# A Novel Timing-based Network Covert Channel Detection Method

# A Novel Timing-based Network Covert Channel Detection Method

**Shoupu Lu[1,2], Zhifeng Chen[1*], Guangxin Fu[3] and Qingbao Li[1]**

[1]State Key laboratory of Mathematical Engineering and Advanced Computing, Zhengzhou, Henan Province, 450001, China

[2]Henan University of Economics and Law, Zhengzhou, Henan Province, 450000, China

[3]P.O.Box 5111, Beijing, China

[*]Corresponding author's e-mail: xiaohouzi06@163.com

**Abstract.** Network stealth events are endless, and covert timing channel is one of the most difficult means to prevent. In order to further improve the detection rate of covert timing channel, several typical network covert timing channel construction algorithms are analyzed. On the basis of the above analysis, a detection method based on IPDs multidimensional features was proposed in this paper. IPDs of covert timing channels from three dimensions: shape, change rule and data statistics are analyzed. Respectively, polarization feature, autocorrelation feature, clustering feature are proposed, and the three features are unified into a model. The threshold method is used to determine whether the channel to be detected is a normal channel. Experiments show that the method can detect the existing covert timing channel with less time cost and compared with the traditional detection method has a certain rate of improvement.

## 1. Introduction

At present, the network theft behavior shows a high degree of concealment and camouflage, which poses a serious threat to the computer network security[1]. Among the many network stealth problems, the network covert channel technology is one of the most difficult to detect. The network covert channel is a way of violating the original intention of the system designer to steal information. It can reside in the user's computer for a long time and send sensitive data to the outside world using the normal network communication channel to bypass the operation System defense mechanism[2]. Therefore, the detection of network covert channel for user privacy data protection and the country's cyberspace security has important significance.

In view of the detection of network covert channels, academia and industry have carried out extensive research and developed corresponding tools and systems. Covert channels are usually classified into two broad categories: storage-based (that hide secrets within legitimately communicated data) and timing-based (that simply manipulate the timing of access to certain hardware or software resources). In the detection of covert timing channel, the typical method is network scene detection method proposed by Helouet[4]. But the actual network of the scene is too complex, this detection algorithm performance overhead is too large. In order to achieve higher detection accuracy and smaller false positive rate, the detection algorithm based on the rules and the statistics have been proposed[5,6]. In[7], Cabuk et al. used the fingerprints left by covert channel embedding process into

the traffic stream and designed one metric around the distribution of the inter-packet arrival time. The first metric was built on the premise that covert timing channels would create regular patterns on inter-packet delays sequence(IPDs) whereas a regular traffic might be possibly bursty and hence lack such patterns. By observing the distribution of the traffic for these patterns, they are able to detect the covert channels. In [8], the authors investigate the p-values of two nonparametric statistical tests as detection parameters called K-L test and Welch's T-test. Although the mentioned parameters and statistical tests are shown to detect one or two given CTC algorithms, these parameters might not be effective in detecting other covert timing channel. Therefore, Shrestha et al. in [10] utilize a combination of the parameters that are introduced in [8] and [9] to be able to detect the majority of covert timing channel algorithms using Support Vector Machine (SVM) classification. However, this approach requires extensive calculations and previous knowledge of network IPD metrics to treat the classifier. Fahimeh et al. in [11] present and leverage three different non-parametric statistical tests that can be used to generate very different statistical test scores for overt and covert traffic IPDs. Based on the shape detection and the law detection, the shape of IPDs can be described by first-order statistics. The law of IPDs can be described by high-order statistics. Based on this, they proposed the use of entropy and modified conditional entropy to detect covert channels. The entropy rate can describe the shape of IPDs, and the modified condition entropy can describe the regularity of IPDs. Though these method can effectively detect the vast majority of covert timing channel, they have a high false positive rate.

In response to these problems, in this paper, several typical network covert timing channel construction algorithms are analyzed. On the basis of the above analysis, a method based on IPDs multidimensional feature is proposed to detect network covert timing channel. IPDs is described by three dimensions: shape feature, change rule and statistical feature, and use the threshold method to determine whether the channel to be detected is a normal channel.

## 2. Analysis of IPDs Distribution in Classical Covert timing Channel

IPDs distribution of the normal channel usually follows a random distribution. In order to enable the receiver to identify the hidden information, the covert timing channel algorithm needs to change IPDs distribution of the normal channel, so that its distribution has clear characteristics. Therefore, the potential randomness of IPDs distribution of the normal channel will be broken, and values of the covert channel IPDs will be within a certain range. This range is closely related to the network connection jitter and bit error rate. Assuming that IPDs of the normal channel are taken as random variable X, the network jitter is defined as $\lambda(\mu, \delta) = \{\lambda_1, \lambda_2, \dots, \lambda_n\}$, Where μ is the mean value and δ is the standard deviation, then the value of the covert channel IPDs is the random variable Y, and Y is defined as:

Where $\Delta_c$ is the delay added in the covert channel.

Assuming that the embedded string is $B_{Lc}$, and the length is $L_c = n + m$,including $n$ bits 1 and $m$ bits 0, thus, $B_{Lc}$ can be defined:

$$B_{Lc} = b_m^{(0)} b_n^{(1)} \ , 2 \le n \le L_c \ , 0 \le m \le L_c - 2$$

IPDs distribution of four classical covert channels is analyzed below.

### *2.1. On-Off Covert Timing Channel*

The sender and receiver of this covert channel [8] share a time interval $T_c$. If the sending covert bit is 1, wait for a period of time $T_c$ to send a packet. If sending the covert bit is 0, send a packet within $T_c$. In order to define IPDs of such covert channels, the embedded string has at least two hidden bits 1, that is, at least two packets within the specified $T_c$. IPDs distribution is defined as follows:

$$\{y_1, y_2, \dots, y_n | y_i \in \{T_c, 2T_c, 3T_c, \dots, (k+1)T_c\}\}$$

Where $k$ represents the number of consecutive hidden bits 0.

*2.2. L-bits to N-packets Covert Timing Channel*

This concealment algorithm embeds the hidden information of L bits into N different IPDs, which can improve the capacity of the channel and reduce the bit error rate [12]. IPDs distribution satisfies the following definition:

$$\{y_1, y_2, \dots, y_n | y_i \in \{D, D+d, D+2d, \dots, D+2^L \cdot d\}\}$$

Where $1 \le L \le L_c$ ,$1 \le N \le L_c$, $D$ is the base delay and $d$ is the additional delay due to the need to allocate the L-bits string to N intervals.

*2.3. Jitterbug*

Shah et al. [13] designed a network covert timing channel called Jitterbug, which manipulates network packets of legitimate channels and adds additional latency to these network packets. The sender and receiver share a value $w$. If covert bit 0 is sent, IPD can be divisible by $w$. If covert bit 1 is sent, IPD can be divisible by $[w/2]$. IPDs distribution satisfies the following definition:

$$\{y_1, y_2, \dots, y_{L_c} | y_i \in \{\frac{w}{2}, w, \frac{3w}{2}, \dots, \frac{nw}{2}\}\}$$

*2.4. Time Replay Channel*

Time Replay type covert timing channel [14] uses IPDs on the overt channel as it's a priori data to mimic the randomness of the overt channel IPDs distribution. However, IPDs distribution of the two types channel is not exactly the same. Mainly because of the existence of network jitter, covert channel IPDs can not take the a priori data on the boundary. Assuming covert bit 0 and covert bit 1 are respectively encoded into IPDs of two prior data, IPDs of the prior data are defined as $S_0 = \{T_{cb1}^1, T_{cb1}^2, \dots, T_{cb1}^n\}$, $S_1 = \{T_{cb2}^1, T_{cb2}^2, \dots, T_{cb2}^n\}$, and $S_0$, $S_1$are sorted in ascending order. Since there exists network jitter of $\lambda(\mu, \delta)$, then it may appear $S_0 + \lambda(\mu, \delta) \ge S_1$, causing the recipient to be wrong. Assuming that IPDs of the prior data are embedded in the k-coded information, $\{T_{cb1}^1, T_{cb1}^2, \dots, T_{cb1}^{n-r}\}$,$\{T_{cb2}^1, T_{cb2}^2, \dots, T_{cb2}^{m-r}\}$ , …,$\{T_{cbk}^1, T_{cbk}^2, \dots, T_{cbk}^{p-r}\}$ , some of the boundaries of IPDs will be ignored, in order to overcome the impact of network jitter. Therefore, its distribution satisfies the following definition:

$$\{y_1, y_2, \dots, y_{\mathrm{Lc}} | y_i \in \{T_{cb1}^1, T_{cb1}^2, \dots, T_{cb1}^{n-r}\} \cup \{T_{cb2}^1, T_{cb2}^2, \dots, T_{cb2}^{m-r}\} \cup \{T_{cbk}^1, T_{cbk}^2, \dots, T_{cbk}^{p-r}\}\}$$

According to the above analysis, we can draw the following conclusions. Because covert timing channel algorithm randomly modulates the hidden information into IPDs, it leads to a certain degree of difference from IPDs distribution of the overt channel. In fact, covert timing channel compared with the overt channel, on the one hand will add some delay to IPDs, on the other hand will change the distribution of IPDs. Therefore, we can analyse the three dimensions from the shape feature, the data change rule and the statistical feature of the data. By comparing the difference of the Eigenvalues in the channel, we can distinguish the overt channel and the covert channel.

## 3. MULTIDIMENSIONAL FEATURES

According to the above analysis, we propose polarization characteristics, autocorrelation characteristics and clustering characteristics in three dimensions: IPDs shape, data variation rule and data statistics.

*3.1. Polarization Feature Based on IPD Shape*

In smooth network environment, for the overt channel of IPDs, in which most IPD values are small and a few IPD values are large. And the covert timing channel will delay IPDs, the general increasing the value of IPD. Therefore, the overt channel is more polarized in the shape of IPDs.

Economics often use the "Gini coefficient" to determine a country's income distribution fairness. This paper is inspired by the Gini Coefficient, and uses a similar method to define and calculate the degree of polarization of IPDs. Assuming that the original network packet interval is $T =< t_1, t_2, \dots, t_N >$ , the sequence obtained after its sorting is recorded as $T^{\wedge} =< t_1^{\wedge}, t_2^{\wedge}, \dots, t_N^{\wedge} >$ .

Constructor function is $F(x) = \hat{t_x}$ $(x = 1,2, \dots, N)$, and the function is drawn on the Cartesian coordinate system. The point $(1, \hat{t_1})$ and the point $(N, \hat{t_N})$ is connected with a straight line L, then the four straight lines of L, X axis, X = 1, and X = N will form a trapezoid, and the trapezoidal area is S. Connect the points $(i, \hat{t_i})$and points $(i+1, \hat{t_{i+1}})$in turn to form a polygonal line G, then the four lines of the curve G, X axis, X = 1, and X = N will be enclosed in an area, and the area of the area is B. At the same time, the area of the area surrounded by the polygonal line G and the straight line L is denoted as A. Then we can think of A≈S-B. Define the specific gravity of A in S is the polarization coefficient Polar, which can define the polarization coefficient Polar as follows:

$$\text{Polar} = \frac{A}{S} = \frac{|S-B|}{|S|} = |1 - \frac{B}{S}| \quad (1)$$

*3.2. Clustering Feature Based on IPD Data Statistics*

The overt channel is often affected only by the network environment. IPDs is relatively stable over a certain period of time. If the observation is continuous in this period, the difference of each group after clustering is small. And IPDs of the covert channel will be affected by the hidden data sent. In a certain period of time, the difference between the clustering groups is larger. Therefore, the overt channel has more stable clustering feature, and the clustering feature of the covert timing channel is unstable.

Clustering can be described as: The original IPDs $T =< t_1, t_2, \dots, t_N >$ is divided into k similar subsets$\{C_1, C_2, \dots, C_k\}$. Classes generated by clustering are collections of a set of data objects, and objects in the same class are similar to each other, and objects between classes are different from each other. The set can be abstracted as another representation, expressed as a binary:

$$C_i =< mean, amount > \quad (2)$$

$mean$ indicates the cluster center of the cluster, and $amount$ indicates the number of data in the cluster. Respectively, $C_i.mean$ and $C_i.amount$indicate the clustering center vector and the number of cluster elements in the cluster $C_i$. The differences between groups are generally measured using Euclidean distance. The Euclidean distance is positively related to the difference. The degree of similarity between them can be measured by the distance Dis (i, j). Considering two groups $C_i$ and$C_j$, each group has N clusters ($i \neq j$,$k = 1,2, \dots, N$). Therefore, we can obtain the two groups Euclidean distance:

$$\text{Dis} = \sqrt{\sum_{k=1}^{N}(C_i.amount_k - C_j.amount_k)^2} \quad (3)$$

The two IPD sequences are clustered for a certain period of time, getting two groups $C_i$and $C_j$. If the Dis value is relatively small, then the two groups are relatively small differences. In this paper, we chooses the K-means clustering algorithm.

*3.3. Autocorrelation Features Based on IPD Variation*

The autocorrelation coefficient can be used to measure the degree of similarity of a sequence at different moments. Since the distribution of IPDs of the overt channel is highly dependent on the communication mode between the server and the client, there is a relatively high autocorrelation. And the covert timing channel randomly modulates the hidden information into the IPD, so the distribution of IPDs of the covert channel will also show a large randomness. Then the covert timing channel in the data changes rule will have a lower self-correlation. Assuming that the original IPDS is$T =< t_1, t_2, \dots, t_N >$, calculate the mathematical expectations of the sequence $\mu = (t_1 + t_2 + \cdots + t_N)/N$, standard deviation $\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(t_i - u)^2$. Then it can get its autocorrelation function :

$$R(\tau) = \frac{E[(t_i-\mu)(t_{i+\tau}-\mu)]}{\sigma^2} i = 1,2, \dots, N \quad (4)$$

$$Cr = max(R(\tau)) \quad (5)$$

## 4. DETECTION METHOD

Network covert timing channel detection method of the work process is shown in figure 1. IPDs as input parameters, followed by shape feature analysis, statistical feature analysis and data variation analysis to calculate the polarization index Polar, clustering index Dis and autocorrelation index Cr, and induce the three indicators into the channel M0 parameters. Since the covert timing channel requires the hidden information to be modulated into IPDs, it is necessary to change the channel IPDs shape and data statistics. Compared with the covert timing channel, the clustering feature, polarization feature and autocorrelation feature are very different. Therefore, the three features of the overt channel compared to the covert channel should have the following three characteristics:

1) The value of Dis is relatively small
2) The value of Polar is closer to 1
3) Cr value will be relatively large

Use the eigenvalue threshold method to determine whether a channel to be detected is normal. Model the overt channel, analyzing the range of Dis, Polar, Cr in the model, and then select the threshold as the M0 attribute. The channel M1 can be detected by the formula.

$$\begin{cases} M_1.dis \geq M_0.dis \\ M_1.polar \leq M_0.polar \\ M_1.cr \leq M_0.cr \end{cases} \quad (6)$$

If the relationship of the two channels satisfies the formula 6, it indicates that the channel M1 is normal. If the relationship between the two channels does not meet the formula 6, then the M1 may be abnormal, the system will alarm. This method based on the threshold brings more prone to false positives, but the false positives rate in a certain range is acceptable.

According to the above analysis and design, network covert timing channel detection algorithm based on IPDs multidimensional feature is implemented. T denotes IPDs, N denotes the size of the detection window, S denotes the update step size, m denotes the number of consecutive calculations, Dis denotes the distance between two groups, Polar denotes the polarization coefficient, and Cr denotes the autocorrelation coefficient.

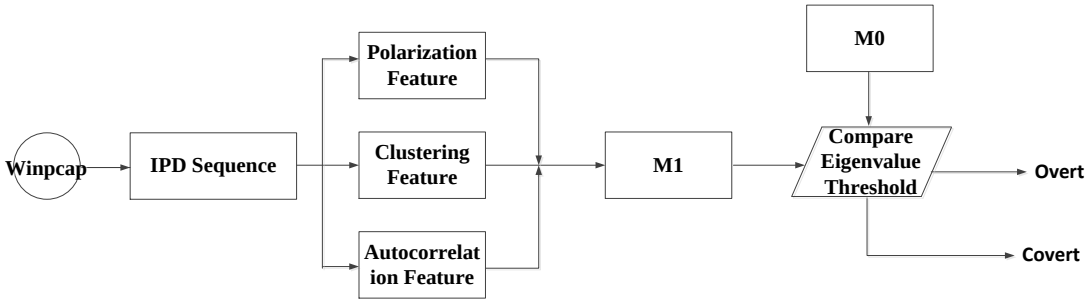


Figure 1. Detection System Work Process.

## 5. EXPERIMENTAL RESULTS AND ANALYSIS

The experimental environment established in this paper is shown in the following figure 2. The IP address is 172.16.1.2-172.16.1.5 is the sender of the covert channel, 172.16.1.6 is the sending end of the overt channel, 172.16.1.7 is the detector, 172.16.6.1 is the receiving end.

### *5.1. Data Set*

In the experiment, four typical covert timing channels are implemented according to the related literature: IPCTC, MBCTC, Jitterbug, Liquid, the first two are active covert timing channels, and the latter two are passive covert timing channels. IPCTC cycles using {40ms, 60ms, 80ms} three different time intervals to construct a covert channel, and MBCTC fits the distribution model based on 20,000 normal HTTP datagram time intervals. For passive covert channels, this paper replays the SSH packets and delays the packets according to the coding algorithm.

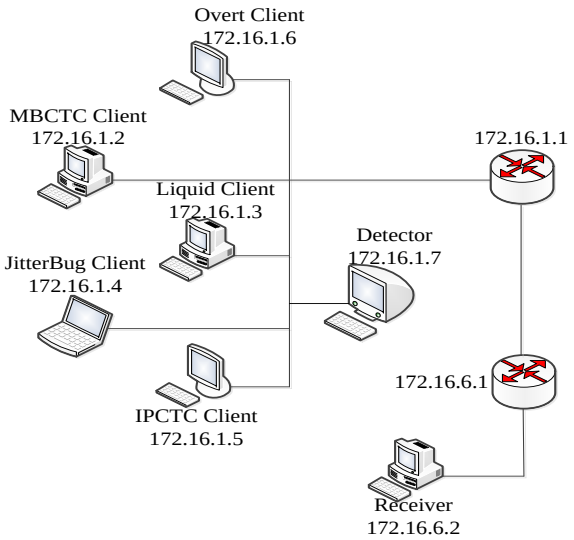


Figure 2. Experiment Environment.

First, the overt channel client sends packets, after collecting a certain number of IPDs, the IPCTC client sends packets. The above process is repeated to form a mixed data stream as shown in figure 3, where the data stream contains a total of 80,000 IPDs, each of which is 10,000 IPDs.

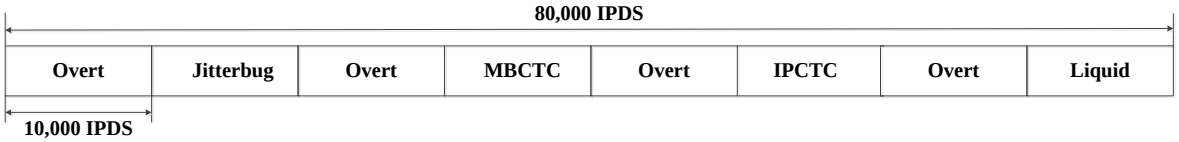


Figure 3. Experimental data flow.

*5.2. Scenarios Parameters and Feature Thresholds*

Table 1. Scenarios Parameters Setting.

| Environment | N | Step_Size | m |
|---|---|---|---|
| P1 | 1000 | 100 | 10 |
| P2 | 1000 | 100 | 20 |
| P3 | 2000 | 100 | 10 |
| P4 | 2000 | 100 | 20 |
| P5 | 4000 | 200 | 10 |
| P6 | 4000 | 200 | 20 |
| D1 | 1000 | 100 | 10 |
| D2 | 1000 | 100 | 20 |
| D3 | 2000 | 100 | 10 |
| D4 | 2000 | 100 | 20 |
| D5 | 4000 | 200 | 10 |
| D6 | 4000 | 200 | 20 |
| C1 | 1000 | 100 | 10 |
| C2 | 1000 | 100 | 20 |
| C3 | 2000 | 100 | 10 |
| C4 | 2000 | 100 | 20 |
| C5 | 4000 | 200 | 10 |

In order to accurately detect the covert channel, we need to select the appropriate detection parameters. First, need to determine the detection window, that is, the number of packets detected by the detector at one time. Taking into account the speed of online detection, the value of the detection window can not take too much, but if the value of the detection window is too small, the clustering method can not work. The other two important parameters are the update step size and the value of m. It is difficult to prove the three parameters value of theoretically, only through the experimental test method to determine the size. The parameters in the test are set in table 1.

It can be seen from the data in figure 4, in the scene of IPDs length of 1000, the step length of 100, m value of 10, the change of the three normalized eigenvalues is not obvious. Especially the polarization coefficient of the overt channel, the value of the partial region has been close to the value of the polarization coefficient of the covert channel. Because the network jitter or delay causes the value of some of IPDs in the overt channel to be too large, resulting in a large polarization coefficient. Therefore, if detection is performed in this scenario, a higher detection rate can not be obtained.

It can be seen from the data in figure 5, the values of the three normalized eigenvalues of the overt and covert channels in this scenario have a distinct dividing line compared to scene 1. Therefore, the overt channel and the covert channel can be distinguished, but the fluctuation range of the three eigenvalues in the scene is relatively large, which leads to the instability of the detection result. So under this scenario, it causes more prone to a high false positive rate.

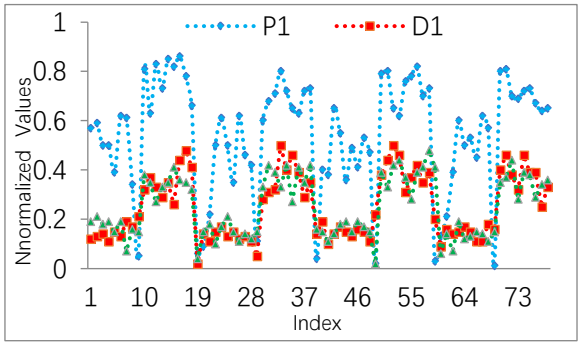


Figure 4. normalized values Under T1 Scenarios.

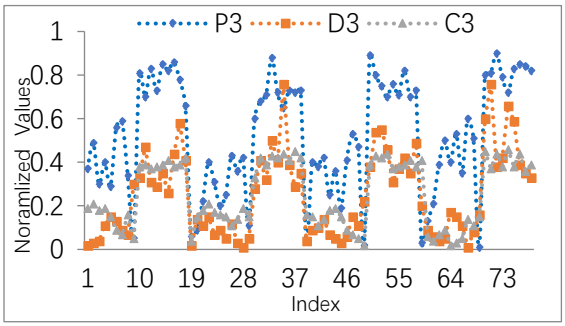


Figure 5. normalized values Under T3 Scenarios.

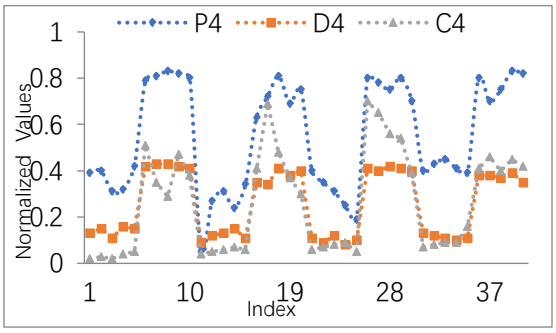


Figure 6. normalized values Under T4 Scenarios.

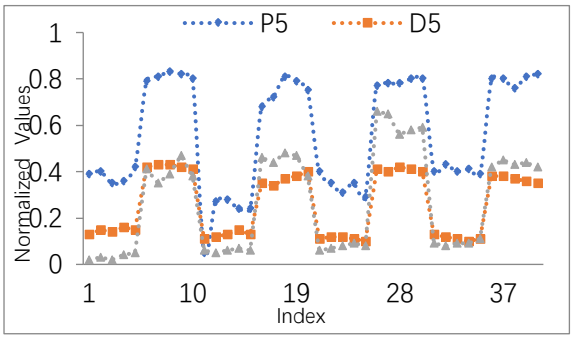


Figure 7. Detection System Work Process.

From the data in figure 6 and figure 7, it can be seen that the values of the three normalized eigenvalues of the overt and covert channels in these two scenarios not only have obvious dividing lines, but also the fluctuation range of the value is stable. Therefore, these two scenarios can be used as scene of detection.

Taking into account the real-time detection, so that the client in figure 2 sends packets to the target machine, set the detection window values are 1000,2000,4000,8000,10,000,16,000. The results obtained are shown in figure 8. When the detection window value is less than 2000, the detection time

is less than 1 second, even if the detection window value is set to 4000, the required detection time is less than 10 seconds. But when the detection window exceeds 8000, the time cost increases dramatically. This is due to the fact that the autocorrelation feature in the algorithm requires more time to deal with high dimensional data. Therefore, the detection window is set to 2000 or less, and it can have good real-time performance.

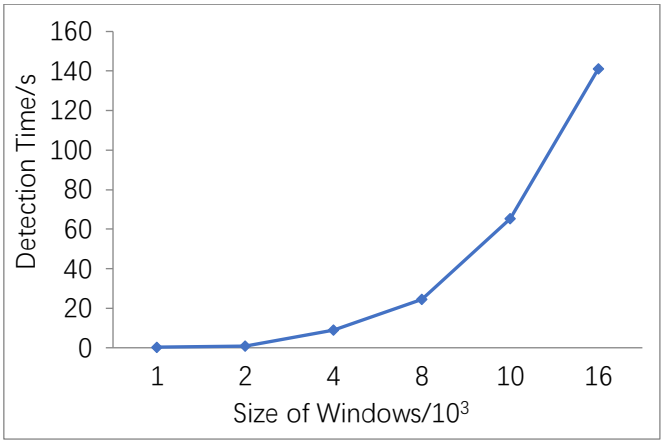


Figure 8. normalized values Under T5 Scenarios.

Based on the above analysis, we can conclude that length of IPDs, step_size and value of m are positively correlated. However, the values of three parameters are higher, the later it finds the time to covert channel. It may cause some information to be leaked before covert channel is detected. This paper defines the environmental parameters of overt channel M0: N = 2000, Step_Size = 200, m = 20, the characteristic threshold is Dis = 0.6, Polar = 0.3, Cr = 0.2.

*5.3. Detection Rate and Time Performance Assessment*

In order to test the effectiveness of the proposed method, the detection rate of the covert timing channel and the false positive rate of the normal channel need to be tested. We collected a network data stream containing 200,000 IPD for experimentation, which included four different time covert timing channels and normal channels. Detection rate and false positive rate are shown in table 2.

Table 2. Detection Rate.

| Channel | Proposed Method | | CCE | | K-S Score | |
|---|---|---|---|---|---|---|
| | TP(%) | FP(%) | TP(%) | FP(%) | TP(%) | FP(%) |
| IPCTC | 100 | 5 | 100 | 9 | 96.3 | 7 |
| MBCTC | 97.3 | 9 | 93.9 | 12 | 91.2 | 11 |
| Jitterbug | 96.2 | 7 | 91.7 | 8 | 86.4 | 18 |
| Liquid | 94.9 | 8 | 90.5 | 12 | 80.8 | 12 |

The proposed method can detect several typical covert timing channels proposed in this paper, which has high versatility. Compared with the traditional detection method, the proposed method can detect several typical time covert channels proposed in this paper, which has a higher detection rate and a relatively low false positive rate. Due to the jitter or delay in the process of network communication, causing IPDs is unstable. The polarization coefficient is sensitive to the network delay and will cause the polarization coefficient to be higher than the set threshold for the overt channel's polarization, which is mistaken covert channel, so there will be a relatively high false positive rate.

To assess the overhead time of the proposed method, test performed 1000 times under parameters N = 2000, Step_size = 200, m = 20 scenes, and an average time overhead is 0.98s. Such an order of magnitude of time overhead in the real-time monitoring system is acceptable.

## 6. CONCLUSION

This paper analyses several typical network covert timing channel construction algorithms. On this basis, a multi - dimensional feature detection method is proposed. Since the attacker modulates hidden information into IPD, it must change its IPDs shape and data statistics, causing changes in the polarization characteristics, clustering characteristics, statistical characteristics of the changes. Therefore, this method can detect network covert timing channel, with a certain degree of blind detection function, and the detection rate better. However, in the actual detection process, the threshold value of the eigenvalue will change with the change of different scenes, which makes it difficult to select the appropriate threshold for online detection, using the method of machine learning to make the detector can adaptively select the threshold is the next step in this research direction.

**Acknowledgments**

This work was supported by the National Natural Science Foundation of China (Grants No. 61802432).

**References**

[1] Archibald R, Ghosal D. (2014) A comparative analysis of detection metrics for covert timing channels. Computers & Security, 45(8):284-292.
[2] Zseby T, Vázquez F I, Bernhardt V, et al. (2016) A Network Steganography Lab on Detecting TCP/IP Covert Channels. IEEE Transactions on Education, 59(3):224-232.
[3] Yao F, Venkataramani G, Doroslovački M. (2017) Covert timing channels exploiting non-uniform memory access based architectures. In: Proceedings of the on Great Lakes Symposium on VLSI 2017. Canada, pp. 155-160.
[4] Hélouët L, Jard C, Zeitoun M. (2003) Covert channels detection in protocols using scenarios. In: Proceedings of Spv' Workshop on Security Protocols Verification, pp. 1-9.
[5] Rezaei F, Hempel M, Shrestha P L, et al. (2015) Detecting covert timing channels using non-parametric statistical approaches. In: IEEE 2015 Wireless Communications and Mobile Computing Conference. Dubrovnik, pp.102-107.
[6] Zhang L, Liu G, Dai Y. (2014) Network Packet Length Covert Channel Based on Empirical Distribution Function. Journal of Networks, 9(6).
[7] Cabuk S, Brodley C E, Shields C. (2004) IP covert timing channels:design and detection. In: 2004 ACM Conference on Computer and Communications Security, Washington, Dc, pp. 178-187.
[8] Berk V, Giani A, Cybenko G. (2009) Detection of covert channel encoding in network packet delays. Rapport Technique Tr.
[9] Pang P, Zhao H, Bao Z. (2015) A probability-model-based approach to detect covert timing channel. In: IEEE International Conference on Information and Automation. Lijiang, pp.1043-1047.
[10] Shrestha P L, Hempel M, Rezaei F, et al. (2016) A Support Vector Machine-Based Framework for Detection of Covert Timing Channels. IEEE Transactions on Dependable & Secure Computing, 13(2):274-283.
[11] Rezaei F, Hempel M, Sharif H. (2017) Towards a reliable detection of covert timing channels over real-time network traffic. IEEE Transactions on Dependable and Secure Computing, 14(3): 249-264.
[12] Brodley C E, Spafford E H, Cabuk S. (2006) Network covert channels: Design, analysis, detection, and elimination. Dissertations & Theses - Gradworks.
[13] Shah G, Molina A, Blaze M. (2009) Keyboards and covert channels. In: Proceedings of the 18th conference on USENIX security symposium. Berkeley, pp. 5.
[14] Tan Y, Zhang X, Sharif K, et al. (2018) Covert timing channels for IoT over mobile networks. IEEE Wireless Communications, 25(6): 38-44.