DOI: 10.15514/ISPRAS-2023-35(3)-4

Development and Implementation of the Digital Steganography Method Based on the Embedding of Pseudoinformation

I.G. Gvozdeva ORCID: 0009-0009-1058-2618 <*gvozdeva-irina@bk.ru*>
A.S. Gromov ORCID: 0009-0000-7130-6785 <*gromov.bo@yandex.ru*>
O.M. Gvozdeva ORCID: 0009-0000-1947-9638 <*olgagvozdevaaa@yandex.ru*>

Penza State University of Architecture and Construction,
28, Herman Titov st., Penza, 440028, Russia.

Abstract. The article provides an overview of the main methods of steganography, on the basis of which a new method was developed, consisting in embedding additional text (pseudo-information) in parallel with the transmitted message. An algorithm of this method has been developed. In this case, the frequency of the bit sequence was obtained in accordance with the generated pseudo-random numbers. In accordance with the algorithm, an application has been developed that allows the sender to encrypt and place the message in a container that is an image, and the recipient to determine the presence of the message and, if there is one, extract it. A computational experiment was also conducted, which showed that an image with a fairly large embedded text does not visually differ from the original image.

Keywords: steganography; cryptography; stegosystem; steganalysis.

For citation: Gvozdeva I.G., Gromov A.S., Gvozdeva O.M. Development and implementation of the digital steganography method based on the embedding of pseudoinformation. *Trudy ISP RAN/Proc. ISP RAS*, vol. 35, issue 3, 2023. pp. 63-70. DOI: 10.15514/ISPRAS-2023-35(3)-4

Разработка и реализация метода цифровой стеганографии на основе встраивания псевдоинформации

И.Г. Гвоздева, ORCID: 0009-0009-1058-2618 <*gvozdeva-irina@bk.ru*>
А.С. Громов, ORCID: 0009-0000-7130-6785 <*gromov.bo@yandex.ru*>
О.М. Гвоздева, ORCID: 0009-0000-1947-9638 <*olgagvozdevaaa@yandex.ru*>

Пензенский государственный университет архитектуры и строительства,
440028, Россия, г. Пенза, ул. Германа. Титова, д. 28.

Аннотация. В статье приведен обзор основных методов стеганографии, на основании которого был разработан новый метод, заключающийся в встраивании дополнительного текста (псевдоинформации) параллельно с передаваемым сообщением. Разработан алгоритм этого метода. При этом частоту последовательности битов получали в соответствии с сгенерированными псевдослучайными числами. В соответствии с алгоритмом разработано приложение, позволяющее отправителю зашифровать и поместить сообщение в контейнер, представляющий собой изображение, а получателю определить наличие сообщения и, если оно имеется, извлечь его. Также был проведен вычислительный эксперимент, который показал, что изображение с довольно большим встроенным текстом визуально не отличается от исходного изображения.

Ключевые слова: стеганография; криптография; стегосистемы; стегоанализ.

Для цитирования: Гвоздева И.Г., Громов А.С., Гвоздева О.М. Разработка и реализация метода цифровой стеганографии на основе встраивания псевдоинформации. Труды ИСП РАН, том 35, вып. 3, 2023 г., стр. 63–70 (на английском языке). DOI: 10.15514/ISPRAS–2023–35(3)–4

1. Introduction

The problem of delivering a confidential message has stood at all times. This problem has given rise to such sciences as cryptography and steganography.

The essence of steganography is that the message is placed in a container in such a way that an uninitiated circle of people sees only the object, not realizing that it may be filled with something. Here are some well-known examples: in ancient Greece, wooden writing boards covered with wax, under which there was a message, the heads of slaves with the message printed hidden under the hair, later, the so-called sympathetic ink, invisible under normal conditions, was widely used [1].

In the modern world, with the development of computer technology, digital data, as a rule, files of multimedia objects (images, video, audio, textures of 3D objects) serve as containers for hiding information. This is due to the fact that digitized objects, which initially have an analog nature, always have quantization noise, and when reproducing these objects, additional analog noise appears [2]. All this contributes to greater invisibility of hidden information.

The advantage of steganography over cryptography is that not only the contents of the transmitted message are hidden, but the very fact of the existence of this message is hidden.

In the science of steganography, such a direction as steganalysis is distinguished. The task of steganalysis is to identify the fact of transmission of hidden information in the analyzed message [3, 4]. Consequently, with the development of steganalysis, new methods are required to make hidden information inaccessible to the uninitiated [5]. In this article, the authors propose a way to embed information in a container that increases the reliability of its protection against unauthorized access.

2. Relevance

Digital steganography as a science was born literally in recent years. It includes the following areas:

- embedding information for the purpose of its hidden transmission;
- embedding of digital watermarks (CVZ) (watermarking);
- embedding identification numbers (fingerprinting);
- embedding titles (captioning).

This work touches on the first direction. Many methods and algorithms of steganography are known today. Here are some of them.

- LSB-steganography (the message is hidden in the lower bits (it is possible to use one or more lower bits) of the container [6, 7].
- The method based on hiding data in the coefficients of the discrete cosine transformation (hereinafter DCP) is a variation of the previous method, which is actively used, for example, when embedding a message in a JPEG format container.
- The method of hiding information using the lower bits of the palette – this method is essentially a variation of the general LSB method, but the information is embedded not in the least significant bits of the container, but in the least significant bits of the palette. As a result, the container capacity is low.
- The method of hiding information in the service fields of the format is a method in which the embedded message is placed in the service fields of the container header. Obvious disadvantages are the low capacity of the container and the ability to detect embedded data using conventional image viewing programs (which sometimes allow you to see the contents of service fields).

As early as 1883, Kergoff wrote that the information security system should provide its functions even with full awareness of the enemy about its structure and algorithms of functioning.

This means that the message embedding model should be sufficiently complex so that the enemy, if he guesses about the presence of a hidden message, even with the presence of powerful computing equipment, would not be able to extract it [8].

In connection with the above, the authors propose a method of concealing information, which aims to increase confidence that the message intercepted by the enemy will not be opened.

This method is based on the fact that false, so-called pseudo-information is embedded in parallel with useful information. When selecting a model for extracting text from a container file, the opponent relies on the result obtained, which represents any characters. And it is not possible to distinguish the symbols belonging to useful information from false information.

3. Algorithm for embedding information in a container and extracting it

Our proposed method is based on the LSB method, the message will be hidden in the lower bits of the image. A broadband method was used to select the sequence of bits. Such transmission methods are used in communication technology to ensure high noise immunity and complicate the interception process. The purpose of broadband methods is similar to the tasks that a stegosystem solves: to try to “dissolve” a secret message in a container and make it impossible to detect it. Since signals distributed over the entire spectrum band are difficult to detect, steganographic methods based on broadband methods are resistant to accidental and intentional distortion. In this work, the method of jumping frequencies was used, when the frequency of bits intended for embedding information changes according to some pseudo-random law. The frequency of using a byte of color is also randomly selected.

Similarly, frequencies are generated for embedding pseudo-information that do not intersect with the received numbers to accommodate the basic information. The generated frequencies are stored and must be transmitted to the receiving party and are a cryptographic key.

A bmp graphic file with an RGB palette model with a coding depth of 24 bits (8 bits per color) was taken as a container file.

The contents of the container file and the file to be hidden are placed in byte arrays.

For embedding, two random of the four lower bits of one of the three components of the color are used. Since 3 bytes form one color, one byte of text will have 12 bytes of graphics. Before starting the implementation of the algorithm, you should check whether the text file fits into the graphic. The following is the embedding algorithm.

- The last two bits in the specified color component are "released". To do this, the corresponding byte of color is multiplied by a byte mask, with zeros in the specified bits using bitwise multiplication. As a result, these two bits will be reset to zero.
- Take the first two bits from the byte-"text". To do this, multiply the byte "text" by the byte mask equal to 192 (11000000).
- In the resulting byte, we will shift to the right. As a result, the first two bits will be in the specified two places.
- Add the received byte to the edited byte obtained in the first paragraph using bitwise addition. As a result, the first two bits of the text are "hidden".
- Further actions will be repeated.
- After reading the next byte of text, and the actions starting from point 1 are repeated.
- The size of the text is recorded in one of the free bytes of the header part of the graphic file.

Similar actions are performed for embedding pseudo-information. When extracting a message from an image, the reverse action is performed in accordance with the available encryption key.

To test the algorithm's operability, an application was developed that allows the sender to hide the message and the recipient to extract it.

4. Requirements for an application that implements the part of the stegosystem in which information is embedded and extracted

We proceed to the formulation of the requirements for the application. The user should be able to perform the following functions:

- selecting a file with a message to embed;
- selecting a text file for embedding pseudo-information;
- selecting an image file for the shorthand algorithm;
- selecting the name of the resulting image file that will contain encrypted data;
- selecting an image file containing encrypted data;
- selecting the name of the resulting file that will contain the extracted data.

Let's show the user's interaction with the application on the use case diagram (*Fig. 1*).

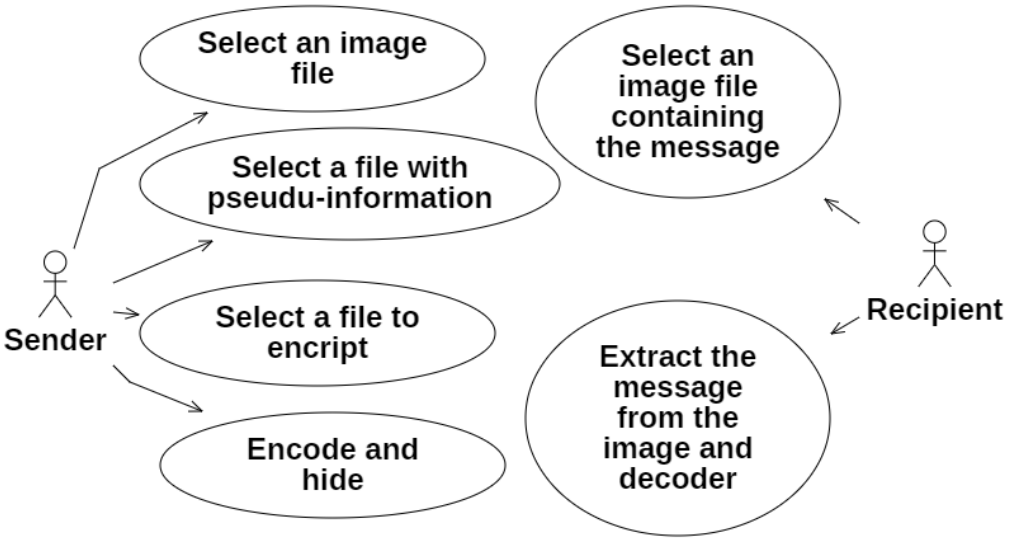

Fig. 1. User interaction with the application

In the implemented system, the text to be embedded is pre-encoded by the byte permutation method, and then by the bit permutation method in accordance with the pseudo-random sequence. Let's describe the logic of the system behavior using the diagram shown in the *Fig. 2*.

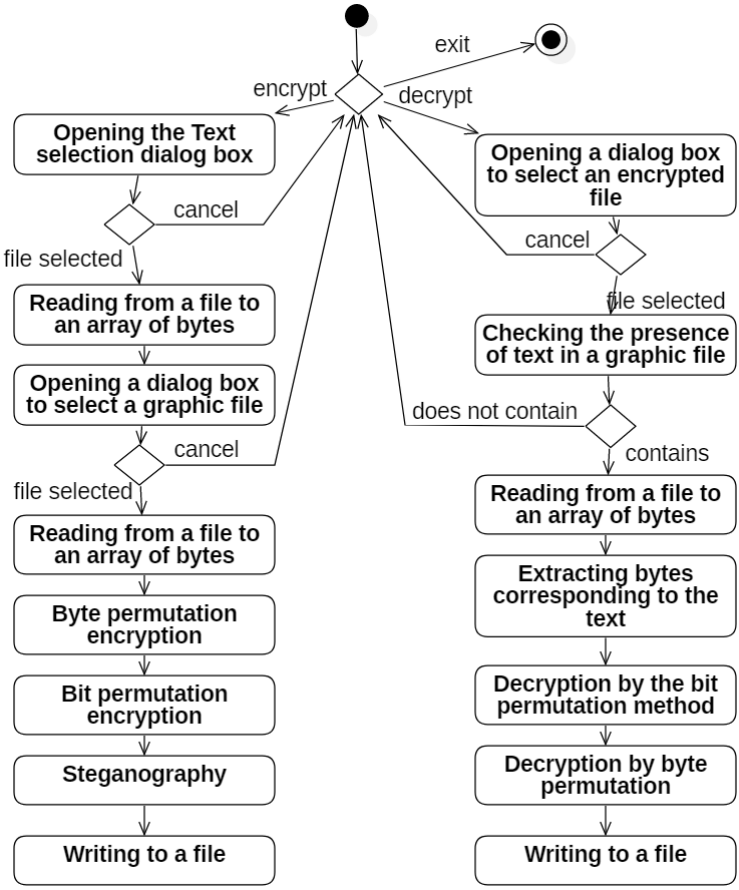

Fig. 2. State diagram

5. Description of the program and test results

In accordance with the designed diagrams, an application was developed.

When the application is launched, the main window appears, providing the user with two functions: encrypt information or decrypt (*Fig. 3*).

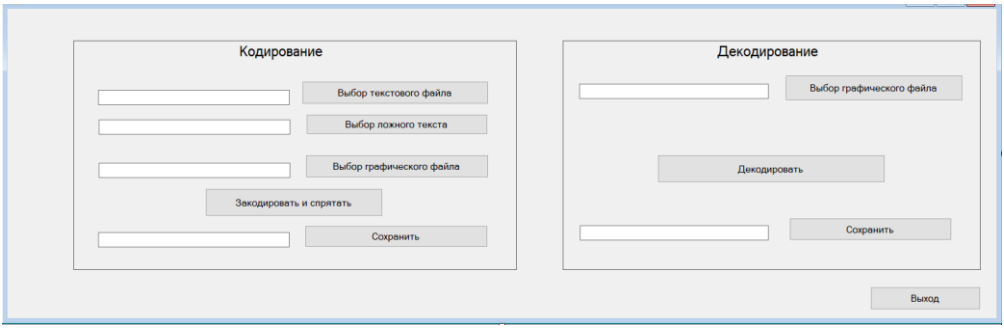

Fig. 3. Main application window

When you click the Select Text File button, a file selection window appears (*Fig. 4*).

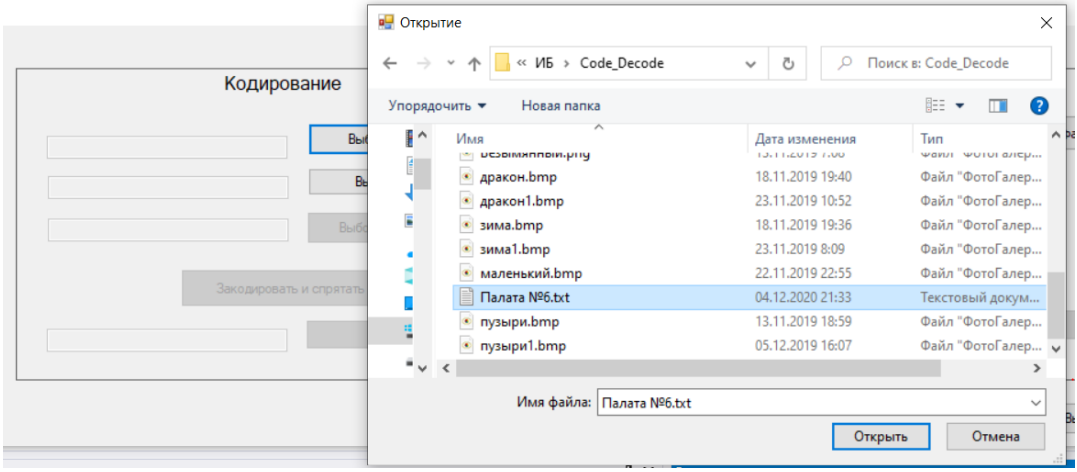

Fig. 4. Selecting a text file

After that, the Select Image File button becomes available. When you click on it, a similar window will open. If the selected image file is too small to contain text information, the file will not open and a message will be displayed about it (*Fig. 5*).

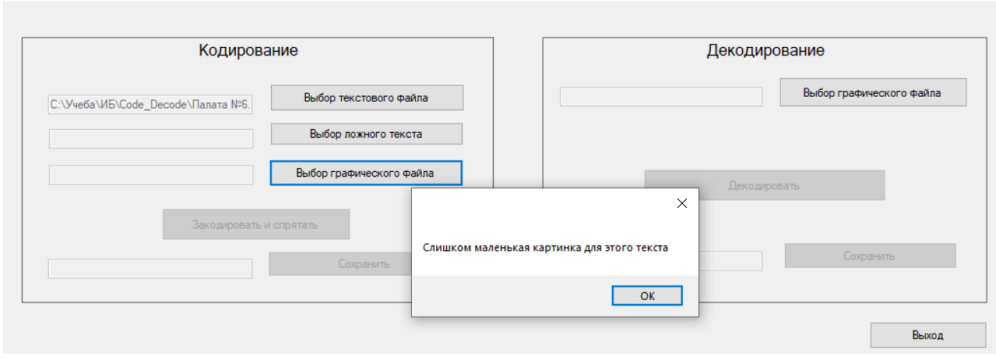

Fig. 5. Insufficient file size

If everything was successful, the Encode and Hide button will be available. When you click on it, the byte and bit permutation methods are applied sequentially, then the steganography method works and the Save button becomes available.

When you select the Select Image File button, it checks whether the file contains hidden text. If not, the file does not open and a message about it is displayed (*Fig. 6*)

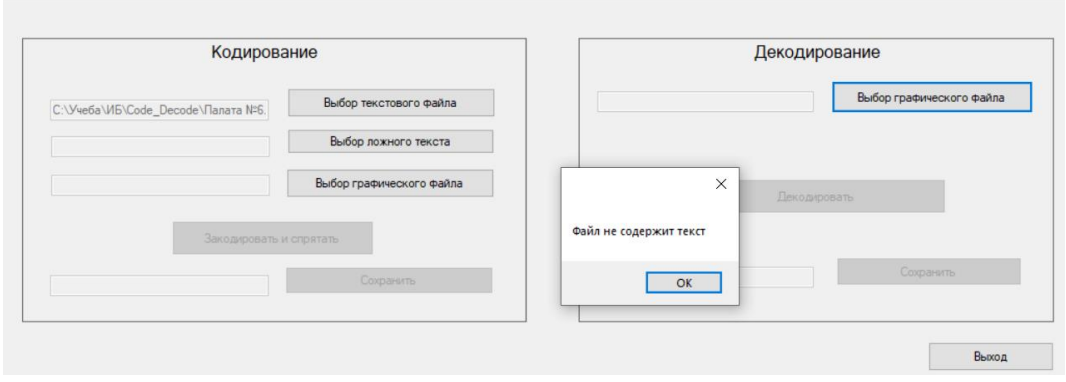

Fig. 6. Opening a file that does not contain text

Further actions are similar to those described above.

Testing was carried out using graphic files "bubbles.bmp" size 2344 KB, "small.bmp" with a size of 1 KB and a text file "Chamber No.6.txt " size 103 KB. The result of the hidden file was recorded in the file "bubbles 1.bmp".

Fig. 7 shows the original graphic file and the file with hidden text.

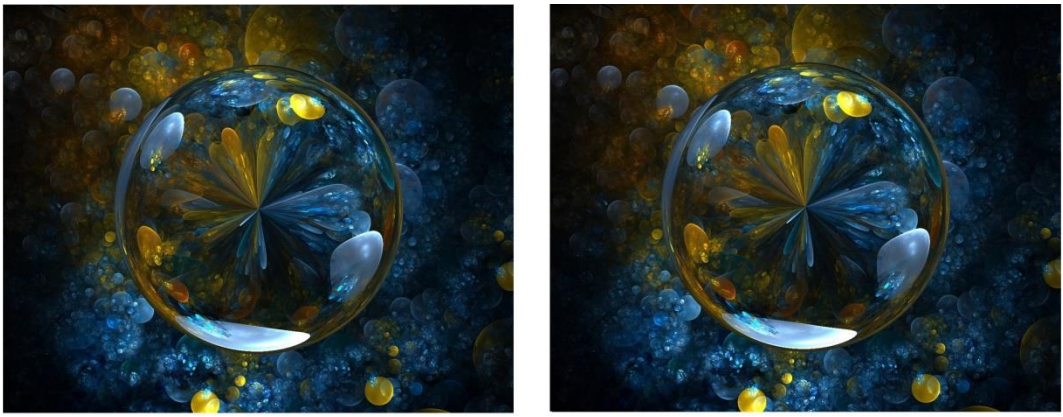

Fig. 7. Image files: with and without embedded text

As can be seen, the replacement of two bits in one color component did not produce any visible changes, which allows using this method along with existing stegosystems.

Conclusion

Thus, as a result of the review of existing methods of steganography, a new method of hiding messages was proposed and implemented, in which the use of embedding pseudo-information is proposed. The results of the development are presented for digital images of the BMP format, however, they can be adapted to other formats.

References

[1]. Gribunin V.G., Okov. I. N. and Turintsev I. V., “Digital steganography” [Text],. Moscow : SOLON-Press, 263 p., 2003./ Грибунин В. Г., Оков И. Н., Туринцев И. В.. Цифровая стеганография [Текст] – Москва: СОЛОН-Пресс, 2003. – 263 с.

[2]. Razinkov E. V., Latypov R. H. “Stability of steganographic systems”, Scientific notes of Kazan.state University, Kazan, Vol. 151, No 2, 2009./ Разинков Е. В., Латыпов Р. Х. Стойкость стеганографических систем // Ученые записки Казанского государственного университета. Сер.: Физико-математические науки. 2009. Т. 151, кн. 2. С.126-132.

[3]. Golubev E. A., Varnovsky N. P. and Logachev O. A., Conference "Mathematics and Security of Information Technologies", Moscow State University, Moscow, Russia, October 2004, 28-29./ Голубев Е. А., Варновский Н. П. и Логачев О. А., Конференция "Математика и безопасность информационных технологий", Московский государственный университет, Москва, Россия, октябрь 2004, 28-29.

[4]. Fridrich J., Du R., Long M. “Steganalysis of LSB encoding in color images”, ICME, 2000.

[5]. Replacement of the least significant bit [Electronic resource]. – Access mode: http://www.nestego.ru/2012/07/lsb.html .

[6]. Provos N., Honeyman P. “Detecting Steganographic Content on the Internet” // Proceeding of the 10 USENIX Security Symposium,. pp. 323–335, 2001.

[7]. Westfeld A. “Attacks on Steganographic Systems: Breaking the Steganographic Utilities EzStego, Jsteg, Steganos and S-Tools-and Some Lessons Learned “/ A. Westfeld, A. Pfitzmann // 3rd International Workshop on Information Hiding (2000)

[8]. Zakalkin P. V., Ivanov S. A., Vershennik E. V. and Kiryanov A. V., “Method of masking transmitted information”, Proceedings of ISP RAS, 32:6 (2020), pp 111–126./

Закалкин П.В., Иванов С.А., Вершенник Е.В., Кирьянов А.В. Способ маскирования передаваемой информации. Труды ИСП РАН, том 32, вып. 6, 2020 г., стр. 111-126.

Информация об авторах / Information about authors

Ирина Геннадьевна ГВОЗДЕВА – кандидат технических наук, специалист кафедры Информационно-вычислительных технологий Пензенского государственного университета архитектуры и строительства. Сфера научных интересов: область математического моделирования и оптимального управления технологическими процессами в строительстве, электрохимии, экологии.

Irina Gennadievna GVOZDEVA – Candidate of Technical Sciences, specialist of the Department of Information and Computing Technologies of the Penza State University of Architecture and Construction. Research interests: mathematical modeling and optimal control of technological processes in construction, electrochemistry, ecology.

Артем Сергеевич ГРОМОВ – студент, обучающийся по специальности «Информационные системы и технологии» в Пензенском государственном университете архитектуры и строительства. Область научных интересов: защита информации.

Artem Sergeyevich GROMOV is a student studying in the specialty "Information Systems and Technologies" at the Penza State University of Architecture and Construction. Research interests: information protection.

Ольга Михайловна ГВОЗДЕВА – студент, обучающийся по специальности «Информационные системы и технологии» в Пензенском государственном университете архитектуры и строительства. Область научных интересов: управление данными.

Olga Mikhailovna GVOZDEVA is a student studying in the specialty "Information Systems and Technologies" at the Penza State University of Architecture and Construction. Research interests: data management.