International Journal of Applied Engineering Research ISSN 0973-4562 Volume 12, Number 18 (2017) pp. 8004-8008
© Research India Publications. http://www.ripublication.com

Steganography Using Unforeseen Media

Prateek Gera[1], G. Sujatha[2], K.Sornalakshmi[3], D.Hemavathi[4] and M.Kavitha[5]

[1]M.Tech(Information Security and Cyber Forensics), SRM University, Chennai, India.
[2,3,4]Assistant Professor (Sr.G), Department of Information Technology, SRM University, Chennai, India.
[5]Assistant Professor (O.G), Department of Information Technology, SRM University, Chennai, India.

[1]Orcid ID: 0000-0002-6531-2684

Abstract

In Steganography, the secret message is hidden into a cover media such as text, audio, video or image due to which attacker would have no indication about the original message that the media contains and which algorithm is used to embed or extract it. This paper will present different types of Medias for sending the data unlike existing media like audio, video and image. The proposed system will try to explore different medias for hiding our data like viruses, unformatted file, multiple compression, executable etc. Before Steganography technique could be applied, AES algorithm will change the secret message into cipher text to ensure the two layer security of message.

Lots of tools are already available and are in use to hide data behind medias like audio, video and image and people are trying to explore new algorithms to hide their data like LSB, DCT etc. but this paper will focus on exploring new media instead of changing algorithms. The reason streams like audio, video and images are being used for such purposes are the availability of redundant data in them, which could be replaced with our private data.

Keywords: Cryptography, Steganography, Multiple Layered Compression, Cover Media, Viruses, AES

INTRODUCTION

Cryptology is a practice and the study of techniques for secure communication in the presence of third parties. It is method of storing and transmitting data in a particular form so that only intended person can read and process it, It is also associated with the method of converting plain text to cipher text(non-readable) and vice-versa.

Proposed System will use AES algorithm to encrypt and decrypt the data.

Steganography refers to information or a file that has been concealed inside a digital picture, video or audio file. If a person views the object he or she will have no indication about any hidden information. Basic methods used for hiding data are Text Steganography, Audio/Video Steganography and Image Steganography in which Image Steganography is the method of hiding text behind image[4], similarly Audio/Video Steganography is to hide the text behind audio and videos plus if we hide our text in an encrypted form like shown in (Figure 1)[4] it would be a nightmare for an eavesdropper to detect and decrypt hidden data **but hiding text in this way is very common and numerous tools are also available to hide the text automatically**. Thus, proposed system will try to explore different media to hide the data.

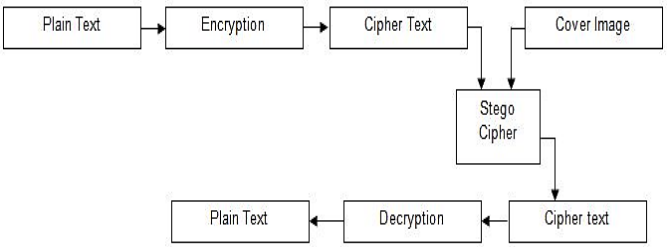

Figure 1: Combination of cryptography and steganography

A. Cryptography versus Steganography

Cryptography is a method which is being used from a long time to secure the data but the problem with cryptography is that even sending cipher data gains attention of a person, so concept of adding both techniques i.e. Cryptography and steganography came like shown in (Figure 1) in which data is encrypted then hidden behind media streams like Audio, Video or image but this technique is being used from several years. So person who has Knowledge of steganography would be suspicious on irregular receipt of image, audio or video.

BACKGROUND

A. Module 1: Hiding data in virus

One of the proposed ways of data hiding is hiding secret data beneath viruses. This technique could provide good level of transparency from an eavesdropper because if somehow he

manages to sniff our communication channel, he would only see that the victim is sending viruses to another person, there would be least possibility to imagine that virus could contain useful data.

Number of viruses could be used to implement this idea, but here we will define only 2 ways to send our data. First way could be sending our data in a virus with dangerous payload like deletion of important OS files or use of self-exploding viruses like "Rombertik virus" which erases the partition sector of the hard drive and forces a machine restart when it senses any detection mechanism initiated by user, so if an eavesdropper tries to analyze our virus it would in turn cause damage to his valuable data and lead him to payback quite a good amount but problem with this technique is it could also be harmful for our receiver

Second technique could be use of simple and small payloads like opening calc and making temp folders etc. but analysis would also be easy if eavesdropper gets suspicious, so for solving this issue proposed system works on an idea of sending user defined number of viruses out of which only one virus would have hidden data.

Other objective of proposed system would be usage of original existing virus signatures so that the eavesdropper's antivirus or any antivirus will detect it as a virus and warn user, thus with the help of this technique we will decrease suspicion level of secret data by eavesdropper and in some cases antivirus would help us in deleting it, but receiver should take care of his antivirus also.

Now proposed system will work on second technique i.e. usage of simple payload but large number of viruses. Our objective is to increase time and complexity for eavesdropper as much as possible so sending large number of viruses with occurrence of data only in one virus will serve our objective.

Although payload of virus is simple in proposed system i.e. like opening a simple calc, but payload can be changed accordingly. Method of hiding cannot be as good as hiding behind an image or audio as those media contain number of redundant data that could be altered, but in virus every data is required for its execution so data could be hidden only at the end of file but this limitation can easily be removed using many methods, some of the methods used by proposed system are:

- Using Padding
- Using large amount of viruses with false data

In proposed system user could generate a number of viruses with only one virus containing original data{ciphered} and other's with false/random data{ciphered}, padding will also be possible in proposed system.

So if eavesdropper gets suspicious also, he needs to work hard to find appropriate virus with data, even if he is successful in that then he also needs to distinguish between encrypted data and padded data which is to be decrypted at the end to access the data.

B. Module 2: Hiding data in raw/unsupported file

Normally all the system recognizes files from their extensions but if a file does not have an extension, its header and footer are used to recognize file type, this technique is also known as file carving. Scalpel is an existing tool for carving files.

Proposed system is based on the fact that a file with no extension and multiple headers and footers as shown in (Figure 2) cannot be recognized by normal system or file carver like scalpel as it will also give lots of false positive and this type of file could be useful for us to hide our data.

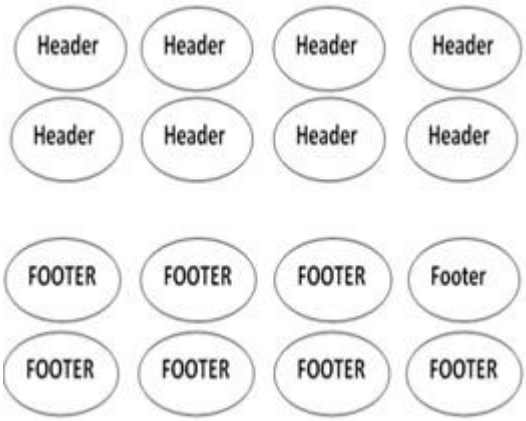

Figure 2: File with only headers and footers

C. Module 3:Hiding data in archive file with multiple layers

Multiple layers are referred as zip file consisting of number of zip file which again further consist of number of zip files and so – on, this technique was earlier used for hiding/sending virus, like every program antivirus also has a memory buffer, when antivirus scans a zipped file, it extracts the zipped file in its memory buffer and after some time memory buffer of antivirus gets flooded which results in antivirus crash. Now in proposed system we are going to use same technique to hide our data instead of virus or even virus with data in it.

In proposed system data will be hidden in 'n^{th}' file of 'x^{th}' level (no of time a file is archived) of archive files. Proposed system is based on the principle that one of the way of saving a data from an intruder is to give him large amount of data which will increase his work drastically, so we will give a zip file with multiple zip file inside it followed by multiple zip files and so on from which one zip file of a particular level will contain our data behind it. Intruder will have to analyze each zip file at each level to find our data which will also be encrypted.

D. Module 4:Chaining of all discussed modules

Although there could be number of ways to hide our data as discussed, but the proposed system will try to implement and test above methods. Moreover proposed system will have option of chaining each module to provide maximum transparency to user's data.

File containing data will be made unsupported/unformatted with large number of headers and footers so that normal program would not read it then that file would be hidden into the bundle of viruses, these viruses will be hidden behind a compressed file which will be inside one of the multiple compressed files at a certain level of compression, thus this will provide a high level of protection, although all these module could also be used separately as per user's requirement.

Even If an intruder manages to capture the message also, he would have to find viruses from multiple level of compression, if somehow he finds viruses, then he has to find out the virus containing original data, if he manages to find original data also he will have to decrypt the cipher text. Objective of Proposed system is not to make a full proof system but to secure file for a time period after which data may have no value even if found.

IMPLEMENTATION

A. Architecture

Proposed system will work with combination of both steganography and cryptography as shown in (figure 3)

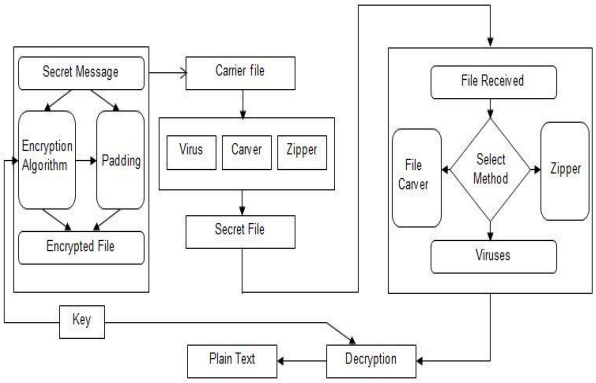

Figure 3: Architecture

Secret message will be encrypted or padded or both according to the user, Encryption algorithm used is AES-Advanced Encryption Standard (AES) is a symmetric-key cryptographic algorithm which was published in 1977. AES allows for three different key lengths: 128, 192, or 256 bits and 128 bits block size. It supersedes the Data Encryption Standard (DES) which is also a symmetric-key cryptographic algorithm [5]. This encrypted file is appended with the selected carrier {i.e. virus, zipper, raw file (carver)} and sent to the receiver.

At receiver's end receiver will open the file in proposed system, choose carrier method with decryption key and retrieve the file.

B. *Module 1 - Hiding data in virus*

Normal text will be converted into cipher text using AES Algorithm using a key taken by user. This cipher text will be inserted in one virus out of 'n' viruses using simple algorithm like dictionary based on user key

Features:

- User defined amount of files
- User can defined filename and extensions
- Automated option will also exist
- Optional feature of random padding of data
- Signatures to save it from Antivirus

Implementation is based on clamav db, system will extract random signature of virus from clamav db and embed it with hex of a file as shown in (Figure 4), and user can also use their own database of signatures.

After embedding, each file will be detected as virus from anti-virus but each file will have pre-defined payload {like executable will open calculator},User will have the option of selecting extension of a file too. System will create file from their hex so that output file will act like usual for example pdf will open as a pdf, vbs as a vbs.

Virus would have header and footer which will serve as markers while adding and separating it from data

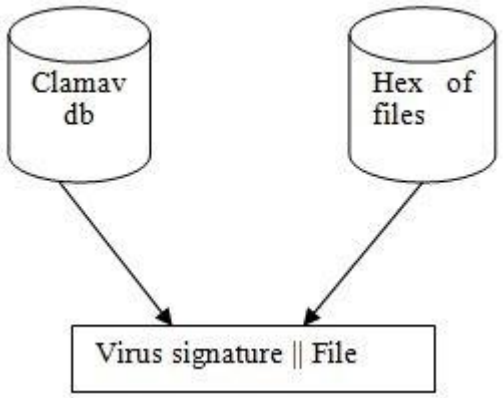

Figure 4: Virus Generation

User will send 'n' number of viruses out of which only one would have original data {Encrypted}, Key from user will be used to select that virus, and similarly at receiver's end same key will be used to identify original data containing virus. Data will be retrieved by removing virus with the help of its header and footer and then decrypted from the provided key (Figure 5)

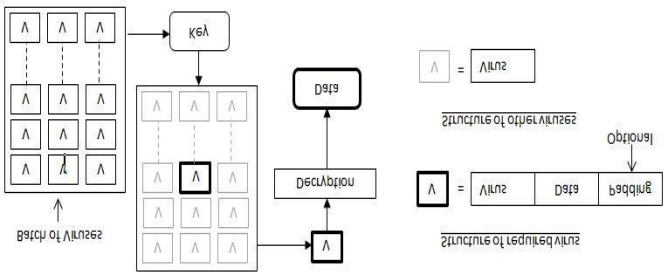

Figure 5: Virus Selection

C. Module 2 - Hiding data in raw/unsupported

As discussed, raw file is a file that is non-readable by a system, file with no extension. For the proposed system, file with multiple headers and footers is required, that type of file could be obtained using tool named as 'Scalpel'. User will provide 3 arguments to hide his data behind this type of media

Arguments

- Header
- Data
- Footer

At receiver side, receiver will provide same header and footer as sender to the proposed system which will retrieve the required data, This technique could be broken by hit-n-trail but using this module with other modules will increase confidentiality of data

D. Module 3 - Hiding data in archive file with multiple layers

Archive file with multiple layers are generally used to crash the program or system reading it, usually it allows the program to work as intended but the archive is carefully crafted so that unpacking it requires unordinary amount of time, space or memory[1].Basic compression usually exploits statistical redundancy E.g. 1110000101 could be written as 3140101,similarly 000000000 could be written as 90.similarly if we compress file containing 0's only of size 256 TB it will get compressed in few MB's.

Algorithm:

- Step1: Create a Gig or Petabyte .txt file full of '0'
- Step2: Compress it.
- Step3: Rename the .zip to .txt then make 16 copies
- Step4: Compress all of it into a .zip file,
- Step5: Rename the renamed .txt files inside the .zip file into .zip again
- Step6: Repeat steps 3 to 5 eight times.

This is the algorithm of basic logic bomb or archive bomb. In proposed system we will try to change functionality of normal archive software to compress it at a greater level. After creating 'n' layers of archive file, data will be hidden behind 'yth' file of 'xth' layer which will be calculated by key of a user

E. Module 4 - Chaining of all discussed modules

Proposed system will have the option of using all the modules together. Secret Data is first encrypted using AES algorithm, encrypted data is saved in a raw file with a header-footer combination known by user only, This raw file is embedded with a virus which is made hidden inside a file in multiple layer archive file as shown in (Figure 6)

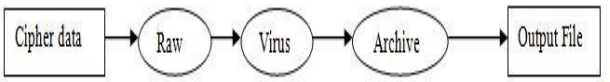

Figure 6: Full Execution

F. Merits of Proposed System

- New methods of sending text
- Provides multiple level of security
- AES + STEGO gives 2 level of primary protection
- Virus automation will help user generates number of signature enabled viruses in just a click
- New way of compressing just with a click
- Using file with multiple extensions for our work

G. Limitation of Proposed System

One of the limitation of this system is text/data will always be embed at last of the media unlike media streams like image or audio. Executable media doesn't have redundant data in it which could be replaced because changing in a single bit of an exe will corrupt the exe, but this issue can be resolved with the option of padding and sending large numbers of files.

CONCLUSION

Conclusion of this project is steganography could be used with several other media's not just basic image, video or audio. Main aim of this project was to explore different media's that could complete our work. Proposed system will be able to integrate secret data within media's like viruses, raw file and multiple layer archive files as data within media's like these create less suspicion over other forms of data{audio, video or image}

REFERENCE

A. Websites:

[1] Xeus, "Zip Bomb", http://xeushack.com/zip-bomb/, May 7, 2014

[2] RohitShane, http://resources.infosecinstitute.com/file-carving/, OCTOBER 4, 2013,"File Carving"

[3] Joel Esler, Douglas Goddard, Nigel Houghton "Creating signatures for ClamAV", http://www.clamav.net/about.html

B. Conference Paper:

[4] An analysis of LSB Based Image Steganography Techniques by K.Thangadurai and G.Sudha Devi,PG and Research Department of Computer Science,Govt., Arts College (Autonomous), Karur, India at 2014 International Conference on Computer Communication and Informatics (ICCCI -2014), Jan. 03 – 05, 2014, Coimbatore, INDIA.

[5] An Efficient Filtering Based Approach Improving LSB Image Steganography using Status Bit along with AES Cryptography by Md. Rashedul Islam1, Ayasha Siddiqa2, Md. Palash Uddin3,Ashis Kumar Mandal4 and Md. Delowar Hossain5 Faculty of Computer Science and Engineering, Hajee Mohammad Danesh Science and Technology University (HSTU),

[6] Dinajpur-5200, Bangladesh at 3rd International Conference On Informatics, Electronics & Vision.