Volume 4, Issue 12, December 2014 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

Design of a Steganographic System for Hiding Information in TCP/IP Packets

Erika Llanes, Roberto Gómez, Maximiliano Canché
UADY, Unidad Multidisciplinaria Tizimín,
ITESM campus Estado de México

***Abstract*—This paper presents the design of a steganographic system for hiding information in network packets, specifically in TCP / IP (Transmission Control Protocol / Internet Protocol) packets. The system uses fields of the packet headers, which are not regularly used or have any intrinsic redundancy at the same protocols to send secret information to the recipient.**

***Key words*—Steganography, network packets, TCP / IP protocol, packet IP, Packet TCP.**

I. INTRODUCTION

Computer security is one of the major current debates; the emergence of new technologies and new applications bring the need for new protection mechanisms. In this context the secret, which is important for safety in a shared communication medium such as the Internet, becomes an issue that deserves more attention.

Computer security covers several aspects such as: equipment protection, stored data integrity, privacy in communications, among others; for each issue there are specific means for reducing exposure to risk. For instance, computer protection can be achieved with voltage regulation systems and stabilization of electric intensity, fire systems, systems against theft etc.

Regarding the privacy and confidentiality of information, specifically for transmission, there are two ways to protect it from people without access privileges: cryptography and steganography. Cryptography involves transforming information so that those who watch can not access their content [1]. Steganography, moreover, is to hide information so that others do not even know that there is a hidden message [2].

The main feature that distinguishes steganography of cryptography is that only the sender and the receiver know that there is a hidden message; Moreover cryptography refers to the transformation of information so that anyone who sees the encrypted message may not know its contents without having the decryption key.

This work is presented as follows: after this introducing section, the use and historical development of steganography are described, from the Greeks to the current computing techniques in order to protect intellectual property. Subsequently a classification of steganographic techniques is presented. Next, the structure of the TCP / IP packets is exposed, and the family of Internet protocols is presented (and described) in general terms. In the final part of this paper the proposed steganographic system called Mukul (which in Mayan language means secret or hidden), mainly composed by ANA and BOB functions, is presented. In that section, the design of steganographic system is detailed and implementation considerations are described. Finally, the tasks in progress and future works are described.

II. THE PROBLEM AND TECHNOLOGY

Like many security tools, steganography has several applications [3]; one of them is the digital watermark, which emerged as a solution to copyright [4]; another is the fingerprinting or fingerprint, which contains copyright information and information of the user who has acquired the rights to use that object. Another utility that we have had throughout our history is in military communications; as an example, it is said that Al-Qaeda planned the attacks of September 11, 2001 by Internet using some steganographic techniques [5], although this has not been proven [6].

Many research efforts within the digital steganography have been directed towards data hiding using image files [7], and conversely, the use of network protocols to hide information received little attention [8]. It is noteworthy that the weaknesses found in the TCP / IP protocols and the amount of information transmitted by Internet allow us finding an ideal channel to send hidden messages providing privacy to the data during its transmission.

A. Steganography

The word steganography comes from the Greek words στεγω (covered) and γραφειν (writing), which literally means hidden writing; and it is defined as the art of hiding information, i.e. discreetly hide data in a given carrier medium [3].

B. Development of Steganography

In [3] the authors state that references to steganography begin in ancient Greece. Herodotus (c 486 -. 425 BC) relates that Histiaeus, around 440 BC, tattooed a message on the shaved head of a slave, then growing his hair. The author notes that

Demeratus wanted to communicate to the city of Sparta that Xerxes had plans to invade Greece. To avoid being caught by spy, he wrote their messages on tables that were then covered with wax, which was the most usual way to send text messages, so that tables seemed not to have been used.

Another ancient author, Eneias of Tactia describes various techniques such as hidden messages in the soles of shoes or written on tablets which were then painted white. He proposed hiding text changing the height of the letters or making small marks above or below each letter [9].

In the nineteenth century Brewster proposed hiding messages "in no more than one point spaces" [3]. In 1860 a French photographer named Dragon, had established a technique for making very small images. Finally, during the Second World War were produced microdots, so small that they had the same size as a spell point or a comma [10].

Other methods have used invisible ink, as Wilkins proposed in a book published in 1694; and although its use spread from the development of advanced techniques of organic chemistry, now it has fallen into disuse due to the existence of products that can make visible any alteration in the paper fibers. [11].

In *Steganographia* (1499) of Tritheim Johannes Heidenberg a quite advanced steganographic system was included, but the general theme of the book, magic and accelerated learning methods, produced that it was never considered a serious [12] publication.

Finally *Schola Steganographica* (1665) of Gasparis Schott, is especially significant in the field of cryptography and steganography since in it the knowledge of the era are discussed regarding the hidden or encrypted writing, some of them discussed above, but in that paper takes a deep shift in the focus of study: Schott away from the esoteric and magical to focus steganography and cryptography from the point of view of art and science [3].

C. Cryptography vs. Steganography

Cryptography and steganography are different and even complementary techniques. While the first one is responsible for making the garbled message from unauthorized agents, the latter provides mechanisms to cause the message to be undetectable, regardless of it is encrypted or not. In fact, with insecure communication channels, the simple destruction of the message by an attacker may be sufficient for his purposes; no matter they can or can not access its content [13]. In this sense, one can say that steganography has been the most widely used technique throughout history to escape censorship.

Steganography and cryptography are means providing discretion and restrict access to certain information. If a spy is listening to the conversation it will be possible to intercept a cryptographic message. A solution to this is steganography, which hides the presence of a message so that the spy who hears all communication may not know that there is a steganographic message. Thus steganography, hiding the presence of a message, and cryptography, which hides the contents thereof, are complementarily used.

A suitable combination of cryptography and steganography can allow that, although the attacker knows completely the hiding mechanism of secret message on the carrier, he only recover some parts whose statistical properties are equal to those of white noise, so that the message originator may repudiate it, preventing it forces him to facilitate their cryptographic keys or being subjected to reprisals, as it will be mathematically impossible to prove the existence of the message [14].

D. Simmons Model

The classic model for the invisible communication was introduced in 1983 by Gustavus J. Simmons [15] as the "problem of prisoners" which formulates it as follows: "A and B are arrested for a crime and they are locked in two different cells, they want develop a plan of escape but unfortunately all communications between them are inspected by W, a security guard, who can detect any encrypted message, so that they must find some technique to hide their message in a communication seen as innocent "[3], see Fig. 1.

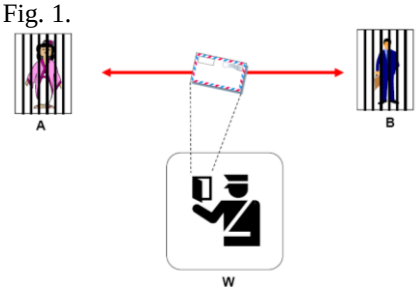

Figure 1. The problem of the prisoner.

A widely used technique in modern times for information hiding is digital watermark or watermarking. The technique involves inserting a message in a visibly or invisibly way inside of a digital object with the purpose of associating the protected object with its owner, for reasons such as copyright protection [4] or for the protection of copying and erased by an intruder. The main purpose of watermarks is to highlight the misuse of data by an unauthorized user. In this discipline it is essential that the brand is not lost even after subjecting the data to different processing.

Another technique used for data hiding is the fingerprint which defends copyright and combats unauthorized copying content; also it contains user information that has acquired the rights to use that object. This technical involves inserting a set of bits in the contents of the product to be protected without perceiving that in the final result. These brands also contain accurate and unique user information.

Traffic volume and easy access to an open channel such as the Internet, facilitate secret communication, this is achieved by carefully selecting those components that contain or display redundancy or irrelevance, to create the presence of a secret channel which B.W. Lampson defined as "the channel which is used to transmit information but was not designed for it" [16]. Packets and network protocols are good candidates for carriers in Internet´s steganography since, by virtue of its ambiguous nature, can be modified to embed secret information without affecting the flow of data [17].

III. STEGANOGRAPHY BASED ON TCP/IP NETWORK PACKETS

A. Network Packets and Protocols

Information stored on electronic media is normally stored using a binary code. The most widely used standards are the ASCII and EBICDIC. In the field of networks to transmit information, bytes are grouped into larger sets known as packets.

To group multiple bytes in packets it is necessary that the containment means are established. These means define both the start and the end of the package as well as other attributes that allow bring coherence and consistency to package [18]. This package creation follows certain rules already established that together are called protocols. There are several protocols for various purposes for the exchange of information on the Internet; the accepted standard is called TCP/IP [18].

The TCP / IP protocol adds the data to be communicated, certain information known as a header that describes the type of information and how to treat it [18].

Some parts or fields of the header of a TCP/IP packet are not usually used in most of the messages transmitted [19]. In this space it is possible hiding information.

To hide information in the headers of each packet of TCP/IP protocol, it is necessary to consider that this protocol actually consists of a set of protocols and all of them are present in both the emitter and receiver. This protocol stack is present in every computer using Internet; since it was created over thirty years ago, some of the fields are no longer used or can be changed without changing the information or affect the correct data transmission [20].

The headings are intended to provide essential data to the receiver of information. IP headers contain information to packets are properly addressed, see Fig. 2.

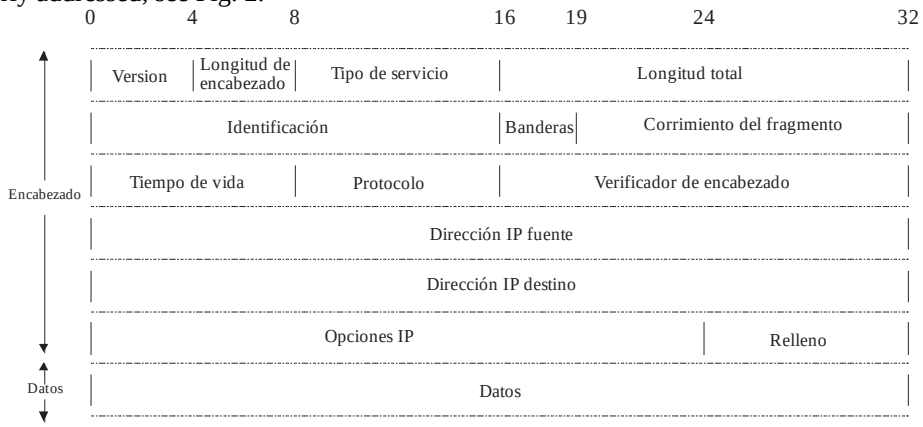

Figure 2. IP packet.

It is important to note that some header fields are subject to change without affecting routing. One of these is the IP identifier. When a data packet is too large, it is necessary to fragment it and all segments should have the same IP identifier, which normally will be increased per unit, but any number can be used and the protocol still will work correctly when assembling the original package [17].

The transmission control protocol, TCP, is used for data reliable transmission; also all computers connected to the Internet use TCP. In the headers of TCP, sequence and acknowledgment numbers are used to indicate how many data are sent and how many were received [21]. At the start of transmission both numbers are arbitrarily assigned, so the sent first packet may contain information hidden in these numbers, because on that one occasion they have no purpose. See Fig. 3.

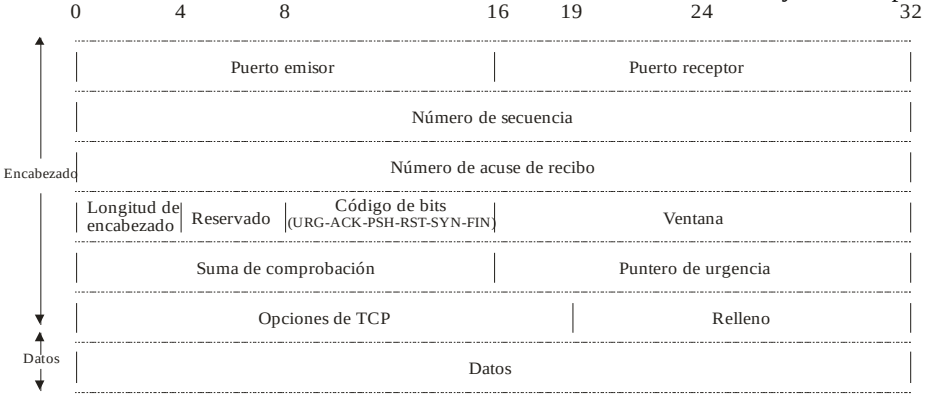

Figure 3. TCP packet.

B. Steganography on TCP/IP Packets

It is assumed that A and B share information openly about a computer network and use hidden data embedded in the TCP / IP protocols to communicate secret information [22].

Data are hidden through a steganographic algorithm that takes as input a secret message Ck, a sequence of network packets Pk (known as secret network packet sequence) and possibly a secret key to generate a stego-sequence of network packets Sk (containing Pk while carrying hidden Ck). The secret message Ck, characterized as a sequence of network packets Sk is sent by the computer network to B and it travels through a non-ideal channel. B estimates the secret information to produce C * k (Fig. 4) [20].

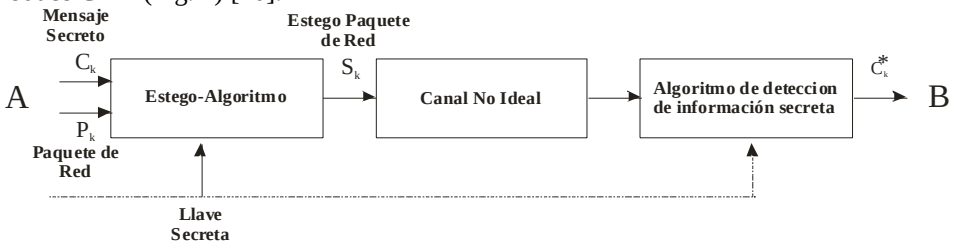

Figure 4. Transfering information from A to B using steganography.

IV. ARCHITECTURE OF THE PROPOSED STEGANOGRAPHIC SYSTEM

In this section the design of a steganographic system called Mukul is described. That system is conceived as a steganographic application over TCP / IP packets to send a secret message from a source host to a destination host using as carriers headers of TCP/IP packets flowing through the network, with the particularity avoid suspicion.

The design of Mukul system is composed of the following modules: ANA, that deals with the concealment and sending the message y BOB, that deals with the reception and retrieval of the message.

A. Description of the ANA Module

The ANA module consists basically of a sniffer and a packet injector. The main tasks of ANA are encode the message to be secretly sent, capture TCP/IP packets flowing through the network and select the appropriate packages to write them secret bits. Through the injector also takes care of sending over the network stego-packages to transmit the secret message to the destination host.

The functionality of this module is described in the following sequence:

It starts reading the secret message in plain text format and encoding it for transmission.

From ANA, through a sniffer, network packets in promiscuous mode are read with the intention to capture all packets that are going through the network.

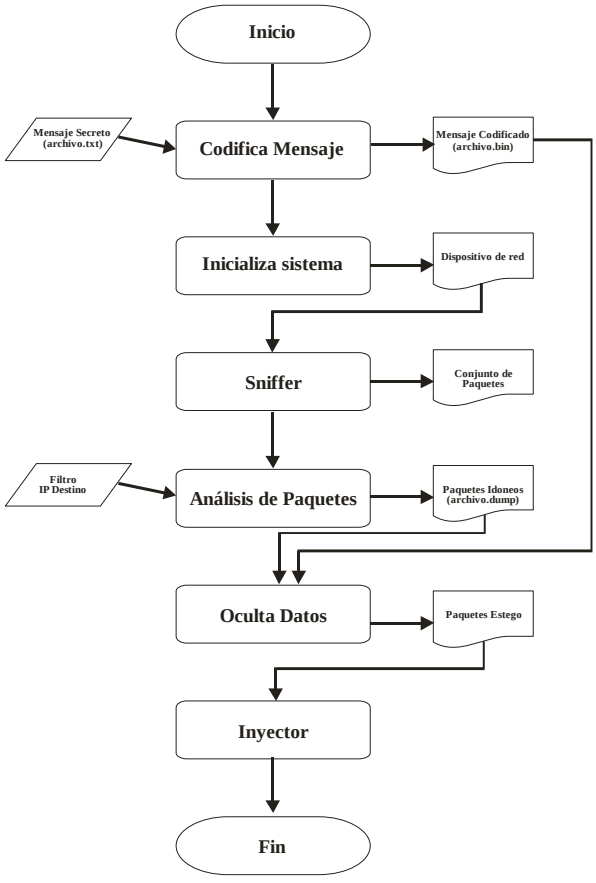

Figure 5.Diagram of the ANA module.

Captured packets are analyzed in order to select those that are suitable for data hiding. The package attributes providing such eligibility are: Are aimed at the destination of the hidden information, have unused bits and do not wear segmented data, i.e. they are unique.

The secret message is received and the possible bits are hidden in the IP and TCP header fields identified as suitable for shipment.

From ANA, through a nozzle the stego-packets are injected to the network with hidden data into headers and other data such as were captured from the network. In Fig. 5, the diagram with ANA functionality is presented.

B. Description of the BOB Module

The BOB module consists basically of a sniffer which is sniffing the network for packets containing secret bits to extract them, decrypt and recover the original message.

The functionality of this module is described below:

In the BOB module, the sniffer is running to capture TCP / IP packets that are directed specifically to this node (normal mode) and whose source node is specifically the transmitting node where ANA is running.

When all packets have been captured at the receiving node, BOB unpacks the hidden message and displays it in clear text. In Fig. 6 the diagram with BOB functionality is displayed.

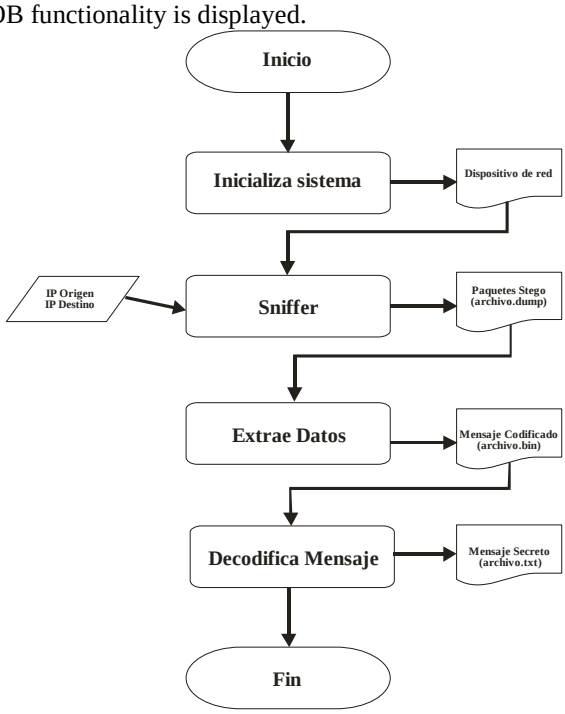

Figure 6. Diagram of BOB module.

C. Fields of TCP / IP Packets Used by Mukul System

To send the hidden information, Mukul uses the fields as it is shown in Table 1:

TABLE I FIELDS USED

Field	Header	Available bits
ToS	IP	4
TTL	IP	1
Id	IP	16
Flag DF	IP	1
Urgent Point	TCP	16

A brief description of each field is done below:

ToS (type of service): A field in the IP header of 8 bits, which defines how the package should be handled by routers. This field is divided into two subfields: precedence (three bits) and type of service (four bits); the remaining bit is not used. Subfield precedence, worth 0-7 defines the priority of the packet in question. In version 4, precedence subfield is not used. Mukul will use the four bits that are not used in IPv4.

TTL (Time to Live): A field in the IP header of 8 bits, designed to contain a time stamp which is decremented by each visited router, the packet will be discarded when the value reaches zero. Currently this field is primarily used to control the maximum number of hops taken by the package (visited routers). Mukul will transmit a hidden bit in this field; if the TTL value> 64 then the encoded value is 1 and if TTL <= 64 the encoded value is 0. The coded value persists unless more than 192 jumps were given in the transmission.

ID (Identification): A field in the IP header of 16 bits, used in fragmentation to identify a package originated in the source host. The IP protocol uses a counter to label the packages. The counter is initialized with a positive number and when the IP protocol sends a packet, it copies the current counter value in the ID field and increments the counter by one and when a packet is fragmented, this value is copied into all fragments in its identification field, which helps the end host in the process of reassembling. In the application the 16 bits are used to transmit hidden data and a condition for selecting a suitable package for the transmission of secret bits is the package does not should be a fragment of a datagram. Furthermore, if the source and destination are on the same network segment, the secret message will be transmitted without danger of being corrupted.

Flags: A field in the IP header of 3 bits used in fragmentation, the first bit is reserved; the second bit is called do not fragment (DF); if the value of this bit is one, the package must not be fragmented, a value of zero means that the packet can be fragmented if necessary. The third bit is called more fragment; a value of one indicates that this package is not the last bit but there are more fragments after him and if zero means it is the last fragment or that is unique. Mukul exploits the redundancy that occurs when there is no fragmentation to hide a bit in DF.

Urgent Point: TCP header field of 16-bit containing the sequence number after which information becomes urgent. Mukul uses 16 bits of the Urgent Point field to hide secret data and sets the URG control bit as 0. This is the field most likely to be detected with only a protocol analyzer.

D. Considerations in the Design of Mukul System

Mukul system intends to demonstrate the possibility of sending hidden information. So, the following assumptions are considered to facilitate the development of the application:

- Communication between ANA and BOB is point to point.
- Communication between ANA and BOB is not interrupted and the packets arrive smoothly to the receiver.
- The ANA module is run from a node specially configured to monitor the network, which is directly connected to the core switch of the Intranet that receives packets from n origins to m destinations but only Z is the target of interest.
- The packets flowing through the network are IPv4.
- Packets destined to BOB are sufficient to hide the secret message.
- BOB does not send ACK to ANA.
- There is no maximum length of message to send.
- Shipping time or number of packets required for shipping are not relevant.
- Information between ANA and BOB is not encrypted.
- No information compaction technique is used.

V. CONCLUSIONS

Currently the international scientific community develops new mechanisms for data protection due to the importance and value of information in institutions.

This paper presents the design of a steganographic system that provides a tool for hiding data in TCP/IP network packages as a secure and imperceptible alternative of data transmission. Currently, this system is in the development stage to be tested in a controlled environment to evaluate its performance.

ACKNOWLEDGEMENTS

The authors acknowledge the facilities and support provided by the Faculty of Mathematics, University of Yucatan, Mexico to complete this work.

REFERENCES

[1] SCHNEIER, Bruce. Applied Cryptography. Second Edition. John Wiley & sons, 1996.11

[2] KIPPER, Gregory. Investigator's Guide to Steganography. Boca Ratón Florida: Auerbach, 2003.

[3] KATZENBEISSER, Stefan and PETITCOLAS, Fabien A.P. Information Hiding techniques for steganography and digital watermarking. Estados Unidos: Artech House Publishers, 2000.

[4] TIRKEL, A. Z., G. A. RANKIN, and R. van SCHYNDEL, "Electronic Watermark," in Digital Image Computing, Technology and Applications—DICTA 93, Macquarie University, 1993, pp. 666–673.

[5] KUNDUR, Deepa and AHSAN, Kamran. Practical Data Hiding in TCP/IP. Scientific Literature Digital Library. 2003.

[6] CASTRO S., Gray-World Team. 2006. How to cook a covert channel. [online] [Last accessed 23rd January 2007] Available from World Wide Web: <http://en.hakin9.org/attachments/pdf/cooking_channels_en.pdf>

[7] JOHNSON, Neil F.; DURIC, Zorac; JAJODIA, Suchil. Information Hiding: Steganography and Watermarking - Attacks and Countermeasures. Estados Unidos: Kluwer Academic Publishers, 2003

[8] AHSAN, Kaonmud. Covert Channel Analysis and Data Hiding in TCP/IP. University of Toronto, Canada, http://gray-world.net/papers/ahsan02.pdf 2002.

[9] MOSKOWITZ, Ivan and KANG, Moo. Covert Channels - Here to Stay?. Naval Research Laboratory, Department of the Navy, USA, http://chacs.nrl.navy.mil/publications/CHACS/1994moskowitz-compass.ps 1994.

[10] HANSMANN, F., Steganos, Deus Ex Machina Communications, <http://www.steganography.com/>, 1996.

[11] WAYNER, Peter. Disappearing Cryptography, Information Hiding: Steganography and Watermarking. 2da ed. San Francisco, Estados Unidos: Morgan Kaufmann Publishers, 2002.

[12] HASTUR, H., Mandelsteg, <ftp://idea.sec.dsi.unimi.it/pub/security/crypt/code/steg.tar.Z>, 1994.

[13] GALLAGHER, Peter, A Guide to Understanding Covert Channel Analysis of Trusted Systems. National Computer Security Center, USA, http://fas.org/irp/nsa/rainbow/tg030.htm 1993.

[14] ARACHELIAN, R., White Noise Storm, <ftp://ftp.csua.berkeley.edu/pub/cypherpunks/steganography/wns210.zip>, 1994.

[15] SBRUSCH, R., 2006. Network covert channels: Subversive secrecy. Tech. rep., SANS Institute, [online]. [Accessed 26th Feb 2007]. Available from the World Wide Web: <http://www.sans.org/reading_room/whitepapers/covert/1660.php>.

[16] CRAVER, S., On Public-Key Steganography in the Presence of an Active Warden, Technical Report RC 20931, IBM, 1997.

[17] KUNDUR, Deepa and AHSAN, Kamran. Practical Internet Steganography: Data Hiding in IP. ACM Workshop on Multimedia Security. 2002.

[18] STEVENS, Richard. TCP/IP Illustrated, Volume 1: The Protocols, New York, Addison-Wesley Professional computing series. 1994

[19] ROWLAND, Craig H. Covert Channels in the TCP/IP Protocol Suite. Peer Reviewed Journal on the internet. 1997.

[20] CAUICH, Enrique; GÓMEZ Roberto; RYOUSKE Watanabe. Data Hiding in Identification and Offset IP fields. Lecture notes in computer science. Springer, New York. Enero 2005.

[21] ABAD, Carlos. IP Checksum Covert Channels and Selected Hash Collision. University of California, USA, http://downloads.securityfocus.com/library/ipccc.pdf 2001.

[22] RFC 2131. 1997. Dynamic Host Configuration Protocol. [online] [Last accessed 23rd January 2007] Available from World Wide Web: < http://www.ietf.org/rfc/rfc2131.txt >

