Journal of Information Hiding and Multimedia Signal Processing ©2015 ISSN 2073-4212
Ubiquitous International Volume **6**, Number **3**, May **2015**

Improving Reliability of Covert Timing Channel to Packet Loss

Lihua Zhang and Guangjie Liu

School of Automation
Nanjing University of Science and Technology
200 Xiaolingwei, Nanjing, China
zlh6013@163.com; gjieliu@gmail.com

Jiangtao Zhai and Yuewei Dai

School of Electronics and Information
Jiangsu University of Science and Technology
No.2 Mengxi Road ,Zhenjiang, Jiangsu, China
jiangtao_zhai@yahoo.com.cn, daiywei@163.com

Received September, 2014; revised February, 2015

Abstract. *Network covert timing channel embeds confidential information into the inter-packet delays of the network traffic generated by a legitimate application. The packets of carrier channel may be lost when the network is in congestion state. It would impact the confidential information transfer in the covert timing channel. In this paper, a novel scheme is proposed to resolve the problem caused by packet loss. Reed-Solomon code and interleaving are introduced into the design of covert communication system to improve the reliability. The experiment results show that the errors caused by packet loss can be corrected when the packet loss rate is in the range of error correction capability. The reliability of covert timing channel is improved significantly while the detection resistance is not degraded.*

Keywords: Covert timing channel, Packet loss, Reliability, Reed-Solomon code, Interleaving

1. **Introduction.** Network covert channel exploits overt traffic as carrier to transfer confidential information to an unauthorized recipient. In general, network covert channel can be classified into two types: covert storage channel and covert timing channel [1]. In a covert storage channel, the sender encodes the confidential information symbols into the unused or reserved bits of packet frame. However, most of covert storage channels can be detected and eliminated easily. In a covert timing channel, the sender modulates the inter-packet delays (IPDs) according to the confidential information and the receiver recovers the confidential information by observing the packets' arrival time. The IPDs may be affected by the network inherent noise, such as network delay jitters and packet loss. Jitters will result in the decoding mistakes. Existing researches proposed several methods to reduce the symbol error rate of covert timing channel due to jitters [2, 3, 4]. When packet loss occurs, some symbols of confidential information will be missed [5]. To make matters worse, the following symbols will be disorder and the confidential information decoding may be unsuccessful. In this paper, an encoding and decoding scheme is proposed to improve the reliability of covert timing channel over packet loss channel.

Automatic Repeat Request (ARQ) technique is generally used in network data transmission. If the sender does not receive an acknowledgment before the timeout, it usually re-transmits the packet until the sender receives an acknowledgment or exceeds a predefined number of re-transmissions. However, ARQ technique is not applicable in covert communication system. It is because of the low bandwidth of covert channel and the lack of a feedback channel. In this case, forward error correction (FEC) techniques become useful.

In FEC techniques, the sender prevents losses by transmitting some amount of redundant information, which allows the reconstruction of missing data at the receiver without further interactions. Besides reducing the time needed to recover the missing packets, such an approach generally simplifies both the sender and the receiver since it might render a feedback channel unnecessary. The two main categories of FEC codes are block codes and convolution codes. Block codes work on fixed size blocks of bits or symbols of predetermined size, while convolution codes work on bit or symbol streams of arbitrary length. The Reed-Solomon (RS) code is particularly useful for burst-error correction and effective for channels that have memory [6]. It can be used efficiently on channels where the set of input symbols is large. This paper adopts the idea of RS code in the encoding process and proposes a reliable timing channel which can resistant packet loss of normal traffic.

If a group of consecutive errors occurs in the symbol sequence of confidential information due to several consecutive lost packets, the RS code may be not enough power to correct the bursts of errors. For instance, the error correction capability of the RS (7, 3) is two. If there is one burst of errors consisting of 5 consecutive symbols, the RS (7, 3) is not be able to correct this burst of errors. Interleaving which converts bursts of errors into random-like errors is an effective means to combat bursts of errors. In this paper, interleaving is first used to improve the reliability of network covert timing channel combined with the RS code.

The remainder of this paper is organized as follows. We review the related work of covert timing channel in Section II. Section III introduces the basics of Reed-Solomon codes. In Section IV, we give a detailed description of the reliable covert timing channel scheme. In Section V, we evaluate the performance of the proposed scheme by the simulation experimental results. Finally, the paper is concluded and some future directions are pointed out in Section VI.

2. **Related Work.** Network covert storage channels hide the confidential information bits in the packet header by changed the unused or optional fields, such as IP checksum, IP TTL, TCP initial sequence number, and so on. Traffic normalization which standardize the header fields of packets is an effective way to eliminate network covert storage channels. Therefore, this paper focuses on the covert timing channel.

Network covert timing channels modulate the transmission time of the packets to covertly transmit the confidential information symbols. In initial researches, the sender and receiver agree a priori on a timing interval [7]. The sender either transmits a single packet or remains silent during each interval to indicate which confidential information symbol 1 or 0 to be sent. It is a challenge to implement end-to-end synchronization between the sender and the receiver [8]. Another kind of covert timing channel inserting the confidential information symbols in the IPDs [9]. The sender and receiver agree on a set of timing intervals $t1$, $t2$, a set of associations for each timing interval $(t1, 0)$ and $(t2, 1)$. The sender sleeps for $t1$ units of time and sends a packet at the end of this interval to transmit 0. Likewise, the sender sends a packet after sleeping for $t2$ units of time to transmit 1. The receiver records the binary string of confidential information over the covert channel by recording the inter-arrival time t and comparing t to both $t1$ and $t2$. Both of the covert traffic can be reliably differentiated from normal traffic by the e-Similarity measure.

Shah et al. built a hardware keylogger named Keyboard JitterBug [9] which is deployed between the keyboard and the computer. Sensitive information can be leaked over the Internet without compromising the host. If a user is typing in an interactive network application, each keystroke will be sent in its own network packet. The Keyboard JitterBug adds timing information to keypresses in the form of small jitters that are unnoticeable to the user. The receiver reconstructs the original information by observing the packet timings.

The schemes mentioned above tried to taken steps to better hide the presence of covert timing channels. To make the covert traffic close to legitimate traffic, Gianvecchio et al. introduced model-based covert timing channels which endeavor to evade detection [10]. The scheme first derives a model based on the statistical properties of the observed legitimate traffic. According to the model, the sender chooses the appreciate distribution functions and traffic generation libraries to generate covert traffic. Liu et al. proposed a covert timing channel with distribution matching [11] which treats normal traffic as the flow with the fixed-length fragment, and calculates the histogram of the IPDs in each fragment. The confidential information symbols are modulated into the IPDs by the binary coding method.

Recent researches have considered introducing some coding algorithm into the design of covert timing channels. Liu et al. modulated confidential information in the IPDs by using spreading codes in order to increase the robustness of the covert timing channel to the network perturbations [2]. Amir et al. improved the robustness of covert timing channel by using different coding algorithms [3]. Wu et al. introduced Huffman coding scheme to compress the transferred data by exploiting redundancy of confidential information [4].

To defend against covert timing channels, various algorithms have been proposed to disrupt and detect them. Giles and Hajek proposed a device named network jammer which added random delays to the IPDs of all network traffic [12]. Although the jammer limits the usefulness of covert timing channels, it disrupts the expected performance of legitimate traffic, especially to the interactive traffic and real-time traffic. Cakuk et al. proposed regularity test which examined whether the variance of IPDs sequence is relatively constant [8]. The regularity test is based on the fact that the variance of the IPDs sequence changes over time for most legitimate traffic, while the variance of the IPDs sequence for covert timing channel remains relatively constant when the encoding scheme does not change over time. Peng et al. proposed the Kolmogorov-Smirnov (KS) test which is able to detect IPDs sequence of covert timing channel [13]. The KS test is a nonparametric test that can be used to compare two samples by the KS statistic which is defined as the maximum distance between the empirical distribution function of two samples.

Gianvecchio et al. proposed a scheme to detect covert timing channels based on Shannon entropy [14]. In a sample of IPDs, each IPD may be mapped to one of M symbols. Each symbol can be mapped to a bin in a histogram of IPDs. The bins range is determined based on a large number of legitimate traffic IPDs, such that each bin contains the same number of samples. In other words, each bin should have equal probability for each IPD of legitimate traffic. To detect the presence of covert timing channel, a sample set of IPDs is mapped to the bins. The entropy H is calculated using Shannon entropy. The sample set whose entropy is less than legitimate set is identified the IPDs of a covert timing channel. In summary, it can be found that few work has been done to resolve the problem caused by packet loss in covert timing channel. In this paper, we proposed a novel encoding and decoding scheme which combines RS code and interleaving to improve the reliability of covert timing channel over packet loss channel.

3. Proposed Scheme.

3.1. **The Framework.** In traditional covert timing channels, the sender modulates IPDs of adjacent packets according to the confidential information symbols directly. The impact of network noise such as packet loss to covert timing channel is not took into account. In fact, the noise can cause fatal impact to the covert transmission of confidential information. As shown in Fig. 1, if a packet is lost, two adjacent IPDs will combine into one. The minor error will result in a demodulation mistake of the IPD and disorder of all the subsequent received symbols. The decoding of confidential information at receiver end may be unsuccessful. It can be regarded as Butterfly Effect in covert timing channel.

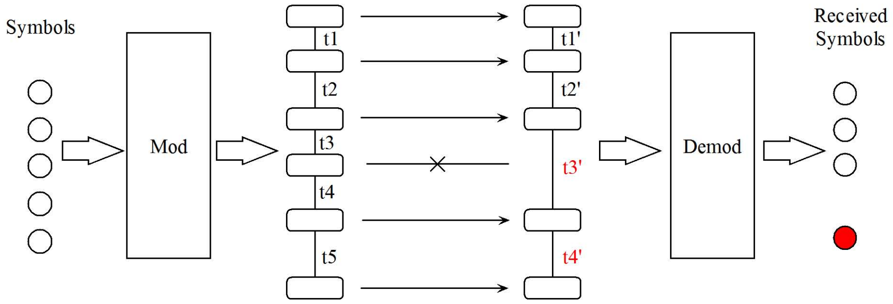


Figure 1. The impact of a lost packet to covert timing channel

In order to mitigate the impact of packet loss to covert timing channel, we introduce RS code and interleaving in the encoding and decoding process. Framework for the proposed covert communication scheme is shown in Fig. 2. The symbols of confidential information are encoded by RS code defined over $GF(8)$ with the code rate of $r = k/n$ where k and n denote the number of information symbols and output symbols of the RS encoder respectively. The encoder maps k input symbols into n RS code symbols. Each code symbol consists of 8 bits. The code symbol sequence is converted into the binary symbol sequence and input into Interleaver. Interleaver mixes up the code symbols from different codewords so that bursts errors are spread across multiple codewords when the codewords are reconstructed at the receiver. Then, bit error number for each codeword is more likely to be in the range of RS error correction capability.

The Symbol-to-Ipds modulator embeds the code symbols from Interleaver into IPDs of legitimate traffic. Aiming to evade detection, IPDs are generated by the inverse transform of offline legitimate traffic IPDs' cumulative distribution function(CDF). Transmitter sends out packets of normal application with the generated IPDs.

Arrival moments of covert traffic packets are the signal observable to covert channel receiver. After transmission on the network, the IPDs sequence is interfered by network noise. In order to recover the correct confidential information symbols, receiver needs to record the IPDs of covert traffic at first. The demodulator converts the recorded IPDs sequence to codewords. Then, the codewords are input to the de-interleaver whose output is input into RS decoder. Finally, confidential information is obtained by receiver.

3.2. **Encoding of RS Code.** By adding t check symbols to the data, RS code can detect any combination of up to t erroneous symbols, and correct up to $t/2$ symbols. As an erasure code, it can correct up to t known erasures, or it can detect and correct

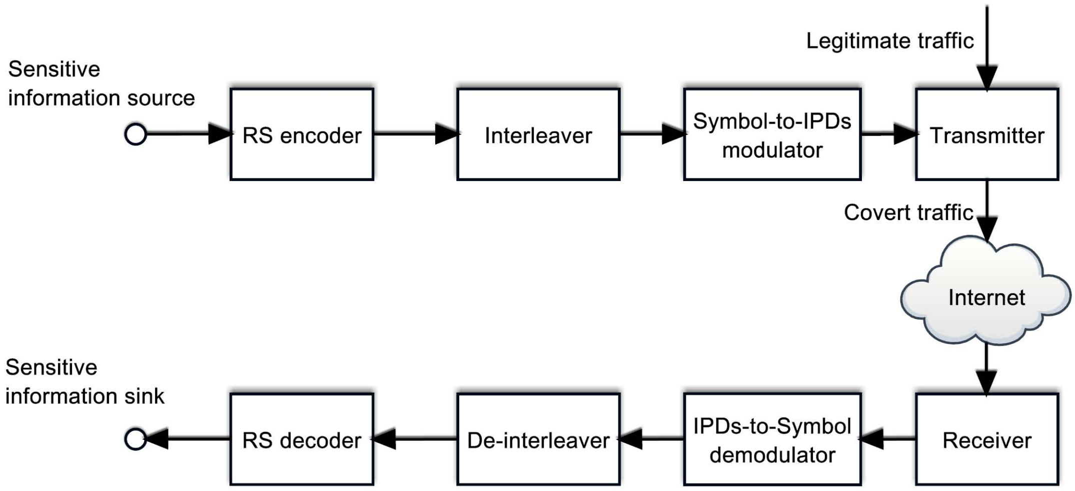


Figure 2. Framework for the proposed covert communication scheme

combinations of errors and erasures. RS code $C(n,k)$ $(r \triangleq n-k)$ is defined by the generator polynomial expresses as Eq.(1).

$$G(X) \triangleq \prod_{i=0}^{r-1} (X - \alpha^{1+i}), \tag{1}$$

where α is a primitive element of $GF(q)$. A polynomial of degree less than n is a codeword polynomial if and only if it is divisible by the generator polynomial $G(X)$. A codeword polynomial $C(X)$ satisfies

$$C(\alpha^{1+i}) = 0,\ i = 0,\ 1,\ 2, \dots,\ n-k-1. \tag{2}$$

In practice, a message polynomial $M(X)$ of degree up to $k-1$ is systematically encoded to a codeword polynomial $C(X)$ via a linear feedback shift register circuit which divides $X^r M(X)$ by $G(X)$ and results in its remainder polynomial $\Psi(X)$ of degree up to $r-1$. The encoding procedure can be expressed as Eq.(3). The minimum Hamming distance of RS code $C(n,k)$ is $d_{min} = n-k+1$, a feature known as maximum distance separable.

$$C(X) \triangleq X^r M(X) - \Psi(X) \tag{3}$$

3.3. **Principle of Interleaving.** Block interleaver is a kind of interleavers which are commonly used. The coded symbols in blocks from the RS encoder are input into a block interleaver. The block interleaver permutes the symbols, and then feeds the rearranged symbols to the modulator. The usual permutation of the block is accomplished by filling the columns of an M-row by N-column $(M \times N)$ array with the encoded sequence. After the array is completely filled, the symbols are then fed to the modulator one row at a time and transmitted over the channel.

At receiver end, the deinterleaver performs the inverse operation. It accepts the symbols from the demodulator, deinterleavers and feeds them to the decoder. Symbols are entered into the deinterleaver array by rows, and removed to the columns.

3.4. **Generation Mechanism of IPDs.** To evade detection, the modulator generates IPDs which mimic the statistical features of normal IPDs. The symbols which come from the interleaver map to IPDs based on the offline normal IPDs. The sender and receiver share a group of offline normal IPDs before the transmission of confidential information.

The offline normal IPDs are sorted to obtain the sequence $T = t_1, t_2, ..., t_T$. If an IPD is used to carry k symbols, the sequence T is partitioned into 2^k subsequences denoted as $T_1, T_2, ..., T_{2^k}$.

$$T_1 = \{t_1, t_2, ..., t_{T_1}\}, \tag{4}$$

$$T_2 = \{t_{T_1+1}, t_{T_1+2}, ..., t_{T_1+T_2}\}, \tag{5}$$

$$T_{2^k} = \{t_{\sum_{i=1}^{2^k-1} T_i+1}, t_{\sum_{i=1}^{2^k-1} T_i+2}, ..., t_T\}. \tag{6}$$

All of subsequences subject to the constraint condition expressed as Eq.(7).

$$\min_{T_1, T_2, ..., T_{2^k}} \left\{ \begin{array}{l} \left|F(t_{T_1}) - 1/2^k\right| + \left|F(t_{T_2+T_1}) - 2/2^k\right| \\ +... + \left|F(t_{\Sigma_{i=1}^{2^k-1} T_i}) - (2^k-1)/2^k\right| \end{array} \right\}, \quad T_1, T_2, ..., T_{2^k} \in \mathbf{T} \tag{7}$$

When the sizes of subsequences $T_1, T_2, ..., T_{2^k}$ are determined, the boundaries set of corresponding subsequences B expressed as Eq.(8) can be obtained.

$$\mathbf{B} = \{B_0,\ B_1,\ \cdots,\ B_{2^k-1},\ B_{2^k}\}, \quad B_0 = 0,\ B_1 = F(t_{T_1}),\ \ldots,\ B_{2^k} = 1 \tag{8}$$

The k bits binary symbols of confidential information $\{s_{i+1}, s_{i+2}, ..., s_{i+k}\}$ are firstly converted into random variable R which belongs to the interval $(0, 1)$ by Eq.(9).

$$R = B_d + (B_{d+1} - B_d) \cdot r, \tag{9}$$

where r is an pseudo uniform random number in the interval $(0, 1)$ and $d \in \{0, 1, ..., 2^k-1\}$ is the decimal number of k bits binary symbols of confidential information $\{s_{i+1}, s_{i+2}, ..., s_{i+k}\}$. Then, R is input into $F^{-1}(\cdot)$ to obtain corresponding IPD denoted as D. $F^{-1}(\cdot)$ is defined as

$$D = F^{-1}(R) = arg \min_{t \in T} |F(t) - R|. \tag{10}$$

Assume the IPDs observed by receiver end is denoted as D'. R' can be obtained as

$$R' = F(D'), \tag{11}$$

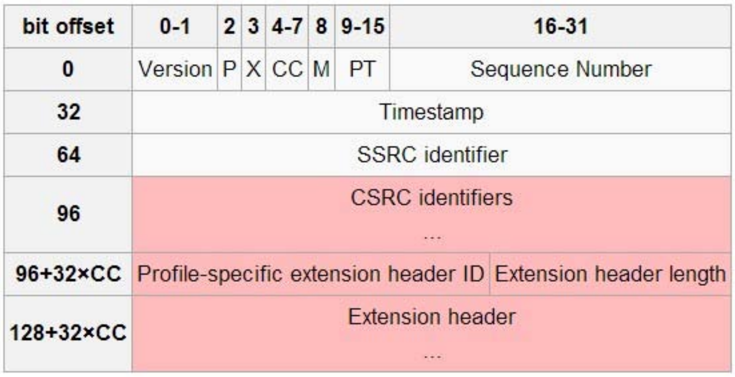
where $F(\cdot)$ the empirical distribution function of offline normal IPDs. The decimal number of symbols carried by the current IPD can be recovered according to Eq.(12).

$$d' = \left\{ \begin{array}{ll} i, & R' > B_i \ and \ R' \leq B_{i+1} \\ 0, & R' \geq B_0 \ and \ R' < B_1 \end{array} \right. \tag{12}$$

Finally, receiver obtains confidential information symbols by converting the decimal number d' to binary form.

3.5. **Processing Method for Packet Loss.** In a covert timing channel where packet loss may occur, the essential prerequisite for correct decoding is that location of packet loss can be determined. As we all know, TCP provides reliable delivery while the UDP has no handshaking dialogues, and is unreliability. Most of real time applications, such as Voice over IP (VoIP), video teleconference applications and television services, use UDP for transport because they don't require the reliability of TCP. Real-time Transport Protocol (RTP) is designed for end-to-end, real-time, transfer of stream data. The RTP header format is shown in Fig. 3.

UDP packet loss frequently occurs on the bottleneck node from local area networks (LANs) to WANs. This is because heavy traffic concentrates on the bottleneck node,



bit offset	0-1	2	3	4-7	8	9-15	16-31
0	Version	P	X	CC	M	PT	Sequence Number
32	Timestamp						
64	SSRC identifier						
96	CSRC identifiers ...						
96+32×CC	Profile-specific extension header ID						Extension header length
128+32×CC	Extension header ...						

Figure 3. The RTP header format

thereby causing traffic overload against the capacity of the WAN. The receiver can locate packet loss through the sequence number in ATP header. Then, the IPDs sequence needs to be corrected by the receiver.

The received IPDs sequence is denoted as $\{D'_1, D'_2, \ldots, D'_m\}$ where m is the number of recorded IPDs. $\{l_1, l_2, \ldots, l_n\}$ is defined as the state sequence. $l_i = 1$ indicates that the corresponding packet is lost, otherwise $l_i = 0$. The variable k is defined to record number of packets which are lost successively. The default value for k is one. The processing progress can be expressed as

$$D'_i \rightarrow \underbrace{D''{}_i,\ \ldots,\ D''{}_i}_{k} \quad if\ l_i = 1,\ l_{i+1} = 1,\ \ldots,\ l_{i+k-1} = 1, \tag{13}$$

where $D''{}_i = D'_i/k$. The supplements of IPDs may be not the same with the corresponding IPDs generated by sender. The errors caused by processing of packet loss will be corrected by RS code.

3.6. **Correction Algorithm based on RS Code.** Because of packet loss, the IPDs are interfered so that some confidential information symbols are recovered incorrectly. The error pattern can be described in polynomial form as

$$e(X) = \sum_{i=0}^{n} e_i X^i \tag{14}$$

The received corrupted-codeword polynomial is the sum of the transmitted codeword polynomial and the error pattern polynomial.

$$R(X) = C(X) + e(X) \tag{15}$$

Whether R is a valid member of the codeword set is determined by the result of a parity check performed on R. If the R is a member, the result S has value 0. Any nonzero value of S indicates the presence of errors. The computation of the results can be described as

$$S_i = R(X)\,|_{X=\alpha^i} = R(\alpha^i), \quad i = 1,\ \cdots,\ n-k. \tag{16}$$

Suppose there are v errors in the codeword at location $X^{j_1},\ X^{j_2},\ \ldots,\ X^{j_v}$. Then, the error polynomial can be written as

$$e(X) = e_{j_1}X^{j_1} + e_{j2}X^{j_2} + \cdots + e_{jv}X^{j_v} \tag{17}$$

To correct the interfered codeword, each error value e_{j_l} and its location X_l^j must be determined. When a nonzero syndrome vector has been computed, it signifies that an error has been received. Next, it is necessary to determine the location of the error or errors. An error locator polynomial can be defined as

$$\sigma(X) = (1 + \beta_1 X)(1 + \beta_2 X) \cdots (1 + \beta_v X) \tag{18}$$

The roots of $\sigma(X)$ are $1/\beta_1,\ 1/\beta_2,\ \ldots,\ 1/\beta_v$. The reciprocal of the roots of $\sigma(X)$ are the location numbers of the error pattern $e(X)$.

According to the autoregressive modeling techniques, we form a matrix from the syndromes, where the first t syndromes are used to predict the next syndrome. That is,

$$\begin{bmatrix} S_1 & S_2 & S_3 & \cdots & S_{t-1} & S_t \\ S_2 & S_3 & S_4 & \cdots & S_t & S_{t+1} \\ & & & \ddots & & \\ S_{t-1} & S_t & S_{t+1} & \cdots & S_{2t-3} & S_{2t-2} \\ S_t & S_{t+1} & S_{t+2} & \cdots & S_{2t-2} & S_{2t-1} \end{bmatrix} \begin{bmatrix} \sigma_t \\ \sigma_{t-1} \\ \vdots \\ \sigma_2 \\ \sigma_1 \end{bmatrix} = \begin{bmatrix} -S_{t+1} \\ -S_{t+2} \\ \vdots \\ -S_{2t-1} \\ -S_{2t} \end{bmatrix} \tag{19}$$

To solve for the coefficients of the error locator polynomial $\sigma(X)$, we should take the inverse of the matrix. The roots of $\sigma(X)$ are the reciprocals of the error locations. Once these roots are located, the error locations will be known. In general, the roots of $\sigma(X)$ may be one or more of the elements of the field. We determine these roots by exhaustive testing of the $\sigma(X)$ polynomial with each of the field elements. Any element X that yields $\sigma(X) = 0$ is a root, and allows us to locate an error. An error had been denoted e_{jl} , where the index j refers to the error location and the index l identifies the lth error. Since each error value is coupled to a particular location, the notation can be simplified by denoting e_{jl} simply as e_l . Preparing to determine the error values e_l , associated with locations β , any of the four syndrome can be used. The estimated error polynomial is formed, to yield $\hat{e}(X)$.

4. **Experimental Results.** Reliability, efficiency and undectability are main properties for covert channel. The scheme proposed in this paper aims to improve reliability of covert timing channel when some of medium packets are lost. There is a trade-off between the reliability and efficiency. The improvement of reliability will result in the decrease of transmission rate. However, it is necessary to improve reliability, otherwise the lost packets not only cause the error of the adjacent IPDs, but also lead to the dislocation of the following confidential information symbols. In this section, we have validated the effectiveness of our proposed scheme. Reliability of covert timing channel is measured by the bit error rate (BER). Detection resistance is measured by corrected conditional entropy (CCE) [14]. The CCE is used to detect the differences in the regularity of a channel. It is defined as

$$CCE(X_m|X_1, ..., X_{m-1}) = CE(X_m|X_1, ..., X_{m-1}) + perc(X_m) \cdot EN(X_1), \tag{20}$$

where $CE(X_m|X_1, ..., X_{m-1})$ is the conditional entropy of detected sequence, $perc(X_m)$ is the percentage of unique patterns of length m and $EN(X_1)$ is the entropy with m fixed at 1.

We collect 226, 432 IPDs which are extracted from the traffic of YY Voice Application. These IPDs generated by normal communication are regarded as legitimate samples. The steganographic samples are generated to carry confidential information symbols based on the CDF of legitimate samples. The experiments are performed in three groups. In the first group, confidential information symbols are embedded into IPDs directly. In the second group, confidential information symbols are encoded by the RS code before modulated. In the third group, RS code and interleaving are both introduced in the encoding and decoding process.

RS code is characterized by the length of codeword $n = 7$, message length $k = 3$. The error correction capability of $RS(7, 3)$ is 2. One symbol of codeword is expressed as 3 bits binary symbols. Therefore, each block consists of 21 bits binary symbols. The maximum number of binary symbols which are demodulated incorrectly for a block recovered correctly by receiver is 6. Because one lost packet may cause two demodulation errors, the theoretical upper bound of packet loss rate that the $RS(7, 3)$ can resistance is 14%.

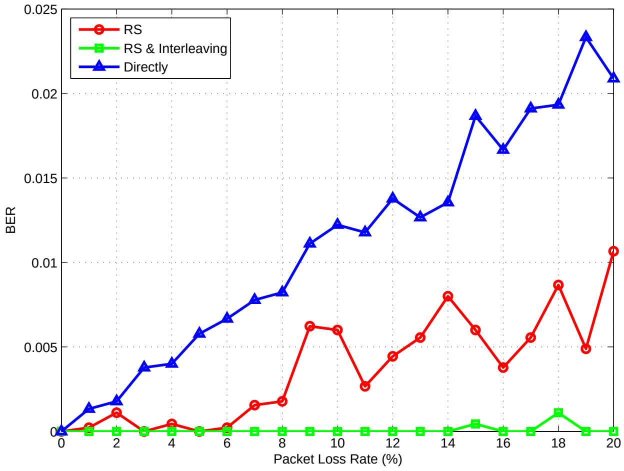


Figure 4. The BER performance comparison

The BER performances for three groups experiments are presented in Fig.4. As the packet loss rate increases, BER for the first group experiment rises faster than the other groups. Error correction capability is enhanced by using interleaving technology. When the packet loss rate is in the range of theoretical error correction capability, BER performance of the proposed scheme is improved significantly.

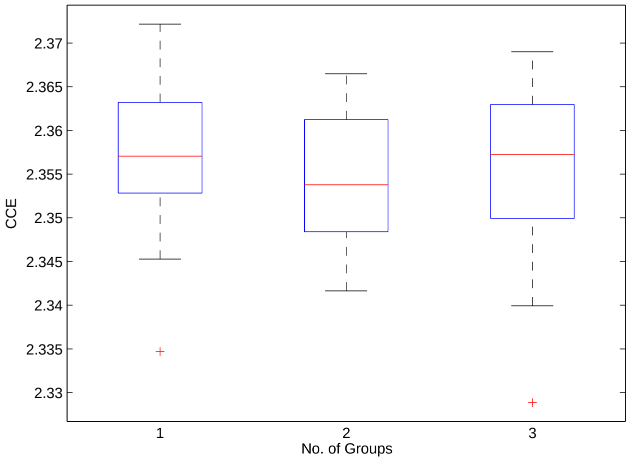


Figure 5. Boxplots of CCE test scores

Fig.5 illustrates boxplots of CCE test scores for three groups of steganographic IPDs when the packet loss rate is 5%. The bin sizes for computing of CCE test scores is $Q = 5$.

It is observed from Fig.5 that the CCE test scores for three groups are very close and indistinguishable. Introducing of RS code and interleaving doesn't cause a significant change in CCE test scores. Detection resistance of covert timing channel is not degraded for the proposed scheme in this paper.

5. **Conclusions and Future Work.** In this paper, the impact of packet loss to the transmitting of confidential information symbols in covert timing channel is analyzed. The encoding and decoding scheme which can resist the impact of packet loss is proposed. In order to improve the reliability of the proposed scheme, RS code and interleaving are introduced into the design of covert communication system. Simulation experiment results show that the confidential information symbols can be recovered from the received steganographic IPDs in low bit error rates when the packet loss rate is in the range of error correction capability. As we all know, the network is a dynamic system. The packet loss rate of carrier channel is always change. The design of adaptive covert timing channel is an interesting future directions for this work.

Acknowledgment. This study is supported by Nature Science Foundation of China (Grant No.61170250, 61103201 and 61472188) and the fundamental research funds for the central universities (No.30920140121006).

REFERENCES

[1] S. Zander, G. Armitage, and P. Branch, A survey of covert channels and countermeasures in computer network protocols, *IEEE Communications Surveys and Tutorials*, vol. 9, no. 3, pp. 44-57, 2007.

[2] Y. Liu, D. Ghosal, F. Armknecht, A.-R. Sadeghi, S. Schulz, and S. Katzenbeisser, *Robust and Undetectable Steganographic Timing Channels for i.i.d Traffic*, LNCS 6387, Springer, pp. 193-207, 2010.

[3] A. Houmansadr and N. Borisov, *CoCo: coding-based covert timing channels for network flows*, LNCS 6958, Springer, pp. 314-328, 2011.

[4] J. Wu, Y. Wang, L. Ding, and X. Liao, Improving performance of network covert timing channel through Huffman coding, *Mathematical and Computer Modelling*, vol. 55, no. 1-2, pp. 69-79, 2012.

[5] T. Nagano and A. Ito, Packet Loss Concealment of Voice-over IP Packet using Redundant Parameter Transmission Under Severe Loss Conditions, *Journal of Information Hiding and Multimedia Signal Processing*, vol. 5, no. 2, pp. 285-294, 2014.

[6] Y. Wu, Novel Burst Error Correction Algorithms for Reed-Solomon Codes, *IEEE Trans. on Information Theory*, vol. 58, no. 2, pp. 519-529, 2012.

[7] C. G. Girling, Covert Channels in LAN's, *IEEE Trans. on Software Engineering*, vol. SE-13, no. 2, pp. 292-296, 1987.

[8] S. Cabuk, C. E. Brodley, and C. Shields, IP Covert Timing Channels: Design and Detection, *Proc. of the 11th ACM conference on Computer and communications security*, Washington, DC, USA, pp. 178-187, 2004.

[9] G. Shah, A. Molina, and M. Blaze, Keyboards and covert channels, *Proc. of the 15th conference on USENIX Security Symposium*, Vancouver, B.C., Canada, pp. 59-75, 2006.

[10] S. Gianvecchio, H. Wang, D. Wijesekera, and S. Jajodia, Model-based covert timing channels: Automated modeling and evasion, *Recent Advances in Intrusion Detection*, Cambridge, MA, USA, pp. 211-230, 2008.

[11] G. Liu, J. Zhai, and Y. Dai, Network covert timing channel with distribution matching, *Telecommunication Systems*, vol. 49, no. 2, pp. 199-205, 2010.

[12] J. Giles and B. Hajek, An information-theoretic and game-theoretic study of timing channels, *IEEE Trans. on Information Theory*, vol. 48, no. 9, pp. 2455-2477, 2002.

[13] P. Peng, P. Ning, and D. S. Reeves, On the secrecy of timing-based active watermarking trace-back techniques, *IEEE Symposium on Security and Privacy*, Los Alamitos, California, USA, pp. 335-349, 2006.

[14] S. Gianvecchio; and H. Wang;, An Entropy-Based Approach to Detecting Covert Timing Channels, *IEEE Trans. on Dependable and Secure Computing*, vol. 8, no. 6, pp. 785-797, 2011.