A Technique to Limit Packet Length Covert Channels

Anna Epishkina and Konstantin Kogos
Department of Cybernetics and Information Security, National Research Nuclear University MEPhI (Moscow Engineering Physics Institute), Moscow, Russian Federation

Keywords: Information Security, Covert Channels, Binary Channel, Multi-Symbol Channel, Packet Size, Capacity.

Abstract: We designed the technique to estimate and limit the capacity of the covert channel based on traffic padding and random increase of packets lengths. It was applied to two types of packet size covert channels namely binary and multi-symbol channels. The method to choose the parameter of counteraction tool is given, it takes into account an allowable value of covert channel capacity and error level. The investigation carried out is significant because such type of covert channels could be constructed even if traffic encryption is used. The novelty of the investigation undertaken is that the covert channel capacity is limited preliminary, whereas state of the art methods focus on detecting active IP covert channels.

1 INTRODUCTION

1.1 Timing and Storage Covert Channels

The covert channel is a communication channel that was not intended for information transfer at all, such as the service program's effect on the system load (Lampson, 1973). TCSEC postulates that the covert channel is a communication channel which allows the transfer of data and violation of security policy (Department of defence trusted computer system evaluation criteria, 1985).

Presently, the most popular covert channels are built on packet switching data networks because of some features available in the TCP/IP protocol suite (Zander, 2007). Moreover, traditional security measures based on traffic encryption also permit the design of different types of covert channels.

Covert channels are divided by the data transfer technique into two classes such as, timing and storage channels. Storage channel allows the direct or indirect storage recording by one process and the direct or indirect reading of it by another. Timing channel allows one process to signal information to another process by modulating of system resources (e.g. CPU time) usage so that the change in response time observed by the second process would provide information.

The first technique to design a storage channel in the IP network is to modulate packet header fields, e.g. TTL (Zander, 2006), IP ID (Ahsan, 2002), ToS (Handel, 1996). The second technique is based on the modification of the packet length. Different timing channels in the IP networks use alteration of the inter-packet delays (Berk, 2005 and Sellke, 2009), e.g. by JitterBug (Shah, 2009) and packet transfer rate (Yao, 2009). In addition, the packet reordering could be used to build a timing channel (Kundur, 2003). Timing channel is a channel with noise since a packet timing is a random variable whose distribution depends on the network load (Bovy, 2002).

It is very important to examine the methods of data leakage using covert channels since it was proved the ability to construct absolutely invisible covert channels by a violator who knows system security tools (Grusho, 1999). As a capacity of undetectable packet length covert channels can be higher than a capacity of timing channels, these channels can lead to a more serious security threats. Moreover, a violator can create such channels even if traffic encryption is utilized and storage channels of another type can not be build in such case. We present the investigation of this type of covert channels and propose the technique to estimate and limit their capacity.

1.2 Our Contribution

The way to eliminate such covert channels is to send packets with the fixed length, but in this case, the residual capacity of communication channel reduces

Epishkina, A. and Kogos, K..
A Technique to Limit Packet Length Covert Channels.
In *Proceedings of the 7th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (IC3K 2015) - Volume 3: KMIS*, pages 144-151
ISBN: 978-989-758-158-8

unacceptable. Therefore, it is quite important to investigate covert channel limitation methods. Since a technique to choose the quantitative characteristics of countermeasures in order to keep balance between capacity of covert and communication channels has not yet been proposed, we offer an approach to gain it. This paper describes the technique to estimate and limit the capacity of binary and multi-symbol packet size covert channels based on traffic padding. Since these covert channels can be constructed even if data encryption is used and there are complicated undetectable covert channels which have no noise in contrast to timing channels, our investigation is of current importance.

This paper is organized as follows. It gives an analysis of different types of packet size covert channels. The investigated covert channel schemes and the counteracting techniques are shown. Then the capacity of the covert channel is estimated and the technique to generate dummy packets and increse packets lengths is given. The main results are summed up in the conclusion.

2 RELATED WORK

Data link layer frames lengths modulation in order to hide data transfer was examined and the scheme where sender and receiver share the rule used to compose a byte of the covert message depending on the frame length was given (Padlipsky, 1978 and Girling, 1987).

Covert channel in which a sender and a receiver shared the periodically updated matrix with elements representing unique unsorted packets lengths was investigated (Yao, 2008). The sender using the bits of hidden transmitted message determines the matrix row and randomly chooses a packet length from it. The receiver finds the gained packet length in the matrix and recovers bits of the message according to the row number. This type of covert channel is detectable since packet length distribution in the case where covert channel is not equal to the same distribution of normal traffic (without covert channel).

A protocol-independent covert channel with the following properties was built (Ji, 2009b). Before the transmission starts, sender and receiver form the dynamically updated reference of packets lengths by fixing packets lengths in normal traffic. In order to transfer a hidden message the sender transmits a special packet. The length of the special packet is chosen from the reference using the algorithm known to the sender and the receiver. The length of the next packet is a sum of length of the previous packet and the number corresponding to the message bits. The reference is updated by adding the length of the transmitted packet. The receiver recovers the message bits by evaluating the difference of the packets lengths gained. The disadvantage of the covert channel is as follows: the lengths of hidden messages are added to the reference, therefore the packet length distribution with the covert channel is not equal to packet length distribution of the normal traffic and this type of covert channels is detectable in case of the large volume of data transferred via covert channel.

There is another protocol-independent covert channel (Ji, 2009a). At the beginning of the data transmission sender and receiver fix packets lengths in the normal traffic and form a non-updating reference of packets lengths. This technique is practically useful due to a small space and time complexity of the decoding, as the sender stores the whole reference and receiver holds only the lower and upper bounds of each basket. To transmit the hidden data the sender randomly chooses the packet length from the basket, the receiver determines the number of the basket and recovers the message bits. The regularity in the distribution of the transmitted message bits could cause a highly probable detection of the channel.

This technique was improved and high capacity covert channel based on the alteration of packets lengths and information content was designed (Hussain, 2011). Sender and receiver share periodically updated matrix with elements representing unsorted packets lengths in normal traffic. The sender determines the matrix row using bits of hidden message and randomly chooses the packet length from it. If the chosen length belongs to the stego-column, then the data is transferred in the information content of the packet, otherwise the data is transferred in the number of matrix row. The receiver finds the length of the message in the matrix, detects the transfer method using the matrix row and recovers bits of the message. The disadvantage of the channel is that information content of the packet has to be used as the hidden container and it is more complicated in comparison with the other techniques.

The method was improved and realized using TCP (Edekar, 2013). Packets lengths in the shared matrix are unique and each matrix element a is associated with the binary vector (v, y), where $v = 1 \Leftrightarrow a$ belongs to the stego-column and $y = 1 \Leftrightarrow a$ belongs to the stego-row. Then if $(v, y) = (1,1)$ the packet is ignored; if $(v, y) = (1,0)$ data is transferred in the packet information content;

if $(v, y) = (0,0)$ data is transmitted in the number of the matrix row; if $(v, y) = (0,1)$ data is transferred in the number of the matrix row and the packet information content.

The way to eliminate covert channels based on length of transferred packets modulation is to equalize packets lengths and send packets with maximum length. However, the technique essentially diminishes the capacity of a communication channel. To limit a covert channel capacity, the random increase of packets lengths and generation of dummy packets can be used. In order to make traffic nontraceable this method was realized using IPsec (Kiraly, 2008).

Also it was suggested to use information theory to estimate a capacity of covert channels with noise (Millen 1987 and Ventakraman 1995). The authors determine network covert channels and analyse techniques to audit and limit a capacity of covert channels utilizing indirect routing.

However, a technique to choose the quantitative characteristics of the methods in order to keep balance between capacity of covert and communication channels has not yet been proposed. Therefore we present an approach to gain it.

3 TRAFFIC PADDING TO LIMIT BINARY COVERT CHANNEL

3.1 The Counteraction Tool

The investigation proposes a technique to limit the capacity of covert channel based on traffic padding. After k data packets have been sent, random length dummy packets are created, where k is the parameter of a counteraction tool. Let μ be the capacity of a communication channel, then a counteraction tool decreases the capacity of a communication channel and it equals $k\mu/(k+1)$.

3.2 Binary Covert Channel Design

Let the lengths of transferred packets possess the values from $l_{fix}+1$ to $l_{fix}+L$. For example, l_{fix} can be equal to 20 bytes and corresponds to the length of IP header. The disjoint sets L_0 and L_1 are given and

$$\begin{cases} L_0 \bigcup L_1 = N_{L+l_{fix}} \setminus N_{l_{fix}}, \\ |L_0| = |L_1| \end{cases} \quad (1)$$

where N_a stands for the set of positive integers from 1 to a.

Further, we consider a method to build a binary covert channel. In order to transfer "0" the sender communicates a packet with length $l \in L_0$, to transfer "1" the sender communicates a packet with length $l \in L_1$. It is obvious that the capacity of such a channel without counteraction is equal to one bit per packet. A large-scale site loses about 26 Gb of data annually if there is a covert channel with such a capacity (Fisk, 2002).

If the symbol distribution in a transmitted message simulates a uniformly distributed random sequence (e.g. cryptographic keys sending), a random equiprobable choice of packets length from L_0 and L_1 leads to equally probable random distributed lengths of transmitted packets. Moreover, L_0 and L_1 could be periodically changed multisets so that a random choice of packets length from L_0 and L_1 induces the distribution of packets lengths to be close to the empirically obtained distribution of normal traffic.

To build such a covert channel, the sender must have one of the following possibilities: to modify lengths of transmitted messages, to form packets with undefined length or to buffer packets to be sent and transfer them to a channel at a specified moment.

After a dummy packet is received, the mismatch between the hidden sender and hidden receiver takes place. To negotiate this fact SOF packets (Cabuk, 2004) are utilized after transferring $T-1$ packets within a covert channel. A receiver fixes $T-1$ packets gained after SOF packet and waits for the next SOF packet. Thus, T is the parameter of a covert channel which estimates the synchronization frequency.

3.3 The Capacity of Binary Covert Channel

The capacity C of the investigated covert channel is

$$C = \max_{T} I(X, Y) \quad (2)$$

where $I(X, Y)$ is the mutual information of random variables describing the input and output properties of the covert channel properly, the dimensionality of covert channel capacity is one bit per packet.

Let us consider the case when synchronization is done more rarely than a dummy packet sending, i.e. $T > k$. After a dummy packet is received, mismatch between sender and receiver occurs, therefore

identification of the following received bits would be wrong until the next synchronization happens. Consequently, in order to build a covert channel the inequality $T < k+1$ is required.

Let the synchronization be not less frequent than dummy packet sending, i.e. $T < k+1$. Corresponding choice of parameters is explained in Figure 1 $(T=3, k=5)$.

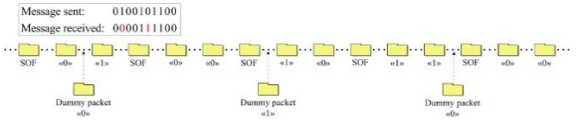

Figure 1: The scheme of data transfer in binary covert channel.

Since each T-th packet sent via a covert channel is not a data packet but a synchronization packet, the mutual information can be calculated using the following formula

$$I(X,Y) = \frac{T-1}{T}\, I^*(X,Y) \tag{3}$$

where $I^*(X,Y)$ is a mutual information of random variables describing the input and output properties of a covert channel without synchronization accordingly.

The mutual information $I^*(X,Y)$ is equal to the form of

$$I^*(X,Y) = H(Y) - H(Y|X) \tag{4}$$

where corresponding entropies are

$$H(Y) = -\sum_{y\in\{0,1\}} p(y)\log_2 p(y), \tag{5}$$

$$H(Y|X) = -\sum_{x\in\{0,1\}}\left(p(x)\left(\sum_{y\in\{0,1\}}(p(y|x)\log_2 p(y|x))\right)\right). \tag{6}$$

Since cardinalities of sets L_0 and L_1 are equal and length of passing through the covert channel dummy packets is chosen randomly and equiprobable, then $H(Y)=1$.

Whereas the values of conditional probabilities $p(y|x), x, y \in \{0,1\}$ depend on the number of packets sent via a covert channel between the moment of synchronization and the moment of dummy packet receiving, the mutual information $I^*(X,Y)$ can be found using the following formula

$$I^*(X,Y) = \frac{k-(T-1)+\sum_{i=0}^{T-2}\left(1-H_i(Y|X)\right)}{k} \tag{7}$$

where $H_i(Y|X)$ is the conditional entropy of random variable Y compared to random variable X and it is evaluated when i packets received between a moment of synchronization and the dummy packet's arrival.

The mutual information $I(X,Y)$ could be estimated as

$$\begin{aligned} I(X,Y) = {} & \frac{T-1}{T} - \frac{(T-1)^2}{kT} - \\ & - \frac{(T-1)^2\log_2(T-1)}{kT} + \\ & + \frac{1}{2kT}\left(\sum_{i=0}^{T-2} f^+(i) + \sum_{i=0}^{T-2} f^-(i)\right) \end{aligned} \tag{8}$$

where

$$\begin{aligned} f^+(i) &= (T+i-1)\log_2(T+i-1), \\ f^-(i) &= (T-i-1)\log_2(T-i-1). \end{aligned} \tag{9}$$

In order to analyze functions $f^+(i)$ and $f^-(i)$ we will examine analogue variable $\tilde{i}, \tilde{i} \in [0, T-2]$ instead of discrete variable i, in which case $f^+(\tilde{i})$ is a strictly increasing and $f^-(i)$ is a strictly decreasing defined and continuous functions in the interval $[0, T-2]$. Then the values of the following forms

$$\sum_{i=0}^{T-2} f^+(i), \sum_{i=0}^{T-2} f^-(i) \tag{10}$$

could be approximated by means of functions $f^+(\tilde{i})$ and $f^-(i)$ integrating in the interval $[0, T-2]$ accordingly

$$\begin{aligned} \sum_{i=0}^{T-2} f^+(i) \approx {} & \int_0^{T-2} f^+(\tilde{i})d\tilde{i} + f^+(T-2) - \\ & - \sum_{j=0}^{T-3} \frac{f^+(j+1) - f^+(j)}{2} = \\ = {} & \int_0^{T-2} f^+(\tilde{i})d\tilde{i} + \frac{f^+(0) + f^+(T-2)}{2}, \end{aligned} \tag{11}$$

$$\sum_{i=0}^{T-2} f^{-}(i) \approx \int_{0}^{T-2} f^{-}(\tilde{i})d\tilde{i} + \\ + \sum_{j=0}^{T-3} \frac{f^{-}(j) - f^{-}(j+1)}{2} = \\ = \int_{0}^{T-2} f^{-}(\tilde{i})d\tilde{i} + \frac{f^{-}(0)}{2}. \quad (12)$$

Values of the integrals

$$\int_{0}^{T-2} f^{+}(\tilde{i})d\tilde{i}, \int_{0}^{T-2} f^{-}(\tilde{i})d\tilde{i} \quad (13)$$

could be found using the variable substitution and integration by parts. It follows that

$$\sum_{i=0}^{T-2} f^{+}(i) + \sum_{i=0}^{T-2} f^{-}(i) \approx \\ \approx -\frac{(T-2)(T-1)}{\ln 2} + (T-1)\log_2(T-1) + \\ + (2T-3)(T-1)\log_2(2T-3). \quad (14)$$

Then if k is a continuous variable $k \in [T;+\infty]$, then $I(X,Y)$ as a function from k, is defined and continuous at the interval $[T;+\infty]$ and is the hyperbola

$$I(X,Y) \approx A(T) + \frac{B(T)}{k}, \quad (15)$$

$$A(T) = \frac{T-1}{T}, \quad (16)$$

$$B(T) = \frac{(T-1)^2}{T} - \frac{(T-2)(T-1)}{2T\ln 2} + \\ + \frac{(2T-3)(T-1)}{2T}\log_2\frac{2T-3}{T-1}. \quad (17)$$

Functions $A(T)$ and $B(T)$ are positive strictly increasing and negative strictly decreasing functions from T accordingly. Graphs of function $I(X,Y)$ from k where $T = 2,3,4$ are illustrated in Figure 2.

To build a covert channel the parameter T is chosen while $I(X,Y)$ has a maximum value. Figure 2 shows that when $k \in \{2,3,4\}$, parameter T should be equal to 2 and when $k \in \{5,6,7,8\}$, parameter T should be equal to 3.

The criteria (Department of defence trusted computer system evaluation criteria, 1985) postulate that the functioning of a covert channel with capacity less than 100 bits per sec can be acceptable in some cases. Let the communication capacity be equal to 100 Mbit per sec (e.g. 100Base-T). The maximum

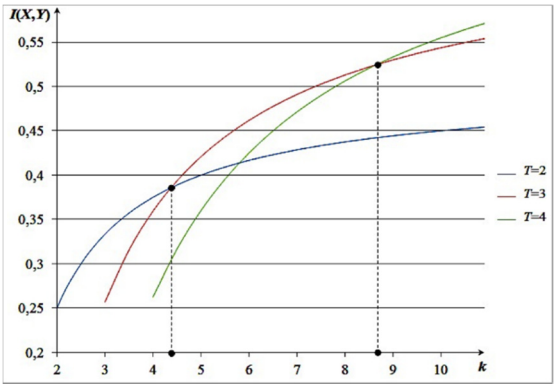

Figure 2: $I(X,Y)$ function as a function from k graph, $T = 2,3,4$.

length of IP packet is 65535 bytes, from which 20 bytes is a header length. Let $l_{fix} = 20$ bytes and $L = 65515$ bytes. In this case, to limit the capacity of covert channel, parameter k should be equal to 3.

4 PACKET LENGTH RANDOM INCREASE TO LIMIT MULTI-SYMBOL COVERT CHANNEL

4.1 The Counteraction Tool

We propose to increase transmitted packets lengths randomly, i.e. we add some padding bits to the length of each packet before sending. The number of added bits is randomly chosen from the interval [0,*a*]. Chosen uniform distribution brings to the hidden message receiver the largest entropy while he recovers the message. The left cell boundary of [0,*a*] is equal to 0 because in this case there is minimal extra load of communication channel. Let μ be the capacity of a communication channel, then a counteraction tool decreases the capacity of a communication channel and it equals

$$\frac{2l_{fix} + L + 1}{2l_{fix} + L + 1 + a}\mu \quad (18)$$

4.2 Multi-Symbol Covert Channel Design

Let the lengths of transmitted packets be from $l_{fix} + 1$ to $l_{fix} + L$ bits, $n|L$ and sets $W_0, ..., W_{\frac{L}{n}-1}$ are given,

$$W_i = N_{l_{fix}+(i+1)n} \setminus N_{l_{fix}+in} \quad (19)$$

where $i \in N_{\frac{L}{n}-1} \bigcup \{0\}$.

Then $\frac{L}{n}$-symbol covert channel can be constructed as follows. In order to transmit symbol «*i*» the sender forwards a packet with length *l* from the set W_i, $i \in N_{\frac{L}{n}-1} \bigcup \{0\}$. If there are no countermeasures, the capacity of such channel is $\log_2 \frac{L}{n}$ bits per packet.

If the sequence of symbols in a transmitted message is a sequence of uniformly distributed random numbers (e.g. cryptographic keys are transferred via covert channel), then random equiprobable choice of packets lengths from equinumerous sets $W_0,...,W_{\frac{L}{n}-1}$ leads to a uniformly distributed random sequence of packets lengths in covert channel in the set $N_{l_{fix}+L} \setminus N_{l_{fix}}$. It is quite important to analyse such covert channels since a large Internet site looses 4 Gb of data daily if there is covert channel with the capacity equal to 8 bits per packet (Fisk, 2002).

4.3 The Capacity of Multi-Symbol Covert Channel

Let us suppose that the packet length is *l* and it is increased in α bits before packet sending, $\alpha \in N_a \bigcup \{0\}, l + \alpha > l_{fix} + L$, therefore the packet with the length equal to $\tilde{l} = (l + \alpha) \bmod L$ will be sent.

As the sets $W_0,...,W_{\frac{L}{n}-1}$ are equinumerous, then $H(Y) = \log_2 \frac{L}{n}$. The values of conditional probabilities of symbol "*y*" recovering when symbol "*x*" is sent $p(y|x), x, y \in N_{\frac{L}{n}-1} \bigcup \{0\}$ depend on the value of the integer part plus one of the following fraction $\left\lceil \frac{a}{n} \right\rceil$, so if $nx < a \le n(x+1), x \ge 1$, then $\forall i \in N_{\frac{L}{n}-1} \bigcup \{0\}$:

$$p(i|i) = \frac{n+1}{2(a+1)}; \quad (20)$$

$$p\left((i+j) \bmod \frac{L}{n} \middle| i\right) = \frac{n}{a+1}, j \in N_{x-1}; \quad (21)$$

$$p\left((i+x) \bmod \frac{L}{n} \middle| i\right) = \\ = \frac{2n(a-xn+1)+(a-xn+n)(xn+n-a-1)}{2n(a+1)}; \quad (22)$$

$$p\left((i+x+1) \bmod \frac{L}{n} \middle| i\right) = \\ = \frac{(a-xn)+(a+1-xn)}{2n(a+1)}; \quad (23)$$

$$p(j|i) = 0, \\ j \in N_{(L \setminus n)-1} \bigcup \{0\} \setminus \{(i+k) \bmod \frac{L}{n}, \\ k \in N_{x+1} \bigcup \{0\}\}. \quad (24)$$

It is easy to check that the sum of conditional probabilities is equal to one.

In order to construct the covert channel the parameter *n* should have the value when the mutual information $I(X,Y)$ is maximum. The value of the mutual information $I(X,Y)$ decreases when the parameter *n* increases. Hence, maximum capacity of covert channel is

$$C = \log_2 \frac{L}{a+1}. \quad (25)$$

On the other hand, the error level in the covert channel enlarges when *n* decreases. Hence, the value of *n* should be chosen the least, when the error level in the covert channel is allowable.

Let the allowable error level in the covert channel be $p_{error} \le 0.5$, then $n = a / 2p_{error}$ and $\forall i \in N_{\frac{L}{n}-1} \bigcup \{0\}$:

$$\begin{cases} p(i|i) = 1 - p_{error}; \\ p\left((i+1) \bmod \frac{L}{n} \middle| i\right) = p_{error}; \\ p(i|j) = 0, \\ j \in N_{\frac{L}{n}-1} \bigcup \{0\} \setminus \left\{i, (i+1) \bmod \frac{L}{n}\right\}. \end{cases} \quad (26)$$

The covert channel capacity is

$$C = \log_2 \frac{2Lp_{error}}{a} - H(\zeta), \quad (27)$$

where

$$H(\zeta) = -(p_{error} \log_2 p_{error} + \\ +(1-p_{error}) \log_2 (1-p_{error})) \quad (28)$$

is the entropy of a random variable with Bernoulli distribution (the success probability equals p_{error}).

Let the value of covert channel capacity $v_{\max}$ be known, then its dimensionality is one bit per packet, such as functioning of a covert channel with capacity less than $v_{\max}$ has no influence upon security. Then if the allowable error level in the covert channel $p_{error} \le 0.5$ the parameter of counteraction tool a should be

$$a = \left\lceil \frac{Lp_{error}}{2^{v_{\max}+H(\zeta)-1}} \right\rceil. \quad (29)$$

If there are no restrictions on the error level the parameter of counteraction tool a should be

$$a = \left\lceil \frac{L}{2^{v_{\max}}} - 1 \right\rceil. \quad (30)$$

Therefore, we propose quantitative characteristics of a technique to limit a covert channel capacity that allows decreasing it to a given critical value. This is practically useful since e.g. the functioning of a covert channel with capacity less than 100 bits per sec can be acceptable in some cases.

5 CONCLUSIONS

In this work, the capacity of network covert channels was estimated using the information theory statements. The counteraction tool utilizes dummy packets generation and random increase of packets lengths. The authors suggested a technique to select the parameter of the counteraction tool when an allowable covert channel capacity is given. The novelty of the method is that the capacity of covert channel is limited in advance in contrast to the other approaches, which detect the active covert channel.

REFERENCES

Ahsan, K., Kundur, D., 2002. Practical data hiding in TCP/IP. In *Proc. of the 2002 ACM Multimedia and security workshop*.

Berk, V., Giani, A., Cybenko, G., 2005. *Detection of covert channel encoding in network packet delays: Technical report TR2005-536*. New Hampshire: Thayer school of engineering of Dartmouth College.

Bovy, C.J., Mertodimedjo, H.T., Hooghiemstra, G., Uijterwaal, H., Mieghem, Van P., 2002. Analysis of end-to-end delay measurements in Internet. In *Proc. of ACM Conference Passive and Active Measurements*.

Cabuk, S., Brodley, C.E., Shields, C. 2004. IP covert timing channels: design and detection. In *Proc. of the 11th ACM conference on computer and communications security*, pp. 178–187.

Department of defence trusted computer system evaluation criteria, 1985. Department of defence standard.

Edekar, S., Goudar, R., 2013. Capacity boost with data security in network protocol covert channel. In *Computer engineering and intelligent systems*, Vol. 4, No. 5, pp. 55–59.

Fisk, G., Fisk, M., Papadopoulos, C., Neil, J., 2002. Eliminating steganography in Internet traffic with active wardens. In *Proc. of the fifth International workshop on information hiding*, pp. 18–35.

Girling, C.G., 1987. Covert channels in LAN's. In *IEEE Transactions on software engineering*, Vol. 13, No. 2, pp. 292–296.

Grusho, A.A., 1999. On the existence of hidden channels. In *Discrete mathematics and applications*, Vol. 11, No. 1, pp. 24-28.

Handel, T., Sandford, M., 1996. Hiding data in the OSI network model. In: *Proc. of the first International workshop on information hiding*, pp. 23–38.

Hussain, Mehdi, Hussain, M., 2011. A high bandwidth covert channel in network protocol. In *Proc. of the 2011 International conference on information and communication technologies*, pp. 1–6.

Ji, L., Liang, H., Song, Y., Niu, X., 2009a. A normal-traffic network covert channel. In *Proc. of the 2009 International conference on computational intelligence and security*, pp. 499–503.

Ji, L., Jiang, W., Dai, B., Niu, X., 2009b. A novel covert channel based on length of messages. In *Proc. of the 2009 Symposium on information engineering and electronic commerce*, pp. 551–554.

Kiraly, C., Teofili, S., Bianchi, G., Cigno, R. Lo, Nardelli, M., Delzeri, E., 2008. Traffic flow confidentiality in IPsec: protocol and implementation. In *The International federation for information processing*, Vol. 262, pp. 311–324.

Kundur, D., Ahsan, K., 2003. Practical Internet steganography: data hiding in IP. In *Proc. of the 2003 Texas workshop on security of information systems*.

Lampson, B.W., 1973. A Note on the Confinement Problem. In *Communications of the ACM*, pp. 613–615.

Millen, J.K., 1987. Covert channel capacity In *Proc. of the IEEE Symposium on Security and Privacy*, pp. 60–66.

Padlipsky, M.A., Snow, D.W., Karger, P.A., 1978. *Limitations of end-to-end encryption in secure computer networks: Technical report ESD-TR-78-158*. Massachusetts: The MITRE Corporation.

Sellke, S.H., Wang, C.-C., Bagchi S., Shroff N.B., 2009. Covert TCP/IP timing channels: theory to implementation. In *Proc. of the 28th Conference on computer communications*, pp. 2204–2212.

Shah, G., Molina, A., Blaze, M., 2009. Keyboards and covert channels. In *Proc. of the 15th USENIX Security symposium*, pp. 59–75.

Venkatraman, B.R., Newman-Wolfe, R.E., 1995. Capacity estimation and auditability of network covert channels. In *Proc. of the IEEE Symposium on Security and Privacy*, pp. 186–198.

Yao, Q., Zhang, P., 2008. Coverting channel based on packet length. In *Computer engineering*, Vol. 34, No. 3, pp. 183–185.

Yao, L., Zi, X., Pan, L., Li, J., 2009. A study of on/off timing channel based on packet delay distribution. In *Computers and security*, Vol. 28, No. 8, pp. 785–794.

Zander, S., Armitage, G., Branch, P., 2006. Covert channels in the IP time to live field. In *Proc. of the 2006 Australian telecommunication networks and applications conference*, pp. 298–302

Zander, S., Armitage, G., Branch, P,. 2007. A survey of covert channels and countermeasures in computer network protocols. In *IEEE Communications surveys and tutorials*, Vol. 9, No. 3, pp. 44–57.