A Security Real-time Privacy Amplification Scheme in QKD System

Bo Liu, Baokang Zhao*, Bo Liu, Chunqing Wu
(National University of Defense Technology, Hunan, China
liub0yayu@gmail.com, {bkzhao, boliu, chunqingwu}@nudt.edu.cn)

Abstract: Quantum Key Distribution (QKD) technology, based on the laws of physics, can create unconditional security keys between communication parties. In recent years, researchers draw more and more attention to the QKD technology. Privacy amplification is a very significant procedure in QKD system. In this paper, we propose the real-time privacy amplification (RTPA) scheme which converts the weak secret string to a uniform key that is fully secret from Eve. Our detailed proofs show the security of our RTPA scheme. In order to prevent the potential man-in-middle attacks, we employ an authentication procedure to RTPA scheme (ARTPA) with the ε-XOR almost universal hash functions. We implement our ARTPA scheme based on CLIP system, which is connected to the quantum communication system. Considering the privacy amplification and authentication overhead and the finite size effect on the security of final keys, we set the secret key length be 256k before privacy amplification and the authentication tag length be 60. Our experimental results show the efficiency of the proposed ARTPA scheme.

Key Words: privacy amplification, authentication, quantum key distribution, security

Category: H.1.0, H.1.1, E.4

1 Introduction

Quantum Key Distribution (QKD) technology is an important practical application of quantum information [Bennett and Brassard, 1984], [Ekert, 1991], [Gisin et al., 2002]. QKD system can create unconditionally secure keys between communication parties. The security of generated keys from QKD systems is based on the laws of physics, rather than the computational complexity of mathematical problems assumed by current cryptography methods [Mayers, 2001], [Biham et al., 2006], [Shor and Preskill, 2000]. Since the first QKD scheme is proposed by Bennett and Brassard in 1984 (called BB84 protocol), QKD has attracted much research [Heid et al., 2007], [Qi et al., 2007], [Lo et al., 2012]. QKD protocols typically can be divided into two phases: the quantum phase and the classical phase [Ma et al., 2009], [Ma et al., 2011]. Quantum signals are distributed and measured in the quantum phase, which generates correlated data shared between Alice and Bob. The classical phase aims to gain the identical and private keys between Alice and Bob by classical communication protocols. The

*The corresponding author: Dr. Baokang Zhao(bkzhao@nudt.edu.cn).
classical phase is also called QKD post-processing phase. QKD post-processing procedures include three steps: data pre-processing, error correction and privacy amplification [Ma and Lütkenhous, 2012]. After the procedure of error correction [Brassard and Salvail, 1994], Alice and Bob have own a weak uniform key $W$. But Eve may have partial knowledge about the keys by eavesdropping or other ways.

Therefore, in order to gain the absolute security keys, we must ensure the keys are privacy amplified. Privacy amplification (PA) [Bennett et al., 1995] is a technology, through a public channel, to improve the information confidentiality. Privacy amplification converts the weak secret string $W$ to a uniform key $R$ that is full secret from Eve.

Privacy amplification technology typically uses a random and public hash function to shorten the weak secret key and reduce the amount of information obtained by Eve as much as possible. By sacrificing partial key information of Alice and Bob, privacy amplification makes the knowledge obtained by Eve be meaningless.

Though majority research about privacy amplification focusing on theoretical study and proof of security ([Renner and König, 2005], [Watanabe, 2007], [Horváth et al., 2011], [Fung et al., 2012], [Singh et al., 2012]), implementing an efficient privacy amplification scheme in QKD systems has been more and more significant. In order to continuously gain security keys from QKD systems, a real-time privacy amplification scheme must be implemented. Hash functions are transmitted through public channels in privacy amplification procedure. In order to prevent the potential man-in-middle attacks, authentication must be employed to the public hash functions.

In this paper, we focus on the design and implementation of the real-time privacy amplification scheme in QKD systems. We propose a real-time privacy amplification scheme (RTPA) and implement RTPA based on CLIP system [Zou et al., 2012], which is connected to the quantum communication system. Our quantum communication system is a robust two-way quantum key distribution system based on phase encoding which conducts BB84 protocol [Sun et al., 2010]. In order to enhance the security of our RTPA scheme, we employ an authentication scheme (ARTPA) with the $\varepsilon$-XOR almost universal hash functions [Krawczyk, 1995]. Our experimental results show the efficiency of the proposed RTPA scheme for generating unconditional security keys in quantum cryptography.

The rest of the paper is organized as follows. We give a brief introduction to privacy amplification, Shannon entropy, Rényi entropy [Renner and Wolf, 2004] in Section 2. We propose the RTPA scheme in Section 3. We present the security analysis of RTPA scheme in Section 4. The RTPA scheme with authentication procedure is given in Section 5, followed by the experimental results and perfor-
2 Preliminaries

Privacy amplification technology usually applies two-universal hash functions to shorten the weak secret keys and reduce the amount of information obtained by Eve. In this section, we will give a brief introduction to two-universal hash functions, Shannon entropy and Rényi entropy.

2.1 Two-Universal Hash Function

Universal Hash function has been widely used for data encryption, authentication and other security protection scenes [Lee et al., 2010], [Singh et al., 2013], [Yanai et al., 2012], [Claycomb et al., 2011]. In this paper, we focus on such two-universal hash functions described as follows: $f$ is called a two-universal hash function if 
\[
\Pr[f(x_a) = f(x_b) | x_a \neq x_b] < \frac{1}{2^r},
\]
where $f \in F$, $F : X \rightarrow Y$, $X = \{0,1\}^n$, $Y = \{0,1\}^r$, $n > r$, $x_a \in X$ and $x_b \in X$. 
$Pr[f(x_a) = f(x_b) | x_a \neq x_b]$ is the collision probability when a random string $X$ is mapped to $Y$. $f$ is two-universal that means, for any distinct $x_a, x_b \in X$, the random variables $f(x_a)$ and $f(x_b)$ are independent and uniformly distributed [Renner and König, 2005].

2.2 Shannon Entropy and Rényi Entropy

As we have known, the information can be measured in exactly the same manner as the uncertainty is measured.

The widely accepted measure of information is the one introduced by Shannon. Shannon entropy measures the uncertainty of a random variable. The Rényi entropy is a one-parameter generalization of the Shannon entropy.

Let $X$ and $Y$ be random variables over finite set $\{0,1\}^n$ and $\{0,1\}^r$. Then the Shannon entropy of $X$, $H(X)$, and the Rényi entropy of $X$, $R(X)$, are defined by

\[
H(X) = - \sum_{x \in X} \Pr[X = x] \log_2 \Pr[X = x], \tag{2.1}
\]

\[
R(X) = - \log_2 \sum_{x \in X} \Pr[X = x]^2. \tag{2.2}
\]

The conditional Shannon entropy of $X$ given $Y = y$, $H(X|Y = y)$, and the Shannon entropy of $X$ conditioned on $Y$, $H(X|Y)$, are defined by

\[
H(X|Y = y) = - \sum_{x \in X} \Pr[X = x|Y = y] \log_2 \Pr[X = x|Y = y], \tag{2.3}
\]
\[ H(X|Y) = - \sum_{x \in X} \Pr[Y = y] H(X|Y = y). \]  

(2.4)

The conditional Rényi entropy of \( X \) on \( Y \), \( R(X|Y) \), is defined by

\[ R(X|Y) = - \sum_{y \in Y} \Pr[Y = y] R(X|Y = y). \]  

(2.5)

The mutual information [Ribeiro et al., 2008] between \( X \) and \( Y \), \( I(X,Y) \) is defined by

\[ I(X,Y) = H(X) - H(X,Y). \]  

(2.6)

Assume that, Alice sends a random string \( X \) to Bob through quantum channel, Bob receives a random string \( Y \) and Eve gets a random string \( Z \) by eavesdropping or other ways. After the procedure of privacy amplification, we can get

\[ I(X,Z) \approx 0. \]  

(2.7)

3 The Proposed RTPA Scheme

As we researched both classical and quantum privacy amplification theoretical study in [Bennett et al., 1995], [Dodis et al., 2009], [Chandran et al., 2010], [Horváth et al., 2011] and our previous work in [Liu et al., 2013], we approach the RTPA scheme (Real-Time Privacy Amplification scheme) in QKD systems.

In this section, we begin by providing some notations which will be used later. Then, we give the detailed description of the RTPA scheme.

3.1 Notations

We assume that Alice sends a random key string \( X \) to Bob through quantum channel. \( Y \) is the string received by Bob. After the procedures of data preprocessing and error correction, Alice and Bob own the same weak security key string \( W \), and its length is \( n \). Random variable \( V \) means the information about keys obtained by Eve, and its length is \( t \). The security parameter is \( s \) and the final key string is \( R \), and its length is \( l \).

In order to generate the absolute security keys, we have

\[ l = n - t - s. \]  

(3.1)

In this paper, we assume that the security parameter \( s \) is 30.
Algorithm 1 Privacy amplification algorithm

Input: weak security key \( W \) with length \( n \), random string \( a \) with length \( n+l-1 \)

Output: final key \( R \) with length \( l \)

1: for \( i = 0; \ i < l; \ i++ \) do
2: \( r_i = 0; \)
3: for \( j = 0; \ j < n; \ j++ \) do
4: \( r_i = r_i \oplus w_j a_{l-1-i+j} \)
5: end for
6: end for

3.2 Hash Function Construction

In RTPA scheme, the key step is the hash function construction. The hash function \( f \) is randomly chosen from class \( H_3 \) of two-universal hash functions. The hash function construction method is described by Toeplitz matrix construction method [Mansour et al., 1993], [Krawczyk, 1994], [Fung et al., 2010].

The Toeplitz matrix can be defined as

\[
H_r = \begin{bmatrix}
  a_0 & a_{-1} & a_{-2} & \cdots & a_{-n+1} \\
  a_1 & a_0 & a_{-1} & \cdots & \vdots \\
  a_2 & a_1 & \ddots & \vdots & \vdots \\
  \vdots & a_2 & \ddots & \ddots & \vdots \\
  a_{l-1} & \cdots & a_{l-n} & \cdots & a_0
\end{bmatrix}_{n \times l}, \tag{3.2}
\]

where \( a_i \in \{0, 1\} \). A construction of a Toeplitz matrix \( H_r \) is two-universal and it requires a random seed string \( a \) with length of \( n + l - 1 \) to define.

We can gain the final secure key by

\[
R = H_r W. \tag{3.3}
\]

The whole procedure of privacy amplification based on Toeplitz matrix construction method is given in algorithm 1.

3.3 The Proposed RTPA Scheme

The proposed RTPA Scheme consists of the following three phases: initialization phase, data communication phase and key generation phase.

1. Initialization Phase:

Based on the QKD system and the RTPA scheme requirements, the following steps should be performed to bootstrap the RTPA scheme.
– Alice and Bob get the privacy amplification input key string $W$ after error correction procedure.

– Alice and Bob calculate the security parameter $s$, according to the quantum bit error rate and other information.

– Alice and Bob determine the final key length $l$.

– Alice generates a random seed string $a$ with length of $n + l - 1$, which is used for hash function construction.

2. Data Communication Phase:
Alice transmits the random seed string $a$ to Bob over the public channel. Since the seed string is not encrypted, Eve may get the hash function totally. But this affects nothing about the security of the final key.

3. Key Generation Phase:
The hash function construction and privacy amplification method has been detailed in Section 3.2. In this phase, Alice and Bob employ $R = H_rW$ to get the final keys.

3.4 The Implementation of RTPA Scheme

![Figure 1: The architecture of RTPA scheme](image)

The architecture of RTPA scheme is given in Fig. 1. The implementation of RTPA scheme mainly consists of five modules: parameter controller, RNG module, seed generation module, privacy amplification module and communication module.
The parameter controller carries out the security parameter $s$ and controls the generation of the seed string $a$ and final key $R$.

The RNG module (Random Number Generation module) generates real random numbers for seed generation module.

The seed generation module provides a random seed string $a$ for privacy amplification module and communication module.

4 Security Analysis of RTPA Scheme

Since we have given the brief introduction about two-universal hash functions, Shannon entropy and Rényi entropy in Section 2, we will give the security analysis of RTPA scheme in this section.

**Definition 1 (Security).** A privacy amplification scheme is secure with $\xi$ if the mutual information between Alice and Eve is less than $\xi$ after the procedure of privacy amplification.

In our RTPA scheme, with the security parameter $s$, we can get $\xi \leq 2^{-s}/\ln 2$.

The detail security proofs are given as follows.

From (2.1) and (2.2), we can get

$$R(X) \leq H(X).$$  \hfill (4.1)

Combining (2.4), (2.5) and (4.1), we would have that

$$R(X|Y) \leq H(X|Y)$$  \hfill (4.2)

Suppose that $F$ is a random variable, with distribution probability of $p_F(F = f) = p_F(f)$, where $f$ is a two-universal hash function, $F : \{0, 1\}^n \rightarrow \{0, 1\}^r$.

From (2.5), we would have that

$$R[F(X)|F] = \sum_F p_F(f) R[F(X)|F = f]$$
$$= \sum_F p_F(f) \{-\log_2 p_c[F(X)|F = f]\}$$  \hfill (4.3)

where $p_c[F(X)|F = f]$ is the collision probability of $F(X)$ on the condition of $F = f$. $p_c[F(X)|F = f]$ can be defined as

$$p_c[F(X)|F = f] = \sum_{F(X)} \{\Pr[F(X)|F = f]\}^2.$$  \hfill (4.4)

Now, we can get

$$\sum_F p_F \{-\log_2 p_c[F(X)|F = f]\}$$
$$\geq -\log 2 \{\sum_F p_F(f)p_c[F(X)|F = f]\}.$$  \hfill (4.5)
Then, we have
\[
\sum_f p_F(f)p_c[F(X)|F = f] 
\leq p_c(X) + [1 - p_c(X)]2^{-r}.
\]
Because \(p_c(X) = \sum_{x \in X} \Pr[X = x]^2\), combing with (2.2), we can get
\[
p_c(X) = 2^{-R(X)}.
\]
From (4.6) and (4.7), we would have that
\[
-\log 2\{\sum_f p_F(f)p_c[F(X)|F = f]\} 
\geq r - \log_2[1 + 2^{r-R(X)}].
\]
As we have known that \(\ln z \leq z - 1, z > 0\), combing with (4.2), (4.3) and (4.8), we can get
\[
H[F(X)|F] \geq R[F(X)|F] \geq r - \frac{2^{r-R(X)}}{\ln 2}.
\]
In (4.9), let \(X\) be the key string owned by Alice and Bob after the procedure of error correction. \(F(X) = R\) means the final key string of Alice and Bob after the procedure of privacy amplification. Assume that Eve obtains the information string about keys is \(V\), with length of \(t\).
As we have known that \(r = n - t - s > 0\).
Suppose a hash function \(e : \{0,1\}^n \to \{0,1\}^t\) and \(v = e(x)\), where \(v \in V, x \in X\). Then, we would have
\[
\Pr(V = v) = 2^{-t},
\]
\[
\Pr(X|V = v) = 2^t/2^n,
\]
\[
p_c(X|V = v) = \sum_x [\Pr(X|V = v)]^2 
= \Pr(X|V = v),
\]
\[
R(X|V = v) = -\log 2p_c(X|V = v).
\]
From (4.10), (4.11), (4.12) and (4.13), we can get
\[
H[R|F, V = v] \geq r - \frac{2^{r+t-n}}{\ln 2},
\]
\[
H[R|FV] \geq r - \frac{2^{r+t-n}}{\ln 2} = r - \frac{2^{-s}}{\ln 2}.
\]
Because

\[ I(R, FV) = H(R) - H(R|GV), \]  

and \( H(R) = r \), we would have

\[ I(R, FV) \leq \frac{2^{-s}}{\ln 2}. \]  

(4.17)

\( I(R, FV) \) is the mutual information between Eve and Alice after the procedure of privacy amplification. It is less than \( 2^{-s}/\ln 2 \).

In all, the proposed RTPA scheme is security with \( \xi \leq 2^{-s}/\ln 2 \).

5 RTPA Scheme With Unconditional Security Authentication

We have proved the security of our RTPA scheme detailed in Section 4. As we have known, Alice and Bob exchange data over public channels. Eve may not only eavesdrop but also modify the transmitted data. When Alice transmits a seed string to Bob, it can be known to Eve totally. In order to ensure the integrity of the privacy amplification procedure, we must protect the integrity of the seed string.

In this section, we start with the basic concept about unconditional security authentication technology in QKD systems. Then, we propose the RTPA scheme with unconditionally secure authentication.

5.1 Unconditional Security Authentication in QKD System

As we have known, Alice and Bob exchange messages through public channel in RTPA scheme. It needs unconditional security authentication in order to thwart man-in-middle attacks [Bennett and Brassard, 1984], [Abidin, 2010]. Unconditional security authentication is also called ITS authentication (Information-Theoretically Security).

In paper [Carter and Wegman, 1979] and [Wegman and Carter, 1981], Wegman and Carter have proposed several strongly universal hash families which can be used in authentication (we call it as WCA in this paper). The security of the WCA scheme as used in QKD was studied in [Cederlof and Larsson, 2008]. In this paper, we employ WCA scheme to ensure the security of RTPA.

The WCA scheme conducts a \( \varepsilon \)-XOR almost universal hash function to generate authentication tags.

**Definition 2 (\( \varepsilon \)-XOR almost universal hash function).** A hash function \( f \) is \( \varepsilon \)-XOR almost universal if for any input \( x_a \neq x_b \) and output \( y \), \( \Pr_{f \in F}[f(x_a) \oplus f(x_b) = y] \leq \varepsilon \).

It has been proved that we can use simplified Toeplitz matrices generated by a LFSR to construct a WCA authentication scheme [Krawczyk, 1994], [Krawczyk, 1995].
5.2 Simplified Toeplitz matrices construction

Assume that the scale of simplified Toeplitz matrix $H_s$ is $m \times p$, $m > p$. The seed string $a'$ construction method can be described in Algorithm 2.

Algorithm 2 LFSR based Toeplitz matrix seed string construction algorithm

**Input:** seed string $h$ with length $p$, random string $c$ with length $p$, $m$

**Output:** seed string $a'$ with length $m + p - 1$

1: for $i = 0$; $i < p$; $i + +$ do
2: $a'_i = h_i$
3: end for
4: for $i = p$; $i < m + p$; $i + +$ do
5: for $j = 0$; $j < p$; $j + +$ do
6: $b = (a'_{i-p+j} \& c_j) \oplus b$
7: end for
8: $a'_i = b$
9: end for

5.3 RTPA Scheme With Unconditional Security Authentication

As we have discussed in Section 5.1 and 5.2, we propose the RTPA scheme with unconditional security authentication (we call it as ARTPA in this paper), which is shown in Fig. 2.

In ARTPA scheme, we have to generate an authentication tag $t$ for the seed string $a$, suppose the authentication hash function is $H_s$, the length of $t$ is $p$, we can get

$$t = H_s a \oplus r,$$

where $r$ is the secret key string with length of $p$.

Suppose that the length of $a$ is $m$, we have that

$$p = s + 1 + \log_2 m,$$

where $s$ is the security parameter.

Also, we will have

$$\varepsilon \leq m/2^{p-1}.$$

The proposed ARTPA scheme can be divided into three phases: authentication tag generation phase, data communication phase and authentication tag confirmation phase.
Let $h$ be the seed string used for constructing the authentication hashing function and its length is $2p$.

1. **Authentication Tag Generation Phase**
   Alice generates a random seed string $a$ by RTPA module used for privacy amplification.
   Alice gets $2p$ bits length seed string from the key pool to construct the authentication hash function $H_s$, gets $p$ bits length secret key string to encrypt the tag message and conducts (5.1) to calculate the tag $t$.

2. **Data Communication Phase**
   Alice transmits the random seed string $a$ and authentication tag $t$ to Bob over the public channel. Since the seed string is authenticated and the authentication tag is encrypted by one-time pad method, which avoids the man-in-middle attacks like modifying or confusing the seed string.

3. **Authentication Tag Confirmation Phase**
   Bob gets $2p$ bits length seed string from the key pool to construct the authentication hash function $H_s$, gets $p$ bits length secret key string to encrypt the tag message and conducts (5.1) to calculate the tag $t'$. By comparing $t$ and $t'$, Bob will know whether Eve modifies or confuses the seed string $a$ or not.
6 Performance Evaluation

6.1 Experiment Setup

To evaluate the performance of ARTPA scheme, we implemented our schemes based on CLIP [Zou et al., 2012], which is connected to the quantum communication system [Sun et al., 2010]. Our quantum communication system conducts BB84 protocol. CLIP provides the weak security keys for our scheme. The performance evaluation metrics are the privacy amplification and authentication overhead, fraction of secret key removed by authentication and throughput. The experiment environment is shown in Fig. 3.

![Figure 3: The experiment environment of ARTPA](image)

6.2 Privacy Amplification Overhead

Various hash function constructed for different input key lengths, will lead to different time overhead per privacy amplification process. While the input key length should be long enough in order to gain an absolute security key, the privacy amplification overhead will be very high. In this experiment, we focus on the privacy amplification calculate time of different hash function scale. The result is shown in Fig. 4.

The complexity of the privacy amplification algorithm is $O(n^2)$. As shown in Fig. 4, it will cost 100 times more overhead when the hash function scale increase 10 times.

Considering the finite size effect on the security of final keys, we conduct the ARTPA scheme with the hash function scale of $10^5$ length, though it costs little time when the hash function scale is less than $10^5$. 
6.3 Authentication Overhead

In this test, we focus on the authentication overhead with different tag lengths. We set the seed string length for authentication be $10^5$, and the authentication security parameter $s = 10, 20, \cdots, 60$. From (5.2), we get the tag length $p = 30, 40, \cdots, 80$. The experimental results are shown in Fig. 5.

From (5.3), we can get $\varepsilon \leq 9.3 \times 10^{-4}$ with the tag length $p = 30$, $\varepsilon \leq 8.8 \times 10^{-13}$ with $p = 60$ and $\varepsilon \leq 8.5 \times 10^{-19}$ with $p = 80$. We know that the authentication is more effective with longer tag lengths. But with longer tag length, it will cost more overhead. For example, when $p = 80$, the overhead is 687.104ms, 171.761ms higher than the overhead with $p = 60$.

Considering both the security and authentication overhead, we set $p = 60$ at last.
6.4 Fraction of Secret Key Removed by Authentication

The length of secret key used for authentication is $3p$. When we set $p = 60$, the secret key loss is 180 bits. In Fig. 6, we give the fraction of secret key removed by authentication. In our ARTPA scheme, the secret key length is $2.56 \times 10^5$ and the fraction is less than $1.8 \times 10^{-3}$.

6.5 Throughput

In the throughput test, the secret key length $n$ before privacy amplification changes from $10^3$ to $10^6$. The authentication tag length $p$ is 60. The experiment result is shown in Fig. 7.
As shown in Fig. 7, the authentication procedure affects little in the throughput tests. The throughput of ARTPA scheme is 150.3kbps with \( n = 10^4 \). The throughput decreases to 6.97kbps quickly when \( n = 10^6 \). The throughput is 31.2kbps with \( n = 2.56 \times 10^5 \), which is conducted in our ARTPA scheme.

7 Conclusion

In this paper, we have designed and implemented a real-time privacy amplification scheme (RTPA) in QKD system. Our detailed proofs show the security of our RTPA scheme. In order to prevent the potential man-in-middle attacks, we employ an authentication procedure to RTPA scheme (ARTPA) with \( \varepsilon \)-XOR almost universal hash functions. We implement our ARTPA scheme based on CLIP system, which is connected to the quantum communication system. Considering the privacy amplification and authentication overhead and the finite size effect to the security of final keys, we set the secret key length before privacy amplification be \( 2.56 \times 10^5 \) and the authentication tag length be 60. Our experimental results show the efficiency of the proposed ARTPA scheme for generating unconditional security keys in quantum cryptography.

Acknowledgements

The work described in this paper is partially supported by the grants of the National Basic Research Program of China (973 project) under Grant No.2009CB320503, 2012CB315906; the project of National Science Foundation of China under grant No. 61070199, 61103189, 61103194, 61103182, 61202488, 61272482; the National High Technology Research and Development Program of China (863 Program) No. 2011AA01A103, 2012AA01A506, 2013AA013505, the Research Fund for the Doctoral Program of Higher Education of China under Grant No. 20114307110006, 20124307120032, the program for Changjiang Scholars and Innovative Research Team in University (No.IRT1012), Science and Technology Innovative Research Team in Higher Educational Institutions of Hunan Province(“network technology”); and Hunan Province Natural Science Foundation of China (11JJ7003).

References


