Practical Development and Deployment of Covert Communication in IPv4

Vishal Bharti
Chandigarh University
29 PUBLICATIONS **56** CITATIONS
SEE PROFILE

Itu Snigdh
Birla Institute of Technology, Mesra
40 PUBLICATIONS **87** CITATIONS
SEE PROFILE

H. Benalla
University of Constantine 1
126 PUBLICATIONS **1,022** CITATIONS
SEE PROFILE

Rahman Mahmood
University of Diyala
3 PUBLICATIONS **13** CITATIONS
SEE PROFILE

PRACTICAL DEVELOPMENT AND DEPLOYMENT OF COVERT COMMUNICATION IN IPV4

[1]VISHAL BHARTI, [2]ITU SNIGDH
[1]Department of Computer Science & Engineering
Birla Institute of Technology,
Mesra, Ranchi, INDIA

[2]Department of Computer Science & Engineering
Birla Institute of Technology,
Mesra, Ranchi, INDIA
Email: [1]mevishalbharti@yahoo.com, [2]itu_snigdh@yahoo.com

ABSTRACT

Steganography is the science of hiding information or transmitting secret messages in a given host carrier for the purpose of enhancing value through undetectable covers. The paper focuses on the existent methods used with ipv4 and studies the various algorithms. Transfer of data over the internet crosses the different layers of the TCP and IP protocols. Each layer has its own characteristics, which indicate the scenarios in which it can best be used. Information can be hidden in the transport and network layers by usage of optional fields, semantic changes, and improper but acceptable construction of protocol data units (packets). Protocol steganography is the scheme to bypass the firewall. The TCP/IP has the following hidden channels namely the manipulation of HTML, XML, HTTP, UDP, TCP, IP, ICMP, Ethernet (CSMA/CD) in the four layers respectively.

KEYWORDS: *Network Security, Data Hiding, Covert Channel, TCP/IP, network security, toral automorphism, IPSec, packet sorting.*

1. INTRODUCTION

1.1 Network Security:

Network security involves the protection of an agency or internal network from threats posed by authorized or unauthorized connections. With the explosion of the public Internet and e-commerce, private computers, and computer networks, if not adequately secured, are increasingly vulnerable to damaging attacks. Hackers, viruses, vindictive employees and even human error all represent clear and present dangers to networks. And all computer users, from the most casual Internet surfers to large enterprises, could be affected by network security breaches. The Internet has undoubtedly become the largest public data network, enabling and facilitating both personal and business communications worldwide. Despite the costly risks of potential security breaches, the Internet can be one of the safest means by which to conduct business. As more and more communication is taking place via e-mail; mobile workers, telecommuters, and branch offices are using the Internet to remotely connect to their corporate networks; and commercial transactions completed over the Internet, via the World Wide Web, now account for large portions of corporate revenue.

1.2 Threats to Data

As with any type of crime, the threats to the privacy and integrity of data come from a very small minority of vandals. A single hacker working from a basic computer can generate damage to a large number of computer networks that wreaks havoc around the world. Furthermore, with the recent pervasiveness of remote connectivity technologies, businesses are expanding to include larger numbers of telecommuters, branch offices, and business partners. These remote employees and partners pose the same threats as internal employees, as well as the risk of security breaches if their remote networking assets are not properly secured and monitored.

1.3 Levels of Security

Data hiding techniques have become much more open and public in the last few years. This has created some confusion in the terminology and the differences between the different techniques. In (Petitcolas, Anderson et al. 1999, [3]) it is suggested a classification of the information hiding techniques as can be seen in Figure 1.1 – A Classification of information hiding techniques.

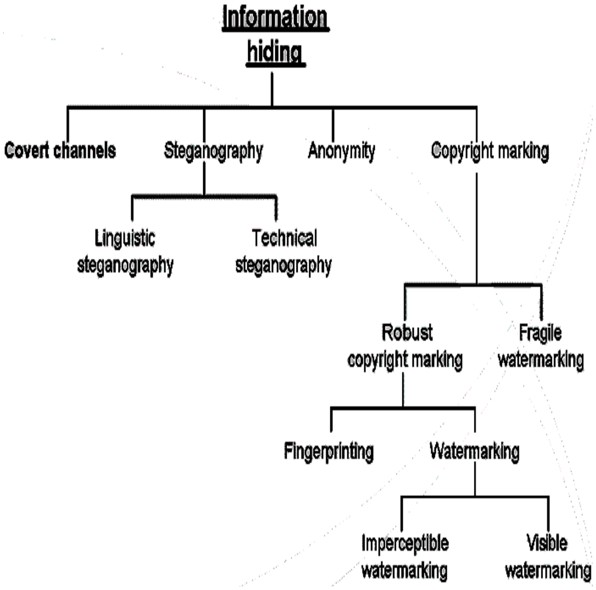

Figure 1.1: A Classification of information hiding techniques

Steganography it is the art of hiding information in ways that prevents the detection of hidden messages. In this case, the existence of the message it is not known, hence a successful attack will consist in detecting the existence. In the linguistic Steganography, the carrier medium is text whilst in the technical Steganography the carrier medium in non-text, such as graphic files, photos, and so on. [12]

In the case of the copyright marking, the most important characteristic it is the robustness against possible attacks [2]. In fact, there are many copyright marks that are not hidden, such as a company logo, and whose main purpose it is to differentiate one product, or service, against others and thus give a unique identity.

Cryptography is commonly used in information hiding. However, it is not included within this area as the message is not hidden, and is thus it is clear that there is a message, but it cannot be read unless the required decoding method is known.

1.4 Covert Channels

Covert channels can be regarded as one of the main sub-disciplines of data hiding. In data hiding, the two communicating parties are allowed to communicate with each other based on the security policy of the system while exploiting the features as associated with covert channel definition. A covert channel is one that is used for information transmission, but that is not designed nor intended for communications [11]. A resource state variable, for instance, is any system variable that can be used by a covert channel to signal information from one point to another within that system e.g. a variable showing *file status* at several points (states) in a system. By definition, the existence of covert channels must be non-detectable.

Covert channels are classified into *covert storage channels* and *covert timing channels*. Communication in a covert storage channel entails the writing of hidden data into a storage location (not meant for communication) by the transmitting party, and the subsequent retrieval of that information by the receiving party. In contrast, communication in a covert timing channel requires that the transmitting party signal information by modulating its own system resources such that the manipulation affects the response time observed by the receiving party.

Covert storage channels in case of TCP/IP utilize the *reserved fields, pad fields* and *undefined fields* of the frames. The fields identified, as means to covertly send information, can easily be detected through the implementation of automated mechanisms. Such mechanisms only monitor such fields, which would discard such frames utilizing these fields irrespective of their purpose.

2. STEGANOGRAPHY

Steganography is the art of hiding information in ways that prevent detection of hidden messages. In other way we can say that Steganography is the art and science of hiding information such that its presence cannot be detected. In Greek steganography means "covered writing" Steganography and cryptography are cousins in the spy craft family; the goal of the cryptography system is to conceal the content of the messages, the goal of information hiding or steganography is to conceal their existence.

2.1 Background to Steganography

Steganography is not a modern concept. An ancient example comes from the Histories of Herodotus. Xerxes planned to invade Greece and a warning message had to be passed to Sparta. Text was written on wax covered wooden tablets. The wax was removed, the message written on the underlying wood and covered with wax again. These "unused" tablets passed inspection easily. A formula for the information hiding process might look like this:

cover medium + embedded message + stegokey = stego-medium

The cover medium is any innocent looking digital image in which the secret message will be embedded. The stegokey is any additional information required to imbed the information. The resulting image is called the stego-image or stego-medium and is the final image to be sent. [12]

Cryptography can and should be used in conjunction with steganography. If a hidden message is discovered the attacker will need to decipher the code before the message is revealed. A simple comparison between steganography and cryptography is shown in Figure 2.1

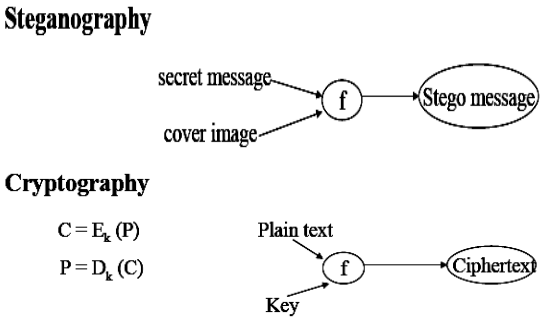

Figure 2.1: Steganography Vs Cryptography

2.2 TCP/IP Based Steganography

One of the most common ways of passing messages in modern times is through the use of the Internet. TCP/IP packets used to transport information across the Internet have unused space in their packet headers. The TCP packet header has 6 reserved bits and the IP packet header has 2. These provide excellent covert communication channels. The network packets in TCP/IP suite are shown in Figure 2.1.

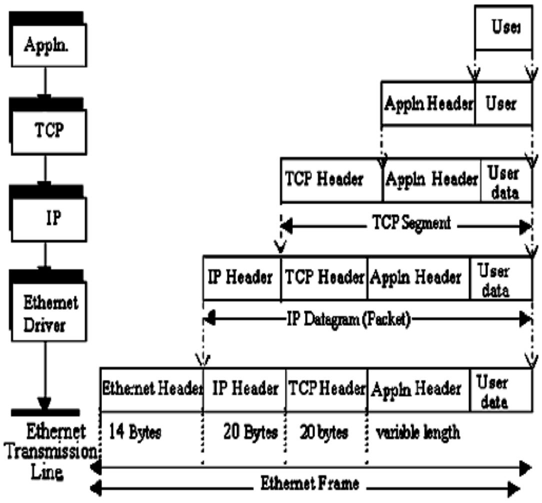

Figure 2.1: Network packets in TCP/IP

IP header - Possible Hidden Channels

♦PAD (padding bits) -bandwidth 31 bits/packet
♦IP identification -16 bits/packet
♦Fake source IP address -32 bits/packet
♦Usage of IP destination address as a flag -8 bits/packet
♦Usage of the unnecessary fields (ToS, options, some flags for example Don't Fragment - DF for the fragmented packet) -various bandwidth.

TCP Header- Possible Hidden Channels

♦PAD (padding bits) -bandwidth 31 bits/packet
♦Usage of chosen ISN (initial SN) -32 bits per connection
♦Usage of urgent pointer, when URG=0-16 bits/packet
♦Usage of reserved bits -6 bits/packet
♦Existence of data, when RST=1
♦Port numbers as analphabet (→)

One noteworthy feature of IP for our purposes is that it allows fragmentation and reassembly of long datagram, requiring certain extra header fields. TCP, on the other hand, does aim to provide a reliable channel to its clients. It has a stream oriented interface, and keeps its reliability properties even within networks exhibiting packet loss, reordering and duplication. Its features for implementing reliability and flow control give scope for data hiding coding [1] [2] [5] [6] [13].

The TCP/IP header can serve as a carrier for a steganographic covert channel if a header field can take one of a set of values. A schematic view of TCP/IP header is shown in Figure-2.2. The fields in italics provide an ample medium to hide data. The intruder should not be able to distinguish whether the header was generated by an unmodified TCP/IP stack or by a steganographic encoding mechanism. TCP/IP steganography exploits the fact that few headers are altered in transit. As mentioned above, IP packets can be fragmented, but (unless we are hiding data in the fragmentation-related headers) no information is lost. The time-to live field in the IP header is decremented each time the packet passes through a router, but the initial values used by IP stacks are well known, so this field gives little scope for steganography. [13]

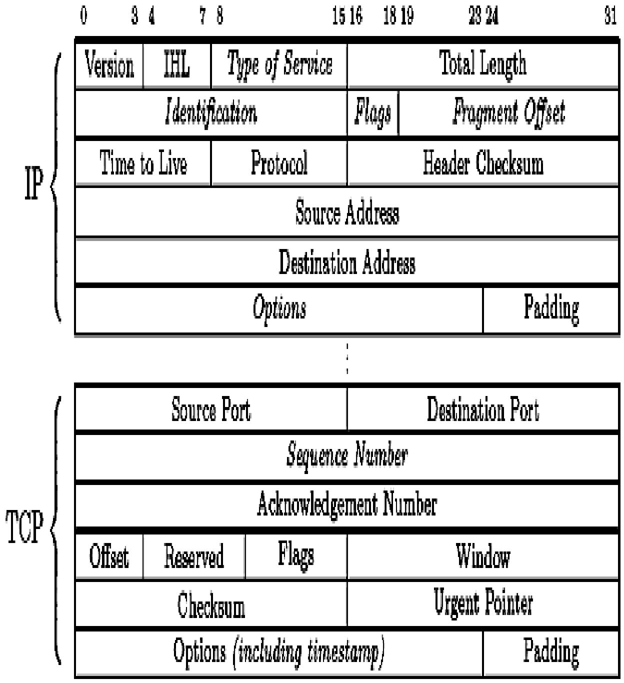

Figure 2.2: Basic TCP/IP header structure

3. DATA HIDING METHODS

3.1 Method One - Manipulation of the IP Identification Field

The identification field of the IP protocol helps with re-assembly of packet data by remote routers and host systems. Its purpose is to give a unique value to packets so if fragmentation occurs along a route, they can be accurately re-assembled. The first encoding method simply replaces the IP identification field with the numerical ASCII representation of the character to be encoded. This allows for easy transmission to a remote host which simply reads the IP identification field and translates the encoded ASCII value to its printable counterpart.

3.2 Method Two - Initial Sequence Number Field

The Initial Sequence Number field (ISN) of the TCP/IP protocol suite enables a client to establish a reliable protocol negotiation with a remote server. As part of the negotiation process for TCP/IP, several steps are taken in what is commonly called a "three way handshake" as was described earlier. For our purposes the sequence number field serves as a perfect medium for transmitting clandestine data because of its size (a 32 bit number). In this light, there are a number of possible methods to use. The simplest is to generate the sequence number from our actual ASCII character we wish to have encoded.

3.3 Method Three - The TCP Acknowledge Sequence Number Field - Bounce

This method relies upon basic spoofing of IP addresses to enable a sending machine to "bounce" a packet of information off of a remote site and have that site return the packet to the real destination address. This has the benefit of concealing the sender of the packet as it appears to come from the "bounce" host. This method could be used to set up an anonymous one-way communication network that would be difficult to detect especially if the bounce server is very busy. This method relies on the characteristic of TCP/IP where the destination server responds to an initial connect request (SYN packet) with a SYN/ACK packet containing the original initial sequence number plus one (ISN+1).

As discussed above, other protocols can be used in a similar manner and in some cases may provide a more reliable channel of data transmission as the packet can hold much more data.

3.4 Method Four – HICCUPS

HICCUPS (**HI**dden **C**ommunication system for **C**orr**UP**ted network**S**), takes advantage of imperfections of transmission medium environment – interferences and noise in communication medium – natural susceptibility to data distortion. HICCUPS is a steganographic system with bandwidth allocation for shared medium networks [4].

The "hearing" of all frames with data transmitted in medium and the possibility of sending corrupted frames with improper correction codes values are two important network features for HICCUPS. In particular, wireless area networks use air connection with variable bit error rate (BER) that creates opportunity to inject "synthetic" corrupted frames. In general the novelty of HICCUPS includes:

1. Usage of secure telecommunications network armed with cryptographic mechanisms to provide steganographic system and
2. Proposal of new protocol with bandwidth allocation for steganographic purposes based on corrupted frames.

Crucial information from a hidden communication system's perspective is exchanged in hidden channels. The remaining communications channels at Medium Access Control layer of IEEE LAN Reference Model are open for crypto analytical attack and left for penetration as bait or honey pot; they also serve normal data exchange.

Proposed system is destined to be implemented in a network environment with the following three properties:

P1: shared medium network with possibility of frame's interception,
P2: publicly known method of cipher initiation for instance by initialization vectors,
P3: integrity mechanisms for encrypted frames for instance one-way hash function, Cyclic Redundancy Code – CRC (note: CRC is rarely strong enough for protecting integrity, but it is used in IEEE 802.11 for such purpose).

3.5 Method Five-Through the packet sorting (IPSec)

This method presented by [1] [13] deals with the use of packet ordering to convey covert information. The possible ways to arrange objects in a set is surprisingly complex and offers a correspondingly large opportunity for steganography. Changing the order of the packets requires no change in the packet content (i.e. the payload and the headers are not affected).

Therefore no major modifications are expected either in the protocol definition/design or in the overall system in order to implement a data-hiding scenario. The sorting/resorting process holds a surprisingly large amount of information. Based on these facts, data hiding feasibility is explored in the TCP/IP protocol based on packet sorting and resorting processes at source and destination, respectively. The packet sorting and resorting processes require some reference in order to relate packet numbers to their actual order. The natural packet ordering is needed so that the stego packet ordering (sorting) can be undone (resorting). This reference is not available at the transport layer using TCP. Sequence number field and acknowledgement number field point to the number of octets of data and are not directly related to the packet number. Moreover, the data packetized at the transport layer can be broken down into fragments at the Internet layer and would further complicate the notion of a packet order.

3.6 Method Six-manipulation of the 6 bit flags in the TCP header

At the transport layer, TCP is intended to provide a reliable process-to-process communication service in a multi-network environment. TCP is, therefore, a connection-oriented and reliable transport protocol. The header of the TCP protocol has a 6-bit field labeled as code bits (*URG, ACK, PSH, RST, SYN, FIN*). These bits determine the purpose and contents of the TCP segment. These six bits tell the TCP header node how to interpret other fields in the header. There are 64 possible combinations for these six bits, out of which 29 combinations are considered to be valid as per the rules set forth by the protocol. [1] [13]. the *redundancy* condition within these possible code bit combinations is to be detected. One of the redundancy conditions is when the URG bit is not set, the Urgent pointer field (16 bit) of the TCP header becomes redundant and therefore can be used to have a storage covert channel. Likewise, redundancy conditions exist for all those possible cases wherein the URG bit is not set thereby making the urgent pointer field redundant. The SYN bit set can also have possible combinations either with the ACK bit set or the URG/PSH (not both at the same time) set to 1. Thus, the remaining bits are meaningless for the protocol enabling covert data transmission possibilities through TCP header.

3.7 Method Seven-Sequential transmission of 0s and 1s through DF bit

A covert communication can be done by employing the protocol suite of the network when both the sender and the receiver are well aware of the fact that the network administrator is very security cautious and the TCP/IP software is configured properly as per the security policy of the organization. The requirement of such covert communication is that both the ends should have knowledge of the *MTU* (maximum transmission unit) of their network and they are aware of the fragmentation strategy, which follow the standard design considerations of IP [13]. The embedding algorithm avails the fragmentation strategy of the Internet protocol and DF bit is used to send covert bit.

The two cases when a Complete datagram; moderate size; fragmentation not allowed since DF bit is set ;and when a Complete datagram; moderate size; fragmentation is possible, since DF bit is not set; but fragmentation will not take place since MTU is known can be sent to transmit ‘0’ and ‘1’ respectively. But the limitation is that both the ends need to remain on the same network, need to know the MTU and the covert information gets over unsuspicious only if this strategy is not used very often.

3.8 Method Eight- Manipulation of 16 bit identification field with DF bit

Each datagram represents unique multiple bit information if the combination of the DF bit with the identification field. Since the DF bit and identification field are independent, multiple covert channels within single packet can be obtained. The 16 bit identification field facilitates hosts to have more users in secret communications. This method provides one to many covert communication for multiple recipients of a single covert message.

The constraint is that all recipients need to connect to the same network and have a prior knowledge of the MTU.

3.9 Method Nine- Combination of identification field, version header and IP header

This scheme overrules the restriction of the prior knowledge of the MTU, it uses the identification field with the assumption that datagram must not contain “OPTIONS” field in the IP header. It utilizes the combination of identification field and the version and the internet header length field of the IPv4 header.

The XOR operator is used to encode and decode covert information. [1] [13]. This data hiding scenario is resistant to packet filtering firewalls and hence the recipients need not be on the same network.

3.10 Method Ten- Chaotic mixing and Toral Automorphism

This method uses the same identification field and the generation of the identifier is based on chaotic mixing which provides a structure of sorted sequences from an original sequence, the elements of which are statistically uncorrelated. The restriction to the scheme is that both ends should have a suitable method for exchanging keys. The identification field is used for embedding secret information and the deciphering is done with the help of three key parameters. The random numbers generated through the chaotic mixing leads to the covert information becoming non detectable against packet filtering and firewalls for recipients across the internet.

4. PROPOSED SCHEME

Assumptions:-

1> The algorithm is based on the assumption that network is ideal i.e.

what is sent is what is received (WISIWIR).

2> Keys are known to both sender and receiver

Table-1: Algorithm Symbols

P	Number of packets
n	Any arbitrary number
K	Main Key
k	Subkey
A[2][2]	Toral Automorphism matrix
O	Original Packet Sequence
S	Sender Packet Sequence
R	Receiver Packet Sequence

Table-2: Function list for algorithm

Read(x)	Read value into x
Seq_num(i)	Return the sequence number of packet i
Cal(P/K)	Return the floor value of P/K

<u>Sender end:-</u>

```
Read (n)
Set k=P*n
Initialize 'A' as any Toral Automorphism Matrix
Set k=A [2] [2];
for i=1 to P do
     {
     Set O[i] =Seq_num (i);
      }

M=Cal (P/K);
For i=1 to P do
{
If i<=K then
      {
      S [(i+3) mod K] = [(O [I] +1) mod K]
      }
else
   S[i] = (M-1) K+O[i];
}
```

<u>Receiver end:-</u>

```
M=Clear [P/K]
For i=1 to P
{
If (i<=K) then
R [(i+3) mod K] = (S[i] +1) mod K;
else
   {
   R[i] = (M-1) K+S(mod (k-i))
   }
}
```

5. POTENTIAL APPLICATIONS—DATA HIDING IN TCP/IP

Associating supplementary information sent via covert mechanisms employing packet header manipulation algorithms find the following application scenarios:

1. Enhanced filtering criteria in packet filtering routers (firewalls).
2. A client server architecture wherein several clients make a request to the FTP server, say of a library. Moreover, serving the request by transferring a digital image to the user, say, can have the same user information or library information tied to the content packets.

3. A logging process for the above application scenario based on the user or application specific-information completes the picture (i.e. logging of valid user), maintaining the record of user requests based on user information and ultimately serving the user requests by having either the user information or the server / source (library) information tied to the content packets to avoid unlawful use such as copyright violation.

4. Adding value to content delivery networks. A content delivery network is an overlay network to the public Internet or private networks, built specifically for the high performance delivery of content. Use of supplementary covert data adds intelligence to networking wherein the network makes path decisions based on more than simple labels such as IP address.

6. CONCLUSION AND FUTURE WORK

Our work puts forward various existing data hiding scenarios are put together along with a new one. The existence of covert channels is a phenomenon that exists almost everywhere. Covert channels find interesting applications in network security and in facilitating various network processes which are inline with modern concepts. The covert channel exploration in TCP/IP suite therefore has much potential in network environment.

Data can be embedded in the lower layers of the protocol i.e. the physical and data link layers with the limitation that such a procedure requires low level control of hardware which appears to be difficult. Moreover the messages may be stripped over by devices that connects networks at higher layers (e.g. IP router).and hence it is required that the recipient is on the same LAN [11]. Embedding at the presentation or the application layers requires a wide access to the machine from where the embedding is to be done to enable the anticipation of the most likely executed and used applications and hence their modification to carry the sublime messages over the traffic they generate.

The format of the files sent over HTTP or FTP may also provide repositories for embedded messages, thus building a high bandwidth channel. The remaining layers namely the network, transmission and session layers are the places where TCP/IP is most commonly used.

This paper studies a number of previously proposed schemes for embedding data within the TCP and IP protocol headers, and proposes an appropriate algorithm based on the chaotic mixing sequence to create a steganographic covert channel.

We propose to expand our work in implementation of the algorithm on group management protocols that have remained untouched. This exploratory research is not complete in the sense that all protocols are not evaluated. The analysis does not cover IPv6 which can be another potential avenue. Similarly, UDP (user datagram protocol) has also not been covered. Packet header manipulation approach can also be applied on routing protocols like RIP (routing information protocol), BGP (border gateway protocol, OSPF (open shortest path first).

BIBLIOGRAPHY

[1] Kamran Ahsan and Deepa Kundur, “Practical data hiding in TCP/IP”. *In Proc. Workshop on Multimedia Security at ACM Multimedia*, December 2002.

[2] Rowland, C. H. (1996). "Covert Channels in the TCP/IP Protocol Suite." First Monday Peer-Reviewed Journal on the Internet.

[3] R. J. Anderson and A. P. Petitcolas, “On the limits of steganography” *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 4, pp. 474–481, May 1998.

[4] Szczypiorski, K, “HICCUPS: Hidden communication system for corrupted networks”. International Multi-Conference on Advanced Computer Systems. 2003 http://krzysiek.tele.pw.edu.pl/pdf/acs2003-hiccups.pdf.

[5] D. E. Comer, “*Internetworking with TCP/IP, Principles, Protocols and Architectures”.* New Jersey 07458: Prentice Hall, Upper Saddle River, fourth ed., 2000.

[6] S. M. Bellovin, “Security problems in the TCP/IP protocol suite,” *Computer Communication Review*, vol. 19, pp. 32–48, April 1989.

[7] Andy Oram The McCain-Kerrey “Secure Public Networks Act”
http://www.cpsr.org/cpsr/nii/cyber-rights/web/mccain-kerrey.html

[8] Quinion, Michael. “Weird Words Section. Steganography”, 23 Oct 1999
http://www.quinion.com/words/weirdwords/ww-ste1.htm ,12 Jan 2002.

[9] John Bartlett. “The Ease of Steganography and Camouflage”, Sans Institute, 17 March’2002

[10] Pierre Richer, “Steganalysis: Detecting hidden information with computer forensic analysis”, SANS Institute 2003.

[11] Steven J. Murdoch and Stephen Lewis, “Embedding Covert Channels into TCP/IP”, University of Cambridge, http://www.cl.cam.ac.uk/users/fsjm217, srl32g/

[12] Jonathan Househam, “Steganography and related topics including Tempest, with a focus on Image Steganography”, 28 March 2004

[13] D. Kundur and K. Ahsan, “Practical Internet Steganography: Data Hiding in IP” *Proc. Texas Workshop on Security of Information Systems*, College Station, Texas, April 2003