İleri Teknoloji Bilimleri Dergisi
Cilt 2, Sayı 3, 53-60, 2013

Journal of Advanced Technology Sciences
Vol 2, No 3, 53-60, 2013

A HYBRID STEGANOGRAPHIC APPROACH VIA WEB ADRESSES

Esra ŞATIR[1*] Oğuzhan KENDİRLİ[2]

[1] Duzce University, Fac. of Engineering, Dep. of Computer Engineering, 81620, Duzce, TURKEY
[2] Duzce University, Vocational College of Cumayeri, Dep. of Mechatironics, 81620, Duzce, TURKEY

Abstract: With the rapid growth of information technology and internet, security has become a critical issue. Therefore, data hiding gained importance for delivering secret messages. Data hiding techniques hide messages such as images, videos, texts, etc. A data hiding technique is different from cryptology. A cryptographic scheme encrypts the message and then the message is sent, which is more secure and unpredictable, to the receiver’s side. Since the message has a meaningless and uncommon content, the communication makes the observer aware of the exchange, so there is always a threat from a malicious attacker. Steganography is the art of writing secret data in such a way that no one except the intended receiver knows about the existence of secret data. Successful steganography depends on the carrier medium not to raise attention. In this study, a steganographic scheme that employs URL of web pages, has been proposed. Images have been used as the carriers. LZW coding and DES cryption algorithms have been used to increase the security. Experimental results showed that the proposed method is feasible for any communication between two parties. Since the communication is performed via only a web address, it does not raise suspicion in case of an observation.

Key Words- Steganography, data compression, cryptography, DES

1. INTRODUCTION

With the widespread use of the Internet and the booming growth of the computer industry, people nowadays can easily retrieve multimedia contents with their own computers or mobile phones over the Internet or mobile channels, under the ubiquitous computing environments. Multimedia related researches and applications have greatly increased in the last twenty years. The multimedia data is supposed to be transmitted over the Internet or the wireless networks, and the ease of delivery over the ubiquitous computing environments tends to get the multimedia contents infringed upon at any time. Data hiding is one of the useful schemes for delivering secret messages [1]. Data hiding technique hides messages such as images, videos, maps, text, etc. in the digital media imperceptibly. Data hiding technique is different from a cryptographic scheme. A cryptographic scheme encrypts the message and then the message is sent, which is more secure and unpredictable, to the receiver’s side. The receiver uses the cryptographic scheme to decrypt the message when he/she receives the encrypted messages. During the communication since everybody is aware of the exchange of messages there is always a threat from a malicious attacker [2]. However, sometimes both cryptography and steganography can be used to achieve two levels of security [3].

[*] esrasatir@duzce.edu.tr

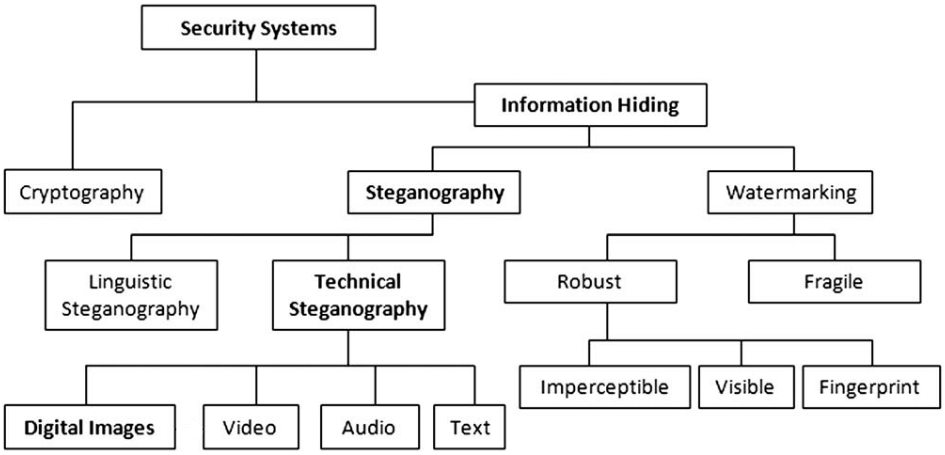


Figure 1. The different embodiment disciplines of information hiding [4].

Generally, data hiding includes digital watermarking and steganography. Watermarking is different from steganography in its main goal. Watermarking is used for copyright protection, broadcast monitoring, transaction tracking, and similar activities. It can be observed as steganography that is concentrating on high robustness and very low or almost no security. Steganography is the art of writing secret data in such a way that no one except the intended receiver knows about the existence of secret data. Successful steganography depends upon the carrier medium not to raise attention [5].

Image steganography, where information is embedded within an image has been widely used during the last decade due to the lowering of the cost of image storage and communication and also due to the weaknesses of the human visual system (HVS) [6]. Accordingly, it can be claimed that images are the most widespread carrier mediums. There are various ways of using images as carriers. For instance, properties of images can be manipulated including luminescence, contrast and colours. In some methods, the purpose is to minimize changes to the image and in some, the purpose is to store the message in a random way so as to make it more difficult to detect [7].

In general, image steganography methods can be separated into two categories: spatial domain and frequency domain based methods. In the first case, the secret message is embedded directly in the intensity of the pixels. In the second case, images are firstly transformed to frequency domain and then, the secret message is embedded in the transform coefficients. Many image file formats, such as jpeg, bmp, and gif, have been used so far in the literature for image steganography. 8-bit and 24- bit images are the most typical carriers, the first due to their small size and the second due to the high payload they offer and to the fact that the large number of colors they contain make the changes from the secret message undetectable from the human visual system [6].

The goal of a steganographic algorithms is to be more statistical undetectable. Nowadays, the detectability of secret messages is mostly influenced by two factors:

1. The selection rule used to choose the imperceptible parts of cover object that can be modified during embedding the secret bits.
2. It is better to embed as many bits of secret message as possible by changing the least number of the cover object [8].

By considering the second item, we preferred to camouflage secret message, instead of embedding it into a multimedia object as in traditional multimedia steganography. Here, the purpose is to leave the cover image unchanged by only using it as a platform that has the all possible coordinates where the characters of secret message are mapped. After mapping the characters of secret message to the cover image, we obtain a matrix that holds the coordinate information of the secret characters. Then we process this matrix via DES (Data Encryption Standart) algorithm to increase security and LZW (Lempel–Ziv–Welch) coding algorithm to increase capacity and complexity. After these operations we obtain a random array in order to use in a URL (Uniform Research Locator) of a web page that has the cover image.

Namely, we use two major elements here: One of them is the cover image (our map) while the other is the web address (coordinates) of the web page where the used cover image is demonstrated. Thus, firstly we aim to provide an unsuspicious medium for communication by using only a web address between the two parties. Secondly, we aim to render the stego-medium (cover image) resilient to any kind of attack by making no change on it.

This study has been organized as four sections. In the second section, the proposed method has been explained by both mentioning the embedding and extracting procedures. In the third section, results of the performed experiments have been provided. Finally, a general outcome has been pointed out in the last section titled as “Conclusion”.

2. THE PROPOSED METHOD

2.1. Embedding Procedure

Step 1. Let the secret message be a set of letters:

$$S = \{s_1, s_2, \dots, s_l\}$$

And let *A* be the set of numbers codes that corresponds to the ASCII (American Standard Code for Information Interchange) codes of the letters in *S*:

$$A = \{a_1, a_2, \dots, a_l\}$$

In this step, it is obvious that we obtain each element *(a)* of *A* by converting each letter *(s)* of *S* to its ASCII code equivalent.

Step 2: The operations in this step consist of ranking the candidate cover images according to the compression performances they provide. Firstly, let’s represent each image in the form of a matrix called *I* that has the dimensions of $m \times n$:

$$I = I_{m\times n} = \begin{bmatrix} i_{1,1} & \cdots & i_{1,n} \\ \vdots & \ddots & \vdots \\ i_{m,1} & \cdots & i_{m,n} \end{bmatrix}$$

For every image in the image base repeat the following sub-steps:

a) Let’s compare each element of *A* to each element of *I*, respectively from right to the left: *if* $a=i$ then we construct a second matrix that holds the coordinates (namely lines and columns) where this equation forms:

$$C = C_{l,2} = \begin{bmatrix} c_{1,1} & c_{2,1} \\ \vdots & \vdots \\ c_{1,l} & c_{2,l} \end{bmatrix}$$

Here, *l* denotes the length of *S*.

b) We encrypt *C* via DES by concatenating each line. Thus we obtain the set or array C':

$$C' = C \oplus K \qquad (1)$$

$$C' = \{c_1, c_2, \dots, c_{l'}\}$$

(Here, notice that *l'* is different from *l*, m and *n* since it changes according to the used key; *K* for encryption. The employed keys has the length of 64 bits. In Equation 1, $\oplus$ operator denotes the usage of *K* for encryption.)

c) We compress C' by employing LZW coding to reduce the size of this encrypted set (Refer to Satir E., and Isik H., 2012 for additional information about LZW coding in steganography). After LZW compression, we obtain a smaller set *T*. Namely, $|T| \leq |C'|$.

Step 3. Finally, we find the image that corresponds to the smallest *T* and we add *T* to the suitable web address that gives no error message when we type it on the address bar of the concerning web browser. Thus we have a modified URL; *U* that contains symbols of the web address. The aim of employing the smallest *T* is to represent *U* (that points to the web address which has the chosen image) in a reasonable length. That is to say, we avoid to use long web addresses as much as possible. Then we upload the image to the corresponding web site and send *U* (the modified URL) to the recipient by employing any communication channel like Skype, Facebook and etc.

Thus a web address becomes sufficient to share the secret message via internet. The web address contains the compressed and encrypted coordinates. As mentioned above, the web page to which this web address points to, carries the cover image (the map).

2.2. Extracting Procedure

As mentioned above, we have a web address that contains coordinates and a web page to which this web address points to. In fact, the entire operation is to extract the coordinate information. Then, by employing the cover image in the web page, secret message can be obtained via these coordinates.

Step 1. Get the sent web address; *U*. Decompose the added part and thus obtain *T*. Then decompress *T*, by employing LZW coding algorithm and thus we have C'; the encrypted content.

$$U = \{u_1, u_2, \dots, u_d\}$$
$$T = \{t_1, t_2, \dots, t_c\}$$

$c<d$ and $T \subset U$.

After decompression we have a bigger set, namely:

$|C'| \geq |T|$

Step 2. Decrypt C' by employing DES algorithm in order to obtain the real coordinate matrix *C*:

$$C = C' \ominus K \tag{2}$$

In Equation 2, $\ominus$ operator denotes the usage of *K* for decryption. Now we can proceed on the cover image via the elements of this matrix by getting them as (*x*,*y*) locations.

Step 3. Find each element (*a*) of *A* by using each element of *C* on *I*. Namely, get the concerning pixel value from *I* whose index corresponds to *C(x, y)*. Repeat this operation till the end of *C*. Thus we obtain *A* whose elements are the ASCII codes of *S*. Then by employing these ASCII codes, the original secret message *S* is extracted.

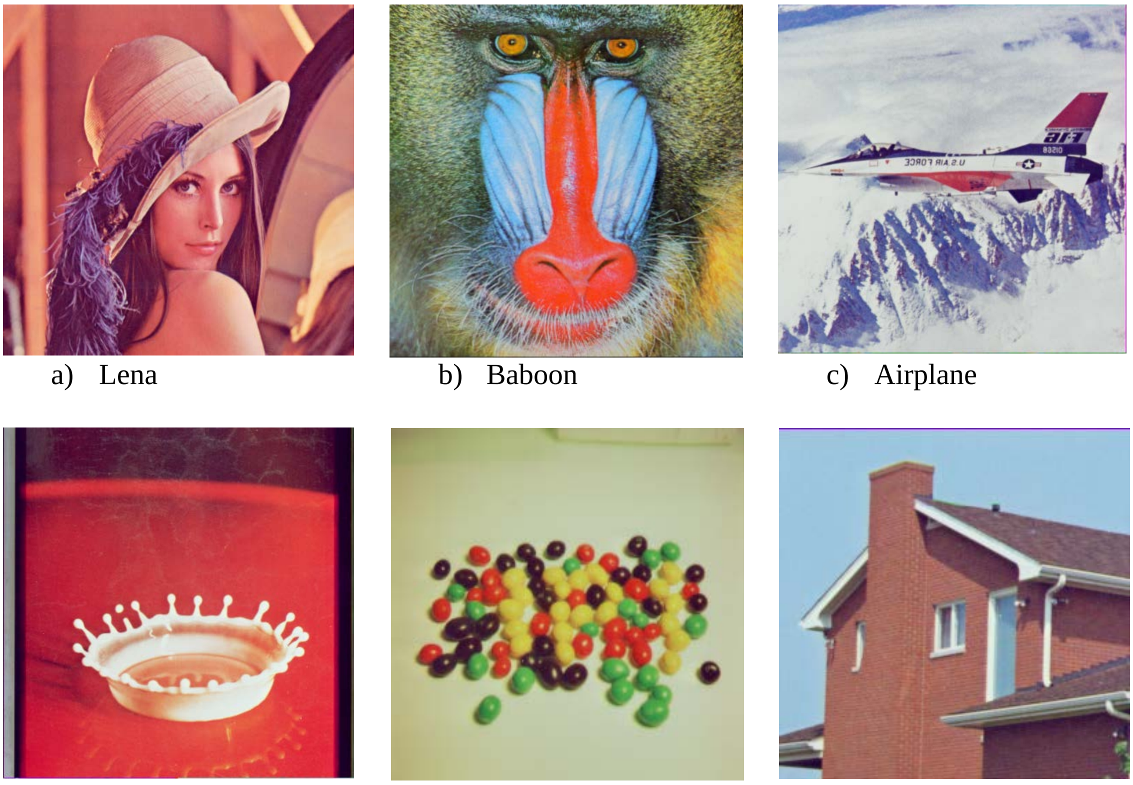
3. EXPERIMENAL RESULTS

In this section, the results of the performed experiments have been presented. The experiments have been conducted by employing the software written in *C#* programming language. Here we used the randomly generated Lorem Ipsum patterns [9] as the secret messages; *S*. As seen in Table 1, lengths of each *S* have been changed from 5 to 25, by incrementing the length five at a time. Thus, we aim to perform an unbiased evaluation. Table 1 contains the inputs and outputs of each experiment like the used *S* (secret message), *A* (ASCII codes of the characters in *S*), *C* (coordinate matrix after hiding *S* in *I*), *C'* (encrypted *C*) and finally, *T* (the compresses *C'* that is added to the chosen URL). As it can be seen in Table 1, generally, *T* is reduced by means of the compression. Thus, length of *T* is hold shortened as much as possible to make the URL of the web page seem innocent and unsuspicious.

Table 1. Details of the performed experiments

	S	***A***	***C***	***C'***	***x***	***T***	***x'***
S1	**{L,o,r,e,m}**	{76,111,114,101,109}	0 19 1 206 2 205 2 219 3 213	s09J/0nJNhzvMx+HY0nyzTeSkO6ITzjj9KfwNREiJaU=	44	s09J/0nJNhzvMx+HYąyzTeSkO6ITzjj9KfwNREiJaU=	43
S2	**{L,o,r,e,m, ,i,p,s,u}**	{76,111,114,101,109,32,105,112,115,117}	0 110 0 215 . . : : 4 133 5 12 5 258	/WOrS5siCmTO+++2g/uNv6LRDpYwkSVIerp5LyG9wMf2rGZSCL8HOM/+hXd/Cv3thdrhdpFzjsc=	76	/WOrS5siCmTO+Č2g/uNv6LRDpYwkSVIerp5LyG9wMf2rGZSCL8HOM/+hXd/Cv3thdrĿpFzjsc=	74
S3	**{L,o,r,e,m, i,p,su,m, ,d,o,l}**	{76,111,114,101,109,32,105,112,115,117,109,32,100,111,108}	0 110 0 215 . . : : 7 104 7 164 7 187	/WOrS5siCmTO+++2g/uNv6LRDpYwkSVIerp5LyG9wMf2rGZSCL8HOM/+hXd/Cv3ts5/C9A3AgvwGkgEZ2AGeHdsKJQW9NBOiuZwLugSzUKwOoY04HM0CbQ==	120	/WOrS5siCmTO+Č2g/uNv6LRDpYwkSVIerp5LyG9wMf2rGZSCL8HOM/+hXd/Cv3ts5Ĺ9A3AgvwGkgEZ2AGeHdsKJQW9NBOiuZwLugSzUKwOoY04HM0CbQ==	118
S4	**{L,o,r,e,m, i,p,su,m, ,d,o,l,o,r, ,s,i,t}**	{76,111,114,101,109,32,105,112,115,117,109,32,100,111,108,111,114,32,115,105,116}	0 110 0 215 . . : : 14 184 14 245 14 488	/WOrS5siCmTO+++2g/uNv6LRDpYwkSVIerp5LyG9wMf2rGZSCL8HOM/+hXd/Cv3ts5/C9A3AgvwGkgEZ2AGeHdsKJQW9NBOiuZwLugSzUKxoy59xs5fNQxLHpaU/3TnifgHh+YgUIMg//emB4wVK7/bk2jUy6h2hJi82cLG0xK4=	172	/WOrS5siCmTO+Č2g/uNv6LRDpYwkSVIerp5LyG9wMf2rGZSCL8HOM/+hXd/Cv3ts5Ĺ9A3AgvwGkgEZ2AGeHdsKJQW9NBOiuZwLugSzUKxoy59xĿfNQxLHpaU/3TnifgHh+YgUIMď/emB4wVK7/bk2jUy6h2hJi82cLG0xK4=	168
S5	**{L,o,r,e,m, i,p,su,m, ,d,o,l,o,r, ,s,i,t, ,a,m,e,t}**	{76,111,114,101,109,32,105,112,115,117,109,32,100,111,108,111,114,32,115,105,116,32,97,109,101,116}	0 110 0 215 . . : : 41 311 42 72 42 112	/WOrS5siCmTO+++2g/uNv6LRDpYwkSVIerp5LyG9wMf2rGZSCL8HOM/+hXd/Cv3ts5/C9A3AgvwGkgEZ2AGeHdsKJQW9NBOiuZwLugSzUKxoy59xs5fNQxLHpaU/3TnifgHh+YgUIMg//emB4wVK7/bk2jUy6h2hZh/PFylZU7zJw94r2ZhnvFt63/Rsl5OS0W9mBGnObCvCLZShA/asCA==	216	/WOrS5siCmTO+Č2g/uNv6LRDpYwkSVIerp5LyG9wMf2rGZSCL8HOM/+hXd/Cv3ts5Ĺ9A3AgvwGkgEZ2AGeHdsKJQW9NBOiuZwLugSzUKxoy59xĿfNQxLHpaU/3TnifgHh+YgUIMď/emB4wVK7/bk2jUy6h2hZh/PFylZU7zJw94r2ƜnvFt63/Rsl5OS0ŘĎGnObĻįĭhA/asCA==	206

All of the used images have been demonstrated in Figure 2 a, b, c, d, e and f. According to the proposed method, the chosen images for obtaining the shortest *T* are Airplane for S_1 and Baboon for the rest.



a) Lena b) Baboon c) Airplane

d) Splash e) Jelly beans f) House

Figure 2. Chosen images a) Lena b) Baboon c) Airplane d) Splash e) Jelly beans f) House

As an example, the constructed URL for S_1 has been given below:

http://s15.postimg.org/o8a9y3hff/s09J/0nJNhzvMx+HY%C4%85yzTeSkO6ITzjj9KfwNREiJaU=/Airplane.png

This web address is the only part that will be sent to the recipient via any communication channel like Facebook, Skype and etc. Once the recipient get this URL, he/she can obtain the secret message by applying the extracting procedure of the proposed method (refer to subsection 2.2).

4. CONCLUSION

With the rapid growth of information technology, nowadays, people can easily retrieve multimedia contents with their own computers or mobile phones over the Internet or mobile channels. Accordingly, data hiding has become one of the useful schemes for delivering secret messages. Data hiding technique hides messages such as images, videos, texts, etc. in the digital media, imperceptibly. Here, steganography is the art of writing secret data in such a way that no one except the intended receiver knows about the existence of secret data. So it is different from a cryptographic scheme since a cryptographic scheme renders the message meaningless and suspicious.

In this study, an unsuspicious and a covert communication has been targeted by means of the proposed method. Experimental results show that the implementation of the proposed method successfully performed in terms of embedding and extracting. Bu still there are some issues which need to be handled for a faster and efficient application. For instance, here the imperceptibility has been provided but the capacity issue still needs to be tackled for a more efficient and faster communication. Besides, the proposed method is targeted to be a standard algorithm in case of applying it any kind of secret message.

5. REFERENCES

[1]. Huang, H. C., and Fang, W. C., (2010). Techniques and applications of intelligent multimedia data hiding, *Telecommunication Systems*, 44(3-4):241-251.

[2]. Weng, C. Y., Tso, H. K., Wang, S. J., (2012). Steganographic data hiding in image processing using predictive differencing, *Opto-Electronics Review*, 20(2):126-133.

[3]. Swain, G., and Lenka, S. K., (2013). Steganography using two sided, three sided, and four sided side match methods, *CSI Transactions on ICT*, 1(2):127-133.

[4]. Cheddad, A., Condell, J., Curran, K., McKevitt, P., (2010). Digital image steganography: Survey and analysis of current methods, *Signal Processing*, 90(3):727-752.

[5]. Satir, E., and Isik, H., (2012). A compression-based text steganography method, *Journal of Systems and Software*, 85(10): 2385-2394.

[6]. Ioannidou, A., Halkidis, S., T., Stephanides, G., (2012). A novel technique for image steganography based on a high payload method and edge detection, *Expert Systems with Applications*, 39(14): 11517–11524.

[7]. Bailey, K., and Curran, K., (2006). An evaluation of image based steganography methods using visual inspection and automated detection techniques, *Multimedia Tools and Applications*, 31(3):327.

[8]. Fu, Y., Zhang, R., Ma, S., Qu, Z., Nıu, X., Yang, Y., (2009). Fast coding in digital steganography, *The Journal of China Universities of Posts and Telecommunications*, 16(6):92-96.

[9]. http://www.tr.lipsum.com/feed/html, (Last Accessed: 28.11.2013).