Efficient Post-Quantum Secure Network Coding Signatures in the Standard Model

Dong Xie[1,2], HaiPeng Peng[1,2], Lixiang Li[1,2] and Yixian Yang[1,2]
[1] Information Security Center, State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, 100876 Beijing, P.R. China
[e-mail: xdlyn@126.com]
[2]National Engineering Laboratory for Disaster Backup and Recovery, Beijing University of Posts and Telecommunications, 100876 Beijing, P.R. China
[e-mail: penghaipeng@bupt.edu.cn]
*Corresponding author: Haipeng Peng

Received October 11, 2015; revised December 16, 2015; revised January 26, 2016; accepted March 30, 2016; published May 31, 2016

Abstract

In contrast to traditional "store-and-forward" routing mechanisms, network coding offers an elegant solution for achieving maximum network throughput. The core idea is that intermediate network nodes linearly combine received data packets so that the destination nodes can decode original files from some authenticated packets. Although network coding has many advantages, especially in wireless sensor network and peer-to-peer network, the encoding mechanism of intermediate nodes also results in some additional security issues. For a powerful adversary who can control arbitrary number of malicious network nodes and can eavesdrop on the entire network, cryptographic signature schemes provide undeniable authentication mechanisms for network nodes. However, with the development of quantum technologies, some existing network coding signature schemes based on some traditional number-theoretic primitives vulnerable to quantum cryptanalysis. In this paper we first present an efficient network coding signature scheme in the standard model using lattice theory, which can be viewed as the most promising tool for designing post-quantum cryptographic protocols. In the security proof, we propose a new method for generating a random lattice and the corresponding trapdoor, which may be used in other cryptographic protocols. Our scheme has many advantages, such as supporting multi-source networks, low computational complexity and low communication overhead.

Keywords: Network coding, signature scheme, lattice, post-quantum cryptography.

This research was supported by the Asia Foresight Program under NSFC Grant (Grant No. 61411146001), the Beijing Natural Science Foundation (Grant No. 4142016), Scientific Research Project of Beijing Municipal Commission of Education (Grant Nos. KZ20150015015, KM201510015009), and the National Natural Science Foundation of China (Grant Nos. 61573067, 61472045).

1. Introduction

The traditional manner of data transmitting in communication network is "store-and-forward", i.e., every intermediate node plays the role of transponder and do not do any calculations for the received data packets. In order to guarantee correct transmission of packets, every intermediate node need to verify the integrity of the received packet before forwarding it. However, this multicast routing pattern does not achieve the optimal throughput. In 2000, Ahlswede et.al. [1] proposed a brand-new method for achieving the optimal throughput in computer networks, called network coding. In a nutshell, network coding is a message-switching technique which simultaneously has the function of encoding and routing. The core idea is to allow intermediate nodes to perform a linear or nonlinear operation on some received packets, and then forward it to the downstream nodes. In addition to increase network throughput, network coding also improves the robustness of the network. That is to say, even if a large fraction of packets are discard in transmission, the destination nodes still can accurately recover the original files once it has received sufficiently many correct packets.

Although network coding has so many advantages, it is extremely vulnerable to pollution attacks by even a single malicious intermediate node. If there is no verification procedure for packets, the malicious node can modify received messages and forward them to its downstream nodes. It can cause pollution diffusion to the whole network and thus the destination nodes can not recover the original files. Up to now, all existing method for solving this security issue can be categorized into two types: information-theoretic solution [2][3][4] and cryptographic solution [5][6]. Information-theoretic solutions add some redundant information into an original file and the destination nodes can accurately reconstruct the file only when the proportion of modified messages to the whole file is sufficiently low. This restriction implies that information-theoretic solutions is only suitable for a relatively weak class of adversaries. Cryptographic solutions are that the source nodes use some cryptographic protocols, such as signature schemes and Message Authentication Codes (MACs), to generate some additional verification information which are transmitted together with original packets. However, standard signature schemes or MACs can not be directly applied to network coding settings. Because the intermediate nodes is not only a simple "store-and-forward", but also need to perform operations on some received packets. Thus, the intermediate nodes can not produce a valid verification information on a combined packets if we use the standard cryptographic primitives mentioned above. Since the related homomorphic schemes have the property that any intermediate node can produce a valid verification information for the combined packet without knowing the private key, they are suitable for solving security issues in network coding no matter what kind of adversaries it face.

1.1 Urgent Demand of Post-Quantum Cryptographic Schemes

There exist some network coding signature schemes based on some traditional number-theoretic primitives, such as the discrete logarithm problem and the integer factorization problem. However, all of these traditional hard problems can be easily solved on a sufficiently large quantum computer running Shor's algorithm [7][8]. Though current existing quantum computers are too small to attack some real cryptographic schemes, we must to research post-quantum cryptographic schemes before this disaster comes. So in recent ten years, post-quantum cryptography becomes a hot research topic in the field of network security. As one of the most promising candidates, lattice-based cryptography, has many advantages described below:

• **Resistance to quantum attacks:** Unlike more widely used and known public key cryptography such as the RSA or Diffie-Hellman cryptosystems which are easily attacked by a quantum computer, some lattice-based cryptosystems appear to be resistant to attack by both classical and quantum computers.
• **High efficiency:** The operations of lattice-based cryptographic schemes can be extremely efficient and conceptually simple. It usually requires only linear matrix/vector arithmetic operations modulo some small primes. By contrast, the analogous operations in traditional number-theoretic cryptosystems are much more complex.
• **Resistance to unforeseen structural attacks:** Lattice-based cryptography connects the average-case complexity of hard lattice problems to their complexity in the worst-case [9] [10], which provide strong theoretical evidence that their random instances are indeed asymptotically hard. Notice that random instances of some number-theoretic hard problems may suffer from some unforeseen structural attacks.

1.2 Related Work

From cryptographic perspective, existing solutions for resisting malicious attacks in network coding systems can be divided into two categories: Homomorphic Hashes and Homomorphic Signatures. In the scheme of [11] or [12], the sender computers a homomorphic hash of each block of the file, and any intermediated node can verify whether the received packet is the linear combination of the original files. The drawback of this method is that both the public keys and the authentication information are large. Agrawal and Boneh [13] designed a keyed homomorphic hash function, i.e., homomorphic MAC, which can be used to mitigate pollution attacks. Esfahani et al. [14] proposed a dual-homomorphic MAC for NC-enabled wireless sensor networks. Wang [15] pointed out some drawbacks of some existing cryptographic protocols and proposed a secure an efficient homomorphic authentication scheme for secure network coding. Chen et al. [16] proposed an efficient symmetric key based authentication scheme for P2P live streaming system with network coding. In addition, there are some other related studies [17][18][19] about that topic. In the aspect of homomorphic signatures, Charles et al. [20] introduced a homomorphic signature scheme for network coding system based on the discrete logarithm problem. The drawback of this scheme is that public/private key pairs must be updated when the next file is transmitted. Zhao et al. [21] presented a scheme that the sender computed a authentication information for a file, but their scheme only handle a single file every time and both the authentication information and public keys are large. The schemes in [20][21] cannot prevent inter-session pollution efficiently. Boneh et al. [5] proposed a homomorphic signature scheme in the random model that can be viewed as authenticating linear subspaces. The public key of their scheme has constant size and their construction can directly take into account the distribution of multiple files using a single public key (in contrast to [20][21]). Subsequently, Agrawal et al. [22] mainly focused on the integrity of packets in the setting of multi-source network coding, and presented a generic construction for this setting. Catalano et al. [23] introduced two homomorphic network coding signatures with security proofs in the standard model. Their schemes achieve communication and computational efficiency comparable to those of the random oracle implementation [5] and outperform the two related constructions [24][25]. Liu and Wang [26] proposed a novel homomorphic signature scheme using a dynamic public key technique, which not only can resist intra-generation pollution attacks but also can prevent inter-generation pollution attacks. Chen et al. [27] introduced an improved homomorphic signature scheme. Zhang et al. [31] presented a hybrid-key cryptographic method and the source node produce a number of MACs and a signature for every transmitted message. There are also some other related research

results about this direction, e.g., [28][29][30].

However, almost all above-mentioned linear homomorphic signature schemes for network coding system (also known as network coding signature schemes) are based on traditional number-theoretic primitives (e.g., the discrete logarithm problem and the integer factorization problem) . These constructions are vulnerable to quantum cryptanalysis [7][8]. As one of the best promising tools of designing cryptographic protocols resistance to quantum attack, lattice theory, has many advantages just as described in Section 1.1. Up to now, there are several network coding signature schemes based on lattice. Boneh and Freeman [32] presented a network coding signature scheme in the random model based on the Small Integer Solution (SIS) problem. Different from previous number-theoretic schemes whose linear combination coefficients are chosen from relatively large finite field, their scheme is the first one that authenticates vectors defined over binary field. Wang et al. [33] designed an efficient lattice-based network coding signature scheme in the random model, in which both the public key size and the signature size are shorter than those in [32]. Unfortunately, these two constructions only support the case of single source network coding systems. Through the improvement of [32], Zhang et al. [34] proposed a scheme which can support multi-source network coding settings. Recently, Jing [35] improved the scheme [32] and constructed an efficient network coding signature scheme in the random oracle for the multi-source case.

1.3 Our Contribution

However, we note that the lattice-based network coding signature schemes mentioned above are all in the random oracle. In this paper we propose an efficient post-quantum secure network coding signature scheme in the standard model. Specifically, our contributions mainly consist of the following aspects:

(1) We present a lattice-based network coding signature scheme over binary field in the standard model. Although some signature and encryption schemes can be proved secure in the random oracle model, it is not enough to cover a practical implementation [36]. The problem with random oracle model is that it turns out to be very difficult to build a really "random" oracle. And any implementation of the random oracle may results in insecure schemes.

(2) We give a new sampling method for generating a random lattice and the corresponding short basis. The algorithm takes a random matrix B as input, and outputs a matrix $\mathbf{C}$ (with small norm), a random lattice $\mathbf{\Lambda}^{\perp}(\mathbf{A})$ and its short basis $\mathbf{T}_{\mathbf{A}}$ such that $\mathbf{A}=\mathbf{BC}$. As a matter of independent interest, the proposed sampling algorithm may be useful in many other lattice-based cryptographic constructions.

(3) In general, almost all existing lattice-based cryptosystems are time-consuming. Compared to schemes based on traditional number-theoretic primitives, our scheme has low computational complexity (just need linear matrix/vector arithmetic operations) when it generates the combined signature at each intermediate node. In addition, our scheme has low communication overhead compared to existing lattice-based network coding signature schemes.

1.4 Organization

Section 2 recalls some basic background knowledge, including the fundamentals of network coding, lattice theory and the formal definition of network coding signature scheme. Section 3 presents a new sampling algorithm used in our security proof, which may also be applied in other cryptographic protocols. Section 4 describes our scheme in detail and proves some

important properties, including correctness, unforgeability, and privacy. Section 5 analyzes the efficiency of our scheme. Finally, we conclude this paper in Section 6.

2. Preliminaries

For any positive integer N, $[N]$ denotes the set $\{1,2,\cdots,N\}$. Let F_q denote the finite field of order q. Vectors are assumed to be in column form and are written using bold lower-case letters (e.g. $\mathbf{x}$). Similarly, we use bold capital-case letters (e.g. $\mathbf{A}$) to represent matrices. Given two matrices $\mathbf{A}_1 \in F_q^{n\times m_1}$ and $\mathbf{A}_2 \in F_q^{n\times m_2}$, we use $[\mathbf{A}_1\|\mathbf{A}_2]$ to denote the $n\times(m_1+m_2)$ matrix formed by concatenating $\mathbf{A}_1$ and $\mathbf{A}_2$. For a matrix $\mathbf{A} \in F_q^{n\times m}$, we use $\|\mathbf{A}\|$ to denote the maximum norm of column vector $\mathbf{a}_i$ of the matrix, i.e., $\|\mathbf{A}\| = \max_{i\in[m]}\{\|\mathbf{a}_i\|\}$. If the column vectors of $\mathbf{A}$ are linearly independent, let $\overline{\mathbf{A}} = \{\bar{\mathbf{a}}_1,\cdots,\bar{\mathbf{a}}_m\}$ denote the Gram-Schmidt orthogonalization of vectors $\mathbf{a}_1,\cdots,\mathbf{a}_m$ taken in that order.

2.1 Network Coding

The concept of network coding was initially proposed by Ahlswede et al. [1]. Without loss of generality, we just recall the fundamentals of the single source network coding here [5] [22]. A file is represented by an ordered sequence of r-dimensional vectors $\hat{\mathbf{v}}_1,\hat{\mathbf{v}}_2,\cdots,\hat{\mathbf{v}}_k \in F_q^r$, where $\hat{\mathbf{v}}_i$ is a block of the file and q is a prime. Prior to transmission, the source node **S** creates the augmented vectors $\mathbf{v}_1,\mathbf{v}_2,\cdots,\mathbf{v}_k$ given by:

$$\mathbf{v}_i = (\hat{\mathbf{v}}_i, \underbrace{0,\cdots,0,1}_{i},\overbrace{0,\cdots,0}^{k}) \in F_q^{k+r}.$$

Namely, the augmented vector $\mathbf{v}_i$ is formed by appending a vector which is the i column of k-dimensional identity matrix. Then **S** sends $\{\mathbf{v}_i\}_{i\in[k]}$ to some intermediate nodes.

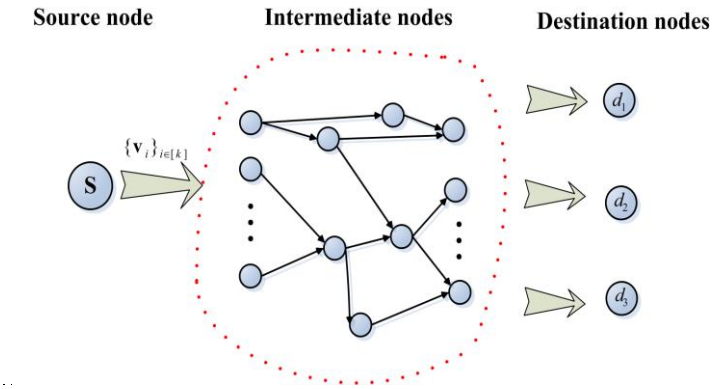


Fig. 1. The common topology of a single source multicast network system.

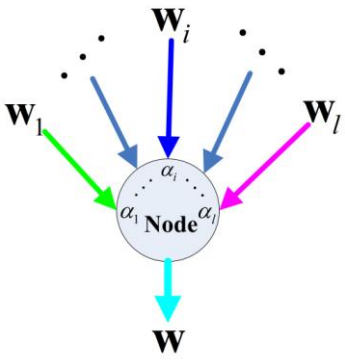


Fig. 2. The encoding process of each intermediate node.

Fig. 2 shows the encoding process of each intermediate node when it receives l packets. It first chooses randomly l weight coefficients $\alpha_i \in \{0,1\}$ and computes the output $\mathbf{w} = \sum_{i=1}^{l} \alpha_i \mathbf{w}_i$. When any destination node receives k linearly independent vectors $\mathbf{w}_1, \mathbf{w}_2, \cdots, \mathbf{w}_k$, it can accurately recover the original file using Gaussian elimination. Specifically, let α_{ij} be the j-th weight coefficient of node i. It is easy to see that each packet transmitted in the network can be viewed as a linear combination of the augmented vector $\mathbf{v}_i$. I.e.,

$$\mathbf{w}_i = \sum_{j=1}^{k} \alpha_{ij} \mathbf{v}_j = \begin{pmatrix} \sum_{j=1}^{k} \alpha_{ij} \hat{\mathbf{v}}_j \\ \alpha_{i1} \\ \alpha_{i2} \\ \vdots \\ \alpha_{ik} \end{pmatrix}.$$

Let $\mathbf{w}_i^{a} = \begin{pmatrix} \mathbf{w}_i^1 \\ \vdots \\ \mathbf{w}_i^{r-1} \\ \mathbf{w}_i^r \end{pmatrix} \in F_q^{\,r}$ and $\mathbf{w}_i^b = \begin{pmatrix} \mathbf{w}_i^{r+1} \\ \vdots \\ \mathbf{w}_i^{r+k-1} \\ \mathbf{w}_i^{r+k} \end{pmatrix} \in F_q^k$. We have $\mathbf{w}_i^a = (\hat{\mathbf{v}}_1, \hat{\mathbf{v}}_2, \cdots, \hat{\mathbf{v}}_k)\mathbf{w}_i^b$. So,

$$\left(\mathbf{w}_1^a, \mathbf{w}_2^a, \cdots, \mathbf{w}_k^a\right) = \left(\hat{\mathbf{v}}_1, \hat{\mathbf{v}}_2, \cdots, \hat{\mathbf{v}}_k\right) \cdot \left(\mathbf{w}_1^b, \mathbf{w}_2^b, \cdots, \mathbf{w}_k^b\right).$$

Thus, the destination nodes can recover the original file through the above linear equation. We stress that the dimension of augmented vector should be small. Because the augmented data increases communication overhead and the destination nodes just need a small number of packets to recover the original files.

2.2 Lattice and Hard Assumption

Informally, a m-dimensional lattice is a set of points in R^m with a periodic structure. It also can be viewed as a algebraic additive subgroup of R^m. Let $\mathbf{B} = \{\mathbf{b}_1, \mathbf{b}_2, \cdots, \mathbf{b}_n\}$ be a set of n linearly independent vectors in R^m. The lattice $\Lambda(\mathbf{B})$ is the set of all integer linear combinations of these vectors, i.e., $\Lambda(\mathbf{B}) = \{\sum_{i=1}^{n} x_i \mathbf{b}_i \,|\, x_i \in Z\}$. We say that $\mathbf{B}$ is a basis for $\Lambda(\mathbf{B})$, and the positive integers n and m are the rank and dimension of the lattice respectively. If $m = n$, we say the lattice is full-rank. In the lattice-based cryptography, we always focus on the integer lattice where the lattice points are contained in Z^m. For any positive integers n, $m(\geq n)$ and $q \geq 2$, let $\mathbf{A} \in Z_q^{n \times m}$ be a matrix. The two kinds of random lattice related to $\mathbf{A}$ are defined as follows:

$$\Lambda^{\perp}(\mathbf{A}) = \{\mathbf{e} \in Z_q^m : \mathbf{A} \cdot \mathbf{e} = \mathbf{0} \bmod q\};$$

$$\Lambda_{\mathbf{u}}^{\perp}(\mathbf{A}) = \{\mathbf{e} \in Z_q^m : \mathbf{A} \cdot \mathbf{e} = \mathbf{u} \bmod q\}.$$

In fact, the lattice $\Lambda_{\mathbf{u}}^{\perp}(\mathbf{A})$ is a coset of $\Lambda^{\perp}(\mathbf{A})$. That is to say, $\Lambda_{\mathbf{u}}^{\perp}(\mathbf{A}) = \Lambda^{\perp}(\mathbf{A}) + \mathbf{t}$, where the vector $\mathbf{t}$ satisfies that $\mathbf{A} \cdot \mathbf{t} = \mathbf{u} \bmod q$.

For any real $s>0$ and vector $\mathbf{c}\in R^n$, the n-dimensional Gaussian function $\rho_{s,\mathbf{c}}(\mathbf{x})$ on R^n centered at $\mathbf{c}$ with parameter s is defined as $\rho_{s,\mathbf{c}}(\mathbf{x})=\exp(-\pi\|\mathbf{x}-\mathbf{c}\|^2/s^2)$, where $\mathbf{x}$ is an n-dimensional vector in R^n. For a n-dimensional lattice Λ, the discrete Gaussian distribution is defined as $D_{\Lambda,s,\mathbf{c}}(\mathbf{x})=\rho_{s,\mathbf{c}}(\mathbf{x})/\rho_{s,\mathbf{c}}(\Lambda)$, where $\mathbf{x}$ is a vector in Λ. We omit $\mathbf{c}$ and s when they are taken to $\mathbf{0}$ and 1, respectively. For a positive real $\varepsilon>0$, the smoothing parameter $\eta_\varepsilon(\Lambda)$ is the smallest real s such that $\rho_{1/s}(\Lambda^*\setminus\{0\})\le\varepsilon$, where Λ^* is the dual lattice of Λ, defined by $\Lambda^*=\{\mathbf{z}\in R^m\,|\,\forall\mathbf{y}\in\Lambda,\langle\mathbf{z},\mathbf{y}\rangle\in Z\}$.

Some advanced lattice-based cryptographic constructions require generating a matrix $\mathbf{A}$ (statistically close to uniform distribution) together with a short basis of the lattice $\Lambda^\perp(\mathbf{A})$. Next, we state some important lemmas that will be used in our paper.

Lemma 1([38]). There exists a Probabilistic Polynomial Time (PPT) algorithm **TrapGen** $(1^n,1^m,q)$ that, on input positive integers n, q, and $m\ge 6n\log q$, outputs a matrix $\mathbf{A}\in Z_q^{n\times m}$ statistically close to uniform over $Z_q^{n\times m}$ and a basis $\mathbf{T}\in Z^{m\times m}$ of the lattice $\Lambda^\perp(\mathbf{A})$ such that $\|\overline{\mathbf{T}}\|\le O(\sqrt{n\log q})$ with overwhelming probability.

Lemma 2([38]). Given a basis $\mathbf{T}$ of m-dimension lattice Λ, a parameter $s>\|\overline{\mathbf{T}}\|\cdot w(\sqrt{\log n})$, and a vector $\mathbf{c}\in R^m$, there is a PPT algorithm that output a sample from a distribution that is statistically close to $D_{\Lambda,s,\mathbf{c}}$.

Lemma 3([38]). Let n, $q\ge 2$, $m>2n\log q$ be three positive integers. For a random matrix $\mathbf{A}\in Z_q^{n\times m}$, let $\mathbf{T}$ be a basis of $\Lambda^\perp(\mathbf{A})$ and $s>\|\overline{\mathbf{T}}\|\cdot w(\sqrt{\log n})$. Then,

(1) Given a vector $\mathbf{v}\in Z^n$, there is a PPT algorithm **SamplePre**(**A, T,** s**,** **v**) that outputs a sample **u** from a distribution that is statistically close to $D_{\Lambda_{\mathbf{v}}^\perp,s}$ and **u** satisfies $\|\mathbf{u}\|\le s\sqrt{m}$ with overwhelming probability.

(2) For any $\mathbf{t}\leftarrow D_{Z^m,s}$, the distribution of syndrome $\mathbf{u}=\mathbf{A}\cdot\mathbf{t}\bmod q$ is statistically close to uniform over Z_q^n.

Lemma 4([35] [39]). For an arbitrary basis $\mathbf{T}\in Z^{m\times m}$ of the lattice $\Lambda^\perp(\mathbf{A})$ about a random matrix $\mathbf{A}\in Z_q^{n\times m}$, an additional matrix $\mathbf{A}'\in Z_q^{n\times m'}$ and the parameter $s>\|\overline{\mathbf{T}}\|\cdot w(\sqrt{\log n})$,then

(1) There is a deterministic polynomial time algorithm **ExtBasis**$(\mathbf{T},\mathbf{B}=\mathbf{A}\|\mathbf{A}')$ that outputs a new short basis $\mathbf{T}'$ of the lattice $\Lambda^\perp(\mathbf{B})$ such that $\|\overline{\mathbf{T}}\|=\|\overline{\mathbf{T}'}\|$.

(2) There is a PPT algorithm **RandBasis**$(\mathbf{T},s)$ that outputs another short basis $\mathbf{T}'$ of the lattice $\Lambda^\perp(\mathbf{A})$, which is independent of the original basis $\mathbf{T}$ and is still short.

Lemma 5([32]). Let $\Lambda\subseteq Z^m$ be a lattice and $s\in R$ be a parameter . For $i=1,2,\cdots,k$, let $\mathbf{t}_i\in Z^m$ and let X_i be mutually independent random variables sampled from $D_{\Lambda+\mathbf{t}_i,s}$. Let $\mathbf{c}=(c_1,\cdots,c_k)\in Z^k$, and define $g=\gcd(c_1,\cdots,c_k)$, $\mathbf{t}=\sum_{i=1}^k c_i\mathbf{t}_i$. If $s>\|\mathbf{c}\|\cdot\eta_\varepsilon(\Lambda)$ for

some negligible ε, then $Z=\sum_{i=1}^{k} c_i X_i$ is statistically close to $D_{g\Lambda+\mathbf{t},\|\mathbf{c}\|s}$.

Similar to existing lattice-based network coding signature schemes, the security of our scheme is also based on the problem of finding short vectors in $\Lambda^{\perp}(\mathbf{A})$ for a random matrix **A**. This is known as the Small Integer Solution (SIS) problem, and is defined as follows.

Definition 1([32][33]). Given positive integers n, m, q, a real constant β and a random matrix $\mathbf{A}\in Z_q^{n\times m}(m\geq n)$, the $\mathrm{SIS}_{q,m,\beta}$ problem is find a nonzero vector $\mathbf{u}\in Z^m$ such that $\mathbf{A}\cdot\mathbf{u}=0 \bmod q$ and $\|\mathbf{u}\|\leq\beta$.

2.3 Network Coding Signature Scheme

In this subsection we first describe the formal definition of general network coding signature scheme, and then provide two security games related to unforgeability and privacy. For the sake of convenience, every original file represented by a set of block vectors is associated with an identifier [32] [35]. Here we state that the intermediate nodes in the network combine the block vectors tagged the same identifier. We adapt the model of [32] and consider the multi-source case. Throughout this paper, let n be the security parameter and $L\geq 1$ be the maximum number of linear combinations that can be authenticated. There is a trusted Private Key Generator (PKG), which can distribute public/private key pairs for source nodes. A network coding signature scheme is a tuple of polynomial time algorithms Π = (**KeyGen**, **Sign**, **Com**, **Verify**) with the following syntax.

•**KeyGen**(1^n, L). This PPT algorithm takes the security parameter n and L as inputs. It outputs a public/private key pair (pk_i, sk_i) for the source node i. (This is run by the PKG.)

•**Sign**(id, v_i, sk_i) For the i-th source node, this PPT algorithm takes as input the identifier $id\in\{0,1\}^n$, a message v_i and the secret key sk_i, and outputs a signature δ_i. (This is run by the source nodes.)

•**Com**($\{pk_i\}_{i=1}^{L}$, id, $\{\alpha_i\}_{i=1}^{l}$, $\{v_i,\delta_i\}_{i=1}^{l}$) This PPT algorithm takes as input public keys of all source nodes, an identifier id, $\{\alpha_i\}_{i=1}^{l}\in\{0,1\}^l$ and $l(\leq L)$ message-signature pairs $\{v_i,\delta_i\}_{i=1}^{l}$. It outputs a combined signature δ on the combined message $\sum_{i=1}^{l}\alpha_i v_i$. (This is run by the intermediate nodes.)

•**Verify**($\{pk_i\}_{i=1}^{L}$, id, v, δ) This deterministic algorithm takes as input the public keys of all source nodes, an identifier id, a message v and a signature δ, and outputs either 0 (reject) or 1 (accept). (This is run by the intermediate nodes and destination nodes.)

For correctness, we require that both the original signatures (generated by **Sign**) and the combined signatures (generated by **Com**) are accepted. Specifically, we require that the following two conditions hold:

1. For all id and v_i, if $\delta_i\leftarrow$ **Sign**(id, v_i, sk_i) then **Verify**($\{pk_i\}_{i=1}^{L}$, id, v_i, δ_i) =1.

2. For all id and all sets of triples $\{\alpha_i,v_i,\delta_i\}_{i=1}^{l}$, if it holds that **Verify**($\{pk_i\}_{i=1}^{L}$, id, v_i, δ_i)= 1 for all i, then **Verify**($\{pk_i\}_{i=1}^{L}$, id, $\sum_{i=1}^{l}\alpha_i v_i$, **Com**($\{pk_i\}_{i=1}^{L}$, id, $\{\alpha_i\}_{i=1}^{l}$, $\{v_i,\delta_i\}_{i=1}^{l}$)) =1.

For the security of network coding signature scheme, we also consider the properties of unforgeability and privacy of combined signature [32][35]. For the unforgeability, the security model allows an adversary to make adaptive signature queries on files what he can choose

arbitrarily, but he must query all the blocks in a file at once. Formally, we give the definition of existential unforgeability of network coding signatures under chosen file attacks.

Definition 2([5][32]). A network coding signature scheme Π = (**KeyGen**, **Sign**, **Com**, **Verify**) is *unforgeable* if the advantage of PPT adversary in the following security game is negligible in the security parameter n:

• The challenger runs **KeyGen**$(1^n, L)$ to get (pk_i, sk_i), and gives pk_i to the adversary.

• Proceeding adaptively, the adversary specifies a sequence of signature queries on files represented by $V_i = \{v_{i1}, v_{i2}, \cdots, v_{ik}\}$. For each file V_i, the challenger choose id_i uniformly from $\{0,1\}^n$ and give the adversary the identifier id_i and the j-th signature $\delta_{ij} \leftarrow$ **Sign**(id, v_{ij}, sk_i) for $j = 1,2,\cdots,k$.

• The adversary outputs an identifier id^*, a new message v^*, and a signature δ^*.

If **Verify**($\{pk\}_{i=1}^{L}$, id^*, v^*, δ^*) =1, then the adversary wins the game. In fact, there are two types of forgers: One is $id^* \neq id_i$ for all queried i, and the other is $id^* = id_i$ for some index i but v^* is not a linear combination of message blocks $v_{i1}, v_{i2}, \cdots, v_{ik}$.

The definition of privacy for network coding signatures captures the idea that given signatures on a number of combined messages in one of two different files, the adversary cannot tell which file the combined signatures came from even the adversary knows the secret keys. This property was called *weakly context hiding*, which is introduced in [32] in the case of single source. Next we give the formal definition for the case of multi-source settings.

Definition 3([32]). A network coding signature scheme Π = (**KeyGen**, **Sign**, **Com**, **Verify**) is *weakly context hiding* if the advantage of any PPT adversary in the following security game is negligible in the security parameter n:

• The challenger runs **KeyGen**$(1^n, L)$ to get (pk_i, sk_i), and gives pk_i and sk_i to the adversary.

• The adversary outputs $(V_0, V_1, f_1, f_2, \cdots, f_k)$ where V_b is represented by a vector set $\{v_1^{(b)}, v_2^{(b)}, \cdots, v_k^{(b)}\}$ for b =0,1. The functions $f_1, f_2, \cdots, f_k$ are satisfying $f_i(v_1^{(0)}, \cdots, v_k^{(0)}) = f_i(v_1^{(1)}, \cdots, v_k^{(1)})$ for $i = 1, \cdots, k$. In response, the challenger generates a random bit $b \in \{0,1\}$, a random identifier $id \in \{0,1\}^n$ and signs $v_i^{(b)}$ using the corresponding private sk_i. Subsequently, the challenger uses **Com** to derive signatures δ_i on $f_i(v_1^b, \cdots, v_k^b)$ and sends $\delta_1, \delta_2, \cdots, \delta_k$ to the adversary. The functions $f_1, f_2, \cdots, f_k$ can be output adaptively after V_0, V_1 are output.

• The adversary outputs a bit b'.

If $b = b'$, the adversary wins the game. The advantage of the adversary is defined as the probability that the adversary wins the game.

3. A New Trapdoor Sampling Algorithm

We present a new trapdoor sampling algorithm in this section. Our constructing method is very similar to that of **SuperSamp** in [40], which samples a random superlattice with a short basis. In our construction, the algorithm takes a random matrix $\mathbf{B} \in Z_q^{n \times m}$ as input, and outputs

a matrix $\mathbf{C} \in Z_q^{m\times m}$, a matrix $\mathbf{A} \in Z_q^{n\times m}$ and a short basis $\mathbf{T} \in Z^{m\times m}$ of the lattice $\Lambda^{\perp}(\mathbf{A})$, where $\mathbf{A} = \mathbf{BC} \bmod q$ and $\|\mathbf{C}\| \le \sqrt{nq}$. Next, we state our theorem.

Theorem 1. There is a PPT algorithm **ProductSamp** that on input 1^n, 1^m, $q \ge 2$ with $m = O(n\log q)$, and a random matrix $\mathbf{B} \in Z_q^{n\times m}$, outputs a pair $(\mathbf{A}, \mathbf{T}) \in Z_q^{n\times m} \times Z^{m\times m}$ and $\mathbf{C} \in Z_q^{m\times m}$ such that (1) $\mathbf{A} = \mathbf{BC} \bmod q$; (2) $\mathbf{A}$ is statistically close to uniform over $Z_q^{n\times m}$; (3) $\mathbf{T}$ is a short basis of the lattice $\Lambda^{\perp}(\mathbf{A})$; (4) $\|\widetilde{\mathbf{T}}\| \le O(\sqrt{n\log q})$; (5) $\|\mathbf{C}\| \le \sqrt{nq}$.

Proof. Let $\mathbf{B} = [\mathbf{B}_1 \| \mathbf{B}_2]$, where $\mathbf{B}_1 \in Z_q^{n\times(m-n)}$ and $\mathbf{B}_2 \in Z_q^{n\times n}$. Without loss of generality, we assume that $\mathbf{B}_2$ is invertible. In fact, such decomposition can be found with overwhelming probability if we permute the columns of $\mathbf{B}$. Algorithm **ProductSamp** works as follows:

(1) Let $\mathbf{B}_1 = [\mathbf{B}_{11} \| \mathbf{B}_{12}]$, where $\mathbf{B}_{11}$ is the first square matrix of $\mathbf{B}_1$ and $\mathbf{B}_{12} \in Z_q^{n\times(m-2n)}$. Compute $(\mathbf{A}_1, \mathbf{T}) \leftarrow \mathbf{TrapGen}(1^n, 1^{m-n}, q)$ and let $\mathbf{A}_2$ be $\mathbf{B}_{11}$. Since $\mathbf{B}_2$ is invertible, the matrix $\mathbf{C}$ can be computed as

$$\begin{bmatrix} \mathbf{0}_{(m-n)\times(m-n)} & \mathbf{I}_{(m-n)\times n} \\ \mathbf{B}_2^{-1}\mathbf{A}_1 \in Z_q^{n\times(m-n)} & \mathbf{0}_{n\times n} \end{bmatrix}.$$

(2) Compute the short basis $\mathbf{T} \leftarrow \mathbf{ExtBasis}(\mathbf{T}_1, \mathbf{A})$. Output $\mathbf{A}$, $\mathbf{T}$ and $\mathbf{C}$.

Now, we prove that this algorithm satisfies the required properties above. First,

$$\mathbf{B}_1\mathbf{I}_{(m-n)\times n} = [\mathbf{B}_{11} \| \mathbf{B}_{12}] \begin{bmatrix} \mathbf{I}_{n\times n} \\ \mathbf{0}_{(m-2n)\times n} \end{bmatrix} = \mathbf{B}_{11} = \mathbf{A}_2.$$

Thus,

$$\mathbf{BC} = [\mathbf{B}_1 \| \mathbf{B}_2] \begin{bmatrix} \mathbf{0} & \mathbf{I}_{(m-n)\times n} \\ \mathbf{B}_2^{-1}\mathbf{A}_1 & \mathbf{0} \end{bmatrix} = [\mathbf{A}_1 \| \mathbf{A}_2] = \mathbf{A} \bmod q.$$

$\mathbf{A}$ is statistically close to uniform over $Z_q^{n\times m}$ because $\mathbf{A}_1$ is statistically close to uniform over $Z_q^{n\times(m-n)}$ and $\mathbf{B}$ is a random matrix (Lemma 1). (3) and (4) can be obtained directly from Lemma 4. (5) holds because every entry of matrix $\mathbf{C}$ is not larger than q. This completes the proof.

4. Our Network Coding Signature Scheme

In this section we present our lattice-based network coding signature scheme in the standard model. Our scheme achieves the desired properties of correctness, unforgeability and privacy. In our scheme, n is the security parameter. Let $L \ge 1$ be the maximum number of linear combinations and the maximum number of source nodes. Suppose that the augmented message blocks transmitted in the network are represented by n-dimensional binary vectors. We set the Gaussian parameter $s = O(\sqrt{n\log q})w(\sqrt{\log n})$. Now we first give the specific scheme as follows:

- **KeyGen**$(1^n, L)$. The algorithm takes the security parameter n and L as inputs:

  (1) Choose parameters m and q, where $m = O(n\log q)$ and $q = poly(n)$.

(2) Sample a random matrix $\mathbf{A} \in Z_q^{n \times m}$ and its corresponding trapdoor short basis $\mathbf{T}_1 \in Z^{m \times m}$ using the **TrapGen** algorithm.

(3) Generate $L-1$ independent short basis $\mathbf{T}_i (1 < i \le L)$ of the lattice $\Lambda^{\perp}(\mathbf{A})$ using the **RandBasis** algorithm.

(4) Choose a random matrix $\mathbf{B} = [\mathbf{b}_1, \mathbf{b}_2, \cdots, \mathbf{b}_m] \in Z_q^{n \times m}$.

Let $(\mathbf{A}, \mathbf{T}_i)$ be the public/private pair of the i-th source node. Output the common public key **A** and **B**, and send $\mathbf{T}_i$ to the i-th source node secretly.

• **Sign**(id, $\mathbf{v}_i$, $\mathbf{T}_i$). For the i-th source node, the algorithm takes an identifier $id \in \{0,1\}^n$, a message $\mathbf{v}_i \in F_2^n$ and the secret key $\mathbf{T}_i$ as inputs:

(1) For $i = 1,2,\cdots,n$, let $\mathbf{h}_i = id \| 00\cdots00 \| BitSring(i) \in \{0,1\}^m$, where the number of '0' added in front of $BitSring(i)$ is $m - n - \lceil \log n \rceil$ and $BitSring(i) \in \{0,1\}^{\lceil \log n \rceil}$ is the binary representation of i.

(2) Let $\mathbf{H} = [\mathbf{h}_1, \mathbf{h}_2, \cdots, \mathbf{h}_n] \in F_2^{m \times n}$. Output the signature $\boldsymbol{\delta}_i \leftarrow$ **SamplePre**(**A,** $\mathbf{T}_i$**,** s**,** $\mathbf{BHv}_i$).

• **Com**(**A**, **B**, id, $\{\alpha_i\}_{i=1}^l$ $\{\mathbf{v}_i, \boldsymbol{\delta}_i\}_{i=1}^l$). The algorithm takes the public key **A** and **B**, an identifier id and $l (\le L)$ message/signature pairs tagged the same identifier id as inputs:

(1) Choose l encoding coefficients $\{\alpha_i\}_{i=1}^l \in \{0,1\}^l$.

(2) Output the combined signature $\boldsymbol{\delta} = \sum_{i=1}^l \alpha_i \boldsymbol{\delta}_i$ on the message $\mathbf{v} = \sum_{i=1}^l \alpha_i \mathbf{v}_i$.

• **Verify**(**A**, **B**, id, **v**, $\boldsymbol{\delta}$). This algorithm takes the public key **A** and **B**, an identifier id, a message **v** and signature $\boldsymbol{\delta}$ as inputs:

(1) Check that $\mathbf{A}\boldsymbol{\delta} = \mathbf{BHv} \bmod q$ and $\|\boldsymbol{\delta}\| \le Ls\sqrt{m}$.

(2) Output 1 (accept) if and only if the above two conditions hold. Otherwise, output 0 (reject).

4.1 Correctness

Since the parameter $s = O(\sqrt{n \log q}) w(\sqrt{\log n}) \ge \|\widetilde{\mathbf{T}}_i\| w(\sqrt{\log n})$, two important algorithms **SamplePre** and **RandBasis** in our scheme can work correctly with overwhelming probability from Lemma 2 and Lemma 4. From Lemma 3, the signatures produced by the **Sign** algorithm can be accepted by the **Verify** algorithm obviously. If the signatures are generated by the **Com** algorithm, we have

$$\mathbf{A}\boldsymbol{\delta} = \sum_{i=1}^l \alpha_i \mathbf{A}\boldsymbol{\delta}_i = \sum_{i=1}^l \alpha_i \mathbf{BHv}_i = \mathbf{BH}\sum_{i=1}^l \alpha_i \mathbf{v}_i = \mathbf{BHv} \bmod q;$$

$$\|\boldsymbol{\delta}\| = \left\| \sum_{i=1}^l \alpha_i \boldsymbol{\delta}_i \right\| \le Ls\sqrt{m} \text{ because } \alpha_i \in \{0,1\} \text{ and } \|\boldsymbol{\delta}_i\| \le s\sqrt{m}.$$

Thus, they also can be accept by the **Verify** algorithm.

4.2 Unforgeability

Here we show the existential unforgeability of our lattice-based network coding signature

scheme under chosen file attacks. Given an adversary that breaks our proposed signature scheme, we can construct an challenger that simulates the signature scheme and solves the SIS problem.

Theorem 2. If there is a PPT adversary that can win the security game defined in Definition 2 with advantage ε, then there is a challenger that can solve the $\mathrm{SIS}_{q,m,\beta}$ problem with the same advantage, where $\beta = qO(n\sqrt{m\log q})w(\sqrt{\log n}) = ploy(n)$.

Proof. Suppose that there is a PPT adversary that wins the game of existential unforgeability with advantage ε. Our aim is to construct a challenger that takes a random instance $\mathbf{B} \in Z_q^{n\times m}$ of the SIS problem as input and outputs a nonzero vector $\mathbf{u}$ such that $\mathbf{Bu} = \mathbf{0} \bmod q$ and $\|\mathbf{u}\| \le \beta$. The simulation step is as follows:

• The challenger computes $(\mathbf{A}, \mathbf{T}_1, \mathbf{C}) \leftarrow \mathbf{ProductSam}(1^n, 1^m, q, \mathbf{B})$, $\mathbf{T}_i \leftarrow \mathbf{RandBasis}(\mathbf{T}_1, s)$ and outputs the public key **A** and **B**.

• The adversary adaptively makes a polynomial (in n) number of queries. For the i-th query, he chooses a file represented by k vectors $\mathbf{v}_{i1}, \mathbf{v}_{i2}, \cdots, \mathbf{v}_{ik} \in F_2^n$ and the challenger does the following:

(1) Choose a random id_i from $\{0,1\}^n$.

(2) Compute the $\mathbf{h}_i$ and **H** according to the **Sign** algorithm.

(3) Output the signature $\boldsymbol{\delta}_{ij}$ using the algorithm **SamplePre**(**A**, $\mathbf{T}_i$, s, $\mathbf{BHv}_{ij}$).

(4) Output the signed data $\{\boldsymbol{\delta}_{ij}\}_{j\in[k]}$ and sends it to the adversary.

• Eventually the adversary outputs an identifier id^*, a non-zero vector $\mathbf{v}^*$, and a signature $\boldsymbol{\delta}^*$.

In fact, the distribution of the challenger's outputs is statistically indistinguishable from the distribution of the outputs in the real signature scheme. In the real scheme, the public key **A** is sampled from the algorithm **TrapGen** and **B** is chosen uniformly at random. In the simulation step, **A** is the output of the algorithm **ProductSamp** and **B** is a random instance of the SIS problem. From the result in Section 3, we can easily know that the distribution of public keys (**A**,**B**) is statistically indistinguishable in real and simulated execution. In addition, the output distributions of the signatures in both executions are statistically indistinguishable because all signatures are generated by the algorithm **PreSample** using the trapdoor short basis of $\Lambda^{\perp}(\mathbf{A})$.

If the adversary outputs a forgery ($\mathbf{v}^*$, $\boldsymbol{\delta}^*$) for the identifier id^*, we can solve the SIS solution for a random instance $\mathbf{B} \in Z_q^{n\times m}$. The forgeries can be divided into the following two different classes:

(1) $id^* \neq id_i$ for all queried i, i.e., the adversary never makes signature queries for any block message tagged by the identifier id^*. Naturally we have $\mathbf{A}\boldsymbol{\delta}^* = \mathbf{BH}^*\mathbf{v}^* \bmod q$. From Theorem 1, we can obtain that $\mathbf{B}(\mathbf{C}\boldsymbol{\delta}^* - \mathbf{H}^*\mathbf{v}^*) = \mathbf{0} \bmod q$, where $\mathbf{H}^* \in F_2^{m\times n}$ is generated by the identifier id^* using the same method in the **Sign** algorithm.

Let $\mathbf{u} = \mathbf{C}\boldsymbol{\delta}^* - \mathbf{H}^*\mathbf{v}^*$. Obviously, we have

$$\mathbf{Bu} = \mathbf{0} \bmod q$$

and

$$\|\mathbf{u}\| \le \|\mathbf{C}\boldsymbol{\delta}^*\| + \|\mathbf{H}\mathbf{v}^*\| \le (q \cdot s + 1)\sqrt{m}\sqrt{n} = qO(n\sqrt{m\log q}) \cdot w(\sqrt{\log n}) \le \beta .$$

Similar to previous lattice-based network coding schemes [32] [33], we can obtain that $\mathbf{u} \neq \mathbf{0}$ with overwhelming probability.

(2) $id^* = id_i$ for some index i but $\mathbf{v}^*$ is not a linear combination of $\mathbf{v}_{i1}, \mathbf{v}_{i2}, \cdots, \mathbf{v}_{ik}$. We have $\mathbf{A}\boldsymbol{\delta}^* = \mathbf{B}\mathbf{H}^*\mathbf{v}^* \bmod q$. Since the adversary have requested the signatures of L vectors $\mathbf{v}_{i1}, \mathbf{v}_{i2}, \cdots, \mathbf{v}_{ik}$, the challenger can output a combined signature $\boldsymbol{\delta}$ on the combined messages $\mathbf{v}$. Thus, we also have $\mathbf{A}\boldsymbol{\delta} = \mathbf{B}\mathbf{H}^i\mathbf{v} \bmod q$, where $\mathbf{H}^i = \mathbf{H}^*$ because $id^* = id_i$. Hence,

$$\mathbf{A}(\boldsymbol{\delta}^* - \boldsymbol{\delta}) = \mathbf{B}\mathbf{H}^*(\mathbf{v}^* - \mathbf{v}) \bmod q ,$$

i.e.,

$$\mathbf{B}[\mathbf{C}(\boldsymbol{\delta}^* - \boldsymbol{\delta}) - \mathbf{H}^*(\mathbf{v}^* - \mathbf{v})] = \mathbf{0} \bmod q .$$

Let $\mathbf{u} = \mathbf{C}(\boldsymbol{\delta}^* - \boldsymbol{\delta}) - \mathbf{H}^*(\mathbf{v}^* - \mathbf{v})$. Obviously, we have

$$\mathbf{B}\mathbf{u} = \mathbf{0} \bmod q$$

and

$$\|\mathbf{u}\| = \|\mathbf{C}(\boldsymbol{\delta}^* - \boldsymbol{\delta})\| + \|\mathbf{H}^*(\mathbf{v}^* - \mathbf{v})\| \le (2qs + 1)\sqrt{n}\sqrt{m} = qO(n\sqrt{m\log q})w(\sqrt{\log n}) \le \beta .$$

From [32] [33], we also can obtain that $\mathbf{u} \neq \mathbf{0}$ with overwhelming probability.

4.3 Privacy

In order to guarantee the privacy of signature packets in our scheme, we also consider the property of *weakly context hiding*. Every intermediate node generate a combined signature $\boldsymbol{\delta}$ on a combined message $\mathbf{v}$ using the **Com** algorithm when it receives l message/signature pairs $(\mathbf{v}_i, \boldsymbol{\delta}_i)_{i\in[l]}$. The *weakly context hiding* means that the combined signature does not leak any information about $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_l$ beyond what is revealed by $\mathbf{v}$.

Theorem 3. Our network coding signature scheme Π = (**KeyGen**, **Sign**, **Com**, **Verify**) is *weakly context hiding*.

Proof. In the privacy game, suppose that the challenger runs the algorithm **KeyGen** to get the common public key **A** and all private keys $\{\mathbf{T}_i\}_{i\in[L]}$ and give them to the adversary. Let $(V_0, V_1, f_1, f_2, \cdots, f_k)$ be the adversary's output in the challenge phase, where $V_b = \{\mathbf{v}_1^{(b)}, \mathbf{v}_2^{(b)}, \cdots, \mathbf{v}_k^{(b)}\}$ for b =0,1. Let $\mathbf{c}_i = f_i(\mathbf{v}_1^{(0)}, \mathbf{v}_2^{(0)}, \cdots, \mathbf{v}_k^{(0)}) = f_i(\mathbf{v}_1^{(1)}, \mathbf{v}_2^{(1)}, \cdots, \mathbf{v}_k^{(1)})$ for all $i = 1,2,\cdots,k$. For $j = 1,2,\cdots,k$, let $\boldsymbol{\delta}_j^b$ be the challenger's signature on the message $\mathbf{v}_j^b$. For $i = 1,2,\cdots,k$, let $\mathbf{d}_i^b$ be a combined signature on $\mathbf{c}_i$ computed using the **Com** algorithm applied to the signature $\{\boldsymbol{\delta}_j^b\}_{j\in[k]}$ and the function f_i. The challenger chooses a random bit b and gives the adversary the signatures $\{\mathbf{d}_i^b\}_{i\in[k]}$.

Suppose $b = 0$. By the definition of the algorithm **Sign** every signature $\boldsymbol{\delta}_j^0$ is generated from a distribution statistically close to $D_{\mathbf{t}_j + \Lambda^\perp(\mathbf{A}), s}$ and these signatures are mutually independent, where $\mathbf{t}_j \in Z^m$ is an arbitrary solution to $\mathbf{A}\mathbf{t}_j = \mathbf{B}\mathbf{H}\mathbf{v}_j \bmod q$. Therefore, by Lemma 5 the combined signature $\mathbf{d}_i^0 = f_i(\boldsymbol{\delta}_1^0, \cdots, \boldsymbol{\delta}_k^0) = \sum_{j=1}^k c_i\boldsymbol{\delta}_j^0$ is statistically close to

$D_{\mathbf{t}+g\Lambda^{\perp}(\mathbf{A}),\|\mathbf{c}\|s}$, where $\mathbf{t}=\sum_{j=1}^{k}c_j\mathbf{t}_j$ and $g=\gcd(c_1,c_2,\cdots,c_k)$. Since the same holds for $b=1$, the distribution of $\mathbf{d}_i^0$ and $\mathbf{d}_i^1$ is statistically close. Consequently, the advantage of any PPT adversary in the privacy game defined in Section 2.3 is negligible.

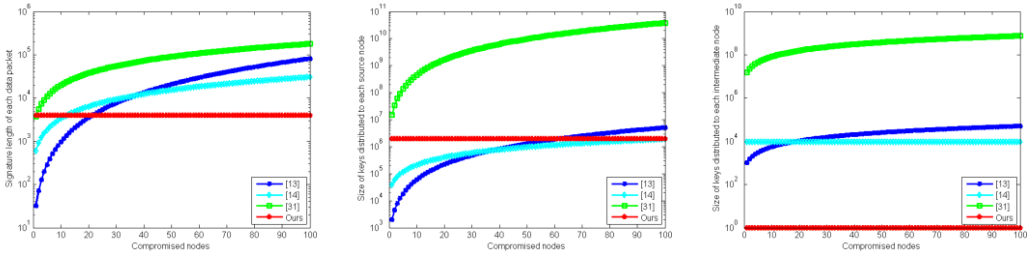
Note that in our scheme we set $s=O(\sqrt{n\log q})w(\sqrt{\log n})$. Since $\mathbf{c}$ is a 0/1 vector and $\eta_{\varepsilon}\left(\Lambda^{\perp}(\mathbf{A})\right)\le w\left(\sqrt{\log m}\right)$ for some negligible ε [35][38]. Thus, $s\ge\|\mathbf{c}\|\eta_{\varepsilon}\left(\Lambda^{\perp}(\mathbf{A})\right)$ and the condition of Lemma 5 holds. This completes our proof.

5. Efficiency

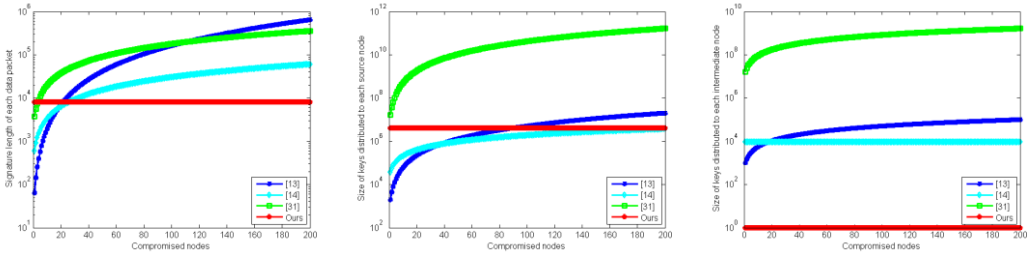
On the one hand, we provide a comparison of our scheme to previous lattice-based network coding signature schemes [32-35] in terms of model supporting, multi-source supporting, public key size, signature length, signing cost, respectively. In the **Sign** algorithm, source nodes mainly use three time-consuming algorithms, **SamplePre**, **ExtBasis** and **RandBasis**. For the sake of convenience, we denote the time cost to run once **SamplePre** algorithm, once **ExtBasis** algorithm and once **RandBasis** algorithm by T_{sp}, T_{eb} and T_{rb}, respectively. Because the length of id is the same for each scheme, and therefore we omit it in the comparison of the signature length. In the **Com** and the **Verify** algorithms, they mainly involve simple addition and multiplication operations over a finite field. Thus, each intermediate node in our scheme has low computational complexity. In fact, each intermediate node requires more complex operations in some number-theoretic protocol for secure network coding schemes. **Table 1** shows that compared to other related schemes, our scheme has low communication overhead. The size of all signatures in our scheme is very small, which achieves the minimum value of those in the five schemes. It is a pity that in order to design secure network coding scheme in the standard model, the public key size of our proposed scheme is two times of that in [33] [34] [35]. Although the signing cost is more than Wang's

Table 1. Comparison of existing lattice-based network coding signature schemes

Scheme	Model	Multi-source supporting	Public key size	Signature length	Signing cost
[32]	Random	No	$mn+mn\log q$	$2m+2m\log q$	$\mathrm{T}_{sp}+\mathrm{T}_{eb}$
[33]	Random	No	$mn\log q$	$m\log q$	T_{sp}
[34]	Random	Yes	$mn\log q$	$2m\log q$	$\mathrm{T}_{sp}+\mathrm{T}_{eb}$
[35]	Random	Yes	$mn\log q$	$m\log q$	$\mathrm{T}_{sp}+\mathrm{T}_{rb}$
Ours	Standard	Yes	$2mn\log q$	$m\log q$	$\mathrm{T}_{sp}+\mathrm{T}_{rb}$



(a) $q = 2^8$, $N = 1000$, $CN = 100$.



(b) $q = 2^{16}$, $N = 2000$, $CN = 200$.

Fig. 3. The comparison of communication overhead between some existing schemes based on number-theoretic assumptions and our lattice-based scheme.

Table 2. The operation time of each intermediate node.

Time / Scheme	[13]	[14]	[31]	Ours
Verification	6.35×10^{-5}s	2.97×10^{-5}s	1.2×10^{-2}s	4.09×10^{-4}s
Combination	3.25×10^{-4}s	2.5×10^{-3}s	3.3×10^{-4} s	2.78×10^{-5}s
Total	3.89×10^{-4}s	2.53×10^{-3}s	1.23×10^{-2}s	4.37×10^{-4}s

(a) $q = 2^8$, $r = 100$, $k = 50$.

Time / Scheme	[13]	[14]	[31]	Ours
Verification	5.33×10^{-4}s	4.79×10^{-5}s	1520.2s	8.96×10^{-4}s
Combination	2×10^{-3}s	6.9×10^{-3}s	2.2×10^{-3}s	8.59×10^{-5}s
Total	2.73×10^{-3}s	6.95×10^{-3}s	1520.2s	9.82×10^{-4}s

(b) $q = 2^{16}$, $r = 200$, $k = 100$.

[33], our scheme is in the standard model. In addition, our scheme can support multi-source network system. Thus from the perspective of communication overhead and computational complexity, our scheme is competitive.

On the other hand, we provide a comparison between our lattice-based scheme and some typical schemes based on number-theoretic assumptions in terms of communication overhead and computational complexity. We use a desktop which has a 8-core Intel(R) Core (TM) i7-4770 processor running at 3.40 GHz and 8 GB of RAM. Let CN be the maximal number of compromised nodes that the network system can tolerate and N be the number of nodes in the network. Agrawal et al. [13] proposed a homomorphic MAC for checking the integrity of network coded data, and its key distribution protocol was based on the cover-free family constructed from polynomials [42]. In their scheme, the source node has t^2 keys for generating tags and each intermediate node has t keys for correctness verification. From [42], we know that CN, N and t must satisfy $t-1 \geq CN \cdot \lceil \log_t N \rceil$. Because $N \geq t$ usually, $t \geq CN+1$. In order to minimize the communication overhead, we set $t = CN+1$ in our

simulations. Esfahani et al. [14] presented a dual-homomorphic MAC for network coding-enabled wireless sensor networks. Zhang et al. [31] proposed a hybrid-key cryptographic approach to network coding authentication, called MacSig scheme. We assume that the lengths of the seeds used in [14] and [31] are 500bit. All other relevant parameters in our experiments are the same as the three schemes mentioned above. In our scheme, we set $n = 200$ and $m = 500$. Note that in practical network coding setting, the order of the finite field is equal to 2^8 or 2^{16}. Thus, we consider these two cases. **Fig. 3** shows the comparison of communication overhead and **Table 2** shows the computational complexity of each intermediate node. From **Fig. 3** we can see that when the number of compromised nodes is larger than 20, the signature length of each data packet in our scheme is the shortest one. The size of keys distributed to each source node in our scheme are smaller than that in [31], which are very similar to that in [13] and [14]. In addition, our scheme does not distribute any private key to intermediate nodes. Simultaneously, the signature length and the private keys distributed to source nodes or intermediate nodes are not influenced by the number of compromised nodes. In **Table 2**, we investigate the time of processing data packets of each intermediate node, including the verification time and the combination time. This experiment was perform 1000 times and took the average value. Because the verification procedure in [31] needs the modular exponentiation operation, the time overhead is huge when the modulus q is very large. In all, from an experimental point of view, our proposed scheme is competitive compared to some number-theoretic schemes for secure network coding.

6. Conclusion

In this work, we propose a lattice-based network coding signature scheme in the standard model. In order to prove the security, we introduce a new trapdoor sampling method **ProductSamp** for generating random lattice and the corresponding short basis, which may also be used in many other cryptographic protocols. In fact, our scheme can achieve existential unforgeability under full chosen-message attacks [41], where the adversary can make adaptive queries on individual message blocks within a given file, possibly even interleaving those queries across several files.

Although our scheme can prevent multisource network system from pollution attacks, there is still much work to be done in order to improve the capability of the scheme. Note that ideal lattice can be used to decrease the public key size, and our future work mainly focuses on designing network coding signature schemes using that technique.

References

[1] Rudolf Ahlswede, Ning Cai, Shuo-Yen Robert Li, and Raymond W. Yeung, "Network information flow," *IEEE Transactions on Information Theory*, vol. 46, no. 4, pp. 1204-1216, July, 2000. Article (CrossRef Link).

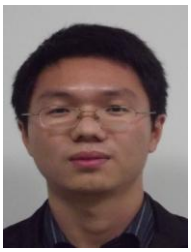
[2] J. Feldman, T. Malkin, C. Stein, and R.A. Servedio, "On the capacity of secure network coding," in *Proc. of 42th Annual Allerton Conference on Communication,Control, and Computing*, pp. 63-68, September 29-October1, 2004. Article (CrossRef Link).

[3] S. Jaggi, M. Langberg, S. Katti, T.Ho, D. Katabi, and M. Médard, "Resilient network coding in the presence of byzantine adversaries," in *Proc. of IEEE Conf. on Computer Communications*, pp.616-624, May 6-12, 2007. Article (CrossRef Link).

[4] T. Ho, B. Leong, R. Koetter, M. Médard, M. Effros, and D. R. Karger, "Byzantine modification detection in multicast networks with random network coding," *IEEE Transactions on Information Theory*, vol. 54, no. 6, pp. 2798-2803, June, 2008. Article (CrossRef Link).
[5] D. Boneh, D. Freeman, J. Katz, and B. Waters, "Signing a linear subspace: Signature schemes for network coding," in *Proc. of 12th International Conference on Practice and Theory in Public Key Cryptography*, pp. 68-87, March 18-20, 2009. Article (CrossRef Link).
[6] D. Catalano, D. Fiore, and B. Warinschi, "Efficient network coding signatures in the standard model," in *Proc. of 15th International Conference on Practice and Theory in Public Key Cryptography*, pp. 680-696, May 21-23, 2012. Article (CrossRef Link).
[7] P. W. Shor, "Algorithms for quantum computation: discrete logarithms and factoring," in *Proc. of 35th Annual Symposium on Foundations of Computer Science*, pp. 124-134, November 20-22, 1994. Article (CrossRef Link).
[8] P. W. Shor, "Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer," *SIAM Journal on Computing*, vol. 26, no. 5, pp. 1484-1509, October, 1997. Article (CrossRef Link).
[9] M. Ajtai, "Generating hard instances of lattice problems," in *Proc. of 28th Annual ACM Symposium on Theory of Computing*, pp. 99-108, May 22-24, 1996. Article (CrossRef Link).
[10] D. Micciancio and O. Regev, "Worst-case to average-case reductions based on Gaussian measures," *SIAM Journal on Computing*, vol. 37, no. 1, pp. 267-302, February, 2007. Article (CrossRef Link).
[11] C. Gkantsidis and P. R. Rodriguez, "Network coding for large scale content distribution," in *Proc. of 24th Annual Joint Conference of the IEEE Computer and Communications Societies*, pp. 2235-2245, March 13-17, 2005. Article (CrossRef Link).
[12] M. N. Krohn, M. J. Freedman, and D. Mazieres, "On-the-fly verification of rateless erasure codes for efficient content distribution," in *Proc. IEEE Symposium on Security and Privacy*, pp. 226-240, May 9-12, 2004. Article (CrossRef Link).
[13] S. Agrawal and D. Boneh, "Homomorphic MACs: MAC-based integrity for network coding," in *Proc. of 7th International Conference on Applied Cryptography and Network Security*, pp. 292-305, June 2-5, 2009. Article (CrossRef Link).
[14] A. Nascimento and J. Rodriguez, "Dual-homomorphic message authentication code scheme for network coding-enabled wireless sensor networks," *International Journal of Distributed Sensor Networks*, vol. 2015, Article ID 510251, 2015. Article (CrossRef Link).
[15] Y. Wang, "Insecure "provably secure network coding" and homomorphic authentication schemes for network coding," *IACR Cryptology ePrint Archive*, 60, 2010. Article (CrossRef Link).
[16] C. Cheng, T. Jiang, and Q. Zhang, "TESLA-based homomorphic MAC for authentication in P2P system for live streaming with network coding," *IEEE Journal on Selected Areas in Communicati- ons*, vol. 31, no. 9, pp. 291-298, September, 2013. Article (CrossRef Link).
[17] A. Esfahani, D. Yang, G. Mantas, A. Nascimento, and J. Rodriguez, "An improved homomorphic message authentication code scheme for RLNC-enabled wireless networks," in *Proc. of 19th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)*, pp. 80-84, December 1-3, 2014. Article (CrossRef Link).
[18] A. Esfahani, A. Nascimento, J. Rodriguez, and J. C. Neves, "An efficient MAC-signature scheme for authentication in XOR network coding," in *Proc. of 9th IEEE Symposium on Computers and Communication (ISCC)*, pp. 1-5, June 23-26, 2014. Article (CrossRef Link).
[19] W. Wang and L. Hu, "A generic homomorphic MAC construction for authentication in network coding," *Security and Communication Networks*, vol. 7, no. 2, pp. 429-433, February, 2014. Article (CrossRef Link).
[20] D. Charles, K. Jain, and K. Lauter, "Signatures for network coding," *International Journal of Information and Coding Theory*, vol. 1, no. 1, pp. 3-14, February, 2009. Article (CrossRef Link).
[21] F. Zhao, T. Kalker, M. Médard, and K. J. Han, "Signatures for content distribution with network coding," in *Proc. of IEEE International Symposium on Information Theory*, pp. 556-560, June 24-29, 2007. Article (CrossRef Link).

[22] S. Agrawal, D. Boneh, X. Boyen, and D. M. Freeman, "Preventing pollution attacks in multi-source network coding," in *Proc. of 13th International Conference on Practice and Theory in Public Key Cryptography*, pp. 161-176, May 26-28, 2010. Article (CrossRef Link).
[23] D. Catalano, D. Fiore, and B. Warinschi, "Efficient network coding signatures in the standard model," in *Proc. of 15th International Conference on Practice and Theory in Public Key Cryptography*, pp. 680-696, May 21-23, 2012. Article (CrossRef Link).
[24] N. Attrapadung and B. Libert, "Homomorphic network coding signatures in the standard model," in *Proc. of 14th International Workshop on Theory and Practice in Public Key Cryptography*, pp. 17–34, March 6–9, 2011. Article (CrossRef Link).
[25] Catalano Dario, Fiore Dario, and Warinschi Bogdan, "Adaptive pseudo-free groups and applications," in *Proc. of Advances in Cryptology-EUROCRYPT 2011*, pp. 207-223, May 15–19, 2011. Article (CrossRef Link).
[26] G. Liu and B. Wang, "Secure network coding against intra/inter-generation pollution attacks," *Communications, China*, vol. 10, no. 8, pp. 100-110, August, 2013. Article (CrossRef Link).
[27] C. Cheng, T. Jiang , Y. Liu, and M. Zhang, "Security analysis of a homomorphic signature scheme for network coding," *Security and Communication Networks*, vol. 8, no. 18, pp. 4053-4060, December, 2015. Article (CrossRef Link).
[28] H. He, R. Li, Z. Xu, and W. Xiao, "An efficient ECC-based mechanism for securing network coding-based P2P content distribution," *Peer-to-Peer Networking and Applications*, vol. 7, no. 4, pp. 572-589, December, 2014. Article (CrossRef Link).
[29] X. Wu, Y. Xu, C. Yuen, and L. Xiang, L, "A tag encoding scheme against pollution attack to linear network coding," *IEEE Transactions on Parallel and Distributed Systems*, vol. 25, no. 1, pp. 33-42, January, 2014. Article (CrossRef Link).
[30] Y. Zou, J. Zhu, L. Yang, Y. C. Liang, and Y. D. Yao, "Securing physical-layer communications for cognitive radio networks," *IEEE Communications Magazine*, vol. 53, no. 9, pp. 48-54, September, 2015. Article (CrossRef Link).
[31] P. Zhang, Y. Jiang, C. Lin, H. Yao, A. Wasef, and X. S. Shen, "Padding for orthogonality: Efficient subspace authentication for network coding," in *Proc. of the 30th IEEE International Conference on Computer Communications*, pp. 1026-1034, April, 2011. Article (CrossRef Link).
[32] D. Boneh and D. M. Freeman, "Linearly homomorphic signatures over binary fields and new tools for lattice-based signatures," in *Proc. of 14th International Workshop on Theory and Practice in Public Key Cryptography*, pp. 1-16, March 6-9, 2011. Article (CrossRef Link).
[33] FengHe Wang, YuPu Hu, and BaoCang Wang, "Lattice-based linearly homomorphic signature scheme over binary field," *SCIENCE CHINA: Information Sciences*, vol. 56, no. 11, pp.234-242, November, 2013. Article (CrossRef Link).
[34] Peng Zhang, Jianping Yu, and Ting Wang, "A homomorphic aggregate signature scheme based on lattice," *Chinese Journal of Electronics*, vol. 21, no. 4, pp. 701-704, October, 2012.
[35] Zhengjun Jing, "An efficient homomorphic aggregate signature scheme based on lattice," *Mathematical Problems in Engineering*, vol. 2014, pp. 1-9, 2014. Article (CrossRef Link).
[36] R. Canetti, O. Goldreich, and S. Halevi, "The random oracle methodology, revisited," *Journal of the ACM*, vol. 51, no. 4, pp. 557-594, July, 2004. Article (CrossRef Link).
[37] J. Alwen and C. Peikert, "Generating shorter bases for hard random lattices," *Theory of Computing Systems*, vol. 48, no. 3, pp. 535-553, April, 2011. Article (CrossRef Link).
[38] C. Gentry, C. Peikert, and V. Vaikuntanathan, "Trapdoors for hard lattices and new cryptographic constructions," in *Proc. of the 40th annual ACM symposium on Theory of computing*, pp. 197-206, May 17-20, 2008. Article (CrossRef Link).
[39] D. Cash, D. Hofheinz, E. Kiltz, and C. Peikert, "Bonsai trees, or how to delegate a lattice basis," *Journal of Cryptology*, vol. 25, no. 4, pp. 601-639, October, 2012. Article (CrossRef Link).
[40] S. D. Gordon, J. Katz, and V. Vaikuntanathan, "A group signature scheme from lattice assumptions," in *Proc. of Advances in Cryptology-ASIACRYPT*, pp. 395-412, December 5-9, 2010. Article (CrossRef Link).

[41] X. Boyen, X. Fan, and E. Shi, "Adaptively secure fully homomorphic signatures based on lattices," *IACR Cryptol. ePrint Archive*, 916, 2014. Article (CrossRef Link).
[42] R. Kumar, S. Rajagopalan, and A. Sahai, "Coding constructions for blacklisting problems without computational assumptions," in *Proc. of Advances in Cryptology-CRYPTO*, pp. 609-623, January, 1999. Article (CrossRef Link).



Dong Xie is a Ph.D. Candidate in Computer Science at the Beijing University of Posts and Telecommunications, Beijing, China. He received his M.S. degree from Hangzhou Normal University, Hangzhou, China, in 2013. His currently research interests include network security, lattice-based cryptography, homomorphic signature, homomorphic encryption.

Haipeng Peng received his M.S. degree in system engineering from Shenyang University of Technology, Shenyang, China, in 2006, and the Ph.D. degree in signal and information processing from Beijing University of Posts and Telecommunications, Beijing, China, in 2010. He is currently an associate professor at the School of Computer Science and Technology, Beijing University of Posts and Telecommunications. His research interests include information security, network security, complex networks and control of dynamical systems. Dr. H. Peng is the co-author of 50 scientific papers and over 10 Chinese patents.

Lixiang Li received the M.S. degree in circuit and system from Yanshan University, Qinhuangdao, China, in 2003, and the Ph.D. degree in signal and information processing from Beijing University of Posts and Telecommunications, Beijing, China, in 2006. She is currently a professor at the School of Computer Science and Technology, Beijing University of Posts and Telecommunications. Her research interests include swarm intelligence, information security and network security. Dr. L. Li is the co-author of 70 scientific papers and 10 Chinese patents.

Yixian Yang received the M.S. degree in applied mathematics in 1986 and the Ph.D. degree in electronics and communication systems in 1988 from Beijing University of Posts and Telecommunications, Beijing, China. He is the Managing Director of information security center, Beijing University of Posts and Telecommunications, Beijing, China. His research interests include network security, information security and coding theory. Dr. Y. Yang is the co-author of 300 scientific articles and 50 patents.