# Methods of information systems protection

UDC 003.26 **doi: 10.20998/2522-9052.2020.2.20**

V. Pevnev, Yu. Voikov

National Aerospace University "Kharkiv Aviation Institute", Kharkiv, Ukraine

## RESEARCH AND PROTOTYPING METHODS OF STEGANOGRAPHY USING MOSAIC

**Abstract.** The paper considers the possibility of using a mosaic composed of many miniature images to hide the fact of text information transmission. The results of the study of optimal mosaic construction when using steganographic methods of hiding information are presented. The use of the proposed algorithms implies the use of cryptographically secure pseudorandom number generator with dynamically changing parameters for concealing information. The generators are used to determine the location of information bits in a certain mini-images, as well as in a specific pixels and color channels. Using the proposed algorithm, no more than twenty-five percent of the color channels are changed, provided that one bit per color is used. This paper is supplemented with illustrations and examples of how the algorithm works.

**Keywords:** mosaic; steganography; pseudo-random sequence generators; pixel; color channel; information bit.

### Introduction

Every day IT increases influence in all of the sides of our life. Even in Art it becomes more and more popular and helps to create new masterpieces. One of the examples is a painting but how it can help to create a picture? There is a new trend that became popular last years – a photographic mosaic. It is a compilation of small photos grouped together in a way, that they look like a certain big picture. You can see the example of a photographic mosaic bellow on the Fig. 1.

**Fig. 1.** A photographic mosaic made from different small pictures

To implement this concept, a certain set of images of the same size is used and a mosaic of existing images is formed on the basis of the original image. There are many different principles and parameters that provide effective comparison, and choose the most appropriate variant from all possible by comparing it with the original [1].

It is necessary to understand that it is impossible to achieve a hundred percent match, though it is not required. One of the main ideas of the mosaic is that the viewer is aware that any large composition consists of many small parts and each of them is a complete work.

It should also be noted that the issue of user data security and confidentiality is more pressing than ever. In recent years, many ways of intercepting transmitted information have been created. This is due to the fact that the information is in the least secure state at the very moment of transmission from one hand to another, which greatly simplifies the work for malicious users.

There are various ways of combating unauthorized access to information, but steganography (the implicit transfer of information) is particularly notable. It allows you to hide the fact of data transfer, so that other users will not even know that the information was successfully transmitted. However, in order to implement this approach, it is necessary that the sender and the receiver have previously accepted arrangements. For example, when hiding information in an image, it is necessary that all participants have the original image or know how the used algorithm of steganorgaphy works.

The obvious problem with this method is the possibility of information leakage about the method of hiding data, which will instantly compromise all the messages transmitted in this way. Steganography is also easily recognized by simple comparison algorithms, which can negate the whole point of hiding information in another object. In such a case, you can additionally secure the transmitted information using encryption algorithms, or create an algorithm that allows you to use the transmitted object not only as a container for information, but also as an encryption tool. Thus, even if the fact of transfer is known, the information cannot be read by a third party.

This paper considers various methods for comparing a set of pixels, analysis and algorithm for hiding information in a mosaic, as well as a description of the requirements for a pseudo-random sequence generator.

### Image analysis

All of the raster type digital images and displays consist of small dots. Typically, they have particular structure – RGB [1]. It stands for Red, Green and Blue. A set of RGB colors forms one point that is called a "pixel" and image is made of rows and columns of pixels. Then pixels are placed a matrix with a size n by m (Fig. 2).

| | | |
|---|---|---|
| 1,1 | … | 1,n |
| … | | |
| m,1 | … | m,n |

**Fig. 2.** Pixel structure of a RAW(uncompressed) digital raster image where n is column number and m is row number

The combination of these colors can give us majority of existing color that can be recognized by human's eye. The only difference is the number of bits that are used to store the value of each color. The bigger amount of values leads to a wider color range we get. It is also called a "color depth", which is easily shown on smooth gradient transitions. Commonly, there are 8 bits for each color so it gives us the following result:

$$8b \cdot 3 = 24b = 16777216 \quad (1)$$

Based on this structure there are many different types of image analysis but here would be described only easy and effective ones. It is important due to service requirements. It should be efficient and fast enough so user won't have to wait a lot of time until the output image is done. Using mosaic in stenography gives us big advantage of not sending the original image to the receiver. Since the database contains all of the images used in the mosaic it can be easily compared with originals. The original image is divided into patches, which can hold from 1 pixel to several hundreds or even thousands, in the end each patch will be replaced by the most suitable image from the database. Choosing how many pixels a patch will hold directly affects the quality and size of the mosaic (Fig. 3).

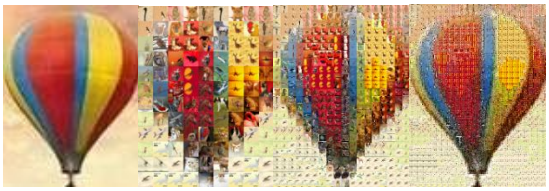

**Fig. 3.** Original image and mosaics with different patch size

Next step is to choose the most appropriate images for the database in order to replace created patches. The most accurate is a pixel-by-pixel comparison, but it takes too much time, moreover, it is almost impossible to find the perfect picture for the required patch, so it is necessary to calculate the overall picture characteristic.

All images from the database are processed and the required metadata for each of them is calculated, allowing comparison by properties that we need. Then, the same metadata is calculated for each patch of the original image. For each patch a certain number of best images is selected, after which the patch and these images are further split into 4 parts, each of which as a result receives its own metadata and another comparison is made. The results select the most suitable image from the available ones. There are several criteria for comparing the similarity of images: Saturation; Contrast; Average values of color channels; Boundaries of color transitions (Edges).

Perhaps the main way to compare pixels similarity is measuring mean values of the color channels and calculating the difference, the rest help to improve the accuracy of the final selection. For example, partial coincidence of borders of transitions of colors allows to achieve selection not only on color, but also on various lines, figures and the original image (Fig. 4).

**Fig. 4.** Grey scale image and applied edge detection filter image [1]

One of the main ideas of using stenography in combination with image mosaics is that there is no need to convey the original image. All original images are stored in the application database.

As a source of images, you can use any set of small pictures of the same shape. The easiest to use are square images with the side equal to $2^n$ pixels.

This size allows you to quickly make calculations on changing the size of the image without the need for additional cropping of the photo.

## Steganography

Steganographic system is an association of methods and algorithms used for creation of the hidden channel for information transfer [2]. In other words, steganography refers to the process of information transfer, which hides the very fact of transfer.

The majority of experts in the field of steganography, describe three groups of methods with different approaches to the implementation of the implicit information transfer. These are the following methods [3-7]: usage of special properties of computer data formats; usage of less important audio and visual information; linguistic steganography.

Methods based on the use of special properties of computer formats are divided into groups [8]: computer data formats reserved for field expansion; special formatting text files; hiding in unused places of CD-disk; removing the identifying header file. All researchers pay great attention to the methods of using redundancy of audio and visual information. At the same time, audio steganography methods include [9]: LSB (Least Significant Bit); even coding; phase coding; spread spectrum method. In a general case, the following scheme represents steganographic system work (Fig. 5).

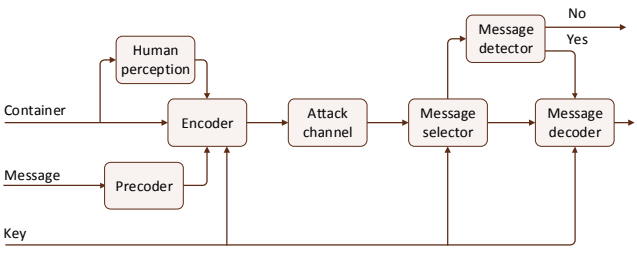


**Fig. 5.** Steganographic system work scheme

It contains the following elements: ***Human perception*** – actions or algorithms that source data which has minor influence on human perception (ability to see or hear) in order to make the message existence almost invisible; ***Precoder*** – changes the message to a suitable for encoder form; ***Encoder*** – encodes the message into container; ***Attack channel*** – data transfer channel; ***Message selector*** – selects the data where the message could be stored; ***Message detector*** – detects if the message it is hidden inside data; ***Message decoder*** – decodes the message from data.

Depending on container we need to know which data we can modify so a user will not recognize any changes. Then we can modify message if encoder requires different format of data. After it, encoder inserts message into the container in a way that it became invisible for a user.

Data that contains hidden message can be damaged by channel noise or other reasons so it might be necessary to use noise-resistant coding algorithm.

When encoding is completed the container with hidden message goes to untrusted channel. We assume that anyone can see our container there. After receiver gets the container, message selector retrieves the data where the message could be stored. Then message detector analyses if this data has a hidden message and id so decoder convert it to original message.

Like most objects that were created from real-world information, photographs are not of ideal quality, there are always all kinds of noises and artifacts, so they are widely used in steganography for the implicit transmission of information. However, in order to record information in an image, it is necessary to change it, which, if implemented incorrectly, can be visible even to an average user. In order to avoid this situation, it is necessary to understand what effect on the final image will have a change of a certain bit in the color channel (Fig. 6).

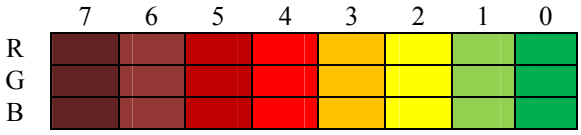


**Fig. 6.** Impact of changes in bits on the final image

Therefore, changing a lower bit has almost no effect on the final result, and vice versa, changing a higher bit can have a significant effect on the final image. On this basis, it can be concluded that the safest bit is the lowest bit, but if more information needs to be packed, the next highest bit can be used.

## Encoding

One of the key features of the steganography algorithm is how and in what sequence it places information in the container. Thus, when information is written sequentially to the lower bits of pixels in a row, the information can be easily read as soon as the user realizes that the picture is different from the original.

For reliable data protection, a distributed recording of information across all pixels with a normal distribution of random values is the best option (Fig. 7). This approach helps to minimize the risks of encrypted information disclosure. You can also use the following method to ensure better information privacy [10].

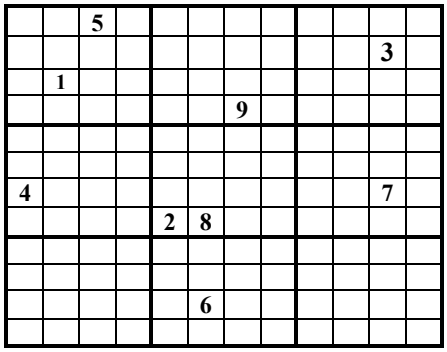


**Fig. 7.** Example of a distributed recording with normal distribution probability

When converting a message to a bit sequence, the number of zeros and units is compared. If the number of units is greater than the number of zeros, the message is inverted (Fig. 8). Then a pseudorandom bit sequence is generated. The bits of the original container data sequence are selected based on the indexes of the PRBS elements whose value is equal to zero. The message bit values are written to the container bits using the XOR logical operation as shown below (Fig. 9).

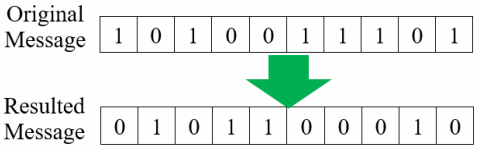


**Fig. 8.** Example of message changing

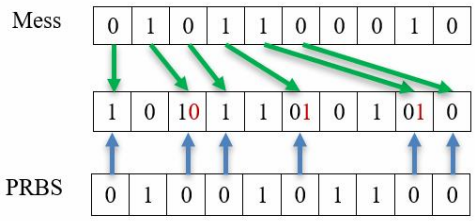


**Fig. 9.** Encoding algorithm example

This way, in the worst-case scenario, the number of modified bits will not exceed twenty-five percent of the bits used for encoding.

However, a random sequence of sufficient length and stability must be generated to allocate the bits.

## Pseudorandom number generator

Generators of true random numbers are usually based on natural processes, such as pickups in the transmission channel, so that they can produce a crypto-resistant sequence of any length, but they cannot reproduce the same sequence a second time.

The most common types of PRBS generators are shift register-based and congruent generators. The congruent generators are based on numerical sequences, in which each member depends on one or more previous ones. Generators based on shift registers are based on the generating polynomial. These generators are not

crypto-resistant enough, because they are easy enough to hack, but have a big advantage in simplicity and speed of operation.

Successful use of a generator directly depends on its crypto resistance, so it is critical for a reliable generator. Improved generators are perfectly suited for this task. The main features of these generators are the use of dynamically changing parameters. Depending on the type of generator it can be parameters a and c for the congruent generator(2) and generating polynomials for the generator based on the shift registers:

$$X_{n-1} = (aX_n + c) \bmod m . \qquad (2)$$

In addition to using dynamic parameters to improve the cryptographic stability of the output sequence, there are a modification of the algorithm of generating PRBS, which consists in using the operation XOR over the generated sequences of a certain length, obtained from two registers with different digit capacity. The used registers change their parameters at least every $2N$ - 1 bit, where N is the product of the two registers' digit capacity.

## Conclusion

In the paper we investigate the possibility of using steganography in combination with a mosaic of miniature images for covert information transfer. The crucial parameters of the images that most affect the selection of the most suitable image for a certain area of the original drawing have been described. It was pointed out that the main criterion is the average value of the color channel. The algorithm of division of the original image into patches and the system of rounds of selection of the most suitable pictures to replace the selected area were also described, and examples were given.

The described approach allows to achieve a reliable result and also has flexibility that allows to easily add various criteria for evaluation of the image, as well as choose in which round it will be applied.

In such a way, it is possible to perform resource-intensive operations on the most promising variants of pictures only. The basic principles of steganography for understanding the significance of the next steps were presented, and the variant of hiding information in a mosaic was suggested and described. This method allows to achieve no more than 25 percent of lower bit changes, which is about 3.125 percent of all information, provided that the lower bits are fully filled with information. In the real situation, when encoding a message is used a significantly smaller number of bits, which in turn dramatically reduces the percentage of changes in the mosaic data, and visual detection of the transfer is almost impossible.

Also, the principle of uniform pseudo-random distribution of information bits throughout the mosaic makes it almost impossible to use brute-force attack to evaluate the hidden message. The concept of using the PRBS generator was described, as well as security problems that arise when using simple implementations of generators that do not have sufficient crypto resistance. To generate and restore a pseudo-random sequence, it was proposed to use an improved generator with dynamic parameters [2]. It has significant resistance to known crypto attacks, which allows its use in modern cryptosystems. As a result, the proposed system allows you to secretly transfer information inside the mosaic, and has sufficient protection to protect information from unauthorized reading in case of detection of the transfer of information.

The presented steganographic algorithm can be easily adapted to work with usual raster images, sound files and other types of data which could be presented as an array of numerical values.

### References


1. Inad, Aljarrah, Abdullah, Al-Amareen, Abdelrahman, Idries and Osama, Al-Khaleel (2014), “Image Mosaicing Using Binary Edge Detection”, *Proc. of the Int. conf. on Comp. Technology and Information Management*, Dubai, UAE, 2014, pp. 187-190, available to: https://www.researchgate.net/figure/Gray-scale-image_fig1_261551383
2. Pevnev, V. and Frolov, V. (2018), “Results of studies of generators of pseudorandom sequences with dynamic parameters”, *Control, navigation and communication systems*, Vol. 4 (50), pp. 139-143, DOI: https://doi.org/10.26906/SUNZ.2018.4.139.
3. Neil F., Johnson, Zoran, Duric and Sushil, Jajodi (2012), [Information Hiding: Steganography and Watermarking-Attacks and Countermeasures, Springer Science & Business Media, 137 p.
4. Abdelrahman Desoky, *Noiseless Steganography: The Key to Covert Communications*, CRC Press, 2016 – 300 p.
5. Jessica, Fridrich (2010), *Steganography in Digital Media: Principles, Algorithms, and Applications*, Cambridge University Press, 437 p.
6. Frank Y., Shih (2017), *Digital Watermarking and Steganography: Fundamentals and Techniques*, CRC Press, 200 p.
7. Frank Y., Shih (2017), *Multimedia Security: Watermarking, Steganography, and Forensics*, CRC Press, 423 p.
8. Kazuhiro, Kondo (2012), *Multimedia Inf. Hiding Technologies and Methodologies for Controlling Data*, IGI Global, 497 p.
9. Barney, Warf (2018), *The SAGE Encyclopedia of the Internet*, SAGE, 1120 p.
10. Pevnev, V., Novakov, Y., Tsuranov, M. and Kharchenko V. (2017), “The Method of Data Integrity Assurance for Increasing IoT Infrastructure Security”, *Proc. of the 31 th Int. Conference on Information Technologies* (InfoTech-2017), pp. 27-36.


Received (Надійшла) 10.03.2020
Accepted for publication (Прийнята до друку) 13.05.2020

### Відомості про авторів / About the Authors

**Пєвнєв Володимир Яковлевич** – кандидат технічних наук, доцент, доцент кафедри комп’ютерних систем, мереж та кібербезпеки, Національний аерокосмічний університет імені М.Є. Жуковського «ХАІ», Харків, Україна;
**Volodymyr Pevnev**– Candidate of Technical Science, Associate Professor, Associate Professor of Computer Systems, Networks and Cyber security Department, National Aviation University “Kharkiv Aviation Institute”, Kharkiv, Ukraine;
e-mail: v.pevnev@csn.khai.edu; ORCID ID: http://orcid.org/0000-0002-3949-3514.

**Войков Юрій Володимирович** – магістрант кафедри комп'ютерних систем, мереж і кібербезпеки Харківського національного аерокосмічного університету ім. М.Є. Жуковського, Харків, Україна;
**Yurii Voikov** – Master's Degree in Computer Systems, Networks and Cyber Security, Kharkiv National Aerospace University. M.E. Zhukovsky, Kharkiv, Ukraine;
e-mail: y.voykov@student.csn.khai.edu; ORCID ID: http://orcid.org/0000-0002-0926-4833

**Дослідження та прототипирування методів стеганографії з використанням мозаїки**

В. Я. Пєвнєв, Ю. В.Войков

**Анотація**. Предметом вивчення в статті є можливість використання мозаїки в сукупності зі стеганографічними алгоритмами для прихованої передачі повідомлення. З кожним днем IT збільшує вплив на всі сторони нашого життя. Навіть в мистецтві він стає все більш популярним і допомагає створювати нові шедеври. З'явився новий напрямок, що стало популярним в останні роки - фотографічна мозаїка. Це збірка маленьких фотографій, згрупованих так, щоб вони виглядали як якась велика картина. **Мотивація**. В останні роки було створено безліч способів перехоплення інформації, що передається. Це пов'язано з тим, що інформація в момент передачі з одних рук в інші знаходиться в найменш безпечному стані, що значно спрощує роботу для зловмисників. Стеганографія дозволяє приховати факт передачі даних, так що інші користувачі навіть не знатимуть, що інформація була успішно передана. Однак для реалізації такого підходу необхідно, щоб відправник і одержувач попередньо мали домовленості, такі як наявність у обох учасників однакової оригінальної картинки. Очевидною проблемою цього методу є можливість витоку інформації про спосіб приховування даних. **Метою** цього дослідження є отримання алгоритму приховування повідомлення всередині мозаїки, що забезпечує мінімальну видимість, надійний захист даних навіть у тому випадку, якщо зображення мозаїки і алгоритм відомі зловмиснику і в першу чергу не потребує відправки оригінального зображення. **Результати**. Описаний підхід дозволяє добитися якісного результату, а також має значну гнучкість, що дозволяє легко додавати різні критерії оцінки зображення, а також вибирати в якому раунді воно буде застосовуватися. Таким чином можна виконувати ресурсомісткі операції тільки над найбільш перспективними варіантами картинок. Запропоновані алгоритми мають на увазі використання криптографічно-безпечного генератора псевдовипадкових чисел з динамічно змінними параметрами для приховування інформації. Генератори псевдо-випадкової послідовності використовуються для визначення місця розташування інформаційних бітів в певних міні-зображеннях, а також в окремих пікселях і колірних каналах. При використанні запропонованого алгоритму змінюється не більше двадцяти п'яти відсотків колірних каналів за умови, що використовується один біт на колір. **Висновки**. Запропонована система дозволяє приховано передавати інформацію в мозаїці, а також володіє достатньою захищеністю для захисту інформації від несанкціонованого читання в випадки виявлення факту передачі інформації. У поєднанні з даним підходом можна використовувати додатково будь-який алгоритм шифрування для забезпечення ще більшої захищеності даних. Представлений алгоритм кодування і приховування даних можна легко адаптувати для роботи зі звичайними растровими зображеннями, звуковими файлами та іншими типами даних, які представлені у вигляді масиву значень.

**Ключові слова:** мозаїка; стеганографія; генератор псевдо-випадкової послідовності; піксель; канал кольору; інформаційний біт.

**Исследование и прототипирование методов стеганографии с использованием мозаики**

В. Я. Певнев, Ю. В. Войков

**Аннотация.** Предметом изучения в статье является возможность использования мозаики в совокупности со стеганографическими алгоритмами для скрытной передачи сообщения. С каждым днем IT увеличивает влияние на все стороны нашей жизни. Даже в искусстве он становится все более популярным и помогает создавать новые шедевры. Появилось новое направление, ставшее популярным в последние годы - фотографическая мозаика. Это сборник маленьких фотографий, сгруппированных так, чтобы они выглядели как некая большая картина. **Мотивация**. В последние годы было создано множество способов перехвата передаваемой информации. Это связано с тем, что информация в момент передачи из одной рук в другие находится в наименее безопасном состоянии, что значительно упрощает работу для злоумышленников. Стеганография позволяет скрыть факт передачи данных, так что другие пользователи даже не будут знать, что информация была успешно передана. Однако для реализации такого подхода необходимо, чтобы отправитель и получатель предварительно имели договорённости, такие как наличие у обоих участников одинаковой оригинальной картинки. Очевидной проблемой с этим методом является возможность утечки информации о способе сокрытия данных. **Целью** этого исследования является получение алгоритма скрытия сообщения внутри мозаики, обеспечивающего минимальную видимость, надежную защиту данных даже в том случае, если изображение мозаики и алгоритм известны злоумышленнику и в первую очередь не требующего отправки исходного изображения. **Результаты**. Описанный подход позволяет добиться качественного результата, а также обладает гибкостью, что позволяет легко добавлять различные критерии оценки изображения, а также выбирать в каком раунде оно будет применяться. Таким образом можно выполнять ресурсоёмкие операции только над наиболее перспективными вариантами картинок. Предлагаемые алгоритмы подразумевают использование криптографически-безопасного генератора псевдослучайных чисел с динамически изменяющимися параметрами для сокрытия информации. Генераторы псевдо-случайной последовательности используются для определения местоположения информационных битов в определённых мини-изображениях, а также в отдельных пикселях и цветовых каналах. При использовании предложенного алгоритма изменяется не более двадцати пяти процентов цветовых каналов при условии, что используется один бит на цвет. **Выводы**. Предложенная система позволяет скрытно передавать информацию в мозаике, а также обладает достаточной защищенностью для защиты информации от несанкционированного чтения в случаи обнаружения факта передачи информации. В сочетании с данным подходом можно использовать дополнительно какой-либо алгоритм шифрования для обеспечения ещё большей защищённости данных. Представленный алгоритм кодирования и скрытия данных можно легко адаптировать для работы с обычными растровыми изображениями, звуковыми файлами и другими типами данных, которые представлены в виде массива значений.

**Ключевые слова:** мозаика; стеганография; генератор псевдослучайной последовательности; пиксель; цветовой канал; информационный бит.