A. S. Vasiyra, Cand. Sc. (Eng.), Prof.; V. V. Lukichev

METHODS FOR PATTERN DATA INPUT ON VEIVLVET- FACTORS ON THE BASIS OF CRITERIA OF STENOGRAPHIC STABILITY

There had been considered the peculiarities of input, determining the secrecy and stability of hidden data. With an aim of development of efficient stenographic method, there had been synthesized the criteria, used as the target function during input.

***Key words**: stenography JPEG-algorithm, method of pattern input, secrecy, robustness, veivlvet-transformation, method of support vectors.*

Image stenography is a sphere which has been rapidly developing recently. Its objective may be described as secrete data hiding, which is stable to different transformation. Correspondingly, practical tasks to be solved within its limits, to this or that degree concern the aspects of secrecy and robustness [1,2].

Since the violation of secrecy may cause the total message loss, this described peculiarity determine the main restrictions during the designing of stegosystem. It is necessary to note, that the relative character of this factor stipulates for existence of large number of criteria, efficiency of which is different for different methods of input.

The second important aspect is the requirements to stability. The wide spread diagrams of excessive coding with protection against failures allows to solve the problem of robustness with efficiency, which is determined by the information restored [3].

Thus the designing of any stegosystem may be considered as the task of conditional optimization where the target function connects the robustness with the stage of secrecy and the restrictions determine the sphere of criteria adequacy. Such a universal approach allows to assure the high adaptability to the conditions of direct functioning of stegosystem.

The diagram of blind input is widely spread in image stenography where only the stegocontainer is transferred. This stipulates for peculiarity of stegoanalysis, the task of which consists in binary classification of images on the basis of peculiarities, which undergo the most changes during input. Most prosperous are the criteria on the basis of method of support vectors (SVM – support vector machines) [4], the greatest advantage of which is the efficiency of classification of dots – characteristics in multi – dimensional space of peculiarities.

The compression methods are the most popular among the image processing methods. The biggest factor of compression is possible to be achieved by compression methods with losses [5]. Compression method JPEG is very widely used, though there are many efficient formats on the basis of wavelet – transformation (for ex. JPEG 2000) introduced. Such a situation is stipulated for by the inertness of development of software in the sphere, which, in turn, allows to prognosticate the considerable duration of transformation.

Thus, JPEG is considered to be the main factor of influence on the stegocontainer. On the other hand, the stenographic usage of wavelet – transformation allows to hope for obscurity of the entered changes. With the help of the developed criteria, there had been suggested to research the complex connection between the secrecy and robustness of the above usage in the sphere of wavelet – transformation.

Factors had been decided to be modified according to the disseminated approach to vector quantization. The pattern diagram of input is its variability, for which the value of the secret portion of data depends on the correlation of factors set with the specific standard value [6].

The main advantage of pattern diagram is the possibility of multy – variant representation of secrete data portion, which allows to improve their stability.

During the development of modern methods of concealment in the majority of cases there optimizes one of the peculiarities of secrecy or robustness. Using of criteria, which unites the specific qualities, must improve the efficiency of stegoprotection. Aspect of actuality is not covered only by this criteria: there had been suggested the adaptive way of its improvement. To do this, it is necessary to consider the peculiarities of each object of stenographic manipulation, which bares the elementary part of secret data.

It had been supposed that the peculiarities of an approach, suggested in this paper, will ensure the high efficiency of stegomethod on its basis. The development of such a method is the objective of this research.

Criteria of steganographic efficiency. Complex efficiency evaluation of steganomethod is suggested to be executed with the usage of independent factors of secrecy and robustness. Measure of robustness is determined as the part of the saved elementary portions of secret data after stegoimage processing. The stegoanalytical criterium, suggested in [4] had been selected as the criteria of secrecy. It uses SVM for image classification.

The main idea of method for support vectors is to transfer the initial vectors to the space of more high dimension and search for dividing hyper surface with maximal space in this surface (fig.1). Two parallel hyperspaces are built on both sides of the hyperspace, dividing the classes. The dividing hyper surface is that one, which maximizes the distance to two parallel hyper surfaces. The algorithm works on the difference or distance between these average failure of classifier.

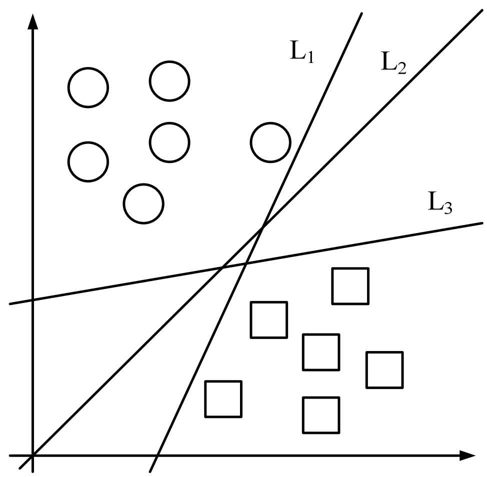


Fig.1. Some classifying lines (hyper surfaces)

Let the points be $\{(x_1,c_1),(x_2,c_2),...,(x_n,c_n)\}$ where c_i accepts the meaning 1 or -1 , depending on the class point x_i , belongs to. Each x_i , - is the p – dimensional vector, usually normalized by values [0,1] or [-1, 1]. If the points are not normalized, than the point with the bigger deviation from the average values of coordinates influences the classifier greatly. Let us consider it as the training collection with the set class for each element. It is necessary for them to be classified by the algorithm of method of support vectors in the same way. To do this, we built the dividing hyper surface which looks like: $w \cdot x - b = 0$

Vector w – perpendicular to the dividing hyper surface. Parameter b depends on the shortest distance of parameter b equals zero, hyper surface goes through the beginning of coordinates, which restricts the decision.

During the optimum division the support vectors and hyper surfaces are parallel to the optimal (fig.2). If is possible to show that these parallel hyper surfaces may be described by the following equations (with the accuracy to the normalization): $w \cdot x - b = 1$, $w \cdot x - b = -1$.

If the learning collection is dividable linearly, this allows to choose the hyper surfaces in such a way, that there should be no single point of learning selection and then maximizing the distance between hyper surfaces. The width of the strip between them is easy to find geometrically, it equals

$\frac{2}{\|w\|}$, in such a way to minimize $\|w\|$.

To exclude all the points from the strip it is necessary for the conditions for all the i to come true:

$$\left[\begin{array}{l} w\cdot x_i - b \ge 1 \\ w\cdot x_i - b \le -1 \end{array}\right.$$

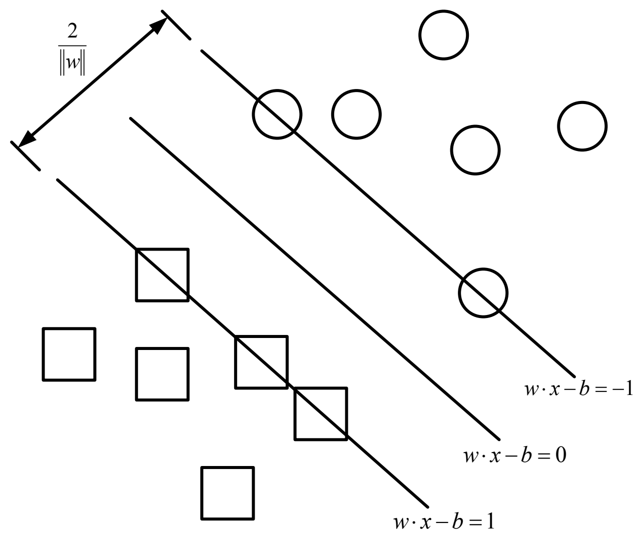


Fig.2. Optimal dividing hyper surface for the method of support vectors, built on points of two classes.

It may also be written as:

$$c_i(w\cdot x_i - b) \ge 1,\ \ 1 \le i \le n \tag{1}$$

In case of linear division, the problem of building the optimal dividing hyper surface is reduced to the minimization of $\|w\|$, on condition of (1). This is the task of quadratic optimization which looks like: $\begin{cases} \|w\|^2 \to \min \\ c_i(w\cdot x_i - b) \ge 1, \quad 1 \le i \le n. \end{cases}$

According the theorem of Kuna – Taker, this task is equivalent to dual task of search of saddle point of Lagrange function.

$$\begin{cases} L(w,b;\lambda) = \frac{1}{2}\|w\|^2 - \sum_{i=1}^{n} \lambda_i c_i (w_i \cdot x_i) \to \min_{u,b} \max_{\lambda} \\ \lambda_i \ge 0, \quad 1 \le i \le n \\ \left[\begin{array}{l} \lambda_i = 0 \\ w\cdot x_i - b = c_i, \end{array}\right. \quad 1 \le i \le n \end{cases} \tag{2}$$

where $\lambda = (\lambda_1, ..., \lambda_n)$ – vector of dual variable.

Let us reduce this task to the equivalents tasks of quadratic programming, which contains the dual variables only.

$$\begin{cases} -L(\lambda) = -\sum_{i=1}^{n}\lambda_i + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\lambda_i\lambda_j(c_i c_j(x_i \cdot x_j - b) - 1) \to \min_\lambda \\ \lambda_i \geq 0, \quad 1 \leq i \leq n \\ \sum_{i=1}^{n}\lambda_i c_i = 0. \end{cases} \tag{3}$$

Lets assume that the above task is solved then w and b may be found accordingly to the formula: $w = \sum_{i=1}^{n}\lambda_i c_i x_i$, $b = w \cdot x_i - c_i$.

Classification algorithm may be written as:

$$a(x) = sign\left(\sum_{i=1}^{n}\lambda_i c_i x_i \cdot x - b\right) \tag{4}$$

We underline that the summing up does not concern the whole sampling, but the support vector, for which $\lambda_i \neq 0$.

For the algorithm to be able to work in case when the classes are linearly non-dividable, we allow it to make failures in the learning collection. We introduce the set of additional variables $\xi_i \geq 0$, characterizing the value of the failure on the objects x_i, $1 \leq i \leq n$. We take (2) as the initial point, soften the restrictions of inequity and also introduce the penalty for the total error into the minimizing functional:

$$\begin{cases} \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n}\xi_i \to \min_{\varpi,b,\xi_i} \\ c_i(w \cdot x_i - b) \geq 1 - \xi_i, \quad 1 \leq i \leq n \\ \xi_i \geq 0, \quad 1 \leq i \leq n. \end{cases}$$

Factor C – parameter of method tuning which allows to regulate the relations between the maximization of the width of the dividing belt and minimization of the total error.

Analogically, following the theorem of Kuna-Taker, we reduce the task to search of Lagrange function.

$$\begin{cases} L(w,b;\xi,\lambda,\eta) = \frac{1}{2}\|w\|^2 - \sum_{i=1}^{n}\lambda_i(c_i(w_i \cdot x_i) - 1) - \sum_{i=1}^{n}\xi_i(\lambda_i + \eta_i - C) \to \min_{u,b,\xi}\max_{\lambda,\eta} \\ \xi_i \geq 0, \lambda_i \geq 0, \eta_i \geq 0, \quad 1 \leq i \leq n \\ \left[\begin{array}{l} \lambda_i = 0 \\ c_i(w \cdot x_i - b) = 1 - \xi_i, \end{array}\right. \quad 1 \leq i \leq n \\ \left[\begin{array}{l} \eta_i = 0 \\ \xi_i = c_i, \end{array}\right. \quad 1 \leq i \leq n. \end{cases}$$

Analogically, we reduce this task to the equivalent one:

$$\begin{cases} -L(\lambda) = -\sum_{i=1}^{n}\lambda_i + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\lambda_i\lambda_j c_i c_j(x_i \cdot x_j) \to \min_\lambda \\ 0 \leq \lambda_i \leq C, \quad 1 \leq i \leq n \\ \sum_{i=1}^{n}\lambda_i c_i = 0. \end{cases}$$

In practice for building the support vector machine it is necessary to solve this task, but not (3) as it appears impossible to guarantee the linear delimitation of dots into two classes in general case.

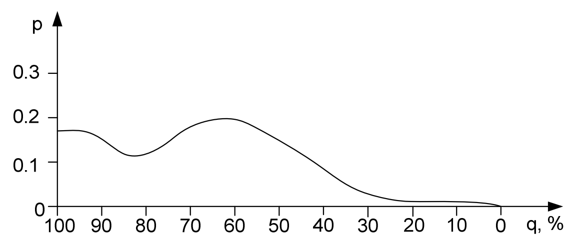
Thus, for each stegoanalytical criteria the connection between the PSNR and entropy of detecting $e^{\det} = -p\log p - \overline{p}\log\overline{p}$ is straight line, where $\overline{p} = 1 - p$, p - is a possibility of correct classification. For the given criteria there had been experimentally fixed the high correlation between these factors. So this connection is considered further as set.

The evaluation of efficiency of embedding stipulates for the construction of consequences of specific influences from part of the third person. When using the JPEG – compression the result depends on compression parameters set by user. Quantization of DKP factors is described by relation:

$$dct_{i,j}^{jpeg} = \frac{Q_{i,j}}{q}\mathrm{round}\left(\frac{dct_{i,j}}{Q_{i,j}}q\right), \tag{5}$$

where $Q_{i,j}$ – corresponding element of quantization matrix Q, $i,j = 1\ldots 8$, q – parameter, set by user, determining quality and size of compressed image [5].

It is of course impossible to foresee the value q in each specific case, but the usage of statistic distribution of f_q allows to proceed to backgrounded evaluation. Fig. 3 reflects the typical distribution f_q.



Fig, 3. Typical distribution of values of parameter q

Since the results of processing JPEG – algorithm (quantization) depends on values of DCT coefficients, the stability of embedded data for different blocks of image will be different. The same concerns the quantative measure of distortion of embedding. During the JPEG compression the image blocks 8×8 are processed independently. So, under condition of independent imbedding to these blocks it is possible to receive the stegosystem which is adaptive to the secrecy requirements and robustness.

Criteria of embedding efficiency must be integral, since instead of q value, the distribution f_q is known. Thus, the possibility f_{q_i} and complex characteristics of systems efficiency z_i corresponds to the definite i th condition of quantification, which is totally determined by q_i. If z_i shall be determined as product of stegoanalytical entropy of detecting $e_i^{\det}$ and the factor of robustness $r_i = 1 - \mathrm{BER}_i$, where BER – index of common errors, the criteria for general efficiency of embedding may be presented be presented by expression :

$$E = \sum_i z_i f_{q_i} = \sum_i e_i^{\det} r_i f_{q_i}\ . \tag{6}$$

Considering that the value of factors $e_i^{\det}$ and r_i are dependable embedding energy $d = \left\| I^{org} - I^{stg} \right\|^2$ (distortion of stegoimage I^{stg} in comparison with original I^{org}), the previous expression looks like:

$$E(d)=\sum_i e_{i,d}^{\det} r_{i,d} f_{q_i} . \tag{7}$$

For the case of continuous changing of quantization conditions we have:

$$E(d)=\int e^{\det}(q,d)r(q,d)f(q)dq . \tag{8}$$

The suggested adaptive approach however requires the additional definition of criterion of embedding efficiency. In respect of above supposition $e^{\det}(q,d)$ is the unique function. For the majority of popular stegomethods it concerns also index of robustness $r(q,d)$. In case of adaptive embedding, arguments (q,d) are not sufficient for adequate representation of robustness level since each of objects of steganographic manipulation may experience polysemantic influence.

That is why the key moment of maximization $E(d)$ will be search $r(q,d,\Omega)$, where $\Omega=\{\Omega_j\}$, $j=1\ldots m$, Ω_j – vector of state of the j - th object. The final task of stegomethod is formalized:

$$\max_d\left(\max_\Omega \int e^{\det}(q,d)r(q,d,\Omega)f(q)dq\right). \tag{9}$$

Apparently, the embedding efficiency will be determined not only by methods of optimization during the solution of the above set task. The way of embedding (diagram), in the first place sets limitations and influences the result [2]. Though the suggested approach may be connected to any diagram, the decision had been made to use the pattern one. This choice is explained by high level of manipulation freedom.

Experiment. The objective of the experiment is to compare the efficiency of embedding data by developed method and methods, which are widely used in practice. The following had been selected for comparison: method of the last significant bit (LSB) [7], pattern method on the basis of integer – valued veivlet – transformation (IWT) [6] and method which operates in the sphere of **DCT** [8]. Secret data were embedded into selected images by stegokey. Methods' efficiency had been determined in respect of two dependences secrecy of stegomanipulations and robustness of embedded data on parameter q, which sets the compression ration. Since the development of method had been conducted on basis of suggested criterion of embedding efficiency, the comparison with other methods in regard of this criterion and mentioned above relationships will allow determining the criterion adequacy.

According to the described peculiarities of designing of stegomethod for the setting and solutions of task of embedding optimization, it is necessary to determine the distribution $f(q)$ and function of stegoanalytical entropy of detecting $e^{\det}(q,d)$. Dependence $f(q)$ is set by expert recognition of popular and wide used images in the grading of gray with the size 256×256, which, depending on the requirements of considered web – pages, had been processed by JPEG algorithm with different value of parameter q. During the determination of $e^{\det}(q,d)$, for each q_i (value q_i changed from 1 to 0.65 with the step 0.05) there had been formatted the learning and text sampling. The first was used for training SVM according to the suggested in [4] vector of characteristics, and on the second there had been determined the average probability of correct detecting depending on values of distortion d. The images in the learning and text sampling do not coincide. Each sampling partially consist of original images (400 in number), the remaining – stegoimages, received from the original with the help of described pattern diagram of embedding.

In the result of the described stages of optimization of embedding 2000 bit of secret data into the Haara wavelet-factors according to the criterion E, the quantitative index of efficiency, wch is an average one for 20 images, made up 0.63. In the described in [6] method on the bases of whole number wavelet-basis, the value of criterion is 0.48, the stoganographic efficiency of the method [7] on the basis of ОБЗ – 0.28? the embedding efficiency in the DCTsphere [8] is avaluated at 0.42.

With the aim of demonstration of adequacy of criterion and efficiency of the developed method, there had been given two graphs of dependences of probabilities of detecting p^{det} of q (fig. 4) and robustness of embedding r of q (fig. 5), which obviously show the advantages and disadvantages of each of the above evaluated method.

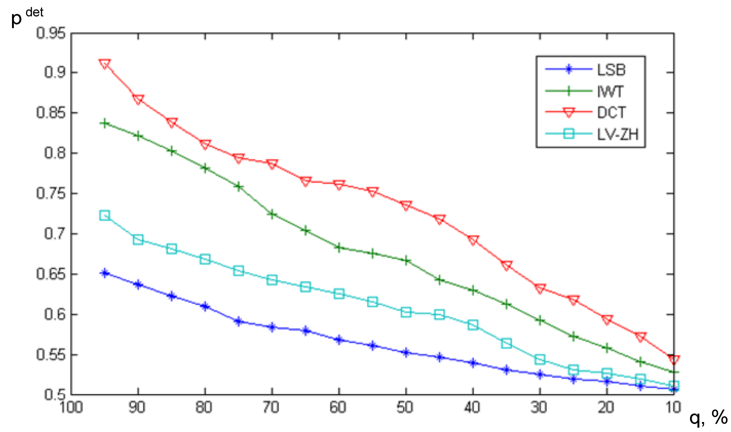


Fig.4. Dependence of probability of detecting p^{det} of the quality parameter JPEG-compression q

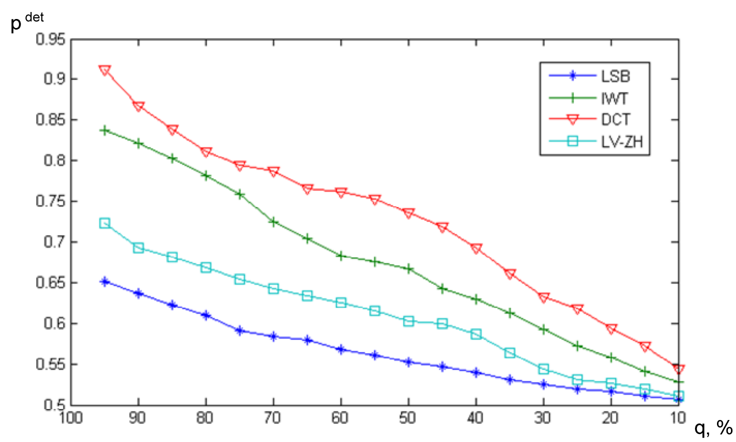


Fig.5. Connection of the robustness r of the embedded data of parameter q

Conclusions.

There had been developed the steganographic method which uses the principle of pattern embedding in the sphere of wavelet-factors. The peculiarity of the method – accounting for requirements to secrecy and robustness to the JPEG-transformation, realized by their unification with the help of the suggested criterion.

The suggested approach allows to improve the general efficiency of data embedding which is experimentally proved in comparison with the popular stegomethods. Disadvantage of this method is a difficulty, which is stipulated for by the differential peculiarity of embedding and necessity of interactive solution of numerical tasks of optimization.

REFERENCES

1. Грибунин В. Г. Цифровая стеганография / В. Г. Грибунин, И. Н. Оков, И. В. Туринцев. – СПб.: Солон-Пресс, 2002. – 272 с.

2. Johnson N. F. Information Hiding: Steganography and Watermarking – Attacks and Countermeasures / N. F. Johnson, Z. Duric, S. Jajodia. – Berlin: Springer, 2001. – 160 p.

3. Glavieux A. Channel Coding in Communication Networks / A. Glavieux. – London: Hermes Science Pub. Ltd., 2007. – 416 p.

4. Zou D. Steganalysis based on Markov Model of Thresholded Prediction-Error Image / D. Zou, Y. Shi, W.Su, G. Xuan // IEEE ICME. – 2006. – № 1. – P. 1365 − 1368.

5. Pennebaker W. JPEG: Still Image Compression Standard / W. Pennebaker, J. Mitchell. – NY.: Kluwer Academic Pub., 1993. – 650 p.

6. Метод вбудовування даних на основі алгоритму вейвлет-стиснення зображень: праці конференції, 25 − 28 вер. 2006 р., Вінниця. Т. 1 / Відп. ред. В. М. Дубовой. – Вінниця: Универсум-Вінниця, 2007. – С. 491 − 495.

7. Wu H.C. Image Steganographic Scheme Based on Pixel-Value Differencing and LSB Replacement Methods / H.C. Wu, N.I. Wu, C.S. Tsai, M.S. Hwang // IEEE Transactions on Image and Signal Processing. – 2005. – № 5. – P. 611 − 615.

8. Quan L. Combination of DCT-Based and SVD-Based Watermarking Scheme / L. Quan, A. Qingsong // IEEE ICSP Conference Record. – 2004. – № 1. – P. 873 − 876.

Anatoliy Vasiura – Director of the Institute, Professor with the department for automatics and information and measuring equipment;

Vitaliy Lukichov **–** Researcher with the department for automatics and information and measuring equipment.

Vinnytsia National Technical University.