Mathematics Subject Classification (2010) 94A60

Informatics and Mathematical Methods in Simulation
Vol. 3 (2013), No. 4, pp. 323-330

A TECHNIQUE TO ESTIMATE A STEGANOGRAPHIC CAPACITY OF A STEGO CHANNEL FORMED WITH LSB MODIFICATION APPROACH

Ivan I. Bobok

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

A polynomial technique for estimating a steganographic capacity (SC) of a stego channel formed with least-significant-bit (LSB) modification approach was developed within the framework of steganalysis efficiency-improvement job based on general approach to analysis of the status and performance of information systems. Digital images of lossy compression formats were used as cover objects. The technique developed is important to implement decoding of additional data embedded in a cover object. With respect to stego transformation with LSB approach, a relationship between the rate-of-change of smallest singular values of image blocks (i.e., submatrices) and SC value was established. It was empirically established that the absolute error of SC obtained by the developed technique is essentially independent of its true value, but dependent on the number of stego messages (images) subjected to analysis.

Keywords: steganalysis, steganographic capacity, singular values, digital image, matrix

Introduction

The state-of-art steganography, recent scientific activity in this area, and publications of some results in open-access sources have boosted the chance of various anti-state and terrorist structures to implement the new developments. Therefore, the issues related to increase in efficacy of steganalysis are of supreme importance [1, 2].

The main problems of steganalysis are:

- Detection of steganographically embedded data, and, if detected;
- Decoding the data embedded inside a cover object.

Currently, the first of these two problems remains the most important, and, when developing the techniques to solve it, a special attention is given to minimization of type-I errors [3].

Today, when organizing a stego communication channel, the most widely used approach is that of LSB modification [1]. However, in current use, it involves some specificity as follows: embedding additional data (AD) in a cover object is often implemented with a low steganographic capacity (SC) [1, 2], thus hampering detection of AD with steganalysis algorithms.

In [4], the author presented a steganalysis technique to detect the embedded data produced with LSB approach, that technique being efficient also for low SC. Digital images (DI) of lossy compression formats were used as cover objects. The basic idea of the technique developed based on general approach to analysis of the status and performance of information systems [4] involves estimating the rate-of- change of two smallest singular values (SVs) of blocks (submatrices) of DI under test, these smallest SVs obtained as a result of standard decomposition [5]. The conclusion of the presence or absence of embedded AD is made after comparison of the argument of the histogram of angular factors interpolating the 7^{th} and 8^{th} SVs of all DI blocks in which the global histogram maximum is achieved, with an experimentally set threshold value. Such use of the technique ensures effective detection of the result of stego embedding with LSB approach [3]; however, to perform decoding operations, it is important to estimate the SC value of a stego channel used for transfer of AI.

Therefore, development of a technique to estimate a steganographic capacity of a least-significant-bit approach with the performed stego transformation is deemed appropriate for organization of steganalysis process.

Aim of the Work and Problem Setting-up

The *aim* of the research is to develop a technique to estimate a steganographic capacity of a stego channel formed with least-significant-bit (LSB) modification approach based on the earlier developed steganalysis technique, by estimating the rate-of-change of two smallest SVs of DI blocks obtained with standard DI matrix decomposition. To accomplish the purpose, the following *problems* are to be solved:

- Establishing qualititative features of SVs of stego message (SM) blocks with increase in SC value;
- Establishing quantitative features of SVs of stego message (SM) blocks with increase in SC value;
- Ensuring a low computational complexity for the technique developed to estimate a steganographic capacity of stego communication channel, and
- Guaranteeing efficient work of the technique developed regardless of the real value of SC when embedding AI with the LSB approach.

Main Body

Let us review characteristic features of changes for the smallest SVs of blocks of lossy compressed (LC) cover objects (digital images) within the process of stego transformation (ST).

Comparison of properties of SVs of blocks of lossy and lossless compressed cover objects makes it possible to foresee the behaviour pattern of properties of SVs of blocks (i.e., submatrices) of LC cover objects within the course of ST. Presence of non-zero SVs within the submatrix is indicative of the submatrix degeneracy (linear dependence of its row vectors (column vectors)). If the smallest SVs are non-zero, but close to zero (as it is often observed in the DI fully restored after compression), the submatrix, being not degenerative, but badly conditioned, contains row (or column) vectors that are close to being linear dependent, e.g., the angle between some row (or column) vectors can be not equal, but close to zero. Any disturbance (in particular, ST), will change these angles in some way. However, apparently, the possibility that the angle will become equal to zero, thus leading to the increase in the number of zero SVs, is much lower than that the angle would remain non-zero after disturbance. Besides, due to specificity of the problem concerned, all row (or column) vectors of 8×8 blocks are geometrically located in the first coordinate orthant of R^8 vector space, and, in the main, no disturbance can lead them beyond the limits of this orthant. Therefore, if the close-to-zero angle between row (or column) vectors of the matrix is disturbed, the possibility of increase in value of the angle is higher, than that of decrease. In order to test the proposed hypothesis, we performed a computing experiment involving 250 JPEG DIs. The experiment was performed as follows. Each DI was decomposed into 8×8 blocks in a standard way. Block columns were normalized. For each normalized B block, an 8x8 $\overline{B}$ matrix was calculated with $\overline{b}_{ij}, i, j = \overline{1,8}$ elements equal to the scalar product of i^{th} and j^{th} columns of B matrix (i.e., cosine of the angle between i^{th} and j^{th} columns). Blocks, where corresponding $\overline{B}$ matrix contained close-to-one (>0.999) elements, were marked, and their total number for the each DI was counted. Furthermore, in marked blocks, the indices of close-to-linearly-dependent columns were saved. With the noise (Gaussian noise with zero expectation value and different variances) superimposed on the image, the matrix was decomposed into blocks again, and the whole above-mentioned procedure was repeated for them. The experiment results showed the following. The number of blocks for which $\overline{B}$ matrix contained close-to-one elements did not increase in 83% of DIs tested. At the stage of verification of the marked blocks, non-increase of the corresponding $\overline{B}$ elements (increase of the angle between column vectors of B) was registered in 93% of all close-to-linear-dependent columns found at the first stage.

The experiment, in case of disturbance, confirmed an increase in the degree of non-singularity of the corresponding block for DI of lossy-compressed format (LCF), decrease of its condition number, increase of deviation-from-zero value for the smallest SVs of the block, thus confirming that, as a result of ST, a block of LC cover object will be «losing» LCF properties of SVs, and «acquiring» those of lossless-compressed format (LLCF) [6].

Therefore, qualitative results of ST of LC cover objects with LSB approach will be as follows:

- Decrease in the quantity of non-zero SVs in blocks compared to LC cover-object blocks, and, the higher is SC of the organized stego channel, the greater will be this decrease, and,
- With the increase of SC, the behavior pattern of SVs of SS blocks will be more and more «matching» to that of SVs for lossless-compressed DI; in particular, for the majority of blocks, the rate of change for the smallest SVs will increase.

Thus, increase in the rate-of-change of the lowest SVs for the overwhelming majority of image submatrices comparing with the corresponding characteristic of LC cover object states that the ST procedure has been held, and the rate-of-change of the lowest SVs itself appears to be that very characteristic of disturbance of image parameters that allows not only to determine the presence of ST [3, 4, 6] results, but also to estimate the SC value.

Let us perform a detailed analysis of the parameters of stego message vs. the value of SC of a channel.

As it follows from above, ST will result in a decrease in number of degenerated blocks when compared to that of LC cover object, and, the higher is SC, the greater will be the number of degenerated blocks [7]. Using estimation of the number of non-degenerated DI blocks as a basis for conclusion on the presence/absence of AI in DI in actual practice causes difficulties due to availability of a vast number of LC cover objects with ≈100% percentage of the blocks not involving zero SVs. This results in great computational complexity in obtaining a threshold value of this number in order to separate the cover image from SM.

Let us review a relationship between the average rate-of-change for the smallest SVs of DI matrix blocks and SC. This relationship is reflected by the plot built according to the results of the computing experiment performed in *MathWorks* MATLAB environment with 500 DIs from the NRCS database [8], the latter being traditional one to test algorithms dealing with images; here the lowest SC was 1/20 bit per pixel (or 5%) (Fig. 1).

The results obtained fully quantitatively confirm the qualitative statement that, with the increase in SC, the behavior of SVs of blocks belonging to SM «tends to» match to that of SVs of lossless-compressed DI. This fact is reflected in strictly monotomic increase of the rate-of-change of smallest SVs with the increase of SC, and the average rate-of-change V vs. steganographic capacity S characteristic is practically linear (Figure 1) and can be approximated by the following relationship:

$$V \approx 0.004 \cdot S + 0.346. \quad (1)$$

This fact provides a possibility in principle for the following:

▪ Estimating quantitatively the average rate-of-change for the arbitrary SC within 5 to 100% range based on the plot obtained, and

▪ Estimating the SC value from average rate-of-change of the smallest SVs of the blocks, when developing SM with LSB approach, thus increasing efficiency of AI decoding (if required).

The results received can have practical importance when not one, but a set of DIs received during stego channel work with maintained SC undergo steganalysis. That is the basis for the *SPS* technique used to estimate SC of the stego channel organized with the help of LSB approach. The main steps are as follows:

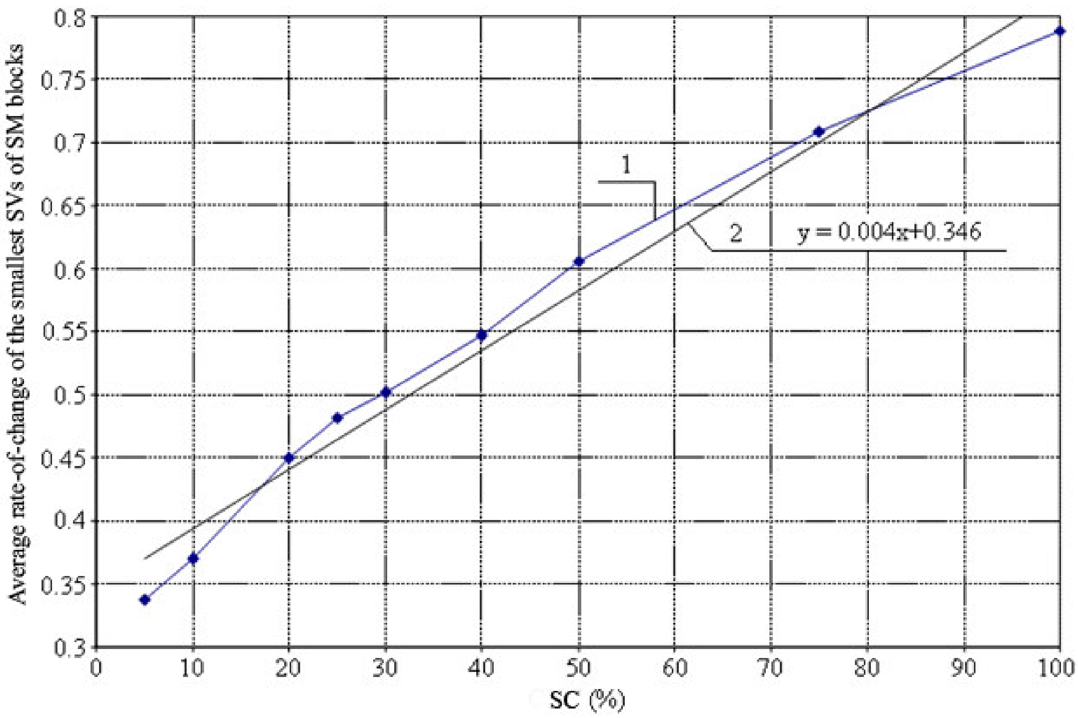

Fig. 1. Average rate-of-change of the smallest SVs of SM blocks vs. SC characteristic: 1 – interpolation spline of degree 1; 2 – linear approximation

Let $F_i,\ i=\overline{1,n}$ be the matrices of DIs under analysis, received from the same stego channel (it is supposed that stego messages were formed with the same SC).

Step 1. Decompose matrices $F_i,\ i=\overline{1,n}$ of digital images under analysis into 8×8 blocks (sumbatrices) $B_j,\ j=\overline{1,N}$, where *N* is the total number of the blocks obtained.

Step 2. For each of thus $B_j,\ j=\overline{1,N}$ submatrices obtained:

2.1. Determine the set of singular values as follows: $\sigma_1^{(j)} \geq \ldots \geq \sigma_8^{(j)} \geq 0$;

2.2. For $\sigma_7^{(j)}, \sigma_8^{(j)}$ determine angular factor $k^{(j)}$ of interpolating polynomial of degree 1.

Step 3. Determine V, the average rate-of-change of the smallest SVs of SM blocks of DIs under analysis:

$$V = \frac{\sum_{j=1}^{N} k^{(j)}}{N}; \tag{2}$$

<u>Step 4.</u> Estimate the SC value in accordance with (1) and using the result of (2):

$$S \approx \frac{V - 0.346}{0.004}. \tag{3}$$

The value of SC obtained in step 4 with formula (3) is approximate. Let us perform a computational experiment to test the technique developed. For this purpose, let us form an I set of DIs (cover objects) from 500 images taken from NCRS base. With this done, let us use LSB approach to form I_1, I_2, I_3, I_4, I_5 sets by stego transformation of the images from I set with steganographic capacity equal to 10, 30, 50, 70, and 90%, respectively. In order to determine the steganographic capacity used during the development of stego messages incorporated into proper sets, let us subject each of I_1, I_2, I_3, I_4, I_5 sets to analysis with SPS method. The test results are given in Table 1.

Table 1.

Results of tests of SPS method

The set under test		I_1	I_2	I_3	I_4	I_5
True value of SC (%)		10	30	50	70	90
SC value obtained/absolute error (%)	for complete set (500 DIs)	5/5	33/3	55/5	72/2	86/4
	for subset (400 DIs)	7/3	34/4	56/6	73/3	86/4
	for subset (300 DIs)	5/5	35/5	55/5	76/6	84/6
	for subset (200 DIs)	4/6	37/7	57/7	78/8	81/9
Average error value		4.75	4.75	5.5	4.75	5.5

As our computational experiment shows, when estimating steganographic capacity with SPS technique, the average of Δ (absolute error) practically does not depend on the true value of SC (including small SC values), but depends on K number of digital images subjected to analysis: the average of Δ (absolute error) denoted hereinafter by Δ_S, is increased with decrease in K (Fig. 2), which is natural if one takes into account the way in which relationship (1) was obtained.

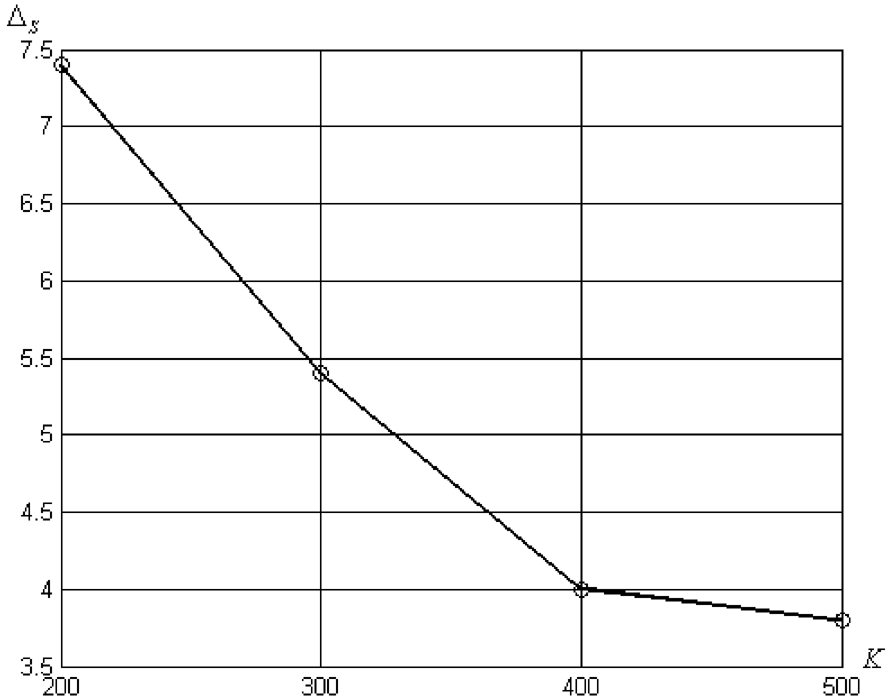

Fig. 2. Average absolute value of steganographic capacity determined with *SPS* method vs. Number of digital images subjected to analysis

Note. The computational complexity of *SPS* technique developed to estimate SC of the stego channel by analysis of a set comprising K number of $n \times n$-pixel digital images (with this channel organized by means of LSB modification approach) is determined by a number of blocks (i.e., submatrices) obtained by standard matrix decomposition. When it concerns one digital image, the computational complexity equals to $\underline{O}(n^2)$. Therefore, the resultant computational complexity of *SPS* technique is determined as $K \cdot \underline{O}(n^2)$.

Conclusions

In this work, a polynomial (of degree 2) technique for estimating a steganographic capacity of a stego channel formed with LSB approach was developed based on general approach to analysis of the status and performance of information systems. Digital images of lossy compression formats were used as cover objects.

During the performance of this work:

1) It was established that strictly monotomic increase of the rate-of-change of smallest singular values with the increase of SC is the qualititative feature of the two smallest singular values of blocks (i.e., submatrices) of a digital image. This fact supports the following hypothesis: with the increase of SC, the behavior pattern of SVs of SS blocks will be more and more «matching» to that of SVs for lossless-compressed DI.

2) With respect to stego transformation with LSB approach, a quantitative correspondence relationship between the rate-of-change of smallest singular values of digital image blocks (i.e., submatrices) and SC value was established.

3) It was empirically established that the absolute error of SC obtained by *SPS* technique is essentially independent of its true value, but dependent on the number of stego messages subjected to analysis.

МЕТОД ОЦІНКИ ВЕЛИЧИНИ ПРИХОВАНОЇ ПРОПУСКНОЇ СПРОМОЖНОСТІ КАНАЛУ, СФОРМОВАНОГО МЕТОДОМ МОДИФІКАЦІЇ НАЙМЕНШОГО ЗНАЧУЩОГО БІТА

І.І. Бобок

Одеський національний політехнічний університет,
просп. Шевченка, 1, Одеса, 65044, Україна; e-mail: onu_metal@ukr.net

У рамках підвищення ефективності стеганоаналізу на основі загального підходу до аналізу стану й технології функціонування інформаційних систем розроблений поліноміальний метод оцінки величини прихованої пропускної спроможності (ППС) каналу зв’язку, організованого за допомогою методу модифікації найменшого значущого біта (LSB-методу). Як контейнери використовувалися цифрові зображення у форматах із втратами. Розроблений метод є важливим при організації декодування вбудованої в контейнер додаткової інформації. Отриманий закон кількісної залежності середньої швидкості зміни найменших сингулярних чисел блоків матриці зображення від величини ППС при організації стеганоперетворення LSB-методом. Емпірично встановлено, що абсолютна похибка одержуваного розробленим методом значення ППС практично не залежить від її дійсного значення, а визначається кількістю зображень-стеганоповідомлень, що аналізуються.

Ключові слова: стеганоаналіз, прихована пропускна спроможність, сингулярні числа, цифрове зображення, матриця

МЕТОД ОЦЕНКИ ВЕЛИЧИНЫ СКРЫТОЙ ПРОПУСКНОЙ СПОСОБНОСТИ КАНАЛА, СФОРМИРОВАННОГО МЕТОДОМ МОДИФИКАЦИИ НАИМЕНЬШЕГО ЗНАЧАЩЕГО БИТА

И.И. Бобок

Одесский национальный политехнический университет,
просп. Шевченко, 1, Одесса, 65044, Украина; e-mail: onu_metal@ukr.net

В рамках повышения эффективности стеганоанализа на основе общего подхода к анализу состояния и технологии функционирования информационных систем разработан полиномиальный метод оценки величины скрытой пропускной способности (СПС) канала связи, организованного при помощи метода модификации наименьшего значащего бита (LSB-метода). В качестве контейнеров использовались цифровые изображения в форматах с потерями. Разработанный метод является важным при организации декодирования внедренной в контейнер дополнительной информации. Получен закон количественной зависимости средней скорости изменения наименьших сингулярных чисел блоков матрицы изображения от величины СПС при организации стеганопреобразования LSB-методом. Эмпирически установлено, что абсолютная погрешность получаемого разработанным методом значения СПС практически не зависит от ее истинного значения, а определяется количеством изображений-стеганосообщений, которые подвергаются анализу.

Ключевые слова: стеганоанализ, скрытая пропускная способность, сингулярные числа, цифровое изображение, матрица

4) It was established that the absolute error of result obtained by SPS technique is dependent on the K number of stego messages obtained with the same SC and subjected to analysis: the more is K number, the less is the error of estimated SC.

References

1. Грибунин, В.Г. Цифровая стеганография [Текст] : монография / В.Г. Грибунин, И.Н. Оков, И.В. Туринцев. — М. : СОЛОН-Пресс, 2002. — 272 с.
2. Стеганография, цифровые водяные знаки и стеганоанализ : [монография] / А.В. Аграновский, А.В. Балакин, В.Г. Грибунин, С.А. Сапожников. — М.: Вузовская книга, 2009. — 220 с.
3. Бобок, И.И. Использование метода анализа ROC-кривых для комплексной оценки эффективности стеганоаналитического метода / И.И. Бобок // Інформатика та математичні методи в моделюванні. — 2012. — Том 2, №3. — С. 221–230.
4. Бобок, И.И. Стеганоаналитический метод для цифрового сигнала-контейнера, хранящегося в формате с потерями / И.И. Бобок // Сучасний захист інформації. — 2011. — №2. — С. 50–60.
5. Гонсалес, Р. Цифровая обработка изображений / Р. Гонсалес, Р. Вудс; пер. с англ. П.А. Чочиа. — М. : Техносфера, 2006. — 1070 с.
6. Бобок, И.И. Детектирование наличия возмущений матрицы цифрового изображения как составная часть стеганоанализа / И.И. Бобок // Вісник Східноукраїнського національного університету ім. В. Даля. — 2011. — № 7(161). — С. 32–41.
7. Бобок, И.И. Стеганоанализ как частный случай анализа информационной системы / И.И. Бобок, А.А. Кобозева // Сучасна спеціальна техніка. — 2011. — № 2. — С. 21–34.
8. NRCS Photo Gallery: *[Електронний ресурс]* // United States Department of Agriculture. Washington, USA. Режим доступа: http://photogallery.nrcs.usda.gov (Дата обращения: 26.07.2012).