A High Bandwidth Covert Channel in Network Protocol

Mehdi Hussain and M. Hussain
Shaheed Zulfikar Ali Bhutto Institute of Science and Technology (SZABIST), Islamabad, Pakistan.
mehdi141@hotmail.com, mhussain@szabist-isb.edu.pk

Abstract

Covert channel is secret communication path, which existence is not expected in original system of design. It allows different users access to the same information, at the same time, but from different points-of-view. It is being more and more studying due to the grooming of internet. Always high-bandwidth network covert channel pose the significant risk of detection over network. Although the existing technique utilized the reserved bits of header, timestamp, initial sequence number, packet length etc for network covert channel, to some extend these have good temper resistance. But when increase their covert data capacity they becomes to fail to persist their original characteristics and vulnerable to network traffic detector. We are motivated to design to achieve a high bandwidth covert channel in network protocols. We simulated our proposed technique in ns2 simulator utilizing TCP protocol. Through experimental results, our proposed model is very high capacity covert channel with temper resistance and time efficient as compared to previous techniques.

Keywords: *covert channel; high bandwidth covert channel; message length covert; Payload covert;*

1. Introduction

Generally, encryption is used to secure communication. However, encryption only prevents to decode the communication from the unauthorized user. Encryption itself makes the communication as suspicious. On the other side, Covert Channels are secret communication paths, where its existence is not in the original design of the system. A covert channel is generally known as a communication channel that is neither designed nor intended to transfer hidden information [1]. This aim of covert channel was first introduced by Lampson in 1973. Girling extended it to network in 1987 [2]. The National Institute of Standards and Technology defines a covert channel as "any communication channel that can be exploited by a process to transfer information in a manner that violates the system's security policy" [3]. Due to the tremendous growth of internet different protocols utilize the covert channel as a vehicle for covert communication.

Alternatively, covert network channels can be used as subversive means of achieving confidentiality and maintaining anonymity. Covert channel can be used to deliver secret commands to system. Suppose a country can digitally watermark a network datagram as it leaves the network of one enemy. They can observe datagram entering other organizations network and thus determine what type of communication links they have? [4]. Covert channel is used for authentication, delay measurements and signaling information [17] [18] [19]. Traditionally covert channel is classified into two main categories. First Storage channels, direct/indirect embedding/encoding data by the sender and direct/indirect extracting/decoding data by the receiver. Second, Timing channels, when sender signaling the information by

modulating the use of resources over time, so the receiver observed it and decode/extract the information [5].

Almost past three decades, Covert channels have been well studied within software systems or on a single machine or even over network protocols. Packet length based is one of the temper resistance covert channel communication. Girling [2] and Padlipsky [24] used the link layer frames length directly for covert communication. Yao [25] also proposed another covert channel based on the packet length known as LAWB. Ji [26] has introduced another method by taking the normal communication packet lengths as a reference. However, all the above packet length techniques are vulnerable to detections due to abnormal network traffic, because newly generated packet sizes are far from the real network communication packet size distribution. In Ji [27] introduced another normal traffic network covert channel known as NTNCC for covert communication technique. Hence, we proposed an efficient network covert channel that could enhance the tamper resistance, will generate normal network traffic plus with large amount of covert information as compare to the above techniques.

In this paper, our focus is to utilize the network protocol packet length and its payload to achieve a high rate of covert channels and also maintaining the normal network traffic behavior. Our proposed scheme is flexible that can easily be used with all network protocols. For experimental results, we achieved covert channel in NS2 with (transmission control protocol) TCP tahoe protocol. The rest of this paper is organized as follows. In section 2 has communication channel model, literature review general and packet length based covert channels techniques and other newly proposed mechanism for network protocol based covert communication. In section 3 we describe our proposed covert channel based on packet length. In section 4 we discuss our experimental results, different scenarios. Finally, in section 5 we summarize our methodology and outline future work.

2. Related Works

In this section we describe the covert channel communication model, related work done so far, and describe our proposed model for covert communication and its utilization in covert channel communication model.

2.1 Communication Channel Model

In Simmons [28] introduced the prisoner problem that is de-facto standard of covert channel communication model. Two people Alice and Bob are prison and wanted to escape from jail. To agree on an escape plan they need to communicate but all their messages are monitored by Wendy the warden. If Wendy finds any signs of suspicious messages Wendy will place Alice and Bob into solitary confinement (making an escape impossible). Alice and Bob must exchange seamless communication containing hidden information, so they hope that Wendy should not be noticed.

For practicality of scenario is explored by communication networks. Alice and Bob are using two networked computers to communicate. They run some innocuous looking overt communication between their computers with a hidden covert channel. For the time being Alice and Bob may well be the same person, for example hacker ex-filtrating restricted information. Wendy can monitor the passing traffic for covert channels or alter the passing traffic to eliminate or disrupt covert channels. Figure 1 show the communication model (Alice sending to Bob).

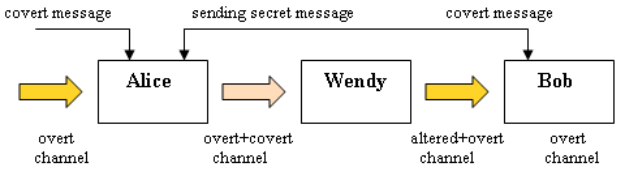

Figure 1: Communication Model

2.2 General Techniques

Generally, covert channel is encoded in the unused or reserved bits of the packet header or protocol header when the protocol does not mandate the content of these bits. Around 1989, Wolf introduced covert channels in header fields of multiple protocols (Bus, Token Ring) [6] over LAN. (Internet protocol) IP header fields, (Type of Service) TOS, (don't fragment) DF fields, checksum and (time to live) TTL fields and traffic class flow label fields in IPv6 are utilized in [7], [8], [9], [15], [16] as covert channel. At TCP header fields like, (initial sequence number) ISN, TCP 16 bits Urgent Pointer, Flags and RST flag are utilized as covert channel in [10], [11], [14]. Packet sorting, reordering and permutation based network covert channel on all type of reliable stream protocol is proposed in [9], [12], [13]. In [20], [21] authors introduced scheme to achieved covert channel in the multiple connection (sockets) and artificial intelligent retransmission of packets also utilized.

2.3 Available Packet Length Based Models

The main advantage of packet length based covert channel scheme is its temper resistance, because these schemes are not tempering the content of the message except its length. It is hard to detect if the packet length distribution is same as normal or real network packet lengths of any application.

Girling [2] and Padlipsky [24] used the link layer frames length for hidden communication. Where each byte of covert message is represent to a certain link layer frame length. So, the minimum 256 message lengths are required to represent single hidden byte. Both parties are agreed on predefine message length distributions. Hidden message is encrypted, where the receiver simple decode the covert byte from the received message length. This covert channel communication is vulnerable to network traffic detection, because predefine (byte mapping to certain length) dictionary encode or decode covert message is static, and its length are not belongs to normal network traffic distribution. This makes the easily vulnerable to statistically network traffic detection.

Yao [25] also proposed a new scheme which is based on packet length, called LAWB. In his scheme the both parities (sender and receiver) has a share a secret matrix, where each element of matrix is representing a unique length. Sender choose row ID as a covert message then further randomly select a column in that row denote as L, than send packet length of L to the receiver. After receiving that packet, receiver search the matrix to find out which rows ID the element L lies in, so that row ID is the covert message. To remove statistical detection, author introduced the periodic matrix transformation on both sides (sender and receiver) simultaneously after predefined transmission. The above procedure is repeated until the whole secret message has been transmitted to the other end. Well, this above Yao [25] technique is still vulnerable due to packet length distribution as compared to the normal or real network packet length distribution.

In Ji [26] proposed a packet length based covert channel to consider the normal network traffic, where sender and receiver capture the normal communicating packet lengths as a

record for *Reference* in covert communication. In this scheme current packet length is randomly selected from the Reference list and next packet length is generated with adding the covert message and sent to the receiver. Sender updates its *Reference* list with newly generated next packet length which has been sent to receiver. On the other side receiver deduce covert message from the received packet length from *Reference* list. So this method would decrease abnormal network traffic to some extent, because newly generated packet length vary from the normal traffic packet length and its *Reference* keeps being updated by appending either the sending lengths or increased lengths. So again the normal length distribution of the *Reference* would be destroyed. This approach is being also vulnerable to network traffic detection.

In Ji [27] introduced another packet length based covert channel technique known as Normal Traffic Network Covert Channel (NTNCC). The author takes the real time packet lengths as references and uses these references to represent covert message on both side sender and receiver. First the reference packet lengths are sorted and generated equal size of buckets, where each bucket is representing a specific length of packet range. So, sender selects required bits of covert data and converts into decimal. Sender then selects equivalent decimal bucket in the reference packet length lists. When bucket found then randomly select a packet length from that bucket and sent to receiver. On the other side receiver simple get the packet length and search into the reference buckets ranges, founded buckets number is the covert data. The strength of this technique is its utilization of normal or real time packets length as references for covert data transmission. Only sender have to maintain the reference list which is further divided into equally partitioned buckets, and receiver just have to maintain the buckets ranges, which shows the efficiency in time and also in space. To our best knowledge the weakest part its reference list staticness, because once it has been initialized with up to (number) N packets length and throughout the transmission it remain constant, not updated. If covert data has minimum variations (homogenous type of data) in content like image or voice data, so specific type of bucket selection occur and same type of packet length ranges used to send covert message, because packets size are first sorted and then equally divided/partitioned into buckets. So this type of staticness can be easily detected by the network traffic detector. The capacity of this packet length scheme is very minimum as compared to our proposed model.

3. Proposed Model

Our proposed technique is also based on packet length and payload to achieve high capacity for data embedding. To consider the normal traffic distribution, we utilize the real network packet length for covert communication. Mazurczyk [20] introduced the intelligent retransmission of protocol filled with (covert data) stego-data. In our proposed scheme also filled covert data into payload of packet to increase the covert data capacity. A detailed scheme is as follows.

In proposed model both (Alice and Bob) generate a reference M x N dimensional master matrix on both sides, where each element of matrix is filled with the real network packet lengths. Each cell is representing a unique length. M, N (integer values) is already known by Alice and Bob.

Terminologies are as follows, C is the covert data bits, Alice want to transmit, Let C = c0 +c1 + c2…. ck. k is the maximum number of bits in covert data. Further C is divided into sub-group of W-bits. Let C = Wi + Wi+1 … Wi+q-1, where Wi be the ith subgroup of the C and q is the maximum length of subgroups of C and i is the simple integer counter for subgroup. Wd is the decimal value of Wi. V is the (covert) stego-column of matrix, pre shared by Alice and Bob. T is the number of packet transmitted to the other end and pre shared by both Alice and Bob. Len is the length of packet.

Step1: Synchronization phase, where Alice and Bob filled the M x N matrix in (checker box, sequential) predefined order with the normal or real network traffic packet lengths.
Step2: Alice selects Wi, the ith subgroup of C, and converts it into decimal Wd value. Find the equivalent Wd row ID into matrix and randomly select a cell in that row. So, a packet length denoted as Len is retrieved.
Step3: If the column of selected cell is matched to V (stego) column, which indicate that sender will send the stego (covert) data of Len size in the payload of that packet.
Step4: If Step 3 fails then, Sender sends the normal data packets of Len size to the receiver.
Step5: Receiver simply find out in his matrix a cell which contain the equivalent size of the received packet length.
Step6: If the column of selected cell is matched to V (stego) column, Receiver extracts the stego (covert) data directly from the packet payload.
Step7: If Step 6 fails, then Stego data is extracted by the row ID of the selected cell.
Step8: After up to T packet transmission, both Alice and Bob reshuffle their matrix in predefine (transposition, checker box) order.
Step9: Above steps repeat until the Alice has covert data to send.

As describe above, main advantage of proposed scheme its capacity improvement. For covert data transmission both packet length and packet payload is used. Another use the normal or real network traffic packets size as references in our covert communication. To remove statistical detection by introducing the periodic (after T predefined packet transmission) matrix transformation or reshuffling simultaneously on both Alice and Bob sides. Each element of matrix has no correlation with its neighbor elements, like sorting or any other sequential ordering etc. So proposed scheme is equally efficient for homogenous (video or audio) or heterogeneous (contains maximum variations) type of covert data. Above scheme is achieving both high capacity and normal-traffic variation behavior efficiencies.

4. Experimental Results

In our experiments, we use the NS2 simulator (ns-allinone-2.31 version) to simulate our proposed model in TCP (tahoe) protocol. We compare proposed model with Ji[26], Ji[27], and Garling [2]. We also generate the synthetic data for study.

We captured the SZABIST server dataset of TCP (protocol) packet sizes for specific hour to use real time packet size characteristics. Figure- 2 depicts the complete scenario, where node 0 and 4 are TCP and node 1 and 3 are UDP sender and receiver. Node 2 and 3 are behaving like routers. Data link between node 0 to 2 and 1 to 2 is 2Mbps, node 2 to 3 link 1.5 Mbps, and 3 to 4 and 3 to 5 has 1.7 Mbps, with 10 ms delay. To increase the packet dropping factor and creating real time router behavior, reduce the data link between 2 to 3 as compare to other data links.

Figure- 3 depicts the average traffic variation of normal Ji[27] and proposed model, we have just plotted a one hour network traffic sizes, where around first 1000 packets taken as references of packet length. We use synthetic data through randomly covert data bits are generated and embedded in simulation, overall normal, previous, and proposed traffic lies in similar type of variation range.

We have compared Ji[27] and proposed technique in different time duration and varying the W bit size 2 to 4. We have generated 5.5, 2.5 and 1 hour traffic from node 1 to 4, for both schemes; figure 4 depict the capacity graph. We use W (bits of covert data) as 2-bit. In Ji [27]

technique produced only 2 bits of covert data per packet throughout TCP transmission. Proposed technique is utilizing the packet payload as covert data, which increase data rate with minor effects of throughput of the data, as shown in the table-1. Overall TCP throughputs is shown in figure-5, which shows that proposed technique does not effect the overall TCP data throughputs of node 1 to 4, but actually application data is suffer from the covert data as depicted in table-1, because covert data is directly proportional to actual data. In figure-5 throughput graph is marked with normal TCP data and covert data with green and red color.

Congestion window and (round trip time) RTT delay graphs verses time are shown in figure 6 and 7. Congestion window and RTT graphs are same for both Ji [27] and proposed technique, because both have same throughput and overall data transmission rate. In proposed technique internal use of packet payload for covert data, which is considering as normal TCP data and it is in-effective for TCP congestion window and RTT.

In table-1, TCP Data, Covert Data, Throughput, Overall Throughput are in bytes. Table-1 shows the 1 hour traffic of TCP packet transmission, where V knows as stego-columns as 3, which indicate to send covert data into packet payload. Simultaneously transposition time of matrix, T time is 360 seconds. Proposed covert data capacity is very high instead of Ji[27] approach.

In table-2 shows the 5.5 hour TCP traffic transmission with W-bit 2. In Table-3 shows the 2.5 hour TCP traffic transmission with W-bit 2. In Table-4 shows the 1 hour TCP traffic transmission with W-bit 3. Table-5 shows the 1 hour TCP traffic transmission with W-bit 4.

Overall throughput and all other characteristics are same as in previous technique. Where Ji [26] and Ji [27] techniques results are same, because Ji [27] is the improved version of Ji [26] in context of packet sizes, its covert data capacity is same. So indirectly our proposed technique has much higher capacity instead of Ji [26] and Ji [27]. Garling [2] uses maximum 8 bits to send covert data in each packet, because its packet range is 256. Its capacity is also very small as compared to our proposed model; quantitative figures are shown in table 6.

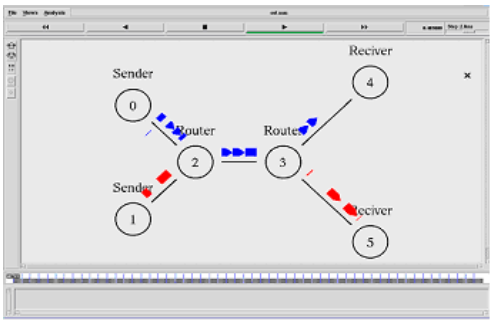

Figure-2 Simulation Scenario

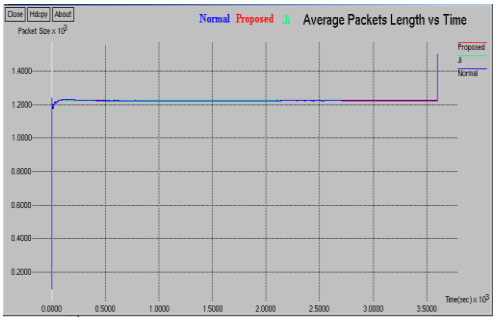

Figure -3 Average Traffic Variations

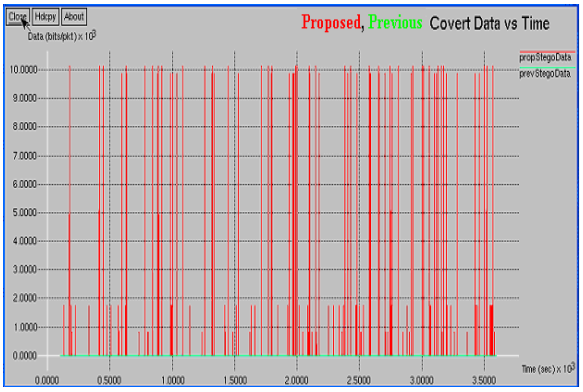

Figure- 4 Capacity Graph

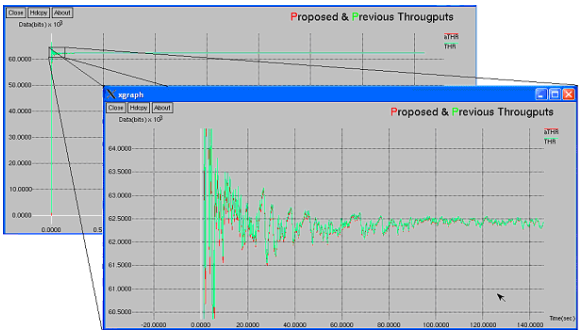

Figure- 5 Throughput Graph

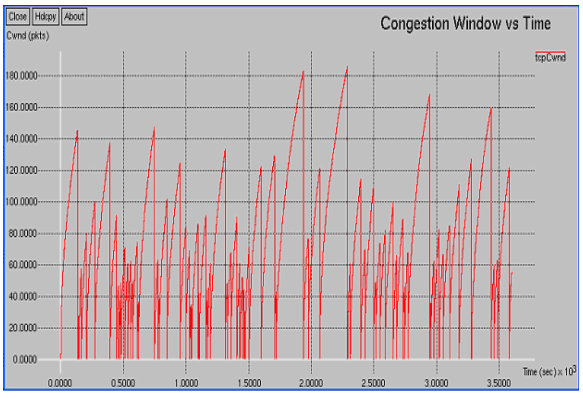

Figure- 6 Congestion Window

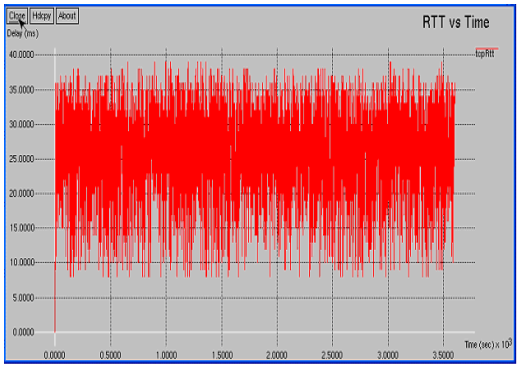

Figure -7 Round Turn Trip Time (RTT)

Table 1. 1 Hour Traffic Characteristics

Parameter	Proposed	Ji [27]
W-bits	2	2
N x M	4x100	400
V(Stego column)	3	X
T(Transposition)	360 sec	X
Packet Range	1 to 1460	1 to 1460
Packet Sent	14400	14400
Traffic Time	3600sec–1 hr	3600sec–1 hr
TCP Data	11097829	11097829
Covert Data	106233	3500
Throughput Except Covert	10991596 (99.42%)	11097829 (100 %)
Overall Throughput	11097829 (100%)	11097829 (100%)
Packet Loss	0.035 %	0.035 %

Table 2. 5 .5 Hour Traffic Characteristics

Parameter	Proposed	Ji [27]
W-bits	2	2
N x M	4x100	400
V(Stego column)	3	X
T(Transposition)	360 sec	X
Packet Range	1 to 1460	1 to 1460
Packet Sent	80000	80000
Traffic Time	20000 sec 5.5 hr	20000 sec 5.5 hr
TCP Data	58208332	58208332
Covert Data	677662	19900
Throughput Except Covert	57540670 (98.58%)	58208332 (100 %)
Overall Throughput	58208332 (100 %)	58208332 (100 %)
Packet Loss	0.022 %	0.022 %

Table 3. 2 .5 Hour Traffic Characteristics

Parameter	Proposed	Ji [27]
W-bits	2	2
N x M	4x100	400
V(Stego column)	3	X
T(Transposition)	360 sec	X
Packet Range	1 to 1460	1 to 1460
Packet Sent	40000	40000
Traffic Time	10 k sec 2.5 hr	10 k sec 2.5 hr
TCP Data	29534246	29534246
Covert Data	416757	9900
Throughput Except Covert	29117489 (98.58%)	29534246 (100 %)
Overall Throughput	29534246 (100 %)	29534246 (100 %)
Packet Loss	0.026 %	0.026 %

Table 4. 1 Hour Traffic Characteristics

Parameter	Proposed	Ji [27]
W-bits	3	3
N x M	8x50	400
V(Stego column)	3	X
T(Transposition)	360 sec	X
Packet Range	1 to 1460	1 to 1460
Packet Sent	14400	14400
Traffic Time	3600sec–1 hr	3600sec–1 hr
TCP Data	10807092	10807092
Covert Data	215960	5250
Throughput Except Covert	10591132 (98.00%)	10807092 (100 %)
Overall Throughput	10807092 (100 %)	10807092 (100 %)
Packet Loss	0.027 %	0.027 %

Table 5. 1 Hour Traffic Characteristics

Parameter	Proposed	Ji [27]
W-bits	4	4
N x M	16x25	400
V(Stego column)	3	X
T(Transposition)	360 sec	X
Packet Range	1 to 1460	1 to 1460
Packet Sent	14400	14400
Traffic Time	3600sec–1 hr	3600sec–1 hr
TCP Data	10781094	10781094
Covert Data	215960	5250
Throughput Except Covert	10404374 (96.50%)	10781094 (100 %)
Overall Throughput	10781094 (100 %)	10781094 (100 %)
Packet Loss	0.023 %	0.023 %

Table 6. Comparing Capacity with Models

Models Wbit-2	**Data Sent bytes**	**Covert Data bytes**
Proposed	58208332	677662
Ji [27]	58208332	19900
Ji [26]	58208332	19900
Garling [2] Wbit-8	58208332	318949

5. Conclusion

In this paper we proposed a high capacity covert channel in network protocol. Our proposed model utilized the normal packet length feature and also packet payload for covert data communication. It is temper resistance and time efficient. Due to after (predefine time) T packet transmission it's reshuffling of normal traffic reference increase its temper resistance as compared to previous technique. Our proposed technique is also effective for homogeneous type of data (video/audio) as well as heterogeneous type data, because due to reshuffling of matrix after specific transmission, correlation between same types of covert data is removed to specific range of packet lengths. Our technique is temper resistance for network traffic detector.

References

[1] B. Lampson, "A Note on the Confinement Problem", Commun. ACM, vol. 16, no. 10, Oct. 1973, pp. 613-615.

[2] C. G. Girling, "Covert Channels in LANs", IEEE Trans. Software Engineering, vol. SE-I3, no. 2, Feb. 1987, pp. 292-296

[3] National Institute of Standards and Technology. Trusted Computer System Evaluation Criteria. 1983.

[4] Ray Sbrusch **, "**Network Covert Channels: Subversive Secrecy", SANS InfoSec Reading Room - Covert Channels 2006

[5] National Computer Security Center, US DoD, "Trusted Computer System Evaluation Criteria," Tech. Rep. DOD 5200.28- STD, National Computer Security Center, Dec. 1985, http://csrc.nist.gov/publications/history/dod85.pdf

[6] M. Wolf, "Covert Channels in LAN Protocols," Proc. Wksp. Local Area Network Security (LANSEC), 1989, pp. 91–101.

[7] T. Handel and M. Sandford, "Hiding Data in the OSI Network Model," Proc. 1st Int'l. Wksp. Information Hiding, 1996 pp. 23–38.

[8] N. B. Lucena, G. Lewandowski, and S. J. Chapin, "Covert Channels in IPv6," Proc. Privacy Enhancing Technologies (PET), May 2005, pp. 147–66.

[9] D. Kundur and K. Ahsan, "Practical Internet Steganography: Data Hiding in IP," Proc. Texas Wksp. Security of Information Systems, Apr. 2003

[10] G. Fisk et al., "Eliminating Steganography in Internet Traffic with Active Wardens," Proc. 5th Int'l. Wksp. Information Hiding, Oct. 2002.

[11] A. Hintz, "Covert Channels in TCP and IP Headers,"2003, http://www.defcon.org/images/defcon-10/dc-10-presentationsdc10-hintzovert.ppt

[12] J. Levy, J. Paduch, and B. Khan, "Superimposing permutational covert channels onto reliable stream protocols," in Proceedings of MALWARE 2008, Alexandria VA, Oct. 2008.

[13] Adel El-Atawy, Ehab Al-Shaer: "Building Covert Channels over the Packet Reordering Phenomenon". INFOCOM 2009: 2186-2194

[14] C. H. Rowland, "Covert Channels in the TCP/IP Protocol Suite," First Monday, Peer Reviewed Journal on the Internet, July 1997.

[15] C. Abad, "IP Checksum Covert Channels and Selected Hash Collision," tech. rep., UCLA, 2001.

[16] S. Zander, G. Armitage, and P. Branch, "Covert Channels in the IP Time To Live Field," Proc. Australian Telecommunication Networks and Applications Conf. (ATNAC), Dec. 2006.

[17] Telvis E. Calhoun, Jr. , Reed Newman , Raheem Beyah, "Authentication in 802.11 LANs using a covert side channel", Proceedings of the 2009 IEEE international conference on Communications, p.1034-1039, June 14-18, 2009, Dresden, Germany

[18] Mario Cola, Giorgio De Lucia, Daria Mazza, Maurizio Patrignani, Massimo Rimondini., "Covert Channel for One-Way Delay Measurements". In Proc. International Conference on Computer Communications and Networks, Aug 2009

[19] Wojciech Mazurczyk and Krzysztof Szczypiorski. "Steganography of VoIP Streams". Lecture Notes in Computer Science (LNCS) 5332, Springer-Verlag Berlin Heidelberg, Proc. of The 3rd International Symposium on Information Security (IS'08), 2008.

[20] Mazurczyk W., Smolarczyk S., Szczypiorski K.: "Hiding Information in Retransmissions", In: Computing Research Repository (CoRR), abs/0905.0363, arXiv.org E-print Archive, Cornell University, Ithaca, NY (USA), May 2009.

[21] Hassan Khan, Yousra Javed, Fauzan Mirza and Syed Ali Khayam "Embedding a covert channel in active network connections", Proceedings of the 28th IEEE conference on Global telecommunications, 2009.

[22] M. Wolf, "Covert Channels in LAN Protocols," Proc. Wksp. Local Area Network Security (LANSEC), 1989, pp. 91–101.

[23] Xiapu Luo Chan, E.W.W. Chang, R.K.C. "CLACK: A Network Covert Channel Based on Partial Acknowledgment Encoding". ICC '09. IEEE International Conference on 14-18 June 2009.

[24] M. A. Padlipsky, D. W. Snow, and P. A. Karger, "Limitations of end-to-end encryption in secure computer networks", Tech. Rep. ESD-TR-78-158, Mitre Corporation, August 1978.

[25] YAO Quan-zhu and ZHANG Peng, "Coverting channel based on packet length", vol.34 No.3 Computer Engineering, February 2008.

[26] Liping Ji, Wenhao Jiang, and Benyang Dai, "A novel covert channel based on length of messages", International Conferenceon e-Business and Information System Security, 2009.

[27] Liping Ji, Haijin Liang, Yitao Song, Xizmu Niu, "A Normal Traffic Network Covert Channel", Computational Intelligence and Security, 2009.

[28] G. J. Simmons, "The Prisoners' Problem and the Subliminal Channel," in Proceedings of Advances in Cryptology (CRYPTO), pp. 51–67, 1983.

Author

Mehdi Hussain completed B.S (Computer Science) from Islamia University of Bahawalpur and currently doing M.S (Computer Science) at SZABIST Islamabad, also working as senior software engineer in Private Software House. Research areas of interest are Image steganography, Image Compression, Information Security and so on.