UDC 004.056.5

I.I. BOBOK, PhD

STEGANALYSIS METHOD FOR DETECTION OF THE HIDDEN COMMUNICATION CHANNEL WITH LOW CAPACITY

Introduction

It is difficult to underestimate the relevance of an effective steganalysis in current conditions. Its main task is the detection of hidden data presence in the information content. The organization of a hidden (steganographic) communication channel for the possible planning, controlling, and committing unlawful actions can lead to significant negative consequences not only for particular people, enterprises, but also for the state and society in general.

One of the most widespread steganographic methods used nowadays for the hidden communication channel organization is the least significant bit method (LSB) [1]. However, there are a large number of known steganalysis methods for its detection [2-4]. The current use of the LSB has some peculiarities: the hidden communication channel of low capacity (less than 20%) has to be organized. The hidden communication channel capacity (HCC) refers to the amount of additional information, which can be embedded into one container item. It is often expressed as a percentage [1]. Taking into account that in this paper, a digital image (DI) is considered as a container and a randomly generated binary sequence is considered as a hidden message, the HCC is estimated by the number of additional information bits per 1 pixel of the DI-container.

The mentioned peculiarity of the modern use of the LSB greatly complicates, and sometimes makes impossible the process of a hidden channel detecting by existing steganalysis algorithms, the majority of which is oriented at level of HCC which is not less than 0.25 bpp [3-5]. Although some research in this field is being carried out [6,7], the final solution for detection of the steganographic channel with low HCC has not been found. The reasons for this are clear: the changes introduced into the container matrix by the LSB method while the low HCC are minor and it is difficult or even impossible to detect them, for example, by image statistics analysis [8] or by change of some formal parameters, which occurred because of steganographic transformation [9]. The following example illustrates this fact. Let us consider the steganographic transformation with the use of the LSB-matching implementation [3] for a Jpeg digital image of size 800×800 pixels, stored with quality factor QF=75 and with HCC of 10%, 5%, 1%. As the result of the additional information embedding the obtained perturbation matrix has the spectral norm of 46.38, 17.76, 6.24, respectively, while the matrix of the input DI-container had the spectral norm of 60951.

Obviously, in order to ensure the possibility of detecting a steganographic channel in conditions of low HCC (less than 0.1 bpp (or 10%)), it is necessary to look for fundamentally new, possibly, unconventional ways of solving this problem.

Aim and tasks of the research

Currently, due to huge volumes of data transmitted through communication channels, information is usually stored in lossy formats (LF). Considering this, it is logical to assume that with a high probability DI in LF will be used as containers while the organization of a hidden communication channel. In particular, the most common format Jpeg with different QF will be used. Having this in mind, the *aim* of the work is to increase the detection efficiency of DI integrity violation by developing a new steganalysis method (methods) for detection of additional LSB-embedded information in the lossy image container under low capacity of hidden communication channel.

The detection effectiveness of DI integrity violation (as a result of steganographic transformation) will be estimated by Type I errors (DI modified by steganographic transformation detected as original one) and Type II errors (original DI detected as DI with integrity violation).

To achieve the aim, the following tasks must be solved:

1. Chose an approach, which will make possible to achieve the aim;

2. Determine the formal parameters of the DI and their features, which should be analyzed during steganalysis under conditions considered in the paper;

3. Develop a steganalysis method (methods), perform an estimation of its algorithmic implementation efficiency.

Main Body

Further, DI in LF (Jpeg format) will be used for hidden communication channel organization.

When the LSB method used not only in the spatial domain of the container, which is most often done in practice [1], but also in the frequency domain, taking into account that the method is not robust against any perturbing actions, the resultant steganographic message will be stored in the lossless format (LLF). Thus, the original Jpeg image, after its successful steganographic transformation by the LSB-method will be resaved in the LLF. This means that an indirect indicator of the presence of embedded additional information in DI is its repeat storing from the LF to the LLF. It should be noted that if the process of steganalysis is based on the detecting of this fact, then the corresponding method will obviously be most efficient in the case of a low HCC, and its efficiency should increase with a decrease of HCC. Indeed, as shown in [6], the process of embedding of additional information in a Jpeg container by the LSB-method modifies its formal parameters in such a way that the parameters of steganographic message move towards properties of DI in the LLF both qualitatively and quantitatively. It becomes more obvious with increase of the HCC. However, with a low HCC (no more than 10%), it is impossible to distinguish between the properties of the formal parameters of the received steganographic message and the properties of the original Jpeg container parameters. Moreover, the lower the HCC, the smaller this difference. Then the idea of steganalysis with a low HCC can be based precisely on the practical inseparability of the formal properties of steganographic message with a low HCC from the Jpeg container, which is performed in steganographic method development section of this work. Here the low HCC will not complicate, but simplify the process of detecting the results of steganographic transformation

In [10], the author proposed a method of separating original DIs saved in LLF from those that were resaved in LLF from LF. The detection of last fact was considered as an indirect indicator of DI integrity violation. The proposed method will be based on a general approach to the solution of the problem of detecting integrity violations of digital information content [11]. For blocks of original DI with $n \times n-$matrix F an equation was set as basis:

$$\angle\left(u_1, \bar{\sigma}\right) \approx \angle\left(v_1, \bar{\sigma}\right) \approx \angle\left(n^o, e_1\right), \quad (1)$$

where $\angle\left(u_1, \bar{\sigma}\right)$, $\angle\left(v_1, \bar{\sigma}\right)$, $\angle\left(n^o, e_1\right)$ are the magnitudes of the angles between the vectors u_1 and $\bar{\sigma}$, v_1 and $\bar{\sigma}$, n^o and e_1 respectively, u_1, v_1 – respectively, left and right normalized lexicographically positive singular vectors of $l \times l-$block B of matrix F, which correspond to its maximum singular number (SN) σ_1; $\bar{\sigma} = \sigma / \|\sigma\|$, wherein $\sigma = \left(\sigma_1, \sigma_2, ..., \sigma_l\right)^T \in R^l$ is a vector, obtained from l SN B $\sigma_1 \geq ... \geq \sigma_l \geq 0$, $n^o = \left(1/\sqrt{l}, 1/\sqrt{l}, ..., 1/\sqrt{l}\right)^T \in R^l$ is a n-optimal vector of space R^l, $e_1 = (1,0,..., 0) \in R^l$ is the first vector of the standard basis R^l. The approximate equality (1) is based on the proof of proximity of normalized vectors $\bar{\sigma}$ and e_1, u_1 and n^o, v_1 and n^o [11]: the angles between the respective vectors are close to zero in most blocks of the original DI.

Let the matrix of the original DI in the LLF be F_T, and the matrix of the corresponding DI, which was obtained from LF is F_J. Each of these matrices is spitted into standard $4 \times 4-$blocks which do not interleave. Let denote an arbitrary $4 \times 4-$block F_T / F_J as B_T / B_J respectively. In [10] proved, that

$$\angle\left(e_1,\bar{\sigma}_T\right) > \angle\left(e_1,\bar{\sigma}_J\right), \quad (2)$$

where $\angle\left(e_1,\bar{\sigma}_T\right)$, $\angle\left(e_1,\bar{\sigma}_J\right)$ are magnitudes of the angles between the vectors $\bar{\sigma}_T$ and e_1, $\bar{\sigma}_J$ and e_1 respectively, $\bar{\sigma}_T, \bar{\sigma}_J$ are normalized vectors of the SN blocks B_T, B_J. Taking into account the research results of the angles $\bar{\sigma}_J$ and e_1 [10] behavior in DI repeat storing in LF with different quality factors the equation (2) allowed to develop a method of separating the original DI in the LLF from those that were resaved in LLF from LF.

Attempts to use this method without any adaptation for steganalysis purpose did not lead to the desired (efficient) result, since, as noted above, after the steganographic transformation, the DI no longer has those formal properties that are specific to the Jpeg container. To ensure the possibility of conducting efficient steganalysis, it is necessary to determine and take into account the following properties of the considered formal parameters of the DI, which will be practically insensitive to small perturbations. Such a perturbation in this case is the process of DI embedding by the LSB-method with a low HCC. The smaller the sensitivity of the selected formal parameters, the higher is the expected efficiency of steganographic messages recognition process. It is desirable that the small perturbations due to embedding of additional information with low HCC to DI do not modify quantitatively and/or qualitatively the properties of the selected parameters of the Jpeg container blocks.

The algorithm proposed in [10] performs the separation of the original DI in LLF from the image resaved in the LLF from LF and based on the analysis of the mean value change of angle $\angle\left(e_1,\bar{\sigma}\right)$ per block in $4\times4-$blocks in corresponding DI group. The DI group, obtained by repeat storing of the analyzed DI in LF with different quality factors. In order to make the conclusion that DI is resaved in LLF from LF it is sufficient to discover that values $\angle\left(e_1,\bar{\sigma}\right)$ in analyzed DI are not the biggest ones in the group of obtained angles for corresponding DIs. However, this feature is not sufficient to distinguish between the steganographic message received on the basis of the Jpeg container and stored in the LLF and the original DI in LLF, despite the fact that the vector $\bar{\sigma}$ is insensitive to perturbation [12]. It follows from the foregoing and confirmed by the experiment carried out in this work.

It was found by experiment that there are some cases when the mean value $\angle\left(e_1,\bar{\sigma}\right)$ per block of staganographic message will be the largest in the group of the obtained angles for resaved to Jpeg DI staganographic message with different *QF*s. However, it will not retain a monotonous decrease in the considered angle with decreasing the value of the quality factor using the lossy compression of analyzed DI (which is characteristic of the original DI, it means that an image has not been steganographically transformed). This is possible for steganographic message in the case when the DI-container was originally saved in the lossy format with a high quality factor (for a Jpeg format with *QF* greater than 80). An embedding of addition information, resulting in relative comparability of the formal properties (mean value $\angle\left(e_1,\bar{\sigma}\right)$ per the image block) of obtained steganographic message with properties of similar parameters of the DI in the LLF, gives the outcome similar to original DI in LLF: $\angle\left(e_1,\bar{\sigma}\right)$ in DI stagographic message is the largest in obtained for corresponding angles group of DI. However, since the lossy compression was applied to the container before the staganographic transformation, the contribution to its block of high frequencies (and possibly middle frequencies) is quantitatively smaller than in blocks of corresponding DI without loss. The process of staganographic transformation will surely violate this relationship, but with a low HCC this violation will not lead to a qualitatively different result, it will be a tiny quantitative change: the process of steganographic message compression as it is repeated for container, will lead to a non-monotonic change of value $\angle\left(e_1,\bar{\sigma}\right)$ with reduction of quality factor, which discussed in detail in [10].

Taking into account the foregoing, the main steps of the proposed LF-container oriented steganalysis method SM1 are as follows.

Let F be the matrix of analyzed lossless DI (Tif).

Step 1. Resave the DI with matrix F in LF Jpeg with different quality factors $QF_i \in \{1,2,3,...,100\}, i = \overline{1,t}$. The result is the DI with matrices $F_i, i = \overline{1,t}$.

Step 2. For $F, F_i, i = \overline{1,t}$ determine: k_{sr}, k_{QF_i}, $i = \overline{1,t}$ which are the mean values per DI $4 \times 4-$blocks of angles $\angle(e_1, \bar{\sigma})$ respectively.

Step 3. *If* among the calculated values k_{QF_i}, $i = \overline{1,t}$, exists $\bar{k}$, that

$$k_{sr} < \bar{k},$$

then F corresponds the DI-steganographic message,

otherwise

if

for $QF_1 < QF_2 < ... < QF_t$ the monotonic increase is broken

$$k_{QF_1} < k_{QF_2} < ... < k_{QF_t},$$

than

F corresponds DI-steganographic message,

otherwise

the steganographic transformation has not been successfully applied to F.

In order to evaluate the efficiency of the proposed method, SM1 was implemented algorithmically with $QF_{i+1} = QF_1 + 5i$, where $QF_1 = 65$, $i = \overline{1,5}$. The choice of quality factors caused by frequency of their practical use. A computational experiment was conducted in the Matlab environment. The experimental set (ES) consisted of 300 DIs in Jpeg format from NRCS database [13], which is considered as a traditional one for working with DI. Steganographic messages were formed on the basis of this set using the LSB method with HCC 100%, 10%, 5%, 1% and 100 original DIs stored in LLF (Tif). The results of the computational experiment are represented in table 1.

Table 1

HCC (%)	100	10	5	1
Type I error (%)	14.7	3.7	2.7	1.7
Type II error (%)	4			

As can be seen from the above results, the developed SM1 method is supposed to be efficient for low HCC values, and its efficiency increases with decreasing of HCC. Such a result has not been observed for any of the analogue methods [3-5]. Moreover, the lowest value of HCC, with which the steganographic channel is efficiently detected by modern steganalysis methods [6, 7], as it follows from open sources, is 0.05 bpp, while the developed SM1 method is effective under HCC conditions, whose value is five times smaller – 0.01 bpp.

The efficiency of SM1 under low HCC conditions may be increased due to next improvement. In [14], it was shown theoretically and practically confirmed that the vector formed from the SN block B of matrix of the DI is more sensitive to perturbation than the vector compiled from the eigenvalues of the matrix BB^T ($B^T B$), which are defined as squares of the SN of B. Considering this, it is proposed to use for examine not the mean value per block of angle $\angle(e_1, \bar{\sigma})$, but angle $\angle(e_1, \bar{\bar{\sigma}})$, where $\bar{\bar{\sigma}}$ is the normalized vector of eigenvalues BB^T ($B^T B$).

Using $\bar{\bar{\sigma}}$ instead of $\bar{\sigma}$ will increase the efficiency of the steganalysis process, because $\angle(e_1, \bar{\bar{\sigma}})$ is

less sensitive to perturbation, it will be less responsive to the embedded additional information (with a low HCC) than $\angle\left(e_1,\overline{\sigma}\right)$. In accordance with this remark, a steganoanalytic method SM2 is proposed. Step 2 of it is different from SM1 and looks as follows.

Step 2 of SM2 method. Determine for DI $F, F_i, i=\overline{1,t}$: the mean values k_{sr}, k_{QF_i}, $i=\overline{1,t}$, which are calculated per $4\times4-$block of DI $\angle\left(e_1,\overline{\overline{\sigma}}\right)$ in the input original DI and DI, obtained after repeat storing an input DI to Jpeg with quality factors $QF_i, i=\overline{1,t}$ respectively.

The SM2 method was implemented algorithmically under the same conditions as SM1. The results of its testing were obtained for DI from the ES and represented in table 2. Results fully confirm the feasibility of using in the DI the mean value $\angle\left(e_1,\overline{\overline{\sigma}}\right)$ as an analyzed parameter: compared to SM1, the Type I error for HCC 0.05 bpp were reduced by 25.9%, and for HCC 0.01 bpp – by 23.5%.

Table 2

HCC (%)	100	10	5	1
Type I errors (%)	14.7	3.7	2.0	1.3
Type II errors (%)	3			

Remark 1. The computational complexity of SM1 and SM2 in the case of examination of DI with $n\times n-$matrix F is evaluated by the number of blocks which F is split into. It is $\underline{O}\left(n^2\right)$ operations, but number of operations in SM2 is greater by $\underline{O}\left(n^2\right)$ than in SM1 due to the calculation of the eigenvalues of the matrix BB^T (B^TB), which corresponds to the block B.

Remark 2. Taking into account the obtained estimates of the efficiency for the developed SM1 and SM2 (Table 1, 2), they can be used after the previous "rough" steganalysis of DI (with a HCC greater than 0.01 bpp) if it did not detect the embedding of additional information.

Conclusions

In the work on the basis of the new general approach to the problem of detecting integrity violations of DI the two steganalysis methods SM1 and SM2 developed. They are used for detection of hidden communication channel based on the LSB for DI-container in lossy format. Such methods are efficient for low hidden communication channel capacity (less than 0.1 bpp). Type I errors for algorithmic implementation SM1 and SM2 under the conditions of HCC 0.01 bpp are 1.7 and 1.3%, respectively, while Type II errors do not exceed 4%. Existing contemporary analogs, which are available in open sources, under conditions of such a low HCC are ineffective. The computational complexity of both methods due to their block organization for the $n\times n-$matrix of the analyzed DI is defined as $\underline{O}\left(n^2\right)$ operations, but the direct number of operations for SM2 is larger than for SM1 by $\underline{O}\left(n^2\right)$. Choice between the two polynomial and efficient under low HCC conditions steganalysis methods SM1 and SM2 should be made depending on the specific conditions of use.

Literature:

1. Стеганография, цифровые водяные знаки и стеганоанализ / А.В. Аграновский, А.В. Балакин, В.Г. Грибунин, С.А. Сапожников. – Москва : Вузовская книга, 2009. – 220 с.

2. Li B. A Survey on Image Steganography and Steganalysis / B. Li, J. He, et al. // Journal of Information Hiding and Multimedia Signal Processing. – 2011. – Vol.2, No.2. – PP.142–172.

3. Zhihua X. A Learning-Based Steganalytic Method against LSB Matching Steganography / Zhihua Xia, Lincong Yang et al. // Radioengineering. – 2011. – Vol.20, No.1. – PP. 102–109.

4. Natarajan V. Universal Steganalysis Using Contourlet Transform / V. Natarajan, R. Anitha // Proceedings of the Second International Conference on Computer Science, Engineering & Applications (ICCSEA 2012), May 25-27, 2012, New Delhi, India. – 2012. – Vol.2. – PP. 727–735.

5. Liu Q.Z. Image complexity and feature mining for steganalysis of least significant bit matching steganography / Q.Z. Liu, A.H. Sung, et al. // Information Sciences. – 2008. – Vol.178, No.1. – PP. 21–36.

6. Бобок И.И. Стеганоаналитический метод для цифрового сигнала-контейнера, хранящегося в формате с потерями // Сучасний захист інформації. – 2011. – №2. – С. 50–60.

7. Ахмаметьєва Г.В. Стеганоаналітичний алгоритм для цифрових контейнерів, збережених в форматах з втратами // Сучасна спеціальна техніка. – 2016. – № 3. – С. 31–8.

8. Жилкин М.Ю. Метод выявления скрытой информации, базирующийся на сжатии данных / М.Ю. Жилкин, Н.А. Меленцова, Б.Я. Рябко // Вычислительные технологии. – Новосибирск : Изд-во ИВТ СО РАН, 2007. – Т. 12. – С. 26–31.

9. Geetha S. Close color pair signature ensemble adaptive threshold based steganalysis for LSB embedding in digital images / S. Geetha, S. Sindhu, and N. Kamaraj // Transactions on Data Privacy. – 2008. – Vol.1, Iss.3. – PP. 140–161.

10. Бобок І.І. Метод виявлення зображень, перезбережених у формат без втрат з формату з втратами // Математичне та комп'ютерне моделювання. – 2017. – Вип.16. – С.5-14.

11. Kobozeva A.A. General Principles of Integrity Checking of Digital Images and Application for Steganalysis / A.A. Kobozeva, I.I. Bobok, A.I. Garbuz // Transport and Telecommunication. – 2016. – Vol. 17, Issue 2. – PP. 128-137.

12. Бобок И.И. Теоретическое развитие общего подхода к проблеме выявления нарушений целостности цифровых контентов, основанного на анализе полного набора формальных параметров // Інформатика та математичні методи в моделюванні. – 2017. – Т.7, №3. – С. 170-177.

13. NRCS Photo Gallery: [Електронний ресурс] // United States Department of Agriculture. Washington, USA. Режим доступа: http://photogallery.nrcs.usda.gov.

14. Бобок І.І. Дослідження властивостей формальних параметрів цифрового зображення в умовах порушення його цілісності // Сучасна специальна техника. – 2017. – №4. – С. 35-43.

Odesa National Polytechnic University *Received 24.08.2018*