Fundamentals of Information Hiding

Lecturer: Dr. Honoured Worker of High School of
Russian Federation, IEEE Member,
Professor of Protected Telecommunication
System Department at SPbSUT
Valery Korzhik

E-mail: val-korzhik@yandex.ru
korzhik1@bk.ru

Lecture 1. Introduction

Steganography (In a broad sense)

Definition. IH –is a family of methods that provide an embedding of some additional information into the main (cover message (CM) or cover object (CO)) under the condition that the quality of the CO is kept acceptable.

Two main parts of IH :

1.Proper steganography (SG).

2.Digital watermarking (WM).

The goal of SG:

Embed some additional information into CO in such a way that it would be impossible to detect by unauthorized users even the fact of presence any additional information in the CO.

The goal of WM : Embed into CO some additional information (typically identification code of CO ‘s owner) in such a way that it would be impossible to remove it by unauthorized user without significant corruption of CO.

(The fact of such embedding can be detected even by unauthorized users.)

Typical CO:

- motionless image
- video images
- audio files
- speech
- printed text
- image documents
- internet protocols
- source codes

Embedding information:

- Images
- text messages and date
- speech

Remark. It is commonly to encrypt the embedded messages in advance with the use of specially chosen ciphers and encrypting keys.

Kerckhoffs' assumption for IH: It is assumed that all users know everything about SG system (including embedding and extraction algorithms) except for *stegokey* (if it is used in SG system).

Unauthorized user that try to break IH system (detect SG presence or to remove WM) is called an attacker, or adversary and his (or her) activity is called an attack on IH system.

Part 1. Stegosystems .

SG versus cryptography (CR):
CR conceals the content of messages still maintaining an opportunity to detect the fact of CR application (unreadible text ctr) , whereas SG conceals the fact of embedding some additional information into “innocent” CO.

The goals of SG application :
1.Alternative to CR under the condition that it is prohibited or limited in use by low.
2.Concealing of users which are needed to save or to transmit some secret information.
3.Retransmission of secret information via unauthorized users.
4. Transmission of secret information (or commands) to authorized users of Internet.
5. Tracing of illegal distributors of some messages.

In the recent paper [75] of the SG researchers leaders has been emphasized that in spite of the promising theoretical results regarding a building of SG systems theory a significant a gap between theory and practice. The researchers formulate open problems which can move theory to practice.

Some examples of the use of steganography by terrorists and criminals:

1. Al Qaeda, September 9-11 (E-mail messages with graphical attachments containing secret information).

   Ref.: S. Betancourt, “Steganography: A New Age of Terrorism”, GSEC, Practical Version, SANS Institute, 2004.

2. Russian spies (2010). The use stegosystems in attachments to E-mail.

   Ref.: http://www.csmonitor.com/Science/2010/0630/

   “How Russian spies hid secret codes in online photos”.

3. Attack on tourist hotel in Mubai (India, 2009). The use of S6 in Emails.

   Ref.: http://www.dnaindia.com/mubai/report_mubai-police fail to crack-july11-suspects-mail-1058716

4. Steganography is used as a mean to distribute child pornography by embedding the contraband images in adult pornographic images.

   The use of e-Bay.

   Ref.: http://news.bbc.co.uk/2/hi/science/nature/2082657.stm

The main attacks on IH :
- detecting of SG
- evaluation of the size of the embedded messages
- “reading” of embedded messages
- removal of embedded messages without significant corruption of CO and even in the case when SG cannot be detected (“at worst case”).

Historical example : A randomization of hand’s positions in the consignment of watches at US customer service on the Second World War time keeping in mind that may be these positions contain some encrypted messages for illegal agents in US *.*

Possible transforms of SG signals :
- natural transforms (filtering, compression, scaling, transmission over noisy channels);
- deliberate transforms (attacks) .

Criteria of SG system efficiency :
-the probability of SG missing
- the probability of false SG detecting
- the probability of bit error in the extracted message by legal users
-quality of CO after embedding of secret message (signal–to-noise ratio or more complex criteria for audio and video messages)
- date embedding rate (the number of embedded bits per one sample of CO) .

Definition. SG system is called *robust* if secret message can be reliablely extracted even after all natural or deliberate transforms which do not corrupt CM significantly .

The main classification of SGs:

a) For legal users:

- with known CO at legal decoder (*informed decoder)*

-with unknown CO at legal decoder (*blind decoder*)

- with the use of CO at legal encoder (*informed encoder*) .

b) For attackers:

- given known SG message (it is true always)

- given known embedded message

- given chosen embedded message

- given known CO (only for SG based on noisy channels - see lecture in the sequel).

The main concept in design of SG:

Find noise components (elements) of CO and replace them to encrypted secret message.

The main problem in design of SG:

Statistic of such complex CO as audio , video or meaningful text is known only partly and it is very hard to be formalized .Thus there exists an opportunity that attacker knows it even better than designer of SG.

Short historical review of SGs .

1. Story described by Herodot (shave off the hairs on the head of the slave, make a tattoo as secret message, wait till the time when the hairs grow again and then send the slave through adversary’s country to sender’s agent , where the slave has to be shaved off again and then a secret message can be be read simply).
2. Egypt’s writing containing cover information for “experts”.
3. Indian “Kama-Sutra” where “secret writing’ has even the number 45…
4. Cover sense in Bible’s stories :

" And he said unto them , Unto you it is given to know the mystery of the kingdom of God : but unto them that are without, all *these* things are done in parables: That seeing they may see, and not perceive; and hearing they may hear , and not understand; lest at any time they should be converted, and *their* sins should be forgiven them.” (The Gospel according to St. Mark).

5. Sympatic ink .
6. Language of gestures (cardsharpers).
7. Arrangement of objects (Stones at the house in Korean’s ancient story).

The main feature of contemporary SG :

It is *digital SG* where all CO are presented in digital form whereas embedding and extraction of secret messages are performed on PC.

If you have at least one of the following problems steganography is the perfect solution for you.

1. Are you afraid that someone else can see your sensitive and secret data stored on your computer?
2. Are you afraid that hackers or other people can penetrate your system and find your valuable information?
3. Your wife, boss and kids can see what you used the internet for?
4. You have many passwords in your mind or spread all over your computer and you can’t organize them?
5. You want to send a secure email to your friend or partner and you want nobody to read or to access it?
6. You want to password protect certain applications to be used only by you? Do you want them not to be visible in the Start Menu, but still have quick access to them?
7. You want to delete an email or a file but you think someone else can restore it and use it against you?
8. You want to combine file encryption with steganography (hide files and folders) to better protect your documents and emails?
9. You want to hide files on your computer so that nobody finds them?
10. Is your current computer security program too complex and too hard to use and want something easier but still keep the same power and security?

IH Courses in other countries

- Course name: Multimedia security
  University: University of Geneva

- Course name: Digital Watermarking
  University: City University of New York

- Course name: Digital Watermarking
  University: University of Ottawa

- Course name: Fundamentals of Watermarking and Data Hiding
  University: University of Illinois

- Course name: Security Information Hiding
  University: University of Surrey

Specific of the current course:

It will contain both well known facts of IH theory and some new results which have been proved by lecturers research group.