HIDING AND RETRIEVAL OF TEXT CONTENT USING DISCRETE COSINE TRANSFORMATION (DCT) AND TRIPLEDES (3Des) METHOD

SIVASANGARI A/P KRISHNAN

THESIS SUBMITTED IN FULFILMENT OF THE DEGREE OF COMPUTER SCIENCE (COMPUTER SYSTEM AND NETWORKING)

FACULTY OF COMPUTER SYSTEM AND SOFTWARE ENGINEERING

2013

ABSTRACT

Computer security features for web applications are widely needed in many fields especially in defense industry. Detection phase is the most important aspect in the security development. Detection and investigation is commonly used to identify the suspected evidences from the computer crime. Steganography technique become the common approach for detect image based evidence. Previously, encryption technique is applied into image which is trial and error is applied to get precise evidences. This will cause the evidence to be not valid if proceed to the court. In this project steganography is used to get precise set of evidences in order to achieve desired proof to proceed to the court. Based on project, steganography is able to produce the output that precise to the desired output. From the output, the text document saved into computer to produce a solid evidence so that the cases related cybercrime can be proceed to the court. This technique overcomes the problem of trial and error in encryption to get the desired evidences. Steganography with encrypted able to produce a strong security environment in the computer to preserve the confidential and private information.

ABSTRAK

Ciri-ciri keselamatan komputer untuk aplikasi web secara meluas diperlukan dalam pelbagai bidang terutama dalam industri pertahanan. Fasa pengesanan adalah aspek yang paling penting dalam pembangunan keselamatan. Pengesanan dan penyiasatan biasanya digunakan untuk mengenal pasti bukti-bukti yang disyaki daripada jenayah komputer. Teknik steganografi menjadi pendekatan yang kukuh untuk mengesan bukti berasaskan imej. Sebelum ini, teknik penyulitan digunakan ke dalam imej dengan kaedah percubaan dan kesilapan digunakan untuk mendapatkan bukti-bukti yang tepat. Ini akan menyebabkan keterangan menjadi tidak sah jika ditujukan kepada pihak mahkamah. Dalam projek ini steganografi digunakan untuk mendapatkan set tepat bukti untuk mendapatkan bukti yang dikehendaki untuk diteruskan ke mahkamah. Berdasarkan projek, steganografi mampu menghasilkan output yang tepat seperti output yang dikehendaki. Dokumen teks akan disimpan ke dalam komputer untuk menghasilkan bukti yang kukuh supaya kes-kes berkaitan jenayah siber boleh diteruskan kepada pihak mahkamah. Teknik ini mengatasi masalah kaedah percubaan dan kesilapan dalam penyulitan untuk mendapatkan bukti-bukti yang dikehendaki. Steganografi dengan kaedah sulitan mampu menghasilkan keselamatan yang kukuh dalam komputer untuk menyelamatkan maklumat sulit.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

3Des	Triple Data Encryption Standard
C	Cover-Object
M	Secret-Messages
LSB	Least Significant Bits
SIS	Steganography Imaging Systems
SMBA	Stirmark Benchmark For Audio
WWII	World War II
JPEG	Joint Photographic Experts Group
TIF	Tagged Image File
PNG	Portable Network Group
GIF	Graphic Interchange Format
PNG	Portable Network Group
BMP	Bitmap Image File
S	Stego-image
SDLC	System Development Life Cycle
ESP	Encapsulating Security Payload Protocol
AH	Authentication Header protocol
M	Object
I	Key
K	Random Number
DCT	Discrete cosine transform

OTP	One-Time Password
XML	Extensible Markup Language
AGILE	Association of Geographic Information Laboratories for Europe
RAD	Rapid Application Development
TCP/IP	Transmission Control Protocol (TCP) /Internet Protocol (IP)
IP sec	Internet Security Protocol

CHAPTER 1

INTRODUCTION

1.0 INTRODUCTION

Chapter 1 basically highlights about the background of the research. The important keys will be discussed in this section to provide basic information about the research. All information that needed will be discussed in the research. This chapter explained root of the research to be expanded as overall research.

1.1 BACKGROUND

Steganography become one of the important techniques for hiding and retrieve private and confidential information. Those content are available in different form of resources such as cipher text, plain text or even images. Steganography works by hides bits of useless or unused data in regular computer files such as plain text with bits of different with image that make the information invisible by others.

Steganalysis techniques will analyze the steganography content over the plain text. User uses plain text to compose messages to send to their recipient via exchange those messages by using text messages through internet. User compose text message in plain text via electronic communication such as Facebook, emails, yahoo messenger, twitter and so on.

Those text messages will be hidden behind the image whereby the image seems just like a normal image. It will be harder for the people who are exploit security such as hacker, session hijacking, and spammer and so on to exploit our personal and confidential information. The user can compose messages by hide text content behind image to form steganography image. The text messages that send to the recipient will be encrypted with password to converse the original text content to encrypted text content to provide extra security features to the text content.

User can view the text messages by the retrieve the text content from the steganography images. The text will be displayed as encrypted text where it is difficult for the hackers to break the security breaches without known the user defined password. The encrypted text will be displayed as original text content only when the user decrypted the text content by password that they used to encrypted text content.

1.2 PROBLEM STATEMENT

The issues of computer crime through the use of a computer or against a computer system have increasing rapidly. Electronic communication application such as Facebook, Twitter, yahoo messenger and emails become one the most famous way to exploit personal information of individual. The criminal will access into an individual text messages to collect their personal information of victim. They hack into victim bank account by using the personal information gain through the text messages for financial gain. The illegal action of the cybercriminal causes a lot of detriment to electronic communication application user.

1.3 OBJECTIVE

The goal of steganography is hidden text content over image. So as a fundamental requirement for the steganography technique, text content will be invisible to the third parties except the electronic communication sender and recipient who compose text messages in the text content format.

Those are objectives for research:-

i. To explore technique of hiding encrypted text content behind image through encryption

ii. To developed search technique in order to retrieve encrypted text content that hidden behind image

iii. To secure text content via product of steganography tool using discrete cosine transformation (DCT) and triple data encryption standard (TripleDes) method.

1.4 SCOPE

The purpose of research is to developed a technique for hiding and retrieve text content behind image. The encrypted and decrypted text content with user defined password will be hide and retrieve from the image. The scope of this research for the implemention of steganography by encrypt and decrypt the message techniques for hiding and retrieve text content includes any type of plain text in .txt and image format.

1.5 ORGANIZATIONS OF THESIS

This thesis consists of six (6) chapters as follow:-

i. Chapter 1 will discuss on introduction to system research.
ii. Chapter 2 will discover on the review of findings by other researcher
iii. Chapter 3 will show method that used to approaches the research
iv. Chapter 4 will design the system and implement functional process into system
v. Chapter 5 will explain about the findings and results from perform of system
vi. Chapter 6 will explain the overall perform of the system

CHAPTER 2

LITERATURE REVIEW

2.0 INTRODUCTION

In this chapter, it covers about the literature review of the hiding and retrieval of encrypted text content by using DCT method via steganography technique and TripleDes method to encrypt and decrypt text content. This chapter comprises about the steganography and its history, types of technique for steganography to explore and extract data. Moreover, it will explain in detail about the techniques, method, software, and hardware which are suitable to be apply into the project.

2.1 HISTORY OF STEGANOGRAPHY

In ancient Greece, people were used to write messages by wood whereby those messages were hidden behind wax so that it seem like an ordinary, the hidden messages were implemented in wax tablets by obsolete tablet. Herodotus tells that the growth of his hair was hidden and exposed by shaving his head again which involves hidden on messenger's body whereby narrative of messages tattooed on a slave's shaved head (Tim.G, 2009).

During and after World War II, hidden messages were written with secret inks on paper under other messages or on other messages of the blank parts were implied by espionage agents by using photographically to produce microdots to send and receive information. A typewriter produced extremely small and even smaller or size of period that produced dots that were hidden within the dots which known as Stego@Text. They needed to be embedded in paper and wrapped with an adhesive which could be identify which is the problem arise from WWII microdots. The embedded microdots would reflect light differently than the papers (David K, 1996).

Velvalee Dickinson who is the Japanese dolls dealer in New York City, sent information to accommodation addresses in neutral South America during World War II. How many of this or that doll to ship was discussed in her letters. The stego-text in this case doll orders. The form of code text which concealed with plaintext was the information about ship movement. The Doll Woman became famous case (Neil F. J, 2009).

The theoretically unbreakable cipher which distinguishable from random texts known as cipher-texts that produced one-time pad. The only private key from any other perfectly random texts where these cipher-texts can be distinguish from private key. A cover-text for a theoretically unbreakable steganography derived from any perfectly random data. In the most of cryptosystems, private symmetric session keys are supposed to be perfectly random. In countries, users of weak crypto where strong crypto is forbidden can safely hide OTP messages in their session keys (James C. J, 2001).

2.2 CONCEPT OF STEGANOGRAPHY

Steganography is a process of hiding information behind images. There are many types of images such as joint photographic experts group (JPEG), tagged image file(TIF), portable network graphic(PNG) and graphics interchange format(GIF) which will corporate with steganography technique for hiding information in protective way. The types of information that will be hides behind images might come from different formats such as audio, document and images.

In the digital steganography technique, the coding of the steganography may include inside of transport layer if it was electronic communication. The cover medium which hide data and also encrypted using steganography key to form steganography medium as a result form the steganography method. Basically steganography techniques, hides personal and confidential data in secure form from acknowledgement of others. Secret data will be replaced with unwanted bits in images (Bret.D, 2002).

Secure communication between two parties who known as sender and receiver to exchange data were established from steganography techniques. The steganography will use smaller memory space to hide data. There is embedded secret data within the images or audio files within the beautiful and attractive images or audio that will divert the mind of person. The secrecy technique to preserve the data to become invisible from other user of electronic devices can also become as one of steganography techniques.

2.3 IMAGE BASED STEGANOGRAPHIC SYSTEMS

Images which provide space to embed data contain quantization noise act as carrier media after digitalization. Alice wants to send a secret message M to Bob. She hides M into a cover-object C, and obtains a stego-object S. Then stego-objecct S sent through the public channel (Mehdi.K, Husrev T. S, Nasir.M, 2004).

Cover-Object: Based on the object that used as carrier to hide messages into many different objects that been employed to hide messages into images.

Stego-Object: Refers to the object which is containing a hidden message.

In a pure steganography framework, the technique for embedding the message is unknown to Wendy and shares as a secret between Alice and Bob. The steganographic algorithm identifies C's redundant bits, then the embedding process creates stego image S by replacing these redundant bits with data from M (Mehdi.K, Husrev T. S, Nasir.M, 2004).

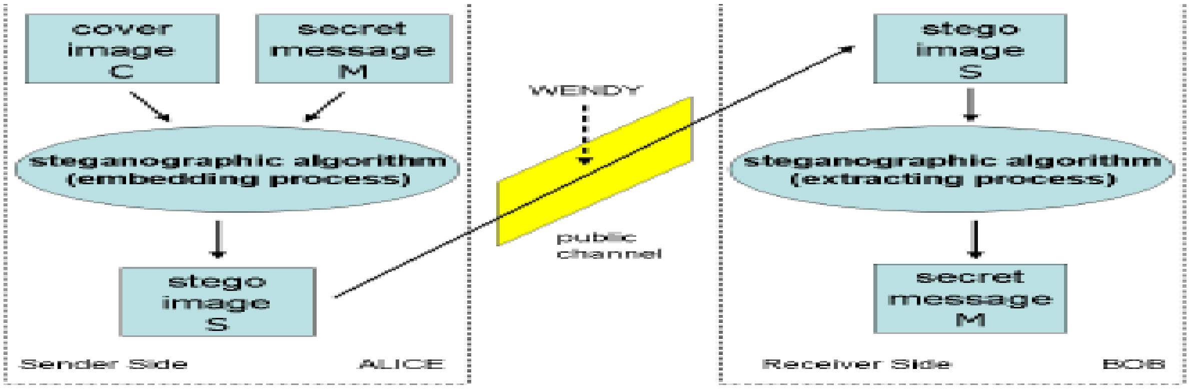


Figure 2.1: Steganographic Model

Source: Domenico.B and Luca.I

The public channel that transmitted over a S which is accepted by Bob only if
S is transmitted over a public channel and is received by Bob only if Wendy has no prejudice on it. Bob can get M through the extracting process if he recovers S. The embedding process where S must similar to C for ignore Wendy's intervention which is known as critical task for a steganography system **(**Domenico.B, Luca.I**)**.

In a cover file whereby M's bit is overwrites by the LSB of a pixel was the simple approach to insert information. I f we choose 24-bit images then each pixels can store 3 bits. The cover image will look similar with result steganography image for human eye
(Johnson, Jajodia, 1998).

If modifying cover image alters its statistical properties then eavesdroppers can identify the distortions in the resulting steganography images. The insertion of high-entropy due to data changes the histogram of color frequencies in a predictable way
(Provos.H, 2003).

Westfeld who proposed F5% algorithm does not overwrite LSB and statistical properties of steganography images were preserved. They recommend encrypt M before embedding since standard steganography systems do not provide strong message encryption. We have to deal with a two-steps protocol: first we must cipher M for obtaining M' and then we can embed M' in C (Westfeld, 2001).

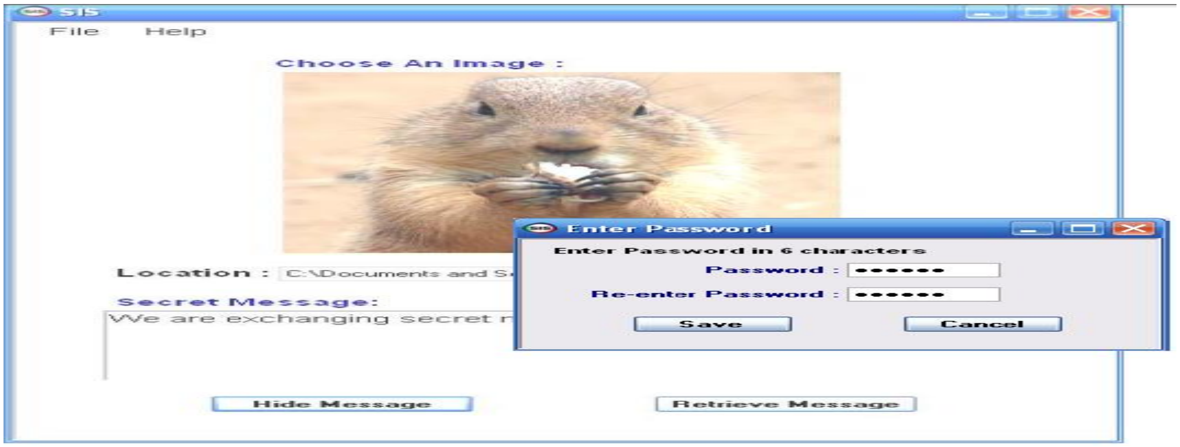
Online way of functioning or work with static images method has been planned either to work with bit streams scattered through mutliple images.The stego image's statiscal properties were preserved by cipher M in a theoretically secure manner that yields random outputs to make steganalysis more difficult. The simplicity shown possibility of using the method in real-time applications like mobile video.

2.4 STEGANOGRAPHY IMAGING SYSTEM (SIS)

Steganography Imaging System (SIS) is a system that eligible for embedding data within the image. The security implement in the system via 2 layers of security to maintain data privacy (Rosziati.I, Teoh.S, 2010).

Steganography basically aids to authorized access and maintain data protected from corruption. The motives of the data security to make sure the privacy of individual information without the corruption in the data that lead to legal issues and dissemination of data. The main advantage of steganography technique because of its simple security mechanism compared to water marking techniques improves the usability of SMBA in large test by allowing the tester to define attack profile profiles which can be saved in XML files (Christian.K, Elke.F, Jean-Luc.D, Andreas.L, 2007).

This research use similar techniques used by the Westfeld but it slightly different with the strong encryption secret key to hide and retrieve the messages from the images.



Location : C:\Documents and S

Secret Message:

We are exchanging secret r

Enter Password in 6 characters

Password : ●●●●●●●

Re-enter Password : ●●●●●●●

Figure 2.2: The secret key required for SIS

Source: Rosziati.I, Teoh.S (2010)

This system also enhanced with two layer security system compared to image based steganography systems which does not implement any security layer which just involve embedding process and extracting process with image that make data invisible. The first layer security as login purpose compared to image based steganography system just hide the secret message within cover image to form stego image via embedding process while in this system the embedding process and retrieving process only involve in the layer 2 of security (Domenico B. ,Luca I.) .

From figure 5, the secret key is required to enter twice for the verification purposes. For simplicity, 6 characters are used for the secret key and embedded inside image together with the data due to reduce size of storing inside the image, only 6 characters are used for the secret key. The data has key in and secret key has been entered, the new stego image can be saved in different image file (Rosziati.I, Teoh.S, 2010).

This new stego image can be used by user to send it via internet or email to other parties without revealing the secret data inside the image. If other parties want to reveal the secret data hidden inside the image, the new stego image file can then be upload again using the system to receive that data that been locked within image using secret key compared to public Vs. private watermarking techniques which implies private watermarking techniques is the watermark is verified in presence of the public where suffer from disclosure of the private parameters to dishonest people while invalidate watermark detection can easily detect by attackers by removing watermarks from the protected data or by adding a false watermark to non-watermarked data (Halder R., Pal S., Cortesi A., 2010).

2.5 PRACTICAL INTERNET STEGANOGRAPHY

The technique to hide data in IP called as practical internet steganography. The internet was established on an already piggybacks over covert channel called as overt channel. The covert communications do not interfere with normal overt information flow via embedding and detection processes. TCP/IP protocol suite involving employ the data hiding and transfer information overtly over a computer network through assumption that communication parties denotes as Alice and Bob compared to steganography imaging system just can send the embedding data over image via internet or email

(Steven J.M, Stephen.L, 2005).

The cover-network packet sequence were taken as stego-algorithm input and overt payload of {P$_k$} while piggybacking C$_k$ contain secret key to generate a stego-network packet sequence {S$_k$} known as data hiding through covert message C$_{k,}$ a sequence of network packet {P$_k$}(Deepa K., Kamran A.).

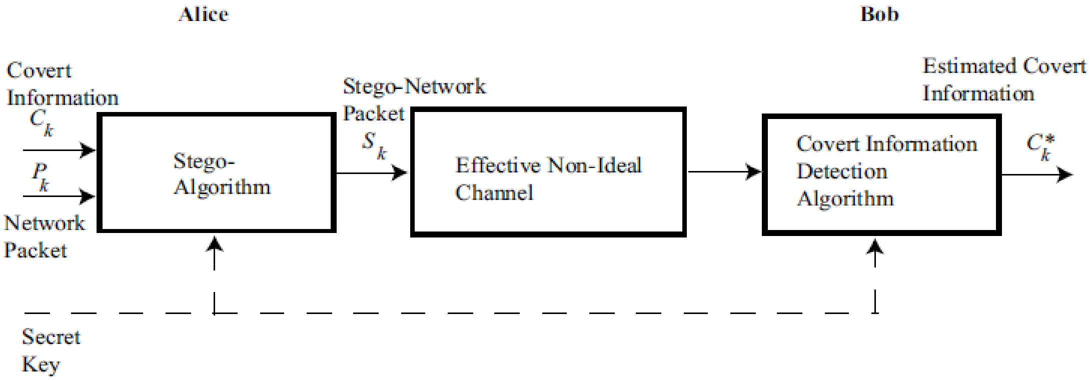

Alice
Bob
Covert Information C_k
P_k
Network Packet
Stego-Algorithm
Stego-Network Packet S_k
Effective Non-Ideal Channel
Covert Information Detection Algorithm
Estimated Covert Information
C^*_k
Secret Key

Figure 2.3: The general covert channel framework in TCP/IP

Source: Deepa.K and Kamran.A