International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 1 Issue 10, December- 2012

Steganography and Steganalysis: Different Approaches for Information Hiding

Mr Chintan Mahant
Student of M.E, CSE Department, PIT Limda,Gujarat, India

Asst. Prof. Yask Patel
Assistant Professor, Computer Science & Engineering Department, PIET Limda, Gujarat, India

Abstract

Video Steganography is a technique to hide any kind of files in any extension into a carrying Video file. This project is the application developed to embed any kind of data (File) in another file, which is called carrier file. The carrier file must be a video file. It is concerned with embedding information in an innocuous cover media in a secure and robust manner. This system makes the Files more secure by using the concepts Steganography and Cryptography. Steganography is the technique of hiding confidential information within any media. Steganography is often confused with cryptography because the two are similar in the way that they both are used to protect confidential information. The difference between the two is in the appearance in the processed output; the output of steganography operation is not apparently visible but in cryptography the output is scrambled so that it can draw attention. Steganlysis is process to detect of presence of steganography. In this article we have tried to elucidate the different approaches towards implementation of steganography using 'multimedia' file (text, static image, audio and video) and Network IP datagram as cover. Also some methods of steganalysis will be discussed.

1. Introduction

Steganography is the art and science of writing hidden messages in such a way that no one, apart from the sender and intended recipient, suspects the existence of the message, a form of security through obscurity. The word steganography is of Greek origin and means "concealed writing" from the Greek words steganos meaning "covered or protected", and graphei meaning "writing". The first recorded use of the term was in 1499 by Johannes Trithemius in his Steganographia, a treatise on cryptography and steganography disguised as a book on magic. Generally, messages will appear to be something else: images, articles, shopping lists, or some other cover text and, classically, the hidden message may be in invisible ink between the visible lines of a private letter.

The advantage of steganography over cryptography alone is that messages do not attract attention to themselves. Plainly visible encrypted messages—no matter how unbreakable—will arouse suspicion, and may in themselves be incriminating in countries where encryption is illegal Therefore, Whereas cryptography protects the contents of a message, steganography can be said to protect both messages and communicating parties.

Steganography includes the concealment of information within computer files. In digital steganography, electronic communications may include steganographic coding inside of a transport layer, such as a document file, image file, program or protocol. Media files are ideal for steganographic transmission because of their large size. As a simple example, a sender might start with an innocuous image file and adjust the color of every 100th pixel to correspond to a letter in the alphabet, a change so subtle that someone not specifically looking for it is unlikely to notice it.

1.1 History of steganography

It is believed that steganography was first practiced during the Golden Age in Greece. An

IJERT

ancient Greek record describes the practice of melting wax off wax tablets used for writing messages and then inscribing a message in the underlying wood. The wax was then reapplied to the wood, giving the appearance of a new, unused tablet. The resulting tablets could be innocently transported without anyone suspecting the presence of a message beneath the wax. An ancient Greek record describes the practice of melting wax off wax tablets used for writing messages and then inscribing a message in the underlying wood. The wax was then reapplied to the wood, International Journal of Computer Applications (0975 – 8887) Volume 9– No.7, November 2010 20 giving the appearance of a new, unused tablet. The resulting tablets could be innocently transported without anyone suspecting the presence of a message beneath the wax. Later on Germans developed microdot technology which FBI Director J. Edgar Hoover referred to as "the enemy's masterpiece of espionage. Microdots are photographs the size of a printed period having the clarity of standard-sized typewritten pages. The first microdots were discovered masquerading as a period on a typed envelope carried by a German agent in 1941. The message was not hidden, nor encrypted. It was just so small as to not draw attention to itself. Besides being so small, microdots permitted the transmission of large amounts of data including drawings and photographs. Another common form of invisible writing is through the use of Invisible inks. Such inks were used with much success as recently as WW-II. An innocent letter may contain a very different message written between the lines. Early in WW-II steganographic technology consisted almost exclusively of invisible inks. Common sources for invisible inks are milk, vinegar, fruit juices and urine. All of these darken when heated.

2. Requirements of hiding nformation digitally

There are many different protocols and embedding techniques that enable us to hide data in a given object. However, all of the protocols and techniques must satisfy a number of requirements so that steganography can be applied correctly.

A. Requirements that Steganography Techniques Must Satisfy

- The integrity of the hidden information after it has been embedded inside the stego object must be correct.
- The stego object must remain unchanged or almost unchanged to the naked eye.
- In watermarking, changes in the stego object must have no effect on the watermark.
- Finally, we always assume that the attacker knows that there is hidden information inside the stego object.
- Robust: Robust marking aims to embed information into a file which cannot easily be destroyed
- Fragile: This steganography involves embedding information into a file which is destroyed if the file is modified.

IJERT

Sr. No	Stegano-graphy Techniques	Cover Media	Embedding Technique	Advantages
1.	Binary File Technique	Binary File	watermark can be embedded by making changes to the binary code that does not affect the execution of the file	Simple to implement
2.	Text Technique	Document	To embed information inside a document we can simply alter some of its characteristic	Alterations not visible to the human eye

3.	Image Hiding: 1) LSB (Least Significant Bit)	Image	It works by using the least significant bits of each pixel in one image to hide the most significa	Simple & easiest way of hiding information.
	2)DCT (Direct Cosine Transform)		Embeds the information by altering the transformed DCT coefficients.	Hidden data can be distributed more evenly over the whole image in such a way as to make it more robust.
	3) Wavelet Transform		This technique works by taking many wavelets to encode a whole image	Coefficients of the wavelets are altered with the noise within tolerable levels
4	4) Sound Technique	MP3 files	Encode data as a binary sequence which sounds like noise but which can be recognised by a receiver with the correct key	Used for watermarking by matching the narrow bandwidth of the embedded data to the large bandwidth of the medium

Table 1 shows comparison of various methods of stegnography

3. Steganography and cryptography

A. Comparison of Steganography and Cryptography:

Steganography and cryptography are closely related. Cryptography scrambles messages so it can't be understood. Steganography on the other hand, hide the message so there is no knowledge of the existence of the message. With cryptography, comparison is made between portions of the plaintext and portions of the cipher text. In steganography, comparisons may be made between the cover-media, the stego-media, and possible portions of the message. The end result in cryptography is the cipher text, while the end result in steganography is the stego-media. The message in steganography may or may not be encrypted. If it is encrypted, then a cryptanalysis technique is applied to extract the message.

B. Combination of Steganography and Cryptography

Those who seek the ultimate in private communication can combine encryption and steganography. Encrypted data is more difficult to differentiate from naturally occurring phenomena than plain text is in the carrier medium. There are several tools by which we can encrypt data before hiding it in the chosen medium. In some situations, sending an encrypted message will across suspicion while an invisible message will not do so. Both methods can be combined to produce better protection of the message. In case, when the steganography fails and the message can be detected, it is still of no use as it is encrypted using cryptography techniques.

4. Steganalysis

Steganalysis is "the process of detecting steganography by looking at variances between bit patterns and unusually large file sizes". It is the art of discovering and rendering useless covert messages. The goal of steganalysis is to identify suspected information streams, determine whether or not they have hidden messages encoded into them, and, if possible, recover the hidden information.Unlike cryptanalysis, where it is evident that intercepted encrypted data contains a message.

Steganalysis generally starts with several suspect information streams but uncertainty whether any of these contain hidden message. The steganalyst starts by reducing the set of suspect information streams to a subset of most likely altered information streams. This is usually done with statistical analysis using advanced statistics techniques.

IJERT

5. Types of Steganography

A. AUDIO STEGANOGRAPHY

Steganographic algorithms can be characterized by a number of defining properties. Three of them, which are most important for audio steganographic algorithms, are defined below.

Transparency evaluates the audible distortion due to signal modifications like message embedding or attacking. In most of the applications, the steganography algorithm has to insert additional data without affecting the perceptual quality of the audio host signal. The fidelity of the steganography algorithm is usually defined as a perceptual similarity between the original and stego audio sequence. However, the quality of the stego audio is usually degraded, either intentionally by an adversary or unintentionally in the transmission process, before a person perceives it. In that case, it is more adequate to define the fidelity of a steganography algorithm as a perceptual similarity between the stego audio and the original host audio at the point at which they are presented to a consumer.

In order to meet fidelity constraint of the embedded information, the perceptual distortion introduced due to embedding should be below the masking threshold estimated based on the HAS/HVS and the host media.

Capacity of an information hiding scheme refers to the amount of information that a data hiding scheme can successfully embed without introducing perceptual distortion in the marked media. In the case of audio, it evaluates the amount of possible embedding information into the audio signal. The embedding capacity is the all included embedding capacity (not the payload) and can be measured in percent (%), bits per second or frame and bits per mega byte or kilo byte audio signal. In the other words, the bit rate of the message is the number of the embedded bits within a unit of time and is usually given in bits per second (bps). Some audio steganography applications, such as copy control, require the insertion of a serial number or author ID, with the average bit rate of up to 0.5 bps. For a broadcast monitoring watermark, the bit rate is higher, caused by the necessity of the embedding of an ID signature of a commercial within the first second at the start of the broadcast clip, with an average bit rate up to 15 bps. In some envisioned applications, e.g. hiding speech in audio or compressed audio stream in audio, algorithms have to be able to embed message with the bit rate that is a significant fraction of the host audio bit rate, up to 150 kbps.

- ***Reconstruction***

The last step is new audio file (stego file) creation. This is done sample by sample. There are two states at the input of this step. Either modified sample is input or the original sample that is the same with host audio file. It is why we can claim the algorithm does not alter all samples or predictable samples. That means whether which sample will be used and modified is depending on the status of samples (Environment) and the decision of intelligent algorithm.

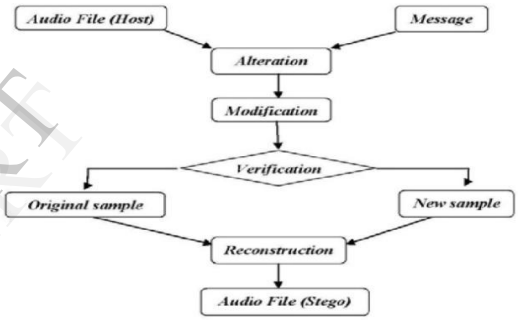

Figure . Audio Stegnography

B. LEAST SIGNIFICANT BIT TECHNIQUES

The most widely used technique to hide data, is the usage of the LSB. Although there are several disadvantages to this approach, the relative easiness to implement it, makes it a popular method. To hide a secret message inside an image, a proper cover image is needed. Because this method uses bits of each pixel in the image, it is necessary to use a lossless compression format, otherwise the hidden information will get lost in the transformations of a lossy compression algorithm.

When using a 24 bit color image, a bit of each of the red, green and blue color components can be used, so a total of 3 bits can be stored in each pixel. Thus, 800 × 600 pixel image can contain a total amount of 1.440.000 bits (180.000 bytes) of secret data. For example, the following grid can be

IJERT

considered as 3 pixels of a 24 bit color image, using 9 bytes of memory:

(00100111 11101001 11001000)

(00100111 11001000 11101001)

(11001000 00100111 11101001)

When the character A, which binary value equals 10000001, is inserted, the following grid results:

(00100111 11101000 11001000)

(00100110 11001000 11101000)

(11001000 00100111 11101001)

In this case, only three bits needed to be changed to insert the character successfully. On average, only half of the bits in an image will need to be modified to hide a secret message using the maximal cover size. The resulting changes that are made to the least significant bits are too small to be recognized by the human eye, so the message is effectively hidden. While using a 24 bit image gives a relatively large amount of space to hide messages, it is also possible to use a 8 bit image as a cover source. Because of the smaller space and different properties, 8 bit images require a more careful approach. Images use three bytes to represent a pixel; an 8 bit image uses only one. Changing the LSB of that byte will result in a visible change of color, as another color in the available palette will be displayed. Therefore, the cover image needs to be selected more carefully and preferably be in grayscale, as the human eye will not detect the difference between different gray values as easy as with different colors.

C. DCT (Direct Cosine Transform)

Input: Message, cover image

Output: steganographic image containing message

While data left to embed do Get next DCT coefficient from cover image

If DCT not equal to 0 and DCT not equal to 1 then get next LSB from message

Replace DCT LSB with message bit

End if

Insert DCT into steganographic image

End while

D. VEDIO STEGANOGRAPHY

The main high resolution AVI fie is nothing but a sequence of high resolution image called frames. Initially I will like to stream the video and collect all the frames in bitmap format . And also collect the following information:

- Starting frame: It indicates the frame from which the algorithm starts message embedding.
- Starting macro block: It indicates the macro block within the chosen frame from which the algorithm starts message embedding.
- Number of macro blocks: It indicates how many macro blocks within a frame are going to be used for data hiding. These macro blocks may be consecutive frame according to a predefined pattern. Apparently, the more the macro blocks we use, the higher the embedding capacity we get. Moreover, if the size of the message is fixed, this number will be fixed, too. Otherwise it can be dynamically changed.

Frame period: It indicates the number of the inter frames, which must pass, before the algorithm period is too small and the algorithm repeats the message very often, that might have an impact onto the coding efficiency of the encoder. Apparently, if the video sequence is large enough, the frame period can be accordingly large. The encoder reads these parameters from a file. The same file is read by the software that extracts the message, so as both of the two codes to be synchronized. After streaming the AVI video file into AVI frames I will like to use the conventional LSB replacement method. LSB replacement technique has been extended to multiple bit planes as well. Recently has claimed that LSB replacement involving more than one least significant bit planes is less detectable than single bit plane LSB replacement.

IJERT

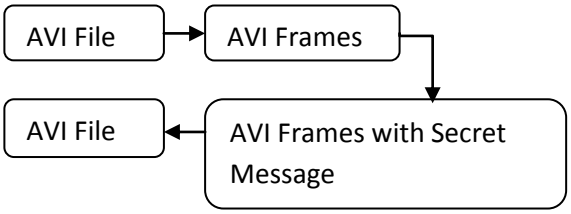

Figure: AVI File

6. Proposed work for Steganography

The main goal of this method is to hide information on the output image of the instrument (such as image displayed by an electronic advertising billboard). This method can be used for announcing a secret message in a public place. In general, this method is a kind of steganography, but it is done in real time on the output of a device such as electronic billboard. Following are the steps involved in embedding the secret information within a cover media.

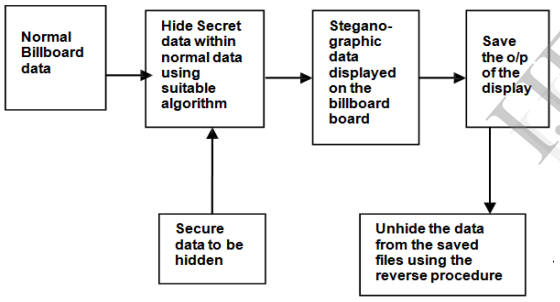

Figure: Block Diagram of Proposed ideas

7. Conclusion

The Module (create, delete & modify advertisemenr) is basically used to create a data base for the advertisement and send it to the second system, where the data is to be hidden. The stego image from the second system is sent to the display board. This image is captured by the camera & the decode module decodes the given information. The results shown consist of the image that is directly saved on to the desktop Steganography transmits secrets through apparently innocuous covers in an effort to conceal the existence of a secret. Digital image steganography and its derivatives are growing in use and application. In areas where cryptography and strong encryption are being outlawed, citizens are looking at steganography to circumvent such policies and pass messages covertly. As with the other great innovations of the digital age: the battle between cryptographers and cryptanalysis, security experts and hackers, record companies and pirates, steganography and Steganalysis will continually develop new techniques to counter each other. In the near future, the most important use of steganographic techniques will probably be lying in the field of digital watermarking. Content providers are eager to protect their copyrighted works against illegal distribution and digital watermarks provide a way of tracking the owners of these materials. Steganography might also become limited under laws, since governments already claimed that criminals use these techniques to communicate.

8. References

[1] Mohammad Shirali-Shahreza , "A new method for real time steganography",ICSP 2006 Proceedings of IEEE .
[2] Yuk Ying Chung, fang Fei Xu , "Development of video watermarking for MPEG2 video" City university of Hong Kong ,IEEE 2006.
[3] Jonathan Cummins, Patrick Diskin, Samuel Lau and Robert Parlett, "Steganography and digital watermarking" School of Computer Science, The University of Birmingham. 2003. www.cs.unibo.it/people/phdstudents/scacciag/home_files/teach/datahide.pdf.
[4] Martín Alvaro, Sapiro Guillermo and Seroussi Gadiel, "Is Image Steganography Natural?" IEEE Transactions On Image Processing, Vol. 14, No. 12, December, 2005.
[5] Cvejic N. and Seppänen T. "Increasing the capacity of LSB based audio steganography", Proc. 5th IEEE International Workshop on Multimedia Signal Processing, St. Thomas, VI, December 2002, pp. 336-338.
[6] Lee, Y. K. and Chen L. H. "High Capacity Image Steganographic Model". IEEE Proceedings Vision, Image and Signal Processing, pp. 288-294, 2000.
[7] Computational Intelligence in Scheduling (SCIS 07), IEEE Press, Dec. 2007, pp. 57-64, doi:10.1109/SCIS.2007.357670.

IJERT