Mohylina Yu.V.
Student;
Bohaichuk V.O.
Student,
Igor Sikorsky Kyiv Polytechnic Institute

STEGO IMAGES DESTRUCTION VIA COMPONENT ANALYSIS METHODS

The spread of global and high-speed access to the Internet, the emergence of social networks aimed at massive multimedia data including digital photos, and world globalization greatly simplify the conditions for creating covert channels of communication between hackers.

To detect hidden messages were used methods of processing images in the spatial domain (median filtration, Gaussian filtration, Wiener filtration, gamma correction) and transform domain (JPEG compression, wavelet compression). However, their practical application limited by the high degree of distortion in images.

To solve this problem in this paper were proposed methods of component analysis such as the principal component analysis and the independent component analysis.

The principal component analysis (PCA) method is a statistical approach first proposed by Carl Pearson in 1901 [1]. This method applies an orthogonal linear transform to a set of inputs that could potentially be redundant and dependent on one another. This conversion translates the multivariate input signals into a new coordinate system for which the largest variance of data would be directed along the first coordinate, the second along the second, etc.

The conversion is set in such a way that all major components of the input signal (vectors along which the data have the largest variance) (principal component – PC) are sorted in descending variance, that is, the first components store the most information about the input and orthogonal to each other. This method was used in many fields, including data visualization [2], audio and video compression [3], and dimensional reduction of dynamic models, bioinformatics [4] and others. Principal component method allows increasing the degree of data compression compared to common spectral transformations, in particular discrete cosine transformations (DCT) and even Two-dimensional discrete wavelet transforms (TDDT) [5]. However, the limitation of the practical application of this method is its high computational complexity.

The independent component analysis (ICA) method used in signal processing for blind source separation (BSS). Components are statistically independent non-Gaussian signals. The simplest example of ICA application is the «party problem» [6], when you can hear and distinguish the essence of each person's conversation individually in one room. Most often, this problem simplified by the absence of delays and reflection. It is also important that the number of signal receivers (microphones) is not less than the number of signal sources. In general, the ICA cannot determine the true number of signal sources, the correct order of the sources, or the scale of the signals among themselves.

One of the most common adaptive stego algorithms is the UNIWARD family of methods. The peculiarity of these methods is the use of the function of estimation of the distortion of the images (1), which is the sum of the corresponding changes between the stego the cover-image in the veil region. In this paper was used SI-UNIWARD method for data hidden.

$$D(X,Y) = \sum_{k=1}^{3} \sum_{u=1}^{n1} \sum_{v=1}^{n2} \frac{\left|w_{uv}^{k}(X) - w_{uv}^{k}(Y)\right|}{\sigma + \left|w_{uv}^{k}(X)\right|} \quad (1)$$

Where X, Y are respectively the original (unfilled) image and the stego with dimensions $n \times n$ pixels; σ is the constant used to stabilize the calculations; W (X), (Y) are the detail coefficients of the TDDT of the original image and the stego.

The study was conducted using the standard digital imaging package MIRFlickr-1M [7]. Test package of 1000 images (1000 images for each degree of filling of the cover-image was filled up) was used. The degree of filling of the SC with steganodata varied from 5 to 30% in step 5 and from 30 to 90% in step 10. To evaluate the efficiency of the destruction methods, we used the metric for estimating the number of pixels of the stego (2) that remained unchanged after filtering or compressing the cover image with losses.

$$\frac{|(C-S)\&(C-S^{*})|}{|C-S|} \quad (2)$$

Where C is the original image, S is the image with built-in stego, S ^ * is the quilting cover after the destruction, & – bit addition.

The calculations were made using software modules created in the Python programming language.

While embedding information in the domain of transformation by the SI-UNIWARD method, it is seen that the methods of component analysis can significantly improve the degree of destruction of the constipated. The principal component method of 20% (75% for PCA vs. 95.5% for median filtration) eliminates more hidden information in the filled cover image with a fill rate of 5% and 15% (75.5% for PCA vs. 90% for median filtration) for the fill rate SC is 90% steady.

With similar characteristics, the method of independent components was better by 19.5% and 14.5%. It should be noted that with increasing the degree of filling of the SC steady-state for the median filter decreases the relative proportion of filtered stegobites: for MFs with 3 × 3 window with the degree of filling the SC steady-state 5% the proportion of stegobites that have not changed is 95%, and for the degree of filling 90 % surviving 90% of pixels. The degree of change in the stegobite fraction, which did not change from the degree of filling of the stego with the steady state for the component methods, is not significant.

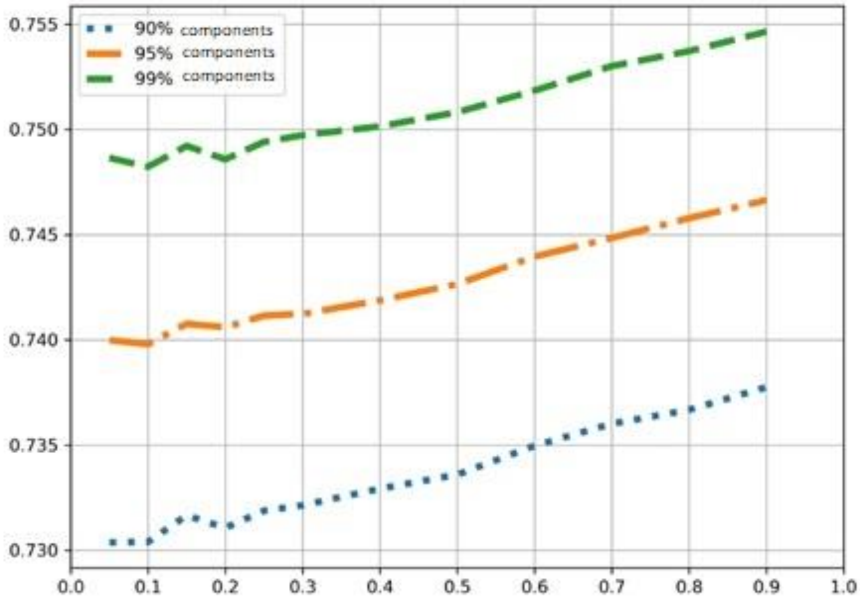

Figure 1. Dependence between number of unchanged pixels (y) and payload of cover images with stego data (x)

The results of the analysis show that the use of the proposed methods can improve the quality of active steganalysis. Thus, the use of principal and independent component methods increases the degree of reliability of the destruction of hidden information by an average of 20%.

It is worth noting that the proposed methods of destruction showed similar results regardless of the methods of information concealment: about 75% of stegobites survived the destruction by component analysis methods using 99% of components for the restoration of the IC, and about 72% of the stegobits survived by the use of 90% of the components for the restoration of the cover images.

References:

1. K. Pearson, «LIII. On lines and planes of closest fit to systems of points in space». The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, т. 2, № 11, pp. 559-572, 1901.

2. Зиновьев А. Ю. Визуализация многомерных данних. Красноярск: Изд. КГТУ, 2000.

3. K. Rao та P. Yip, The Transform and Data Compression Handbook, Baton Rouge: CRC Press, 2001.

4. J.-P. Benzécri, L'Analyse des Données, Paris: Dunod, 1973.

5. D. D. Muresan та T. W. Parks, «ADAPTIVE PRINCIPAL COMPONENTS AND IMAGE DENOISING,» в International Conference on Image Processing 2003, Barcelona, Spain, 2003.

6. S. Getzmann, J. Jasny та M. Falkenstein, «Switching of auditory attention in «cocktail-party" listening: ERP evidence of cueing effects in younger and older adults,» в Leibniz Research Centre for Working Environment and Human Factors, Dortmund, Germany, 2016.

7. M. S. L. M. J. Huiskes, «The MIR Flickr Retrieval Evaluation,» в ACM International Conference on Multimedia Information Retrieval (MIR'08), Vancouver, Canada, 2008.