Journal of Information Hiding and Multimedia Signal Processing ©2016 ISSN 2073-4212
Ubiquitous International Volume **7**, Number **5**, September **2016**

Novel Covert Timing Channels Implemented over Multiple Flows by Using Space-Time Trellis Codes

Lihua Zhang, Guangjie Liu and Weiwei Liu

School of Automation
Nanjing University of Science and Technology
No.200 Xiaolingwei, Nanjing, China
zlh6013@163.com; gjieliu@gmail.com; lwwnjust5817@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
jiangtaozhai@gmail.com, dywjust@163.com

Received December, 2015; revised June, 2016

Abstract. *Existing schemes of covert timing channels mainly exploit a single packet flow between two hosts as carrier and suffer from low reliability in the presence of network interferences. In order to improve the reliability of data transmission, a novel scheme named multiple-input/multiple-output covert timing channel (MIMO) that is implemented over multiple packet flows is proposed in this paper. A Space-Time trellis code (STTC) is adopted to encode secret message data. A receiver recovers secret messages from arriving IPDs by using a decoding mechanism based on Viterbi algorithm. Moreover, the design criterion of the proposed scheme is derived based on the analysis of pairwise error probability. Simulation results for MIMO covert timing channels with 4-ary modulation level and four-state STTC over different numbers of transmission flows demonstrate the significant performance improvement.*
Keywords: Covert timing channel, Reliability, MIMO, Multiple flows, Space-time trellis code.

1. **Introduction.** A network covert channel is one of the parasitic communication systems that clandestinely transfer secret messages and bypass security barriers. There are many network protocols that have been exploited as the carrier of covert channels. Covert channels are not only able to exfiltrate stolen information but also can be used for flow watermarking. Therefore, covert channels techniques have received more attention in recent years. According to the embedding mode, network covert channels could be mainly classified into storage channels and timing channels [1]. In a storage channel, the sender writes a secret message into some fields in a packet that can be read by receiver, such as TTL in IP header [2], reserved fields in TCP header and TCP initial sequence number [3], and so on. However, most of storage channels can be discovered easily by observing the packet headers and eliminated by standardizing the corresponding fields to their defaults [4]. In a timing channel, the sender modulates the sending time of packets according to secret message symbols and the receiver recovers the embedded secret message by observing arriving time of packets. Compared to storage channels, timing channels are more difficult to be detected and eliminated.

Early schemes of covert timing channels aimed to develop efficient methods to embed secret messages into inter-packet delays (IPDs). They cannot evade the detection based on statistical analysis. Then, the performance of detection resistance was taken into account in the designs of covert timing channels and shaping technique was introduced into the procedure of modulation. The transmission rates of covert timing channels are limited by channel interference, such as network jitters, packet losses and disorders. Therefore, forward error correction was exploited to improve the reliability of covert timing channels. However, the adding of some redundancy bits that can provide a level of error detection and correction results in reducing of transmission rate.

In this paper, a novel scheme of MIMO covert timing channels implemented over multiple packet flows is presented. A Space-time trellis code (STTC) that is a method employed to improve the reliability of data transmission by transmitting redundant copies of a trellis code distributed over time and multiple flows is adopted to encode secret message data. A receiver recovers secret messages from the arriving IPDs of the multiple flows by using a decoding mechanism based on Viterbi algorithm. The design criterion of the proposed scheme is derived based on the pairwise error probability analysis. Compared with existing scheme, the reliability of MIMO covert timing channels can be improved by increasing the number of transmission flows.

The remainder of this paper is organized as follows. The related work of covert timing channels is reviewed in Section II. Section III describes the embedding and recovering mechanisms of MIMO covert timing channels. Analysis of pairwise error probability for the proposed scheme is presented in Section IV. Simulation results are reported to evaluate the performance of the proposed scheme in Section V. Finally, the paper is concluded and some interesting areas of future research are discussed in Section VI.

2. **Related Work.**

Covert timing channels transfer secret message symbols by modulated IPDs of carrier packet flows. A lot of work has been done on the design of covert timing channels. On/off timing channel is an early implementation proposed by Cabuk *et al.*, which transmitted a packet to represent 1 and maintained silence to represent 0 during a fixed time interval defined by the sender and receiver before the start of covert communication [5]. Statistical characteristics of IPDs sequence were altered obviously due to the embedding of secret message. In another work, Cabuk *et al.* presented a novel scheme named time-replay covert channel to improve the detection resistance performance [6]. Some inter-arrival time sequences were extracted from the recorded packet flows first. During the covert communication, sending time of packets was adjusted according to secret message symbols based on the offline time sequences.

Shah *et al.* proposed a class of embedding mechanisms named JitterBug that was mainly used for some interactive communication application [7], such as SSH, Telnet, instant messaging, etc. Keyboard JitterBug was a practical implementation that leaked captured messages such as usernames and passwords by adding small delays to keypresses in an interactive network application. It is difficult to distinguish the intentional added delays from inherent network jitters. But the input speed on the keyboard confined the transmission rate of Keyboard JitterBug. Walls *et al.* presented an improved version of JitterBug named Liquid that used a portion of IPDs in the carrier packet flows to smooth out the distortions detected by some detection tests [8]. Each cycle of Liquid IPDs sequence was comprised two parts: transmitting IPDs used for sending secret messages and shaping IPDs used to correct the IPDs probability distortions. JitterBug was used as a base for Liquid to generate transmitting IPDs. In order to provide greater detection resistance, random amounts of sub-millisecond noise were added to each transmitting IPDs, which was absence in JitterBug. Moreover, the sender of Liquid maintained a set

of equally probable bins and determined the bin for each transmitting and shaping IPD that had been sent. Sender kept track of the count for each bin and each generated shaping IPD that was added a additional delay was placed in the bin with smallest count.

Gianvecchio *et al.* proposed a framework to create a model-based covert timing channel (MBCTC) that mimicked the statistical properties of legitimate traffic [9]. IPDs carrying a secret message were generated by using the inverse distribution function of the selected model and recovery was performed by using the cumulative distribution function. The statistical properties of MBCTC traffic were almost the same as that of legitimate traffic. Liu *et al.* proposed a covert timing channel with distribution matching that split network traffic to the flows with fixed length fragment and calculated the probability histogram of IPDs in each fragment [10]. Secret message symbols were embedded into IPDs by a proposed algorithm named binary coding method.

Kothari *et al.* proposed Mimic, a covert timing channel with a shape modeler and a regularity modeler to mimic the shape and regularity of legitimate traffic distributions [11]. The mechanism of shape modeler was similar to the analyzer of MBCTC that fitted a sample of legitimate IPDs to several distributions and selected a distribution model with least root mean squared error. The regularity modeler worked in parallel to the shape modeler and aimed to find the values of nodes that could not exist at each level of the CCE tree, which were used to build regularity tree. IPDs generated by Mimic carried the binary bits of secret message and satisfied the regularity of the medium flow and belonged to the distribution.

Effects of network jitters, packet losses and disorders were not considered in the aforementioned work. In [12], Liu *et al.* applied spreading codes to the modulation of IPDs to increase the Signal-to-Noise Ratio for improving the robustness of covert timing channels at the cost of capacity and camouflage capability. Houmansadr *et al.* tried to use different coding algorithms to eliminate the negative effects of different perturbations [13]. Zhang *et al.* analyzed the effects of packet losses to covert timing channels and proposed a scheme based on Reed-Solomon code and interleaving technique to correct the errors caused by packet losses [14].

Luo *et al.* proposed Cloak, a covert timing channel that used the different combinations of N packets sent over X TCP flows in each round to encode secret messages [15]. There are mainly two possible scenarios for Cloak encoder and decoder to communicate. One of the scenarios is that the encoder connects a remote server with several flows. The decoder eavesdrops at any point of the routing path and recovers the secret messages. The other scenario is that the encoder establish normal links with multiple servers that are dispersed in different locations. Then, the decoder should be located on the common routing path for all servers. In the case of packet losses, Cloak exploited TCP's automatic repeat request technique to provide robust transmission.

The application scenarios of MIMO covert timing channels proposed in this paper are similar to that of Cloak, but the transmission and recovering mechanisms are completely different. Here, secret message data is encoded by a STTC encoder whose outputs are multiple sequences of modulation symbols that are simultaneously transmitted over the same number of transmission flows. STTC that was first introduced by Tarokh *et al.* for wireless communication is able to map binary data to modulation symbols directly for several transmit antennas [16]. So far STTC has been wildly used in various communication systems to provide diversity and coding gain over fading channels. Roy *et al.* used STTC and iterative decoding techniques to obtain high data rates and reliability over underwater acoustic channels [17]. Antonio *et al.* proposed a scheme combining transmit laser selection and STTC for multiple-input/single-output free space optical communication systems over strong atmospheric turbulence channels [18].

3. Proposed Scheme.

3.1. Covert Communication Scenarios.

Two possible scenarios for the MIMO network covert timing channels are depicted in Fig. 1. In Fig. 1(a), the sender establishes a normal session with remote server, consisting of five flows. Secret messages are embedded into the time intervals of the five flows and the receiver eavesdrops at any point of the path and recovers the messages. In Fig. 2(b), the sender establishes normal sessions with multiple servers that are dispersed in different locations and the receiver locates on the common routing path for all servers. The sender can select suitable flows from a large number of flows to different servers for stealthy transmitting secret messages.

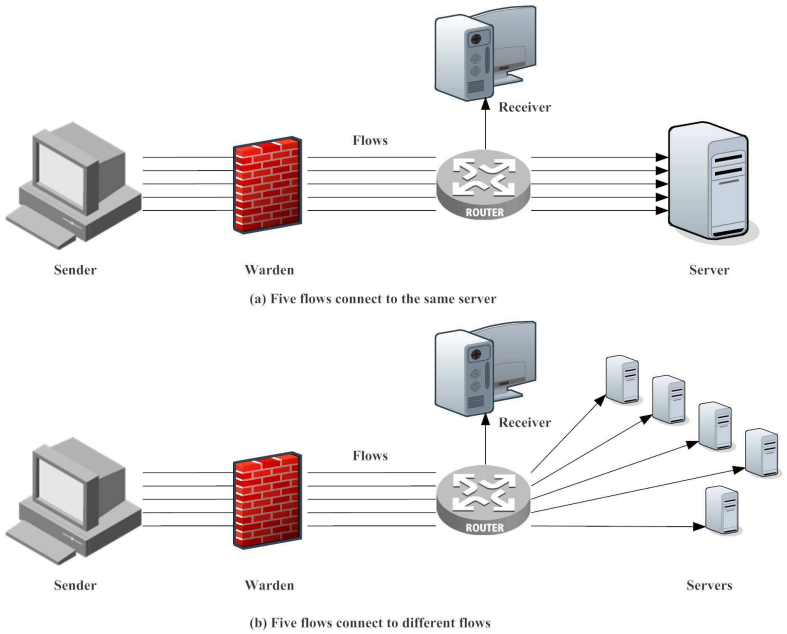


FIGURE 1. Covert communication scenarios for MIMO network covert timing channels

In both scenarios, there is a warden on the senders' network who guards against any network covert channels initiated from inside. The warden could be active or passive. A passive warden attempts to detect network covert channels by analyzing all the traffic sent between the sender and any hosts outside the network. An active warden attempt to interfere with any network covert channels passing through it by altering the packet content or the traffic characteristics. Compared to a active warden, a passive warden does not alter the traffic flows and characteristics.

3.2. Embedding Mechanism.

Assume a MIMO covert timing channel is implemented over n packet flows, secret message data goes through a serial-to-parallel converter, and is divided into m streams of data that are input into a STTC encoder. Fig.2 shows the block diagram of a sender for the MIMO covert timing channel with two transmission flows. Steganographic IPDs are generated in modulators according to the output of STTC encoder. The function of transmitters that send out packets with the intentionally designed IPDs is similar to that of antennas in wireless communication. The binary data of secret message is denoted by $\mathbf{c}$,

$$\mathbf{c} = (\mathbf{c}_0,\ \mathbf{c}_1,\ \mathbf{c}_2,\ \ldots,\ \mathbf{c}_t,\ \ldots) \tag{1}$$

where $\mathbf{c}_t$ is a group of m information bits at time t and given by

$$\mathbf{c}_t = (c_t^1,\ c_t^2,\ \ldots,\ c_t^m). \tag{2}$$

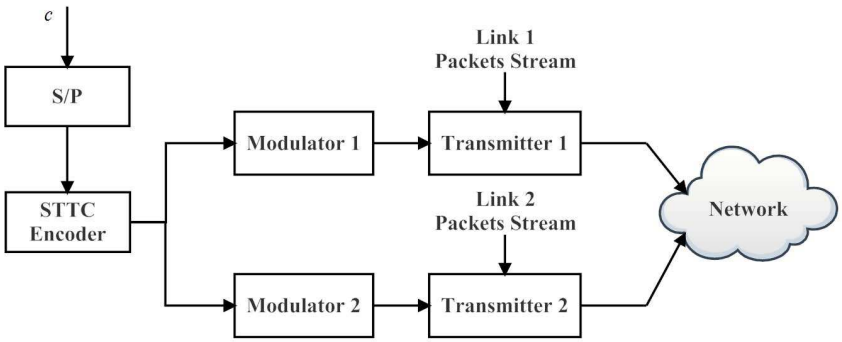


FIGURE 2. The block diagram of a sender for a MIMO covert timing channel with two transmission flows

The structure of STTC encoder that is consisted of m feedforward shift registers is shown in Fig.3. The coefficients of the kth feedforward shift register are denoted by $g^k_{l,\,n}$, $l = 0,\ 1,\ \ldots,\ \eta_k - 1$ and $n = 1,\ 2,\ \ldots,\ N_T$. η_k is the length of the kth feedforward shift registers. The input binary sequence to the kth shift register can be represented as

$$c^k(D) = c^k_0 + c^k_1 D + \cdots c^k_t D^t + \cdots \tag{3}$$

The feedforward generator polynomial for the kth shift register to the ith transmission flow can be written as

$$G^k_i(D) = g^k_{0,i} + g^k_{1,i} D + \cdots g^k_{\eta_k - 1,i} D^{\eta_k - 1} \tag{4}$$

Then, the encoded symbol sequence transmitted over the ith transmission flow can be obtained as

$$x^i = \left[c^1(D)G^1_i(D) + c^2(D)G^2_i(D) + \cdots + c^m(D)G^m_i(D)\right] \bmod M \tag{5}$$

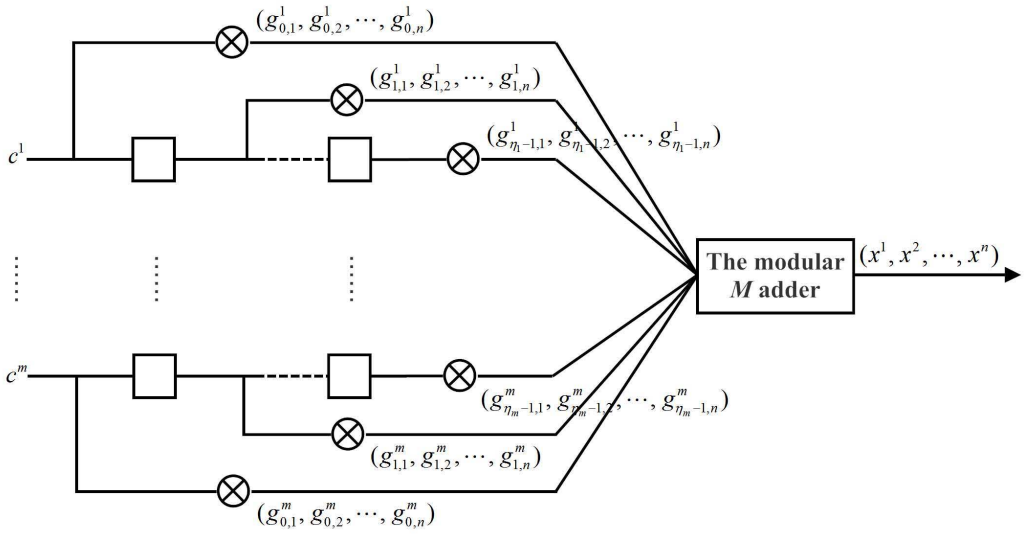


FIGURE 3. The structure of a STTC encoder

Without loss of generality, the STTC encoder with four-state and two transmission flows is taken as example. Assume the coefficients of two shift registers are

$$g^1 = [(0,\ 2),\ (1,\ 0)] \tag{6}$$

$$g^2 = [(2,\ 2),\ (0,\ 1)] \tag{7}$$

The trellis structure for this code is shown in Fig.4. The encoder usually starts at state zero and takes $m = 2$ bits as its input at each time. There are $2^m = 4$ branches leaving from each state corresponding to four different input patterns. Each branch is labeled by $c_t^1 c_t^2 / x_t^1 x_t^2$, where c_t^1 and c_t^2 are a pair of bits input into encoder and x_t^1 and x_t^2 represent two coded symbols transmitted over two transmission flows respectively. The row listed next to a state node in Fig.4 indicates the branch labels for transitions from that state corresponding to the encoder inputs 00, 01, 10 and 11, respectively. Assume that the current state is 00 and the input sequence $\mathbf{c} = (10,\ 00)$, the output sequence generated by the STTC encoder is $\mathbf{x} = (02,\ 01)$, as illustrated by thick line in Fig.4.

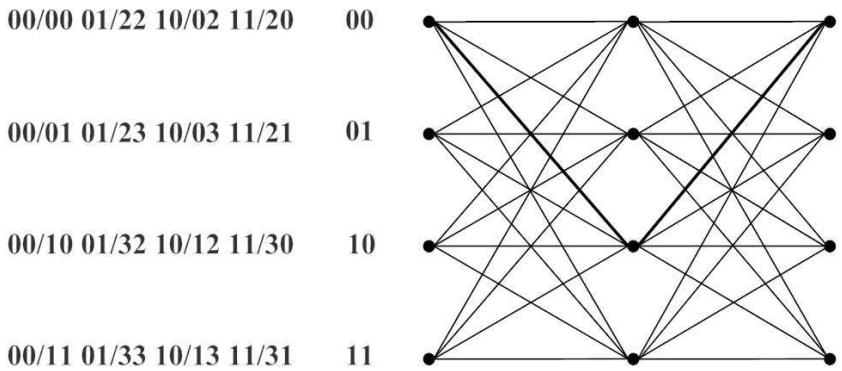


FIGURE 4. The trellis structure for a four-state STTC with two transmission flows

Transmitting IPDs are generated in a manner similar to a MBCTC. Random numbers denoted by v that are generated by a pesudorandom number generator are added to provider even greater detection resistance. Modulator at the ith flow accepts a coded symbol at a time and converts the coded symbol to an IPD as

$$T_t^i = F^{-1}\left(\left(\frac{x_t^i}{M} + v\right) \bmod 1\right), \tag{8}$$

where $F^{-1}(\cdot)$ is the inverse function of cumulative distribution function and M is the modulation order.

3.3. **Recovering Mechanism.** The block diagram of a receiver for a MIMO covert timing channels over two transmission flows is shown in Fig. 5. Demodulators compute a decision statistic based on the arriving IPDs of packets, which are extracted from the corresponding flows by the receivers. The MIMO covert timing channel over n flows can be represented by a channel matrix $\mathbf{H}$. The received IPDs at time t can be expressed as

$$\mathbf{R}_t = \mathbf{H}\mathbf{T}_t + \boldsymbol{\delta}_t, \tag{9}$$

where $\boldsymbol{\delta}_t$ represents the network jitters at time t. For a MIMO covert timing channel proposed in this paper, the channel matrix is deterministic and $\mathbf{H}^H\mathbf{H} = \mathbf{I}$.

Based on the cumulative distribution function denoted by $F(\cdot)$ of IPDs of legitimate traffic, which are shared by sender and receiver, the arriving IPDs of the ith flows are demodulated as

$$\hat{x}_t^i = \left\lfloor M \cdot \left[\left(F\left(R_t^i\right) - v\right) \bmod 1\right] + 0.5 \right\rfloor \tag{10}$$

Viterbi decoder, which aims to find the most likely valid path that starts from state zero and merges to state zero after several time slots, is used to recover secret message data from the output sequences of demodulators. Each node in the trellis shown in Fig. 4 corresponds to a distinct state at a given time, and each branch represents a transition to

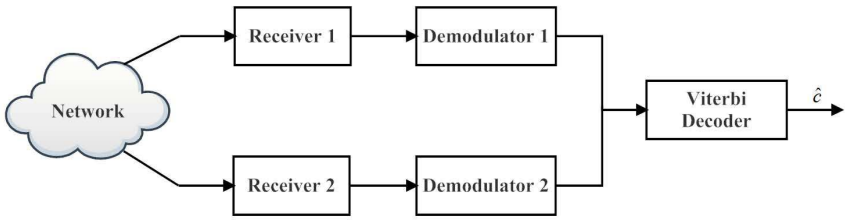


FIGURE 5. The block diagram of a receiver for a MIMO covert timing channel with two transmission flows

some new state at the next time instant. There are several paths formed by branches that reach a node in the trellis. The Hamming distance between the current decoded symbols and the output symbols corresponding to the branch is defined as the metric. The path metric of a valid path is the sum of the branch metrics for the branches that form the path. The most likely path is the one with the minimum path metric and called survivor path.

The Viterbi decoder processes the input sequence in a recursive manner. At the beginning time slot, it computes the partial metric for the single path entering each state and stores the survivor path and its metric for each state. At the next time slot, all survivor paths are extended by one time unit. The partial metrics of the 2^m extended branches entering a state are computed and added to the metrics of the connecting survivor paths. For each state, the metrics of all 2^m paths entering that state are compared and the survivor path along with its metric are stored, all the other paths are eliminated. This procedure is repeated until all the symbols in the input sequence are decoded.

4. **Analysis of Pairwise Error Probability.** The pairwise error probability (PEP) denoted by $P(C^{(0)} \to C^{(1)})$ represents the probability of choosing the codeword $C^{(1)}$ at the receiver side while in fact the codeword $C^{(0)}$ was transmitted at the sender side. For the MIMO covert timing channels proposed in this paper, the PEP can be calculated as

$$\Pr\left\{C^{(0)} \to C^{(1)}\right\} = Q\left(\frac{\left\|(C^{(1)} - C^{(0)})\right\|_F}{\sqrt{2}\sigma_n}\right), \tag{11}$$

where σ_n is the variance of a noise sample. The codeword difference matrix is defined as $\mathbf{D}(C^{(0)}, C^{(1)}) = C^{(1)} - C^{(0)}$ and the codeword distance matrix is defined as $\mathbf{A}(C^{(0)}, C^{(1)}) = \mathbf{D}^H\mathbf{D}$. Suppose $\lambda_i (i = 1, 2, \ldots, N)$ are the eigenvalues of $\mathbf{A}$, $\left\|(C^{(1)} - C^{(0)})\right\|_F^2$ can be calculated as

$$\left\|C^{(1)} - C^{(0)}\right\|_F^2 = tr\{\mathbf{D}^H\mathbf{D}\} = \sum_{i=1}^{N} \lambda_i. \tag{12}$$

The upper bound on the Q function is $Q(x) \leq \frac{1}{2}e^{\frac{-x^2}{2}}$. Therefore, the upper bound of PEP can be expressed as

$$\Pr\left\{C^{(0)} \to C^{(1)}\right\} \leq \frac{1}{2}\exp\left\{-\frac{1}{4\sigma_n^2}\sum_{i=1}^{N}\lambda_i\right\} \tag{13}$$

When the matrix $\mathbf{A}$ is full rank, N equals to the number of transmission flows because none of its eigenvalues is zero. According to the equation 13, it can be found that the increasing of numbers of transmission flows will result in the reducing of PEP when the strength of network interference is fixed. The PEP will increase if some of the eigenvalues λ_i are zeros. Therefore, in order to obtain better error performance of MIMO covert timing channels, the codeword difference matrix $\mathbf{D}$ has to be full rank for any two codewords.

Moreover,the minimum determinant and trace of the codeword distance matrix **A** should be maximized.

5. **Experimental Results.** Simulation experiments are conducted to evaluate the performance of MIMO covert timing channels proposed in this paper. Since there is no open, public network traffic dataset available for testing covert channel schemes, a legitimate traffic dataset that consists of real UDP traffic captured from an online game named World of Warcraft (WOW) is stored in a pcap file that contains 645365 packets from several flows between the host and server by ourself. We implement MIMO covert timing channels with 4-ary modulation level and four-state STTC over various numbers of transmission flows. For a given encoder structure, the generating matrix is determined by minimizing the error probability that is conducted over all possible pairs of paths in the code trellis. The generating matrices used in our experiments are listed in Table 1.

Table 1. Parameters of STTC for different numbers of transmission flows

N_T	v	generating matrix	rank	det	trace
2	2	$g_1 = [(0,\ 2),\ (1,\ 0)]$ $g_2 = [(2,\ 2),\ (0,\ 1)]$	2	4	10
3	2	$g_1 = [(0,\ 2,\ 2),\ (1,\ 2,\ 3)]$ $g_2 = [(2,\ 3,\ 3),\ (2,\ 0,\ 2)]$	2	-	16
4	2	$g_1 = [(0,\ 2,\ 0,\ 2),\ (1,\ 2,\ 3,\ 2)]$ $g_2 = [(2,\ 3,\ 3,\ 2),\ (2,\ 0,\ 2,\ 1)]$	2	-	20

The IPDs of steganographic traffic are generated based on the IPDs of the offline dataset. Steganographic flows will be interfered by the inherent network noise, such as network jitters, packet losses and disorders. The relative interference strength (RIS) of network jitters is measured which can be calculated as

$$RIS = 10\log_{10}\left(\frac{E\left\{|R_i|^2\right\} - E\left\{|J_i|^2\right\}}{E\left\{|J_i|^2\right\}}\right), \tag{14}$$

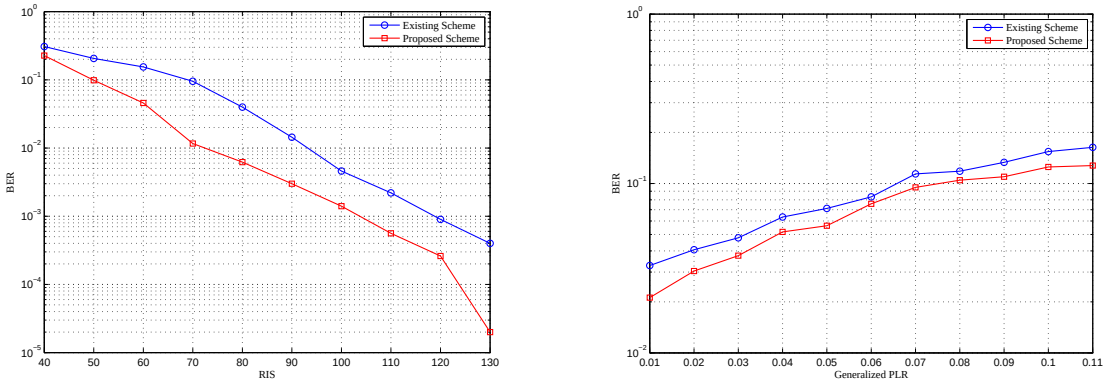
where R_i is the arriving IPDs at the receiver side and J_i denotes network jitters. $E\{\cdot\}$ means the expectation value. The effects of lost and disordered packets to MIMO covert timing channels can be treated with the processing method in [14]. Suppose N_L denotes the number of lost packets and N_D denotes the number of disorder packets, generalized PLR denoted by ρ is defined as

$$\rho = \frac{N_L + N_D}{N_L + N_D + N_R}, \tag{15}$$

where N_R is the number of received packets.

MBCTC is a existing scheme that embeds secret messages into the IPDs of a single flow. Therefore, the error performance of a MIMO covert timing channel implemented over 2 transmission flows is compared to this existing scheme first. Fig.6 shows experiment results under different RISes and generalized PLRs respectively. The average transmission rate is 27bps for these two schemes when they are implemented over the WOW traffic. The BERs of both schemes decrease monotonically with the increasing of RIS and increase monotonically with the increasing of generalized PLR. It is clear that the BER of the proposed scheme is always smaller than the existing scheme.

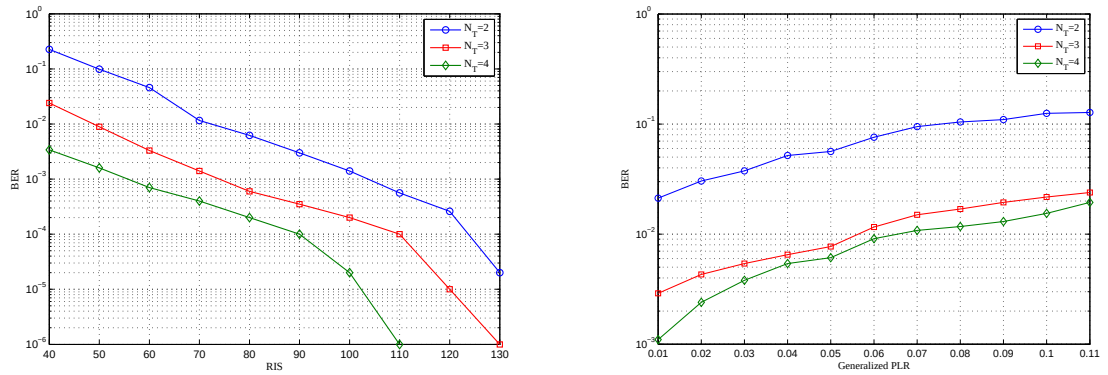
The error performances of MIMO covert timing channels implemented over different numbers of transmission flows in the presence of various relative interference strengths



(a) under different RISes (b) under different generalized PLRs

FIGURE 6. BER performance comparison between a MIMO covert timing channel and a MBCTC

and generalized PLRs are shown in Fig.7. It can be observed that BER of covert timing channels implemented over four transmission flows is lowest when the RIS and generalized PLR are fixed. Experimental results show that increasing the number of transmission flows can provide a significant performance improvement.



(a) under different RISes (b) under different generalized PLRs

FIGURE 7. BER Performances of MIMO covert timing channels implemented over different numbers of transmission flows

Implementing MIMO covert timing channels over multiple transmission flows provides better performance at the expense of a higher decoding complexity, while the increasing of encoding complexity is small. At each time slot, the decoding algorithm need calculate 2^{N_T} branch metrics and find the path with the smallest total metric for every state. Therefore, the decoding complexity grows linearly with the number of states and exponentially with the number of transmission flows. However, the priority of reliability is higher than complexity for covert communication. The designer of covert timing channels should make a trade-off in terms of throughput, complexity and error performance. The number of transmission flows and modulation order should be determined according to the specific requirement.

The detection resistance performance is evaluated by using shape and regularity tests. Kolmogorov-Smirnov (KS) test whose statistic measures the maximum distance between

the empirical distribution functions of two samples is commonly used to examine the distribution shape of IPDs of covert timing channels. The regularity test examines the second-order or higher statistics of detected traffic, while the shape test is based on the first-order statistic [20]. Regularity measures how far the variances of the detected IPDs are spread out. A small regularity indicates that the detected samples tend to be very close to the legitimate samples. Correct conditional entropy (CCE) is another regularity statistic that examines the correlations in the detected samples. A histogram of IPDs with several bins is constructed and the ranges of bins are determined by a large training set of IPDs such that each bin has an equal number of training IPDs. The IPDs of the detected traffic are binned using the histogram. A legitimate traffic is expected to have a nearly uniform distribution among the bins.

In this experiments, IPDs of the MIMO covert timing channel with two transmission flows are selected as steganographic sample. The three aforementioned statistics are computed for the window sizes of 1000 and 2000 and compared with these of legitimate sample. The mean and standard deviation of these statistics are shown in Table 2. It can be observed that the KS distances for steganographic samples are very close to the legitimate samples. This is because the generating of IPDs in MIMO covert timing channels are based on MBCTC, which is able to mimic the statistical properties of legitimate traffic well. The regularity and CCE scores can be improved further by introducing regularity tree into the modulation procedure.

Table 2. Statistics of shape and regularity tests

Detected Sample	Window Size	KS Distance		Regularity		CCE	
		mean	stdev	mean	stdev	mean	stdev
Legitimate sample	1000	0.0022	0.001	0.6273	2.4262	2.1321	0.2726
	2000	0.0012	0.0008	0.6943	2.3918	2.1283	0.2245
Steganographic sample	1000	0.0023	0.0014	0.7185	2.0345	2.3564	0.2111
	2000	0.0015	0.0007	0.5879	1.8575	2.3564	0.2074

6. **Conclusions.** In this paper, we have designed a novel scheme of covert timing channels based on multiple packet flows, which is named as MIMO covert timing channel. In order to mitigate the effects of inherent network noise, Space-time trellis code is adopted to encode secret message data. At the receiver side, Viterbi algorithm is used to recover the embedded secret message. A series of simulation experiments have been conducted and results show that the proposed scheme is able to embed secret message into IPDs of multiple packet flows. Increasing numbers of transmission flows will improve the robustness of covert timing channels. The performance of detection resistance for MIMO covert timing channels is examined by mainly shape and regularity tests methods and the tests scores are very close to legitimate samples.

Acknowledgment. This work is supported by Nature Science Foundation of China (Grant No.61170250, 61103201 and 61472188), the fundamental research funds for the central universities (No.30920140121006, 30915012208) and NSFC of Jiangsu Province (No.BK20150472). The authors also gratefully acknowledge the helpful comments and suggestions of the reviewers, which have improved the presentation.

REFERENCES

[1] S. Wendzel, S. Zander, B. Fechner, and C. Herdin, Pattern-based Survey and Categorization of Network Covert Channel Techniques, *ACM Computing Surveys*, vol. 47, p. 50, 2015.

[2] S. Zander, G. Armitage, and P. Branch, A survey of covert channels and countermeasures in computer network protocols, *Communications Surveys & Tutorials, IEEE*, vol. 9, pp. 44-57, 2007.

[3] H. Zhao and Y. Shi, Detecting Covert Channels in Computer Networks Based on Chaos Theory, *IEEE Transactions on Information Forensics and Security*, vol. 8, pp. 273-282, 2013.

[4] S. J. Murdoch and S. Lewis, Embedding Covert Channels into TCP/IP, *presented at the 7th International Workshop on Information Hiding, Barcelona, Spain*, 2005.

[5] S. Cabuk, C. E. Brodley, and C. Shields, IP Covert Timing Channels: Design and Detection, *presented at the Proceedings of the 11th ACM conference on Computer and Communications Security, Washington, DC, USA*, 2004.

[6] S. Cabuk, C. E. Adviser-Brodley, and E. H. Adviser-Spafford, Network covert channels: design, analysis, detection, and elimination, *Doctor of Philosophy, Purdue University*, 2006.

[7] G. Shah, A. Molina, and M. Blaze, Keyboards and covert channels, *presented at the Proceedings of the 15th conference on USENIX Security Symposium, Vancouver, B.C., Canada*, 2006.

[8] R. J. Walls, K. Kothari, and M. Wright, Liquid: A detection-resistant covert timing channel based on IPD shaping, *J. Computer Networks*, vol. 55, pp. 1217-1228, 2011.

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

[10] G. Liu, J. Zhai, and Y. Dai, Network covert timing channel with distribution matching, *J. Telecommunication Systems*, vol. 49, pp. 199-205, 2010.

[11] K. Kothari and M. Wright, Mimic: An active covert channel that evades regularity-based detection, Computer Networks, vol. 57, pp. 647-657, 2013.

[12] Y. Liu, D. Ghosal, F. Armknecht, A. R. Sadeghi, S. Schulz, and S. Katzenbeisser, Hide and seek in time-robust covert timing channels, *presented at the 14th European Symposium on Research in Computer Security, Saint-Malo, France*, 2009.

[13] A. Houmansadr and N. Borisov, CoCo: coding-based covert timing channels for network flows, *in 13th Information Hiding Conference, ed. Prague, Czech Republic: Springer*, 2011, pp. 314-328.

[14] L. Zhang, G. Liu, J. Zhai, and Y. Dai, Improving Reliability of Covert Timing Channel to Packet Loss, *Journal of Information Hiding and Multimedia Signal Processing*, vol. 6, pp. 544-553, 2015.

[15] X. Luo, E. Chan, P. Zhou, and R. Chang, Robust Network Covert Communications Based on TCP and Enumerative Combinatorics, *IEEE Transactions on Dependable and Secure Computing*, vol. 9, pp. 890-902, 2012.

[16] V. Tarokh, N. Seshadri, and A. R. Calderbank, Space-time codes for high data rate wireless communication: Performance criterion and code construction, *IEEE Transactions on Information Theory*, vol. 44, pp. 744-765, 1998.

[17] S. Roy, T. M. Duman, V. McDonald, and J. G. Proakis, High-rate communication for underwater acoustic channels using multiple transmitters and spaceCtime coding: Receiver structures and experimental results, *IEEE Journal of Oceanic Engineering*, vol. 32, pp. 663-688, 2007.

[18] A. Garcia-Zambrana, C. Castillo-Vazquez, and B. Castillo-Vazquez, Space-time trellis coding with transmit laser selection for FSO links over strong atmospheric turbulence channels, *Optics express*, vol. 18, pp. 5356-5366, 2010.

[19] Z. Chen, J. Yuan and B. Vucetic, An improved space-time trellis coded modulation scheme on slow Rayleigh fading channels, , *presented at 2001 IEEE International Conference on Communications, Helsinki, Finland*, pp.1110-1116, 2001.

[20] S. Gianvecchio; and H. Wang;, An Entropy-Based Approach to Detecting Covert Timing Channels, *IEEE Transactions on Dependable and Secure Computing*, vol. 8, pp. 785-797, 2011.