A Proposed Firewall for Viruses and Steganography Image

Sura Mahmood Abdullah

Computer Science Department / University of Technology
suramahmood84@yahoo.com

Abstract: *The open communication network, Internet, has problems surrounding the security of the Internet sites. Such as hacker intrusion costing organizations a large amount of money and untold losses in productivity; hate groups using the Internet to distribute their malicious works to these sites, and many other types of attacks. A firewalls strategies protect Internet sites from intentional hostile intrusion that could compromise confidentiality or results in data corruption or denial of service.*

This research has two proposals they are:

- *Build a firewall against the unauthorized intruders and viruses which have specified signatures. Depending on the SYN/ACK checkup procedure and last install a virus scanner for the data in all packets of the session by the proposed procedure which is three handshaking proxy procedure.*
- *Build a secure web against the steganographic imags that is hidden in the authorized packets. This would be done by building a firewall, and then examining the suspected authorized packets if they contain steganographic images.*

Keywords: Firewall, Viruses and Steganography Image.

الخلاصة: شبكة الاتصالات المفتوحة، الانترنيت ، لهُ مشاكل تحيط بأمن المواقع على الانترنيت . مثل تسلل القراصنة الذي يكلف المنظمات كميـة كبيرة من المال و خسائر هـائلة في معدل الانتاج، مجموعات الكراهـية يستخدمون الانترنيت في توزيع اعمالهم الخبيثة في هذه المواقع. تقنيات جدار الحماية تحمي مواقع الانترنيت من التدخل العدائي المتعمد التي يمكن ان تضر السرية او تلف في نتائج البيانات او الحرمان من الخدمة.

هذا البحث له اقتراحان هما :-

- بناء جدار حماية ضد التدخل الغير مخول و الفايروسات.اعتمادا على فحص SYN / ACK و التثبيت الاخير لناسخ الفيروس الضوئي للبيانات في كل حزم الدورة.
- بناء شبكة امنة ضد صور إخفاء المعلومات المخبأة في الحزم المخولة لديها و سيتم تنفيذه من خلال بناء جدار حماية و اختبار المشكوك من الحزم المخولة فيما لو كانت تحتوي على صور إخفاء المعلومات.

1. Introduction

Data communications networks have become an infrastructure resource for businesses, corporations, government agencies, and academic institutions. However new technologies introduce new threats, and networking not only puts corporate resources, plans and data at risk, but ultimately the company's reputation and potential survival [1].

A firewall is critical part of the security of any Internet site. Basically, a firewall improves the security of a site by limiting the access of that site to an absolute minimum. It is important to know that although a firewall does not solve all Internet security problems, no Internet site should be without a firewall. Firewalls may give different levels of security and are considered to be of different types and different configurations [2].

2. Firewall technology

Firewall technology in TCP/IP internetworking provides a mechanism to help enforce access policies on communication traffic entering and leaving networks. Now we declare both firewall types and firewall configurations to give clear picture on the proposed work [3].

The common types of the firewalls according to the levels of TCP/IP and OSI stacks are:

a. Network Level Firewall (Packet Filtering Firewall):

A packet filtering is an access control mechanism for network traffic. Instead of processing or forwarding all packets that leave and arrive on the node's network adapters, the packet filters consults its access control rules before handling each packet [4,5]. Works at the network layer of TCP/IP stack and OSI stack are in the same principle [6]. A filter is a program that, in general examines the IP addresses (source and destination addresses), ports numbers, protocol type, and service type fields of every incoming specified access control mechanism [6,7].

b.* Application Level Firewall (Application Proxy)*:

These firewalls work a bit differently from packet filtering firewalls. Application gateway firewalls are software-based when a remote user from the void contacts a network running an application gateway, the gateway blocks the remote connection. Instead of passing the connection along, the gateway examines various fields in the request, if these meet a set of predefined rules, the gateway create a bridge between the remote host and the internal host (in common called proxy) [6,7].

c. Circuit Level Firewall (Circuit Proxy):

A circuit level gateway firewall is a generic proxy that does not know the specifies of the application but performs a more generic set of capabilities [4]. Circuit level gateways work at transport layer of TCP/IP stack and OSI stack in same principle. The circuit level firewalls monitor TCP three handshaking in the TCP connection (session) between packets to determine whether a requested session is legitimate [8].

d.* Stateful Multilayer Inspection Firewall*:

Stateful firewall combines the aspects of the other three types of firewalls. They filter packets at the network layer, determine whether session packets are legitimate and evaluate contents of packets at the application layer. So, this firewall examines all TCP/IP layers and OSI layers in same principle to either accept or reject the requested communication [8].

3. Steganography Theory

One of the newest hot spots in security research is information hiding. Information hiding (literally, covered writing) is the hiding of secret messages within another seemingly innocuous message, called (host signal), data or carrier. Embedding information, which is to be hidden, into media requires two files. The first is the innocent-looking image that will hold the hidden information, called the cover media. The second file is the message that the information to be hidden [9] as shown in Figure (1):

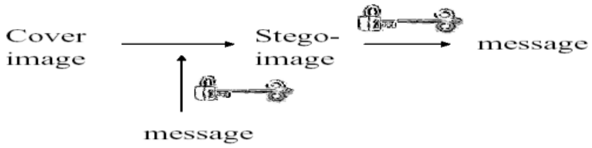


Figure 1: The Hidden Process in an Image

A message may be plaintext, cipher text, images, or anything that can be embedded in bit streams. The most media use is an image for hiding information. When combined, the cover media and the embedded message are a stego-object product. A stego-key (a type of password) may also be used to hide and then later decodes the message [10, 11].

Steganalysis

Hiding information within electronic media requires alterations of the media properties that may introduce some form of degradation or unusual characteristics. These characteristics may act as signatures that broadcast the existence of the embedded message, thus defeating the purpose of steganography. Attacks and analysis on hidden information may take several forms: detecting, extracting, and disabling or destroying hidden information [12].

In general, steganalysis is carried out for breaking the security of a steganographic system. It is assumed that an adversary has the knowledge of the system and has one or many images intercepted from a public channel. Security of the system lies solely on the secrecy of the key. To make steganalysis difficult, a steganographic system is designed to have a large unclustered key-space. In the case of steganographic investigation, it is quite likely that details of the system are unknown. Analysis of non-standard systems is a complex activity and requires *high level of expertise and massive infrastructure for computation of steganographic and cryptographic keys [13, 14].

Steganalysis consists of the following subtasks, one or more of which needs to be accomplished depending upon the nature of application: detection, extraction, removal/disabling/destruction of the message and meaningful modification/insertion. Detection is the first and the most important part of steganalysis. Extraction of secret messages is required for knowing the contents and also the purpose of the hidden communication. This is required for gathering legal evidence for booking an offender or a criminal. An active adversary may extract the message in an attempt to disable or replace it by another message Steganographic detection is carried out based on the knowledge of one or more of the following:

(a) Stego-media
(b) Host media
(c) Embedded data
(d) Steganographic tool

Depending upon the available information, different attacks are devised. The most practical case deals with the analysis of a single or a set of intercepted images (cover/stego) for which no other details are known. This is called a stego-only attack and is relatively difficult to handle. If a set of images is obtained from a single source it is useful to study them for peculiarities and also for similarities. For an isolated image, the job is to identify unique characteristics not found in normal images [10, 14].

4. Anti-virus Techniques

Without anti-virus software, there is no conclusive way to rule out viruses as the source of such problems and then arrive at solutions. Effective anti-virus software must be capable of performing three main tasks: Virus Detection, Virus Removal (File Cleaning) and Preventive Protection. Of course, detection is the primary task ad the anti-virus software industry has developed a number of different detection methods, as follows. Five Major Virus Detection Methods [15, 16]:

- **Integrity Checking (aka Checksumming)**: Based on determining, by comparison, whether virus-attacked code modified a program's file characteristics. As it is not dependent on virus signatures, this method does not require software updates at specific intervals. □***Limitations*** - Does require maintenance of a virus-free Checksum database; allows the possibility of registering infected files; Unable to detect passive and active stealth viruses; Cannot identify detected viruses by type or name.
- **Interrupt Monitoring**: Attempts to locate and prevent a virus "interrupt calls" (function requests through the system's interrupts). □***Limitations*** - Negative effect on system resource utilization; May flag "legal" system calls and therefore be obtrusive; Limited success facing the gamut of virus types and legal function calls.
- **Memory Detection**: Depends on recognition of a known virus' location and code while in memory; Generally successful. □***Limitations*** - As in Interrupt Monitoring, can impose impractical resource requirements; Can interfere with valid operations.
- **Signature Scanning**: Recognizes a virus' unique "signature," a pre-identified set of hexadecimal code, making it highly successful at virus identification. ***Limitations*** - Totally dependent on maintaining current signature files (as software updates from vendor) and scanning engine refinements; May make false positive detection in valid file.
- **Heuristic/Rules-based Scanning**: Faster than traditional scanners, method uses a set of rules to efficiently parse through files and quickly identify suspect code (aka Expert Systems, Neural Nets, etc.). ***Limitations*** - Can be obtrusive; May cause false alarms; Dependent on the currency of the rules set.

5. Firewall proposed to detect viruses

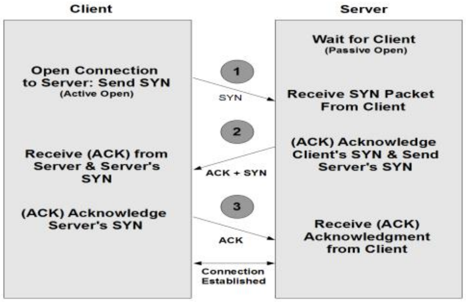

Client
Server
Wait for Client
(Passive Open)
Open Connection to Server: Send SYN
(Active Open)
1
SYN
Receive SYN Packet From Client
2
(ACK) Acknowledge Client's SYN & Send Server's SYN
Receive (ACK) from Server & Server's SYN
ACK + SYN
3
(ACK) Acknowledge Server's SYN
ACK
Receive (ACK) Acknowledgment from Client
Connection Established

Figure 2: Three Way Handshake

In normal case when packets pass the protected site to secure network these packets must pass through a Bastion host and these packets are submitted to a specific procedure.

We would propose a protection procedure for secure network is an approach for defending against all the viruses which have signatures. The three handshaking proxy counters the virus by collecting the data of all the packets of the session in a buffer by making sure that the three-way handshaking is actually completed between the secure authorized site and bastion host before sending a SYN packet to the secure network, destination of connection as shown in Figure (2). To declare this procedure let secure network be S, let Bastion host be B, let author site be A. and note the following points:

1. Suppose A request connection to secure server, at first A send SYN packet to B.

A → SYN → B

2. B receive the SYN packet but it does not pass the SYN packet to S but rather B send SYN/ACK to A directly.

A ◄——— SYN/ACK ◄——— B

3. If A send ACK packet to B then the connection is established between A and B.

A ———► ACK ———► B

4. Get the data of all the packets of that session and store it in a buffer. Then that buffer would be submitted to **Signature Scanning** - Recognizes a virus' unique "signature," a pre-identified set of hexadecimal code, making it highly successful at virus identification. to detect if the data has virus or not this depends if there is a signature of a virus in the data, then the session destroyed and assign that sessions as malicious one then put all the information related to it (such as source and destination IP addresses, ports, protocol type and statues) as unauthorized packets that would be as updating for firewall rules. But if there is no virus the session would be established with the secure network as in the following step.
5. Now B would make a connection establishment between B and S but it would use the SYN and ACK of A.

B ———► SYN ———► S

B◄——— SYN/ACK ◄——— S

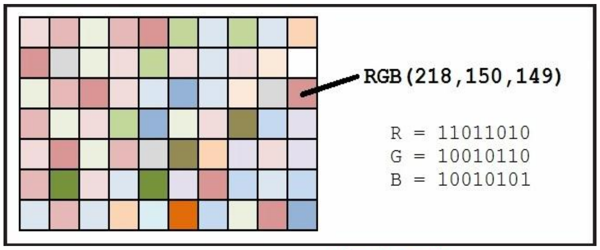
B ———► ACK ———► S

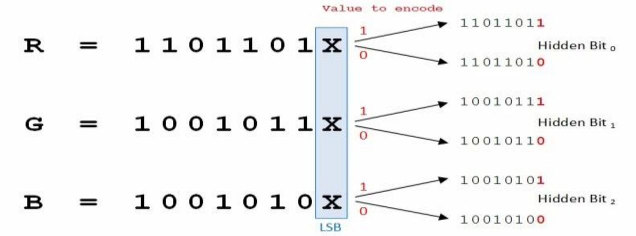
The best anti-virus software in the world cannot protect you if it is not deployed systematically throughout the enterprise (even if "the enterprise" is a single home based computer).

6. Firewall proposed to detect steganography

Two aspects of attacks messages on steganography these are the detection and destruction of the embedded messages. Any image can be manipulated with the intent of destroying some hidden information whether an embedded message exists or not. Detecting the existence of a hidden message will save time in the message elimination phase by processing only those images that contain hidden information. Detecting an embedded message also defeats the primary goal of steganography, that of concealing the very existence of a hidden message. In this proposed system we concentrate on detection the existence of stegographic, and then destroy this stegograhpic image immediately. Here we would use two techniques for detection the first one is the detection with the color images, and the second is the statistical detections when the original images are available in the web secure database:

LSB Encoding Detection in Color Images:





The ratio of number of unique colors to the number pixels is 1:2 for high color BMP images and 1:6 for the low quality JPEG images. Many true color images have a relatively small palette. The embedding process results in creating many very close colors in the image. The presence of too many pairs of colors in the image indicates the presence of hidden data in the image using LSB steganography. The presence of messages in true color images can be tested by using the following technique:

Let U be the number of unique colors in the image. Let P be the number of close color pairs in the image palette.

The colors (R1, G1, B1) and (R2, G2, B2) are selected as a pair if

$$\sqrt{|R1-R2|<=1,|G1-G2|<=1 \text{and} |B1-B2|<=1.}$$

This ratio R is the relative number of close pairs of colors in the image. After embedding the steganogram it will have a different number of close color pairs R. This significant change reflects the process of embedding. If the image already contains a hidden message and if again embedded with a message then the ratio R does not change largely. If the image is a clear image then the ratio increases significantly. So using the relative comparison of R as the decision parameter will increase the efficiency of the detection framework. If R′ > R then the probability of the image containing a hidden message is greater. The ratio R′ is calculated after embedding the input image with a hidden message proportionate to the number of pixels in the image using LSB encoding technique. R′ is actually calculated from U′ and P′ of the embedded image.

Statistical Detection :

As explained previously, the statistical analysis is very useful for detection of embedded data in an image. These analysis or tests can expose abnormalities in an image that are not visible by the human eyes. The statistical tests need the original and the suspected images for getting the correct results. The following statistical tests are used in the proposed system:

a. Average Absolute Difference (AD) Test

The Average of the difference between color of pixels in original image and suspected image can be calculated as:

$$AD = \frac{1}{XY}\sum_{x,y}\left|p_{x,y} - \tilde{p}_{x,y}\right| \qquad 1$$

where

X,Y: The height and width of images.

x, y: The pixel coordinate in x-axis and y-axis.

$P_{X,Y}$: The values of pixel color in location x and y of the original image.

$\tilde{P}_{X,Y}$: The values of pixel color in location x and y of the suspected image.

The absolute value used for difference to get accurate results of summation difference out.

Note: These values definition will be used for below statistical tests.

b. Mean Squared Error (MSE) Test

To calculate the Mean Squared Error between the original image and suspected image, we must know the difference of pixel color in the two images. The result will be the square amount of errors depending on the size of these images.

The equation that used to calculate (MSE):

$$MSE = \frac{1}{XY}\sum_{x,y}\left(p_{x,y} - \tilde{p}_{x,y}\right)^2 \qquad 2$$

Distortion and Removal:

The objective of distortion is to manipulate the carrier to the degree that the embedded message is distorted beyond recognition. The objective of removal is to prevent a hidden message from being transmitted at all. Maintaining some integrity of the carrier while removing the embedded message or, at least, distorting the embedded message beyond recovery is desirable in this form of attack. Simply destroying the container would also destroy the message, however the carrier may be a valid message in itself, and some form of its integrity should be retained. The basic assumption is that a carrier (C) is composed of two components. One (A) is the parts of the carrier that are observable by the human perceptual system. The second (T) is the portion of the carrier that is “invisible” to the human perceptual system (the amount of information that falls below the threshold of perceptibility). If the operation a + b is to denote some composition of a and b, then the carrier can be represented as C = A + T, where T is the amount of the carrier that can be manipulated without observable distortion. The size of T depends upon the structural properties of the carrier. A steganographer treats the amount of information in T as noise in the carrier and takes advantage of it to embed additional information that is below the perceptual threshold. However, an attacker can also manipulate T without noticeable distortion and manipulate, overwrite, or remove any embedded message. A hidden message can be corrupted through distorting the carrier and overwriting or deleting the embedded message. Since many types of embedded messages are below the threshold of the human perceptual system, slight modifications to the carrier may result in an unrecoverable embedded message but maintain a “suitable” carrier.

Conclusions

- The SYN/ACK check up procedure represent the basic procedure in the proposed system, because it protects both the secure server and bastion host from IP Spoofing (IP Impersonation) Attack. This by checking up the SYN/ACK to all the packets enter and leave the protected site along the session (not only in the three handshaking).

- The firewall responsibilities consider the examination of the packets headers only to detect if these packets are authorized or not. So we need another protection method to detect if these authorized packets are stegographic media or not if it is then the good opinions are to destroy these images or remove the hidden messages.

References

[1] Lyles .J .B., Schuba .C .L.,"**A Reference Model for Firewall Technology and it's Implication for Connection Signaling**", Proceeding in Open Signaling Workshop, Columbia University, New York, NY, October 1996.

[2] Goncalves .M., "**Firewalls Complete**", the McGraw-Hill Companies, Inc.1997.

[3] Breedlove .B., et al, "**Web Programming Onleash**", Sams.Net., 1996.

[4] Escamilla T., "**Intrusion Detection: Network Security Beyond the Firewall**", Wiley Computer Publishing, Sons, Inc., 1998.

[5] Comer D E., "**Internetworking with TCP/IP Vol 1: Principles, Protocols, and Architecture**", Third Edition , Prentice-Hall, Inc., 2000.

[6] Goncalves M., Brown S A., "**Check-Point Firewall -1 Administration Guide**", Mc-Graw Hill Companies, Inc., 2000.

[7] Zdnet Research Center Business and Technology, White Papers, "**Intrusion Detection: Deploying the Shomiti Century Tap**", METAS, ZCInc. ZDNet, 2001.

[8] "**Vicom soft Knowledge Share Firewall: Q & A**", Vicom Technology Ltd, 2001.

[9] Amirtharajan R., Akila R.and Deepikachowdavarapu P., "**A Comparative Analysis of Image Steganography**", International Journal of Computer Applications, Vol 2, No.3, pp 41-47, May 2010.

[10] Stalling W., "**Network Security Essentials: Application and Standards**", Prentice-Hall, Fourth Edition, 2000.

[11] Ahsan K., Kundur D. "**Practical Data Hiding in TCP/IP Packets**", Proceeding in ACM workshop on multimedia and security, 2003.

[12] Li B., He H., Huange J., Shi Y., "**A Survey on Image Steganography and Steganalysis**", Journal of Information Hiding and Multimedia Signal Processing, Vol 2, No 2 , pp 142-172, April 2011

[13] Bloom J., Alonso R., "**Smart Search Steganalysis**", Proceeding in SPIE 5020, Security and Watermarking of Multimedia Contents V, 167, Vol 5020, June 2003.

[14] Fisk G., Fisk M., Papadopoulos C. and Neil J., "**Eliminating Steganography in Internet Traffic with Active Wardens**", Proceeding in 5th. International Workshop, pp 18-35, 2004.

[15] Killers B., "**A New Generation of Antivirus Software Offer Maximum Protection**", Pc Magazine , February, Vol. 4, Issue 2, 2004.

[16] Wolder B., "**Internet/Intranet Security**", NSS Group, 2004.