vector $\boldsymbol{v} = (\boldsymbol{x}, -1, \boldsymbol{0})$,

and lattice reduction algorithms can be used to find $\boldsymbol{v}$ and solve the linear modular equations. However, the runtime of lattice reduction algorithms is exponential with the lattice degree, he only work well when the degree is low.

5.2.1 *The Intersection Lattice Method*

In this subsection, a better method will be introduced to solve those modular equations [32–34]. Firstly, let us review the definition of intersection of lattice. Although it can be defined by more general lattices, we consider only the lattices of full rank in $\mathbb{Z}^n$ for some integer n.

Let $\mathcal{L}$ and $\mathcal{K}$ be lattices in $\mathbb{Z}^n$. Then the intersection of $\mathcal{L}$ and $\mathcal{K}$ can be given as followings

$$\mathcal{L} \cap \mathcal{K} := \{\boldsymbol{v} | \boldsymbol{v} \in \mathcal{L} \text{ and } \boldsymbol{v} \in \mathcal{K}\}.$$

From [35] the intersection of $\mathcal{L}$ and $\mathcal{K}$ can be proved a lattice of rank n in $\mathbb{Z}^n$, and it can be computed within $O(n^3)$ computation if the bit size of the elements in the lattice are ignored.

Let $\boldsymbol{a}_i$ be the i-th column vector of matrix $\boldsymbol{A}$, then we define a basic lattice for the new method. For a large constant λ with $\gcd(\lambda, q) = 1$, let $\mathcal{L}_i$ be a lattice of degree $m+2$ generated by row vectors of the following matrix:

$$\boldsymbol{M}_i = \begin{pmatrix} \boldsymbol{I}_{m\times m} & \boldsymbol{0}_{m\times 1} & \lambda \boldsymbol{a}_i \\ \boldsymbol{0}_{1\times m} & 1 & \lambda c_i \\ \boldsymbol{0}_{1\times m} & 0 & \lambda q \end{pmatrix}_{(m+2)\times(m+2).} \tag{4}$$

Since each column vector $\boldsymbol{a}_i$ is uniformly distributed in $\mathbb{Z}_q^m$, every lattices $\mathcal{L}_i = \mathcal{L}(\boldsymbol{M}_i)$ are different with overwhelming probability for all $1 \le i \le k$. With those notions, a new lattice $\mathcal{L}$ can be defined as follows:

$$\mathcal{L} = \bigcap_{i=1}^{k} \mathcal{L}_i. \tag{5}$$

It is obviously that for all $1 \le i \le k$, $\mathcal{L}_i$ contains $\boldsymbol{v} = (\boldsymbol{x}, -1, 0)$ such that

$$\boldsymbol{x}\boldsymbol{A}_{m\times k} \equiv \boldsymbol{c}_k \mod q.$$

Hence, the lattice $\mathcal{L}$ contains vector $\boldsymbol{v}$ that satisfies the Eq. (3). A new method is performed by Algorithm 1 below.

Algorithm 1: LatticeSolve$(\boldsymbol{A}, \boldsymbol{c})$

1: For $1 \le i \le$ k, construct matrix $\boldsymbol{M}_i$ from Eq. 4.
2: Set $\mathcal{L}_i = \mathcal{L}(\boldsymbol{M}_i)$.
3: Compute the new lattice $\mathcal{L}$ from Eq. 5.
4: Reduce basis matrix of the lattice $\mathcal{L}$
5: **if** a vector $\boldsymbol{v}$ belongs to the reduced matrix. **then**
6: **return** $\boldsymbol{v}$.
7: **else**
8: **return** Failure.

REMARK: As shown in [36], the vector $\boldsymbol{x}$ outputted by the reduction algorithms will become larger when m increases, even though one has used the intersection lattice technique. Moreover, if the inner product $\langle \boldsymbol{e}, \boldsymbol{x} \rangle$ is larger than $q/4$, the `LatticSolve` will be infeasible, so does our SA. The attack cannot be applied for large parameters.

5.3 Security Analysis

Since the security of NTRUEncrypt scheme is `IND$-CPA` security, our SA can be proved to be PQSU.

THEOREM 5.1. *$\Pi(\mathsf{KGen}, \mathsf{Enc}, \mathsf{Dec})$ denotes a public encryption scheme, and the subverted scheme $\Pi'(\widetilde{\mathsf{KGen}}, \widetilde{\mathsf{Enc}}, \mathsf{Dec})$ is as defined in the Fig. 4. Suppose $\mathcal{D}$ is an adversary against the PQSU of Π' that makes most k queries to its $\widetilde{\mathsf{Enc}}$ oracle. We have*

$$\mathbf{Adv}_{\Pi,\Pi'}(\mathcal{D}) \le \epsilon,$$

and ϵ is negligible.

Proof. We now give a proof of the undetectability of SA on the scheme Π using a sequence of games. In Game i, S_i denotes the event that $b = b'$.

Fix a distinguishing adversary $\mathcal{A}$, the game $G_1 - G_3$ is described in the Fig. 6.

Games G_1 and G_2 proceed identically except that the vector $\boldsymbol{c}_1$ is taken from $\mathbb{Z}_q^k$ randomly. Because of the `IND$-CPA` security of NTRU, the distribution $\boldsymbol{c}_1 \leftarrow\!\!\$\ \mathbb{Z}_q^k$ and $\boldsymbol{c}_1 \leftarrow\!\!\$\ \mathsf{Enc}_{\mathtt{ntru}}(M', \mathsf{spk})$ are statistically close. $\epsilon_1 = |\Pr[S_2] - \Pr[S_1]|$ is a negligible value.

Game G_3 is the same game as Game G_2, except that we change part of $\widetilde{\mathsf{Enc}}$. In game G_3, the vector $\boldsymbol{x}_1$ is randomly sampled from $\mathbb{Z}^m$. The definition of the algorithm implies that vector $\boldsymbol{x}$ satisfies the equation $\boldsymbol{x}\bar{\mathbf{A}} \equiv \boldsymbol{c}_1 \mod q$. Since the vector $\boldsymbol{c}_1$ is taken from $\mathbb{Z}_k$ randomly, it is hard to distinguish between $\boldsymbol{x}_1 \leftarrow\!\!\$\ \mathsf{LatticeSolve}(\bar{\mathbf{A}}, \boldsymbol{c}_1, q)$ with $\boldsymbol{x}_1 \leftarrow\!\!\$\ \mathbb{Z}_q^k$ when matrix $\bar{\mathbf{A}}$ is fixed. So, adversary $\mathcal{A}$ in game G_3 will hardly note the difference of $\boldsymbol{x}_1$ in both cases, then $|\Pr[S_3] - \Pr[S_2]| = \epsilon_2$ and ϵ_2 is negligible.

Because the vector $\boldsymbol{x}_1$ is sampled from $\mathbb{Z}^m$ randomly in Game G_3, the algorithm $\widetilde{\mathsf{Enc}}$ behaves the same as the original encryption algorithm . Therefore, the probability $\Pr[S_3] = \frac{1}{2}$ and

$$\epsilon_3 = |2\Pr[S_3] - 1| = 0.$$

Let $\epsilon = \epsilon_1 + \epsilon_2 + \epsilon_3$, we can conclude that

$$\mathbf{Adv}_{\Pi,\Pi'}(\mathcal{D}) \le \epsilon,$$

the advantage of the adversary is negligible. Because the NTRU cryptosystems and the LWE scheme are both post-quantum cryptograph, this result will also hold even though the adversary is equivalent with quantum computer. In fact, the algorithm

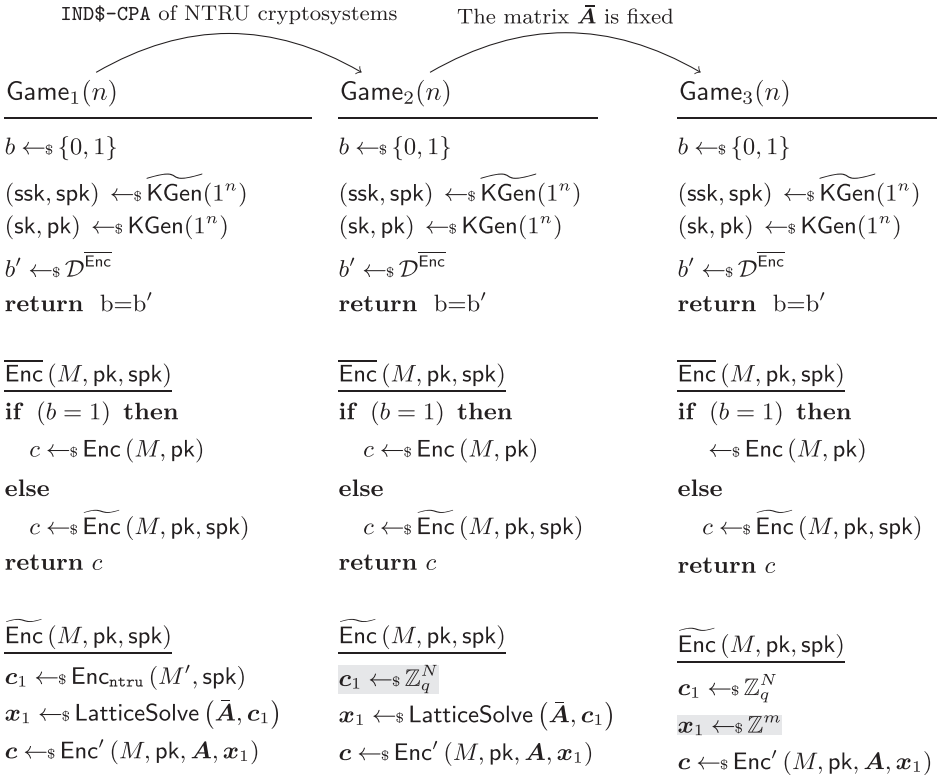

FIGURE 6. The description of game $G_1 - G_3$.

may fail to output an $\boldsymbol{x}$ with negligible probability. Here, we simply assume that will always success. ■

6. EXPERIMENTAL ANALYSIS

In our SA, the property of underlying message recoverability relies on the performance of the lattice algorithm $\mathsf{LatticeSolve}(\boldsymbol{A},\boldsymbol{c})$. The attacker can always recover the underlying message when the lattice algorithm gives right outputs. Then, we implement this algorithm on 2.1 GHz i3 Core PC and run experiments to observe its performance. The set of public LWE parameters we considered is $(n,q,m) = (29,32,145)$. This tuple satisfies the conditions discussed in Regev's LWE cryptosystems. For each instance, we choose matrix $\boldsymbol{A} \in \mathbb{Z}_q^{m\times n}$ and a vector $\boldsymbol{c}$ from $\mathbb{Z}_q^k$ randomly.

To reduce the matrix $\boldsymbol{A}$, we use the BKZ-NTL algorithm [37] of NTL package [38]. Though the BKZ 2.0 [18] algorithm works much better than that original one, the source files of BKZ 2.0 have not been opened yet. We still use the BKZ-NTL algorithm here. The BKZ reduction algorithm seems to be the best lattice reduction algorithm when applied to stages with increasing block size. In [39], Schnorr broke the `Chor-Rivest cryptosystems` successfully by using this technique. However, its runtime is exponential with the degree of the lattice and the blocksize parameter β. Blocksize is an important parameter to balance runtime and output quality. Larger β results in shorter vectors at the cost of increased runtime, and vice versa.

All the results are presented in the Fig. 7. k is the degree of $\boldsymbol{c}_k$ which runs from 2 to 10, and the blocksize β is taken from $\{2,4,6,8\}$. For each parameters set (k,β) we generate 100 random instances. The algorithm successes when it outputs a suitable $\boldsymbol{x}$. The figures indicate that the success rate r (runtime t) and the parameter k satisfy the linear equation below:

$$y = A \cdot k + B,$$

where A and B are constants and y represents success rate or runtime. The figures also show that A and B are related with the blocksize β. Through linear fitting technology, we can estimate A and B for each parameter set (k,β) more specifically. We denote that $r = A_{\mathtt{pr}} \cdot k + B_{\mathtt{pr}}$ and $t = A_{\mathtt{t}} \cdot k + B_{\mathtt{t}}$. As for different parameter set (n,β) the fitting results are listed in the Table 1.

Based on these results we come to realize that, a larger blocksize β results in higher succeed rate but more runtime. Fortunately, the algorithm will succeed with high probability when the parameter pair (k,β) is set to be $(2,6)$ or $(2,8)$. In this case, the backdoor owner can recover the underlying message easily and the SA is arbitrary message recoverable.

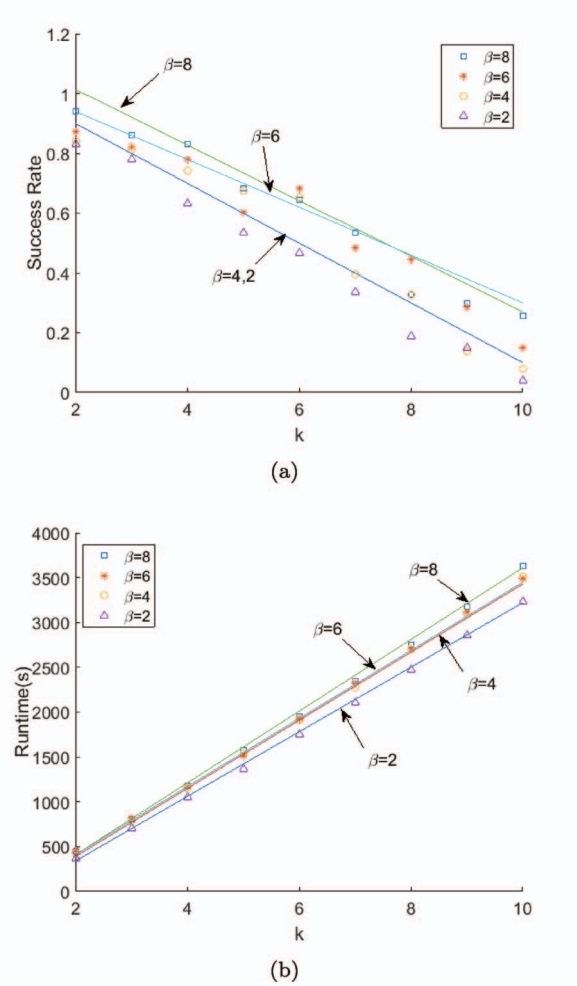

FIGURE 7. Relationship between k and y.

TABLE 1 Constant Parameters in Different Fitting.

(n, β)	Succeed Rate	Runtime
$(29, 8)$	$r = -0.093k + 1.2$	$t = 400k - 390$
$(29, 6)$	$r = -0.08k + 1.1$	$t = 380k - 350$
$(29, 4)$	$r = -0.1k + 1.1$	$t = 380k - 370$
$(29, 2)$	$r = -0.1k + 1.1$	$t = 360k - 380$

7. COUNTERMEASURE

In this part, we will discuss a potential countermeasure to defend the aforementioned SAs on LWE cryptosystems. Essentially, it is similar to exiting SAs against public key cryptosystems; our proposed SA also mainly relies on the randomness vector involved in the encryption algorithm. Therefore, existing approaches [40–42] for constructing subversion-resilient signatures could also be adopted to prevent the quantum-resistance SA in this work. Below we will introduce more details of constructing subversion-resilient scheme based on the `reverse firewalls`.

Generally speaking, the so-called cryptographic reverse firewall [40, 43, 44] is an online external party that intercepts and modifies the ciphertext produced by the encryption algorithm before it is sent out to the outside. Particularly, the ciphertext $\boldsymbol{c} \in \mathbb{Z}^{n+1}$ can be re-randomized by the following operation:

$$\boldsymbol{c}'' = \boldsymbol{c}' + (\boldsymbol{yA}, \boldsymbol{y} \cdot \boldsymbol{b}),$$

where vector $\boldsymbol{y}$ is taken form $\mathbb{Z}^m$ randomly. To make the decryption algorithm correct, the magnitude of the accumulated error $\langle \boldsymbol{e}, \boldsymbol{x} + \boldsymbol{y} \rangle$ should be less than $q/4$. Thus, we ought to choose a smaller $\boldsymbol{x}$ and $\boldsymbol{y}$ or a larger parameter q. By doing so, the attacker can hardly recover the underlying message M' from $\boldsymbol{c}''$. While the receiver can calculate the plaintext M with almost the same probability through the original decryption algorithm.

8 CONCLUSION

In this work, we explored a post-quantum SA against LWE encryption scheme. The analysis and experiments indicated that the subversion of LWE scheme is quite practical, especially when the relative random vector can be calculated in advance. Hence, we claimed that the implementations of LWE encryption scheme as black-box are potentially problematic and even insecure.

Funding

The work was supported by the National Key R&D Program of China under Grants (2017YFB0802300) the National Natural Science Foundation of China [Grant No. 61702541, 11531002, and 61872087] the Young Elite Scientists Sponsorship Program by CAST [Grant No. 2017QNRC001] the Science Research Plan Program by NUDT [Grant No. 2017QNRC001].

REFERENCES

[1] Ball, J., Borger, J. and Greenwald, G. (2013) Revealed: how US and UK spy agencies defeat internet privacy and security. *The Guardian*, 6, 1–10.

[2] Larson, N. and Shane, S. (2013) NSA able to foil basic safeguards of privacy on web. *The New York Times*, 5, 1–8.

[3] Greenwald, G. (2014) No place to hide: Edward Snowden, the NSA, and the US surveillance state. *Intell. Natl. Secur.*, 32, 868–871.

[4] Young, A. and Yung, M. (1997) Kleptography: using cryptography against cryptography. In *Proc. of EUROCRYPT 1997, Germany, 11–15 May*, pp. 62–74. Springer, Berlin.

[5] Young, A. and Yung, M. (1996) The dark side of "black-box" cryptography or: should we trust capstone? In *Proc. of CRYPTO1996, California, USA, 18–22 August*, pp. 89–103. Springer, Berlin.

[6] Bellare, M., Paterson, K.G. and Rogaway, P. (2014) Security of symmetric encryption against mass surveillance. In *Proc. of CRYPTO2014, Santa Barbara, USA, 17–21 August*, pp. 1–19. Springer, Berlin.
[7] Bellare, M., Jaeger, J. and Kane, D. (2015) Mass-surveillance without the state: strongly undetectable algorithm-substitution attacks. In *Proc. of CCS2015, Colorado, USA, 12–16 October*, pp. 1431–1440. ACM, New York.
[8] Ateniese, G., Magri, B. and Venturi, D. (2015) Subversion-resilient signature schemes. In *Proc. of CCS2015, Colorado, USA, 12–16 October*, pp. 364–375. ACM, New York.
[9] Liu, C., Chen, R., Wang, Y. and Wang, Y. (2018) Asymmetric subversion attacks on signature schemes. In *Proc. of ACISP2018, Wollongong, Australia, 11–13 July*, pp. 376–395. Springer, Berlin.
[10] Hoffstein, J., Pipher, J. and Silverman, J.H. (1998) NTRU: a ring-based public key cryptosystem. In *Proc. of ANTS1998, Oregon USA, 21–25 June*, pp. 267–288. Springer, Berlin.
[11] Albrecht, M., Bai, S. and Ducas, L. (2016) A subfield lattice attack on overstretched ntru assumptions. In *Proc. of CRYPTO2016, Santa Barbara, USA, 14–18 August*, pp. 153–178. Springer, Berlin.
[12] Stehlé, D. and Steinfeld, R. (2011) Making NTRU as secure as worst-case problems over ideal lattices. In *Proc. of EUROCRYPT2011, Tallinn, Estonia, 15–19 May*, pp. 27–47. Springer, Berlin.
[13] Yu, Y., Xu, G. and Wang, X. (2017) Provably secure NTRU instances over prime cyclotomic rings. In *Proc. of PKC2017, Amsterdam, Netherlands, 28–31 March*, pp. 409–434. Springer, Berlin.
[14] Yu, Y., Xu, G. and Wang, X. (2017) Provably secure NTRUEncrypt over more general cyclotomic rings. *IACR Cryptology ePrint Archive*, 2017, 304.
[15] Regev, O. (2009) On lattices, learning with errors, random linear codes, and cryptography. *J. ACM*, 56, 34.
[16] Howgrave-Graham, N. (2007) A hybrid lattice-reduction and meet-in-the-middle attack against NTRU. In *Proc. of CRYPTO2007, Santa Barbara, USA, 19–23 August*, pp. 150–169. Springer, Berlin.
[17] Kirchner, P. and Fouque, P.-A. (2015) An improved BKW algorithm for LWE with applications to cryptography and lattices. In *Proc. of CRYPTO2015, Santa Barbara, USA, 16–20 August*, pp. 43–62. Springer, Berlin.
[18] Chen, Y. and Nguyen, P.Q. (2011) BKZ 2.0: better lattice security estimates. In *Proc. of ASIACRYPT2011, Seoul, Korea, 4–8 December*, pp. 1–20. Springer, Berlin.
[19] Coppersmith, D. and Shamir, A. (1997) Lattice attacks on NTRU. In *Proc. of EUROCRYPT1997, Konstanz, Germany, 11–15 May*, pp. 52–61. Springer, Berlin.
[20] Alkim, E., Ducas, L., Pöppelmann, T. and Schwabe, P. (2016) Post-quantum key exchange-a new hope. In *Proc. of USENIX Security Symposium, Vancouver, 16–18 August*, pp. 327–343. USENIX Association.
[21] Kwant, R., Lange, T. and Thissen, K. (2017) Lattice Klepto. In *Proc. of SAC2017, Ottawa, CA, 16–18 August*, pp. 336–354. Springer, Berlin.
[22] Xiao, D. and Yu, Y. (2018) Klepto for ring-LWE encryption. *Comp. J.*, 61, 1228–1239.
[23] Micciancio, D. and Goldwasser, S. (2012) *Complexity of Lattice Problems: A Cryptographic Perspective*. Springer Science & Business Media, Berlin.
[24] Ajtai, M. (1998) The shortest vector problem in L2 is NP-hard for randomized reductions. In *Proc. of STOC1998, Dallas, 24–26 May*, pp. 10–19. ACM, New York.
[25] Khot, S. (2005) Hardness of approximating the shortest vector problem in lattices. *J. ACM*, 52, 789–808.
[26] Micciancio, D. and Regev, O. (2007) Worst-case to average-case reductions based on Gaussian measures. *SIAM J. Comput.*, 37, 267–302.
[27] Rogaway, P. (2004) Nonce-based symmetric encryption. In *Proc. of FSE2004, Delhi, India, 5–7 February*, pp. 348–358. Springer, Berlin.
[28] Chen, C., Hoffstein, J., Whyte2, W. and Zhang, Z. (2018) *NIST PQ Submission: NTRUEncrypt A lattice based encryption algorithm*. https://csrc.nist.gov/Projects/Post-Quantum-Cryptography/Round-1-Submissions.
[29] Peikert, C.*et al.* (2016) A decade of lattice cryptography. *Found. Trends Theor. Comput. Sci.*, 10, 283–424.
[30] Lindner, R. and Peikert, C. (2011) Better key sizes (and attacks) for LWE-based encryption. In *Proc. of CT-RSA2011, San Francisco, USA, 14–18 February*, pp. 319–339. Springer, Berlin.
[31] Gentry, C., Peikert, C. and Vaikuntanathan, V. (2008) Trapdoors for hard lattices and new cryptographic constructions. In *Proc. of STOC2008, Victoria, British, 17–20 May*, pp. 197–206. ACM, New York.
[32] Plantard, T. and Susilo, W. (2009) Broadcast attacks against lattice-based cryptosystems. In *Proc. of ACNS2009, Paris-Rocquencourt, France, 2–5 June*, pp. 456–472. Springer, Berlin.
[33] Yu, Y. and Xiao, D. (2018) Improved broadcast attacks against subset sum problems via lattice oracle. *Inform. Sci.*, 451, 210–222.
[34] Yang, Z., Fu, S., Qu, L. and Li, C. (2017) A lower dimension lattice attack on NTRU. *Sci. China Inform. Sci.*, 61, 059101.
[35] Cohen, H. (2013) *A course in computational algebraic number theory*. Springer Science & Business Media, Berlin.
[36] Gama, N. and Nguyen, P.Q. (2008) Predicting lattice reduction. In *Proc. of EUROCRYPT2008, Istanbul, Turkey, 13–17 April, pp. 31–51*. Springer, Berlin.
[37] Schnorr, C.-P. (1987) A hierarchy of polynomial time lattice basis reduction algorithms. *Theoret. Comput. Sci.*, 53, 201–224.
[38] Shoup, V.N.T.L. *A library for doing number theory*. http://www.shoup.net/ntl/26-8-2018.
[39] Schnorr, C.-P. and Hörner, H.H. (1995) Attacking the Chor-Rivest cryptosystem by improved lattice reduction. In *Proc. of EUROCRYPT1995, Saint-Malo, France, 21–25 May*, pp. 1–12. Springer, Berlin.
[40] Mironov, I. and Stephens-Davidowitz, N. (2015) Cryptographic reverse firewalls. In *Proc. of EUROCRYPT2015, Sofia, Bulgaria, 26–30 April*, pp. 657–686. Springer, Berlin.
[41] Russell, A., Tang, Q., Yung, M. and Zhou, H.-S. (2016) Destroying steganography via amalgamation: kleptographically cpa secure public key encryption. *IACR Cryptology ePrint Archive*, 2016, 530.
[42] Russell, A., Tang, Q., Yung, M. and Zhou, H.-S. (2017) Generic semantic security against a kleptographic adversary. In *Proc. of CCS2017, Dallas, USA, 30 October–03 November*, pp. 907–922. ACM, New York.

[43] Dodis, Y., Mironov, I. and Stephens-Davidowitz, N. (2016) Message transmission with reverse firewallsse-cure communication on corrupted machines. In *Proc. of CRYPTO2016, Santa Barbara, USA, 14–18 August*, pp. 341–372. Springer, Berlin.

[44] Chen, R., Mu, Y., Yang, G., Susilo, W., Guo, F. and Zhang, M. (2016) Cryptographic reverse firewall via malleable smooth projective hash functions. In *Proc. of ASIACRYPT2016, Hanoi, Vietnam, 4–8 December*, pp. 844–876. Springer, Berlin.