DOI:10.1145/2566590.2566610

Methods for embedding secret data are more sophisticated than their ancient predecessors, but the basic principles remain unchanged.

BY ELŻBIETA ZIELIŃSKA, WOJCIECH MAZURCZYK, AND KRZYSZTOF SZCZYPIORSKI

Trends in Steganography

THE MASS MEDIA anointed 2011 as the "year of the hack"[23] due to the numerous accounts of data security breaches in private companies and governments. Indeed, the sheer volume of stolen data was estimated in petabytes (that is, millions of gigabytes).[2]

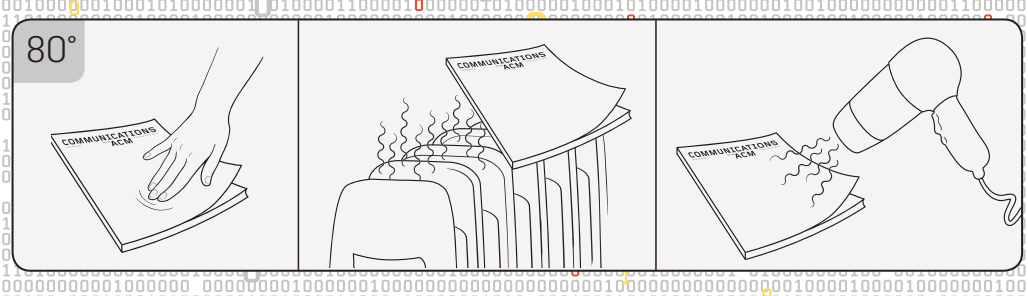
A large fraction of the security breaches that year could be attributed to the so-called Operation Shady RAT.[49] These actions were targeted at numerous institutions around the world and the inflicted damage lasted, in many cases, for months. The mechanism of infection was mainly by means of conning an unaware user to open a specially crafted email message (phishing) and implanting a back door on the victim's computer. The next step was to connect to a website and download files that only seemed to be legitimate HTML or JPEG files. What cybercriminals had actually done was encode commands into pictures or crafted Web pages so they were invisible to unaware third parties, and smuggled them through firewalls into the system under attack. These control commands then ordered a victim's computer to obtain executable code from remote servers, which in turn permitted an outsider to gain access to local files on the compromised host.[32] In numerous cases, the side channel to the confidential resources remained accessible for months, thus deeming the security breach severe. The villains were so daring they did not even put much effort into obscuring the fact that information hiding techniques were involved in the attack. One of the pictures used as a vector for control commands was the famous "Lena," a cropped picture of a Playboy model, which is the standard test image for any digital image processing or steganographic algorithm.

It is very likely we are witnessing the birth of a whole new breed of malware. It all started with the discovery of the well-known Stuxnet[18] computer worm in June 2010 that stirred increased attention because it had been targeted specifically to affect Iranian nuclear power plants.[14] In September 2011 a new worm, called Duqu, was discovered, and it seems to be closely related to Stuxnet.[7] The general characteristics of the malware's structure are the same in both cases, however, unlike its predecessor, Duqu is oriented at gathering information on the infected system. The most stunning intricacy in Duqu's functioning is its employment of special means for transferring the obtained data to the command and con-

» key insights

- **Steganography, or camouflaging the presence of hidden messages in legitimate carriers, has recently become a tool of the trade for malware suppliers, as proven by the recent attacks on major global targets.**
- **Despite the fact steganography has been known for centuries, it recently has proliferated new grounds—digital media, computer networks, and popular telecommunications services.**
- **There is no miracle solution for the abuse of steganography, other than the meticulous search for any loophole that might be exploited for the purpose of embedding of illicit information, or any means of altering the potential carrier in a manner escaping human perception.**

ART BY ANDRIJ BORYS ASSOCIATES


80°
COMMUNICATIONS ACM
COMMUNICATIONS ACM
COMMUNICATIONS ACM

trol centers of the malware's authors. The captured information is hidden in seemingly innocent pictures and traverses the global network as ordinary files, without raising any suspicion.[21,51] A similar functioning mechanism was found in a new variant of the Alureon malware,[3] which was also discovered around the same time.

These facts indicate that in today's world of digital technologies, it is easily imaginable that the carrier, in which secret data is embedded, was not necessarily an image or Web page source code, but may have been any other file type or organizational unit of data—for example, a packet or a frame—that naturally occurs in computer networks. However, we emphasize the process of embedding secret information into an innocent-looking carrier is not some recent invention—it has been known and used for ages by humankind. This process is called *steganography* and its origins can be traced back to ancient times. Moreover, its importance has not decreased since its birth.

Among steganography's applications is providing means for conducting clandestine communication. The purpose of establishing such information exchange may vary; possible uses can fall into the category of legal or illicit activity. Frequently, the illegal aspect of steganography is accentuated, starting from the obvious criminal communication, through information leakage from guarded systems, cyber weapon exchange, up to industrial espionage. On the other side of the spectrum lie legitimate uses, which include circumvention of Web censorship and surveillance,[55] computer forensics (tracing and identification), and copyright protection (watermarking images, broadcast monitoring).

Stuxnet, Duqu, Alureon, and Shady RAT are merely examples of what is becoming a daily routine for security experts. What should ring the alarm bells is the incorporation of steganography into the already versatile armory of rouge hackers. It can be concluded that steganography is becoming the new black among Black Hats.

The inverse of steganography—steganalysis, which concentrates on the detection of covert communication—started to surface fairly recently, what is reflected in the proportion of available software tools concerning information hiding. Programs for the embedding of data considerably outnumber those dedicated to the detection and extraction of embedded content. The largest commercial database of steganographic tools contains 1,025 applications (as of February 2012),[4] up from 111 tools mentioned by Hayati et al.[24] as of year 2007.

Let us take a closer look at the evolution of this technique, with special attention directed toward the class of methods falling in the category of network steganography.

Recent Cases of Steganography Usage

Besides the previously noted cases of steganographic methods' utilization, in the last decade, one can observe intensive research effort related to steganography and its detection methods (steganalysis). This has been caused by two facts: first, industry's and business's interest in DRM (Digital Rights Management) and second, the alleged utilization of the steganographic methods by terrorists while planning the attacks on U.S. on September 11, 2001.[29,45] It is claimed that the rouge organization used images to conceal instructions regarding the plot, which were then posted on publicly available websites.[45] It seems that such communication could have passed unnoticed for as long as three years.[28]

Recent findings suggest that steganography is presently exploited, mainly for illicit purposes.[29,44,50,57] Robin Bryant[11] recollects the case of "Operation Twins," which culminated in 2002 with the capture of criminals associated with the "Shadowz Brotherhood," a pedophile organization responsible for distribution of child pornography with the aid of steganography.

The mushrooming incidents involving the use of information hiding had triggered an official recognition of the problem. In the 2006 Federal Report,[38] steganography had been named among the major threats of the present-day networks, whose significance is predicted to increase. One of the solutions to alleviate the risks connected with this technique is to become acquainted with the evolution of steganography and, consequentially, predict its further development. This need has been recognized by the academic world in the early 1980s, when steganography started to gain popularity.

Steganographic methods have also proven to be useful tools for data exfiltration, for example, in 2008 it was reported[1] that someone at the U.S. Department of Justice smuggled sensitive financial data out of the agency by embedding the data in several image files.

In 2010, the revealing of a Russian spy ring of the so-called 'illegals,' proved that steganography can pass unnoticed for much longer. The compromised group used digital image steganography to leak classified information from U.S. to Moscow.[44]

Steganography and Its Relationship to Cryptography

Steganography, frequently interchangeably and incorrectly referred to as information hiding, is the art of embedding secret messages (steganograms) in a certain carrier, possibly to communicate them in a covert manner. The border between the two fields cannot be visibly demarcated as their definitions are elusive, and there is a lack of a coherent classification of the invented clandestine communication methods, attributing them to specific domains of steganography or information hiding. The arising misconceptions may be attributed to the recent surge of interest in steganography observed in the mass media, which had only shed light on a small fraction of the available techniques. The reports of espionage and terrorist activities emerging during the last decade[29,44,50,57] have mainly promoted these information hiding techniques, which are associated with the Internet and digital image steganography.

The question remains: How to provide rules to distinguish what belongs to the spectrum of steganographic methods? This can be done by means of providing certain conditions that must be fulfilled to consider something steganography. These may be expressed as follows:

- The information undergoing such hidden transmission is embedded in a seemingly innocent carrier, serving as camouflage for the hidden content.
- The purpose of applying a steganographic technique is to communicate information in a covert manner.

▸ The secrecy of the communication is guaranteed primarily by the camouflaging capability of the algorithm applied to the utilized carrier, and how well the processed data blends in with the whole bulk of legitimate entities of the cover (without embedding); this is understood as the capability to withstand detection attempts, which may rely on statistical analysis of the captured traffic or perceptual analysis of a suspicious message.

The best carrier for secret messages must possess two features. Firstly, it should be popular, that is, the usage of such a carrier should not itself be considered an anomaly. Secondly, the steganogram insertion-related modifications of the carrier should not be "visible" to the third party not aware of the steganographic procedure. Thus, if the embedding of additional information causes degradation of the carrier, then their severity should be limited to a level that would not cause suspicion.

Steganography is not only limited to concealing the fact that a message is being sent, and if not detected, make the sender and receiver "invisible." It should also provide anonymity and privacy, which become understandable desires in modern societies. Obviously, the anonymity potential of steganography, while it can be considered as beneficial in the context of protecting privacy, poses a new type of threat to individuals, societies, and states. The trade-off between the benefits and threats involves many complex ethical, legal, and technological issues. In this article, we only consider the latter.

Steganography is often confused with cryptography, due to their common purpose of providing confidentiality. The difference becomes visible once the etymology of these words is known. Steganography is derived from the Greek: "covered writing," whereas cryptography stands for "secret writing." While the first describes the techniques to create a hidden communication channel, the latter is a designation of ongoing overt message exchange, where the informative content is unintelligible to unauthorized parties. To summarize, it is either the method to establish a communication channel that is kept confidential, or the message itself. Either way, the goal of protecting information from disclosure remains common for both techniques—it is the means that permits us to differentiate between the two. Table 1 summarizes differences between cryptography and steganography, and Table 2 summarizes the relationship between steganography and watermarking.

In spite of the common historical background of the stated communication protection methods, only cryptography has managed to sustain an invariably strong position. Steganography experienced its golden age in the times of ancient Greece and Rome, to be gradually marginalized as time passed. Intuitively, any method of protection that relies on its own confidentiality to provide the secrecy of communication, should not be publicized. Therefore, very few accounts proving contemporary exploitation of steganography can be found, which does not point to the conclusion that it is a neglected scientific discipline.

The Origins of Steganography

The inspiration of steganography is strongly related to phenomena observable in the animal and plant kingdoms. Evolution proved long ago that impersonation is good protection and capacitates survival for numerous species. The ability to camouflage one's presence by means of adopting the characteristics of another living organism is referred to as mimicry. This

Table 1. Comparison of characteristics of steganography and cryptography.

		Cryptography	Steganography
Goal		Obfuscate the content of communication	Hide the fact of communication
Characteristics	Secrecy	Ciphertext is illegible	Embedded information is "invisible" to an unaware observer
	Security of communication	Relies on the confidentiality of the key	Relies on the confidentiality of the method of embedding
	Warranty of robustness	Complexity of the ciphering algorithm	Perceptual invisibility/statistical invisibility/compliance with protocol specification
	Attacks	Detection is easy/extraction is complex	Detection is complex/extraction is complex
Countermeasures	Technical	Reverse engineering	Constant monitoring and analysis of exchanged data
	Legal	Cryptography export laws	Rigid device/protocol specification

Table 2. Comparison of characteristics of steganography and watermarking.

		Watermarking	Steganography
Goal		Protect the carrier	Protect secret information from disclosure
Characteristics	Secrecy	Invisibility or perceptual visibility depending on the requirements	Embedded information is "invisible" to an unaware onlooker
	Type of robustness	Robustness against tampering or removal	Robustness against detection
	Effect of signal processing/ random errors/ compression	Must not lead to the loss of the watermark	May lead to the loss of hidden data
	Type of carrier	Digital files—audio, video, text, or images	Any service, protocol, file, environment employing digital representation of data

capability permits certain organisms to improve their chances for survival. The ancient Greeks, who sought inspiration in nature, had considered the ability to simulate its ways as a measure of craftsmanship. Inherently, the ancient people picked ordinary objects as a carrier for the secret message. The vector physical object, possibly even a living organism, had to be transported from one participant of communication to the other without raising suspicion on the way.

It should not be surprising that the first written report of the use of steganography is attributed to the Greek historian Herodotus. The reported method involved camouflaging a secret message within a hare corpse.[16] The animal was meant to imitate a game trophy and was carried by a man disguised as a huntsman. In this way, a message could be passed without raising unnecessary suspicion.

The most notable method quoted in the historian's works is the communication on wooden tablets—these were usually coated with a thin layer of wax, on which text would be embossed. Clandestine passing of information with the aid of such medium could be achieved if the text was carved permanently on the wood (the carrier of the steganograms), and then coated with wax. Such object would then be passed as an unused tablet, and only an aware recipient would know that the letters would become visible if the wax coat was melted.

The Greek methods were fairly easy to implement as they relied on common patterns—the messages that were passed utilized a carrier cover that could be considered common at the time. Alongside the progress of human civilization and of the way people communicated, new opportunities arose. The popularization of parchment, which substituted papyrus, brought about a new cover for steganograms. Its popularity led to the development of complementary steganographic algorithms, capable of exploiting the new cover's properties. Pliny the Elder is considered the inventor of sympathetic inks,[47] as he postulated the use of thithymallus plant's sap to write text, which would become invisible upon drying. A subtle heating process would lead to the charring of the organic substances contained in the ink, which would then turn brown.

The best carrier for secret messages must possess two features. Firstly, the carrier should be popular. Secondly, the steganogram insertion-related modifications of the carrier should not be "visible" to the third party not aware of the steganographic procedure.

The common factor of all of the aforementioned techniques is the operation of adding surplus content (additional features) to a carrier, which otherwise would not physically contain the inserted elements.

A different type of steganography invented in ancient Rome is the semagram, or a secret message that does not take written form. Tacitus, the historiographer of the ancient world, became interested in the Astragali,[48] which were small dice made of bone. Such objects could be threaded onto a string, where the placement of the holes could be attributed meaning. A properly crafted object would pass unnoticed as a toy.

The Medieval Ages had brought about major progress in the art of information hiding. The Chinese invention of paper, upon its introduction to Europe in the Middle Ages, had brought forth the necessity of differentiating between different manufacturers' products. This is how paper watermarking was born.[43] Today, digital watermarking and digital image steganography are based on the same principle. It should be stressed that file watermarking is now considered a separate branch of the information hiding techniques. In their 1999 survey paper, Petitcolas, Anderson, and Kuhn[40] derived a whole field of copyright marking, of which watermarking is a subclass. The current notion refrains from classifying digital watermarking as steganography, due to the lack of an explicit communication aspect and the inferior role of providing "invisibility" to the participants of communication and a larger importance of robustness of such embedded watermarks.

The popularization of paper had further consequences. The steganographic vector was no longer necessarily a physical object, but could take written form, where the carrier text itself would conceal the privileged information. Among the inventions that achieved popularity during medieval times are the textual steganographic methods, particularly the acrostic. This term refers to pieces of writing, whose first letters or syllables spell out a message. The most famous example

of such textual steganography is attributed to a Dominican priest named Francesco Colonna, who, in 1499 hid in his book, *Hypnerotomachia Poliphili*, a love confession which could be spelled out from the first letters of subsequent chapters.[15]

A more sublime carrier is the language itself, as the medieval people had discovered. Here, the embedding process occurs in the linguistic syntax and semantics. Linguistic steganography may be derived from the aforementioned technique of textual steganography, as it relies on the manipulations on the written (possibly even spoken) language with the aim of tricking the perception of an unaware dupe. Following the postulates of Richard Bergmair,[10] linguistic steganography covers within its scope any technique that involves intentional mimicry of typical structures of words, characteristic to a specific language. This may concern the deliberate tampering with grammar, syntax, and the semantics of a natural language. Any action involving modification of those aspects should capacitate maintaining of the innocent appearance of the cover text.

The Renaissance brought about an invention by an Italian scientist Giambattista della Porta who, in the 16th century, detailed how to hide a message inside a hard-boiled egg: write on the shell using ink made from a mixture of alum and vinegar. The solution penetrated the eggshell, leaving no trace on the surface, but a discoloration occurred on the white, leaving the message on its surface, which was only readable once the shell was removed.

Gaspar Schott, a German Jesuit from the Age of Enlightenment, followed the trail marked by his Renaissance predecessors. His work, published in 1680, entitled *Schola Steganographica,* explained how to utilize music scores as a hidden data carrier. Each note corresponded to a letter, which appeared innocent as long as nobody attempted to play the odd-sounding melodies.

The Industrial Revolution, which followed the Age of Enlightenment, brought about new means of communication. Newspapers became a popular and reliable source of the latest information. At some point it became obvious that a newspaper could serve as a perfect steganographic carrier. Since daily papers could be sent free of charge, it was convenient to poke holes over selected letters and thus craft a secret message. This is how "newspaper codes" were born.

The first symptoms of the growing interest in steganography may be traced back to the period of the World War I and II and then the Cold War. These events had brought about such steganographic techniques as microdots—punctuation marks with inserted microscopic negatives of images or texts.[56]

The period during the two World Wars was a true bonanza of hidden communication schemes. World War I witnessed the spectacular return of all sorts of invisible inks.[27] World War II was marked by Hedy Lamarr and George Antheil's patent for spread spectrum communication.[34] They devised a method for guiding torpedoes with a special, multifrequency set of signals, resistant to jamming attempts. The control information was dispersed over a wide-frequency bandwidth that provided cover. The idea of embedding information in a number of different frequencies later found use in the fields of digital image and audio steganography.

The technological development in the 20th century had also accelerated the development of more sophisticated techniques. Among these inventions were the so-called "subliminal channels" based on cryptographic ciphers for the embedding of steganograms. The main principle was to insert content into digital signatures. Gustavus Simmons introduced this concept in 1984, despite the U.S. government's prohibition on publishing of materials on steganography. Simmons proposed the overt and monitored communication conducted between two participants be supplemented with a steganographic channel. This channel would be based on a number of dedicated bits of the message authentication. These, at the cost of reducing the message authentication capability of the digital signature, would serve as the steganographic channel capacity.[46] The steganographic channel established in this way would be visible, yet undetectable. Subliminal channels utilized the cryptographic protocol as the carrier for steganograms.

Contemporary Trends of Development

Modern steganographic techniques utilize the 20th century's inventions—computers and networking. Four main trends of development of the so-called digital steganography can be distinguished: digital media steganography; linguistic steganography; file system steganography; and network steganography.

These four main branches of digital steganography are explained and described here. It must be also emphasized that most of the current research in this area is devoted to digital media and network steganography. The prior is a mature research area with significant achievements, thus the exploration of this field is presently not as dynamic as in the case of the recently sprouted group of techniques falling into the category of network steganography.

Digital media steganography dates back to the 1970s, when researchers focused on developing methods to secretly embed a signature in a digital picture. Many different methods were proposed, including patchwork, least significant bit modifications, and texture block coding.[8] These techniques were intended for both types of images: undergone lossy or lossless compression, like JPEG or BMP, which are the most common image formats. The variety of algorithms for embedding in digital pictures can be grouped according to the type of alterations that were induced. Following Johnson and Jajodia, the modifications are either bit-wise—influencing the spatial domain characteristics of the image, or affect the frequency domain characteristics. Thirdly, specific file format intricacies may be exploited, indeed, a mix of all these techniques is possible. The transform domain provides for the most versatile medium of embedding. Affecting of the image processing algorithms may involve, among others, discrete cosine transform (DCT), discrete wavelet transform (DWT), Fourier transform that may result in alterations of for example, luminance or other measurable property of an image.[20,26] Digital image steganography's position is unfaltering—the survey paper by Cheddad et. al.[13] points to the current interest con-

centrated on employing digital media steganography and watermarking for embedding confidential, patient-related information in medical imagery. Another application of digital image steganography predicted to become popular is the implantation of additional data in printed matter, which, invisible to the naked eye, becomes decodable, when photographed and processed by a cellphone.[13]

Notably, digital image steganography is mostly oriented toward tricking the human visual system into believing the perception of the image has not been manipulated in any way.[8] Similar rules apply to the whole field of digital media steganography, whose primary function is to trick the observer to believe the crafted "forgery" is indeed genuine. The communication aspect of the whole steganographic algorithm is secondary to the process of embedding of the secret data.

Alongside the development of digital image steganography, it appeared the human auditory system is equally prone to delusion as the visual perception. The research focus moved to audio files like MPEGs. The developed techniques included, among others, frequency masking, echo hiding, phase coding, patchwork, and spread spectrum. It also became apparent that error correction coding is a good supplemental carrier for audio steganography—any redundant data can be used to convey the steganogram at the cost of losing some robustness to random errors.[8] This idea later found use in network-protocol based steganography.

Next, steganographers took video files as target carrier. Most of the proposed methods were adaptations of the algorithms proposed for audio and image files. Video-specific solutions involved using either videos I-frames' color space[54] as a steganographic carrier or motion vectors for P-frames and B-frames.[58] Currently, steganography in video files either takes advantage of the existing methods for audio and image files, or makes use of the intrinsic properties of the video transmission, like movement encoding.

Parallel to digital image and audio steganography, information hiding in text was developed—the available methods exploited various aspects of the written word. The first set of techniques altered word spacing, which was even claimed to have been used at the times of Margaret Thatcher to track leakages of cabinet documents.[5] More advanced steganographic methods used syntactic and semantic structure of the text as a carrier. The methods introduced permitted for such displacement of punctuation marks, word order, or alterations of the choice of synonyms, that could be attributed certain meaning. Today, some suggest even SPAM messages may be a carrier of steganography, due to the large amounts of such mail emitted every day.[12] According to work by Bennet,[9] the possible techniques can either rely on the generation of text with a cohesive linguistic structure or the use of natural language text as a carrier. We should note the first technique does not completely fulfill the definition of steganography, where the existence of the carrier should be independent of the existence of the injected hidden content. Thus, a text-lacking rhetorical structure cannot be considered a proper carrier.

Specialists also differentiate between textual steganography and linguistic steganography.[9] The "SPAM method" is a linguistic method, and the embedding occurs with the aid of Context Free Grammars. CFGs have a tree structure, therefore the selection of proper words, or branches, provides encoding for binary data. An example of a textual method would be a substitution technique, where a message's carrier is the set of white spaces and punctuation marks undergoing shifting, repetition, or other modifications.

Parallel to this research, it was revealed that x86 machine code can also be subject to embedding.[17] Some amount of information can be placed in the carrier code, with the aid of careful selection of functionally equivalent instructions. This method exploits the same principle as linguistic steganography, where the choice of words from the set of synonyms can be attributed steganographic meaning.

The invention of a steganographic file system by Anderson, Needham, and Shamir was an eye-opener.[6] It became apparent that information can be steganographically embedded even in isolated computing environments. The main principle of steganogram preparation was similar to invisible inks—one that knew how to search could reveal the encrypted files from a disk. The utilized mechanism relied on the fact that ciphered data resembles random bits naturally present on the disk and only the ability to extract the vectors marking the file boundaries permitted the location process. Another example of a steganographic file system can be found in Pang et al.,[39] whose authors created a steganographic file system implementation on Linux. Their invention preserves the integrity of the stored files and employs a hiding scheme in the disk space with camouflaging with the aid of Dummy Hidden Files and Abandoned Blocks.

Alongside the above-mentioned types of digital steganography, currently the target of increased interest is network steganography. This modern family of methods stems from "covert channels"—a number of techniques intended for monolithic systems, like mainframes. This term was first introduced by Lampson, who identified the problem of information leakage in non-confined programs.[31] The expression "network steganography" was coined by Szczypiorski.[52] Currently, the terms network steganography and covert channels are used interchangeably (and incorrectly), but historically they are sovereign of each other.

A summary of the evolution of the steganographic data carrier is presented in the accompanying figure.

Network Steganography: The Youngest in the Spotlight

Network steganography is the youngest branch of information hiding. It is a fast-developing field: recent years have resulted in multiple new information hiding methods, which can be exploited in various types of networks. The exploitation of protocols belonging to the Open Systems Interconnection (OSI) reference model[59] is the essence of network steganography. This family of methods may utilize one or more protocols simultaneously or the relationships between them—relying on the modification of their intrinsic properties for the embedding of steganograms.

Network steganography is on the rise because embedding secret data into digital media files has been found

to possess two serious drawbacks: it permits hiding only a limited amount of data per one file and the modified picture may be accessible for forensics experts (for example, because it was uploaded to some kind of server). Network-level embedding changes the state of things diametrically; it allows for leakage of information (even very slow) during long periods of time and, if all the exchanged traffic is not captured, then there is nothing left for forensics experts to analyze. As a result, such methods are more difficult to detect and eliminate from networks.

Today, network steganography relies on certain loopholes to conceal its presence. The first is the perceptual inability of the end user to sense minor differences between seemingly identical objects. For example, upon hearing a real-time audio recording transferred through a public network, a person almost certainly will not notice slight alterations of the transmitted voice, especially that he or she will lack any reference for that particular VoIP call. The second loophole permits the passage of steganograms through a network, without raising any alarms in the intermediate nodes. This typically relates to the statistical invisibility—that is, the induced anomalies do not exceed a reasonable threshold typical for network functioning. Typically, three characteristics of communications are utilized for steganographic purposes:

- The communication channel is not perfect—errors are a natural phenomenon and thus it is possible to embed information in a pattern mimicking an ordinary distribution of damaged Protocol Data Units.
- Most protocols bear some quantity of redundant information. The surplus fields can be used for embedding, if this does not induce malfunctioning of the carrier information flow.
- Not every protocol is completely defined. Most of the specifications permit some amount of freedom in implementation, which can be abused.

Network steganography methods, following Jankowski et al.,[25] can be broadly classified according to the number of protocols used for steganographic purposes. The modification of the properties of a single protocol from the OSI model is called intra-protocol steganography, whereas exploitation of relationships between multiple protocols is classified as inter-protocol steganography. Once a protocol or number of protocols are chosen as a carrier for secret data, it is decided how the embedding should be performed. The first possibility is to inject the covert information into the Protocol Data Unit (PDU)— this can be done by means of modification of protocol specific fields or by means of insertion into the payload, or both. Alternatively, or complementary to the previous technique, it is possible to modify the time relations between the PDUs. These changes may impact the order of PDUs, their losses or their relative delays. Hybrid methods utilize both—modification of PDUs and their time relations.

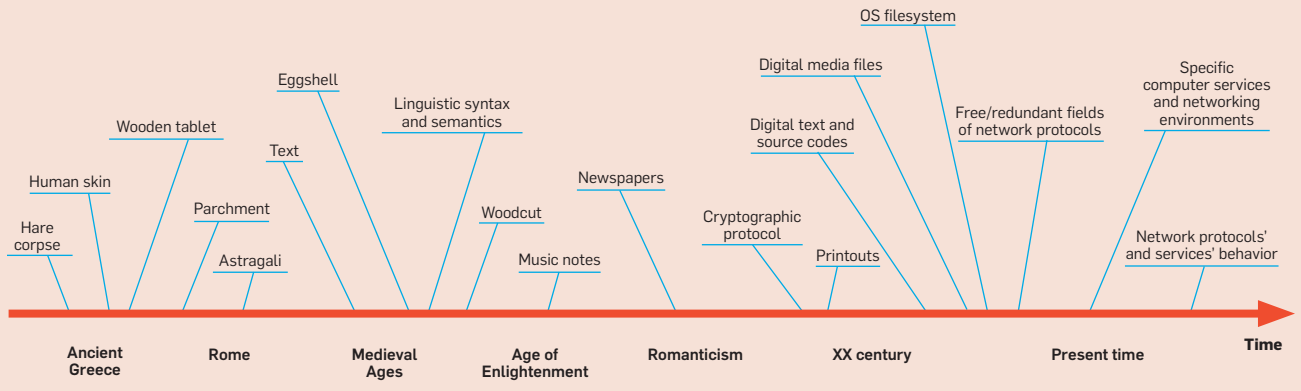
The predecessor of current, more sophisticated network steganography methods, was the utilization of different fields of TCP/IP stack's protocols[42] as a hidden data carrier. The majority of early methods concentrated on embedding in the unused or reserved fields of protocols to convey secret data. Then, more advanced methods were invented, which were targeted toward specific environments or toward specific services. Recent solutions exploit:

- Multimedia, real-time services like IP telephony;[33]
- Popular peer-to-peer services: Skype[35] or P2P file-sharing systems like BitTorrent;[30]
- Social media sites like Facebook;[7]
- Wireless network environments: for example, Wireless Local Area Networks (WLAN),[53] or Long Term Evolution (LTE);[22]
- Cloud computing environments;[41] and
- New network protocols, like Stream Control Transmission Protocol (SCTP).[19]

Although the use of IP telephony service as a hidden data carrier can be considered a fairly recent discovery,[33] the existing VoIP steganographic methods stem from two distinct research origins. The first is the aforementioned, well-established digital media steganography, which has given rise to methods that target the digital representation of the transmitted voice as the carrier for hidden data. The second sphere of solutions target specific VoIP protocol (for example, signaling, transport, or control protocols) fields, or the protocol's behavior.

More recently, Transcoding Steg-

Timeline of the evolution of hidden data carrier.



anography (TranSteg)—intended for a broad class of multimedia and real-time applications like IP telephony—has been proposed.[36] TranSteg is based on the general idea of transcoding (lossy compression) of the voice data from a higher bit rate codec, and thus greater voice payload size, to a lower bit rate codec with smaller voice payload size. This occurs with the least possible degradation in voice quality; compression of the overt data is utilized to make space for the steganogram in the payload field. The achieved steganographic bandwidth is as high as 32kbit/s.

Looking into the P2P services' steganographic applicability, one may encounter a steganographic method named SkyDe (Skype Hide), proposed for Skype by Mazurczyk et al.[35] It utilizes encrypted Skype voice packets as a hidden data carrier. By taking advantage of the high correlation between speech activity and packet size, packets without voice signals can be identified and used to carry secret data. This is achieved by replacing the encrypted silence with secret data bits. The resulting steganographic bandwidth, or hidden-data rate (amount of secret data that can be sent per unit of time, when using a particular method) is about 2kbit/s.

Another recent invention for an Internet P2P service, the StegTorrent, has been introduced for the BitTorrent application.[30] StegTorrent takes advantage of the fact that there are usually many-to-one transmissions in BitTorrent, and that for one of its specific protocols—µTP—the header provides a means for numbering packets and retrieving their original sequence. This allows for sending hidden data with a rate of about 270b/s.

For social media sites like Facebook, Nagaraja et al.[37] proposed creating a botnet communicating over unobservable communication channels. The bots exchanged information with their botmaster by embedding information in images and using the image sharing capabilities to route the secret data to the recipient.

When it comes to the "wireless environment," different standards are targeted by steganographers. For example, for WLANs, Szczypiorski and Mazurczyk have introduced a method called WiPad (Wireless Padding).[53] The technique is based on the insertion of hidden data into the padding of frames at the physical layer of WLANs. It allows data to pass in a covert way with a significantly high data rate of about 1.5Mbit/s. A similar concept was utilized in Grabska and Szczypiorski[22] for LTE and the resulting data rate was about 1.2Mbit/s.

Today, network steganography relies on certain loopholes to conceal its presence.

The cloud computing environment, which Ristenpart et al.[41] views as vulnerable to cross-Virtual Machine information leakage, is a great playground for exercising steganography. They proposed a range of techniques for obtaining classified information by probing the values of shared-cache load, CPU load, keystroke activity, or similar methods.

Other promising future-network protocols, like the SCTP, which is a candidate for new transport layer protocol and might replace TCP (Transmission Control Protocol), and UDP, (User Datagram Protocol) protocols, are also prone to steganography. Detailed analysis in Frączek et al.[19] reveals the most likely places in SCTP transmissions to be utilized for information hiding. Special attention is directed toward steganographic methods that utilize new features, characteristic to SCTP, such as multihoming and multistreaming.

To summarize, various network services and applications can and will become targets of embedding, and the larger the proliferation of a certain service or application, the more attractive it is to piggyback secret data by means of network steganography.

Conclusion

Information hiding covers within its scope various techniques intended for the communication of messages with the aim of keeping some aspect of such exchange secret. This may involve providing security by obscurity for the participants of the dialogue (anonymity), secrecy of the messages (steganography), or protection of the carrier (copyright marking).

The roots of these methods stem from historic times. The need for sending messages that cannot be compromised in case of interception had motivated people to create codes or symbols that appeared innocent, but in fact had different significance than the apparent.

Modern information hiding employs

various embedding techniques—many of these are the result of the transfer of some previously known method into the digital domain. An interesting exception to this notion is network steganography, a family of methods that emerged with the popularization of networked environments. The appearance of new secret data carriers in steganography can be treated as evolutionary steps in the development of information hiding techniques. The growing number of communication protocols, services, and computing environments offers almost unlimited opportunities for displaying a whole spectrum of steganographic methods. It is noteworthy that it is the carrier's properties as well as its popularity that predestine or limit its capability to serve as an efficient medium for clandestine communication, and the emergence of a new technology will likely bring about new information embedding opportunities.

Illicit activities conducted in the virtual world pose a tangible threat to society, as recent cyberwarfare events show. Indisputably, information hiding has joined the arsenal of the utilized weapons, and thus it should be recognized that it poses a large threat to the security of information systems. More importantly, the matter is pressing, because steganalysis techniques are still one step behind the newest steganography methods. There is no "one size fits all" solution available and ready to detect covert communication in our current network security defense systems. Thus, we urge the research community to focus its efforts to discover steganalysis methods that can be practically and promptly deployed in networking environments. C

References

1. Adee, S. Spy vs. spy. *IEEE Spectrum*. (Aug. 2008); http://spectrum.ieee.org/computing/-software/spy-vs-spy/1.
2. Alperovitch, D. *Revealed: Operation Shady RAT*. McAfee, 2011; http://www.mcafee.com/-us/resources/white-papers/wp-operation-shady-rat.pdf.
3. Alureon trojan uses steganography to receive commands. (Sept. 2011); http://-www.virusbtn.com/news/2011/09_26.
4. Analysis, S. and Center, R. World's largest digital steganography database expands again. *SARC Press Release* (Feb. 2012); http://www.sarc-wv.com/news/press_releases/-2012/safdb_v312.aspx.
5. Anderson, R. Stretching the limits of steganography. *Information Hiding*. Springer, 1996, 39–48.
6. Anderson, R. Needham, R. and Shamir, A. The steganographic file system. *Information Hiding*. Springer, 1998, 73–82.
7. Bencsáth, B., Pék, G., Buttyán, L. and Félegyházi, M. Duqu: A Stuxnet-like malware found in the wild, (2011); http://www.crysys.hu/publications/files/-bencsathPBF11duqu.pdf.
8. Bender, W., Gruhl, D., Morimoto, N. and Lu, A. Techniques for data hiding. *IBM Systems Journal 35*, 3&4 (1996), 313–336.
9. Bennett, K. Linguistic steganography: Survey, analysis, and robustness concerns for hiding information in text (2004).
10. Bergmair, R. A comprehensive bibliography of linguistic steganography. In *Proceedings of the SPIE Intl Conf. on Security, Steganography, and Watermarking of Multimedia Contents*, 2007.
11. Bryant, R., Ed. *Investigating Digital Crime*. John Wiley & Sons, 2008, 1–24.
12. Castiglione, A. De Santis, A., Fiore, U. and Palmieri, F. An asynchronous covert channel using SPAM. *Computers & Mathematics with Applications*, 2011.
13. Cheddad, A., Condell, J., Curran, K. and Mc Kevitt, P. Digital image steganography: Survey and analysis of current methods. *Signal Processing 90*, 3 (2010), 727–752.
14. Chen, T. Stuxnet, the real start of cyber warfare? *IEEE Network 24*, 6 (2010), 2–3.
15. Cox, I. *Digital Watermarking and Steganography*. Morgan Kaufmann, 2008.
16. De Sélincourt, A. *Herodotus: The Histories*. Penguin, 1954.
17. El-Khalil, R. and Keromytis, A. Hydan: Hiding information in program binaries. *Information and Communications Security*, (2004), 287–291.
18. Falliere, N., Murchu, L. and Chien, E. W32.Stuxnet dossier. White paper. Symantec Corp., Security Response, 2011.
19. Frączek, W., Mazurczyk, W. and Szczypiorski, K. Hiding information in Stream Control Transmission Protocol. *Computer Communications 35*, 2 (2012), 159–169.
20. Fridrich, J. *Steganography in Digital Media—Principles, Algorithms, and Applications*. Cambridge University Press, 2010.
21. Goodin, D. Duqu spawned by 'well-funded team of competent coders.' World's first known modular rootkit does steganography, too. *The Register*, Nov. 2011.
22. Grabska, I. and Szczypiorski, K. Steganography in LTE. In *Proc. of Intl. Workshop on Cyber Crime*, (San Jose, May 2014).
23. Gross, M.J. Exclusive: Operation shady RAT—Unprecedented cyber-espionage campaign and intellectual-property bonanza. *Vanity Fair* (Aug. 2011).
24. Hayati, P., Potdar, V. and Chang, E. A survey of steganographic and steganalytic tools for the digital forensic investigator. In *Workshop of Information Hiding and Digital Watermarking*, (Canada, 2007).
25. Jankowski, B., Mazurczyk, W., and Szczypiorski, K. PadSteg: Introducing inter-protocol steganography. *Telecommunication Systems: Modeling, Analysis, Design and Management 52*, 2 (2013), 1101–1111.
26. Johnson, N. and Jajodia, S. Steganalysis of images created using current steganography software. *Information Hiding*. Springer, 1998, 273–289.
27. Kahn, D. The history of steganography. *Information Hiding*. Springer, 1996, 1–5.
28. Kellen, T. Hiding in plain view: Could steganography be a terrorist tool? *SANS Institute InfoSec Reading Room*, 2001; http://www.sans.org/reading_room/whitepapers/-stenganography/hiding-plain-view-steganography-terrorist-tool_551.
29. Kelley, J. Terror groups hide behind Web encryption. *USA Today* (May 2001); http://-www.usatoday.com/tech/news/2001-02-05-binladen.htm.
30. Kopiczko, P., Mazurczyk, W. and Szczypiorski, K. StegTorrent: A steganographic method for P2P files sharing service. In *Proc. of Intl. Workshop on Cyber Crime*, (San Francisco, CA, May 2013).
31. Lampson, B. A note on the confinement problem. *Commun. ACM 16*, 10 (Oct. 1973), 613–615.
32. Lau, H. The truth behind the Shady RAT. *McAffe report*, (Aug. 2011); http://-www.symantec.com/connect/blogs/truth-behind-shady-rat.
33. Lubacz, J., Mazurczyk, W. and Szczypiorski, K. Vice over IP. *IEEE Spectrum 47*, 2 (2010), 42–47.
34. Markey, H. and Antheil, G. Secret communication system. Aug. 11 1942, US Patent 2,292,387.
35. Mazurczyk, W., Karaś, M. and Szczypiorski, K. SkyDe: A Skype-based steganographic method. *International J. Computers, Communications & Control 8*, 3 (June 2013), 389–400.
36. Mazurczyk, W., Szaga, P. and Szczypiorski, K. Using transcoding for hidden communication in IP telephony. *Multimedia Tools and Applications* (2011), 1–27; DOI 10.1007/s11042-012-1224-8.
37. Nagaraja, S., Houmansadr, A., Piyawongwisal, P., Singh, V. Agarwal, and Borisov, N. Stegobot: A covert social network botnet. *Information Hiding*. Springer, 2011, 299–313.
38. Networking and Information Technology Research and Development Program. I.W.G. on Cyber Security and I. Assurance. *Federal Plan for Cyber Security and Information Assurance Research and Development*, (Apr. 2006); http://www.nitrd.gov/pubs/csia/-csia_federal_plan.pdf.
39. Pang, H., Tan, K. and Zhou, X. Stegfs: A steganographic file system. In *Proceedings of the 19th Intl. Conf. on Data Engineering*. IEEE, 2003, 657–667.
40. Petitcolas, F., Anderson, R. and Kuhn, M. Information hiding—A survey. In *Proceedings of the IEEE 87*, 7 (1999), 1062–1078.
41. Ristenpart, T., Tromer, E., Shacham, H. and Savage, S. 'Hey, you, get off of my cloud: Exploring information leakage in third-party compute clouds.' In *Proceedings of the 16th ACM Conference on Computer and Communications Security*. ACM, 2009, 199–212.
42. Rowland, C. Covert channels in the TCP/IP protocol suite. *First Monday 2*, 5 (1997).
43. Rudin, B. and Tanner, R. *Making Paper: A Look into the History of an Ancient Craft*. Rudin, 1990.
44. Shachtman, N. FBI: Spies hid secret messages on public websites. *Wired* (June 2010); http://www.wired.com/dangerroom/2010/06/alleged-spies-hid-secret-messages-on-public-websites/.
45. Sieberg, D. Bin Laden exploits technology to suit his needs. *CNN* (Sept. 2001); http://-edition.cnn.com/2001/US/09/20/inv.terrorist.search/.
46. Simmons, G. The prisoners' problem and the subliminal channel. In *Proceedings of Crypto '83: Advances in Cryptology* (1984), 51–67.
47. Singh, S. *The Code Book: The Secret History of Codes and Codebreaking*. Fourth Estate, 2000.
48. Smith, D. Number games and number rhymes: The great number game of dice. *The Teachers College Record 13*, 5 (1912), 39–53.
49. Srivastava, K. Congress wants answers on world's largest security breach. Aug. 2011; http://www.mobiledia.com/news/102480.html.
50. Stier, C. *Russian spy ring hid secret messages on the Web*. (July 2010); http://-www.newscientist.com/article/dn19126-russian-spy-ring-hid-secret-messages-on-the-web.html.
51. Symantec. W32.Duqu—The precursor to the next Stuxnet, (Nov. 2011); http://-www.symantec.com/content/en/us/enterprise/media/security_response/whitepapers/-w32_duqu_the_precursor_to_the_next_stuxnet_research.pdf.
52. Szczypiorski, K. Steganography in TCP/IP networks. In *Proceedings of State of the Art and a Proposal of a New System–HICCUPS*. Institute of Telecommunications' seminar, Warsaw University of Technology, Poland, 2003.
53. Szczypiorski, K. and Mazurczyk, W. Steganography in IEEE 802.11 OFDM symbols. *Security and Communication Networks 3* (2011), 1–12.
54. Wang, Y. and Izquierdo, E. High-capacity data hiding in MPEG-2 compressed video. In *Proceeding of the 9th Intl. Workshop on Systems, Signals and Image Processing* (Manchester, U.K., 2002), 212–218.
55. Wayner, P. *Disappearing Cryptography—Information Hiding: Steganography & Watermarking*. Morgan Kaufmann, 2009.
56. White, W. *The Microdot: History and Application*. Phillips Publications, 1992.
57. Williams, C. Russian spy ring bust uncovers tech toolkit. *The Register* (June 2010); http://www.theregister.co.uk/2010/06/29/spy_ring_tech/.
58. Xu, C., Ping, X. and Zhang, T. Steganography in compressed video stream. In *Proceedings of the First International Conference on Innovative Computing, Information and Control*. IEEE, 2006, 269–272.
59. Zimmermann, H. OSI reference model—the ISO model of architecture for open systems interconnection. *IEEE Transactions on Communications 28*, 4 (1980), 425–432.

Elżbieta Zielińska (ezielinska@tele.pw.edu.pl) is a research assistant at Warsaw University of Technology, Institute of Telecommunications, Warsaw, Poland.

Wojciech Mazurczyk (wmazurczyk@cygnus.tele.pw.edu.pl) is an assistant professor at Warsaw University of Technology, Institute of Telecommunications, Warsaw, Poland.

Krzysztof Szczypiorski (ksz@tele.pw.edu.pl) is a professor at Warsaw University of Technology, Institute of Telecommunications, Warsaw, Poland.