STATISTICAL STEGANALYSIS OF MULTISTAGE EMBEDDING METHODS

Dmytro Progonov

Abstract: *The paper is devoted to comparative analysis of performance the modern methods of statistical steganalysis in case of message hiding in digital images with usage of multidomain and multistage embedding methods. It is considered the case of applying of statistical models of cover images in spatial (SPAM model) and frequency (CC-PEV model) domains, as well as universal CDF model for revealing the stego image with messages, embedded with usage of standard (Discrete Cosine and Wavelet Transforms) and special (Singular Value Decomposition) transforms of cover images and stegodata. It is shown that applying of modern methods of statistical steganalysis allows reliably revealing stego images, formed according to multistage embedding methods. Usage special transforms of cover images, for instance Singular Value Decomposition, allows significantly decrease the performance of statistical steganalysis, which requires development of new statistical models of cover images for steganogram detection.*

Keywords: *statistical steganalysis, multidomain embedding methods, multistage embedding methods, digital images.*

ACM Classification Keywords: *D.4 Operating Systems – Security and protection – Information flow controls.*

Introduction

Today the businesses as well as governments are widely using the information warfare methods and cyber weapons for achieving the competitive advantages in economic, political and military spheres. In most cases the cyber weapons are used for gaining the remote control and/or destruction of adversary's critical infrastructure (ACI) – assets that are essential for functioning of a society and economics, e.g. water supply, electricity generation, transportation systems, public health.

Successful attack on ACI requires usage of protected communication channels, created with applying of cryptographic algorithms, for data transmission between intelligence agencies, spies and bot-nets of infected computers. Due to juristic limitations on usage the cryptographic methods for creation the private protected channels in most countries, it is widely used the specialized communication systems, based on applying of steganographic methods. Peculiarity of steganography-based communication systems (SCS) is embedding of communicational channel into the existed information flows in telecommunication systems (TCS), such as e-mail services, social networks, chats etc. It allows

breaking the existed systems of traffic control like firewalls, deep packet inspection systems etc. Therefore, creation of methods for reliably revealing and destruction of embedded messages (steganograms) is crucial and unresolved issue of the day.

As cover files for message hiding in SCS there can be used the various types of digital data – for instance texts, multimedia data, and service attributes of files etc. Applying of specified types of cover files is highly depends on prescribed requirements of robustness the obtained steganograms to known methods of passive and active analysis, as well as channel capacity (volume of data embedded per each cover file) [Fridrich, 2010]. For obtaining the trade-off between mentioned requirements, in most cases message hiding is provided into multimedia files, especially digital images (DI).

Existed methods of message hiding in DI can be divided into four groups [Katzenbeisser, 2000] [Fridrich, 2010] – via mimicking natural processing of digital images (e.g. stochastic modulation of pixel's brightness, dither quantizers of pixel's brightness), steganalysis-aware methods (for instance, HUGO algorithm, UNIWARD methods), message hiding embedding in spatial domain (LSB-methods) and transformation domain (TD). Considerable disadvantage of practical usage of first three types of steganographic methods is relatively high "sensitivity" to any alteration of steganogram during its storing, processing and transmission in TCS. As a consequence, significant amount of modern SCS is based on stegodata embedding in TD, which gives opportunity to achieve the trade-off between robustness of obtained steganograms to known methods of passive as well as active steganalysis.

For revealing the stego images, formed according to mentioned methods there were developed great amount of passive steganalysis methods. Proposed methods can be divided into three groups – signature [Gribunin, 2002] [Agranovskyi, 2009], statistical [Fridrich, 2010] [Katzenbeisser, 2000] and structural [Progonov, 2014] [Progonov, 2015a] steganalysis. Signature steganalysis is widely used for preliminary investigation of digital images and based on revealing the stego images by detection the distinctive changes (signatures) of cover image attributes or parameters, for instance service attributes, parameters of used graphical format etc. Methods of structural steganalysis are based on usage of methods for hierarchical statistical modeling and multifractal analysis for revealing the alteration the correlation and fractal parameters of digital images, caused by message hiding.

One of the most widespread approaches for revealing the stego images is statistical steganalysis (SS). Methods of SS is based on creation the statistical models (SM) of covers with usage of peculiarities of digital images, for instance significant correlation of brightness the adjacent pixels. Known methods of statistical steganalysis allows revealing with high accuracy the the stego images in case of message hiding in spatial or frequency domain (for instance with usage of two-dimensional discrete cosine or wavelet transforms).

For increase the robustness of stego images to known methods of statistical steganalysis there are proposed to use the special transforms (e.g. Singular Vector Decomposition, SVD) or composition of several transforms the cover images and stegodata (multistage methods) for message hiding. Creation of effective methods for revealing the mentioned steganographic methods requires analysis of

performance the known statistical models. Obtained results of investigation the performance of modern methods of SS in case of forming the stego images with usage of multidomain and multistage methods will be used for creation the improved SM.

Related Works

Known statistical models of digital images can be divided into two groups – simple and rich models. Simple models allow investigating only separate parameters of digital images such as histogram (chi-square test, Pairs-of-Value analysis), co-occurrence matrix of pixel's brightness (Sample-Pairs analysis) or changes the parameters of distribution the pixel's brightness, caused by applying the calibration function to the images (Regular-and-Singular analysis). Simple SM gives opportunity to detect with high accuracy only the known LSB-methods (±1 embedding) or JPEG-based embedding methods (JSteg, F5, OutGuess algorithms) [Fridrich, 2010]. For overcome the mentioned limitation it was proposed to use the rich SM (RSM), obtained by consolidation of several simple statistical models [Fridrich, 2012]. RSM allow noticeably increasing the accuracy of stego image detection in most complicated cases – with usage of modern adaptive steganographic methods, such as HUGO algorithm [Pevny, 2010b], UNIWARD algorithm's families [Holub, 2013a], Synch algorithms [Denemark, 2015] etc.

Creation of effective methods for revealing the stego images with data, embedded in spatial (LSB-methods) or frequency (JPEG-based steganography) domains of cover images, requires development of corresponding RSM only for mentioned domains. It is complicated task, required the deep analysis of modern methods the DI analysis, such as wavelet transform with anisotropic basis functions (e.g. ridgelets, curvelets, bandlets), sparse and redundant dictionaries. For provide the high accuracy of stego image detection regardless of embedding domain there were proposed the universal RSM, based on consolidation of known statistical models of DI in spatial as well as frequency domains.

Development of new methods for message hiding in several domains (multidomain methods [Progonov, 2015b]) and multistage methods, based on composition of several transforms the cover image and stegodata, significantly complicates the detection of formed stego images. Unfortunately in the literature there is no information about the performance of modern statistical stegdetector in case of message hiding in several domains according to multistage methods. Therefore it is represented the interest investigation of performance the modern RSM for detection the stego images, formed with usage of special transforms of cover image, and multistage methods.

The Goal and Contribution

The goal of paper is comparative analysis of accuracy the stego images detection, in case of stegodata embedding in transformation domain of digital images according to multidomain and multistage methods, with usage of modern methods of statistical steganalysis.

Modern Methods of Data Embedding in Transformation Domain of Digital Images

One of the well-known methods for message hiding in transformation domains of digital images are one-stage methods of Dey [Dey, 2011] and Agarwal [Agarwal, 2008], multistage embedding methods of Joseph [Joseph, 2013] and Khan [Khan, 2013], as well as complex methods of Elahian [Elahian, 2011] and Gunjal [Gunjal, 2011]. High robustness of stego images, formed according to one-stage methods, to active steganalysis (e.g. image filtering, lossy compression) is provided by applying to cover image and stegodata, represented as DI, the two-dimensional discrete wavelet transform (DWT) and SVD. Usage of several stage of cover image processing according to multistage methods allows achieving the trade-off between robustness of formed stego images to passive steganalysis and influence of sporadic (channel noise) or intentional (active steganalysis) changes by theirs storage, processing and transmissions in TCS. Peculiarity of complex embedding methods is presence of preliminary stage of cover image (for instance, change the color system) and embedding messages (e.g. encoding) processing for increasing the robustness of formed stego images to passive steganalysis.

Forming of stego images according to mentioned steganographic methods is provided by weighted summation of coefficients the cover image $W(\mathbf{I})$ and stegodata $W(\mathbf{D})$, represented as DI, decomposition in fixed basis of transformation $W(\bullet)$:

$$W(\mathbf{S}) = W(\mathbf{I}) + G \times W(\mathbf{D}),$$

where $W(\mathbf{S})$ – coefficients of formed stego image, G – weighted coefficients, which is depends on energy the hidden message. Processing of separate color channel the cover image and stegodata is provided with usage of standard and special transforms – two-dimensional discrete cosine [Oppenheim, 2010] [Gonsalez, 2008] and wavelet [Gonsalez, 2008] transforms (2D-DCT and 2D-DWT), singular value decomposition [Murphy, 2012].

For obtaining the stego image in spatial domain, it is applied the inverse transform to obtained coefficients $W^{-1}(W(\mathbf{S}))$. Extraction of embedded message is provided according to further formula:

$$W(\mathbf{D}) = (W(\mathbf{S}) - W(\mathbf{I}))/G.$$

For achieving the trade-off between robustness of hidden message to passive ($G \to G_{\min}$) and active ($G \to G_{\max}$) steganalysis there were determined the range of values the weighted coefficient G. Values of G were changed from $G_{\min}$ (lower bound of stegodata reconstruction on receiver's side the SCS) to $G_{\max}$ (appearance the visual distortion of cover image by message hiding) with step Δ_G.

Shaping of stego images $\mathbf{S}_{M\times N}$ according to Agarwal methods is provided by applying the SVD to cover image $\mathbf{I}_{M\times N}$ and stegodata $\mathbf{D}_{M\times N}$, represented as grayscale images with resolution $M \times N$ pixels [Agarwal, 2008]:

$$\mathbf{I}_{M\times N} = \mathbf{U}_{M\times M}(\mathbf{I}) \times \mathbf{Q}_{M\times N}(\mathbf{I}) \times \mathbf{V}^T_{N\times N}(\mathbf{I}),$$

$$\mathbf{D}_{M\times N} = \mathbf{U}_{M\times M}(\mathbf{D}) \times \mathbf{Q}_{M\times N}(\mathbf{D}) \times \mathbf{V}^{T}_{N\times N}(\mathbf{D}),$$

$$\mathbf{V}(\mathbf{S}) = \mathbf{V}(\mathbf{I}) + G \times \mathbf{V}(\mathbf{D}),$$

where $\mathbf{U}_{M\times M}(\mathbf{I}), \mathbf{V}_{N\times N}(\mathbf{I})$ – correspondingly, matrices of left-singular and right-singular vectors of matrix $\mathbf{I}\times\mathbf{I}^{T}$ ($\mathbf{I}^{T}\times\mathbf{I}$); $\mathbf{Q}_{M\times N}(\mathbf{I}) = \mathbf{E}_{M\times N} \times \mathbf{\Lambda}(\mathbf{I})$ – diagonal matrix of singular values; $\mathbf{E}_{M\times N}$ – unit matrix; $\mathbf{\Lambda}(\mathbf{I}) = \left\{\sqrt{\lambda_1}, \sqrt{\lambda_2}, \ldots, \sqrt{\lambda_K}\right\}$ – vector of singular values; $K = \min(M,N)$ – number of singular values.

Simultaneously change of matrices $\mathbf{U}_{M\times M}(\mathbf{I})$ and $\mathbf{V}_{N\times N}(\mathbf{I})$ – shifts of rows $\mathbf{u}_{i\times M}(\mathbf{I}), i \in [1;M]$ and columns $\mathbf{v}_{N\times j}(\mathbf{I}), j \in [1;N]$ on Δ_x positions – does not change the stego image [Bolshakov, 2007], which leads to –ambiguity by message extraction at the receiver side of SCS. For solve the mentioned problem it is proposed to use the singular values of cover image for message hiding:

$$\mathbf{\Lambda}_{K\times 1}(\mathbf{S}) = \mathbf{\Lambda}_{K\times 1}(\mathbf{I}) + G \times \mathbf{\Lambda}_{K\times 1}(\mathbf{D}), \quad \mathbf{\Lambda}_{K\times 1}(\mathbf{D}) = \left(\mathbf{\Lambda}_{K\times 1}(\mathbf{S}) - \mathbf{\Lambda}_{K\times 1}(\mathbf{I})\right)/G.$$

For minimization the color alteration by message hiding according to complex methods of Elahian [Elahian, 2011] and Gunjal [Gunjal, 2011] it is provided the changes of color system the cover image (from RGB to YCbCr or YIQ) as well as stegodata (from RGB to Grayscale). Stegodata are embedded into Y (luma) component (Elahian method) or I (chroma) component (Gunjal method) of cover image. Changes of color system for cover image an stegodata are provided according to standard formulae [Gonsalez, 2008] [Gunjal, 2011].

Also, for increasing of robustness the formed stego images to passive steganalysis it is applied the Arnold mapping (AM) to the embedding messages $\mathbf{D}_{M_D\times N_D}$, represented as grayscale image with resolution $M_D \times N_D \left(M_D = N_D\right)$ [Elahian, 2011]:

$$\begin{pmatrix} x_{i+1} \\ y_{i+1} \end{pmatrix} = \mathrm{mod}\left(\mathbf{R}_{2\times 2} \times \begin{pmatrix} x_i \\ y_i \end{pmatrix}, M_D\right), \mathbf{R}_{2\times 2} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix},$$

where x, y – correspondingly, position of elements the $\mathbf{D}$ (number of row/column); $\mathrm{mod}(a,b)$ – modulo operation; i – current iteration of Arnold mapping. It should be mentioned that period of AM – number of iteration after which the initial view of stegodata is recovered – significantly depends on stegodata size, which complicates the further passive steganalysis of formed stego images.

Stages of processing the cover image and stegodata according to one-stage, multistage and complex methods are represented in Table 1.

Table 1. Stage of cover image and stegodata processing according to one-stage, multistage and complex steganographic methods

		Embedding method					
		Dey	Agarwal	Joseph	Khan	Elahian	Gunjal
Color system change	Cover	RGB → RGB				RGB→YCbCr	RGB→YIQ
	Stegodata	RGB → RGB				RGB → Grayscale	
Stage of cover processing	First	2D-DWT	SVD	2D-DWT	2D-DWT	2D-DWT	2D-DWT
	Second	–	–	SVD	2D-DCT	–	2D-DCT
	Third	–	–	–	SVD	–	–
Stage of stegodata processing	First	2D-DWT	SVD	SVD	SVD	AM	AM
	Second	–	–	–	–	2D-DWT	–
Weighted coefficient G	G_{max}	0.02	0.02	0.10	0.50	1.00	5.00
	G_{min}	0.08	0.08	2.00	4.00	12.00	14.00
	Δ_G	0.02	0.02	1.00	1.00	3.00	3.00

Statistical Steganalysis of Digital Images

Rich statistical models (RSM) can be divided into three groups, depending on domain where digital images are modelling – in spatial (SPAM [Pevny, 2010a], SRM [Fridrich, 2012], PSRM [Holub, 2013b] models) or frequency (CC-PEV [Pevny, 2007], CC-JRM [Kodovský, 2012b] models) domains, as well universal models (CDF [Kodovsky, 2010], J+SRM [Kodovský, 2012b] models), obtained by consolidation of SM the digital images in spatial and frequency domains. In the article we investigated the performance of well-known RSM the digital images in spatial (SPAM model) and frequency (CC-PEV model) domain as well as universal (CDF model) statistical models.

Subtractive Pixel Adjacency Matrix (SPAM) model of digital images is based on modeling the differences between brightness of adjacent pixels with usage of Markov chains. Estimation of parameters the Markov chains for horizontally adjacent pixels the grayscale image $\mathbf{I}_{x,y}$ with resolution $M \times N$ pixels is provided with usage of co-occurrences matrices $\mathbf{C}_{x,y}$ [Pevny, 2010a]:

$$\mathbf{C}^{\rightarrow}(u,v)=\sum_{i=1}^{M}\sum_{j=1}^{N-1}\left(\left[\mathbf{D}_{i,j}^{\rightarrow}=u\right]_I\times\left[\mathbf{D}_{i,j+1}^{\rightarrow}=v\right]_I\right),$$

$$\mathbf{C}^{\rightarrow}(u,v,w)=\sum_{i=1}^{M}\sum_{j=1}^{N-2}\left(\left[\mathbf{D}_{i,j}^{\rightarrow}=u\right]_I\times\left[\mathbf{D}_{i,j+1}^{\rightarrow}=v\right]_I\times\left[\mathbf{D}_{i,j+2}^{\rightarrow}=w\right]_I\right),$$

where

$$\mathbf{D}_{i,j}^{\rightarrow}=\mathbf{I}_{i,j}-\mathbf{I}_{i,j+1},i\in[1;M],j\in[1;N-1],$$

is differences arrays; $[\bullet]_I$ – Iverson bracket; $u,v,w\left(u,v,w\in[-T;T]\right)$ – values of Markov chain elements (differences of brightness the adjacency pixels); T – threshold value, which is used for achieving the trade-off between the accuracy of modeling the differences between brightness of adjacent pixels and number of SPAM model parameters.

Transition probability matrices for Markov chains first ($\mathbf{M}_{u,v}$) and second ($\mathbf{M}_{u,v,w}$) orders are calculated according to further formulae:

$$\mathbf{M}_{u,v}^{k}=\frac{\mathbf{C}^{k}(u,v)}{\sum_{u=(-T)}^{T}\sum_{v=(-T)}^{T}\mathbf{C}^{k}(u,v)},k\in\{\rightarrow,\leftarrow,\uparrow,\downarrow,\searrow,\nwarrow,\nearrow,\swarrow\},$$

$$\mathbf{M}_{u,v,w}^{k}=\frac{\mathbf{C}^{k}(u,v,w)}{\sum_{u=(-T)}^{T}\sum_{v=(-T)}^{T}\sum_{w=(-T)}^{T}\mathbf{C}^{k}(u,v,w)},k\in\{\rightarrow,\leftarrow,\uparrow,\downarrow,\searrow,\nwarrow,\nearrow,\swarrow\}.$$

As parameters of SPAM model is used the averaged transition probability matrices for Markov chains 1st ($\mathbf{F}^{I}$) and 2nd ($\mathbf{F}^{II}$) orders:

$$\mathbf{F}^{I}=\left(\mathbf{M}_{u,v}^{\rightarrow}+\mathbf{M}_{u,v}^{\leftarrow}+\mathbf{M}_{u,v}^{\uparrow}+\mathbf{M}_{u,v}^{\downarrow}+\mathbf{M}_{u,v}^{\searrow}+\mathbf{M}_{u,v}^{\nwarrow}+\mathbf{M}_{u,v}^{\swarrow}+\mathbf{M}_{u,v}^{\nearrow}\right)\Big/8,$$

$$\mathbf{F}^{II}=\left(\mathbf{M}_{u,v,w}^{\rightarrow}+\mathbf{M}_{u,v,w}^{\leftarrow}+\mathbf{M}_{u,v,w}^{\uparrow}+\mathbf{M}_{u,v,w}^{\downarrow}+\mathbf{M}_{u,v,w}^{\searrow}+\mathbf{M}_{u,v,w}^{\nwarrow}+\mathbf{M}_{u,v,w}^{\swarrow}+\mathbf{M}_{u,v,w}^{\nearrow}\right)\Big/8.$$

According to recommendation [Pevny, 2010a] threshold value was chosen equal to $T=3$. Total amount of SPAM model parameters in such case is equal to $d_{SPAM}=686$.

The CC-PEV model was proposed for reliably detection of stego images with message, hidden in frequency domain according to JPEG-based embedding methods [Pevny, 2007]. Peculiarity of CC-PEV model is preliminary stage of digital image calibration, used for suppress the distortion of successive JPEG-compression with different quality factor (quantization tables). Calibration of DI is provided by its decompression in spatial domain, cropping to four rows/columns and further JPEG-compression with initial JPEG Quality Factor (JQF).

At the second stage, initial $\mathbf{I}_{x,y}$ and calibrated $\mathbf{I}_{x,y}^{C}$ images are divided into n_l non-overlapping blocks with size 8×8 pixels. Then to each block the two-dimensional discrete cosine transform (DCT) is applied. At the third stage, further correlation parameters for obtained coefficients the discrete cosine transform inter and intra blocks $d_{ij}(k),k\in[1;n_l]$ [Pevny, 2007] are calculated:

1. Histogram $\mathbf{H}$ of all $64\times n_l$ luminance DCT coefficients;

2. Histograms $\mathbf{h}^{ij}$ of coefficients of 5 individual DCT modes $(i,j) \in \{(1,2),(2,1),(3,1),(1,3),(2,2)\}$;

3. Dual histograms $\mathbf{g}_{ij}^{d}$:

$$\mathbf{g}_{ij}^{d} = \sum_{k=1}^{n_l} \delta\left(d, d_{ij}(k)\right),$$

where $\delta(\bullet,\bullet)$ – Kroneker delta;

4. Functionals V, captured inter-block dependency among DCT coefficients;

5. Blockiness fumctional B_{α}, which is calculated from the decompressed JPEG-image and represented an integral measure of inter-block dependency over all DCT modes over the whole image;

6. Co-occurrence matrix $\mathbf{N}$ of neighboring DCT coefficients;

7. Averaged transition probability matrices $\mathbf{M}$ of Markov chain first order, which are used for modeling the differences between adjacency pixels in horizontal, vertical and diagonal directions.

The parameters of CC-PEV models are obtained by calculating the differences of mentioned parameters for initial and calibrated images. Total number of CC-PEV model's parameters is equal to $d_{CC-PEV} = 548$.

The universal statistical model CDF was proposed in [Kodovský, 2010] by consolidation of SPAM and CC-PEV model for detecting the stego images in case of stegodata embedding in spatial as well as frequency domains. Total number of CDF model's parameters is equal to $d_{CDF} = 1234$.

Results

For analysis the accuracy of stego image detection there were trained and tested the stegdetectors (SD), based on usage the statistical models of digital images in spatial (SPAM model, SD_{SPAM}) and frequency (CC-PEV model, SD_{CC-PEV}) domains as well as universal CDF model (SD_{CDF}). Due to great number of parameters for used RSM – $d_{SPAM} = 686$, $d_{CC-PEV} = 548$, $d_{CDF} = 1234$ – as stegdetector it was used the ensemble of Fisher's Linear Discriminants (FLD) [Kodovský, 2012a]. Separate FLD was tuned for minimization of total detection error P_E on training subset the test packet:

$$P_E = \min_{P_{FA}} \frac{1}{2}\left[P_{FA} + P_{MD}(P_{FA})\right],$$

where P_{FA}, P_{MD} denote, correspondingly, the probabilities of false alarm and missed detection. Assessments of P_{FA} and P_{MD} were provided according to bootstrap estimation algorithm by training each base classifier B_l on pseudo random selected subset of training set [Kodovský, 2012a]:

$$\mathrm{X}_l = \left\{x_m^{(D_l)}, \bar{x}_m^{(D_l)}\right\}_{m \in \mathfrak{M}_l^b},$$

where $x_m, \overline{x}_m$ – training samples of cover and stego images respectively; $D_l, l \in \{1,2...d_{FLD}\}$ – pseudo randomly selected subset of features from general feature space with dimensionality d_{ALL} ($d_{FLD} \ll d_{ALL}$); $\mathfrak{M}_l^b$ – bootstrap sample from set $\{1,2...N^{trn}\}$; N^{trn} – amount of test cover images at training stage.

The total detection error P_E (out-of-bag (OOB) error) for SD after training phase was computed according to formula:

$$P_{E-OOB}^{(n)} = \frac{1}{2N^{trn}} \sum_{m=1}^{N^{trn}} \left[B^{(n)}(x_m) + 1 - B^{(n)}(\overline{x}_m) \right].$$

Analysis of accuracy the stego image detection by usage of statistical stegdetectors was provided for two cases – with utilization of all or separate stegodata at the training/testing stage, as well as usage of true color or grayscale (separate color channels) the test digital images. As indices for analysis the accuracy the stego image detection there were used the standard metrics from ROC-analysis [Murphy, 2012] [Mathews, 1975] – Area-Under-ROC curve (AUC), Sensitivity, Specifity, Matthews Correlation Coefficient. Estimation of mean value and variance of the OOB-error P_E and mentioned metrics was provided by repeating the training and testing stage 10 times.

Investigation of accuracy the steganogram detection by usage of modern RSM was provided on standard image database MIRFlickr-25k [Huiskes, 2008]. For training and testing of stegodetector were used the subset of 9,000 pseudo randomly selected and scaled DI from packet. Cardinalities of training and testing set of digital images were equal to 4,500 images. As stegodata were used three DI – engine's draft, map and portrait. Characteristics of the stegodata are represented in Table 2:

Cover image payload – fraction of changed coefficients of cover image $W(\mathbf{I})$ relatively whole number of coefficients –was changed from 5% to 25% with step 5% and from 25% to 95% with step 10%. Weighted coefficient G, for each investigated embedding method, was changed from $G_{\min}$ up to $G_{\max}$ with step Δ_G (Table 1).

Table 2. Characteristics of used test digital images and stegodata

Characteristics	Cover image	Stegodata		
		Engine's draft	Map	Portrait
Resolution, pixels	512×512	567×463	800×800	565×850
Color system	RGB			
Format	JPEG, TrueColor	BMP		

At the Figure 1-3 it is represented the dependency of AUC metrics on cover image payloads by variation the weighted coefficients G for statistical stegdetector SD_{SPAM}, SD_{CC-PEV} and SD_{CDF}.

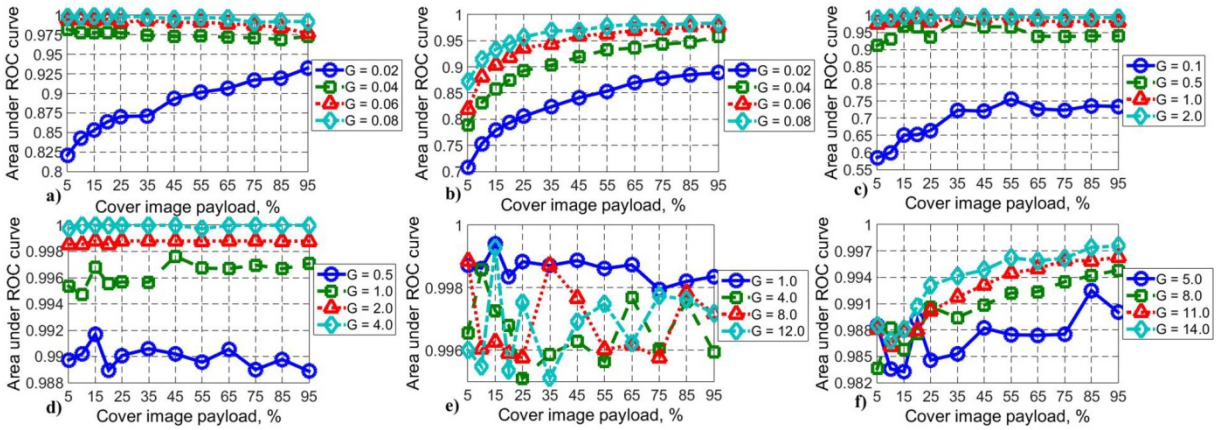

Figure 1. Dependency of AUC metrics on cover image payloads by variation the weighted coefficients G for statistical stegdetector SD_{SPAM}. Message was embedded according to: (a) – Dey method; (b) – Agarwal method; (c) – Joseph method; (d) – Khan method; (e) – Elahian method; (f) – Gunjal method.

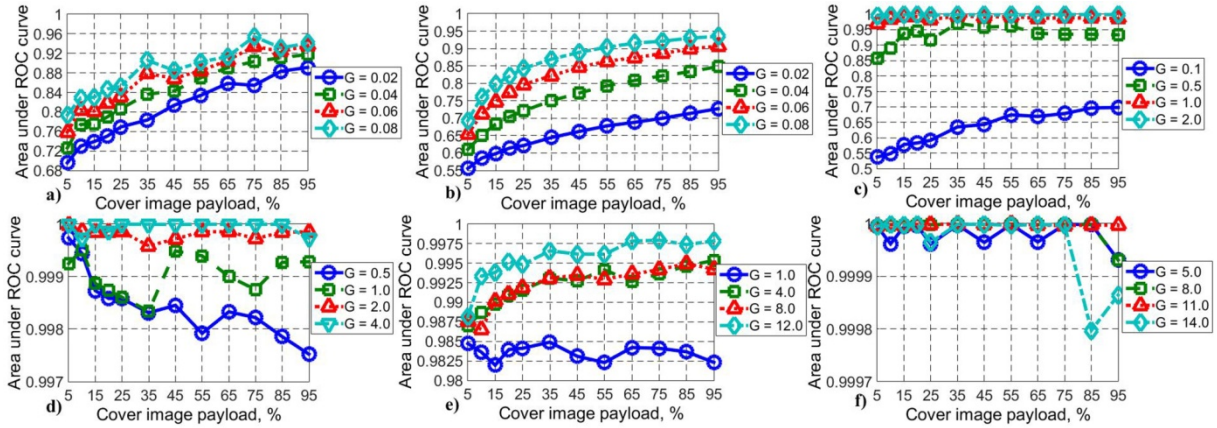

Figure 2. Dependency of AUC metrics on cover image payloads by variation the weighted coefficients G for statistical stegdetector SD_{CC-PEV} (JPEG Quality Factor – 100). Message was embedded according to: (a) – Dey method; (b) – Agarwal method; (c) – Joseph method; (d) – Khan method; (e) – Elahian method; (f) – Gunjal method.

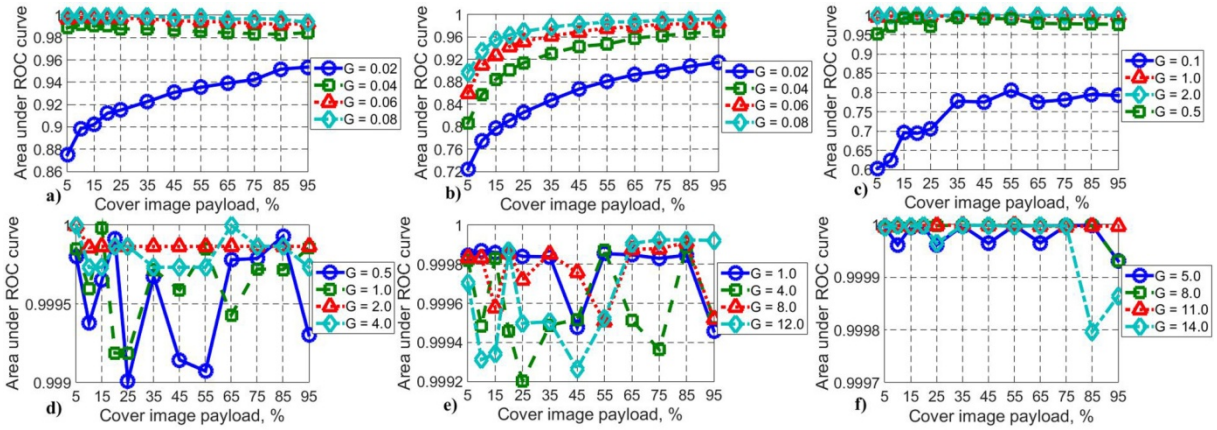

Figure 3. Dependency of AUC metrics on cover image payloads by variation the weighted coefficients G for statistical stegdetector SD_{CDF} (JPEG Quality Factor – 100). Message was embedded according to: (a) – Dey method; (b) – Agarwal method; (c) – Joseph method; (d) – Khan method; (e) – Elahian method; (f) – Gunjal method.

Usage of SPAM model allows detecting with high accuracy (AUC>0.99) the stego images, formed according to multistage Khan method (Fig.1d), as well as Elahian (Fig.1e) and Gunjal (Fig.1f) complex methods, irrespective to the cover image payload and value of coefficient G. It is preliminary unexpected results, since these methods were proposed for increasing the robustness of steganograms to statistical steganalysis. Relatively low robustness of stego images in this case is explained by significant decreasing the correlation between brightness of adjacent pixels (parameters of SPAM models) in comparison with corresponding values for cover images.

Usage of one-stage embedding methods of Dey and Agarwal, as well as multistage Joseph method gives opportunity to significantly decrease the accuracy of stego images detection (Fig. 1a-b), especially in case of low cover image payload ($\Delta_C < 10\%$) and minimal values of coefficient G. Obtained results are explained by simultaneously applying of spectral (2D-DWT) and special (SVD) transform of cover image by message hiding.

Passive steganalysis of DI with usage of CC-PEV model is characterized by relatively low accuracy of stego images detection in case of message hiding with usage of spectral transformation of cover images (Dey and Joseph methods, Fig.2b-c) and low cover image payload ($\Delta_C < 10\%$). Revealed diminution of detection accuracy is connected with peculiarity of CC-PEV model – usage of coefficients the 2D-DCT, obtained for detached blocks, by calculations of model's parameters. Therefore changes of statistical parameters of cover images, caused by message hiding, in these blocks are relatively low, which decrease the effectiveness of applying the CC-PEV model for stego images revealing.

Despite of significantly increasing of dimensionality the feature space by usage of CDF model in comparison with SPAM and CC-PEV models ($d_{SPAM} = 686$, $d_{CC-PEV} = 548$, $d_{CDF} = 1234$), increasing of detection accuracy is relatively small – $\Delta AUC \le 0.055$. For comparison values of AUC metrics in case of low cover image payload, minimum values of weighted coefficient G and usage the statistical stegdetector SD_{SPAM}, SD_{CC-PEV} and SD_{CDF} are represented in Table 3.

It should be mentioned, that lossy JPEG-compression of DI (JPEG Quality Factor is less than 100) lead to additional decreasing the detection accuracy (table 3). It is explained by usage during message hiding of of approximation coefficients the 2D-DWT and the greatest singular values, that corresponds to low-frequency 2D-DCT coefficients. In consequence, alteration of stego images due to JPEG-compression is relatively small.

Applying of universal CDF model allow achieving the high detection accuracy only in case of forming the stego image according to multistage and complex embedding methods (table 3). On the other hand, usage of spectral (2D-DWT) and special (SVD) transform gives opportunity to significantly decrease the detection accuracy.

Table 3. Values of AUC metrics in case of low cover image payload, minimum values of weighted coefficient G and usage the statistical stegdetector SD_{SPAM}, SD_{CC-PEV} and SD_{CDF}

	Statistical model of digital image			
	SPAM	CC-PEV (JQF = 90)	CC-PEV (JQF = 100)	CDF
Dey method	0.843	0.710	0.730	0.898
Agarwal method	0.753	0.542	0.586	0.775
Joseph method	0.585	0.569	0.549	0.623
Khan method	0.990	0.932	0.999	0.999
Elahian method	0.999	0.622	0.984	0.999
Gunjal method	0.984	0.999	0.999	0.999

Conclusion

On the basis on conducted analysis the detection accuracy of stego images, formed according to one-stage, multistage and complex methods, by usage of modern statistical stegdetectors it is established that:

1. Utilization of well-known statistical models of digital images in spatial (SPAM model) and frequency (CC-PEV model) domains, as well as CDF universal model does not gives opportunity to achieve the high detection accuracy in case of message hiding with usage of one-stage Dey and Agarwal methods, as well multistage Joseph methods. It is explained by usage of low-frequency (approximation) coefficients and the greatest singular values, which correspond to image's components with highest energy, (Dey and Agarwal methods) or message hiding at the level of intrinsic noise of digital images (Joseph method). Accurate modelling of mentioned components requires creation a new statistical models.

2. Forming of stego images according to multistage Khan method, as well as Elahian and Gunjal complex methods leads to significant changes of correlation between brightness of adjacent pixels the cover images. It leads to considerable increase of detection accuracy (AUC>0.99), despite of usage of several domains for message hiding and applying the preliminary stage for processing cover image and stegodata.

Acknowledgements

The paper is published with partial support by the project ITHEA XXI of the ITHEA ISS (www.ithea.org) and the ADUIS (www.aduis.com.ua).

Bibliography

[Agarwal, 2008] Agarwal R., Santhanam M.S. Digital watermarking in the singular vector domain. International Journal of Image and Graphics, 2008. Vol. 8. pp. 351-362.

[Agranovskyi, 2009] Agranovskyi A.V., Balakin A.V., Gribunin V.G., Sapognikov S.A. Steganography, Digital Watermarking and Steganalysis. Moskow, Vyzovskaya Kniga, 2009. 220 pp. ISBN 978–5–9502–0401–2. (in Russian)

[Bolshakov, 2007] Bolshakov A.A., Karimov R.N. Methods of multidimensional signals processing. Moskow, Goryachaya Liniya – Telekom. 2007. 522 pp. ISBN 5–93517–287–9. (in Russian)

[Denemark, 2015] Denemark Tomáš, Fridrich Jessica. Improving Steganographic Security by Synchronizing the Selection Channel. Proceedings of the 3rd ACM Workshop on Information Hiding and Multimedia Security. New York, NY, USA, 2015. DOI 10.1145/2756601.2756620.

[Dey, 2011] Dey N., Roy A.B., Dey S. A novel approach of color image hiding using RGB color planes and DWT. International journal of computer application, 2011. Vol. 36, No.5. pp.19-24.

[Elahian, 2011] Elahian A., Khalili M., Shokouhi S.B. Improved robust DWT–watermarking in YCbCr color space. Global journal of computer application and technology. 2011. Volume 1, Issue 3. pp. 300–304. ISSN 2249–1945.

[Fridrich, 2010] Fridrich J. Steganography in Digital Media: Principles, Algorithms and Applications. Cambridge University Press, New York, USA. 2010. 437 p.

[Fridrich, 2012] Fridrich J., Kodovsky J. Rich models for steganalysis of digital images. IEEE Transactions on Information Forensics and Security, 2012. Vol. 7, Issue 3. pp. 868-882.

[Gonzalez, 2008] Gonzalez R., Woods R. Digital Image Processing. International version 3rd edition. Pearson Education Press, 2008. 1103 p.

[Gribunin, 2002] Gribunin V.G., Okov I.N., Tyrintsev S.V. Digital Steganography. Moskow, Solon-Press, 2002. 265 pp. ISBN 5–98003–011–5. (in Russian)

[Gunjal, 2011] Gunjal Baisa L., Mali Suresh N. Secured color image watermarking technique in DWT–DCT domain. International Journal of Computer Science, Engineering and Information Technology (IJCSEIT). 2011. Volume 1, Issue 3. pp. 36–44. DOI 10.5121/ijcseit.2011.1304.

[Holub, 2013a] Holub V., Fridrich, J. Digital image steganography using universal distortion. Puech, W., Chaumont, M., Dittmann, J., and Campisi, P., eds. In 1st ACM IH&MMSec. Workshop, June 17–19, 2013.

[Holub, 2013b] Holub V., Fridrich J. Random Projections of Residuals for Digital Image Steganalysis. IEEE Transactions on Information Forensics and Security, 2013. Vol. 8, Issue 12. pp. 1996-2006.

[Huiskes, 2008] Mark J. Huiskes, Michael S. Lew. The MIR Flickr Retrieval Evaluation. Proceedings of the 2008 ACM International Conference on Multimedia Information Retrieval, Vancouver, Canada. ACM Press, New York, NY, USA. DOI 10.1145/1460096.1460104.

[Joseph, 2013] Joseph A., Anusudha K. Robust Watermarking Based on DWT-SVD. International Journal on Signal & Image Security, 2013. Issue 1, Vol. 1.

[Katzenbeisser, 2000] Katzenbeisser S., Petitcolas P. Information Hiding Techniques for Steganography and Digital Watermarking. Artech House, Boston, USA. 2000. 237 p.

[Khan, 2013] Khan M.I., Rahman M., Sarker I.H. Digital Watermarking for image Authentication Based on Combined DCT, DWT and SVD Transformation. International Journal of Computer Science Issues, IJCSI, 2013. Volume 10, Issue 3, No. 1. pp. 223-230.

[Kodovský, 2010] Kodovský Jan, Pevny Tomas, Fridrich Jessika. Modern steganalysis can detect YASS. Proc. SPIE 7541, Media Forensics and Security II. San Jose, California, USA, 2010. Ed. Memon Nasir D., Dittmann Jana, Alattar Adnan M., Delp Edward J. pp. 1–11. DOI 10.1117/12.838768.

[Kodovský, 2012a] Kodovský J., Fridrich J., Holub V. Ensemble Classifiers for Steganalysis of Digital Media. IEEE Transactions on Information Forensics and Security, 2012. Vol. 7, No. 2. pp. 432-444.

[Kodovský, 2012b] Kodovský J., Fridrich J. Steganalysis of JPEG Images Using Rich Models. Proc. SPIE, Electronic Imaging, Media Watermarking, Security, and Forensics XIV, San Francisco, CA, January 23–25, 2012. DOI: 10.1117/12.907495.

[Mathews, 1975] Mathews B.W. Comparison of the predicted and observed secondary structure of T4 phage lysozyme. Biochimica Et Biophysica Acta (bba), Protein Structure, BIOCHIM BIOPHYS ACTA PROTEIN. 1975. Volume 402, Issue 2. pp. 442–451. DOI 10.1016/0005–2795(75)90109–9.

[Murphy, 2012] Murphy Kevin P. Machine Learning: A Probabilistic Perspective. Massachusetts Institute of Technology Press, 2012. 1071 p.

[Oppenheim, 2010] Oppenheim Alan V., Shaffer Ronald W. Discrete-Time Signal Processing. 3rd edition. Pearson Education Press, 2010. 1046 p.

[Pevny, 2007] Pevny T., Fridrich J. Merging Markov and DCT features for multiclass JPEG steganalysis. Proceedings SPIE, Electronic Imaging, Security, Steganography, and Watermarking of Multimedia Contents IX. San Jose, CA, USA, 2007. Ed. Delp E.J., Wong P.W. pp. 1–14.

[Pevny, 2010a] Pevny T., Bas P., Fridrich J. Steganalysis by Subtractive Pixel Adjacency Matrix. IEEE Trans. on Information Forensics and Security, 2010. Vol. 5, Issue 2. pp. 215-224.

[Pevny, 2010b] Pevný T., Filler T., Bas P. Using high-dimensional image models to perform highly undetectable steganography. Proceedings of Information Hiding, 12th International Workshop. Lecture Notes in Computer Science. Ed. Böhme R., Safavi–Naini R. Calgary, 2010. pp. 161–177.

[Progonov, 2014] Progonov D. O., Kushch S. M. Revealing of steganograms with data, which are hidden in transformation domain of digital images. Visn. NTUU KPI, Ser. Radioteh. radioaparatobuduv., 2014. No. 57, pp. 128-142. (in Ukrainian).

[Progonov, 2015a] Progonov Dmytro, Kushch Serhii. Spectral analysis of Steganograms. Scientific Journal “Radio Electronics, Computer Science, Control”, 2015. Volume 2 (33). pp. 71-81. DOI 10.15588/1607-3274-2015-2-9. (in Ukrainian).

[Progonov, 2015b] Progonov Dmytro, Kushch Serhii. Passive Steganalysis of Multidomain Embedding Methods. International Journal “Information Theories & Applications”, 2015. Volume 22, Issue 1. pp. 86–99. ISSN 1310–0513.

Authors' Information

Dmytro Progonov *– the 3rd year postgraduate student, the Assistant, Faculty of Information Security, Institute of Physics and Technology, National Technical University of Ukraine “Kyiv Polytechnic Institute”; Postal Code 03056, Prospect Peremohy, 37, Kyiv, Ukraine;*

e-mail: progonov@gmail.com.

Major Fields of Scientific Research: Digital Media Steganalysis, Advanced Signal Processing, Machine Learning.