Study of Tolerability against Defacing of Y-00 Quantum Stream Cipher

Takehisa Iwakoshi

Quantum ICT Research Institute, Tamagawa University
6-1-1 Tamagawa-gakuen, Machida, Tokyo 194-8610, Japan

©Tamagawa University Quantum ICT Research Institute 2011
All rights reserved. No part of this publication may be reproduced in any form or by any means electrically, mechanically, by photocopying or otherwise, without prior permission of the copy right owner.
Study of Tolerability against Defacing of Y-00 Quantum Stream Cipher

Takehisa Iwakoshi
Quantum Communication Research Center, Quantum ICT Research Institute, Tamagawa University
6-1-1 Tamagawa-Gakuen, Machida, Tokyo, 194-8610, Japan
E-mail: t.iwakoshi@lab.tamagawa.ac.jp

Abstract—Conventional mathematical stream ciphers generate ciphertext by exclusive-OR of plaintext and running key from pseudo-random-number generators. However, because of this encryption, an attacker can intercept the ciphertexts between legitimate users and deface their plaintexts easily. Realizing tolerability against defacing in a mathematical stream cipher is not easy. On the other hand, a quantum stream cipher Y-00 may have a certain level of tolerability against defacing, since the attacker cannot even correct cipher texts from Y-00 quantum stream ciphers. In this paper, it is shown that there can be difference in tolerability against defacing via the methods in Y-00 realization; intensity-shift-keying Y-00 has a defacing tolerability but phase-shift keying Y-00 does not.

I. INTRODUCTION

In modern communication, block ciphers like AES are employed to protect plaintexts from attackers. Besides, stream ciphers can also be used for faster communication since they can be driven faster than block ciphers can be [1].

However, stream ciphers have several problems compared to block ciphers. For example, typical stream ciphers generate ciphertext \( C_t \in \{0, 1\} \) by exclusive-OR of plaintext \( X_t \in \{0, 1\} \) and running key \( K_t \in \{0, 1\} \) like \( C_t = X_t \oplus K_t \). In such a stream cipher, \( X_t \) is decoded by \( X_t = C_t \oplus K_t \). However, if an attacker intercepts \( C_t \) and replaces it with \( C'_t = C_t \oplus A_t \), the received plaintext by the legitimate receiver becomes a defaced one, \( X'_t = X_t \oplus A_t \). Such a defect is a big threat especially in the case of Known-Plaintext-Attack (KPA), which cannot be prevented even by information-theoretic secure communication, one-time-pad. Thus, there is always a threat that the ciphertexts may be defaced in stream ciphers.

To detect defacing, hash functions can be employed. Hash functions change their output values drastically even input data are slightly changed. It is also expected that recovering the input data from the out put value of the hash function is difficult (pre-image resistance), and finding different inputs with a same output value of the hash function is difficult (second pre-image resistance).

However, vulnerabilities were found in hash functions like MD5 in August 2004 [2], and researchers found a drastically efficient way in finding collisions in SHA-0 in 2004 [3]. Thus, many researchers questioned the security of new hash functions such as SHA-1, RIPEMD-128, and RIPEMD-160 in long term, which were derived from SHA-0, MD5 and so on. It was also suggested SHA-1 should have been abolished before 2010 [4]. Actually, an algorithm was found to yield a collision of out put values with different input data in SHA-1 with \( O(2^{128}) \) steps in June 2009 [5]. Thus, to realize safety assurance, a defacing tolerability and detection are important themes in cryptography.

On the other hand, there is a class of encryption method called physical encryption to realize security of communication by physical phenomena. There are not many ways of physical encryptions but Y-00, which makes an attacker’s cryptoanalysis hard by hiding even ciphertext in quantum noise of coherent light, is a typical example of physical encryption.

Some researchers have been studying to realize near information-security with limited key length by its physical complexity [6-12], independent from computational complexity. However, Y-00 is also a kind of stream cipher, so it is unsure whether it has a defacing tolerability or not. In this article, tolerability against defacing of Y-00 is discussed, in the case of employing Phase-Shift-Keying (PSK) and Intensity-Shift-Keying (ISK).

II. FUNDAMENTALS OF Y-00 PROTOCOL

Gauss proposed a cipher using private randomization 200 years ago to realize highly secure encryption [13, 14]. In 2000, Y-00 protocol was invented to realize a new way of private randomization using quantum noise [6]. Current Y-00 systems employ semi-classical transceivers, which are already realized.

Y-00 protocol expresses ciphertexts using quantum states of coherent light. A PSK Y-00 system expresses ciphertexts by their phases [6], while an ISK Y-00 system expresses ciphertexts by their intensities [10]. Both types of Y-00 are already realized and driven over 2.5 Gbit/sec about 300km distance.

A. Fundamentals of a PSK Y-00 system

A PSK Y-00 system was firstly invented by the proponents of Y-00 protocol [6]. The PSK Y-00 system is easier to understand, so firstly it is explained how the PSK Y-00 system works.

Legitimate users share a secret key \( K_t \) and extract to a \( M \)-ary running key sequence \( K_t^K = \{0, 1, 2, \ldots, M-1\} \) using
PRNG (Pseudo Random Number Generators) like LFSR (Linear Feedback Shift Resister). In the PSK Y-00 system, the phase $\theta$ of coherent light is chosen by eq. (1).

$$\theta = \left[ K^R / M + Pol(K^R) \right] \pi$$

Here, $Pol(K^R) = K^R \mod 2$, for example. A ciphertext is expressed by quantum state of coherent light as $|\alpha^R\rangle$, where $|\alpha|^2$ is the intensity of coherent light.

In the PSK Y-00 system, heterodyne receivers are employed to detect the phases, so the probability of receiving state $|\alpha_t\rangle$ from sent state $|\alpha_s\rangle$ is

$$P(\alpha_t | \alpha_s) = \frac{1}{\pi} \exp(-|\alpha_t - \alpha_s|^2)$$

Thus, the variance of quantum noise is $\sigma_{\text{PSK}}^2 = 1/2$ for the PSK Y-00 system. The number of symbols hidden by quantum noise is $\Gamma = 2\sigma_{\text{PSK}} \Delta_{\text{PSK}}$, where $\Delta_{\text{PSK}} = |\alpha|^2/M$. The original PSK Y-00 system firstly developed has $\Gamma \sim 4.6$, where $|\alpha|^2 = 4.0 \times 10^4$ [photons] and $M = 2048$ [15].

Even legitimate users observe quantum noise, however, they can distinguish $X_t = 0$ from 1 by setting thresholds at $\theta_t^{\text{th}} = [K^R / M \pm 1/2] \pi$ since they know the running key $K^R$. Thus, the legitimate users can receive binary plaintext data with almost zero BER (Bit Error Rate). On the other hand, an attacker who does not know $K^R$ is forced to receive 2M-ary data to eavesdrop the ciphertexts, which has many errors by quantum noise.

When $\Gamma$ is not sufficient, DSR (Deliberate Signal Randomization) is employed. DSR is a technique to enlarge $\Gamma$ using some randomization techniques. It is realized by several methods. A research team in Tamagawa University proposed to use private PRNG [16]. This method is called “keyed DSR”

### B. Fundamentals of an ISK Y-00 system

An ISK Y-00 system was developed from the view point of affinity for conventional optical communication technology [12]. A ciphertext is expressed by the intensity of coherent light $|\alpha_t|^2 = S_t$, where $S_t$ is given by eq. (3)

$$S_t = S_{\text{max}} + \Delta S [K^R + M \cdot Pol(K^R) \cdot X_t]$$

Here, $\Delta S = (S_{\text{max}} - S_{\text{min}})/(2M-1)$, $S_{\text{max}}$ is the maximum intensity of the signal and $S_{\text{min}}$ is the minimum intensity of the signal.

In the ISK Y-00 system, direct detection is employed to detect signals. So the probability of receiving a photon-number state $|n\rangle$ from a sent state $|\alpha_s\rangle$ is

$$P(n | \alpha_s) = \frac{1}{\sqrt{n!}} \exp(-|\alpha_s|^2) |\alpha_s|^n$$

Thus, the variance of quantum noise is $\sigma_{\text{ISK}}^2 = S_t$ for the ISK Y-00 system. The number of symbols hidden by quantum noise is $\Gamma = 2\sigma_{\text{ISK}} / \Delta S$. The prototype ISK Y-00 system has $\Gamma \sim 7.0$ for pure quantum noise, where $\sigma_{\text{ISK}}^2 = \left[ S \right]^2 = 3.1 \times 10^6$ [photons] and $\Delta S = 507$ [photons] [17]. However, circuit noise is added to quantum noise in direct detection, so the experimental result was $\Gamma \sim 255$ [18]. Figure 1 shows the overview of our ISK Y-00 system.

---

**Fig. 1** The overview of our ISK Y-00 system

Same as in the case of the PSK Y-00 system, even the legitimate users observe quantum noise. However, they can distinguish $X_t = 0$ from 1 by setting a threshold at $S_{\text{th}} = S_{\text{max}} + \Delta S [K^R + M/2]$ since they know the running key $K^R$. Thus, the legitimate users can receive binary plaintext data with almost zero BER. On the other hand, the attacker who does not know $K^R$ is forced to receive 2M-ary data to eavesdrop the ciphertexts, which has many errors by quantum noise.

Same as in the case of a PSK Y-00 system, an ISK Y-00 system also can employ DSR. The effect of DSR against defacing attack will be discussed later.

### C. The characteristics of Y-00 protocol as a random cipher

Y-00 protocol realizes the security using the difference of effect of quantum noise on the legitimate users and the attacker. Both the legitimate users and the attacker also observe a Y-00 signal as analog photo-electric current. That is,

$$I(\alpha_s) = f(X_t; K^R; r_t)$$

Here, $r_t$ is noise in a Y-00 signal. This analog photo-electric signal fluctuates by $r_t$, thus the signal corresponds with many possible running key $K^R$. As a result, ciphertexts are randomized. The attacker who does not know $K^R$ is affected by $r_t$ directly. On the other hand, the legitimate users who know $K^R$ can avoid the effect of $r_t$ in getting binary plaintext. That is,

$$X_t = D_{K^R} \{ f(X_t; K^R; r_t) \}$$

The security of asymmetric random cipher like Y-00 is explained like following, briefly.

(a) Chiphertext Only Attack (COA)

The attacker receives analog photo-electric signals randomized by noise and converts them into 2M-ary signals by a digital-analog converter. From this 2M-ary signals, the attacker guesses the secret key $K_S$ which is shared by the legitimate users.

(b) Known Plaintext Attack (KPA)

The attacker receives analog photo-electric signals randomized by noise, and converts them into M-ary signals by a digital-analog converter. From this M-ary signals, attacker guesses the secret key. Otherwise, the attacker tries brute force
attack against $2^K$ patterns of secret keys $K_S$ by special devices for attacking a Y-00 system.

Thus, a cipher randomized by noise realizes the security different from the one of conventional mathematical ciphers’ [19]. In this article, we will show that the tolerability against defacing of this kind of random cipher.

III. TOLERABILITY AGAINST DEFACING Y-00 SIGNALS

Before we start the discussion, we define 2 types of defacing attacks as following;

(a) Unplanned defacing: The attacker defaces the plaintext into another plaintext without certain meaning.

(b) Planned defacing: The attacker defaces the known plaintext into another plaintext which the attacker planned.

Furthermore, we define ciphertext-only defacing as a defacing in case that the attacker has only ciphertexts, beside we define known-plaintext defacing as a defacing in case that the attacker knows the plaintexts correspond to the ciphertexts.

Conventional mathematical stream ciphers including one-time-pad generally encodes a plaintext by

$$C_i = X_i \oplus K_i$$

The attacker can get an exact copy of the ciphertext $C_i$ from such mathematical cipher systems. Thus, defacing is very easy by adding 1 to the ciphertext,

$$C_i' = C_i \oplus 1$$

As explained previously, by this defacing, the attacker can send inverted letters to the legitimate receiver. In the case of known-plaintext defacing, the attacker can target bits to invert. In this way, the attacker can send a meaningfully defaced plaintext to the legitimate receiver. This attack cannot be prevented even information-theoretic secure stream ciphers like one-time-pad.

The above is the defacing attack against typical mathematical stream ciphers. Thus, we are interested in physical ciphers, especially Y-00 protocol. It will be discussed how the difference of the defacing tolerability between the PSK Y-00 system and the ISK Y-00 system.

A. Defacing attack against Y-00 systems

Eliciting the previous definition, we discuss defacing attack against Y-00. In Y-00 protocol, signals are sent as $2^M$-ary optical signals, so the attacker needs to convert them into $2^M$-ary photo-electro signals by an analog-digital converter. To deface a signal, the attacker needs to send a signal with inverted parity bit as an optical signal to the legitimate user. The legitimate receiver discriminates the signal by threshold(s) based on running key $K^R$. From the above, if the attacker makes the legitimate receiver get an inverted bit, the defacing succeeds.

B. Defacing attack against a PSK Y-00 system

In the PSK Y-00 system, heterodyne receivers are employed to receive coherent signals. Thus, the variance of quantum noise is only $\sigma_{PSK}^2 = 1/2$. So the PSK Y-00 system would be better off employing DSR to hide the half of the phase space, which maximizes the security of PSK Y-00. Here, we assume that the PSK Y-00 system employs DSR.

(a) Ciphertext-only defacing

The attacker receives $2^M$-ary phase signals in the middle of communication channel with a heterodyne receiver. Since the legitimate users employ DSR, the variance of noise hides the half of the phase space. The probability distribution of the noise by DSR is expressed by eq. (9) assuming that $\theta_t$ is a sent signal by the legitimate transmitter and $\theta_L$ is a signal the attacker receives.

$$P_{DSR}(\theta_L|\theta_t) = [H(\theta_L - \theta_t + \sigma_{DSR}) - H(\theta_L - \theta_t - \sigma_{DSR})](2\sigma_{DSR})^{-1}$$

Here, $\sigma_{DSR}$ is a half width of DSR noise, which is $0 < \sigma_{DSR} < \pi/2$. $H(x)$ is Heaviside’s step function which is $H(x) = 0$ if $x \leq 0$, otherwise $H(x) = 1$. However, the attacker can invert the signal no matter how $\sigma_{DSR}$ is large.

$$\theta_L \rightarrow \theta_L + \pi$$

Thus, the attacker can send the inverted signal $\theta_L + \pi$ to the legitimate receiver without error. By the above way, attacker succeeds in unplanned defacing of ciphertext with probability 1. This is expressed by Fig. 2.

Fig. 2: A defacing process in a PSK Y-00 system.

A black spot indicates the data bit $\theta_t$ sent by the legitimate transmitter, besides the gray zone is DSR noise $\pm \sigma_{DSR}$. Even if it employs DSR, defacing cannot be prevented.

(b) Known-plaintext defacing

If the attacker knows the plaintext, like $X = 1010101010$, and plans to deface it into $X' = 1111111111$, the attacker can invert the 2nd, 4th, 6th, 8th and 10th bits in the way of eq. (10). Thus, the attacker can succeed in planned defacing with probability 1.

From the above discussion, it is concluded that the PSK Y-00 system does not have a defacing tolerability.

C. Defacing attack against an ISK Y-00 system

In an ISK Y-00 system, there are $2^M$ of symbols between the maximum intensity $S_{max}$ and minimum intensity $S_{min}$ as
previously explained. Here we set an average laser intensity \( S_{\text{ave}} = (S_{\text{max}} + S_{\text{min}})/2 \) to discuss 2 cases of \( S_t > S_{\text{ave}} \) and \( S_t < S_{\text{ave}} \).

(a) Ciphertext-only defacing

Here, the attacker is assumed to be able to get signals without any attenuation. The attacker who does not know running key \( K_t^E \) has to invert the bit by receiving 2M-ary data.

When the attacker tries to invert the parity of the signal, the attacker cannot shift the observed signal \( n_t^E > S_{\text{ave}} \) to \( n_t^E + M \Delta S > S_{\text{max}} \). Thus, the attacker has to shift it to \( n_t^E - M \Delta S \). However, if \( n_t^E < S_{\text{th}} \), then both \( n_t^E \) and \( S_{\text{ave}} - M \Delta S \) indicate same \( X_t \). On the other hand, if a signal \( n_t^E < S_{\text{ave}} \) is received, the attacker cannot shift it to \( n_t^E + M \Delta S < S_{\text{min}} \). Thus, attacker has to shift it to \( n_t^E + M \Delta S \) even if both of \( n_t^E \) and \( n_t^E + M \Delta S \) indicate same \( X_t \) in the case of \( n_t^E + M \Delta S > S_t \). This is the qualitative explanation how the attacker fails in defacing against the ISK Y-00 system. The situation is illustrated in Fig. 3.

![Fig. 3 Qualitative explanation how the attacker fails defacing](image)

Here we calculate the probability that the attacker fails defacing against the ISK Y-00 system. When the signal the attacker observed is \( n_t^E \), the probability that the attacker fails defacing (defacing tolerability) is,

\[
P_{de}(S_t) = \sum_{S_{\text{ave}} < S_t} P(n_t^E | S_t)
\]  

(11)

In the case of \( S_t > S_{\text{ave}} \), the defacing tolerability is,

\[
P_{de}(S_t) = \sum_{S_{\text{ave}} < S_t} P(n_t^E | S_t)
\]  

(12)

So the average of the defacing tolerability is,

\[
\overline{P}_{de} = \frac{1}{2M} \sum_{S_t} P_{de}(S_t)
\]  

(13)

From above discussion, we conclude \( \overline{P}_{de} > 0 \), which means ISK Y-00 has some defacing tolerability.

The profile of \( P_{de}(S_t) \) is shown in Fig. 4. The simulation condition is listed in Tab. 1. \( P_{de}(S_t) \) has meaningful value only when \( S_t \sim S_{\text{ave}} \) but it indicates that \( \overline{P}_{de} \) is not zero.

![Fig. 4 A profile of the defacing tolerability of an ISK Y-00 system without DSR](image)

However, \( \overline{P}_{de} \) is only 0.19%, which is not sufficient to protect plaintext from defacing. Thus, here DSR should be employed in the ISK Y-00 system.

When DSR is employed to the ISK Y-00 system, the signal profile which the attacker receives is explained by eqs. (14) and (15)

\[
P_{de}(S_t \mid S_t) = [H(S_t - S_t + \sigma_{DSR})
\]  

\[ - H(S_t + S_t - \sigma_{DSR})](2\sigma_{DSR})^{-1}
\]  

(14)

\[
P_{de}(S_t \mid S_t) = \int dS_t' P(n_t^E \mid S_t') P_{DSR}(S_t \mid S_t')
\]  

(15)

Equation (15) takes quantum noise into consideration in the final probability distribution the attacker receives, \( P(S_t \mid S_t) \), \( S_t \) is a signal the legitimate transmitter sends, and \( S_t^E \) is a signal the attacker receives. Maintaining BER between legitimate users below \( 10^{-10} \) with signal attenuation rate 1/100 between legitimate users, DSR width \( \sigma_{DSR} \) was adjusted. By this trial, it was obtained that BER between legitimate users \( 8.0 \times 10^{-10} \) and \( \sigma_{DSR} = 92000 \) [photons]. This corresponds to \( \Delta = 2\sigma_{DSR}/\Delta S = 920 \).

Then the defacing tolerability \( P_{de}(S_t) \) was calculated again assuming that the attacker is not affected by any signal attenuation. This is calculated by eq.(16) for \( S_t \leq S_{\text{ave}} \)

\[
P_{de}(S_t) = \sum_{S_{\text{ave}} < S_t} P(n_t^E \mid S_t)
\]  

(16)

and eq.(17) for \( S_t > S_{\text{ave}} \)

\[
P_{de}(S_t) = \sum_{S_{\text{ave}} < S_t} P(n_t^E \mid S_t)
\]  

(17)

The result is shown in Fig. 4. The area where the defacing tolerability has meaningful value was widened. From this profile, \( \overline{P}_{de} \) is 5.6%, which means that the attacker has one defacing error in every 12 letters with probability \( (1-0.056)^{12} = 0.50 \). However, as it is shown in Fig. 5, a defacing tolerability exists only around \( S_t \sim S_{\text{ave}} \). This should be solved by other techniques.

### Table 1 The simulation condition for Fig. 4

<table>
<thead>
<tr>
<th>Photon number ( S_{\text{ave}} )</th>
<th>4×10^6 [photons]</th>
</tr>
</thead>
<tbody>
<tr>
<td>( M )</td>
<td>2048</td>
</tr>
<tr>
<td>( \Delta S )</td>
<td>200 [photons]</td>
</tr>
</tbody>
</table>

Note: \( S_{\text{ave}} \) is the average of the photon number.
(b) Known-plaintext defacing

Assume that the attacker knows the plaintext, like $X = 1010101010$. If the attacker plans to deface $X$ to $X' = 1111111111$, 2nd, 4th, 6th, 8th, 10th bit should be inverted. However, the attacker fails inverting targeted bits with the probability $P_{de} = 5.6\%$. Quantitatively, the probability of the attacker succeeding in defacing of those all 5 letters is $(1-0.056)^5 = 0.75$. Thus, the ISK Y-00 system has a defacing tolerability against planned defacing too, to some degree.

IV. CONCLUSION

In this study, tolerability against defacing of quantum stream cipher Y-00 protocol was evaluated. There are two ways to realize Y-00 protocol; the one is phase-shift-keying Y-00, the other is intensity-shift-keying Y-00. It was found that a phase-shift-keying Y-00 system does not have a defacing tolerability at all, while an intensity-shift-keying Y-00 system has a defacing tolerability to some degree. However, its value is still small and it exists only around the average signal intensity. Thus, another technique for enhancing a defacing tolerability will be the next study.

REFERENCES