Steganography using text embedding in sound files

Steganography is the art and science of writing hidden messages in such a way that no one apart from the intended recipient knows of the existence of the message; this is in contrast to cryptography, where the existence of the message itself is not disguised, but the content is obscured. The word " Steganography" is of Greek origin and means " covered, or hidden writing". Its ancient origins can be traced back to 440 BC. Herodotus mentions two examples of Steganography in The Histories of Herodotus.

Demeratus sent a warning about a forthcoming attack to Greece by writing it on a wooden panel and covering it in wax. Wax tablets were in common use then as re-usable writing surface, sometimes used for shorthand. Another ancient example is that of Histiaeus, who shaved the head of his most trusted slave and tattooed a message on it. After his hair had grown the message was hidden. The purpose was to instigate a revolt against the Persians. Later, Johannes Trithemius's book Steganographia is a treatise on cryptography and steganography disguised as a book on black magic.

Generally, a steganographic message will appear to be something else: a picture, an article, a shopping list, or some other message - the covertext. Classically, it may be hidden by using invisible ink between the visible lines of innocuous documents, or even written onto clothing. During World War II a message was once written in morse code along two-colored knitting yarn[citation needed]. Another method is invisible ink underlining, or simply pin pricking of individual letters in a newspaper article, thus forming a message.

It may even be a few words written under a postage stamp, the stamp then being the covertext. The advantage of steganography over cryptography

alone is that messages do not attract attention to themselves, to messengers, or to recipients. An unhidden coded message, no matter how unbreakable it is, will arouse suspicion and may in itself be incriminating, as in some countries encryption is illegal. Steganography uses in electronic communication include steganographic coding inside of a transport layer, such as an MP3 file, or a protocol, such as UDP.

Steganographic messages are often first encrypted by some traditional means, and then a covertext is modified in some way to contain the encrypted message, resulting in stegotext. For example, the letter size, spacing, typeface, or other characteristics of a covertext can be manipulated to carry the hidden message; only the recipient (who must know the technique used) can recover the message and then decrypt it. Francis Bacon is known to have suggested such a technique to hide messages citation needed.

Modern Steganographic Techniques Concealing messages within the lowest bits of noisy images or sound files. Concealing data within encrypted data . The data to be concealed is first encrypted before being used to overwrite part of a much larger block of encrypted data. This technique works most effectively where the decrypted version of data being overwritten has no special meaning or use. Examples of software that use this technique include Free OTFE and True Crypt.

Embedded pictures in video material (optionally played at slower or faster speed). A new steganographic technique involves injecting imperceptible delays to packets sent over the network from the keyboard. Delays in keypresses in some applications (telnet or remote desktop) can mean a

delay in packets, and the delays in the packets can be used to encode data. There is no extra processor or network activity, so the steganographic technique is " invisible" to the user. This kind of steganography could be included in the firmware of keyboards, thus making it invisible to the system.

The firmware could then be included in all keyboards, allowing someone to distribute a keylogger program to thousands without their knowledge. Historical Steganographic Techniques Steganography has been widely used in historical times, especially before cryptographic systems were developed. Examples of historical usage include: Hidden messages in wax tablets: in ancient Greece, people wrote messages on the wood, then covered it with wax so that it looked like an ordinary, unused tablet.

Hidden messages on messenger's body: also in ancient Greece. Herodotus tells the story of a message tattooed on a slave's shaved head, hidden by the growth of his hair, and exposed by shaving his head again. The message, if the story is true, carried a warning to Greece about Persian invasion plans. Hidden messages on paper written in secret inks under other messages or on the blank parts of other messages. During and after World War II, espionage agents used microdots to send information back and forth.

Since the dots were typically extremely small the size of a period produced by a typewriter (perhaps in a font with 10 or 12 characters per inch) or even smaller the stegotext was whatever the dot was hidden within. If a letter or an address, it was some alphabetic characters. If under a postage stamp, it was the presence of the stamp. The problem with the WWII microdots was that they needed to be written in a special ink, easily detected by holding a suspected paper up to a light and viewing it almost edge on.

The microdot ink would shine when viewed from an extreme angle while the normal ink would not. More obscurely, during World War II, a spy for the Japanese in New York City, Velvalee Dickinson, sent information to accommodation addresses in neutral South America. She was a dealer in dolls, and her letters discussed how many of this or that doll to ship. The stegotext in this case was the doll orders; the 'plaintext' being concealed was itself a codetext giving information about ship movements, etc. Her case became somewhat famous and she became known as the Doll Woman.

In the manga Lone Wolf and Cub, a main plot device is a Yagy? letter which has a message written in mulberry extract. They see the message by placing it in a bowl of silkworms and seeing where they eat. The one-time pad is a theoretically unbreakable cipher that produces ciphertexts indistinguishable from random texts: only those who have the private key can distinguish these ciphertexts from any other perfectly random texts. Thus, any perfectly random data can be used as a covertext for a theoretically unbreakable steganography.

A modern example of OTP: in most cryptosystems, private symmetric session keys are supposed to be perfectly random (i. e. generated by a good RNG). Even very weak ones (e. g. shorter than 128 bits). This means that users of weak crypto (in countries where strong crypto is forbidden) can safely hide OTP messages in their session keys. ? Rumored Usage in Terrorism The rumors about terrorists using steganography started first in the daily newspaper USA Today on February 5, 2001.

The articles are still available online, and were titled " Terrorist instructions hidden online", and the same day, " Terror groups hide behind Web

encryption". In July of the same year, the information looked even more precise: " Militants wire Web with links to jihad". A citation from the USA Today article: " Lately, al-Qaeda operatives have been sending hundreds of encrypted messages that have been hidden in files on digital photographs on the auction site eBay. com". These rumors were cited many times - without ever showing any actual proof - by other media worldwide, especially after the terrorist attack of 9/11.

For example, the Italian newspaper Corriere della Sera reported that an Al Qaeda cell which had been captured at the Via Quaranta mosque in Milan had had pornographic images on their computers, and that these images had been used to hide secret messages (although no other Italian paper ever covered the story). The USA Today articles were written by veteran foreign correspondent Jack Kelley, who in 2004 was fired after allegations emerged that he had fabricated stories and invented sources.

In October 2001, the New York Times published an article claiming that l-Qaeda had used steganographic techniques to encode messages into images, and then transported these via email and possibly via USENET to prepare and execute the September 11, 2001 Terrorist Attack. Despite being dismissed by security experts, the story has been widely repeated and resurfaces frequently. It was noted that the story apparently originated with a press release from " iomart" , a vendor of steganalysis software. No corroborating evidence has been produced by any other source.

Moreover, a captured al-Qaeda training manual makes no mention of this method of steganography. The chapter on communications in the al-Qaeda manual acknowledges the technical superiority of US security services, and

generally advocates low-technology forms of covert communication. The chapter on " codes and ciphers" places considerable emphasis on using invisible inks in traditional paper letters, plus simple ciphers such as simple substitution with nulls; computerized image steganography is not mentioned.

Nevertheless public efforts were mounted to detect the presence of steganographic information in images on the web (especially on eBay, which had been mentioned in the New York Times article). To date these scans have examined millions of images without detecting any steganographic content (see " Detecting Steganographic Content on the Internet" under external links), other than test images used to test the system, and instructional images on web sites about steganography. Counter measures The detection of steganographically encoded packages is called steganalysis. The simplest method to detect modified files, however, is to compare them to the originals. To detect information being moved through the graphics on a website, for example, an analyst can maintain known-clean copies of these materials and compare them against the current contents of the site. The differences (assuming the carrier is the same) will compose the payload.

In general, using an extremely high compression rate makes steganography difficult, but not impossible; while compression errors provide a good place to hide data, high compression reduces the amount of data available to hide the payload in, raising the encoding density and facilitating easier detection (in the extreme case, even by casual observation). There is an obvious need to communicate or to keep certain information unknown to the public or to

anyone except those, who are intended to know that information. Cryptography is the process of achieving this result. ? Encryption

Cryptographic processes are changing the form of the original information (the plain text) so that it becomes unintelligible for anybody except for those, who are intended to understand it. This process is called encoding. The result of the encoding is the cipher text. The process of restoring the original form of the information is the decoding or deciphering. So-called secure methods of encoding and decoding digital information use some information in addition to the original data; this is the cryptographic key, or simply the key. The knowledge of the key is crucial for the decoding process.

Everyone who knows this key (and the applicable method) is able to restore the original data. One of the basic problems of digital cryptographic methods is the communication of the key. If the primary goal of encoding is to communicate the information with someone to the exclusion of all others, then communicating the key itself becomes crucial. This is especially important in cases of those, who can communicate only " openly", for example via Internet. In such cases an asymmetric encoding method will be adopted, using a private/public key pair; the encoding process uses a different key than the decoding.

This method does not require the communication of the complete key itself, but it has following disadvantages: when encoding the original data, the key of the receiver has to be known, it is prone to be broken methodically: a key of this kind, which is " unbreakable" today may be broken tomorrow without unduely large computing capacity. This method will be used typically in electronic commerce, where the partners (the customer and the

seller/provider) usually don't have any other communication with each other, except the process of purchasing.

If the cryptographic strength of the encoding (i. e. the secrecy) is more important than the ease of communication, or the goal of encoding is not (or not only) communicating but storing secret data, then symmetric encoding is better than the private/public key scheme: the same key will be used in the encoding and in the decoding process. The cryptographic strength of this method can be as high as required. The disadvantage of this method is in communication, namely that the key itself too has to be communicated somehow.