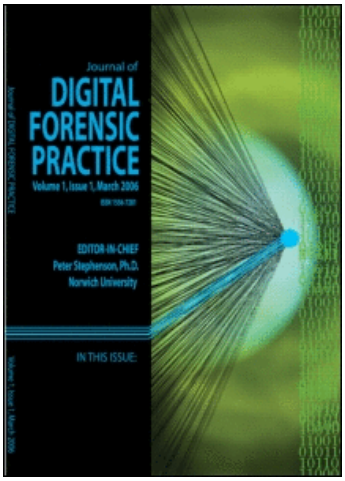


Journal of Digital Forensic Practice
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t716100764

Graphstega: Graph Steganography Methodology

Abdelrahman Desoky[a]; Mohamed Younis[a]
[a] Department of Computer Science and Electrical Engineering, University of Maryland, Baltimore County, Baltimore, MD, United States

To cite this Article Desoky, Abdelrahman and Younis, Mohamed(2008) 'Graphstega: Graph Steganography Methodology', Journal of Digital Forensic Practice, 2: 1, 27 — 36
To link to this Article: DOI: 10.1080/15567280701797087
URL: http://dx.doi.org/10.1080/15567280701797087

Journal of Digital Forensic Practice, 2:27–36, 2008

ISSN: 1556-7281 print / 1556-7346 online
DOI: 10.1080/15567280701797087

Taylor & Francis
Taylor & Francis Group

ARTICLE

Graphstega: Graph Steganography Methodology

Abdelrahman Desoky and Mohamed Younis
University of Maryland, Baltimore County, Department of Computer Science and Electrical Engineering, 1000 Hilltop Circle, Baltimore, MD 21250
United States

ABSTRACT The Graph Steganography Methodology (Graphstega) is the art and science of avoiding the arousal of suspicion in covert communications by concealing a message in a graph-cover. Unlike other approaches, Graphstega does not embed a message as a noise in a cover. Instead the message is camouflaged as plotted data. Graphstega is keyless and the hidden message is anti-distortion. The popular usage of graphs in business, education, news, etc. and the availability of tremendous amount of graphs in electronic and non-electronic format make the investigation and detection of a hidden message extremely difficult. In addition, Graphstega is resilient to contemporary attacks, such as traffic analysis, contrast, and comparison attack, even when launched by an adversary who is familiar with Graphstega.

KEYWORDS Steganography, Graph steganography, traffic attack, contrast attack, comparison attack, graph-cover, image-cover, audio-cover, text-cover

INTRODUCTION

Steganography is the scientific art of avoiding the arousal of suspicion in covert communications. Information hiding and covert transmission of messages have been practiced since the time of ancient civilizations. The Greeks used "cover writing," from which the name "steganography" was derived, to share secrets [1, 2]. The contemporary techniques have increased in sophistication over the years, mostly to counter the growing strength of adversaries in detecting and disrupting such covert communications. Generally, steganography schemes hide a message as a noise in a cover that is assumed to look innocent. Contemporary approaches are categorized based on the steganographic cover type such as text, image, or audio.

Textual steganography can be classified as textual format manipulation (TFM) and textual fabrication (TF) [3]. In TFM, comparing the original text with the modified text will reveal the hidden message [4, 5]. On the other hand, textual fabrication techniques generate an entire text cover for hiding a message rather than manipulating an existing text. Examples of these approaches are null cipher [6], mimic functions [7, 8], NICETEXT and SCRAMBLE [9], and translation-based [10–12]. However, the text cover that is generated by these approaches often has numerous linguistic flaws that can raise suspicion. In addition, revealing the hidden message may be feasible [3, 10].

Address correspondence to Abdelrahman Desoky, University of Maryland, Baltimore County, Department of Computer Science and Electrical Engineering, 1000 Hilltop Circle, Baltimore, MD 21250.
E-mail: iloveitech@yahoo.com.

On the other hand, image steganography is based on manipulating digital images to conceal a message. Such manipulation often renders the message as noise. In general, image steganography suffers from several issues such as the potential of distortion, the significant size limitation of the messages that can be embedded, and the increased vulnerability of detection through contemporary image processing techniques [13]. Audio covers have also been pursued. Example of audio steganography techniques include LSB [14, 15], spread spectrum coding [16, 17], phase coding [16, 18], and echo hiding [19]. In general, these techniques are too complex and, like their image-based counterpart, are still subject to distortion and vulnerable of detection [4, 16].

The pillars of steganography approaches found in the literature are the techniques that are used to hide a message, the secrecy of these techniques, and the use of stega-key. The fundamental assumption is that an adversary will not know that a particular approach is used. The stega-key plays the role of a password in order to prevent the revealing of a hidden message in case the adversary could successfully detect the presence of a hidden message. In other words, contemporary steganography schemes in essence act as a cryptosystem rather than a stegasystem because the goal of steganography is avoiding the arousal of suspicion in covert communications rather than making it difficult for an adversary to decode a hidden message. If an approach relies on a stega-key and an adversary suspects a hidden message, the goal of steganography is defeated regardless of whether or not an adversary is able to reveal a plaintext.

In this article, a novel graph steganography methodology (Graphstega) is presented. Graphstega camouflages a message in two steps. First, it encodes the message in a form that can be plotted by a graph (such as, but not limited to, numerical values). Second, it camouflages the message by representing the steganographic code (the encoded message) as data points in a graph.

The main advantages of Graphstega over all other approaches are as follows:

- Since the message is not concealed as a noise in the steganographic cover the hidden message is antidistortion.
- Graphstega is a public approach that neither relies on the secrecy of its technique nor needs to employ a stega-key.
- Graphstega is also resilient to comparison attack by untraceable or authenticated data. In addition, the naturalistic feature of Graphstega makes the cover look innocent and thus averts the arousal of suspicion.
- A graph cover can be accessible in numerous styles; e.g., chart in a form of bar, pie, scatter, etc.
- Graphstega can convert the graph cover to all other steganographic cover types; e.g., text cover, image cover, and audio cover. Such diverse representation of the graph cover allows flexibility in generating the steganographic cover in order to fool both human and machine examinations.
- Graphstega has a low cost relative to all other schemes.

The remainder of this article is organized as follows. The next section describes the Graphstega methodology in detail, discusses its advantages, and highlights the implementation issues. The following section presents the steganalysis validation. Finally, the article is concluded with a summary and an outline of future research directions.

GRAPHSTEGA METHODOLOGY

Graphstega camouflages a message in two steps. First, it encodes a message in a form that can be camouflaged in a graph. Second, it represents the steganographic code (the encoded message) in a graph as data points. In this section, these steps and the overall communications protocol are demonstrated.

Message Encoding

Graphstega creates an encoded representation of the plaintext message and then camouflages it in a graph. In general, Graphstega does not impose any constraint on the message encoder scheme as long as it generates a set of data values that can be embedded in a graph cover. However, the subject of the graph has to be factored in the selection of the most appropriate encoding scheme. For example, a graph that reports weather changes would restrain the values of the data to a range within which the encoded message has to stay. In addition, the variations in the data values have to be considered, especially when the graph is not shared in a form that allows the recipient to access the table on which the graph is based. For

example, showing a graph in which data varies between 1 and 2000 would most probably make it difficult, and even infeasible, for the recipient to accurately determine the values unless they are annotated on the graph. The same applies if the message is encoded using real numbers, which sometimes lose some precision when plotted. Given the availability of numerous encoding techniques in the literature that can meet these constraints [1–4, 20, 21], the discussion in the balance of this section will be focused on the graph cover rather than the message encoding.

In the examples shown in this article, messages are encoded as follows. First, the plaintext message is converted to a binary string, which is then partitioned into groups of a particular number of bits that is agreed upon among communicating parties. Finally, a decimal representation is generated for the individual groups. For example, in a 7-bit–based grouping, the value of each group would range between 0 and 127 (0000000 to 1111111 in binary). The following describes the encoding of a sample message:

- The plaintext of the message is "Use my secret key."
- The concatenated binary string of the ASCII representation of this message is

  01010101011100110110010100100000
  00110110101111001001000000111001
  10110010101100011011100100110010101
  11010000100000011010110110010101111001

- Slicing this string (from the previous step) into 7 bits each, as set and agreed upon by communicating parties, will result in:

  0101010 1011100 1101100 1010010
  0000011 0110101 1110010 0100000
  0111001 1011001 0101100 0110111
  0010011 0010101 1101000 0100000
  0110101 1011001 0101111 001

- Converting the individual slices (from the previous step) into decimals results in:

  42 92 108 82 3 53 114 32 57 89
  44 55 19 21 104 32 53 89 47 1

It is worth noting that the range of the resulting decimal values can be easily narrowed or widen by partitioning the binary string into groups of less or more than 7 bits. Again, this encoding scheme is just for illustration and many alternative schemes can be employed.

Graph Cover

As mentioned, Graphstega camouflages a message by embedding the encoded message as data points in a graph cover. In other word, the encoded message will be the data represented in a graph. The message may constitute a subset or a full set of the data in the graph. The latter case makes the message's decoding a straightforward exercise, basically by applying the reverse process using all data items. However, the use of a partial data set would require a pre-agreement among the communicating parties on how to pick the data items that are relevant to the decoding of the message. For example, the sender may agree with the recipient on considering only every other data item on the list. The use of a subset of the plotted data can make Graphstega more resilient to attacks, as will be discussed later. Basically, an adversary would have to try all possible subsets of the plotted data in order to identify the relevant items assuming that he will suspect the presence of a hidden message in a particular set of communications traffic and attempt to guess the encoding scheme.

The subject and context of graphs are usually dependent. Obviously, the subject would determine the correctness of the data. For example, a graph that shows the blood pressure over a period of time cannot have data values that are out of the known range for a live human being. The context would most probably influence the choice of the graph style. For example, pie charts would suit high-level summaries, while a Pareto chart captures the relative importance of the differences among groups of data. Therefore, the communicating parties ought to first agree on the subject that each will use to conceal messages. Examples include finance, medicine, math, and economics reports and analysis. The selection criteria include the suitability of the picked subject for concealing the encoded message and for averting suspicion. In general, the picked subject has to fit the communicating parties and provides some ground for justifying the communications. The selected subject also has to suit

the desired frequency of communications. Through the use of some subjects it may be possible to generate a new graph every hour or less. For example, it is customary for a stockbroker to receive a market update every half hour and even more frequent updates on stocks of some monitored companies. On the other hand, some subjects may not justify more than one message per month, season, or even a year. For example, it is not very often that someone will receive an e-mail message or a letter from the utility company about the rate of energy consumption or payment history. Finally, some subjects enable broadcasts or non-private announcements to a set of interested parties and would thus make the association between the sender and the recipient unsuspected. Unsolicited marketing is a perfect example in this category, where a sender e-mails a brochure to multiple recipients in order to sell some unpopular products, promotes a stock of a new company, etc.

One of the obvious concerns about the use of graph cover is the size constraint on the message that can be hidden. One would argue that only small messages may be embedded in a graph. While the concern is legitimate, Graphstega is capable of concealing both short and long messages. For short messages the encoded version will be used as the data plotted in the graph. In this article, MS Excel is employed by Graphstega to generate the graph cover. Figure 1 shows an example that hides the message "Use my secret key" encoded in previously, basically plotting the following set of values: 42 92 108 82 3 53 114 32 57 89 44 55 19 21 104 32 53 89 47 1. Often in that case, it would suffice for the recipient to visually inspect the graph in order to note down the data values and decode the message. For a long message, however, data cannot easily be noted from the graph due to the constraints on the scale and plotting area. In other words, the scale would hinder the determination of the data values with high resolution so that the message can be accurately decoded. Therefore, Graphstega requires that long messages be embedded in a graph that is included as an object in the cover so that the data can be accurately retrieved. Examples include attaching an Excel file to an e-mail message or posting a graph on a web site with an access to a downloadable version of it.

Figure 2 shows a sample graph cover that conceals a long message and lists its relevant characteristics. The message is the Consumer Prices Index of July and August 2007 [22], the size of the long message is 47.5 KB. It is

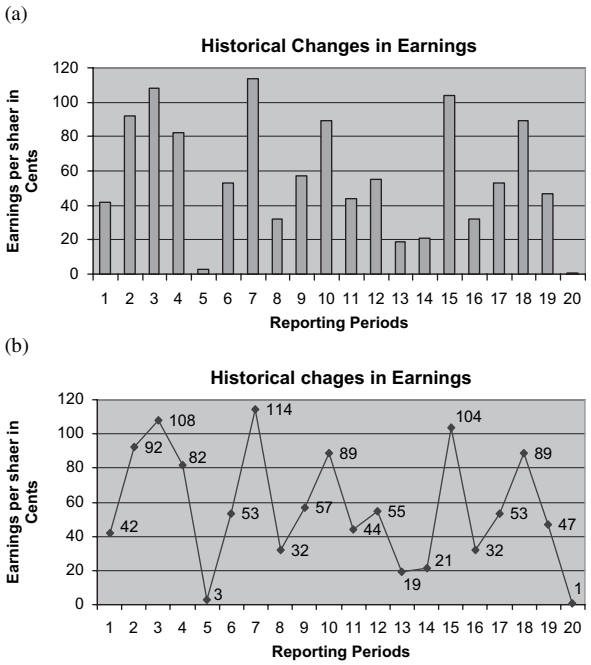
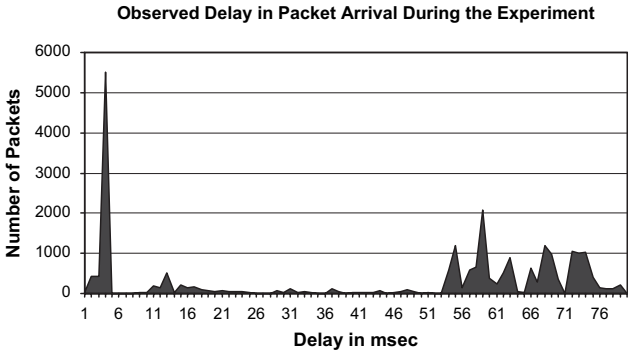


FIGURE 1 **Illustrating the capability of Graphstega in concealing the message "Use my secret key."**



Pages	16
Words	6,652
Characters with spaces	35,801
Lines	894
Paragraphs	701
Message size in bytes	46,923

FIGURE 2 **Example graph cover that conceals a long message.**

encoded by applying the scheme described previous using groups of 8 bits; each decimal number is in range [0, 255]. The frequency of appearance of the decimal numbers in the range [1, 81] is plotted and portrayed as packet delay measures of an experiment. The graph can be inserted as an object in a MS Word document. The recipient can double click the graph in order to open the accompanied Excel file, access the row data, and finally decode the message. It is argued that most published steganography approaches would

highly recommend small messages to avoid detection since they embed them as noise. The bit rate for concealing the above examples of long and short messages is 4.37% and 1.0061%, respectively. Therefore, Graphstega is in fact superior to contemporary steganography when it comes to effectiveness and flexibility in hiding all sorts of messages.

Other Steganographic Cover Types

Graphstega has a very unique feature that sets it apart from all other steganography approaches. Basically, the pursued graph cover can be presented in different cover types. In other words, Graphstega is capable of converting the graph cover to an image cover, text cover, or audio cover without distortion or significant complexity. However, Graphstega's use of the other cover types is somehow constrained by the message size, as explained in this section.

Graph in Image Cover

Converting the graph cover that conceals the message in an image cover is straightforward. Simply the graph can be converted to bitmap, GIF, JPEG, Bitmap, PNG, or any other digital image format. Presenting the graph in an image cover will not cause the loss or the distortion of the hidden message because the message is concealed as data points rather than the pixels. Unlike image steganography, an adversary will not suspect the image cover since no noise is exploited in the message concealing process. Obviously, the data should be visually (or by scaling) identifiable from the image in order to allow the recipient to reveal the encoded message for the legitimate user. Thus, an image cover may be unfit for large messages but suitable for short message.

Presenting the Graph's Data in Text Cover

Graphstega can also employ a text cover. In this case, the data values used in the encoded message will be enumerated and mixed with some text. Obviously, an appropriate subject and text cover need to be generated to suit the data values so that the cover looks legitimate. An example is shown in Figure 3, employing a set of authenticated data that is collected through the use of Internet search engines such as

Book title	Author	Price
My Southern friends	James R. (James Roberts) Gilmore	$42.00
Designing with Solar Power	Deo Prasad (Editor), Mark Snow (Editor)	$92.00
Anthology of Bulgarian Folk Musicians	Todor Bakalov	$108.00
Buen viaje! Level 1, Student Edition	McGraw-Hill	$82.00
Horticulture Magazine. June 1972	Horticulture Magazine Editors	$3.00
The Architect's Handbook of Professional Practice	The American Institute of Architects and Joseph A. Demkin	$53.00
Molecular Anatomy of Cellular Systems	I. Endo, I. Yamaguchi, T. Kudo, H. Osada, and T. Shibata	$114.00
The Jazz Piano Book	Mark Levine	$32.00
Tracking and Monitoring Legislation	TheCapitol.Net and Christopher Davis	$57.00
The LabVIEW Style Book	Peter A. Blume	$89.00
How the Best Get Better, Book and CD set	Dan Sullivan	$44.00
Couples and Family Client Education Handout Planner	Laurie Cope Grand	$55.00
The Well Cat Book	Terri Dvm Mcginnis	$19.00
4 Blondes	Candace Bushnell	$21.00
Differential Equations and Dynamical Systems	J.K. Hale , and J.P. LaSalle	$104.00
Parkett No. 70	Christian Marclay, Wilhelm Sasnal, and Gillian Wearing	$32.00
The Ceo's Guide to Health Care Information Systems	Joseph M. Deluca and Rebecca Enmark Cagan	$53.00
Preparatorio para o Exame de PMP	Rita Mulcahy	$89.00
Particles, Sources, and Fields: Volume 3	Julian Seymour Schwinger	$47.00
A Dollar = $1.00	Carey Molter and Monica Marx	$1.00

FIGURE 3 **A graph cover presented in textual cover to conceal the message "Use my secret key" using book prices from www.amazon.com on Friday, September 7, 2007.**

Google, in which a list of authenticated books from www.amazon.com is used to justify the values. Other examples are shown in Figure 4. The user must have legitimate reasons when using these techniques in order to fool an adversary. Downloading a free trial of software can be a legitimate reason for concealing a message in a software key. It is worth noting that the autogeneration of appropriate textual covers involves numerous techniques from natural language processing and is beyond the scope of this article. Graphstega is resilient to distortion and is capable of passing both human and machine examinations without raising suspicion. As shown by examples, linguistically, the presented text cover is flawless and looks legitimate.

Presenting the Graph's Data in Audio Cover

The use of audio covers is also possible for presenting the graph data. In order for Graphstega to do so, a textual cover has to be generated first, as explained above, and then Graphstega scheme converts the text cover to a voice message. The latter step can be done easily using text-to-speech software that is widely available in the market for a nominal fee or even downloadable for free on the Internet. The audio file can be for an oral presentation during which the graph is explained, news coverage, etc. Again, an audio cover in this case is resilient to both distortion and destruction, implying that the message will not be lost, damaged, or altered during transmission because the message is concealed as data points.

It is important to note that choosing the appropriate subject cover is crucial in this case, and in fact in all other cover types pursued by Graphstega. In addition, the user must have a convincing reason for using the chosen technique for legitimizing the transmittal of a hidden message. For instance, although a bookstore receiving booklists seems innocent, receiving it as an audio file may raise suspicion. Meanwhile, one would not question the motive for sending a commentary of a CEO of startup or an oral summary of a meeting. The ability of Graphstega for employing multiple types of cover makes it a very versatile approach and enables robust communications among the involved parties.

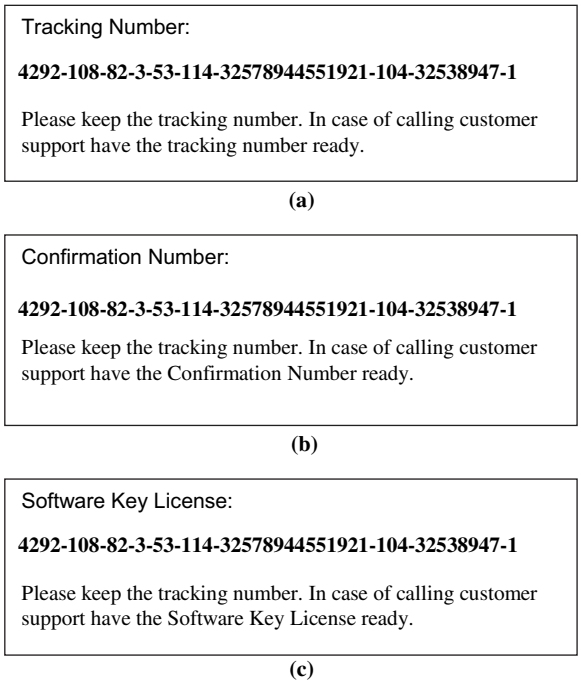


FIGURE 4 **Examples for possible textual covers that can be employed by Graphstega.**

Communications Protocol

Covert communications are done through two steps: concealing a message and then transmitting the hidden message. Contemporary steganography approaches have focused on how to hide a message but not on how the steganographic cover will be delivered to the intended recipient. It is argued that the transmittal process of the hidden message has to be very cautious in order to avoid raising suspicion. At the core of the cover transmittal issue is how to prevent the association between the sender and recipient from drawing suspicion. For example, exchanging e-mail messages would automatically imply a relationship between the communicating parties. Similarly, downloading files from a web site indicates an interest in the accessed material. With advances in monitoring tools for network and Internet traffic, profiles of user's access pattern can be easily established. This issue can be even more serious when the sender always uses the same steganographic cover type; mimic functions, translation-based, image-based, or audio-based. An adversary overseeing this type of communication most probably will suspect the presence of a hidden message, even if the content does not look suspicious, because of the observed traffic pattern and the lack of a justification for the interest in the contents of such message traffic. Therefore, it is very important to rationalize the receiving of the steganographic cover and diversify the format of that cover in order to avoid attracting any attention that may trigger an attack.

Graphstega enables a powerful solution to the issue of cover's transmittal to recipients. The use of graphs allows a legitimate association among the communicating parties and would thus make sharing the cover very ordinary. Graphs are very popular in all sort of reports, articles, educational material, marketing brochures, etc. Such popularity makes the transmission of the cover via e-mail, posting them on web pages, or even downloading articles that include graphs a very natural matter. For example, camouflaging a message in a stock market report, in a form of graph cover, text cover, or any other cover types, that is sent or posted on the Internet from a broker to a client would not be unusual. In fact, such report can be sent to many clients with only one of them being able to reveal the hidden message. In addition, casual message exchange that includes no hidden messages can be pursued in order to avoid the formation of a communications pattern that may draw attention. Explicit message transmission is not the only means for sharing the cover. Web posting, postal mail, and printed articles are samples of other means that can be pursued. In summary, the way of delivering the hidden message can raise suspicion even if using a secure hiding technique. Graphstega averts the suspicion that may arise during covert communications not only by camouflaging a message but also its transmittal. Therefore, Graphstega imposes that the intended users make the appropriate arrangements, techniques, policies, rules, and any other related specifications for achieving its goal. In general, a sender and recipient communicating covertly using Graphstega should agree to the following:

- The specifications and configurations of Graphstega Encoder and Decoder.
- The arrangements for the covert transmission of a hidden message. This step is to establish a legitimate channel for communications among intended users, including picking an appropriate subject and format for the cover.

The Graphstega communications protocol is illustrated in Figure 5.

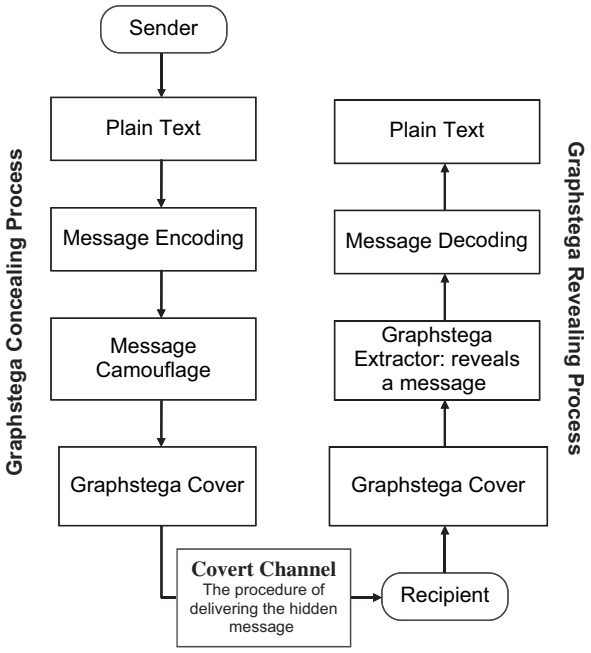


FIGURE 5 **Summary of the Graphstega communications protocol followed by a sender and a recipient.**

STEGANALYSIS VALIDATION

The aim of this section is to show the resilience of Graphstega to possible attacks. Again the success of steganography is qualified with its ability for avoiding an adversary to suspect the presence of a hidden message. It is assumed that an adversary will perform all possible investigations. In addition, the adversary is also aware of Graphstega, as a public methodology, but he does not know the detail and the specification of the actual implementation of Graphstega scheme used.

Traffic Analysis

One of the possible attacks an adversary may pursue is to analyze the communications traffic and the access patterns to publicly available or exchanged documents, images, files, etc. For example, the intelligence community has a number of tools at their disposal for analyzing traffic on the Internet, tracking access to web sites, monitoring checked-out literature from public libraries, etc. The main goal of a traffic attack is to detect unusual or questionable association between a sender and recipient. Traffic analysis intuitively can identify who communicates with whom. The relationship between the communicating parties will be then qualified based on the contents of the message. Traffic attacks can be applied to contemporary steganographic techniques regardless of the steganographic cover type (e.g., image cover, audio

cover, text cover, etc.) and can achieve successful results with relatively low costs. In the context of Graphstega, the subject of the cover is checked rather than its validity and consistency. If someone sends, receives, accesses some materials without a legitimate reason for doing so, suspicion can be raised and further investigation may be warranted. For example, a weather analysis report sent to someone by a physician is definitely unusual. The additional investigations will involve a thorough analysis of the cover, as detailed in the next subsection.

Traffic analysis is deemed ineffective with Graphstega. Graphstega camouflages the transmittal of a hidden message to appear legitimate and thus suspicion is averted. Basically, Graphstega ensures that the involved parties establish a covert communications channel by having a well-defined relationship with each other. Analyzing the traffic between them will not reveal any questionable association and will not trigger any further investigation. Finally, it is noted that if further investigations on a Graphstega cover are triggered by traffic analysis, they would not be successful, as elaborated next. In Graphstega, differentiating between a graph that has or does not have a hidden message is extremely difficult. The popularity of using graphs in both digital and non-digital format would make the investigation of exchanging graphs in all its forms infeasible and would further avert any suspicion about the covert communications.

Contrast and Comparison Attacks

One of the intuitive sources of noise that may alert an adversary is the presence of contradictions in the graph and any associated text, such as finding the value of a product edging up while saying that it has decreased. Contradictions can also be spotted when using data that is factual in nature and has some scientific or technological basis. For example, a graph about variations in the bit rate on a wireless link cannot be in the megabit range. The use of authenticated or untraceable data will definitely counter such an attack. Untraceable data means data that is based on a private context, e.g., expenses in one's own company, and thus cannot be contrasted. Meanwhile, noise in the context of comparison attacks reflects alteration of authenticated or previously used data. The goal is to find any incorrect and/or inconsistent data that may imply the manipulation of the text contents to include a hidden message. The vulnerability of Graphstega to comparison attacks depends on the picked subject. Employing a subject that inherently involves public domain information such as the weather would allow an adversary to access authenticated data as well as historical changes in that data if any.

It is worth noting that the traffic analysis, discussed earlier, can also be pursued as a base for launching comparison attacks in case the data is not publicly accessible. In that case, current data are compared to a record of old data in order to search for any inconsistency over some period of time. Countering such an attack is always a challenge because it requires consistency with data that were previously used over an extended period of time. Contradictions would surely raise suspicion about the existence of a hidden message. Graphstega, as demonstrated through examples, is simply made contrast aware. The freedom in selecting a suitable subject for the cover, enabled by the popularity of graphs, flexibility in messages encoding, and the ability in employing more than one cover type, enables Graphstega to avert such attacks. In addition, the enormous amount of data available, both publicly and privately, provides sufficient sources of correct and consistent data for embedding all sorts of messages.

Discussion

This section attempts to highlight some of the implications of the properties of Graphstega on forensic investigation. As stated above, Graphstega is difficult to crack, especially given the volume of communication traffic that contains graphs, the wide variety of document formats in which graphs may appear, and the various types steganographic covers that Graphstega may employ. Therefore, traffic analysis and large-scale traffic monitoring would not be effective. Instead, a forensic investigator would have to focus on the semantics of the message that contains Graphstega covers and the association among the communicating parties over a period of time.

To illustrate, assume that the graph in Figure 6 is being inspected. The graph obviously looks suspicious because the data used in the graph are not true data and do not make sense. When the investigator applies the revealing process to uncover any hidden message, the result is an unreadable text as shown below.

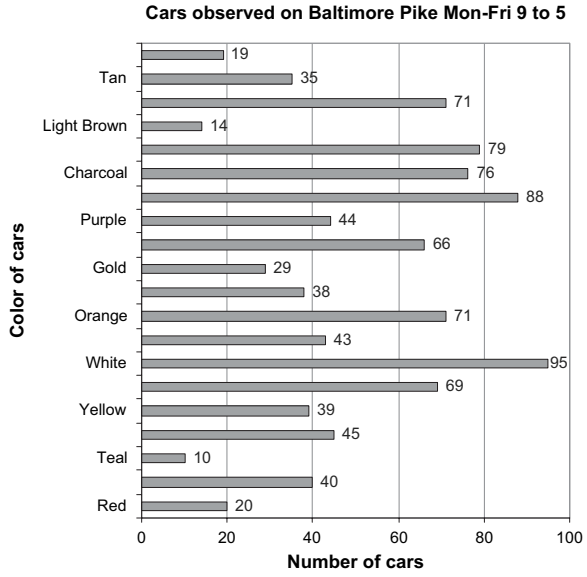


FIGURE 6 **An example of a graph with questionable data.**

The decimal numbers:
45 39 69 95 43 71 38 29 66 44 88 76 79 14 35 19
The corresponding binary numbers:
010110001001101000100101110101010100011001
001010011100100000101010111010111100101110011
1100001101010001000100 10
The text message based on ASCII:
Xš%ÕQ'œ,®¼¹ÃQ□

The unreadable text may or may not be a message. In fact, differentiating between a cover with a hidden message and one without is extremely difficult when Graphstega is used. The only way to determine whether there is a hidden message is the detection or the decoding of an encoded message. Note that a teacher could have made the graph in Figure 6 for her 5th-grade students for educational purposes. The data may look funny and suspicious because a teacher used made-up data.

Although an investigator may assume that if the data are suspicious then there is a hidden message, this is not necessarily true because this resorts to saying that made-up data for all examples such as educational or illustration proposes, or for marketing a product, will be suspicious. This leads to the need for reasoning about the association among the communicating parties and rationalizing the subject of the exchanged messages. If the relationship among a sender and a receiver is questionable, the investigator would have to track the message traffic among these communicating parties for some time. Tracking the communications traffic may enable the detection of the message encoding, if the communicating parties are not aware to counter such an attack. It is worth noting that the tracking duration may be long if the communicating parties change the encoding scheme and the covers frequently.

CONCLUSION

This article has introduced Graphstega, a novel steganography methodology, which employs graphs as a cover for concealing messages. Graphstega, unlike other approaches, does not embed the message as a noise in the cover; instead, the message is encoded and used as the data plotted in a graph. The Graphstega is a keyless approach and can be employed without difficulty using popular software tools such MS Excel. Through multiple examples, the article has confirmed the ability of Graphstega to camouflage both short and long messages. In addition, Graphstega can convert the graph cover to all other steganographic cover types; e.g., text cover, image cover, and audio cover. Such diverse representation of the graph cover allows flexibility in generating the steganographic cover in order to fool both human and machine examinations.

Graphstega camouflages both a message and its transmittal. Graphstega has been shown to be resilient to all contemporary attacks such as traffic analysis and contrast and comparison. The tremendous amount of graphs in electronic and non-electronic format and the high volume of traffic accessing these materials make it impossible to investigate each and every content and transaction. Therefore, graphs are rendered a favorable steganographic cover.

REFERENCES

1. Kipper, G. *Investigator's Guide to Steganography*, pp. 15–16. CRC Press LLC, 2004.
2. Davern, P., and Scott, M., "Steganography Its History and Its Application to Computer Based Data Files." Internal Report Working Paper: CA-0795, School of Computing, Dublin City University (1995). Available from http://computing.dcu.ie/research/papers/1995/0795.pdf. Accessed 23 December 2005.
3. Bennett, K. "Linguistic Steganography: Survey, Analysis, and Robustness Concerns for Hiding Information in Text." CERIAS Tech Report 2004-13, Purdue University, 2004.
4. Johnson, N. F., and Katzenbeisser, S. "A Survey of Steganographic Techniques." In *Information Hiding,* edited by S. Katzenbeisser and F. Petitcolas. Norwood, Mass.: Artech House, 2000.

5. Shirali-Shahreza, M. H., and Shirali-Shahreza, M. "A New Approach to Persian/Arabic Text Steganography." In *Proceedings of 5th IEEE/ACIS International Conference on Computer and Information Science (ICIS-COMSAR 2006)*, Honolulu, Hawaii, July 2006.
6. Kahn, D. *The Codebreakers: The Story of Secret Writing*. Rev. ed. Scribner, 1996.
7. Wayner, P. "Mimic Functions." *Cryptologia* 16, no. 3 (1992): 193–214.
8. Wayner, P. *Disappearing Cryptography*, 2d ed. Morgan Kaufmann, 2002.
9. Chapman, M., and Davida, G. "Hiding the Hidden: A Software System for Concealing Ciphertext as Innocuous Text." In *Lecture Notes in Computer Science*, Vol. 1334 *Proceedings of the International Conference on Information and Communications Security*, 335–345. Springer, 1997.
10. Grothoff, C., et al. "Translation-Based Steganography." Technical Report CSD TR# 05–009, Purdue University, 2005. Available from http://grothoff.org/christian/lit-tech.ps. Accessed 15 August 2007.
11. Grothoff, C., et al. "Translation-Based Steganography." In *Proceedings of Information Hiding Workshop (IH 2005)*, Barcelona, Spain, June 2005.
12. Stutsman, R., Grothoff, C., Atallah, M. and Grothoff, K. "Lost in Just the Translation." In *Proceedings of the 21st Annual ACM Symposium on Applied Computing (SAC'06)*, Dijon, France, April 2006.
13. Martin, A., Sapiro, G., and Seroussi, G. "Is Image Steganography Natural?" *IEEE Transactions on Image Processing* 14, no. 12 (2005): 2040–2050.
14. Cvejic, N., and Seppanen, T. "Increasing Robustness of LSB Audio Steganography Using a Novel Embedding Method." In *Proceedings of the International Conference on Information Technology: Coding and Computing (ITCC'04)*, Las Vegas, Nevada, April 2004.
15. Cvejic, N., and Seppanen, T. "Reduced Distortion Bit-Modification for LSB Audio Steganography." In *Proceedings of the 7th International Conference on Signal Processing (ICSP 04)*, Beijing, China, August 2004.
16. Bender, W., et al. "Techniques for Data Hiding." *IBM Systems Journal* 35, nos. 3 and 4 (1996): 313–336.
17. Kirovski, D., and Malvar, H. "Spread-Spectrum Audio Watermarking: Requirements, Applications, and Limitations." In *Proceedings of the 4th IEEE Workshop on Multimedia Signal Processing*, Cannes, France, October 2001.
16. Ansari, R., Malik, H., and Khokhar, A. "Data-Hiding in Audio Using Frequency-Selective Phase Alteration." In *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing,* vol. 5, 389–392, May 2004.
17. Gruhl, D., Lu, A., and Bender, W. "Echo Hiding." In *Lecture Notes in Computer Science*, Vol. 1174 *Proceedings of First International Workshop on Information Hiding*, 293–315, Springer, 1996.
18. Grune, D. "Two-Level Grammars Are More Expressive Than Type 0 Grammars or Are They?" *SIGPLAN Notices* 28, no.8 (1993): 43–45.
19. van Wijngaarden, A., et al. "Revised Report on the Algorithmic Language ALGOL 68." *Acta Informatica* 5 (1975): 1–236.
20. Koblitz, N. *A Course in Number Theory and Cryptography*. 2d ed. Springer, 1994.
21. Kessler, G. C. "An Overview of Steganography for the Computer Forensics Examiner." *Forensic Science Communications* 6, no. 3 (2004). Available from http://www.fbi.gov/hq/lab/fsc/backissu/july2004/research/2004_03_research01.htm http://www.garykessler.net/library/fsc_stego.html
22. U.S. Bureau of Labor Statistics. "Consumer Prices Index of July 2007." Available from http://www.bls.gov/news.release/cpi.nr0.htm