Part XI

Steganography and Watermarking

STEGANOGRAPHY and WATERMARKING

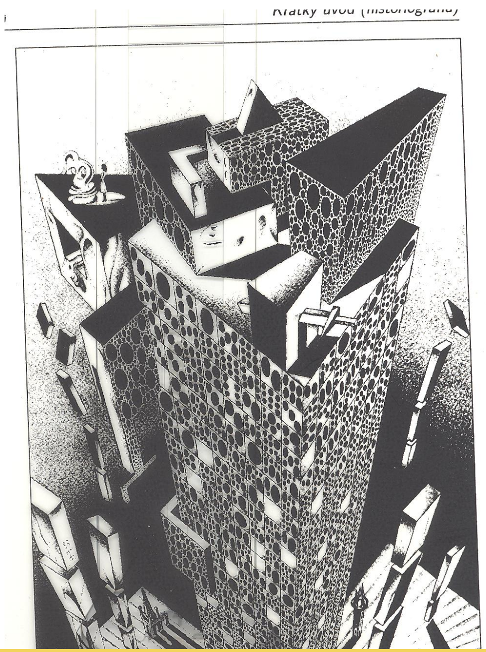
Steganography and **Watermarking** are arts, sciences and technologies of **hiding** information.

Cryptography goals is to make transmitted messages **unreadible** by the third party.

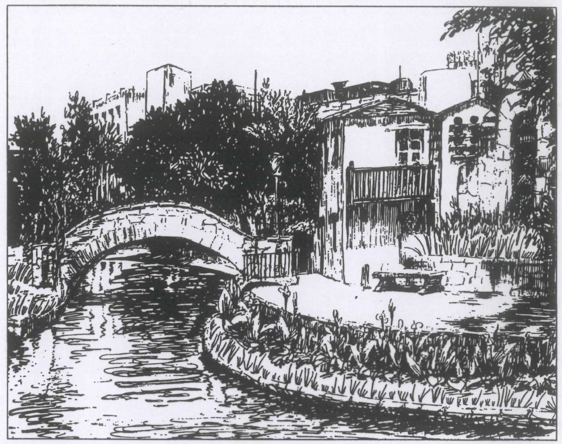
Steganography/watermarking goals is to make transmitted messages **unvisible** by the third party.

FIND NUMBERS

Find two well-known numbers on the folllowing picture



ANALYSIS of a SCENE - I.

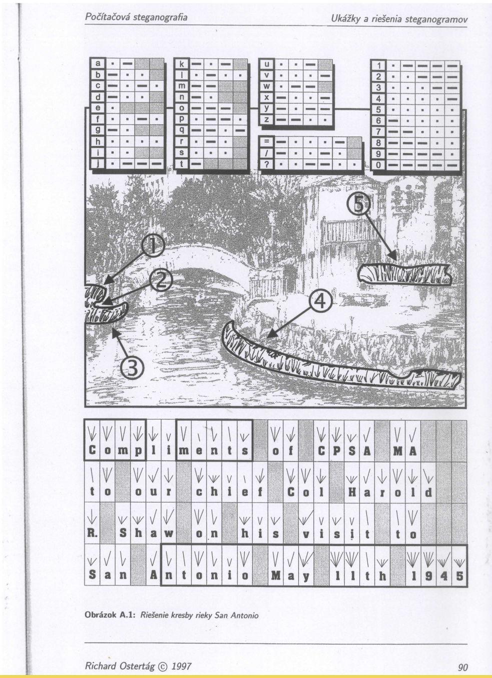


Počítačová steganografia *Krátky úvod (historiografia)*

Obrázok 0.4: *Kresba rieky San Antonio*

rývajú tajnú správu. Každá cenzorská stanica mala svoju banku známok. Cenzori známky, ktoré mohli niesť nejakú tajnú infor-

ANALYSIS of a SCENE - II.



DIGITAL STEGANOGRAPHY and DIGITAL WATERMARKING

PROLOGUE

In this chapter we deal with a variety of methods how to hide information. Hiding of information is much needed in many important cases.

Our main attention will be devoted to methods developed in Steganography and Watermarking.

We will also discuss several anonymity problems and methods to solve them.

Preservation of the anonymity of communicating parties is in many cases also of large importance.

PROLOGUE I - PROBLEMS WITH COPYING of INFORMATION

A very important property of (digital) information is that it is, in principle, very easy to produce and distribute unlimited number of its copies.

This might undermine the music, film, book and software industries and therefore it brings a variety of important problems, concerning protection of the intellectual and production rights, that badly need to be solved.

The fact that an unlimited number of perfect copies of text, audio and video data can be illegally produced and distributed has serious consequences. For example, it is much needed to develop ways of embedding copyright and source information into audio and video data.

Digital steganography and digital watermarking bring techniques to hide important information, in an undetectable and/or irremovable way, in audio and video digital data.

Digital steganography is the art and science of embedding information/signals in such a hidden way, especially in texts, images, video and audio carriers, that only intended recipients can recover them.

Digital watermarking is a process of embedding (hiding) information (through "watermarks") into digital data (signals) - picture, audio or video - to identify its owner or to authentisize its origin in an unremovable way.

Steganography and (digital) watermarking are main parts of the fast developing area of **information hiding.**

INFORMATION HIDING SUB-DISCIPLINES

Covert channels occur especially in operating systems and networks. They are communication paths of networks that were neither designed nor intended to transfer information, but can be used that way.

These channels are typically used by untrustworthy/spying programs to leak (confidential) information to their owner while performing service for another user/program.

Anonymity is finding ways to hide meta content of the message (for example who is the sender and/or the recipients of a message). Anonymity is needed, for example, when making on-line voting, or to hide access to some web pages, or to hide sender.

Steganography – covered writing – from Greek $\sigma\tau\epsilon\gamma\alpha\nu$–$\xi$ $\gamma\rho\alpha\phi$–$\epsilon\iota\nu$

is the art and science of hiding secret messages in innocently looking ones.

Watermarking – is the technique to embed visible and especially imperceptible (invisible, transparent,...) watermarks into carriers in undetectable or unremovable way.

WHY is PROTECTION of INTELLECTUAL RIGHTS so IMPORTANT?

- It is estimated that business and individuals lost a total 63 billions of euro due to forgery alone in the first five years of 21st century.
- Frauds on this scale are also the major source of funding of various criminal activities.
- It is estimated that 40% of drugs in Africa and China are fake.
- It is estimated that most of the fake drugs have little or no medical value.

There are various attempts to deal with this problem.

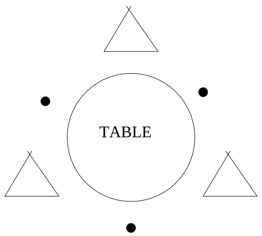
Perhaps the most modern one that is being explored is to write down watermarks into materials using tools of nanotechnology.

ANONYMITY

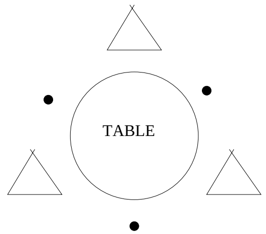
ANONYMITY

THE DINING CRYPTOGRAPHERS PROBLEM - I.

- Three cryptographers have dinner at a round table of a 5-star restaurant.

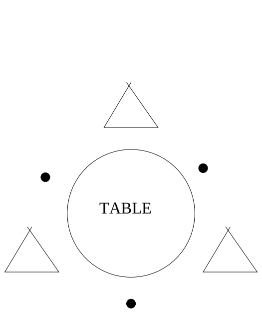


THE DINING CRYPTOGRAPHERS PROBLEM - II.



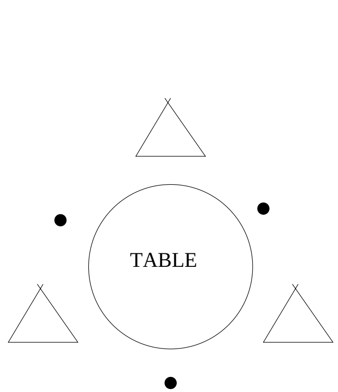
- Their waiter in the restaurant tells the cryptographers that an arrangement has been made that bill will be paid anonymously - either by one of them, or by NSA.
- They respect right of each other to make an anonymous payment, but they would like to know whether NSA payed the dinner.
- How should they proceed to learn whether one of them will pay the bill without learning which one - for other two?

DINNING CRYPTOGRAPHERS - SOLUTION



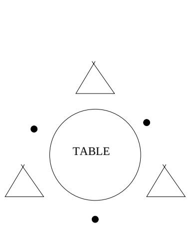
- Protocol
  - Each cryptographer flips a perfect coin between him and the cryptographer on his right, so that only two of them can see the outcome.
  - Each cryptographer who did not pay dinner states aloud whether the two coins he see - the one he flipped and the one his right-hand neighbour flipped - fell on the same side or on different sides.
  - The cryptographer who paid the dinner states aloud the opposite what he sees.

ANALYSIS of the SOLUTION



- Correctness
  - An odd number of differences uttered at the table implies that that a cryptographer paid the dinner.
  - An even number of differences uttered at the table implies that NSA paid the dinner.
  - In a case a cryptographer paid the dinner the other two cryptographers would have no idea he did that.

TECHNICALITIES of SOLUTION



Let three coin tossings made by cryptographers be represented by bits b_1, b_2, b_3. In case none of them payed dinner, what cryptographers make public are values

$$b_1 \oplus b_2, b_2 \oplus b_3, b_3 \oplus b_1$$

and their parity is

$$(b_1 \oplus b_2) \oplus (b_2 \oplus b_3) \oplus (b_3 \oplus b_1) = 0$$

In case one of them payed dinner, say Cryptographer 2, they announce:

$$b_1 \oplus b_2, \overline{b_2 \oplus b_3}, b_3 \oplus b_1$$

and

$$(b_1 \oplus b_2) \oplus (\overline{b_2 \oplus b_3}) \oplus (b_3 \oplus b_1) = 1$$

ANONYMOUS TRANSFER PROTOCOLS

- The term anonymous transfer includes a variety of different tasks.
- Anonymity of an object is the state of being not identifiable with any particular element of a set of subjects known as **anonymity set**.
- **Anonymity set** is a set P of participants able to perform a particular action we are interested in. (For example that a real sender (receiver) is not identifiable within a set of potential senders (receivers).)
- **Cheating** is modeled by an adversary A not in P, who takes a full control of some subset $C \subset P$ of (malicious) participants. (A is assumed to have access to all memories, inputs and outputs of all participants in C - this way one can model the case of cooperating malicious participants).

ANONYMOUS BROADCAST and MANY-TO-ONE COMMUNICATION

- **Anonymous one-to-many or broadcast communication** has one anonymous sender and all parties receive the message that has been sent.
- **Anonymous many-to-one communication** has all parties to send their messages and there is only one receiver.

CHAUM's PROTOCOL for ANONYMOUS BROADCASTING

Let communicating scheme is modeled by an unoriented graph $G = (V, E)$ with $V = \{1, 2, \dots, n\}$ representing nodes (parties) and E edges (communication links). Let n be a large integer.

Protocol: Party P_i performs the following actions (all parties in parallel).

- **For each** $j \in \{1, 2, \dots, n\}$ **it sets** $k_{ij} \leftarrow 0$;
- **If** $(i, j) \in E$, $i < j$, **randomly chooses a key** k_{ij} **and sends it securely to** P_j;
- **If** $(i, j) \in E$, $j < i$, **after receiving** k_{ji} **it sets** $k_{ij} \leftarrow -k_{ji} \bmod n$;
- **It broadcasts** $O_i = m_i + \sum_{j=1}^{n} k_{ij} \bmod n$, **where** $m_i \in \{0, \dots, n-1\}$ **is the message being sent by** P_i;
- **It computes the global sum** $\Sigma = \sum_{j=1}^{n} O_j \bmod n$.

Clearly, $\Sigma = \sum_{j=1}^{n} m_j \bmod n$, and therefore if only one $m_j \neq 0$, all participants get that message.

One can show that to preserve anonymity of a correctly behaving sender P_i it is sufficient that one another participants P_j such that $(i, j) \in E$ behaves correctly.

STEGANOGRAPHY versus WATERMARKING.II

Both techniques belong to the category of information hiding, but the objectives and embeddings of these techniques are just opposite.

In watermarking, the important information is in the cover data. The embedded data - watermarks - are for protection or detection of the cover data origins.

In steganography, the cover data is not important. It mostly serves as a diversion from **the most important information that is in the embedded data.**

Comment: Steganography tools typically embed/hide relatively large blocks of information while watermarking tools embed/hide less information in images or sounds or videos or texts.

Data hiding dilemma: to find the best trade-off between three quantities of embeddings: **robustness, capacity and security.**

STEGANOGRAPHY versus WATERMARKING again

Technically, differences between steganography and watermarking are both subtle and quite essential.

The main goal of **steganography** is **to hide** a message **m** in some audio or video (cover) data **d**, to obtain new data **d'**, in such a way that an eavesdropper **cannot detect** the presence of **m** in d'.

The main goal of **watermarking** is **to hide** a message **m** in some audio or video (cover) data **d**, to obtain new data **d'**, practically indistinguishable from **d**, by people, in such a way that an eavesdropper **cannot remove or replace m** in d'.

Shortly, one can say that **cryptography is about protecting** the content of messages, **steganography is about concealing** its very existence.

Steganography methods usually do not need to provide strong security against removing or modification of the hidden message. Watermarking methods need to to be very robust to attempts to remove or modify a hidden message.

STEGANOGRAPHY versus CRYPTOGRAPHY

- **Cryptography is art, science and technology of presenting information through secret codes.**
- **Steganography is art, science and technology of hiding information.**
- **The goal of cryptography is to make the data unreadable by a third party.**
- **The goal of steganography is to hide the data from a third party.**

Steganography is often used with cryptography to crate a double protection. Data are first encrypted using a cryptography system and then hidden using a steganography tool.

BASIC QUESTIONS

- Where and how can be secret data undetectably hidden?
- **Who and why needs steganography or watermarking?**
- What is the maximum amount of information that can be hidden, given a level of degradation, to the digital media?
- **How one chooses good cover media for a given stego message?**
- How to detect, localize a stego message?

SOME APPLICATIONS of STEGANOGRAPHY

- To have secure secret communications where cryptographic encryption methods are not available.
- To have secure secret communication where strong cryptography is impossible.
- In some cases, for example in military applications, even the knowledge that two parties communicate can be of large importance.
- The health care, and especially medical imaging systems, may very much benefit from information hiding techniques.
- Various secret religious groups and terrorist groups have been reported to use steganography to communicated in public.
- Methods and tools of steganography are consider of increasing importance for national security of world-powers and their developments and study is seen as being of increasing importance.

SOME APPLICATIONS of WATERMARKING

A basic application of watermarking techniques is to provide ownership information of digital data (images, video and audio products) by embedding copyright information into them.

Other applications:

- **Automatic monitoring and tracking of copyright material** on WEB. (For example, a robot searches the Web for marked material and thereby identifies potential illegal issues.)
- **Automatic audit of radio transmissions**: (A robot can "listen" to a radio station and look for marks, which indicate that a particular piece of music, or advertisement, has been broadcast.)
- **Data augmentation** – to add information for the benefit of the public.
- **Fingerprinting applications** (in order to distinguish distributed data)

Actually, watermarking has recently emerged as the leading technology to solve the above very important problems.

All kind of data can be watermarked: audio, images, video, formatted text, 3D models, . . .

STEGANOGRAPHY/WATERMARKING versus CRYPTOGRAPHY

The purpose of both is to provide secret communication.

Cryptography hides the contents of the message from an attacker, but not the existence of the message.

Steganography/watermarking even hide the very existence of the message in the communicated data.

Consequently, **the concept of breaking the system** is different for **cryptosystems** and **stegosystems (watermarking systems).**

- A cryptographic system is broken when the attacker can read the secrete message.
- Breaking of a steganographic/watermarking system has two stages:
  - The attacker can detect that steganography/watermarking has been used;
  - The attacker is able to read, modify or remove the hidden message.

A steganography/watermarking system is considered as insecure already if the detection of steganography/watermarking is possible.

The advantage of steganography over cryptography is that messages do not attract attention to themselves.

CRYPTOGRAPHY and STEGANOGRAPHY

Steganography can be also used to increase secrecy provided by cryptographical methods.

Indeed, when steganography is used to hide the encrypted communication, an enemy is not only faced with a difficult decryption problem, but also with the problem of finding the communicated data.

FIRST STEGANOGRAPHIC METHODS

- First recorded use of steganographic methods was traced to 440 BC. Greek Demaratus sent a warning about an attack by writing it on a wooden desk and then covering it by vax and writing on that an innocent message.
- Ancient Chinese wrote messages on fine silk, which was then crunched into a tiny ball and covered in wax. The messenger then swallowed the ball of wax.
- A variety of steganographic methods was used also in Roman times and then in 15-16 century (ranging from coding messages in music, and string knots, to invisible inks).
- In the sixteenth century, the Italian scientist Giovanni Porta described how to conceal a message within a hard-boiled egg by making an ink from a mixture of one ounce of alum and a pint of vinegar, and then using ink to write on the shell. The ink penetrated the porous shell, and left the message on the surface of the hardened egg albumen, which could be read only when the shell was removed.
- Special invisible "inks" (milk, urine,...) were important steganographic tools since middle ages and even during the Second World War.
- Acrostic - hiding messages in first, last or other letters of words was popular steganographic method since middle ages.
- During the Second World War a technique was developed to shrink photographically a page of text into a dot less than one millimeter in diameter, and then hide this microdot in an apparently innocuous letter. (The first microdot has been spotted by FBI in 1941.)

HISTORY of MICRODOTS

- In 1857, Brewster suggested hiding secret messages "in spaces not larger than a full stop or small dot of ink".
- In 1860 the problem of making tiny images was solved by French photographer Dragon.
- During Franco-Prussian war (1870-1881) from besieged Paris messages were sent on microfilms using pigeon post.
- During the Russo-Japanese war (1905) microscopic images were hidden in ears, nostrils, and under fingernails.
- During the First World War messages to and from spies were reduced to microdots, by several stages of photographic reductions, and then stuck on top of printed periods or commas (in innocuous cover materials, such as magazines).

FIRST STEGANOGRAPHY BOOKS

In the fourth century BC, the Greek Aeneas Tacticus, wrote a book on military techniques, On the defence of fortification in which the whole chapter is devoted to steganographic methods.

In 1499 Johannes Trithemius, opat from Würzburg, wrote 3 out of 8 planned books "Steganographie".

In 1518 Trithemius printed 6 books, 540 pages, on cryptography and steganography called **Polygraphiae.**

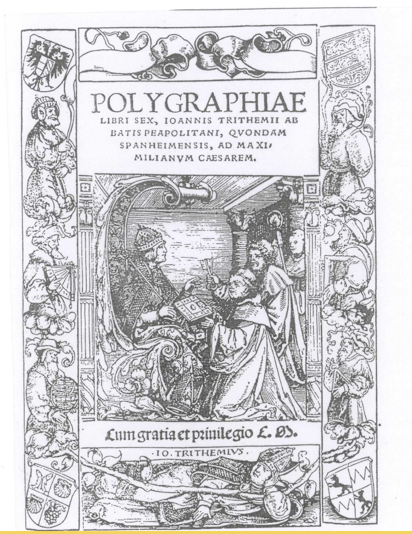
This is Trithemius' most notorious work. It includes a sophisticated system of steganography, as well as angel magic. It also contains a synthesis of the science of knowledge, the art of memory, magic, an accelerated language learning system, and a method of sending messages without symbols.

In 1665 Gaspari Schotti published the book "Steganographica", 400pages. (New presentation of Trithemius.)

TRITHEMIUS

- Born on February 1, 1462 and considered as one of the main intellectuals of his time.
- His book STEGANOGRAPHIA was published in 1606.
- In 1609 catholic church has put the book on the list of forbidden books (to be there for more than 200 years).
- His books are obscured by his strong belief in occult powers.
- He classified witches into four categories.
- He fixed creation of the world at 5206 B.C.
- He described how to perform telepathy.
- Trithemius died on December 13, 1516.

FRONT PAGE of the TRITHEMIOUS BOOK



PHYSICAL versus DIGITAL STEGANOGRAPHY

Steganography that was used before the computer era is usually called **physical steganography** because physical carriers have been used to embed secret messages.

Steganography using enormos potential of digitalization and of modern computers is usually called **digital steganography**.

MODERN DIGITAL STEGANOGRAPHY

THEORY and METHODS

ORIGIN of MODERN - DIGITAL - STEGANOGRAPHY

The origin of modern (digital) steganography has been dated to around 1985 - after personal computers started to be applied to classical steganographic problems.

This was related to new problems at which information needed to be sent securely and safely between parties across restrictive communication channels.

B. Morgen and M. Bary, from a small Dallas based company created and made public two steganographic systems.

Since then a huge spectrum of methods and tools have been discovered and developed for digital steganography.

Some examples:

- Network steganography
- WLAN steganography
- Inter-protocol steganography
- Blog steganography
- Echo steganography

GENERAL STEGANOGRAPHIC MODEL

A general model of a steganographic system:

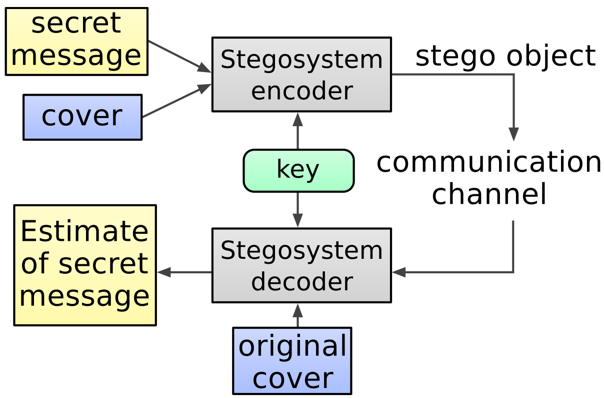


Figure 1: Model of steganographic systems

Steganographic algorithms are in general based on replacing noise component of a digital object with a to-be-hidden message.

Kerckhoffs's principle holds also for steganography. Security of the system should not be based on hiding the embedding algorithm, but on hiding the key.

BASIC CONCEPTS of STEGOSYSTEMS

- **Covertext (cover-data – cover-object)** is an original (unaltered) message.
- **Embedding process** in which the sender, Alice, tries to hide a message by embedding it into a (randomly chosen) covertext, usually using a key, to obtain a stegotext (stego-data or stego-object). The embedding process can be described by the mapping $E : C \times K \times M \rightarrow C$, where C is the set of possible cover – and stegotexts, K is the set of keys, and M is the set of messages.
- **Stegotext (stego-data – stego-object)** is the message that comes out of the embedding process and contains the hidden message.
- **Recovering process** (or extraction process) in which the receiver, Bob, tries to get, using the key only but not the covertext, the hidden message in the stegotext. The recovery (decoding) process D can be seen as a mapping $D : C \times K \rightarrow C$.
- **Security requirement** is that a third person watching such a communication should not be able to find out whether the sender has been active, and when, in the sense that he really embedded a message in the covertext. **In other words, stegotexts should be indistinguishable from covertexts.**

BASIC TYPES of STEGOSYSTEMS

There are three basic types of stegosystems:

- **Pure stegosystems** – no key is used.
- **Secret-key stegosystems** – shared secret key is used.
- **Public-key stegosystems** – public and secret keys are used.

Definition: Pure stegosystem is defined as $S = \langle C, M, E, D \rangle$, where C is the set of possible covertexts, M is the set of secret messages, $|C| \geq |M|$, $E : C \times M \rightarrow C$ is the embedding function and $D : C \rightarrow M$, is the extraction function, with the property that D(E(c,m)) = m, for all $m \in M$ and $c \in C$.

Security of the pure stegosystems depends completely on its secrecy. On the other hand, security of other two stegosystems depends on the secrecy of the key(s) used.

Definition: Secret-key (asymmetric) stegosystem $S = \langle C, M, K, E_K, D_K \rangle$, where C is the set of possible covertexts, M is the set of secret messages with $|C| \geq |M|$, K is the set of secret keys, $E_K : C \times M \times K \rightarrow C$, $D_K : C \times K \rightarrow M$ with the property that $D_K(E_K(c, m, k), k) = m$ for all $m \in M$, $c \in C$ and $k \in K$.

PUBLIC-KEY STEGANOGRAPHY

Similarly as in the case of the public-key cryptography, two keys are used: a public-key E for embedding and a private-key D for recovering.

It is often useful to combine such a public-key stegosystem with a public-key cryptosystem.

For example, in case Alice wants to send a message m to Bob, she encodes first m using Bob's public key e_B, then makes embedding of $e_B(m)$ using process E into a cover and then sends the resulting stegotext to Bob, who recovers $e_B(m)$ using D and then decrypts it, using his decryption function d_B.

TEXT STEGANOGRAPHY

A variety of steganography techniques allow to hide messages in formatted texts.

- **Acrostic.** A message is hidden into certain letters of the text, for example into the first letters of some words.
  Tables have been produced, the first one by Trithentius, called Ave Maria, how to replace plaintext letters by words.
- An improvement of the previous method is to distribute plaintext letters randomly in the cover-text and then use a mask to read it.
- The presence of errors or stylistic features at predetermined points in the cover data is another way to select the location of the embedded information.
- **Line shifting encodings.**
- **Word shifting encodings.**
- **Data hiding through justifications.**
- Through features encoding (for example in the vertical lines of letters b, d, h, k).

Text steganography (a really good one) is considered to be very difficult kind of steganography due to the lack of redundancy in texts comparing to images or audio.

ACROSTIC

Amorosa visione by Giovanni Boccaccio (1313-1375) is said to be the **world largest acrostic.**

Boccaccio first wrote three sonnets (1500 letters together) and then he wrote other poems such that the initials of the successive tercets correspond exactly to the letters of the sonnets.

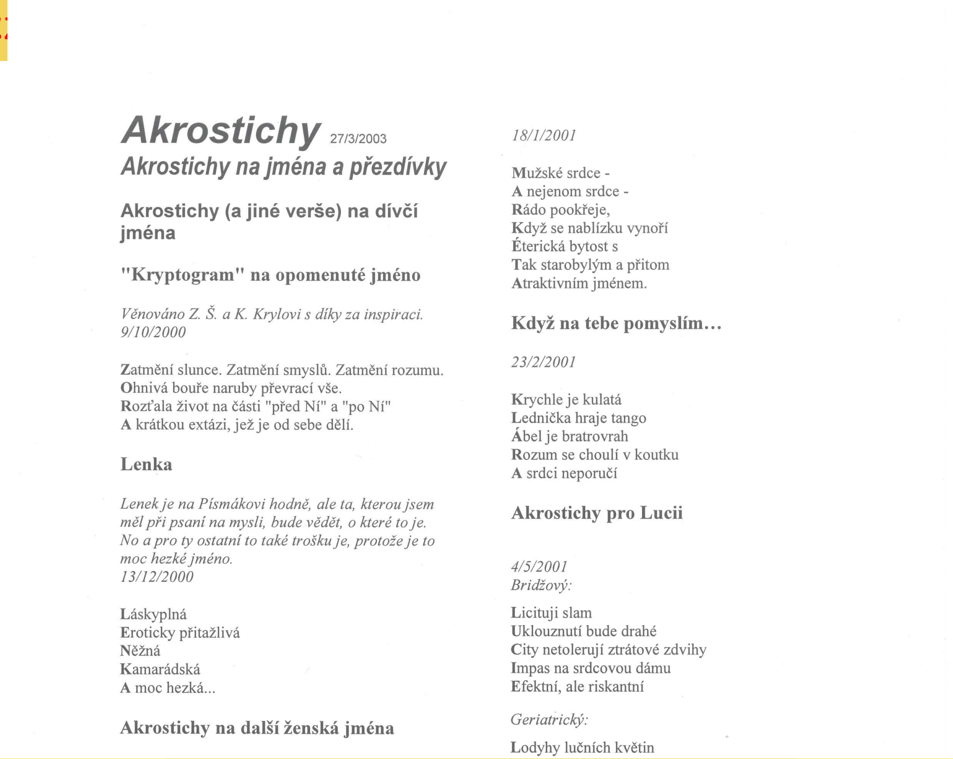
In the book **Hypnerotomachia Poliphili**, published **by an anonymous** in 1499, and considered as one of the most beautiful books ever,the first letters of the 38 chapters spelled out as follows:

Poliam frater Franciscus Columna peramavit

with the translation

Brother Francesco Colonna passionately loves Polia

C


Akrostichy 27/3/2003
Akrostichy na jména a přezdívky

Akrostichy (a jiné verše) na dívčí jména

"Kryptogram" na opomenuté jméno

Věnováno Z. Š. a K. Krylovi s díky za inspiraci.
9/10/2000

Zatmění slunce. Zatmění smyslů. Zatmění rozumu.
Ohnivá bouře naruby převrací vše.
Rozťala život na části "před Ní" a "po Ní"
A krátkou extázi, jež je od sebe dělí.

Lenka

Lenek je na Písmákovi hodně, ale ta, kterou jsem měl při psaní na mysli, bude vědět, o které to je. No a pro ty ostatní to také trošku je, protože je to moc hezké jméno.
13/12/2000

Láskyplná
Eroticky přitažlivá
Něžná
Kamarádská
A moc hezká...

Akrostichy na další ženská jména

18/1/2001

Mužské srdce -
A nejenom srdce -
Rádo pookřeje,
Když se nablízku vynoří
Éterická bytost s
Tak starobylým a přitom
Atraktivním jménem.

Když na tebe pomyslím…

23/2/2001

Krychle je kulatá
Lednička hraje tango
Ábel je bratrovrah
Rozum se choulí v koutku
A srdci neporučí

Akrostichy pro Lucii

4/5/2001
Bridžový:

Licituji slam
Uklouznutí bude drahé
City netolerují ztrátové zdvihy
Impas na srdcovou dámu
Efektní, ale riskantní

Geriatrický:

Lodyhy lučních květin

PERFECT SECRECY of STEGOSYSTEMS

In order to define secrecy of a stegosystem we need to consider

- probability distribution P_C on the set C of covertexts;
- probability distribution P_M on the set M of secret messages;
- probability distribution P_K on the set K of keys;
- probability distribution P_S on the set $\{E_K(c, m, k), | c \in C, m \in M, k \in K\}$ of stegotexts.

The basic related concept is that of the relative entropy $D(P_1 \| P_2)$ of two probability distributions P_1 and P_2 defined on a set Q by

$$D(P_1 \| P_2) = \sum_{q \in Q} P_1(q) lg \frac{P_1(q)}{P_2(q)},$$

which measures the inefficiency of assuming that the distribution on Q is P_2 if it is really P_1.

Definition Let S be a stegosystem, P_C the probability distribution on covertexts C and P_S the probability distribution of the stegotexts and $\varepsilon > 0$. Stegosystem **S** is called – ε-secure against passive attackers, if

$$D(P_C \| P_S) \leq \varepsilon$$

and **perfectly secure** if $\varepsilon = 0$.

PERFECTLY SECURE STEGOSYSTEMS

A perfectly secure stegosystem can be constructed out of the ONE TIME-PAD CRYPTOSYSTEM

Theorem There exist perfectly secure stegosystems.

Proof. Let n be an integer, $C_n = \{0, 1\}^n$ and P_C be the uniform distribution on C_n, and let $m \in C_n$ be a secret message.

The sender selects randomly $c \in C_n$, computes $c \oplus m = s$. The resulting stegotexts are uniformly distributed on C_n and therefore $P_C = P_S$ from what it follows that

$$D(P_{C_n} \| P_S) = 0.$$

In the extraction process, the message m can be extracted from s by the computation

$$m = s \oplus c.$$

INFORMATION HIDING in NOISY DATA

Perhaps the most basic methods of steganography is to utilize the existence of redundant information in communication channels/media.

Images and digital sounds naturally contain such redundancies in the form of noise components.

For images and digital sounds it is natural to assume that a cover-data are represented by a sequence of numbers and their least significant bits (LSB) represent noise.

If cover-data are represented by numbers

$$c_1, c_2, c_3, \ldots,$$

then one of the most basic steganographic methods is to replace, in some of c_i's chosen using an algorithm and a key, the least significant bits by the bits of the message that should be hidden.

Unfortunately, this method does not provide high level of security and it can change significantly statistical properties of the cover-data.

ACTIVE and MALICIOUS ATTACKS

At the design of stegosystems special attention has to be paid to the presence of active and malicious attackers.

- Active attackers can change cover during the communication process.
- An attacker is malicious if he forges messages or initiates a steganography protocol under the name of one of the communicating parties.

In the presence of a malicious attacker, it is not enough that stegosystem is robust.

If the embedding method does not depend on a key shared by the sender and receiver, then an attacker can forge messages, since the recipient is not able to verify sender's identity.

SECURITY of STEGOSYSTEMS - less strong requirements

Definition A steganographic algorithm is called secure if

- Messages are hidden using a public algorithm and a secret key. The secret key must identify the sender uniquely.
- Only the holder of the secret key can detect, extract and prove the existence of the hidden message. (Nobody else should be able to find any statistical evidence of a message's existence.)
- Even if an enemy gets the contents of one hidden message, he should have no chance of detecting others.
- It is computationally infeasible to detect hidden messages.

STEGO – ATTACKS

Stego-only attack: Only the stego-object is available for stegoanalysis.

Known-cover attack: The original cover-object and stego-object are both available.

Known-message attack: Sometimes the hidden message may become known to the stegoanalyser. Analyzing the stego-object for patterns that correspond to the hidden message may be beneficial for future attacks against that system. (Even with the message, this may be very difficult and may even be considered equivalent to the stego-analysis.)

Chosen-stego attack: The stegoanalysis generates a stego-object from some steganography tool or algorithm from a chosen message. The goal in this attack is to determine corresponding patterns in the stego-object that may point to the use of specific steganography tools or algorithms.

Known-stego attack The steganography algorithm is known and both the original and stego-objects are available.

BASIC STEGANOGRAPHIC TECHNIQUES

Substitution techniques: substitute a redundant part of the cover-object with the secret message.

Transformed domain techniques: embed the secret message in a transform space of the signal (e.g. in the frequency domain).

Spread spectrum techniques: embed the secret messages adopting ideas from the spread spectrum communications.

Statistical techniques: embed messages by changing some statistical properties of the cover-objects and use hypothesis-testing methods in the extraction process.

Cover generation techniques: do not embed the message in randomly chosen cover-objects, but create covers that fit a message that needs to be hidden.

DIGITAL COVER DATA

A **cover-object** or, shortly, a **cover c** is a sequence of numbers $c_i, i = 1, 2, \ldots, |c|$.

Such a sequence can represent digital sounds in different time moments, or a linear (vectorized) version of an image.

$c_i \in \{0, 1\}$ in case of binary images and, usually, $0 \leq c_i \leq 256$ in case of quantized images or sounds.

An image C can be seen as a discrete function assigning a color vector c(x,y) to each pixel p(x,y).

A color value is normally a three-component vector in a **color space**. Often used are the following color spaces:

RGB-space – every color is specified as a weighted sum of a red, green and a blue component. A vector specifies intensities of these three components.

YCbCr-space It distinguishes a luminance Y and two chrominance components (Cb, Cr).

Note A color vector can be converted to YCbCr components as follows:

$$Y = 0.299\ R + 0.587\ G + 0.114\ B$$
$$Cb = 0.5 + \frac{(B - Y)}{2}$$
$$Cr = 0.5 + \frac{(R - Y)}{1.6}$$

REPRESENTATION of IMAGES

- Images typically use either 8-bits or 24-bits colors.
- When 8-bits are used the color palette has 256 colors.
- When 24-bits are used each pixel is represented by three primary colors, each represented by an 8-bit.
- The size of an image file is directly related to the number of pixels and granularity of colors.
- A typical 640×480 pix image using 256 colors requires a file of 307 KB.
- A high-resolution 1024×768 pix file with 24-bit color image requires 2.36 MB file.

BASIC SUBSTITUTION TECHNIQUES

- **LSB substitution** – the LSB of an binary block c_{k_i} is replaced by the bit m_i of the secret message.

  The methods differ by techniques how to determine k_i for a given i.

  For example, $k_{i+1} = k_i + r_i$, where r_i is a sequence of numbers generated by a specific pseudo-random generator.
- Substitution into parity bits of blocks. If the parity bit of block c_{k_i} is m_i, then the block c_{k_i} is not changed; otherwise one of its bits is changed.
- Substitution in binary images. If image c_i has more (less) black pixels than white pixels and $m_i = 1 (m_i = 0)$, then c_i is not changed; otherwise the portion of black and white pixels is changed (by making changes at those pixels that are neighbors of pixels of the opposite color).
- Substitution in unused or reserved space in computer systems.

LSB SUBSTITUTION in IMAGES - EXAMPLE

As already mentioned, representation of images usually use for each pixel either 8-bit representation of a palette of 256 colors, or 24-bit representation of three bytes representing RGB coloring.

Example: Let LSB technique be used to hide "101101101" in RGB representation of three pixels:

10010101 00001101 11001001

10010110 00001111 11001010

10011111 00010000 11001011

The outcome will be the following representation of these three pixels

10010101 00001100 11001001

10010111 00001110 11001011

10011111 00010000 11001011

Observe that actually only 4 LSB have been changed – less than 50%

CAT in a TREE

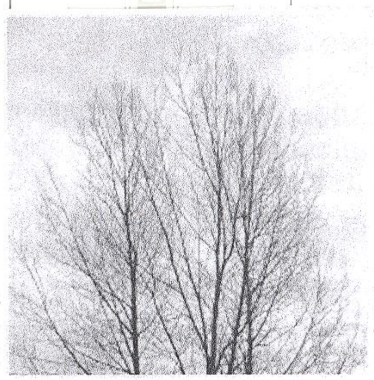


Image of a tree with a steganographically hidden image. The hidden image is revealed by removing all but the two least significant bits of each color component and a subsequent normalization.

Image of a cat extracted from the tree image above.

PROFESSIONAL EMBEDDINGS

Cover figure and stego figure:

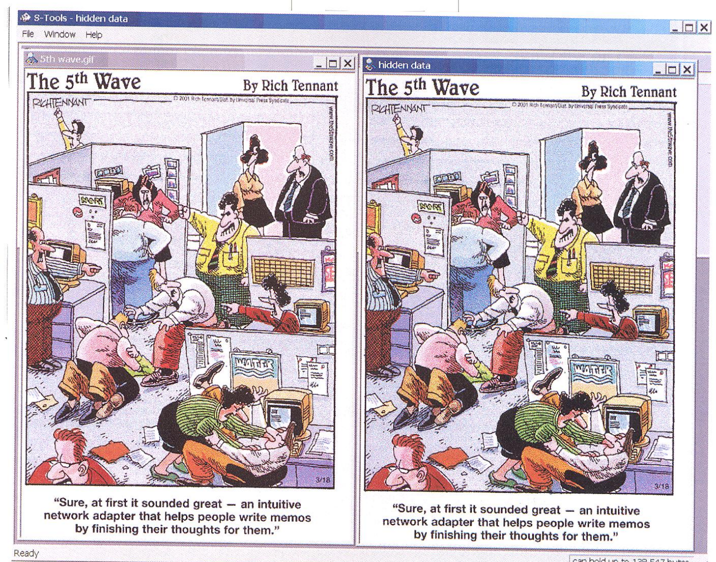


FIGURE 2. The original image file (left) and image file with embedded text (right), side by side.

LSB SUBSTITUTION PLUSES and MINUSES

Bits for substitution can be chosen (a) randomly; (b) adaptively according to local properties of the digital media that is used.

Advantages:

(a) LSB substitution is the simplest and most common stego technique and it can be used also for different color models.

(b) This method can reach a very high capacity with little, if any, visible impact to the cover digital media.

(c) It is relatively easy to apply on images and radio data.

(d) Many tools for LSB substitutions are available on the internet

Disadvantages:

(a) It is relatively simple to detect the hidden data;

(b) It does not offer robustness against small modifications (including compression) at the stego images.

AUDIO STEGANOGRAPHY

Audio based staganography has several advantages:

- Audio files are generally larger than images.
- Our hearing can be easily fooled.
- Slight changes in amplitudes can store vast amounts of information.

Examples of audio steganography:

- Echo hiding embeds data by creating an artificial echo to the source audio.
- Phase hiding of data.

SHOW EXAMPLE: !!!!!!!!!!!!!!!!!!!!!!!!

ROBUSTNESS of STEGANOGRAPHY

Steganographic systems are extremely sensitive to cover modifications, such as

- image processing techniques (smoothing, filtering, image transformations, ...);
- filtering of digital sounds;
- compression techniques.

Informally, a stegosystem is **robust** if the embedded information cannot be altered without making substantial changes to the stego-objects.

Definition: Let S be a stegosystem and P be a class of mappings $C \to C$. S is P-robust, if for all $p \in P$

$$D_K(p(E_K(c,m,k)),k) = D_K(E_K(c,m,k),k) = m$$

in the case of a secret-key stegosystem and

$$D(p(E(c,m))) = D(E(c,m)) = m$$

in the case of pure stegosystem, for any message m, cover c, and key k.

- There is a clear tradeoff between *security* and *robustness*.
- Some stegosystems are designed to be robust against a specific class of mappings (for example JPEG compression/decompression).
- There are two basic approaches to make stegosystems robust:
  - By foreseeing possible cover modifications, the embedding process can be robust so that possible modifications do not entirely destroy embedded information.
  - Reversing operations that has been made by an active attacker.

STEGANALYSIS - ART of DETECTING HIDDEN MESSAGES

The main goal of a passive attacker is to decide whether data sent to Bob by Alice contain secret message or not.

The detection task can be formalized as a statistical hypothesis-testing problem with the test function $f : C \to \{0,1\}$:

$$f(c) = \begin{cases} 1, & \text{if } c \text{ contains a secret message;} \\ 0, & \text{otherwise} \end{cases}$$

There are two types of errors possible:

- Type-I error - a secret message is detected in data with no secret message;
- Type-II error - a hidden secret message is not detected

In the case of ε-secure stegosystems there is well know relation between the probability β of the type II error and probability α of the type I error.

Let S be a stegosystem which is ε-secure against passive attackers, β the probability that the attacker does not detect a hidden message and α the probability that the attacker falsely detects a hidden message. Then

$$d(\alpha,\beta) \leq \varepsilon,$$

where $d(\alpha,\beta)$ is the binary relative entropy defined by

$$d(\alpha,\beta) = \alpha \lg \frac{\alpha}{1-\beta} + (1-\alpha)\lg \frac{1-\alpha}{\beta}.$$

NETWORK STEGANOGRAPHY

Network steganography utilizes communication protocol's elements and their basic functionality as a cover for hidden data.

Typical network steganography methods involve modification of the properties of a single network protocol or a relation between several network protocols to enable secret communication.

A use of network steganography is usually very hard to detect.

WATERMARKING

WATERMARKING

Historically, a (physical) watermark is the replication of an image, logo, or text on paper stock so that the source of the document can be, at least partially, authenticated.

Digital watermarking is a process of embedding information (a digital watermark) into digital data (called often signal) which may be used to verify its authenticity or the identity of its owner. This should be done in such a way that if a signal is coppied so is the embedded watermark.

DIGITAL WATERMARKING

Digital watermarking seems to be a promising technique to deal with the following problem:

Problem Digitalization allows to make unlimited number of copies of intellectual products (books, art products, music, video,...). How to make use of this enormous potential digitalization has and, at the same time, to protect intellectual rights of authors (copyrights, protection against modifications and insertion into other products), in a way that is legally accepted?

Solution Digital watermarking tries to solve the above problem using a variety of methods of informatics, cryptography, signal processing, ... and in order to achieve that tries to insert specific information (watermarks) into data/carrier/signal in such a way that watermarks cannot be extracted or at least detected and if data with one or several watermarks are copied, watermarks should not change.

BASIC APPLICATIONS

- **Copyright protection - ownership assertion** For example, if a watermark is embedded into a music (or video) product, then each time music (video) is played in public information about author is extracted and tandem are established. Another example: annotation of digital photographs
- **Source tracing.** Watermarks can be used to trace or verify the source of digital data.
- **Insertion of additional (sensitive) information** For example, personal data into röntgen photos r of keywords into multimedia products.

ROBUSTNESS of STEGANOGRAPHY

Steganographic systems are extremely sensitive to cover modifications, such as

- image processing techniques (smoothing, filtering, image transformations, ...);
- filtering of digital sounds;
- compression techniques.

Informally, a stegosystem is **robust** if it resists a designated class of transformations.

Definition Let S be a stegosystem and P be a class of mappings $C \rightarrow C$. S is P-robust, if for all $p \in P$

$$D_K(p(E_K(c, m, k)), k) = D_K(E_K(c, m, k), k) = m$$

in the case of a secret-key stegosystem and

$$D(p(E(c, m))) = D(E(c, m)) = m$$

in the case of pure stegosystem, for any m, c, k.

- There is a clear tradeoff between *security* and *robustness*.
- Some stegosystems are designed to be robust against a specific class of mappings (for example JPEG compression/decompression).
- There are two basic approaches to make stegosystems robust:
  - By foreseeing possible cover modifications, the embedding process can be robust so that possible modifications do not entirely destroy embedded information.
  - Reversing operations that has been made by an active attacker.

DETECTING SECRET MESSAGES

The main goal of a passive attacker is to decide whether data sent to Bob by Alice contain secret message or not.

The detection task can be formalized as a statistical hypothesis-testing problem with the test function $f : C \rightarrow \{0, 1\}$:

$$f(c) = \begin{cases} 1, & \text{if } c \text{ contains a secret message;} \\ 0, & \text{otherwise} \end{cases}$$

There are two types of errors possible:

- Type-I error - a secret message is detected in data with no secret message;
- Type-II error - a hidden secret message is not detected

In the case of ε-secure stegosystems there is the following relation between the probability β of the type II error and probability α of the type I error.
Let S be a stegosystem which is ε-secure against passive attackers, β the probability that the attacker does not detect a hidden message and α the probability that the attacker falsely detects a hidden message. Then

$$d(\alpha, \beta) \leq \varepsilon,$$

where $d(\alpha, \beta)$ is the binary relative entropy defined by

$$d(\alpha, \beta) = \alpha \lg \frac{\alpha}{1 - \beta} + (1 - \alpha) \lg \frac{1 - \alpha}{\beta}.$$

HISTORY of WATERMARKING

Paper watermarks appeared in the art of handmade paper marking 700 hundred years ago.

Watermarks were mainly used to identify the mill producing the paper and paper format, quality and strength.

Paper watermarks was a perfect technique to eliminate confusion from which mill paper is and what are its parameters.

Legal power of watermarks has been demonstrated in 1887 in France when watermarks of two letters, presented as a piece of evidence in a trial, proved that the letters had been predated, what resulted in the downfall of a cabinet and, finally, the resignation of the president Grévy.

Paper watermarks in bank notes or stamps inspired the first use of the term watermark in the context of digital data.

The first publications that really focused on watermarking of digital images were from 1990 and then in 1993.

EMBEDDING and RECOVERY SYSTEMS

in WATERMARKING SYSTEMS

Figure 2 shows the basic scheme of the watermarks embedding systems.

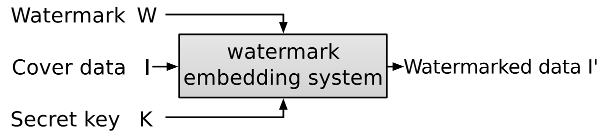


Figure 2: Watermark embedding scheme

Inputs to the scheme are the watermark, the cover data and an optional **public or secret** key. The output are watermarked data. The key is used to enforce security.
Figure 3 shows the basic scheme for **watermark recovery schemes.**

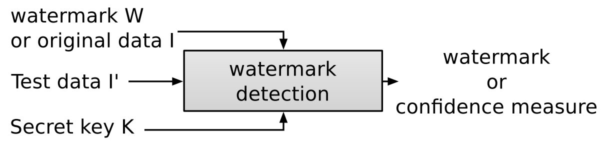


Figure 3: Watermark recovery scheme

Inputs to the scheme are the **watermarked data**, the **secret or public key** and, depending on the method, the **original data and/or the original watermark**. The **output** is the **recovered watermark W** or some kind of **confidence measure indicating how likely it is for the given watermark at the input to be present in the data.**

TYPES of WATERMARKING SCHEMES

Private (non-blind) watermarking systems require for extraction/detection the original cover-data.

- Type I systems use the original cover-data to determine where a watermark is and how to extract the watermark from stego-data.
- Type II systems require a copy of the embedded watermark for extraction and just yield a yes/no answer to the question whether the stego-data contains a watermark.

Semi-private (semi-blind) watermarking does not use the original cover-data for detection, but tries to answer the same question. (Potential application of blind and semi-blind watermarking is for evidence in court ownership,...)

Public (blind) watermarking – neither cover-data nor embedded watermarks are required for extraction – this is the most challenging problem.

SECRET SHARING by SECRET HIDING

A simple technique has been developed, by Naor and Shamir, that allows for a given n and $t < n$ to hide any secret (image) message m in images on transparencies in such away that each of n parties receives one transparency and

- no t - 1 parties are able to obtain the message m from the transparencies they have.
- any t of the parties can easily get (read or see) the message m just by stacking their transparencies together and aligning them carefully.

APPENDIX

APPENDIX

WATERMARKS

Historically, a watermark is a replication of an image, logo, or text on paper stock so that the source of the document can be, at least partially, determined.

STEGANOGRAPHY TOOLS

There are a number of software packages that perform steganography on just about any software platform.

They usually hide information in image or audio files.

In case of images, systems gets as input an image and text to be hidden (and key) and provide stego-image hiding a given text.

The intended receiver who knows the key takes corresponding steganalysis tool and for a given stego-image and stego-key gets the hidden data/message.

SIGNAL PROCESSING TERMINOLOGY

In some applications of steganography the following signal processing technology is used.

- **Payload** - message to be secretly communicated;
- **Carrier** - data file or signal into which payload is embedded
- **Package - stego file - covert message** - the outcome of embedding of payload into carrier.
- **Encoding density** - the percentage of bytes or other signal elements into which the payload is embedded.

TO REMEMBER !!!

There is no use in trying, she said: one cannot believe impossible things.

I dare to say that you have not had much practice, said the queen,

When I was your age, I always did it for half-an-hour a day and sometimes I have believed as many as six impossible things before breakfast.

Lewis Carroll: Through the Looking-glass, 1872