Wide Area Key Distribution Network Based on a Quantum Key Distribution System

Hua Dong 1,2,3,†, Yaqi Song 4,† and Li Yang 1,2,3,*

1 State Key Laboratory of Information Security, Institute of Information Engineering, Chinese Academy of Sciences, Beijing 100093, China; donghua@iie.ac.cn
2 State Key Laboratory of Cryptology, P.O.Box 5159, Beijing 100878, China
3 School of Cyber Security, University of Chinese Academy of Sciences, Beijing 100049, China
4 China Academy of Electronics and Information Technology, Beijing 100041, China; songyaqi@iie.ac.cn
* Correspondence: yangli@iie.ac.cn; Tel.: +86-133-0132-0389
† These authors contributed equally to this work.

Received: 16 January 2019; Accepted: 6 March 2019; Published: 14 March 2019

Abstract: The point-to-point quantum key distribution (QKD) system is limited by the transmission distance. So, the wide area QKD network with multiple endpoints is the research focus of this study. The relay-node scenario and key relay protocols provide the solutions to the QKD network. The early key relay protocols require the relay nodes to be reliable. Once the relay nodes become compromised, the whole network is insecure. In this paper, we extend the chain structure of the public-XOR(exclusive OR)-key scheme with two endpoints to the complex network with multiple endpoints. The relay nodes in our scheme do not need encryption actions, decryption actions, or storage XOR keys, which simplifies the system compared with other key distribution schemes based on trusted relay nodes. Our scheme not only improves the practical performance and simplifies the system’s complexity, but it also ensures that the security is not reduced. Specifically, we rigorously demonstrate that an eavesdropper can never access the key shared by the users of the network as long as the process of generating XOR keys and destroying the original keys is secure. In addition, we discuss the information leakage of the practical QKD network from the perspective of the unicity distance.

Keywords: quantum key distribution; key distribution network; relay node; unicity distance

1. Introduction

To ensure the security of communications on the internet, an effective solution is to use cryptographic techniques. The security of the most classical cryptosystems depends on the computational complexity of the algorithms. With the fast development of quantum computation and quantum algorithms, communication security based on the above algorithms is greatly compromised. In conventional cryptography, a typical cryptosystem not based on the computational assumptions is the one-time-pad, which is unconditionally secure and free from the threat of quantum computing. In one-time-pad, the length of the keys and the messages is the same, and the keys can only be used once. Both communicating parties need to share the keys in advance. The security of one-time-pad should be completely dependent on the confidentiality of the pre-shared keys. It is too expensive to realize and too difficult to apply in practice. Thus, how to realize the unconditionally secure key distribution becomes a critical problem, which motivates the development of quantum key distribution (QKD). In 1984, Bennett and Brassard proposed the concept of QKD [1], which achieves the unconditionally secure key distribution based on quantum mechanics. It enables both parties of communication to generate and share a random and secure key. Then a series of QKD protocols have been proposed [2–5]. More importantly, the unconditional security of QKD protocol has been strictly proved [6–9].
The early implementations of QKD systems mainly focus on the communication between two endpoints. The point-to-point QKD system is limited by the transmission distance, which is limited by the key rate. So, the development of a wide area QKD network is a challenging issue. Some countries and regions have deployed QKD networks for last years. The first experiment of QKD network was designed by Townsend [10]. Some QKD network projects [11–18] have been completed successfully, such as the defense advanced research projects agency (DARPA) quantum network [11] set up by the USA in 2004, secure communication based on quantum cryptography (SECOQC) designed by the European project [13], the commercial telecommunication fiber network in China [12], the SwissQuantum QKD network project launched in Geneva [14], and the live video conference through a high-speed QKD network in Tokyo [16]. Satellite communications are also building QKD networks, such as the international QKD channel created by the QUESS space mission [19,20] and the world’s first space-ground quantum network [21,22]. These projects supported by different underlying devices can be divided into three categories: (1) satellite communications; (2) quantum relaying; (3) classical trusted relay. In the light of the present situation, the satellite-based method is not realistic, and the research of quantum relay equipment is not perfect. Thus, determining how to establish a practical and secure QKD network based on the trusted relay is the key problem.

The approach of using trusted relays in QKD networks was proposed in 2002 [23]. This allows both endpoints to access a series of trusted relays to expand the arbitrary distance. The BBN key relay protocol has been operating continuously in the DARPA quantum network since 2003 [24]. In the protocol, the endpoints create a new random number \( R \) and send \( R \) through the trusted relay nodes by one-time-pad encryption. Each trusted relay node of the chosen path decrypts the ciphertext with the QKD keys shared with the upstream node and encrypts \( R \) with the QKD keys shared with the downstream node. In 2012, Los Alamos National Laboratory set up a trusted hub-and-spoke QKD network [18]. To communicate, each node sends a one-time pad to the hub, which it then uses to communicate securely over a classical link. The entire network is secure only if the central hub is secure. In these relay nodes-based key distribution schemes, if one of the relay nodes gets compromised, the whole network is insecure. Additionally, the relay nodes need encryption actions, decryption actions and storage of function, which increase the complexity of the system. In [25], Schartner and Rass proposed a re-encryption key distribution scheme for relay QKD system, which is introduced in Section 2.1. Compared with BBN key relay protocol, the re-encryption scheme reduces the damage in case of the compromised node and reduces the storage memory of keys for relay nodes. However, in re-encryption scheme, the relay nodes still need to keep the XOR (exclusive OR) value of the QKD keys secretly, and the successive communications between adjacent relay nodes is also required. In 2013, “Beijing–Shanghai Line” of quantum secrecy communication is based on the public-XOR-key scheme, where each trusted relay node publicly announces the XOR key it holds, enabling the endpoints to share a key. Compared with the re-encryption scheme, the scheme simplifies the system’s complexity and eases the traffic of the relay nodes. In 2014, William Stacey et al. proposed a simplified trusted relay (STR) protocol for point-to-point [26]. It is similar to public-XOR-key scheme. The two endpoints of STR are connected by a series of trusted relays that announce the XOR keys, but the rest of the classic QKD post-processing needs to be performed based on data from its original terminal, which comes at the cost of a low key rate.

Contributions: considering the key rate and the requirements of the long-distance communications with multiple endpoints, the public-XOR-key scheme is more efficient and practical than the BBN protocol, re-encryption protocol and STR. We extend the chain structure of the public-XOR-key scheme with two endpoints to the complex network with multiple endpoints. Our contributions are as follows.

- Our scheme can reduce the memory complexity and heavy traffic of communications for relay nodes in real implementation. It is appropriate for the remote communications of the complex networks with multiple endpoints. Compared with other key distribution schemes based on relay nodes, our scheme reduces the complexity of the system and eases the traffic of the relay nodes.
• Compared with the re-encryption schemes that does not announce the XOR value, our scheme not only improves the actual performance and simplifies the complexity of the system, it also does not weaken the security of the system. To analyze the security, we build the threat model and the security model of the public-XOR-key scheme. We rigorously demonstrate that the scheme is as secure as the re-encryption scheme under the same adversary model. In addition, we analyze the information leakage of practical QKD network from the perspective of Shannon’s ciphertext-only attack model by exploring the theory of unicity distance and applying it to the model of practical QKD networks.

2. Preliminaries

In this section, we review two cryptographic primitives that are necessary for understanding the proposed complex QKD network scheme, including the re-encryption key distribution scheme and the theory of unicity distance for discussing information leakage of practical key distribution scheme.

2.1. Re-Encryption Scheme for QKD Network

In re-encryption scheme, a secret message is encrypted by sender’s local random key $s$. The delivery of the random key $s$ can be executed by the following protocol. After the protocol, two endpoints can share the key $s$. The following is the main steps of the re-encryption protocol with two endpoints.

Protocol 1: re-encryption scheme for the relay QKD network:

1. Each node executes QKD protocol with its neighbor nodes, which generates $n$ pairs of QKD keys $Q_k$ as shown in Figure 1.
2. Alice encrypts the random key $s$ with her QKD key $Q_{k1}$, which is shared with the first relay node. Calculate the first ciphertext $c_1 = s \oplus Q_{k1}$, where $\oplus$ denotes bitwise module-2 addition.
3. The relay nodes do the XOR operation with the QKD keys, which are shared with the upstream relay nodes and the downstream relay node, to obtain the XOR keys $a_i$, i.e., $a_i = Q_k \oplus Q_{k+1}$, then immediately dismiss the QKD keys, $Q_k$, $Q_{k+1}$. When relay node $i$ receives the incoming ciphertext, it calculates $c_{i+1} = c_i \oplus a_i$. It is easy to see that $c_1 = s \oplus Q_{k1}$.
4. When Bob receives the incoming ciphertext, $c_n = s \oplus Q_{kn}$, he decrypts the ciphertext with his QKD key $Q_{kn}$ to obtain $s$. Then Alice and Bob share the common key $s$. Alice can transmit the message secretly to Bob with the key latter.

2.2. The Theory of Unicity Distance

In order to discuss the information leakage of the practical QKD network, we study it from a perspective of the unicity distance. Shannon laid the foundations of communication theory of secrecy system in [27]. In the second part of the paper, he put forward the concept of unicity distance, which gave a way of calculating approximately how many ciphertexts are required to obtain the key of a secrecy system. The adversary tries each possible key to decrypt the intercepted ciphertexts, and records the keys which make the deciphered messages meaningful in the original language. Only one of the record keys is the correct one, others are spurious keys.
In Shannon’s ciphertext-only attack model, suppose there is a symmetric cryptosystem 
\((M, C, K, E, D)\), where \(|C| = |M|\). In the cryptosystem, \(M\) denotes plaintext space, \(C\) denotes ciphertext space, \(K\) denotes key space, \(E\) denotes the encryption algorithms and \(D\) denotes the decryption algorithms. For \(m \in M, k \in K\), the encryption algorithm is \(E_k(m_i) = c_i\). Suppose the plaintexts are in language \(L\), the rate(entropy) is

\[
H_L = \lim_{n \to \infty} \frac{H(A^n)}{n}.
\]  

(1)

The redundancy of the language \(L\) is defined as

\[
R_L = 1 - \frac{H_L}{\log_2 |A|},
\]  

(2)

where \(A\) is the word set, \(|A|\) is the number of words in language \(L\), \(A^n\) is the totality of all plaintext \(n\)-alphabets, and \(H_L\) denotes the average information bits of the each word in language \(L\). \(R_L\) measures the degree of statistical constraint imposed by language \(L\).

For \(Y \in C^n\), \(K(Y)\) is defined as a set of all possible key with the given ciphertext \(Y\), \(p(Y)\) is the probability of the events \(Y\) occurring. The expected number of spurious key is

\[
S_n = \sum_{Y \in C^n} p(Y) (|K(Y)| - 1) = \sum_{Y \in C^n} p(Y)|K(Y)| - 1.
\]  

(3)

When \(n\) is sufficiently large, the relationship between the redundancy of the language \(R_L\) and the plaintext \(M\) is

\[
H(M^n) \approx n(1 - R_L) \log_2 |M|.
\]  

(4)

Since \(H(K|C^n) = H(K) + H(M^n) - H(C^n)\), \(H(C^n) \leq n \log_2 |C|\), where \(C^n\) is the totality of all ciphertext \(n\)-alphabets. Then

\[
H(K|C^n) \geq H(K) - nR_L \log_2 |M|.
\]  

(5)

According to Jensen’s inequality and Equation (3), the conditional entropy is

\[
H(K|C^n) = \sum_{Y \in C^n} p(Y)H(K|Y) \leq \sum_{Y \in C^n} p(Y) \log_2 |K(Y)|
\leq \log_2 \sum_{Y \in C^n} p(Y)|K(Y)| = \log_2 (S_n - 1).
\]  

(6)

Therefore, the expected number of spurious key \(\overline{S_n}\) satisfies that

\[
\overline{S_n} \geq 2^{\frac{H(K)}{|M|nR_L}} - 1.
\]  

(7)

When the length of the ciphertext is \(N_0\), the expectation of the number of spurious key is null. \(N_0\) is defined as the unicity distance of the system,

\[
N_0 \approx \frac{H(K)}{R_L \log_2 |M|},
\]  

(8)

where the entropy \(H(K)\) measures the uncertainty of the value of \(k\).

### 3. Public-XOR-Key Scheme for QKD Network

In this section, we extend the chain structure of the public-XOR-key scheme with two endpoints to the complex network with multiple endpoints. The scheme is secure as long as the XOR keys are generated securely. In this scheme, there are no re-generation keys and the relay nodes need not perform the encryption operation or store any key material, which reduces the complexity of the
system. There is no need to communicate between adjacent relay node, which avoids communication failure caused by network congestion.

3.1. The Public-XOR-Key Scheme with Chain Structure

First, we introduce the chain structure of two endpoints. In the scheme, each relay node has only two connections with the upstream node and the downstream node. Assume the whole chain network contains \( n + 1 \) nodes. Each node executes QKD protocol with the neighbor nodes. The \( n + 1 \) nodes make up \( n \) point-to-point QKD systems. There are two endpoints, \( n - 1 \) relay nodes and \( n \) pairs of QKD keys \( Q_k \) for \( i = 1, 2, ..., n \).

Protocol 2: public-XOR-key scheme for the relay QKD system:

1. Each node executes QKD protocol with the neighbor nodes, which generates \( n \) pairs of QKD keys \( Q_k \) for \( i = 1, 2, ..., n \) as shown in Figure 2.

2. Each relay node does the XOR operation with the QKD keys, which are shared with the upstream relay node and the downstream relay node, to obtain the XOR keys. For relay node \( i \), its XOR key is \( a_i = Q_k_i \oplus Q_{k_{i+1}} \), where \( Q_k_i \) is the QKD key shared by relay node \( i - 1 \) and relay node \( i \), \( Q_k_{i+1} \) is the QKD key shared by relay node \( i \) and relay node \( i + 1 \). Then each relay node immediately dismisses the QKD keys and publishes the XOR keys. We assume that the process of dismissing QKD keys is secure.

3. Bob calculates the final key with all of the public-XOR-key and his QKD key \( Q_k_n \). That is, \( \bigoplus_{i=1}^{n-1} a_i \oplus Q_k_n = Q_k_1 \). Then Alice and Bob share the common key \( Q_k_1 \). Alice can transmit the message secretly to Bob with the key \( Q_k_1 \) later.

Remark 1. There is a security assumption in the protocol that the process of dismissing QKD keys is secure. For the practical realization of the assumption, we give a kind of method as follows. Use the physical isolation to protect XOR operation and the process of dismissing the original QKD keys from the outside world to access its physical devices, so that even if the relay node is compromised, the adversary can not get the original key.

3.2. The Complex QKD Network Scheme with Multiple Endpoints

We have described the public-XOR-key scheme in the chain network with two endpoints. The following approach is used to solve the key distribution in a wide area network with multiple endpoints. A communication chain is formed between any two endpoints, and multiple communication chains constitute all kinds of network structures. No matter how complex the structures are, there will always be a chain between the two endpoints. As long as the pair of endpoints is in a connected graph, the key distribution scheme is similar to that of Protocol 2. The difference between the communications in the chain network and other complex networks is that some relay nodes have more than two connections
with other nodes. Additionally, the simplest model is the chain structure (each connecting relay node has only two ports). In general networks, there are not only two ports connecting the relay nodes. Since there may be various topological structures in the actual network, we abstract the complex network into simplified model with three-port relay nodes, as shown in Figure 3. The more complex multiple endpoints models are similar to this, as long as it conforms to the principle of one-time padding.

Firstly, consider a simplified model that a relay node has three connections that are connected with three endpoints: Alice, Bob and Charlie. Each pair of endpoints is in a chain network. When they execute the original process of Protocol 2, the communications of Alice and Bob, Alice and Charlie are shown in Figure 3.

$Q_k^A$ is the QKD key shared by Alice and the relay node, $Q_k^B$ is the QKD key shared by Bob and the relay node, $Q_k^C$ is the QKD key shared by Charlie and the relay node. The relay node calculates and publishes

$Q_k^{AB} = Q_k^A \oplus Q_k^B,$

$Q_k^{AC} = Q_k^A \oplus Q_k^C,$

$Q_k^{BC} = Q_k^B \oplus Q_k^C,$

and then dismisses the QKD keys immediately. Bob calculates $Q_k^{AB} \oplus Q_k^B$, which is equal to $Q_k^A$. Then Bob and Alice share the same key $Q_k^A$. Charlie calculates $Q_k^{AC} \oplus Q_k^C$, which is equal to $Q_k^A$. Then Charlie and Alice share the same key $Q_k^A$.

![Figure 3. The communications of three endpoints with Protocol 2 with repeated keys.](image)

However, a key distribution scheme is secure if and only if the shared keys are not known by the third party. In the above scenario, Alice, Bob and Charlie share the same key $Q_k^A$. If Bob or Charlie is the internal attacker, he can get the keys shared by Alice and the other user, which threatens the security.

In order to guarantee the security against the internal attacker, all of the QKD keys should be used only once. The improved key distribution scheme for three endpoints is shown in Figure 4. $Q_k^{A_1}$ and $Q_k^{A_2}$ are the QKD keys shared by Alice and the relay node, $Q_k^{B_1}$ and $Q_k^{B_2}$ are the QKD keys shared by Bob and the relay node, $Q_k^{C_1}$ and $Q_k^{C_2}$ are the QKD keys shared by Charlie and the relay node. Each QKD key is independent of the others. The relay node calculates and publishes

$Q_k^{A_1B_1} = Q_k^{A_1} \oplus Q_k^{B_1},$ $Q_k^{A_2C_2} = Q_k^{A_2} \oplus Q_k^{C_2},$ $Q_k^{B_2C_1} = Q_k^{B_2} \oplus Q_k^{C_1},$ and then dismisses the QKD keys immediately. Alice calculates $Q_k^{A_2C_2} \oplus Q_k^{A_2}$ to get the key $Q_k^{C_2}$, which is the final key shared with Charlie. Bob calculates $Q_k^{A_1B_1} \oplus Q_k^{B_1}$ to get the key $Q_k^{A_1}$, which is the final key shared with Alice. Charlie calculates $Q_k^{B_2C_1} \oplus Q_k^{C_1}$ to get the key $Q_k^{B_1}$, which is the final key shared with Bob. Since all QKD keys are used only once and are independent of each other, the third party can never access the keys he should not have.
Figure 4. The QKD network scheme with three endpoints based on the public-XOR-key scheme. The solid arrow stands for the QKD process. The dotted arrow indicates the key shared by each pair of endpoints. Relay node calculates and publishes $Q_{k1}^{A} = Q_{k1}^{A} \oplus Q_{k1}^{B}$, $Q_{k2}^{A} = Q_{k2}^{A} \oplus Q_{k2}^{C}$, and then dismisses the QKD keys immediately. Finally, Alice and Bob share the key $Q_{k1}^{A}$, Alice and Charlie share the key $Q_{k2}^{C}$, Bob and Charlie share the key $Q_{k2}^{B}$.

3.3. The Advantages of the Public-XOR-Key Scheme

The function of wide area key distribution network based on relay nodes is to distribute keys for multiple endpoints. The relay nodes are set up for helping the pairs of remote endpoints to share keys. Therefore, the less functions and devices the relay nodes have, the simpler the whole system is. In real implementation, our scheme has the following practical advantages.

1. The relay nodes do not need to store the QKD keys in our scheme. In the BBN key relay scheme, each relay node recovers the plaintext with QKD keys shared with the upstream node, encrypts it with QKD keys shared with the downstream node and then transfers it to the downstream node. Once the relay nodes have been broken, the adversary can access all plaintext transferred through this relay node. The relay nodes destroy the original QKD keys securely right after the QKD keys shared with upstream node and downstream node are immediately sent to the cipher machine to do XOR operations. The adversary has no chance to access the original QKD keys or plaintext. Compared with BBN scheme, our scheme reduces the storage of QKD keys and increases the security of the system.

2. The relay nodes do not need to store the XOR results of the QKD keys secretly in our scheme. Key buffers and corresponding preventive measures are needed for the secret storage, such as anti-electromagnetic radiation devices and video surveillance system. In our scheme, the XOR results are public. The way of publishing depends on the actual network. For example, the XOR results are saved in the relay nodes without protection, sent to endpoints directly, or broadcast to the users over the network. It is difficult for the actual networks to ensure the absolute security of the data storage. The system of our scheme simplifies the defence.

3. There are no re-generation keys in our scheme. In re-encryption scheme and BBN scheme, one of the endpoints firstly generates a local random number and encrypts the random number with the help of relay nodes. Finally, the pair of endpoints shares the random number. In our scheme, the keys shared by endpoints are exactly the QKD keys and there is no need to generate random numbers from point to point.

4. In our scheme, the shared keys do not need to be encrypted by each relay node over the path. When two endpoints want to share keys in re-encryption scheme, the ciphertext of the random
number generated by one endpoint is re-encryption and transferred by one relay node after another. There are multiple endpoints to communicate simultaneously in reality. Some relay nodes may be included in plenty of communications. The communications may hang or fail due to the heavy traffic. In our scheme, the relay nodes no longer participate in the communications after publishing XOR results. When endpoints want to share keys, either of the endpoints just calculates the key directly, which eases the traffic of the relay nodes.

4. The Security of the Public-XOR-Key Scheme

We described the complex QKD network scheme with multiple endpoints and analyzed its advantages in Section 3. Compared with the re-encryption scheme that does not announce the XOR value, our scheme can improve the actual performance and simplify the complexity of the system. In this section, we analyze our scheme from the security perspective. We analyze the security of the scheme by building the threat model of adversaries and strictly demonstrate its security under the threat model. As described in Section 3.2, there is always a chain between any two endpoints in all kinds of networks. So, we can use the chain QKD network as an example and build the security model. In addition, by comparing the security of our scheme with that of the re-encryption scheme, we prove that our scheme has the same security with that of the re-encryption scheme under the same assumption of adversary’s capacity, while our scheme achieves simplification in implementation.

4.1. The Security Analysis of the Public-XOR-Key Scheme

Threat model: we model the capacity of the adversary, Eve. Assume the destruction of the original QKD keys in the relay nodes is secure and Eve cannot access the original QKD keys even if she compromises the relay nodes. Assume the relay nodes never play the role as the internal attacker. QKD protocols have been proved to be unconditionally secure in theory. However, limited by the development of technology, there are attacks on QKD system in practice [28–32]. As a result of this defense, Eve cannot get all the information about QKD keys. However, there still exists some leaked information. In the chain network with $n$ point-to-point QKD systems, the amount of leaked information may be large. Assume leaked information about each point-to-point QKD system is the same, which is denoted as $Q_0$. The information obtained by Eve in the point-to-point QKD system is

$$I_0 = 1 - H(p),$$

where $H(p)$ is the entropy of QKD key $Q_k$ in each point-to-point QKD system. For $Q_k \in \{0, 1\}$, Eve knows that it has bias “0” or “1” with the probability of $p$. To simplify the following calculating, assume $p(Q_k = 0) = p$ for each $i = 1, 2, ..., n$, where $0 < p < 1$.

The amount of information about the shared key $Q_k$ obtained by Eve is

$$I_{\text{key}} = 1 - H(Q_k | a_i),$$

where $a_i$ is the public XOR value. A new parameter $b_i \in \{0, 1\}$ is defined from the public $a_i$,

$$b_i = a_1 \oplus a_2 \oplus ... \oplus a_i = Q_k \oplus Q_{k+1} \oplus Q_{k+2} \oplus ... \oplus Q_{k+n}, \quad i = 1, 2, ..., n - 1.$$

The number of elements $b_i$ is $n - 1$ and they form a Boolean vector $b$. The random variable of $b_i$ is defined as $B_i$, the random variable of $b \equiv (b_1b_2...b_{n-1})$ is $B \equiv (B_1B_2...B_{n-1})$, and the random variable of $Q_k$ is $Q_k$. $H(Q_k | a_1)$ can be treated as $H(Q_k | B)$ for more obvious analysis. That is, the information obtained by Eve is $I_{\text{key}} = 1 - H(Q_k | B)$. From the definition of $b_i$, it can be seen that the QKD key $Q_k$ is encrypted by other different QKD keys $n - 1$ times, which is similar to the random key $s$ encrypted by QKD keys in the re-encryption scheme.

Analyzing the security of the public-XOR-key scheme is equivalent to calculating how much information Eve obtains. We first analyze the conditional entropy of $Q_k$ in the threat model.
Theorem 1. In the assumption model of key distribution network using public-XOR-key scheme, the value of the conditional entropy $H(QK_1|B)$ is

$$H(QK_1|B) = \sum_{w=0}^{n-1} C_{n-1}^w \left[ p^{w}(1-p)^{n-w} + (1-p)^{n-w} p^w \right] H\left( \frac{p^{n-w}(1-p)^w}{p^{w}(1-p)^{n-w} + (1-p)^{n-w} p^w} \right).$$

(13)

where $p$ is the probability of $Qk_i = 0$ for all $i = 1, 2, ..., n$, and $C_{n-1}^w$ is the binomial coefficient.

Proof of Theorem 1. According to the definition of conditional entropy,

$$H(QK_1|B) = \sum_b p(B=b) H(QK_1|B=b) = -\sum_b \left[ p(B=b) \sum_{QK_1} p(QK_1 = Qk_1|B=b) \log p(QK_1 = Qk_1|B=b) \right].$$

(14)

Then we analyze the properties of the terms in the sum. For $Qk_1 \in \{0, 1\}$, the corresponding random variable satisfies that

$$p(QK_1 = 0) = p, \ p(QK_1 = 1) = 1 - p.$$

(15)

For $B \in \{0, 1\}^{n-1}$, there are $2^{n-1}$ values of $B$. Consider classifying $B$ by the number of zeros or ones. This kind of classification is exactly the hamming weight of the $B$, which is creative in cryptanalysis. Then Equation (14) becomes

$$H(QK_1|B) = -\sum_{QK_1} p(QK_1 = Qk_1|B=b) \sum_{QK_1} p(QK_1 = Qk_1|B=b) \log p(QK_1 = Qk_1|B=b)$$

(16)

$$= -\sum_{w=0}^{n-1} C_{n-1}^w p(B=b) \sum_{QK_1} p(QK_1 = Qk_1|B=b) \log p(QK_1 = Qk_1|B=b),$$

where $b_w$ denotes a particular vector with hamming weight $w$. According to Bayesian formula, the conditional probability of $QK_1 = 0$ in Equation (16) is

$$p(QK_1 = 0|B=b_w) = \frac{p(B=b_w|QK_1 = 0) \cdot p(QK_1 = 0)}{p(B=b_w)}.$$

(17)

According to the formula of total probability,

$$p(B=b_w) = p(B=b_w|QK_1 = 0) p(QK_1 = 0) + p(B=b_w|QK_1 = 1) p(QK_1 = 1).$$

(18)

Since $B_i = QK_i \oplus Qk_{i+1}$, the distribution of $B_i$ depends on $Qk_{i+1}$ and $QK_i$. When the random variable $QK_i = 0$, there is $B_i = Qk_{i+1}$. Then

$$p(B=b_w|QK_1 = 0) = p(QK' = b_w|QK_1 = 0) = p(QK' = b_w),$$

(19)

where $QK' = (QK_2, QK_3, ..., QK_n)$ is a random variable with the length of $n-1$. The difference between the random variables $QK$ and $QK'$ is that $QK'$ does not contain $QK_1$. Therefore, the random variables $QK'$ and $QK_1$ are independent of each other and the second equality of the above equation holds. Similarly, when $QK_1 = 1$, there is $B_i = Qk_{i+1} \oplus 1$. Then

$$p(B=b_w|QK_1 = 1) = p(QK' = \overline{b_w}|QK_1 = 1) = p(QK' = \overline{b_w}),$$

(20)
where $\vec{b} \equiv (\vec{b}_1 \ldots \vec{b}_i \ldots \vec{b}_n)$, $\vec{b}_i = b_i \oplus 1$. According to Equation (18),

$$p(B = b_w) = p \cdot p(QK' = b_w) + (1 - p) \cdot p(QK' = \overline{b_w}).$$  \hspace{1cm} (21)

When $QK' = b_w$, there is $w$ number of ones, and $n - w - 1$ number of zeros in the random variable $QK'$. When $QK' = \overline{b_w}$, there are $w$ zeros and $n - w - 1$ ones. Then

$$p(B = b_w) = p \cdot p(QK' = b_w) + (1 - p) \cdot p(QK' = \overline{b_w})$$  \hspace{1cm} (22)

$$= p^{n-w}(1-p)^w + p^w(1-p)^{n-w}.$$  

Therefore, the conditional probability of $QK_1 = 0$ is

$$p(QK_1 = 0|B = b_w) = \frac{p^{n-w}(1-p)^w}{p^{n-w}(1-p)^w + p^w(1-p)^{n-w}}.$$  \hspace{1cm} (23)

Similarly, the conditional probability of $QK_1 = 1$ is

$$p(QK_1 = 1|B = b_w) = \frac{p^w(1-p)^{n-w}}{p^{n-w}(1-p)^w + p^w(1-p)^{n-w}}.$$  \hspace{1cm} (24)

Then the conditional entropy $H(QK_1|B)$ is

$$H(QK_1|B) = \sum_{w=0}^{n-1} \left( C_{n-1}^w \left[ p^{n-w}(1-p)^w + (1-p)^{n-w} \right] H \left( \frac{p^{n-w}(1-p)^w}{p^{n-w}(1-p)^w + (1-p)^{n-w}} \right) \right).$$  \hspace{1cm} (25)

The equivocation $H(QK_1|B)$ measures the uncertainty of the QKD key $QK_1$ when given the public XOR keys. If and only if $p = 0$ or $p = 1$, the conditional entropy is null. Actually the security of point-to-point QKD systems guarantees that $0 < p < 1$. Therefore, $H(QK_1|B)$ cannot be null and Eve can never confirm the value of $QK_1$. The public-XOR-key scheme is secure against the adversary under the above assumption.

4.2. Comparison of the Security of the Public-XOR-Key Scheme and Re-Encryption Scheme

We have already analyzed the security of the public-XOR-key scheme. Now we contrast its security with that of re-encryption scheme. We first model the re-encryption scheme and analyze the information obtained by Eve in the scheme, then contrast it with the result in Section 4.

The endpoint transfers the message encrypted by local random key $s$. We denote the cryptosystem of QKD network using the re-encryption scheme as $(S, C, QK, E'', D'').$ A local random key $s$ is encrypted $n$ times with $QK_i$ by Alice and the relay nodes. To simplify the model, suppose the local random key, QKD keys, ciphertexts are all binary words with the length of 1. The encryption algorithm is the simplest XOR, that is

$$E''_{QK_i}(s) = s \oplus QK_i = c_i.$$  \hspace{1cm} (26)

The ciphertexts $c_i$, for $i = 1, 2, ..., n$ are available to Eve. The assumption of the capacity of Eve about the QKD keys is the same as the decryption of the threat model in Section 4.

For a chain network, which contains $n$ point-to-point QKD systems, the local random key $s$ is encrypted $n$ times with $QK_i$. The QKD keys can be analyzed as a Boolean vector $QK$ with the length of $n$. The ciphertext can be analyzed as a Boolean vector $c$ with the length of $n$. Let $QK, C \in \{0,1\}^n$, $S \in \{0,1\}$ denote the random variables of the vectors of QKD keys, ciphertexts and local random key, respectively. The equivocation of the local random key $H(S|C)$ measures the uncertainty of the local random key $s$ when given ciphertexts. When $H(S|C)$ is null, Eve gets the random key $s$. 
Theorem 2. In the assumption model of key distribution network using re-encryption scheme, the value of the conditional entropy $H(S|C)$ is

$$H(S|C) = \frac{1}{2} \sum_{w=0}^{n} \left( C^w_n \left[ p^{n-w}(1-p)^w + (1-p)^{n-w}p^w \right] H\left[ \frac{p^{n-w}(1-p)^w}{p^{n-w}(1-p)^w + (1-p)^{n-w}p^w} \right] \right).$$

(27)

Proof of Theorem 2. According to the definition of conditional entropy,

$$H(S|C) = \sum_c p(C = c) H(S|C = c)$$

$$= - \sum_c \left[ p(C = c) \sum_s p(S = s|C = c) \log p(S = s|C = c) \right].$$

(28)

Then we analyze the properties of the terms in the sum. In order to guarantee the randomness of the local random key $s$, there is

$$p(S = 0) = p(S = 1) = \frac{1}{2}.$$  

(29)

For the Boolean vector $C \in \{0,1\}^n$, there are $2^n$ values. Classify $C$ by the hamming weight. Equation (28) becomes

$$H(S|C) = - \sum_c \left[ p(C = c) \sum_s p(S = s|C = c) \log p(S = s|C = c) \right],$$

(30)

where $c_w$ denotes a particular vector of length $n$ with hamming weight $w$. According to Bayesian formula, when $C = c_w$, the conditional probability of $S = 0$ in Equation (30) is

$$p(S = 0|C = c_w) = \frac{p(C = c_w|S = 0) \cdot p(S = 0)}{p(C = c_w)}.$$  

(31)

Since $c_i = s \oplus Qk_i$, the distribution of $c_i$ depends on $Qk_i$ and $s$. When random variable $S = 0$, there is $QK_i = C_i$. Then

$$p(C = c_w|S = 0) = p(QK = c_w|S = 0) = p(QK = c_w) = p^{n-w}(1-p)^w.$$  

(32)

When random variable $S = 1$, there is $QK_i = C_i \oplus 1$. Then

$$p(C = c_w|S = 1) = p(QK = \overline{c_w}|S = 1) = p(QK = \overline{c_w}) = p^w(1-p)^{n-w}.$$  

(33)

In Equations (32) and (33), the second equality holds because the random variables $QK$ and $S$ are independent of each other. According to the formula of total probability,

$$p(C = c_w) = p(C = c_w|S = 0) p(S = 0) + p(C = c_w|S = 1) p(S = 1)$$

$$= \frac{1}{2} p^{n-w}(1-p)^w + \frac{1}{2} p^w(1-p)^{n-w}.$$  

(34)
From Equations (32) and (34), the conditional probability of $S = 0$ is

$$p(S = 0|C = c_w) = \frac{p^{n-w}(1-p)^w}{p^{n-w}(1-p)^w + p^w(1-p)^{n-w}}. \quad (35)$$

Similarly, the conditional probability of $S = 1$ is

$$p(S = 1|C = c_w) = \frac{p^w(1-p)^{n-w}}{p^{n-w}(1-p)^w + p^w(1-p)^{n-w}}. \quad (36)$$

Therefore, the conditional entropy $H(S|C)$

$$H(S|C) = \frac{1}{2} \sum_{w=0}^{n} \left[ C_n^w \left( \frac{p^{n-w}(1-p)^w}{p^{n-w}(1-p)^w + p^w(1-p)^{n-w}} \right) H \left( \frac{p^{n-w}(1-p)^w}{p^{n-w}(1-p)^w + p^w(1-p)^{n-w}} \right) \right]. \quad (37)$$

In this model, the information about the Alice’s random key $s$ obtained by Eve is $I_s = 1 - H(S|C)$. Then we contrast the information leakage of the public-XOR-key scheme and re-encryption scheme.

**Theorem 3.** For the adversary with the same capacity, the information leakage of re-encryption scheme and public-XOR-key scheme are the same, i.e., $I_{key} = I_s$.

**Proof of Theorem 3.** Define two functions $\Delta$ and $f_p(w)$ as follows,

$$\Delta \equiv I_{key} - I_s = H(S|C) - H(QK_1|B),$$

$$f_p(w) = \left[ \frac{p^{n-w}(1-p)^w}{p^{n-w}(1-p)^w + (1-p)^{n-w}p^w} \right] \cdot H \left( \frac{p^{n-w}(1-p)^w}{p^{n-w}(1-p)^w + (1-p)^{n-w}p^w} \right).$$

It can be seen that $f_p(i) = f_p(n-i)$. The difference between $I_{key}$ and $I_s$ is

$$\Delta = H(S|C) - H(QK_1|B)$$

$$= \sum_{w=0}^{n} \left[ \frac{1}{2} C_n^w f_p(w) \right] - \sum_{w=0}^{n-1} \left[ C_n^{w-1} f_p(w) \right]$$

$$= \left( \frac{1}{2} C_n^0 - C_{n-1}^0 \right) f_p(0) + \sum_{w=1}^{n-1} \left( \frac{1}{2} C_n^w - C_{n-1}^w \right) f_p(w) + \frac{1}{2} C_n^n f_p(n) \quad (38)$$

$$= -\frac{1}{2} f_p(0) + \frac{1}{2} f_p(n) + \sum_{w=1}^{n-1} \left( \frac{1}{2} C_n^w - C_{n-1}^w \right) f_p(w)$$

$$= \sum_{w=1}^{n-1} \left( \frac{1}{2} C_n^w - C_{n-1}^w \right) f_p(w).$$

Sum the $i$th term and the $(n-i)$th term in Equation (38),

$$\sum_{w=i,n-i} \left( \frac{1}{2} C_n^w - C_{n-1}^w \right) f_p(w)$$

$$= \left( \frac{1}{2} C_n^i - C_{n-1}^i \right) f_p(i) + \left( \frac{1}{2} C_n^{n-i} - C_{n-1}^{n-i} \right) f_p(n-i) \quad (39)$$

$$= \left[ \left( \frac{1}{2} C_n^i - C_{n-1}^i \right) + \left( \frac{1}{2} C_n^{n-i} - C_{n-1}^{n-i} \right) \right] f_p(i).$$
we proposed is equivalent to that of the re-encryption scheme. The public-XOR-key scheme cannot be treated as the key in Shannon’s model, the QKD keys of the system. The unicity distance as the plaintext in Shannon’s model. The unicity distance, where $H_{\text{c}}(p) = H(1 - H(p))$, where $H(p)$ is the entropy of the QKD key in each point-to-point QKD system. It seems reasonable that the physical interpretation of $H(p)$ and $H_{\text{c}}$, $I_0$ and $R_L$ are equivalent. When $n$ is sufficiently large, according to Equation (4), the relationship between $H(p)$ and the quantum keys with the length of $n$ is

$$H(K^n) \approx nH(p).$$
The unicity distance $N'_0$ of the relay QKD system is

$$N'_0 \approx \frac{1}{I_0} = \frac{1}{1 - H(p)}, \quad (45)$$

which means that if Eve attacks $N'_0$ point-to-point QKD systems and gets $N'_0$ ciphertext, she can obtain the random key $s$. The contrast between the parameters of Shannon’s model and those of the model of the QKD network is shown in Table 1.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Shannon’s Model</th>
<th>Model of QKD Network</th>
</tr>
</thead>
<tbody>
<tr>
<td>Keys</td>
<td>$K$</td>
<td>$S$</td>
</tr>
<tr>
<td>Plaintext</td>
<td>$M$</td>
<td>$QK$</td>
</tr>
<tr>
<td>Ciphertext</td>
<td>$C$</td>
<td>$C$</td>
</tr>
<tr>
<td>Rate of language</td>
<td>$H_L$</td>
<td>$H(p)$</td>
</tr>
<tr>
<td>Redundancy of language</td>
<td>$R_L$</td>
<td>$I_0$</td>
</tr>
<tr>
<td>Unicity distance</td>
<td>$1/R_L$</td>
<td>$1/I_0$</td>
</tr>
</tbody>
</table>

However, in Sections 4 and 5 we show that the conditional entropy $H(S|C)$ cannot be null as long as $0 < p < 1$, which reflects that Eve cannot obtain the exact key $s$ no matter how many ciphertexts she has in the real world. This result conflicts with the unicity distance that was analyzed above.

We hold the view that there are two reasons for the difference. Firstly, although $H(p)$ and $H_L$, $I_0$ and $R_L$ have very similar physical interpretation, actually they are not equivalent. In Shannon’s model, $H_L$ denotes the rate of each word in language $L$, which is an average value with a sufficiently large $n$. In the model of QKD network, $H(p)$ is the real entropy of the QKD key in each point-to-point QKD system rather than an average value. Since each QKD key encrypts the same random key $s$, with the known $c_i$ each point-to-point QKD system is not completely independent. Suppose the rate of each point-to-point QKD system is $H_K$, $H_K = \lim_{n \to \infty} \frac{H(K^n)}{n} \neq H(p). \quad (46)$

The second reason is that the unicity distance $N_0$ is an approximate and minimum value in theory. In the derivation of the value of unicity distance, there are some approximations and inequalities. When $n$ is sufficiently large, Equation (4) holds. In practical QKD network, $n$ cannot be large enough, the value of which is usually just several dozens. Equation (8), denoting the approximation of $N_0$, actually is its minimum value. Even if Eve gets the ciphertext with the length of $H(K)/R_L \log_2 |M|$, there still exist spurious keys and she cannot get the real key.

Through the above analysis, we find the reasons why the unicity distance of the practical QKD network is different from that in theory. It indicates that the theoretical value of unicity distance has a certain gap with the practical system. In addition, the unicity distance of the relay QKD system is infinity, which indicates the adversary in the model can never obtain the entire secret message.

6. Discussion and Conclusions

Considering the key rate and the requirements of multiple-endpoint long-distance communication, the public-XOR-key scheme is more efficient and practical than the BBN protocol, re-encryption protocol and STR. In this paper, we extend the chain structure of the public-XOR-key scheme with two endpoints to the complex network with multiple endpoints. In the complex network with multiple endpoints, the practical contributions are as follows: (1) The relay nodes do not need to store the QKD keys. (2) The relay nodes do not need to store the XOR results of the QKD keys secretly, which reduces the storage. (3) The relay nodes do not participate in the transmission of the key, which
reduces the communication traffic to a certain extent. A particularly important contribution is the proof of the security of the scheme. On the one hand, compared with the security of re-encryption key distribution scheme whose XOR keys are not announced, the security of the key distribution scheme we proposed and the re-encryption scheme is equivalent. Our scheme can not only improve the practical performance and simplify the system’s complexity, but also ensure that the security is not reduced. Even though the adversary obtains partial information about the QKD keys in our scheme, she cannot confirm the shared key. On the other hand, we discuss the information leakage of the practical key distribution networks from the perspective of the unicity distance. The unicity distance of the practical QKD network is infinity, which bears out that our scheme is secure and Eve can never obtain the shared key.

The major advantages of our scheme are as follows.

- The superiority of the practical performance: (i) The relay nodes need not store the QKD keys. Even if Eve attacks the relay nodes, she cannot access the QKD keys and plaintext. The scheme reduces the storage of QKD keys and increases the practical security of the system. (ii) The relay nodes need not store the XOR results of the QKD keys secretly, which reduces the storage. It is difficult for actual network to ensure the absolute security of the data storage. The secret storage is avoided in our scheme. (iii) There is no re-generation keys in our scheme. In our scheme, the keys shared by endpoints are exactly the QKD keys and it is no need to generated random numbers from point to point. (iv) The relay nodes do not participate in the transmission of the key, which reduces the communication traffic to a certain extent.

- The superiority of security: (i) we analyze the information leakage of public-XOR-key scheme and re-encryption scheme in the threat model and prove the security of the schemes. Even though the adversary obtains partial information about the QKD keys, she cannot confirm the shared key. (ii) We discuss the information leakage of the practical key distribution networks from the perspective of the unicity distance. The unicity distance of the practical QKD network is infinity, which bears out that our scheme is secure and Eve can never obtain the shared key. The security analysis of complex network structures are worthy for the further study.

In addition, we analyze the reason why the unicity distance of the practical system is different from that in theory. In fact, this system is a typical example of that the same plaintext is encrypted with different keys. It can be analyzed as Shannon’s ciphertext-only attack model because the encryption and decryption algorithms are XOR, where the order of operations can be exchanged. The unicity distance of the systems with other complicated encryption and decryption algorithms is left as an open question for future work.

Author Contributions: Conceptualization, L.Y.; methodology, L.Y.; validation, H.D. and Y.S.; formal analysis, H.D. and Y.S.; investigation, H.D. and Y.S.; writing—original draft preparation, Y.S.; writing—review and editing, H.D.; project administration, L.Y.

Funding: This research was funded by National Natural Science Foundation of China under Grant No. 61672517 and National Cryptography Development Fund under Grant No. MMJJ20170108.

Conflicts of Interest: The authors declare no conflict of interest.

References


© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).