Mathematics Subject Classification (2010) 94A08, 94A62

IMPROVEMENT OF NOISE IMMUNITY STEGANOGRAPHY ALGORITHM

Iryna I. Borysenko

Odessa National Polytechnic University,
1, Ave. Shevchenko, Odessa, 65044, Ukraine; e-mail: boris_enko@ukr.net

The research field is computer steganography, namely, stego systems, which ensure secret transmission of confidential information. The article considers the issue of improving the noise immunity of the algorithm, developed to create stego messages in the conditions of the ideal communication channel.

Keywords: steganography, information hiding, graph theory

Introduction

One of the topical, but at the moment incompletely solved problems is copyright, protection of intellectual property rights and confidential data of digital format. An important direction in solving this issue is the development of the methods of hiding information, in particular, methods of digital steganography [1].

General feature of all steganographic methods is the fact that the secret message (SM) is embedded in innocuous looking object or carrier. Digital images, videos, sound files, and other computer files that contain perceptually irrelevant or redundant information can be used as "covers" or carriers to hide secret messages. The cover object with the secretly embedded message is then called the stego object. After embedding a secret message into the cover image, a so-called stego image is obtained. This stego object is then transferred to other end, there we have detector algorithm which extract the message from cover object.

In [2] a new staganographic algorithm was suggested (we will call it a *Stego_Graph*) for sending and decoding a secret message, based on the application of the graph theory. The offered algorithm was developed for such information-hiding systems, where the capacity (C) is maximized while providing the required secrecy of a stego channel, whereas the requirements to the noise immunity are at the minimum.

The aim of this work is the improvement of noise immunity of *Stego_Graph*, provision of the possibility of its use in the systems, where noise is present.

Recently was developed a general approach to the analysis of information systems, based on the perturbation theory and matrix theory [3]. These theories selected as tool to improve the noise immunity a *Stego_Graph*, that exploits cover image. A mathematical model for digital image is a matrix.

Therefore, to meet the set aim, it is necessary to solve the following *tasks*:

- Detailed analysis of *Stego_Graph*, enabling to point out the reserves to improve its noise immunity;
- Analysis of the cover image matrix singular vector perturbation after embedding SM;
- Modification of the procedure of embedding SM into the cover image to change the region of singular vector perturbation localization;
- Practical confirmation of the improvement of noise immunity of the *Stego_Graph* by a computing test and the comparative analysis of the results of its modifications work.

Brief overview of Stego_Graph

Procedure of embedding SM is based on changing the brightness of some pixels in the primary image – $f(x,y)$. Practical tests showed that to provide reliability of the perception of a stego image (the embedding process not degrade the visual quality of the cove image), the calculated brightness of pixel $f'(x,y)$ must be within

$$f(x,y)-\delta \le f'(x,y) \le f(x,y)+\delta, \qquad (1)$$

where δ – maximum acceptable value of the deviation in the pixel brightness from the primary value.

Research [2] offers an iteration algorithm – Algorithm 1 to divide an image into sub-regions and determines threshold value for each sub-region.

To illustrate the work of the algorithm it is possible to consider the third block of image *Pout.tif* dimension 8x8. Suppose $\delta = 15$. This value is given experimentally for this image. In the result of the Algorithm 1 work, the image will be segmented into four sub-regions with gradation of brightness 214-184, 183-153, 152-122, 121-106 with corresponding values of thresholds $T_1 = 199$, $T_2 = 168$, $T_3 = 137$, $T_4 = 114$. Calculating T_4 we applied operation of rounding-off. Then we may perform threshold transformation of the image with an adaptive threshold. The matrix of the primary image and the results of the transformation with an adaptive threshold are shown in Figs. 2(a) and 2(b) respectively. Various tints of grey represent four abovementioned sub-regions.

Thus, in the result of the threshold transformation we receive a binary matrix and, further, we will use it to embedding a secret message.

To prepare SM for embedding into the cove image the graph theory is used. Basic data and definitions on the graph theory can be found e.g. in [4].

A secret message is represent as a marked graph-tree and then matrix S us made out of the tree according to the rule

$$s_{ij} = \begin{cases} 0, \text{ if vertex } v_i \text{ don't adjacent } v_j, \\ 1, \text{ if vertex } v_i \text{ adjacent } v_j \end{cases} \qquad (2)$$

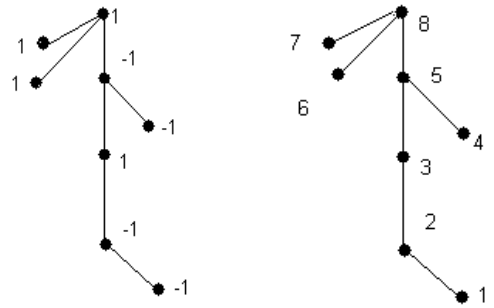
Algorithm of forming SM matrix – Algorithm 2 is shown in [2].

Suppose we have a given message with the length of 8 bits: $\{1,1,1,-1,-1,1,-1,-1\}$, then building a graph for this message and its matrix is demonstrated in Fig. 1.

The elements of the main diagonal are not used by the algorithm of hiding information, therefore, in Fig. 1(c), showing matrix S, on the matrix diagonal graph vertex are numbered. As S is symmetrical, it is possible to use one of its triangular sub-matrixes – lower or upper. Analyzing the structure of lower sub-matrix S (to be specific) we notice that first 7 bits of the secret message are built into the elements s_{ij} $j = \overline{1,n-1}$, $i = \overline{j+1,n}$, beginning from element s_{87}.

Thus, moving upwards along the diagonal, which is lower than the main one, it is possible to restore all the primary sequence, using a unit in the matrix cells as a sign switch. For the eighth bit it is possible to identify any cell of the sub-matrix, lying lower than the second diagonal, say, to be specific – s_{81}.

In general SM, presented by sequence 1 and -1, may start with any sign so this information should be also reflected in the matrix, e.g. in the following way. If $s_{87} = 1$, the sequence starts with a plus, if $s_{87} = 0$, with as minus.



1	1	0	0	0	0	0	0
1	2	1	0	0	0	0	0
0	1	3	0	1	0	0	0
0	0	0	4	1	0	0	0
0	0	1	1	5	0	0	1
0	0	0	0	0	6	0	1
0	0	0	0	0	0	7	1
0	0	0	0	1	1	1	8

a **b** **c**

Fig. 1. Message tree and its matrix: a – tree of the primary message; b – marked tree of the primary message; c – message tree matrix

The process of SM embedding into the cover image consists in the following. The matrix of the primary image F is to be segmented into blocks with dimension 8x8 in a standard way, so that the union of all blocks makes up matrix F.

Information is embedding into every block, therefore, we will be limited to the description of SM embedding into one of them into the lower triangle.

For the marked cover image block we make threshold transformation under Algorithm 1, receiving in the result a matrix, and defining it as G. From the message to be plunged into the cover image, we single out the sequence of 8 bits and for the highlighted sequence matrix S is made under Algorithm 2.

Information embedding into the block of cover image will occur in the result of the correction of the brightness of the container pixels (Algorithm 3 [2]) , that will take place only provided that it is found out that there is a non-correspondence of value of elements $g_{i,i-1}$ and $s_{i,i-1}$, $i=\overline{2,8}$ and elements g_{81} and s_{81} of matrixes G and S. If $g_{i,i-1} \neq s_{i,i-1}$, $i=\overline{2,8}$, the value of the pixel brightness $f_{i,i-1}$ of matrix F should be raised or lowered depending on the value of $s_{i,i-1}$. For example, if $s_{i,i-1}=0$, it means that $g_{i,i-1}=1$ and, consequently, $f_{i,i-1}$ should be lowered. New value $f_{i,i-1}$ depends on the threshold value T of the sub-region $f_{i,i-1}$ belongs to, and it will be determined as follows. If $s_{i,i-1}=0$, new brightness value $f'_{i,i-1}$ will be $f'_{i,i-1}=T$, otherwise $f'_{i,i-1}=T+1$.

We may illustrate the above for the third block of image *Pout.tif* and SM $\{1,1,1,-1,-1,1,-1,-1\}$. The stage of creating a stego message are shown in Fig. 2. The matrix of the stego message (stego image) – Fig. 2(d)) highlights the elements that were changed.

Similarly, the following 8 bits of information are embedded into the upper triangular sub-matrix of matrix F. Thus, 16 bits of information are embedded into one block of the cover image.

The procedure of extracting SM from the cover image is very easy and consists in identifying the sign of SM elements. If an element of matrix G' of the cover image threshold transformation, where a bit of information was embedded, equals 1, the sign of the SM element should be changed for the opposite in relation to the previous elements.

213	214	212	197	137	112	110	109
189	210	213	210	165	119	110	108
159	204	210	212	189	137	119	107
133	186	204	209	194	153	128	107
137	175	201	210	195	145	126	107
189	164	175	199	175	131	121	108
214	186	164	154	142	121	113	107
213	200	177	131	119	113	112	106

a

1	1	1	1	0	0	0	0
0	1	1	1	0	1	0	0
0	1	1	1	0	0	1	0
0	0	1	1	0	0	0	0
0	1	1	1	0	1	0	0
0	0	1	0	1	0	1	0
1	0	0	0	1	1	0	0
1	1	1	0	1	0	0	0

b

1	0	0	0	0	0	0	0
1	2	0	0	0	0	0	0
0	1	3	0	0	0	0	0
0	0	0	4	0	0	0	0
0	0	0	1	5	0	0	0
0	0	0	0	0	6	0	0
0	0	0	0	0	0	7	0
0	0	0	0	0	0	1	8

c

213	214	212	197	137	112	110	109
200	210	213	210	165	119	110	108
159	204	210	212	189	137	119	107
133	186	199	209	194	153	128	107
137	175	201	210	195	145	126	107
189	164	175	199	168	131	121	108
214	186	164	154	142	114	113	107
199	200	177	131	119	113	115	106

d

Fig. 2. Stages of SM embedding into container: a – matrix F of the primary image; b – matrix G of threshold transformation F; c – matrix of message S; d – matrix F' of stego message

Modification of the basic algorithm.
The development of *Stego_Graph_1* **and** *Stego_Graph_2*

Stego_Graph has high reliability of perception, as the cover image matrix is slightly changed under the embedded of a secret message into it, good capacity – 0.25 bit/pixel, and the volume of correctly restored information reaches 100%, but *Stego_Graph* may be applied only under conditions of the ideal communications channel. In case of the appearance of insignificant noise the volume of correctly restored information decreases to 65%. Detailed analysis of *Stego_Graph* enabled to single out hidden reserves to improve its noise immunity.

It is to be borne in mind that following Algorithm 3 of embedding SM into the cover image, a cover image pixel is given value $T+1$, if 0 of the characterizing cover image matrix should be transformed into 1, and T otherwise, where T is the threshold value. Therefore, the received value either does not differ from the threshold one at all, or it is by a unit higher. Application of slight noise, which is equivalent to, e.g. ± 2 gradations of brightness results in the fact that a pixel may be found in the previous sub-region. To make a pixel more stable it is logical to give it the value which will correspond to the middle of the required sub-region, but this may entail the correction of brightness by the value more than δ, which, in its turn, will certainly lead to the breach in reliability of perception. This refers to pixels with the values, close to the values of the limits of the region they belong to.

Regarding the third block of image *Pout.tif*, we may consider its threshold transformation, shown in Fig. 3.

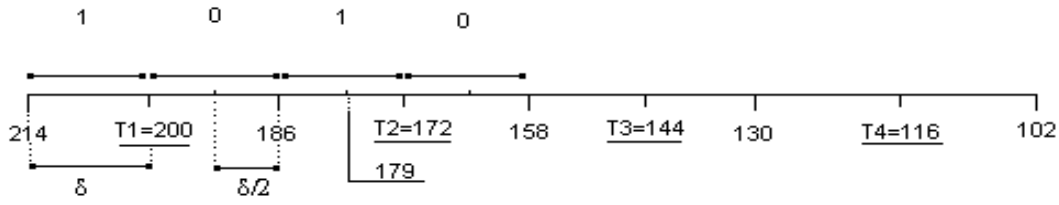


Fig. 3. Scale of threshold transformation of the third block of *Pout.tif*

We may set several agreements: the right limit does not belong to the current region (except the last), δ will be even (to be specific 14), the last region, if it is lower than 2δ should be determined as 2δ. All these agreements will ensure integer value of the threshold in all of the regions, which is important to ensure reliability of stego perception, if δ reaches maximum acceptable values.

Threshold transformation suggests transformation of the pixels brightness values into 0 or 1. Addressing Fig. 3, it is easy to notice that zeroes and units of different regions alternate. For a pixel, whose characterizing value is 1 (except the first sub-region), it does not matter which sub-region it will be transferred to. The point is to save reliability of a stego message perception. For example, we may consider the region with limits value 186-158 and threshold $T=172$. If a pixel with brightness value meets condition $179=186-\delta/2\le f(x,y)\le 186$ should be lowered so that its characterizing value is zero, it should be given value $T1-\delta/2=200-7=193$. However, if a pixel has the brightness meeting condition $T2<f(x,y)<T2+\delta/2$, it should be given value $T2-\delta/2=172-7=165$. At the same time the difference between the primary and new values will meet condition (1), i.e. it will be within $\pm\delta$.

Algorithm 3. Correction of cover image pixel brightness in *Stego_Graph_1*:

Step 1. Determine limits $L(z)$ and $L(z+1)$ of the region $f_{i,j}$ belongs to. Determine the value of threshold T_k

Step 2. If $g_{i,j}\neq s_{i,j}$ and $s_{i,j}=0$, then

If $k\neq 1$ (not the first region) and $L(z)-f_{i,j}\le\delta/2$, then $f_{i,j}=T_{k-1}-\delta/2$,

Otherwise $f_{i,j}=T_k-\delta/2$.

If $g_{i,j}\neq s_{i,j}$ and $s_{i,j}=1$, then

If $k\neq n$ (not the last region) and $f_{i,j}-L(z+1)\le\delta/2$, then $f_{i,j}=T_{k+1}+\delta/2$, otherwise $f_{i,j}=T_k+\delta/2$.

For the first and the last regions of the current block information embedding is made in the same way as in *Stego_Graph*.

Embedding SM into *Stego_Graph_2* we use information excess, each value of SM is recorded thrice, i.e. we create blocks of 000 and 111 type, and while decoding the sign is determined by the majority of similar symbols in the block.

Assessment of stego image sensitivity to noises in the communication channel

Solving the problem of improving noise immunity of the basic algorithm we used general approach to the analysis of information systems, based on the perturbation and matrix theories, that is shown in [3]. Here we will briefly describe its conception, used when developing the algorithms of *Stego_Graph_1* and *Stego_Graph_2*.

As the mathematical model for the cover image is a matrix, and all of the transformations of the image may be represented in the equivalent matrix view, then the set of parameters, strictly determining and comprehensively characterizing any image may be presented by the multitude of singular numbers and singular vectors of its matrix. Singular decomposition of matrix F can be shown as:

$$F = U\Sigma V^T, \tag{3}$$

where U and V are matrixes of left and right singular vectors, respectively (SV), whereas Σ – is the matrix of singular numbers (SN). In general singular decomposition of matrix is determined ambiguously. However, it is possible to perform normal singular decomposition, which is the only one.

Any stego transformation will perturb cover image matrix F, and consequently, somehow will perturb the corresponding SNs and SVs, which enables to reduce the task of the analysis of stego transformation to the analysis of SN and SV perturbations. Further, we will need two important statements, grounded in [5]. Firstly, SVs, corresponding to low SNs, receive significant perturbations even at low perturbations of the cover image, on the contrary, SVs, corresponding to high SNs are noted for their strong noise immunity. Secondly, SNs perturbations are comparable to the data perturbation – ΔF, i.e. cover image matrix SNs are no insensitive to the perturbing effects, regardless of whether the stego image is sensitive or not to the perturbing effects, therefore, it is reasonable to analyze only SVs whereas their aggregate perturbation can be used as a rate of sensitivity of the stego image to perturbing effects. Basing on the first statement, we made a conclusion that the algorithm of *Stego_Graph* should be modified so that the cover image matrix under the embed of SM, receives such perturbation that it will lead to the perturbation of not only SVs, corresponding to small SNs but also SVs that are more resistant to noises. The second statement was used to assess the noise immunity of *Stego_Graph* and its modifications.

In MATLAB environment we carried out a computing test, during which new stego methods were practically implemented and their noise immunity was assessed. To show the results we further use the third block of *Pout.tif*, which is typical of other blocks of *Pout.tif*, as well as for other images.

Owing to the specifics of SM embedding under the *Stego_Graph* cover image matrix algorithm receives rather low perturbations and, here, mostly SVs are perturbed, corresponding to low SNs – Fig. 4(a)) (6,7,8 SVs – SM is mostly in these vectors and under the smallest noise information is distorted). Embedding SM under *Stego_Graph_1* and *Stego_Graph_2* considerable perturbation is also made on SVs, corresponding to 5, 4 and 3 SNs, i.e. more information is transferred to the vectors, which are more resistant to additional effects. Analyzing the behaviour of SVs (their aggregate perturbation) of the cover image matrix, it is possible to make a conclusion that the most resistance stego image to the additional noises is *Stego_Graph_2*, which is confirmed by Fig. 4(b)).

Perturbations in the communication channel were simulated with the help of the additive Gauss noise, whose application was made in the standard way *imnoise*, with mean equaling zero and dispersions D=0.00001, D=0.0001 and D=0.001. As it was expected, the largest volume of correctly restored information was received under *Stego_Graph_2*.

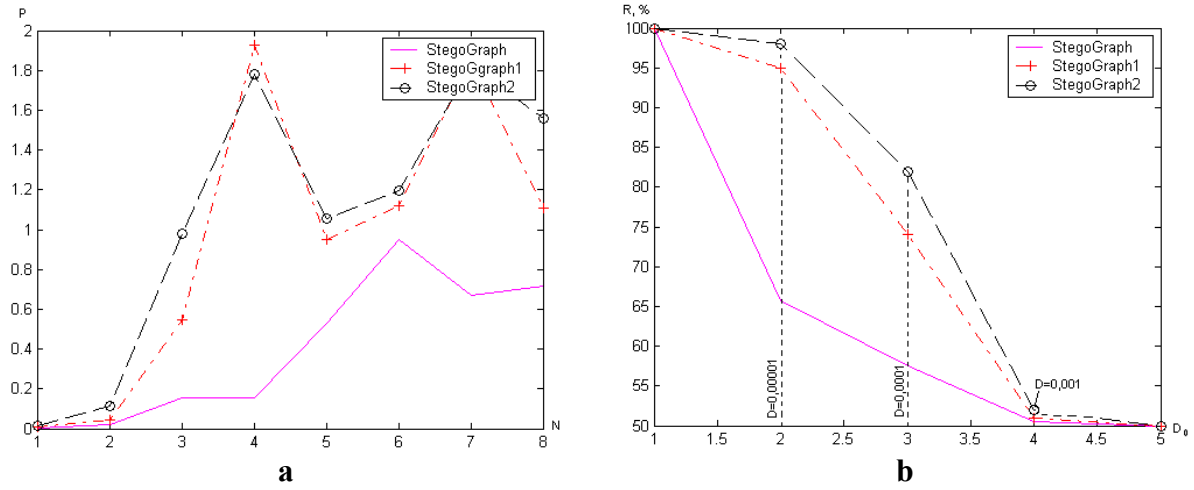


Fig. 4. Representation of embedded and restored information using algorithms *Stego_Graph*, *Stego_Graph_1*, *Stego_Graph_2*: a – Representation of embedded information by perturbed SVs (P – rate of difference between SV cover image matrix and stego image matrix, N – SV number); b – volume of correctly restored information ($R,\%$) under D_0 (mean perturbation of pixels)

Conclusion

The work considered the algorithms for stego systems, whose task is hidden transmission of confidential information. A special feature of these systems is hiding the very fact of message transmission. Therefore, one of the most important characteristics is the reliability of a stego message perception. It is logical to conclude that the less perturbations we make for the cover image, the more reliable stego channel is. It is this principle that was used to make up the algorithm of *Stego_Graph*, and the algorithm, e.g. shown in [6], where cover image pixels are, as a rule, replaced by the found, necessary pixels, but they are not corrected. All these algorithms, based on low perturbations of coding, are very sensitive to noises, because, as it was shown in the fourth section, low perturbations of the cover image under stego transformation lead to perturbations of only low-resistant SVs that correspond to low SNs. Under the impact of external effects these vectors receive additions perturbations and a significant part of SM, being in them, is lost.

To create noise resistant algorithm it is required to seek and perturb the vectors that correspond to high SNs. SVs, corresponding to the highest SNs are the most reliable. Even under strong noises, when the image is seriously distorted, it is possible to extract information from these vectors almost with no losses. However, unfortunately, we may not embed information into these vectors as reliability of a stego message perception is affected. This is explained by the fact that SVs, corresponding to the highest SNs of the image matrix match low-frequency, and the lowest – high-frequency components of the primary cover. Frequency sensitivity of the human eyesight system is in the fact that a person is much more sensitive to low-frequency, than to high-frequency signal component. Owing to the abovementioned the task of the development of noise resistant algorithm is the search of a compromise between its characteristics, as the improvement of one parameter, e.g. the capacity value is provided by other parameters, such as secrecy of information transmission or immunity to perturbing effects.

References

1. Хорошко В.А. Методы и средства защиты информации / В.А. Хорошко, А.А. Чекатков. – К.: Юниор, 2003. – 501 с.
2. Борисенко И.И. Особенности применения многоуровневого порогового преобразования изображения в компьютерной стеганографии // Праці УНДІРТ. Теоретичний та науково-практичний журнал радіозв'язку, радіомовлення і телебачення. – 2006. – № 4. – С.53-59.
3. Кобозева А.А. Анализ информационной безопасности / А.А. Кобозева, В.А. Хорошко. – К.: Изд.ГУИКТ, 2009. – 251 с.
4. Ф. Харари. Теория графов / Пер. с англ. В.П. Козырев; под ред. Г.П. Гаврилова – М.: Мир, 1973. – 300 с.
5. Кобозєва А.А. Загальний підхід до оцінки властивостей стеганографічного алгоритму, заснованого на теорії збурень // Информационные технологии и компьютерная инженерия. – 2008. – № 1. – С. 164-171.
6. S. Hetzl and P. Mutzel. A graph-theoretic approach to steganography. In J. Dittmann et al., editor, Communications and Multimedia Security. 9th IFIP TC-6 TC-11 International Conference, CMS 2005 volume 3677 of Lecture Notes in Computer Science, pages 119-128, Salzburg, Austria, September 19-21 2005.
7. Иванов Б.Н. Дискретная математика. Алгоритмы и программы: уч. пособ. / Б.Н. Иванов. – М: Лаборатория Базовых Знаний, 2003. – 288 с.

І.І. Борисенко

ПІДВИЩЕННЯ СТІЙКОСТІ ДО ЗАВАД СТЕГАНОГРАФІЧНОГО АЛГОРИТМА

Областю наукових досліджень є комп'ютерна стеганографія, а саме стеганосистеми, які забезпечують таємну передачу конфіденційної інформації. В статті розглядається питання підвищення стійкості до завад алгоритму, який був розроблений для створення стеганоповідомлень в умовах ідеального каналу зв'язку.

Ключові слова: стеганографія, приховування інформації, теорія графів

И.И. Борисенко

ПОВЫШЕНИЕ ПОМЕХОУСТОЙЧИВОСТИ СТЕГАНОГРАФИЧЕСКОГО АЛГОРИТМА

Областью научных исследований является компьютерная стеганография, в частности стеганосистемы, которые обеспечивают тайную передачу конфиденциальной информации. В статье рассматривается вопрос повышения помехоустойчивости алгоритма, который был разработан для создания стеганосообщений в условиях идеального канала связи.

Ключевые слова: стеганография, сокрытие информации, теория графов