Hindawi
Security and Communication Networks
Volume 2017, Article ID 9130683, 13 pages
https://doi.org/10.1155/2017/9130683

Research Article

Protecting Information with Subcodstanography

Mirko Köhler, Ivica Lukić, and Višnja Križanović Čik

Josip Juraj Strossmayer University of Osijek Faculty of Electrical Engineering, Computer Science and Information Technology Osijek, Kneza Trpimira 2b, 31000 Osijek, Croatia

Correspondence should be addressed to Mirko Köhler; mkohler@etfos.hr

Received 20 July 2016; Revised 11 October 2016; Accepted 17 January 2017; Published 13 February 2017

Academic Editor: An Braeken

In modern communication systems, one of the most challenging tasks involves the implementation of adequate methods for successful and secure transfer of confidential digital information over an unsecured communication channel. Many encryption algorithms have been developed for protection of confidential information. However, over time, flaws have been discovered even with the most sophisticated encryption algorithms. Each encryption algorithm can be decrypted within sufficient time and with sufficient resources. The possibility of decryption has increased with the development of computer technology since available computer speeds enable the decryption process based on the exhaustive data search. This has led to the development of steganography, a science which attempts to hide the very existence of confidential information. However, the stenography also has its disadvantages, listed in the paper. Hence, a new method which combines the favourable properties of cryptography based on substitution encryption and stenography is analysed in the paper. The ability of hiding the existence of confidential information comes from steganography and its encryption using a coding table makes its content undecipherable. This synergy greatly improves protection of confidential information.

1. Introduction

Every confidential information that is sent through an unprotected communication channel should be protected from unauthorized access by third parties. The process of sending confidential information between the sender and the recipient through an unprotected communication channel is shown in Figure 1.

The techniques for secure communication in the presence of third parties are studied within the cryptography. The confidential information could be protected by encryption process. The confidential information is converted into encrypted form by applying an adequate encryption algorithm. The content of the encrypted information is undecipherable without the adequate cryptography key.

However, the encrypted information is not hidden. With sufficient time and computing resources, the encryption algorithm can be decrypted by applying a proper decryption algorithm, and the content of confidential information can be revealed. Each new developed encryption algorithm possesses some advantage; however, each encryption algorithm can be decrypted within sufficient time and with sufficient resources. The only exception is the "one time pad" algorithm which uses a temporary card or pad that is immediately destroyed after its use.

Another approach for protecting the confidential information is to hide the existence of the information using steganography. Steganography is a science of writing secret messages. Thus, nobody besides the sender and the intended recipient knows about the existence of the message. If the existence of hidden information in some steganography file is revealed, its content can easily be read. The advantage of steganography is the fact that a hidden message does not attract attention. On the other side, encrypted messages, no matter how undecipherable, will induce suspicion. Thus, while cryptography protects the contents of the message, the steganography hides the message itself.

The existence of the stego-images, images with embedded information, can be detected by a specialized algorithm, named steganalysis. Some previously conducted researches show that the most of the steganographic algorithms have been detected by steganalysis algorithms and that more robust information protection approaches should be used.

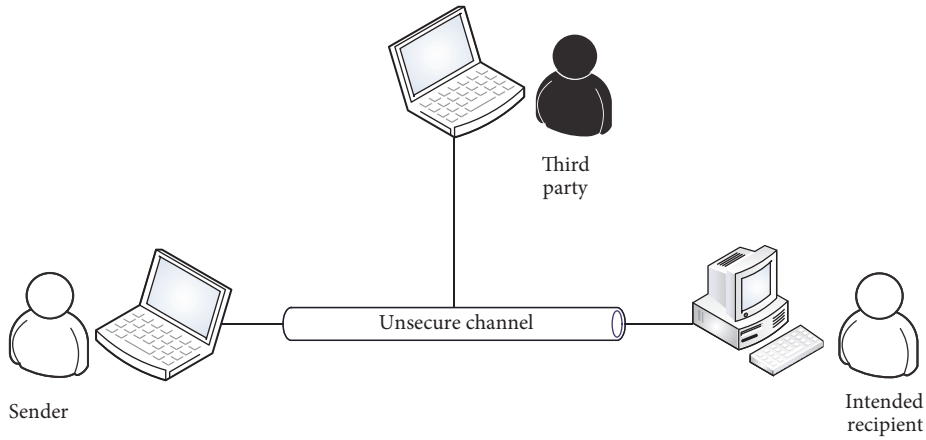


Figure 1: Persons involved in information exchange.

In this paper, encryption and steganography techniques have been analysed. Their strengths and weaknesses have been explained. Furthermore, in order to achieve greater security of information which travels through an unsecure communication channel, favourable properties of those techniques have been considered, and a novel method for integrating cryptography based on substitution encryption using a coding table and steganography is proposed. Finally, this novel method, named subcodstanography, is presented and its features have been explained.

2. Information Security with Cryptography and Steganography

In cryptography, encryption is the process of transforming information using encryption algorithm to make it unreadable and incomprehensible to everyone except those who possess knowledge of the encryption algorithm and encryption key. The result of encryption is encrypted information. The most popular methods of encryption are unkeyed cryptosystems, secret key cryptosystems, and public key cryptosystems [1]. Although these algorithms provide a certain level of security, there are ways in which they can be decrypted and the hidden information read.

Steganography is the science of writing hidden messages in such a way that no one except the sender and the intended recipient suspects the existence of the hidden message. Steganography derived from the Greek words steganos and grapheine, meaning hidden writing. Modern steganography uses the advantages of digital technology and often involves hiding a message within a certain multimedia files such as pictures, audio, or video files. These files are an integral part of usual communication and as such do not induce attention to them.

Different types of steganography exist. In this paper, the application of steganography in digital images is examined. Other types of steganography, such as linguistic or audio, are not included in the analysis. The most of the developed steganography techniques were set up to exploit the structures of the most popular image formats (GIF, JPEG, and PNG), as well as of the bitmap format (BMP) due to its simple data structure.

Image files typically contain unused or less important bits which can be manipulated by various steganography techniques and replace them with the confidential information. Such files can be exchanged without provoking doubt about true purpose of communication.

2.1. Steganography Techniques. Through the history, various steganography techniques were applied [2]. With the advancement of technology and computer power, new techniques were created and introduced, as shown in [3–5]. The main categorization of steganographic techniques in digital images focuses on spatial domain methods, frequency domain methods, and adaptive methods. The list of steganographic methods and techniques is presented in the list below based on the classification given in [6].

Image Steganography Classification

Spatial domain methods are as follows:

LSB encoding
LSB replacement
LSB matching
LSB matching revisited
Pixel value differencing method
Singular value decomposition method
Histogram shifting method

Frequency domain methods are as follows:

Discrete Fourier transform (DFT) method
Discrete cosine transform (DCT) method
Discrete wavelet transform (DWT) method
Integer wavelet transform (IWT) method

The spatial domain techniques generally use a direct least significant bit (LSB) replacement method. This technique, although simpler, has a larger impact compared to the other

two types of techniques. BMP and GIF based steganography apply LSB techniques. Although their resistance to statistical counterattacks and compression are reported to be weak, as stated, for example, in [7], using BMP instead of JPEG images is proposed, for example, in [8]. Moreover, JPEG images were avoided since the changes as small as flipping the LSB of a pixel in a JPEG image can be detected, as shown in [9].

Furthermore, the frequency domain based techniques generally use a discrete cosine transform (DCT), discrete Fourier transform (DFT), and discrete wavelet transform (DWT). Finally, adaptive techniques can either be applied in the spatial or frequency domains.

In researches related to digital images steganography in the spatial domain, different techniques were proposed. For instance, instead of proposing an embedding technique, a spatial domain technique in producing a fingerprinted secret sharing steganography for robustness against image cropping attacks is used in [10]. In addition, another used data hiding scheme histogram-based data hiding, given in [11], proposed lossless data hiding, using the difference value of adjacent pixels. Moreover, an alphabet punctuation for hiding messages is exploited in [12]. Furthermore, several proposed methods are based on the least significant bit (LSB) replacement approach. One of them is the colour palette based steganography. The LSBs are modified based on their positions in the palette index.

In researches related to digital images adaptive steganography, several techniques related to the LSB replacement have been analysed. For example, in [13, 14], the proposed methods take into consideration statistical features of an image before interacting with its LSB coefficients. Furthermore, an adaptive technique applied to the LSB substitution method is proposed in [15], in which the correlation among neighbouring pixels is exploited and used to estimate the degree of smoothness. Moreover, in [16], the image embedding based on segmenting homogenous grayscale areas using a watershed method is proposed.

The importance of robustness in steganography system design provokes different opinions. In [17], steganography is defined as a process that should not consider robustness since then it is difficult to differentiate it from watermarking. In [18], on the other hand, robustness is defined as a practical requirement for a steganography system.

Furthermore, there has always been a trade-off between robustness and payload. In [19], the three kinds of encoding format (Hexadecimal, Based-64 and ACSII code) in the proposed system were compared and analysed. Among them, ACSII encoding format is proven to be the most efficient for encrypting large plaintext message.

In addition, robustness against high quality of image is also an important issue. The frequency domain and adaptive steganography techniques are not too prone to attacks, especially when the hidden message is small. It is so because they alter coefficients in the transform domain, and, thus, image distortion is kept to a minimum.

However, these methods have a lower payload compared to spatial domain algorithms. It can be noticed that, compared to the embedding in the 1st LSB, embedding in the 4th LSB generates more visual distortion to the cover image as the hidden information is seen as unnatural [20]. The trade-off between the payload and the quality of image distortion is present. However the payload, embedding up to the *n*th LSB, is analogous with respect to the recovered embedded image.

2.2. Substitution Method. The substitution method replaces some or all redundant components in media files with confidential information. The aim is to replace the file with redundant components by already encrypted information.

In order to detect hidden message, the true content of the file must first be suspected. After that, a part of the media file in which the true content is hidden should be determined and the distribution of information bits in the file encoded. At this point, the decryption can begin.

In order to understand this principle, it is important to know the structure of file used in steganography. For example, the detailed description of bitmap's RGB (i.e., Red-Green-Blue) system, the impact of inserting additional information on visual information in image, and the explanation of the detailed process of inserting information in a BMP image are given in [21].

A secondary safety measure is presented in this paper. In simple terms, it is a function that inserts series of bits from the first step (encryption) into an original file (carrier), such as a bitmap image. The original file is selected from the multimedia database. The way in which these series of bits are inserted into original file could be changed every time for a new hiding of information to prevent steganalysis methods. Steganalysis methods detect the used steganography method and key and are described in the next chapter.

2.3. Steganalysis Attacks and Countermeasures. Steganalysis is the process of detecting steganography files based on studying variation patterns of bits. The objectives of steganalysis are to identify suspicious data sets, such as files which can carry hidden information and to extract them from a steganography file. Unlike cryptanalysis, where the existence of encrypted messages is evident, steganalysis usually starts with several suspicious data sets that might contain a secret message. Using various advanced methods of statistical analysis, presented in [22, 23], a steganalyst can reduce the set of suspicious data until the right steganography file is found. Information could be hidden on any public source on the Internet, and this greatly complicates the process of steganalysis.

Steganalytical analysis and attack on hidden communication includes various activities such as detection, isolation, and disabling or deleting hidden information. The type of attack depends on the resources available to the steganalyst. The first type of attack is carried out if the steganalyst disposes only with a steganography file that carries the message, while the second type of attack is carried out when beside steganography file steganalyst possesses original file as well. The third type of attack can be carried out if the steganalyst has both the steganography file and the algorithm used to insert a secret message [21].

The analysis of repeating patterns can be used to identify the steganography method or even hidden information. The

examination of patterns compares the original steganography carrier with the steganography file that contains a hidden message. Such attack is called an attack with a known carrier. Therefore, each new message should use a new original carrier file (original file). The used original file should be deleted after inserting the information. This prevents the second type of steganalytical attack and represents the third step in increasing the security process.

2.4. Cryptography and Steganography Interaction. The existing steganographic methods rely on the secret key and the robustness of the steganographic algorithm. However, no single steganalysis algorithm is constantly superior, as proven in [24]. Moreover, the existing steganographic techniques do not address the issue of encryption of the payload prior to embedding. The interaction between the cryptography and steganalysis is not yet very well researched, as noted in [25].

Several researches covering the interaction between the cryptography and steganography have been conducted in the recent years. For instance, in [26], frequency domain steganography, that is, discrete wavelet transforms (DWT) based steganography, is used, and the filter bank cipher is used to encrypt the secret text message. In addition, in [19], the frequency domain steganography, that is, the discrete cosine transform (DCT) based steganography, is used. In [27], a secret message is embedded in more than one JPEG format image. Hence, in order to recover the secret message, a steganalyst has to determine all stego-objects and unravel the algorithm used to hide the secret message in them. In [28], a combination of cryptography and steganography was achieved by using the DES algorithm and the LSB technique.

The encryption of the payload prior to embedding is discussed in [29, 30], and the basic notes that should be observed by a steganographer are defined. First, in order to eliminate the attack of comparing the original image file with the stego-image, it is advised that a completely new image is created and destroyed after generating the stego-image. The cover image must be carefully selected. Also, a familiar image should not be used. Hence, steganographers should create their own images [31]. Furthermore, in order to avoid a visual perceptual attack, the generated stego-image must not have visual distortions. An alteration made up to the 4th LSB of a given pixel could yield a dramatic change in its value, and this would thwart the perceptual security of the transmission. Finally, the secret data must be composed of balanced bit values, since the expected probabilities of bit 0 and bit 1 for a typical cover image are the same [32]. In some cases, encryption provides such a balance, for example, in the case of the parity check.

3. Subcodstanography Method

In this paper, a novel method for integrating cryptography based on substitution encryption using coding table and steganography is proposed. The features of the proposed approach are given as follows.

(a) The existence of hidden information must first be suspected which is achieved by using steganography.

(b) After that, the used steganography method and key must be discovered. Each new hidden information uses a new steganography key and a new steganography file.

(c) If the first two steps are detected and the hidden message is read, as a result a series of encrypted data is obtained. In this step, the used encryption algorithm, which is also different for each encryption, must be detected.

(d) In the end, if the identification of an encryption algorithm is successful, the used encryption key must be discovered.

In this paper, it is assumed that each hidden information has different steganography methods, steganography keys, encryption algorithms, and encryption keys. If a proposed method is carried out, and if one message is decrypted, only the information in that message can be read, while the hidden information in other messages remains inaccessible, since all the above-mentioned steps are different for each message.

The process of information hiding is presented step by step in the following chapter. Inserting information into bitmap image is selected, and a coding table is used for encryption. The entire process is shown in Figure 2 and named "subcodstanography."

3.1. Coding Table Construction for Substitution Encryption. In this paper, the encryption algorithm is used as the first security measure for protecting information. The encryption proposed in this paper uses a coding table. The coding table is secret and randomly generated for each new encryption. It contains a sequence of bits which are used instead of the letters, numbers, or other characters. For each symbol ("A", "j", " ", "2") there are several different sequences, depending on how many bits are used to represent each symbol in the coding table. For each symbol there may be several different sequences, so the commonly used symbols cannot be found or searched with the help of statistical value. After all symbols in the message are replaced with a series of bits from the coding tables, the first security measure of encryption is completed.

A coding table consists of all possible ASCII symbols. As previously stated, each symbol is composed of several different sequences, which are generated by a random function, depending on how many bits are used to store a single symbol in the coding table, as shown in Figure 3. In the table, four different combinations of bits display the letter "A", as well as the character "!", which is the most rarely used.

This table uses a symmetric key encryption, in which all symbols are represented with the same number of different coded values. It was chosen in this application for simplicity reasons.

From the example in Figure 3, it is visible that the one ASCII symbol (letter "l") is presented with different combination of bits in different places in the message. The attackers can neither find the patterns nor the most frequently used characters, using the familiar values of the statistical occurrence of certain characters in a particular language.

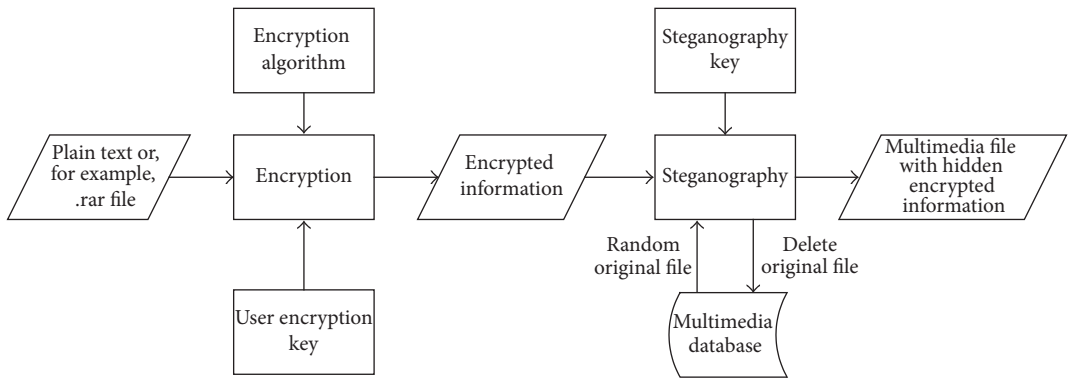


Figure 2: Process of hiding information.

Plain text: Hello world!

Encoding table

" "	0100001111	1000110010	1110001100	0011100101
"A"	1010101010	0110000001	1000001101	0111011001
		...		
"H"	1000111111	0111000011	0001111000	0000100000
		...		
"d"	1010000110	1001001110	1000000111	1110000000
"e"	0000011100	0000000001	1000001110	0111000100
		...		
"l"	1100111001	0101011100	0001111100	1000000110
"o"	0001111100	1000111010	1001010101	1010001011
"p"	1111001111	0110011100	1110110000	1110010101
"r"	1000101101	1111111100	1100000010	0000111100
		...		
"w"	0011110001	1100111000	0011000111	0011111001
		...		
"!"	0101101001	1110011111	1110011010	1001110100

"H"	"e"	"l"	"l"	"o"	" "
0111000011	0000011100	0101011100	1100111001	0001111100	1000110010

"w"	"o"	"r"	"l"	"d"	"!"
0011110001	1001010101	1111111100	0001111100	1001001110	1110011111

Figure 3: Example of coding table.

For example, when the ASCII characters are presented with 8 bits, there are 256 combinations. If each symbol in the coding table is presented with 10 bits, there are 1024 combinations. Thus each symbol could be presented with four different coded values. Depending on the needs and the level of security that should be achieved, by increasing the number of bits in the coding table more combinations for each symbol can be obtained. It is recommended to use an asymmetric coding table, in which the symbols that appear more often are coded with more mutually different combinations than the symbols that appear less frequently.

Another advantage of this encryption method is that for each new encryption a new coding table is used, with new randomly generated sequences of bits. Therefore, the message cannot be encrypted without coding table, unless possessing a lot of time and resources. The sender and the intended recipient do not know which coding table is used in the process of encryption. Only the computer encryption program knows which coding table is used.

3.2. Inserting Encrypted Information into Image. The next step after encrypting confidential information using a coding table is inserting encrypted information into the BMP image. Steganography is used to hide an encrypted message inside a media file. The selected carrier is bitmap file, in the given example. The bitmap file was selected for its size, which allows an insertion of larger amounts of data. The technique of information inserting into a bitmap file is explained in [21]. In that paper, different algorithms for inserting encrypted information were presented, and the effects on changes in visual information of the bitmap file are explained. Various steganography methods of inserting information and new

bitmap file are used for each new process of information hiding. The process increases security since the place inside the media file where the information was hidden is not predefined.

In the previous chapter, steganalysis method was presented. To detect the hidden information, either the steganography and the original files or the used steganography algorithm, should be known. The aim is to contact any of these data. The original file is deleted from the multimedia database after the insertion of confidential information. Thus modified bits using steganalysis method are hard to find.

Using the methods of exhaustive data search, all possible combinations of bits inserted in the media file may be attempted to be found. This will produce a set of incomprehensible bits. Only one among all combinations is the real message, and it is still undecipherable because of the encrypted inserted message.

The size of encrypted information (SEI) is calculated by multiplying the number of symbols (NS) and the number of bits in the coding table (NBCT), as shown in the following equation:

$$\text{SEI} = \text{NS} \times \text{NBCT}, \tag{1}$$

where SEI is the size of encoded information, NS the number of symbols, and NBCT the number of bits in the coding table.

For example, to hide the message "Hello World!," which is coded with a 10-bit symmetric coding table, the size of encrypted information can be calculated according to (1), and the result is 120 bits.

The number of bits that can be used for steganography (NSB) in the bitmap file can be calculated using the following equation:

$$\text{NSB} = \text{height} \times \text{width} \times \text{color} \times \text{bits}, \tag{2}$$

where NSB is the number of steganography bits, height × width is the bitmap dimensions, color is the number of colors, and bits is the number of bits useful for steganography. The total number of steganography combinations of the bitmap file can be calculated using the following equation:

$$\text{NSC} = \binom{\text{NSB}}{\text{SEI}}, \tag{3}$$

where NSC is the number of steganography combinations.

If the information is inserted inside the 24-bit bitmap file size of 100 × 100 pixels, then according to (2) 120.000 bits are suitable for steganography manipulation, where the number of colors is three and the number of bits is four [21]. However, to insert the message "Hello World!" only 120 bits are necessary. Equation (3) indicates that there are $4.47 \cdot 10^{410}$ different combinations in which the encrypted message "Hello World!" can be inserted into the image. Thus, using the method of the overall search, $4.47 \cdot 10^{410}$ different 120 bit encrypted messages should be analysed. Each message should be decoded, and without a coding table this task is almost impossible.

The inability to decrypt a message lies in the fact that the length of the message inserted into a bitmap is also unknown. In this example, it is stated that the message is 120 bits long; the length of the message is unknown. Likewise, when a certain message is read, it is unknown how many bits are used in the message to present each ASCII symbol, because the number of bits used in the coding table is not known. Therefore, the variation of bit length in the coding table must be taken into consideration. In this case, the message length is 120 bits, and therefore 8, 10, and 12 bits (or some other number divisible by 120) must be assumed for encoding. This means that the number of encrypted messages in this case should be multiplied by three and the total number of combinations is $13.41 * 10^{410}$. It is important to note for messages longer than 120 bits that there is a greater number of the dividers and thus the coding table may use more than 12 bits to represent each ASCII symbol.

3.3. Decoding Process. The subcodstanography method is explained in the last chapter, unlike the decoding process. Sender neither knows the steganography method nor applied key nor encryption algorithm. The intended recipient is also not familiar with this information. That reduces the possibility that someone else reads the information. It follows that information about encryption method must be placed inside steganography file along with hidden message. There are several ways how this can be done. It is important that this information must be hidden and unknown to both, the sender and recipient.

One way of sending steganography key is to insert data in the file title. The software solution for this method stores all possible keys for all used algorithms. These keys possess their own codeword, which is implemented in a title of steganography file (in this method). When steganography file is received, decoding program first reads the steganography key from the file title and recognizes the used algorithm and location of bits in a file where the message is hidden. Hence, the file title should be unobtrusive. It is possible to hide the key in several places in the file to ensure its consistency and to prevent manipulation with steganography file title.

Besides the hidden message, the codeword for used encryption algorithm is embedded in steganography file. When this codeword is inserted into a steganography file, it is necessary to send it to the recipient. When recipient receives the sent file, decoding program uses program part for codeword recognition and finds the matching algorithm. The last step in the decryption process is to read the encrypted message. The process of decryption is shown in Figure 4.

The used bitmap file can be sent together with a number of other images. The same procedure refers to other types of multimedia files as well. Furthermore, since many files are sent, the location of the hidden message is unknown, and, in this way, it becomes very difficult to find the hidden message. File titles of all images sent with steganography file should be sent in the series and in that way mask information hidden in the title of a steganography file.

When the media files are received, they are loaded into decoding program which traces the files containing hidden message. After that, program checks whether the steganography key attached in the file name matches the one in

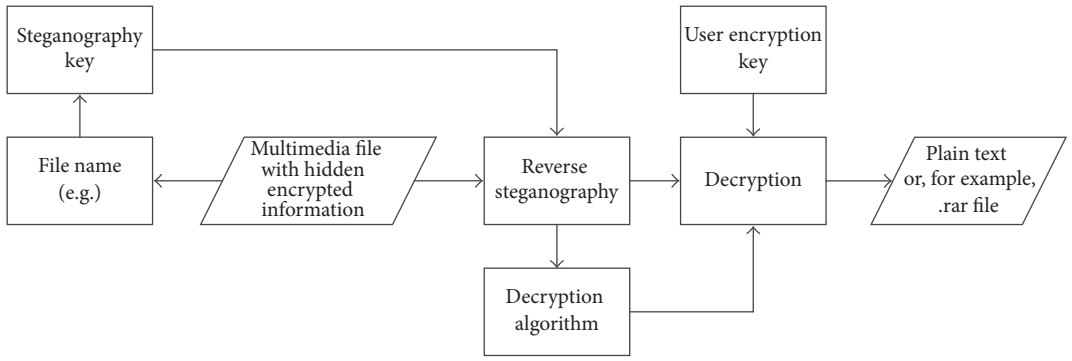


Figure 4: Information retrieving process.

the image and finds the codeword of the used encryption algorithm. If the keys are equal and if the label of the algorithm is recognized, hidden message is read from the file. The secret information is decrypted using the algorithm codeword and the decryption key. This completes the process of transmitting the secret information and obtaining of the original information. Another advantage of this encryption method is that, for each new encryption, a new coding table with the new randomly generated sequences of bits is used. Therefore, the message cannot be decrypted without the coding table.

4. Evaluation of the Proposed Subcodstanography Scheme

For the purpose of evaluating the proposed scheme, the impact of inserting different amount of information into BMP image will be shown. As it is already shown in [21], BMP images are suitable for importing a large amount of information. In this paper, the analyses of the three conducted experiments are presented. In the first experiment, only the last bit of red colour component of each pixel in image has been replaced with a 10-bit code word, as presented in Figure 3. In the second experiment, the same amount of coded information has been inserted into the place of the third least significant bit of each pixel of blue colour component. In the third experiment, the RAR file has been inserted into the places of all four least significant bits of each colour component in the picture. In this experiment, each group of eight bits in RAR file has been coded with a 10-bit code word, as presented in Figure 3.

The chosen BMP image is named "original.bmp". Its size is 400 × 400 pixels. It has been stored with a 24-bit color depth. The header has been separated and the proposed changes have been made. Each image obtained from experiment has been saved as a separate BMP file under a name "experiment" followed by the ordinal number of the experiment. The original image is shown in Figure 5.

In order to present the impact of insertion of information into picture, (5) has been used. The resulting images have been obtained using binary subtraction process conducted between images given from experiments and the original image. The subtraction process between the original image and each image obtained from experiments has been conducted using the following method:

Figure 5: The original image.

$$A = \text{width} \times \text{height}; \quad \text{for (all } i < A)$$
$$\text{difference} = \text{original XOR experiment}. \tag{4}$$

The percentage change value (CV) of picture content is calculated using the following equation:

$$\text{CV} = \frac{\sum 2^i \times b_i}{255 \times 3 \times \text{width} \times \text{height}} \times 100\,(\%)\,, \tag{5}$$

where b_i presents the value of bit on ith place ($i = 0, \ldots, 7$). The value of individual bits is described in [21].

4.1. The First Experiment. The information that has been inserted into default image original.bmp is in fact *Lorem Ipsum*, which contains 2,644 characters with spacing. These

Figure 6: The image resulting from the first experiment.

Figure 7: The image resulting from the second experiment.

characters have been coded with the 10-bit code words, so the total number of bits that has been inserted into picture is 26,440.

The obtained image is shown in Figure 6. The percentage change value of the original image is 0.0132%. From Figure 6 it is obvious that the conducted process of insertion of information into original image has not affected the visual perception of the image. Without the original image it is not possible to detect the inserted information since there is nothing to compare the changed image with.

4.2. The Second Experiment. The same information used in the first experiment has been inserted into default image original.bmp in the second experiment as well. These characters have also been coded with the 10-bit code, and the number of bits that has been inserted into picture is the same as in the previous experiment. The only difference between the processes conducted in the first and in the second experiments is in the chosen colour component and the significance of the changed bit.

The obtained image is shown in Figure 7. The percentage change value of the original image is 0,0987%. From Figure 7 it is obvious that the conducted process of insertion of information into original image also did not affect the visual perception of the image.

4.3. The Third Experiment. The information that has been inserted into original image is the RAR file having size of 171 kB. The information has also been coded with the 10-bit code words and inserted into the places of all four least significant bits of each colour component. In this experiment, every group of eight bits of RAR file have been coded with ten bits.

The obtained results are shown in Figure 8. The percentage change value of the original image is 2,7854% but from

Figure 8: The image resulting from the third experiment.

Figure 8 it is clear that the process of inserting information into original image has not affected the visual perception of the image.

4.4. The Evaluation of the Proposed Schemes. By observing experiments and the given results it is clear that the additional information can be inserted into original image without compromising visual information of the image. The third experiment has shown that it is possible to insert almost 50% of additional information in order to have percentage change value in the image lower than 3%. The overview of the given results is presented in Table 1.

Table 1: The comparison of the proposed schemes.

Experiment	Number of inserted bits	Number of used bits in steganography process	Result CV(%)
The first	26,440	1	0,0132
The second	26,440	1	0,0987
The third	1,712,597	12	2,7854

4.5. Comparison of the Proposed Scheme with Other Existing Steganography Schemes. In this chapter, the proposed scheme has been categorized and compared to other existing steganography schemes from several different aspects.

4.5.1. Categorization of Proposed Scheme. The main categorization of the proposed steganographic scheme is given in the list below. The proposed scheme fits into spatial domain methods group within image steganography techniques. In order to avoid the introduction of distortions in image which may be percepted by human vision, in the proposed scheme the information embedding process has been carried out in those bits which carry least weight (LSBs).

Scheme categorization is as follows:

(i) Text steganography

(ii) Video steganography

(iii) Audio steganography

(iv) Image steganography

(a) Frequency domain methods

(b) Spatial domain methods

The proposed scheme base

LSB encoding/replacement

4.5.2. Complexity of Implementation. There are a number of ways to hide information in digital images. Common approaches include the following processes:

(i) Insertion

(ii) Masking

(iii) Filtering or

(iv) Transformation

The proposed scheme uses the process of information insertion in digital image. The implementation of the proposed scheme is less complicated approach compared to other spatial domain methods since, within the process of secret message embedding, the original image does not have to be divided into carefully chosen blocks. Also, unlike other image steganography methods in the frequency domain, the proposed scheme does not have to convert time and space frequency components into frequency domain.

4.5.3. Classification Considering Digital Images Used for Steganography. Digital images used for steganography differ in their format types. Moreover, information can be hidden in these images in many different ways. To hide the secret information, the following techniques can be used:

(i) The straight message insertion in the image

(ii) The selectively embedding the message into noisy areas of image that draw less attention or

(iii) The insertion of message scattered randomly throughout the image

Each of these techniques can be applied, with varying degrees of success, to different image files. In this paper, embedding information in the places of the least significant bits (LSB) of BMP image is chosen due to its simple data structure. To hide an image in the LSBs of each byte of image, the following images can be used:

(i) The coloured 24-bit images or

(ii) The grayscale 8-bit images

The chosen method uses a 24-bit image instead of 8-bit image. In this way it is possible to store three bits in each pixel. Hence, as it is proven in the conducted experiments in this paper, when the message is compressed before embedding into image, a large amount of information can be hidden. The resulting stego-image looks identical to the original image and it is not possible to visually discern it. The 8-bit images are not as forgiving to LSB manipulation because of colour limitations. To effectively hide information in 8-bit images, the original image must be carefully selected so that the stego-image will not uncover the existence of an embedded message.

4.5.4. Comparison of Positions of Embedded Bits in Digital Image. The following analysis is based on the results given in [21]. Its objective is placed to define the best scenarios for hiding information in original image based on the positions of the replaced bits. For conducted experiments, the BMP original image sized 400×400 pixels and stored with a 24-bit color depth is selected (Figure 9). Every image obtained from experiments (Figures 10(a)–10(d)) has been binary subtracted from the original image and the results are saved as new BMP file (Figures 11(a)–11(d)).

The following experiments have been conducted in [21].

(a) Experiment Number 1. The stochastic information of 160.000 bits is inserted in the default image on the position of last bit of blue colour component in every pixel (Figure 10(a)).

(b) Experiment Number 2. The stochastic information of 480.000 bits is inserted in the default image on the position

Figure 9: The chosen original image [21].

of the sixth bit of blue, the seventh bit of green, and the eighth bit of red colour component in every pixel (Figure 10(b)).

(c) Experiment Number 3. The stochastic information of 1.920.000 bits is inserted in the default image on the positions of the last four bits of each colour component in every pixel (Figure 10(c)).

(d) Experiment Number 4. The stochastic information of 1.920.000 bits is inserted in the default image on the positions of all odd bits of each colour component in every pixel (Figure 10(d)).

The given results show that using more bits from the most significant bits part is not recommended. An alteration made up to the 4th LSB of a given pixel could yield a dramatic change in its value, and this would thwart the perceptual security of the transmission. As shown in experiment number 4, by inserting information on position of all odd bits of each component in every pixel, CV is around 33%, what is clearly visible on coded image (Figure 10(d)).

4.5.5. Comparison Based on PSNR Values. The peak signal-to-noise ratio (PSNR) presents the ratio between the maximum possible power of a signal and the power of corrupting noise that affects the fidelity of signal representation. Since signals have a very wide dynamic range, the PSNR is expressed in terms of the logarithmic decibel scale.

The PSNR is most easily defined via the mean squared error (MSE). The PSNR (in dB) is defined as

$$\mathrm{PSNR} = 10 \cdot \log_{10}\left(\frac{\mathrm{MAX}_I^2}{\mathrm{MSE}}\right). \tag{6}$$

Table 2: Comparison of PSNR values.

Experiment	Peak signal-to-noise ratio (PSNR) value
Experiment number 1	58.00 dB
Experiment number 2	42.69 dB
Experiment number 3	31.84 dB
Experiment number 4	8.83 dB

For example, for $m \times n$ monochrome image I and its noisy approximation K, the MSE is defined as

$$\mathrm{MSE} = \frac{1}{mn}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}\left[I(i,j) - K(i,j)\right]^2. \tag{7}$$

MAX_I is the maximum possible pixel value of the image.

The PSNR values are calculated for the results of the experiments numbers 1–4 presented in Figures 10(a)–10(d) and shown in Table 2. Higher PSNR values indicate that the reconstruction is of higher quality.

The given results of experiment 1–experiment 3 fit into the typical range for a 24-bit depth images. The results of experiment 4 do not fit which proves that an alteration made up to the 4th LSB of a given pixel could yield a dramatic change in its value and is not recommended since it would thwart the image perception.

4.5.6. Comparison Based on Number of Used Coding Bits. When the ASCII characters are presented with the 7 bits, there are 128 different combinations. Generally, the secret data should be composed of balanced bit values, since the expected probabilities of bit 0 and bit 1 for a typical original image are the same. In some cases, the encryption provides such a balance, for example, in the case of the parity check. In the proposed scheme characters are presented with 8 bits (7 information bits and 1 parity bit). There are 256 different combinations, and the encryption provides a balance.

4.5.7. Comparison Based on Encryption of the Payload Prior to Information Embedding. The existing steganographic techniques do not address the issue of encryption of the payload prior to embedding. In the proposed scheme the encryption of the payload is conducted prior to embedding. Each symbol in the coding table is presented with 10 bits, so there are 1024 different combinations. Thus each symbol could be presented with four different coded values. Depending on the needs and the level of security that should be achieved, by increasing the number of bits in the coding table more combinations for each symbol can be obtained, as presented in Table 3.

Another advantage of this encryption method is that for each new encryption a new coding table is used, with new randomly generated sequences of bits. Therefore, the message cannot be encrypted without coding table, unless when possessing a lot of time and resources.

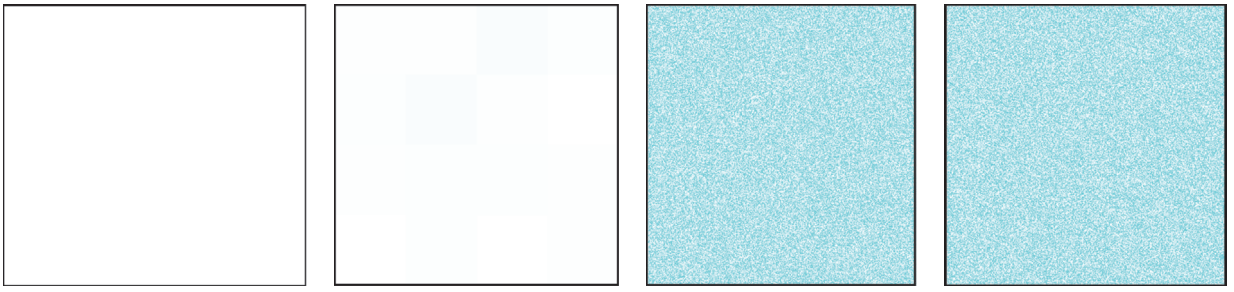
It is recommended to use an asymmetric coding table, in which the symbols that appear more often are coded with more mutually different combinations than the symbols that appear less frequently.

Table 3: Comparison of code word lengths.

Other schemes		Proposed scheme		
Number of bits per symbol	Number of combinations	Number of bits per symbol	Number of combinations	Number of code words for each symbol
7	128	10	1024	8
8	**256**	**10**	**1024**	**4**
8	256	11	2048	8

(a) (b) (c) (d)

Figure 10: The given image.



(a) (b) (c) (d)

Figure 11: The binary subtracted image.

This table uses a symmetric key encryption, in which all symbols are represented with the same number of different coded values. It was chosen in this application for simplicity reasons. The advancement of this method could be conducted using different number of coded values.

5. Conclusion

In this paper, the idea of combining encryption and steganalysis methods is examined in a way in which the favourable properties of both methods are exploited and a method for a secure transfer of confidential information over an unsecured communication channel called "subcodstanography" is proposed. Encryption method that uses a coding table is proposed. In this method, plain text is encrypted by one of many possible codes that are changed with each new encoding. The advantage of the coding table is reflected in the fact that a different number of bits can be used to represent the same ASCII symbol. By increasing the number of bits, the encrypting system increases the number of combinations that can be used to replace an individual ASCII symbol applying for the same character in the encrypted message.

Since it is proven that the ACSII encoding format is the most efficient for large plaintext message encryption, it can be concluded that the combination of embedding information presented using ASCII in BMP image presents an adequate choice when the picture size is large and when the transmission speed is not crucial.

The next step in protection of confidential information in the proposed method is insertion of encrypted message into carrier file. Multimedia files were used for the carrier file, because they have widespread use and enough bits on which information can be inserted and are unobtrusive. In this paper, a steganography technique uses information insertion in a bitmap file. This method is chosen because it is well explained and documented in the previously published paper.

The advantage of this method is explained within the demonstrated example in which the overall search of

steganography file gives $13.41 \cdot 10^{410}$ different combinations which might carry a message. Stated number of combinations is for a 100×100 pixels bitmap image, while, for larger bitmaps, this number exponentially increases since the greater number of bits can be used to hide the information. After all possible combinations are given, only coded data are available. It is virtually impossible to decode data if the coding table is unknown because $13.41 \cdot 10^{410}$ different messages should be decoded. Coding table is created by random code generation and changed each time for new message.

The proposed subcodstanography scheme is evaluated in Section 4. It is shown that inserting information into BMP image does not affect the visual perception of the original image.

It is important to note that the information sender does not know the applied encryption and steganography method used in the hiding information process. Applied encryption and steganography method is responsibility of computer program developed according to the method presented in this paper. Although, a substantial increase in security is achieved, it is very important to secure the program from unauthorized access.

Competing Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

References

[1] R. Oppliger, *Contemporary Cryptography*, Artech House, Norwod, Mass, USA, 2005.

[2] D. Kahn, *The Codebreakers*, The Macmillan Company, New York, NY, USA, 1973.

[3] R. Radhakrishnan, M. Kharrazi, and N. Memon, "Data masking: a new approach for steganography?" *Journal of VLSI Signal Processing*, vol. 41, no. 3, pp. 293–303, 2005.

[4] X. Wu and N. Memon, "Context-based, adaptive, lossless image coding," *IEEE Transactions on Communications*, vol. 45, no. 4, pp. 437–444, 1997.

[5] M. Kharrazi, H. T. Sencar, and N. Memon, "Benchmarking steganographic and steganalysis techniques," in *Security, Steganography, and Watermarking of Multimedia Contents VII*, vol. 5681 of *Proceedings of SPIE*, pp. 252–263, San Jose, Calif, USA, January 2005.

[6] S. Jindal and N. Kaur, "Digital image steganography survey and analysis of current methods," *International Journal of Computer Science and Information Technology & Security*, vol. 6, 2016.

[7] N. Provos and P. Honeyman, "Hide and seek: an introduction to steganography," *IEEE Security and Privacy*, vol. 1, no. 3, pp. 32–44, 2003.

[8] N. F. Johnson and S. Jajodia, "Exploring steganography: seeing the unseen," *Computer*, vol. 31, no. 2, pp. 26–34, 1998.

[9] J. Fridrich, M. Goljan, and D. Hogeg, "Steganalysis of JPEG images: breaking the F5 algorithm?" in *Proceedings of the Information Hiding: 5th International Workshop (IH '02)*, Lecture Notes in Computer Science, Springer, Noordwijkerhout, The Netherlands, 2002.

[10] V. M. Potdar, S. Han, and E. Chang, "Fingerprinted secret sharing steganography for robustness against image cropping attacks," in *Proceedings of the 3rd IEEE International Conference on Industrial Informatics (INDIN '05)*, pp. 717–724, IEEE, Perth, Australia, August 2005.

[11] Z. Li, X. Chen, X. Pan, and X. Zeng, "Lossless data hiding scheme based on adjacent pixel difference," in *Proceedings of the International Conference on Computer Engineering and Technology (ICCET '09)*, pp. 588–592, January 2009.

[12] M. H. Shirali-Shahreza and M. Shirali-Shahreza, "A new approach to Persian/Arabic text steganography," in *Proceedings of the 5th IEEE/ACIS International Conference on Computer and Information Science (ICIS '06)*, pp. 310–315, July 2006.

[13] M. Kharrazi, H. T. Sencar, and N. Memon, "Performance study of common image steganography and steganalysis techniques," *Journal of Electronic Imaging*, vol. 15, no. 4, Article ID 041104, 2006.

[14] R. Tzschoppe, R. Bäuml, J. B. Huber, and A. Kaup, "Steganographic System based on higher-order statistics," in *Security and Watermarking of Multimedia Contents V*, vol. 5020 of *Proceedings of SPIE*, pp. 156–166, Santa Clara, Calif, USA, January 2003.

[15] C. C. Chang, P. Tsai, and M. H. Lin, "An adaptive steganography for index-based images using codeword grouping?" in *Advances in Multimedia Information Processing—PCM 2004*, vol. 3333 of *Lecture Notes in Computer Science*, pp. 731–738, Springer, Berlin, Germany, 2004.

[16] J. Kong, H. Jia, X. Li, and Z. Qi, "A novel content-based information hiding scheme," in *Proceedings of the International Conference on Computer Engineering and Technology (ICCET '09)*, pp. 436–440, Singapore, January 2009.

[17] I. Cox, *Information Hiding, Watermarking and Steganography*, Public Seminar, Intelligent Systems Research Centre, University of Ulsterat Magee, Derry, Northern Ireland, 2009.

[18] S. C. Katzenbeisser, "Principles of steganography," in *Information Hiding Techniques for Steganography and Digital Watermarking*, S. Katzenbeisser and F. A. P. Petitcolas, Eds., Artech House, Inc, Norwood, Mass, USA, 2000.

[19] P. P. Aung and T. M. Naing, "A novel secure combination technique of steganography and cryptography," *International Journal of Information Technology, Modeling and Computing*, vol. 2, no. 1, pp. 55–62, 2014.

[20] A. Cheddad, J. Condell, K. Curran, and P. Mc Kevitt, "Digital image steganography: survey and analysis of current methods," *Signal Processing*, vol. 90, no. 3, pp. 727–752, 2010.

[21] M. Köhler, I. Lukić, and N. Slavek, "Impact of inserting a stochastic noise on the visual information in a bitmap," in *Proceedings of the 34th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO '11)*, Opatija, Croatia, May 2011.

[22] I. Avcibas, M. Kharrazi, N. Memon, and B. Sankur, "Image steganalysis with binary similarity measures?" *EURASIP Journal on Applied Signal Processing*, vol. 17, 2005.

[23] J. J. Harmsen and W. A. Pearlman, "Steganalysis of additive-noise modelable information hiding," in *Security and Watermarking of Multimedia Contents V*, vol. 5020 of *Proceedings of SPIE*, Santa Clara, Calif, USA, January 2003.

[24] G. Cancelli, G. Doërr, M. Barni, and I. J. Cox, "A comparative study of ± steganalyzers," in *Proceedings of the IEEE 10th Workshop on Multimedia Signal Processing (MMSP '08)*, pp. 791–796, Queensland, Australia, October 2008.

[25] The CRYSTAL Project, http://www1.inf.tu-dresden.de/~aw4/crystal/.

[26] S. Saraireh, "A secure data communication system using cryptography and steganography," *International Journal of Computer Networks & Communications*, vol. 5, no. 3, pp. 125–137, 2013.

[27] K. Challita and H. Farhat, "Combining steganography and cryptography: new directions?" *International Journal on New Computer Architectures and Their Applications*, vol. 1, no. 1, pp. 199–208, 2012.

[28] R. Nivedhitha and T. Meyyappan, "Image security using steganography and cryptographic techniques," *International Journal of Engineering Trends and Technology*, vol. 3, no. 3, 2012.

[29] D.-C. Lou and C.-H. Sung, "A steganographic scheme for secure communications based on the chaos and Euler theorem," *IEEE Transactions on Multimedia*, vol. 6, no. 3, pp. 501–509, 2004.

[30] A. Cheddad, J. Condell, K. Curran, and P. M. Kevitt, "Securing information content using new encryption method and steganography," in *Proceedings of the 3rd International Conference on Digital Information Management (ICDIM '08)*, pp. 563–568, London, UK, November 2008.

[31] K. Curran, X. Li, and R. Clarke, "An investigation in to the use of the least significant bit substitution technique in digital watermarking?" *American Journal Applied Sciences*, vol. 2, no. 3, 2005.

[32] Y.-S. Chen and R.-Z. Wang, "Steganalysis of reversible contrast mapping watermarking," *IEEE Signal Processing Letters*, vol. 16, no. 2, pp. 125–128, 2009.