A Simple Network Steganography Method based on the IPDs' Frequency Characteristic

Fang Yang

School of Electrics and Information Engineering, Jiangsu University of Science and Technology, Zhenjiang, 212003, China

Abstract

Network covert channel exploits the normal traffic as the carrier to transmit information secretly over the network. The existing covert timing channels have high security. However, they are sensitive to the jitters in the network. Thus, a new network steganography method based on the IPDs' frequency characteristic is proposed in this paper to increase its robustness while maintaining the security. It embeds secret information into the DCT coefficients of the IPDs by comparing the parity of the shared random numbers and the DCT coefficients, making the covert traffic more invisible and stable. When the covert traffic is received, a DCT transform is performed on the extracted IPDs and the secret information is got as the xor result of the parity of the random numbers and the DCT coefficients. The experimental results show that our scheme has stronger robustness and better invisibility than the existing methods.

Keywords: *covert timing channel, frequency domain, discrete cosine transform(DCT), robustness, invisibility*

1 Introduction

Network covert channel is a hidden communication technique, which utilizes the legitimate traffic as the vehicle to transfer the secret information covertly over the network. In recent years, network covert channel has become a hot research topic in the field of information security due to the fine properties of network traffic. There are two broad types of network covert channel: covert storage channel and covert timing channel. The former embeds the secret information into the redundancies of network protocols[1]. It's simple to implement, however, it can be detected by the existing methods easily[2]. The latter delivers the secret information by exploiting the time-relevant events of network packets and it has better stealthiness than covert storage channel[3]. Thus, the research on covert timing channel becomes an increasingly important issue.

In recent years, the research is mainly conducted to improve the stability and robustness of the covert timing channel. However, most of the existing methods would either generate abnormal covert traffic or reveal distinct properties compared with the normal case. As the embedding of the secret information alters the original carrier traffic in some way, the covert traffic is deviated from the normal one in some respects. This will make it possible to be detected by the existing methods[4]. Although some model-based covert timing channels have achieved high security, they are not stable in fact of uncertain disturbances[5]. In addition, most of the research has been conducted in time domain while few have been done in the frequency domain. Thus, we begin to consider a new covert timing channel based on the frequency domain. Discrete Cosine Transform (DCT) is an orthogonal transformation method put forward by Ahmed *et al.* in 1974. As Discrete Cosine Transform (DCT) is considered to be one of the most common and effective method of signal transformation, it is used in this paper to transform IPDs from time domain into frequency domain where the secret message is embedded.

The remainder of this paper is organized as follows. Section 2 reviews some selected steganographic methods and analysis of DCT. The proposed scheme is introduced in Section 3. In Section 4, experimental results are presented and analyzed. Finally, the whole paper is concluded.

2 Related works

2.1 Review of covert timing channels

Generally, covert timing channel can be divided into three sub-classes: on-off covert channel, inter-packet time intervals based covert channel, packet sorting and combination based ones. There is a brief introduction of several typical covert timing channels below.

Cabuk *et al.* proposed an IP covert timing channel called IPCTC[6]. It is a binary on-off channel which transmits a bit ‘1’ or ‘0’ by sending a packet during a certain time interval or not. It has improved considerably on the properties of stability and robustness by overcoming the problem caused by uncertainties in network transmission. But it can be easily detected because of the regularity of the time intervals.To imitate the statistical feature of normal traffic, Cabuk developed a more advanced method based on the replay attack, called Time-Replay covert timing channel(TRCTC), in which a sample of legitimate traffic is used as input and is replayed later to transmit information[7].

Girling proposed a covert timing channel which embeds time delay into the original traffic. Based on this method, Shah *et al.* designed a keyboard device named Jitterbug, to create a loosely-coupled covert channel capable of leaking information on a keyboard over the network[8]. Jitterbug conveys the information by adding small delays to the original time intervals of key-presses, inevitably altering the properties of the normal traffic.

Gianvecchio proposed a flexible framework called model-based covert timing channel(MBCTC)[9]. MBCTC employs the parametric estimation to construct the statistical model of the normal traffic, which is then utilized to generate inter packet delays containing the secret information. Liu *et al.* presented a more feasible method to fit the histogram distribution property of the normal traffic, termed as covert timing channel with distribution matching (CTCDM), which improves the accuracy of the transmission[10].

2.2 A brief introduction of DCT

DCT is considered to be one of the best methods of signal transformation. Let $c(x)$ be the IPD sequence and $C(u)$ be the DCT coefficients, the number of IPD is N, then the formula of DCT will be defined as Eq.(1):

$$
\begin{aligned}
C(0) &= \frac{1}{\sqrt{N}}\sum_{x=0}^{N-1} c(x) \\
C(u) &= \sqrt{\frac{2}{N}}\sum_{x=0}^{N-1} c(x)\cos\frac{(2x+1)u\pi}{2N} \\
c(x) &= \sqrt{\frac{1}{N}}C(0)+\sqrt{\frac{2}{N}}\sum_{u=0}^{N-1} C(x)\cos\frac{(2x+1)u\pi}{2N}
\end{aligned}
\qquad (1)
$$

There are two significant characteristics in DCT, they are listed as follows:

1. Energy concentration. The energy of the IPD sequence remain unchanged but redistributed with a few low frequency coefficients representing most of the energy after DCT transform. If the secret information is inserted into these low frequency coefficients, the covert channel would achieve high robustness, but low invisibility. When the secret information is inserted into those high frequency coefficients, there would be high invisibility and low robustness. In this case, it is compromised to embed secret message into the intermediate frequency coefficients.

2. Stability. A slight disturbance imposed on the IPD sequence will be dispersed into frequency domain, so there would be no significant impact on

DCT coefficients, and vice versa. This keeps the invisibility and robustness of the imposed covert timing channel based on frequency domain.

All the steganographic systems have some common features, and the most important points are invisibility, robustness and capacity. It is agreed universally that the contradictory unity of these three characteristics makes a system unable to have optimum invisibility, optimum robustness and optimum capacity simultaneously. Improved capacity will inevitably generate decrease of invisibility and robustness, and vice versa. To reach a relatively good performance of the whole system, it is decided to embed 40 bits of secret information into the frequency domain of every 200 IPDs.

3 The proposed scheme

The model of our scheme is demonstrated as follows. Initially, a sample of normal traffic is collected, and the time intervals of the adjacent packets are extracted to get the IPDs. A DCT transform is conducted in IPDs of the normal traffic with the window number of 500, and the size of each window is 200. There are 40 bits of secret information embedded into each window and the promissory frequency coefficients are from No. 101 to No. 140, which are agreed by the sender and the receiver both. Then the encrypted IPDs are got after the transform of IDCT and the covert traffic is sent as the encrypted IPDs. When the receiver receives the traffic, extract the IPDs and conduct a DCT transform first, then get the secret message according to the decoding method.

The process of our scheme is presented in the following Fig. 1:

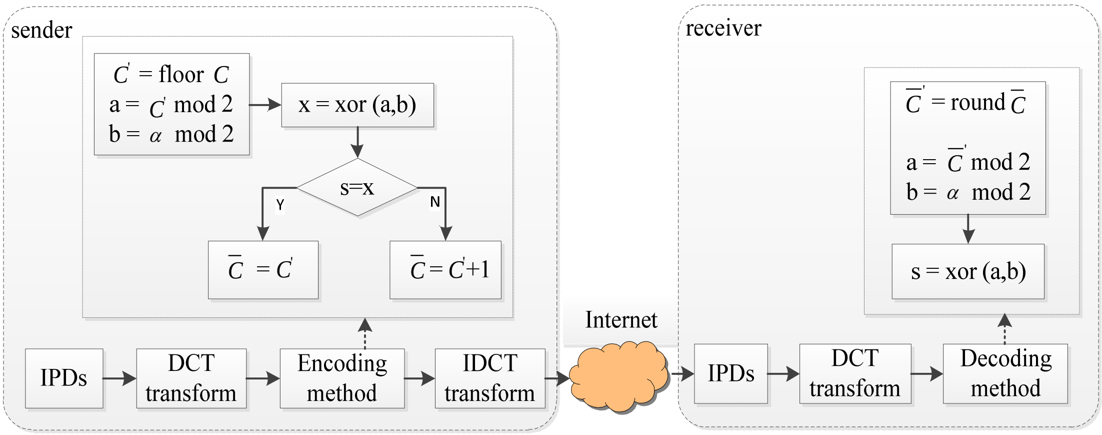


Figure 1 The process of the proposed covert timing channel

3.1 Encoding

When the IPD sequence is transformed into DCT domain, encode the secret message into the DCT coefficients. First, a set of random numbers ranging from 0 to 10 are generated and shared by the sender and the receiver. Round the random number and the corresponding DCT coefficient to determine their parity. Xor the parity of the random number with the parity of DCT coefficient. If the secret bit is the same as the xor result, make the modified coefficient be the integer of the original DCT coefficient. Otherwise, make the modified coefficient be the integer of the original DCT coefficient plus 1.

Let C be the primitive DCT coefficient and $C^{'}$ be the integer of C. Let a be the remainder of the coefficient divided by 2. Let α be the corresponding random number and b be the remainder of α divided by 2. Then the modified coefficient $\overline{C}$ is calculated by $C^{'}$, a, b and the secret information s, which is denoted by Eq.(2):

$$\overline{C}=\begin{cases} C^{'}, & s=a\oplus b \\ C^{'}+1, & s\neq a\oplus b \end{cases} \tag{2}$$

3.2 Decoding

To get the secret information, a DCT transform is conducted in the IPDs after the receiver extracts the time intervals of the covert traffic. Round the DCT coefficients to integers to determine the parity, then xor the parity of the random number with the parity of DCT coefficient to get the secret message.

Let $\overline{C}$ be the DCT coefficient of the IPDs in covert traffic and $\overline{C}^{'}$ be the integer of $\overline{C}$. Let a be the reminder of $\overline{C}^{'}$ divided by 2 and b be the remainder of the random number α divided by 2. Then the secret information s is calculated by a and b, as is denoted by Eq.(3):

$$s=a\oplus b \tag{3}$$

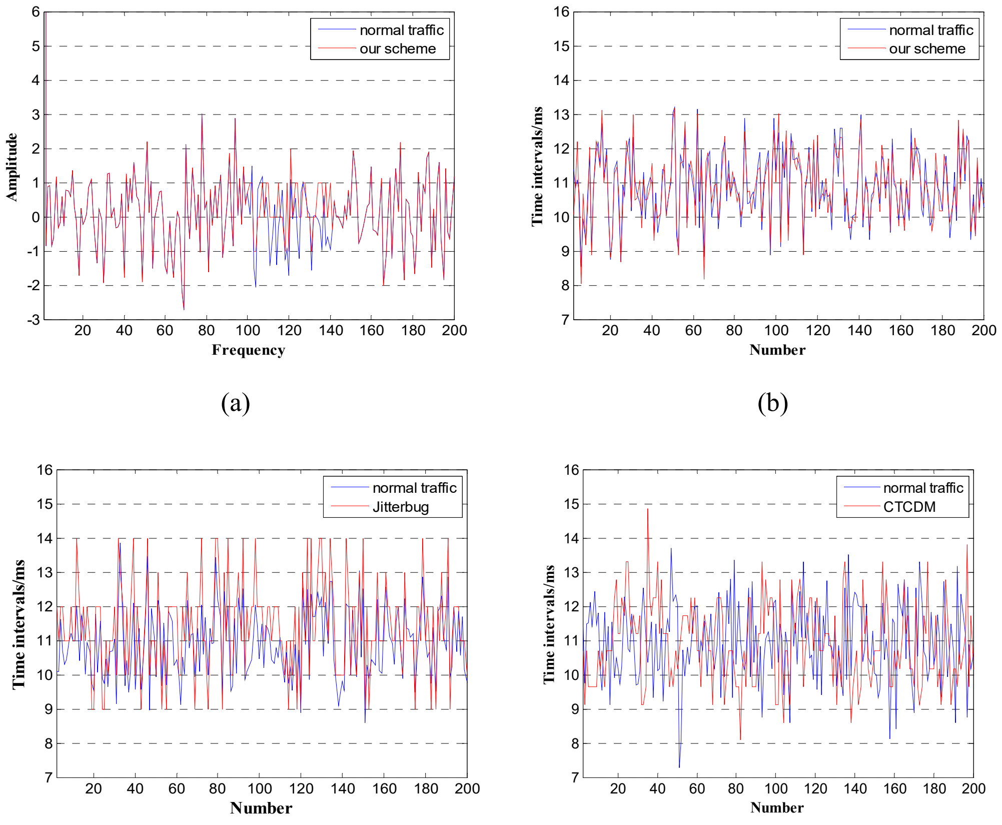
4 Experiments and analysis

The proposed scheme is protocol-independent, which means that any network application can be used as its carrier. Therefore, it is widely applicable in different scenarios. This scheme can be used in any applications. In this paper, the normal traffic of YY audio is captured from an intermediate router during the communication of the two hosts within our campus network of Jiangsu University of Science and Technology.

In this section, experiments are performed to evaluate the concealment of the proposed scheme compared with Jitterbug and CTCDM. The experimental results are analyzed in some significant properties, such as the distribution of IPDs, entropy values and robustness.

4.1 The distribution of IPDs

In this experiment, the whole size of the normal traffic is 100,000 and the normal traffic is departed into 500 windows. The secret information is inserted into the DCT coefficients from number 101 to 140 in each window. The performance of the modified time interval is different from the original one after secret message encoding. The comparison between normal and covert traffic is presented in Fig. 2:



(a) (b)

(c) (d)

Figure 2 The comparison between normal and covert traffic

There are four pictures in Fig. 2. Fig. 2(a) represents the comparison of IPDs in DCT domain between normal traffic and our scheme. It is clear that they only differs slightly where the secret message is embedded. Fig. 2(b), (c), (d) represents the comparison of IPDs in time domain between normal traffic and covert traffic of our scheme, Jitterbug and CTCDM respectively. The distribution of IPDs of our scheme matches well with that of the normal traffic, while that of Jitterbug and CTCDM not. So the proposed scheme performs better in IPD distribution than Jitterbug and CTCDM.

4.2 Entropy test

The concept of 'Entropy' is originated in physics and is used to measure the chaos of a system. The higher the entropy is, the messier the system presents. Entropy is defined by the formula in Eq.(4):

$$H = -\sum_{i=1}^{L} p_i \log p_i \qquad (4)$$

The purpose of entropy test is to detect whether the traffic is covert traffic or not by comparing the entropy value of IPDs of normal traffic and the unknown traffic. In entropy test, the whole traffic of 100,000 IPDs is divided into several windows with w IPDs in each window. IPDs are divided into L bins to compute the entropy value in each window and L is set to be 50. The entropy values of normal and covert traffic are presented in Tab. 1:

Table 1 The comparison of mean and variance of entropy values of IPDs between normal traffic and covert traffic

Window size	*w*	200	500	1000
Normal traffic	Mean	4.7430	4.7309	4.7206
	Variance	0.0170	0.0171	0.0182
Our scheme	Mean	4.7248	3.8046	3,6028
	Variance	0.0160	1.3005	2.0241
Jitterbug	Mean	2.2726	1.9293	1.8239
	Variance	0.0067	0.1552	0.3000
CTCDM	Mean	3.1350	2.5937	2.4552
	Variance	0.0056	0.4422	0.7434

It can be seen in Tab. 1 that entropy values of the normal traffic decrease slightly as window size increases. This pattern applies to the three covert traffic as well with growing gaps. And it is obviously that the entropy value of our scheme is closer to that of the normal traffic than Jitterbug and CTCDM, which means better performance.

The comparison of entropy values of IPDs between normal traffic and covert traffic with different windows is presented in Fig. 3:

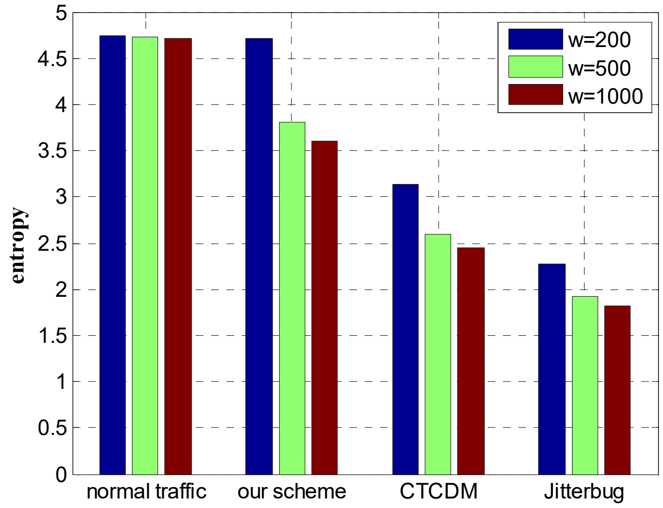


Figure 3 The comparison of entropy values of IPDs between normal traffic and covert traffic with different windows

The threshold value can be set according to Fig. 3 in entropy test. Jitterbug can be detected when threshold value ranges from 2.5 to 4.5, and CTCDM can be detected when threshold value ranges from 3.3 to 4.5. However, it is difficult to set a threshold value to detect the covert traffic of our scheme. It is apparently that our scheme has better performance of detection resistance than the other two covert channels.

4.3 Robustness test

Most timing covert channels are easily affected by network noise, such as jitter and delay. In this experiment, noise of different powers is injected into covert traffic to figure out its robustness. The additive white Gaussian noise is utilized to simulate the channel noise in some aspects. The power of noise is measured by signal-to-noise ratio (SNR) when the power of signal is fixed. The comparison of bit error rate (BER) of the secret message between covert channels is shown in Fig. 4:

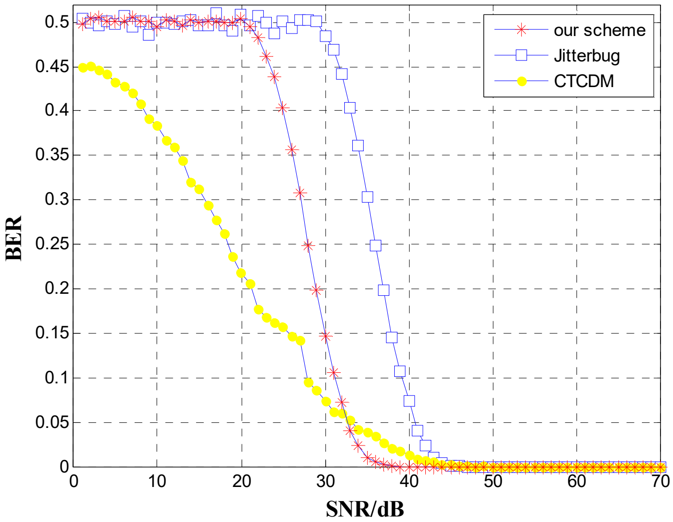


Figure 4 The comparison of BER between covert channels with SNR ranging from 0 to 70

From the result, it can be seen that distortion raised by the noise stays at a high level in Jitterbug till the SNR is 30dB. It achieves relatively well accuracy when the SNR is larger than 48dB. In the method of CTCDM, the value of BER decreases smoothly and slowly until SNR is 50db. While in our scheme, BER begins to go down when SNR is 20dB and reaches 0.01 when SNR is 35dB. Thus, the proposed scheme is robust when the power of noise is less than 10^{-4} of the signal. It's obviously that our scheme has better robustness than the other two methods.

From the above experimental results and analysis, it is manifest that both of Jitterbug and CTCDM can be detected by certain methods, such as entropy test, since the transmission behavior or properties of the original carrier has been significantly altered. However, the performance of our scheme matches perfectly well with the normal one, thus our scheme performs well in detection resistance. Moreover, our scheme has better robustness than the other two typical methods. Therefore, it can be easily concluded that our scheme possesses better performance than the existing methods.

5 Conclusions and future work

In this paper, a network covert timing channel based on the IPDs' frequency characteristic is proposed. The secret information is embedded into the agreed DCT coefficients by making them of the same parity as the shared random numbers or not. From the experimental results, it is indicated that the performance of our scheme is quite close to normal traffic, thus our scheme can successfully evade detection. In addition, our scheme has better robustness than the other two covert channels of

Jitterbug and CTCDM. Hence, the proposed scheme outperforms the existing methods in performance of detection resistance and robustness.

In the future, the capacity of our scheme will be further studied. Besides, the conditions of Internet and packet loss will be considered to evaluate the proposed scheme.

Acknowledgement

The authors greatly acknowledge the anonymous reviewers for their helpful comments and suggestions during the review process, which contribute to improve the quality of the original manuscript.

This work is supported by the NSF of Jiangsu province(Grant No.: BK20150472).

References

[1] W. Mazurczyk, M. Karas, K. Szczypiorski. SkyDe: a Skype-based Steganographic Method[J]. International Journal of Computers, Communications and Control(IJCCC). 2013, 8(3): 389-400.

[2] S. Grabski, K. Szczypiorski. Network steganalysis: Detection of steganography in IEEE 802.11 wireless networks[C]. International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT). 2013:13-19.

[3] J.F. Lalande, S. Wendzel. Hiding Privacy Leaks in Android Applications Using Low-Attention Raising Covert Channels[C]. Eighth International Conference on Availability, Reliability and Security (ARES). 2013: 701-710.

[4] H. Zhao, M. Chen. WLAN covert timing channel detection[C]. Wireless Telecommunications Symposium (WTS). 2015: 1-5.

[5] A. Houmansadr, N. Borisov. CoCo: Coding-Based Covert Timing Channels for Network Flows[J]. Information Hiding, Lecture Notes in Computer Science. 2011, 6958: 314-328.

[6] S. Cabuk, C.E. Brodley, C. Shields. IP Covert Channel Detection[C]. ACM Transactions on Information and System Security (TISSEC). 2009, 12(4): 1-29.

[7] S. Cabuk. Network Covert Channels: Design, Analysis, Detection and Elimination[D]. PHD thesis, Purdue university, USA. 2006.

[8] G. Shah, A. Molina, M. Blaze. Keyboards and covert channels[C]. 15th USENIX Security Symposium. 2006: 59-75.

[9] S. Gianvecchio, H. Wang, D. Wijesekera, *et al.*. Model-Based Covert Timing Channels: Automated Modeling and Evasion[J]. Recent Advances in Intrusion Detection, Lecture Notes in Computer Science. 2008, 5230: 211-230.

[10] G. Liu, J. Zhai, Y. Dai, *et al.*. Network Covert Timing Channel with Distribution Matching[J]. Telecommunication Systems. 2012, 49(2): 199-205.

Author Fang Yang(1989-), female. Postgraduate student(2013-2016), School of Electrics and Information Engineering, Jiangsu University of Science and Technology. Research direction: Network information security.