Network-based Steganography using Encryption in TCP/IP Header

Joshi Rana
M.Tech Student
ITM University
Gurgaon

Amanpreetkaur
Asst. Professor
ITM University
Gurgaon

Nitin Malik
Asst. Professor
ITM University
Gurgaon

ABSTRACT

Steganography, introduced in 2003, is a techniques used for hidden communication between two covert parties. It is an art of hidden communication. It also relates to the areas like network protocols and security for practical data hiding in communication networks using Transmission Control Protocol/Internet Protocol (TCP/IP). The typical steganographic method utilizes digitized media files (images, video and audio files) as a cover medium for hiding data, network steganography uses communication protocols such as TCP/IP. Such methods make it harder to detect and eliminate. In a typical steganography using network the modification of a single network protocol occurs. Such modification can be applied to the Protocol Data Unit. Network steganography sheltersa broad spectrum of techniques.

General Terms

Steganography

Key words

Network Steganography, cryptography, TCP/IP, Covert channel, Steganalysis

1. INTRODUCTION

With the growth in information technology due to internet, machines using digital multimedia applications like digital cameras, videos have given immense opportunities for scientific and mercantile (commercial) purposes. Internet has been widely used for transferring, storing and retrieving of data. There is a need to protect this data by certifying or insuringtheir security, privacy and capacity. Otherwise, this may lead to many attacks like hacking and replications affecting the information confidentiality, integrity, its value, efficiency and its accessibility. [1]
The need arises in mostly all sectors like government, military, conglomerate and even in individuals to hide the information so that anyone cannot perceive its occurrence or existence. The information is made mysterious and this is capable with cryptography and Steganography.

Steganographyis a dualistic word which is a merger or amalgamation of SteganndGraphine, where in Greek stegano means to hide or covered and graphine means to write. It is a form of art and science that came in its existence in the fourteenth century and is still being implemented. It is a method in which the information or data is keep under wraps of cover medium secretly without being affected by any outsider [4][5]. Network steganography shelters a broad spectrum of techniques. [2][3]. The typical steganographic method utilizes digitized media files (images, video and audio files) as a cover medium for hiding data, network steganography uses communication protocols such as TCP/IP [6].

1.1 Steganalysis

The method to detect the hidden data in transmission is called as steganalysis. In packet length based Steganography, the length of the transmitted packet is being modified to hide the data and analysing various data packets, and it's possible to detect the data present in it. The detector has to study a large number of packets to detect the anomaly or hidden data.

1.2 Steganography vs. cryptography

They are used to protect the hiding files. In cryptography the data is visible but not in any meaningful form and only by knowing the cryptographic algorithm, the hidden data can be deciphered**.**In cryptography everybody knows that there is hidden information present but only the right algorithm can reveal i.e. in cryptography a message can be easily seen n recognised as cryptic message but only the one who has information as how the data is encrypted will come to know how to decrypt it.
Whereas the data in Steganography is written in plain text but is hidden in the cover medium such as that it's hard to detect and uses the non-prominent area of the text or image or video i.e. any medium which is being used as a Steganographic cover medium, it has various examples but a simplest can be a picture of a man has a pose in which his figure points upwards means he is happy and downward points mean he is sad. So only the sender and receiver know the code and can detect the hidden information. [10][13]

To hide the information within any media involve vital features like a cover medium or file which is required for hiding data , a secret data that need to be hidden and a key or code word that may be used by sender and receiver for encryption and decryption.
In short, steganography can be signified as:[4]
Secret data + cover medium = stegogramme
Stegogramme + stego key = stego-medium.

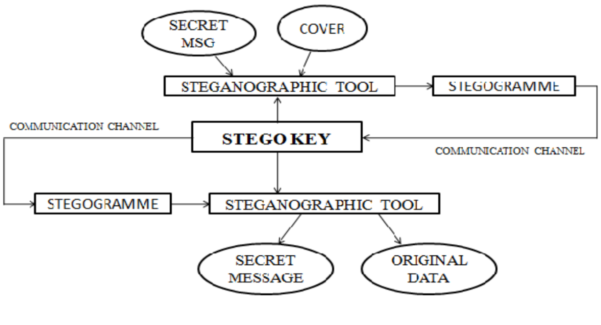

Fig. 1: Steganography concept

Steganography started in Greece, where wax tablets were used to hide the written content. Secret message was inscribed underneath the *timber wood* and then the wax was deposited over the message written on the timber, which made it look like an unused new wax tablet.

The resultant tablet of wax was safely conveyed to its destination. In Persia, reliable slave’s head was shaved and the secret information was inscribed on their head and with the regeneration of hair on slave’s head, Slave was sent to the destination where his head is again shaved to reveal the information.

One of the Nazi spies used to hide their information by writing on that on their handkerchief with a solution of copper sulphate and it remainsnon-existent until showed to ammonia fumes.
During World WarII, Nazis used technique known as “microdots”, where hidden message is so small that it appears like a dot at the end of paragraph so that nobody pays attention to it. Another method was “invisible ink” which was used for coverting the information and information bearing medium needs to be heated to reveal the information. Commonly used medium were milk, vinegar, fruit juices.

Now, this game of sending secret data has changed all together. Digital media and over all connected world has made this medium a very fast and secure means of communicating information like copyright the digital image by the means of watermarking, audio files, text files, videos, network protocols etc.[12]

UDP and TCP protocols can be used for packet length network steganography, but UDP protocol is more suitable for packet length steganography as the distribution of packet length is variable.[14][16]

1.3 TCP/IP Protocols

A key knowledge of TCP/IP protocol is required; it has the IP header and TCP header function.
TCP/IP has 2 protocols TCP and UDP. Both the protocol have alike primary functions but have different ways to connect primarily between hosts. TCP is a “connection oriented” and reliable whereas the UDP has “connectionless” and thus TCP has a feature that makes the data appear intact or in order at the receiver.[16]

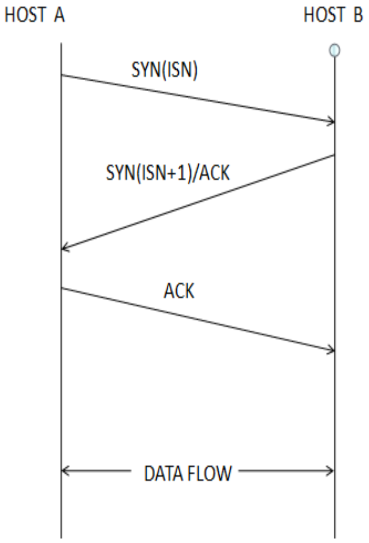

Fig. 2: Three way handshake

This is done by a procedure of “three-way handshake” followed by the TCP protocol. [14]

Step 1:
Host A directs a synchronize bit to proclaim a new connection and in ISN to track packets.
Step 2:
Respond of host B holds the sequence number of communication and initial sequence number +1, indicating it has acknowledge last packet and is ready to receive another.
Step 3:
Final acknowledgement is sent to the host B along with sequence number to indicate that connection is complete and data will flow now.

Now this entire succession occurred in just few milliseconds and each packet is acknowledged individually by both sides. Due to this handshake mechanism, TCP is a connection oriented and UDP is not as only TCP packets display such a connection process.

There are numerous methods obtainable in case of TCP/IP in order to set up or establish the covert channels where information can be furtively or secretly carried between hosts. These types of methods can be used in various fields like:

- Packet filter bypassing
- Network sniffers
- “dirty word” search engines

Number of fields in the header of TCP/IP is free and optional and can be worked as transmission of information covertly. For that the information in TCP/IP header is manipulated in such a way like encoding ASCII value of range 0-255.

Data needed to be hidden can be sent to a remote host secretly via TCP/IP, as it contains numerous areas where the information can be stored.

There may be many areas in TCP/IP header field where the data can be hidden but after long analysis it’s proposed that the data be hidden in any one of the mandatory field rather than the optional fields because there is any chance of filtering out these through packet filtering or fragment re-assembly.

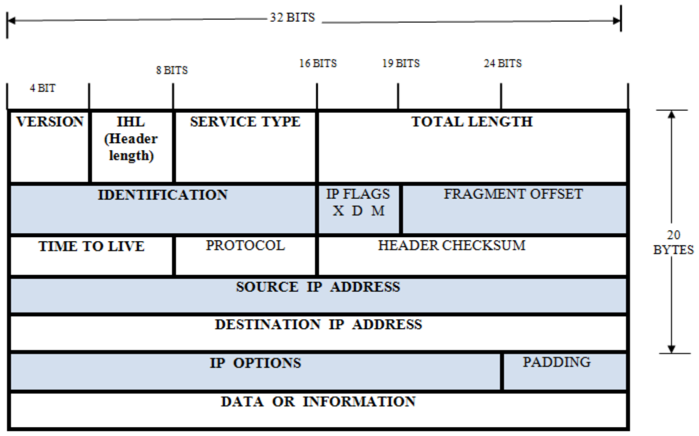

Fig. 3:IP Header:

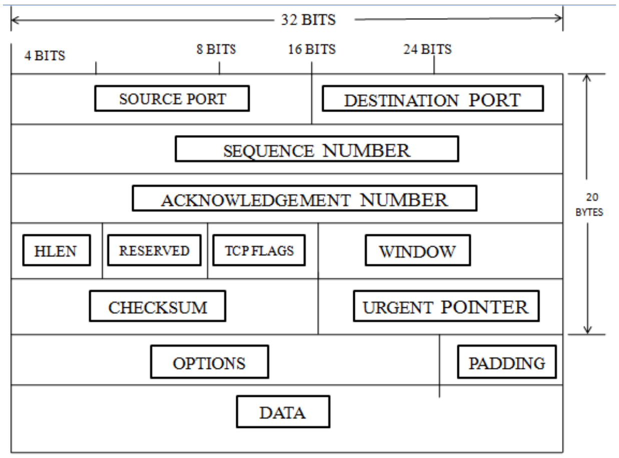

Fig. 4: TCP Header

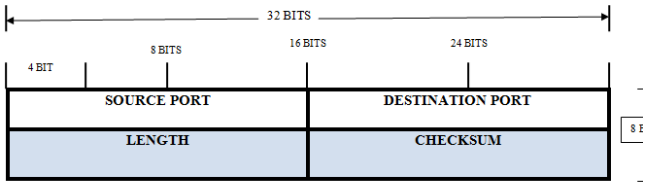

Fig. 5: UDP Header

2. METHODOLOGY

2.1 Modifying IP Packet Identification Field –

The IP packet identification field helps in the reassembly of the data packets at the remote host. It provides a unique value to the packet so if the packet is broken up in the route it's again re-assembled accurately to provide meaningful information. So in IP packet, identification field is modified by replacing the IP identification field by the numerical ASCII value of the characters that's to be encoded.
ASCII (American Standard Code for Information Interchange) is a 7 bit code with parity to make an 8-bit code representing all the characters.When the user sends ASCII value to remote host embedded in the identification field. Now the remote host simply reads the IP ID fields and convert the encoded value which can be a printable character.

Here the packet is sent with the ASCII code embedded in the packet ID field with the appropriate source and destination host info and the port information which the host is listening to.Now this method can be easily discarded and detected by firewall where the header information is written and again rewritten every time it passes a firewall or router and hence public network loss of data may occur.

2.2 Modifying - Initial Sequence Number Field (ISN)

The TCP/IP protocol suite establishesreliable protocol arbitration with a remote server.In the arbitration process of TCP/IP, several steps as in "three way handshake" asreferred to earlier.The sequence number field serves as a medium for transmitting data a 32 bit [7]. Itgenerate the sequence number from the actual ASCII character which it wish to encoded. A more "realistic" looking sequence number

2.3 Modifying - TCP Acknowledge Sequence Number "Bounce"

Bouncing Packet of TCP IP Field is also a well know technique for Information concealment this method uses spoofing of IP addresses for sending machine to "bounce" a packet [8] of a remote site and the packet from that site is returned to the real destination address. This helps in hiding the sender of information package as it appears to have come from a "bounce" host. It wouldbe useful to establish an anonymous single sided communication network whichcan be very difficult to detect if the bounce server is busy.

This method depends on the characteristic of TCP/IP in which the destination serverreplies to an initial connection request (SYN) with a SYN/ACK which has the incremented sequence number to the original ie (ISN+1) [8]. In this following method, sender constructs a packet that contains the below mentioned information which is forged:

- A fake SOURCE IP address.
- A fake SOURCE port.
- A fake DESTINATION IP address.
- A fake DESTINATION port.
- A TCP SYN number with encoded data.

The DESTINATION IP address is server you wish to BOUNCE information offand the SOURCE IP is the address of server you plan to communicate WITH [8].

A packet sent from client's computer system & routed through to the fake destination IP address in the headierof ("bounce server"). That bounce server receives the packet & sends either a SYN/ACK or SYN/RST depending on which state theport is forwhich the packet was destined. The return packet then is sent to forged source address with the ISN number +1. The listening destination server then decodes the hidden information by transforming the reverted sequence number minus one back into the ASCII Character. A step-by-step representation

- Sending Client: A
- BounceServer:B
- ReceivingServer:C

Step 1: Client A sends a forged packet & encoded information to bounce server B. This has the address of server which is Receivingie C.

Step 2: Bounce server B receives the packet and outlaysansuitable SYN/ACK or SYN/RST packet depending on port status. Since B thinks the packet came from C, the packet is back to receiving server C. An acknowledgment sequence number (encoded sequence number +1) is sent to server C.

Step 3: Server C, which is expecting to receive a packet from B (on a pre-determined port) decodes the data out to a disk. This method tricks the remote server to send the packet along with encapsulated data back to the receiver with a forged source IP address, which it thinks is legitimate. At the reciever , the packet appears to come from the bounced server, & it actually does. As a note, if this receiving system is behind a filter that allows communication to some of the sites, this method can bounce packets off of trusted sites which then relays them to a system behind packet filter with a authenticsource address. This can be used in communicating with receiving servers which are heavily scrutinized networks.[17]

Bouncing a packet from site (.mil, .gov, .com, etc.) is a worthwhile technique to hide operations in ordinary traffic. If the bounce site is using round-robin DNS (stable IP address) then the receiving server is just passively listening to a pre-determined port to decode (send out forged source address & source port as 1234 so the bounce server returns the packet on port 1234). By this technique, the sender can bounce packets off thousands of Internet hosts whereas the receiving server listens on pre-defined port number irrespective of IP address.
If any of the network sites is having a correctly configured router, it would not allow a forged packet, many routers are not configured by this protection in mind and will fortunately pass the data.[7]

3. APPLICATIONS

1) The major advantage is a method to transmit data which is hidden uses normal means of communication
2) Only sender and receiver knows about the information
3) Data may be easily hidden into what looks like a normal text and different form of encryption also do not make the hidden data obvious [7]
4) Can wrongly be used by extremists , and other to damage the system or leak information
5) Steganography can a solution which makes possible to send news and information without being censored
6) Steganography can be used to store information on a location.
7) Secretly transmit messages without the worry of anyone knowing there has been a transmission. For example, any picture of may conceals the plans for the company's latest innovation.
8) Governments and businesses are interested in two types of hidden communications: those that support national security and those that do not. Digital steganography provides vast potential for both types.
9) MMS is a technology that allows a user of a properly enabled mobile phone to create, send, receive and store messages that include text, images, audio and video clips. Steganography can be used in the context of MMS.
10) Various technologies are there to hide the data on network. It is hard to detect. It becomes a new and popular way to protect or transfer the information

4. RESULTS AND CONCLUSIONS

While running these methods there are few tools available on the net to modify the TCP/IP protocol for the proof of concept like NETCROSS, Covert.TCP,OpenPuff or Socat. Any of the three methods described can be used each has its advantage method one has ID field of same size as ASCII and is not removed by most of the filters, in SYN field the hidden data can be much more as regards to ID filed modification in IP. This is because it’s of 32 bit. Similarly in the third method where the data is bounced of the server and is received at the receiver, in this method the address of the sender is anonymous so anybody cannot detect who the sender is. [15]
TCP Dump of the port and TCP/IP packets shows the data one character at a time. And can be detected as the sequence no is not in order, but it can also be improved by a much more complex algorithm to scale the SYN field as such that it is incrementing and is different for every character at times. Still its detection is impossible to novice eyes and can only be detected with experts and if the future scope is implemented then it would be a tedious task to steganalyse the data sent from an unknown sender to an unknown receiver.

Now, steganography of the new era is growing with stupendously greater possibility for mischief. With the latest technological advancements, the limitation on the length of the Secret message has been removed. Consider example involving the use of Skype Where it requires a carrier which can be an MP3 song or a video—there was no such obligation for the transmission of a photograph. The data were concealed in the bits of a digital VOIP conversation. In this novel era of steganography, the scapegoat that co-conspirators are using is not the carrier but the whole communication protocol with anadvantagelonger the communicators talk, the longer can be the secret message which issent [1]. It makes the data nearly impossible to detect.

In Cryptography whenever we use a fixed length of group for bits we would be working on block cipher.
CAST -128 or (Cast 5) is a block cipher used in GPG which has been used here. This method has been approved by some of the governments like Canadian government. The major advantage is that its royalty free and can be used for commercial and non-commercial basis

What is being implemented is an encryption along with steganography on the TCP/IP header to double secure the data as an eavesdropper who could access the data easily (if experienced) can read the ASCII character. But now it’s not possible as catching only some data will not help in gaining the whole information. And if whole information is captures then too without the pass phrase it’s worth a penny.

5. ACKNOWLEDGEMENT

We would like to give special thanks to the Head of Department Prof SwarnAhuja for his guidance and inspiration.

6. REFERENCES

[1] “Hide and Seek: An Introduction to Steganography”, IEEE Computer Society, May/June 2003.
[2] ElizbietaZielinska and Krzysztof Szczypiorski, “Direct Sequence Spread Spectrum Steganographic Scheme for IEEE 802.15.4”, IEEE International Conference on Multimedia Information Networking and Security, 2011.
[3] Pukhraj Singh. “Whispers On The Wire: Network Based Covert Channels”, http://gray-world.net/papers/pukhrajsingh_covert.doc
[4] Michael T. Raggo , “Steganography”, DefCon 12, Aug. 2004
[5] Donovan Artz, “Digital Steganography: hiding data within data”, IEEE Internet Computing, May/June 2001, pp. 1089-7801
[6] “Steganography”, http://en.wikipedia.org/wiki/Steganography
[7] Craig H. Rowland, “Covert Channels in the TCP/IP Protocol Suite”, First Monday, Vol. 2, No. 5, May 1997, http://firstmonday.org/ojs/index.php/fm/article/view/528/449
[8] PrabhakarMateti, “TCP Exploits”, http://www.cs.wright.edu/people/faculty/pmateti/InternetSecurity/Lectures/TCPexploits/

[9] “Network Steganography and Anomaly Detection”, http://stegano.net/network-steganography.html

[10] Tariq Jamil, “Steganography, the art of hidden information in plain sight”, IEEE Potentials, Feb/March 1999, ISSN 0278-6648

[11] Arvindkumar and K.M. Pooja, “Steganography data hidden technique”, International Journal of Computer Application, Volume 9, November 2007.

[12] Neil F. Johnson, “Information Hiding: Steganography & Digital Watermarking”, http://www.jjtc.com/Steganography

[13] S. K. Pal, P. K. Saxena and S. K. Muttoo, “Image Steganography For Wireless Networks Using The Hadamard Transform”, International Conference on Signal Processing and Communications, 2004.

[14] Anand S Nair, Arijit Sur and Sukumar Nandi, “Detection of Packet Length Based Network Steganography”, IEEE International Conference on Multimedia Information Networking and Security, 2010

[15] WbStego Steganography Tool, http://wbstego.wbailer.com/, March 2004.

[16] Anand S Nair, Abhishek Kumar, Arijit Su and Sukumar Nandi, “Length Based Network Steganography using UDP Protocol”, IEEE, 2011.

[17] WojciechFrączek, WojciechMazurczyk and Krzysztof Szczypiorski, “How Hidden Can Be Even More Hidden?, IEEE International Conference on Multimedia Information Networking and Security, 2011.