NATIONAL AVIATION UNIVERSITY
INSTITUTE OF AERONAVIGATION

RADIO-ELECTRONIC COMPLEXES DEPARTMENT

RADIO-ELECTRONICS DEPARTMENT

IEEE UKRAINE SECTION JOINT SP/
AES CHAPTER (KIEV)

STATISTICAL METHODS OF SIGNAL AND DATA PROCESSING (SMSDP-2010)

PROCEEDINGS
OCTOBER 13-14, 2010 KIEV, UKRAINE

Publishing of National Aviation University «NAU-Druk»
Kiev 2010

NATIONAL AVIATION UNIVERSITY
INSTITUTE OF AERONAVIGATION

RADIO-ELECTRONIC COMPLEXES DEPARTMENT

RADIO-ELECTRONICS DEPARTMENT

IEEE UKRAINE SECTION JOINT SP/
AES CHAPTER (KIEV)

STATISTICAL METHODS OF SIGNAL AND DATA PROCESSING

(SMSDP-2010)

PROCEEDINGS

OCTOBER 13-14

Kiev
National Aviation University
"NAU-Druk" Publishing House
2010

UDC 621.391: 519.2 (063)

Матеріали конференції складаються із тез доповідей науковців з Іспанії, Мексики, Польщі, України, Росії та ін.

Матеріали можуть бути корисними для науковців та викладачів і тих, хто працює у цій галузі.

Statistical Methods of Signal and Data Processing (SMSDP-2010): Proceedings. Kiev, Ukraine, October 13-14, 2010 / General Chairman I. Prokopenko. – Kiev: National Aviation University "NAU-Druk" Publishing House, 2010. – 180 p.

ISBN 978-966-598-662-1

The book of materials of International Conference "Statistical Methods of Signal and Data Processing (SMSDP-2010)" contains articles contributed by researchers from Spain, Libya, Mexico, Poland, Ukraine, Russian etc.

The materials presented can be used in further scientific and academic process at universities and colleges, are of interest both for the general reader and professionals.

General Chairman of SMSDP-2010 Prof. *Igor Prokopenko*

ISBN 978-966-598-662-1

International Conference on

STATISTICAL METHODS OF SIGNAL AND DATA PROCESSING (SMSDP-2010)

October 13-14, 2010 Kiev, Ukraine

ORGANIZED

by IEEE Ukraine Section Joint SP/AES Chapter (Kiev) and National Aviation University – Radio-electronic Complexes Department and Radio-Electronics Department

OBJECTIVES AND GROUNDS

The Conference is planned to be a forum of ideas exchange and results discussion in the field of methods, algorithms and means of data and signal processing. Submissions that contain novel results is signal detection, parameter estimation, and other kind of statistical processing of signals and data in telecommunication, remote sensing and information systems are welcome.

SESSIONS

1. Statistical models of signals and fields.
2. Detection of signals. Signal and interference parameters estimation.
3. Signal and data processing in remote sensing.
4. Discrete signals processing in telecommunication systems.
5. Data and signal processing in information systems.

Technical program committee Co-Chairs:

Krzysztof Kulpa	Warsaw University of Technology, Poland
Igor Prokopenko	National Aviation University, Ukraine
Felix Yanovsky	National Aviation University, Ukraine

Chairman of SMSDP-2010

Prof. **Igor Prokopenko**, IEEE Member
prokop-igor@yandex.ru

Co-Chair

Prof. **Felix Yanovsky**, IEEE Fellow
felix.yanovsky@ieee.org

SMSDP-2010 Coordinator

Dr. **Rustem Sinitsyn,** IEEE Member
comandor@i.com.ua

CONTENT

Session 1. STATISTICAL MODELS OF SIGNALS AND FIELDS

Session 2. DETECTION OF SIGNALS. SIGNAL AND INTERFERENCE PARAMETERS ESTIMATION

Session 3. SIGNAL AND DATA PROCESSING IN REMOTE SENSING

Session 4. DATA AND SIGNAL PROCESSING IN INFORMATION SYSTEMS

Session 5. SIGNAL COMPRESSION, CODING AND CRYPTOGRAPHY

Use of Complex Discrete Signals for Steganographic Information Security

Alexander Kuznetsov
Information and Computing Center, Kozhedub Air Force University, Kharkov, Ukraine
kuznetsov_alex@rambler.ru

Roman Serhiienko
Cathedra of Artillery Reconnaissance, Army Academy, L'viv, Ukraine

Vladislav Kovtun
Cathedra of Information Technologies Security, National Aviation University, Kiev, Ukraine

Anatolii Botnov
Cathedra of Information Technologies Security, Kharkov National University of Radio Electronics, Kharkov, Ukraine

The steganographic methods of information embedding in stable images for the secure transfer of information are investigated. The method of steganographic security based on the use of complex discrete signals and technology of direct expansion of spectrum is examined, its efficiency is probed from point of the provided security, carrier capacity and amount of introduced distortions in a container-image.

Keywords-component; container-image; steganography; noise-immunity; erroneous extraction

I. RAISING THE PROBLEM IN A GENERAL VIEW AND ANALYSIS OF LITERATURE

Important direction in development of modern facilities of information secutity are the steganographic systems, which provide a keeping in a secret from the opponent of not only informative content of transferrable information but also fact of passing of messages [1, 2]. Most perspective are steganographic methods the construction of which is based on the developed mathematical apparatus of discrete signals theory and noise-immune communication [3 - 7].

The purpose of this paper is research of steganographic method of embedding of information in stable images [2], based on the use of complex discrete signals and technology of direct expansion of spectrum, estimation of its efficiency from point of the provided security, carrier capacity and amount of the insertion distortions in a container-image.

II. DIRECT EXPANSION OF SPECTRUM (DIRECT SEQUENCE) IN A COMMUNICATION THEORY

For the construction of modern and noise-immune digital communication networks the methods of theory of discrete signals, cross-correlation and spectral analysis are utilized [3 – 5]. From point of the effective use of frequency-time and energy resources of communication channels as most perspective ones are considered the broadband systems with complex discrete signals and direct expansion of spectrum [3, 4].

As a discrete signal we will consider an informative signal which is a set of separate values taken at certain moments of time. Further we will consider a complex discrete signal as binary pseudorandom sequence (BPRS) $\Phi_i = (\varphi_{i_0}, \varphi_{i_1}, ..., \varphi_{i_{n-1}})$ lengths of n from set $\Phi = \{\Phi_0, \Phi_1, ..., \Phi_{M-1}\}$ with power of $|\Phi| = M$. The elements of binary BPRS take values:

$$\varphi_{i_z} = \begin{cases} +1 \\ -1 \end{cases}, \; z = 0, ..., n-1 \qquad (1)$$

For the construction of noise-immune broadband connection the concept of correlation of discrete signals to statistical relationship of two or a few BPRS is utilized. The mathematical measure of correlation (conformity) of two discrete signals $\Phi_i, \Phi_j \in \Phi$ is the coefficient of correlation $\rho(\Phi_i, \Phi_j)$ [3, 4]:

$$\rho(\Phi_i, \Phi_j) = \frac{1}{n} \sum_{z=0}^{n-1} \Phi_{i_z} \Phi_{j_z} \qquad (2)$$

Two signals Φ_i, Φ_j are orthogonal, if coefficient of correlation $\rho(\Phi_i, \Phi_j) = 0$. If $\rho(\Phi_i, \Phi_j) \approx 0$ we will name signals Φ_i and Φ_j as a quasi-orthogonal ones [5, 7].

In a modern digital communication theory the large bands of weak-correlated discrete signals are utilized for the construction of the wide-band noise-immune systems of data communication. Transferred messages in such cannels acquire the appearance of noise-like sequences, and due to high power of ensembles of discrete signals and direct expansion of frequency spectrum high authentication, noise-immunity and secrecy of the digital connection channels are provided [3 – 5].

For communication of data in a broadband communication network informative signal $x(t) = \begin{cases} +1 \\ -1 \end{cases}$ modulated by means of multiplying by an extending code signal $g(t) = \Phi_i \in \Phi$ – pseudorandom sequence from the ensembles of discrete signals which were considered earlier. Code signal on the statistical properties is similar to noise, therefore the extended signal

$$y'(t) = y(t) + e(t) \quad (3)$$

is poorly distinguished from noises in a communication channel, which allows to carry out the hidden transmission.

At a reception in a demodulator the resulted signal $y'(t) = y(t) + e(t)$ as mixture of the transferred sequence $y(t)$ and errors in a communication channel $e(t)$ is multiplied by the synchronized copy of extending signal $g(t)$. In other words, on a receiving side the calculation of coefficient of correlation is carried out (2), the value of which is determined by a decision-making rule:

$$\rho\left(y'(t), g(t)\right) = \frac{1}{n}\sum_{z=0}^{n-1} x(t)\Phi_{i_z}\Phi_{i_z} + \frac{1}{n}\sum_{z=0}^{n-1} e(t)\Phi_{i_z}$$

Taking into account a pseudorandomess of Φ_i, which are used as $g(t)$, the second term can be neglected (the amount of «+1» is approximately equal to the amount «-1») in right part of equation, i.e.

$$\rho\left(y'(t), g(t)\right) \approx \rho\left(y(t), g(t)\right) = x(t)\frac{1}{n}\sum_{z=0}^{n-1}\left(\Phi_{i_z}\right)^2 = x(t), \quad (4)$$

i.e. the value of informative signal on a receiving side is determined on expression

$$x(t) = \begin{cases} +1, \text{при } \rho\left(y'(t), g(t)\right) \approx +1; \\ -1, \text{при } \rho\left(y'(t), g(t)\right) \approx -1; \end{cases} \quad (5)$$

where the sign «$\approx$» is imposed by the presence of errors $e(t)$, caused natural or intentional noise in a communication channel.

III. Direct Expansion of Spectrum in Steganography

In the method of Smith-Comiskey [2], as well as in the considered above communication networks with direct expansion of spectrum, information message is bitwise modulated by multiplying the ensemble of orthogonal signals. After that the modulated message is embedded into a container – stable image.

We will present an information message m which is built in a digital container-image as blocks m_i of equal length, i.e. $m = (m_0, m_1, ..., m_{N-1})$ where every block m_i is a sequence (vector) of n bits: $m_i = (m_{i_0}, m_{i_1}, ..., m_{i_{n-1}})$. Let's consider a container-image as an array of data C by a dimension of $K \cdot L$, divided onto subblocks length of $k \cdot l = n$. As array cells C can be taken, for example, raster information of in-use image. Confidential key data is a set of base functions $Key = \Phi = \{\Phi_0, \Phi_1, ..., \Phi_{M-1}\}$, where all the base functions $\Phi_i = (\varphi_{i_0}, \varphi_{i_1}, ..., \varphi_{i_{n-1}})$ are interorthogonal discrete signals with length, equal to the size n block of report m_i, i.e. for any $i, j \in [0, ..., M-1]$ equality is true

$$\rho\left(\Phi_i, \Phi_j\right) = \frac{1}{n}\sum_{z=0}^{n-1}\Phi_{i_z}\Phi_{j_z} = \begin{cases} +1, \text{при } i = j; \\ -1, \text{при } i \neq j; \end{cases}$$

The purpose of steganographic transformation of information is embeding of every separate block of message m_i in the proper block of container-image. In the block of data of digital representation a dimension of $K \cdot L$ elements can be embedded $K \cdot \frac{L}{n}$ blocks of information message, i.e. to $K \cdot L$ bits. As key data (array of base functions $Key = \Phi$) we will utilize ensembles of orthogonal Walsh-Hadamard discrete signals.

Embedding of information message is carried out as follows. Every block of message $m_{i_j}, j = 0, ..., n-1$ compared with the separate block of container-image. Every informative bit of block $m_{i_j}, j = 0, ..., n-1$ appears as an informative signal $m_{i_j}(t) = \begin{cases} +1, m_{i_j} = 1; \\ -1, m_{i_j} = 0; \end{cases}$ and by analogy with (3) modulated by extending code signal (by base functions), i.e. BPRS $\Phi_j \in \Phi$. As a result, for every informative block m_i the modulated informative signal is formed

$$E_i(t) = \sum_{j=0}^{n-1}\sum_{z=0}^{n-1} m_{i_j}(t)\Phi_{j_z}. \quad (9)$$

Derived block of message E_i is pixelwise added to a subblock C_i of container, and steganogram (filled container) is formed by means of aggregation of data arrays S_i, $i = 0, ..., N-1$:

$$S_i = C_i + E_i \cdot G, \quad (10)$$

where $G > 0$ – amplification of extending signal factor, defining the «energy» of the embedded bits of informative sequence.

The operation of decoding consists in renewal of the hidden message by projection of every block S_i, derived steganographic image S on all of base functions $\Phi_j \in \Phi$, $i = 0, ..., N-1$. For this purpose every block S_i is presented in form vector $S_i = (S_{i_0}, S_{i_1}, ..., S_{i_{n-1}})$, $i = 0, ..., N-1$. To extract j-th bit of image from i-th block of steganographic image it is necessary to calculate the coefficient of correlation between Φ_j and the accepted block S_i (presented as a vector):

$$\rho\left(S_i, \Phi_j\right) = \frac{1}{n}\sum_{z=0}^{n-1} S_{i_z}\Phi_{j_z} = G \cdot \frac{1}{n}\sum_{z=0}^{n-1} E_{i_z}\Phi_{j_z} + \frac{1}{n}\sum_{z=0}^{n-1} C_{i_z}\Phi_{j_z}, \quad (11)$$

where C_i is one-dimensional array, i.e. the proper block of container, presented in form of vector.

Let's assume that array C_i has a random statistical structure, i.e. the second element in right part of expression (11) close to the zero and they can be ignored. Then we have:

$$\rho(S_i,\Phi_j)\approx G\cdot E_i\cdot\Phi_j = G\cdot\sum_{l=0}^{n-1}\sum_{z=0}^{n-1} m_{i_x}(t)\cdot\Phi_{l_z}\Phi_{j_z}. \quad (12)$$

By analogy with (6) we notice that all of sequences from a set of Φ interortogonal, i.e. if $l\neq j$ we have $\rho(\Phi_i,\Phi_j)=0$. Thus, all the terms in the right part of (12) in case of $l\neq j$ can be ignored. Thus:

$$\rho(S_i,\Phi_j)\approx G\cdot m_{i_j}(t)\cdot\frac{1}{n}\sum_{z=0}^{n-1}(\Phi_{j_z})^2 = G\cdot m_{i_j}(t). \quad (13)$$

Since $G>0$ and $n>0$ sign of $\rho(S_i,\Phi_j)$ в (13) depends only on $m_{i_j}(t)$, whence it appears:

$$m_{i_j}(t)=sign(\rho(S_i,\Phi_j))=\begin{cases}-1, \text{при } \rho(S_i,\Phi_j)<0;\\ +1, \text{при } \rho(S_i,\Phi_j)>0;\\ ?, \text{при } \rho(S_i,\Phi_j)=0;\end{cases} \quad (14)$$

If $\rho(S_i,\Phi_j)=0$ in (14) then we assume, that built-in information was lost.

Considered steganographic system inherits all of advantages of broadband communication networks: resistance to unauthorized extraction of built-in reports and its destruction or modification.

IV. ESTIMATION OF STEGANOGRAPHIC SYSTEM EFFICIENCY

Taking into account the purpose of steganographic system, we will introduce following efficiency indexes.

1. *Throughput* – the ratio of V embedded in the container of information to the general capacity D of container

$$Q=\frac{V}{D}. \quad (15)$$

2. *Key data capacity* (bits)

$$l_{Key}=\log_2(|Key|), \quad (16)$$

where $|Key|$ – power of set of key data.

3. *Security of steganographic method* we will estimate as an inverse value to power of set of secret key data. We can consider it as a probability index of secret key selection:

$$W=\frac{1}{|Key|}=2^{-l_{Key}}. \quad (17)$$

4. *Value of added distortions* is percentage ratio of arithmetical mean value of all the absolute values Δ-changes of container data to maximum value $\Delta_{\max}$:

$$I=\frac{\Delta_{mean}}{\Delta_{\max}}\cdot 100=\frac{100}{\Delta_{\max}\cdot D}\cdot\sum_{i=1}^{D}|\Delta_i|, \quad (18)$$

where Δ_i – Δ-changes i-th element of container.

5. *Probability of erroneous extraction* of message information data

$$P_{err}=\lim_{D\to\infty}\frac{V_{err}}{D}=1-\lim_{D\to\infty}\frac{V-V_{err}}{D}, \quad (19)$$

where V_{err} – capacity of extracted data.

By using of introduced indexes (15) – (19) we will estimate efficiency of the considered steganographic method of information security. Experimental researches we will conduct by embedding of information in bitmapped images (color model of RGB) with an 8-bit encoding of every color. The achieved results are presented on figures 3 – 6.

From the presented on a fig. 1. dependences follows, that the increase of carrying capacity of steganographic channel conduces to the strong increase of introduced distortions in a container-image. Unnoticeable for an observer distortions (that are lying below than threshold of sensitiveness of the visual system of human) are brought in only at $Q\leq 0.005$. It corresponds to building not more than 10 bits in one block of image, i.e. to modulating of to ten informative signals $m_{i_j}(t)$, $j=0,...,9$ in expression (9). Dependences, presented on a fig. 2, 3 testify that amplification factor used in expressions (10) - (13) allows substantially reduce the probability of erroneous extraction of informative data. Unfortunately, it is achieved at expence of the strong increase of the introduced distortions in an utilized container-image. Dependences are achieved at $Q=0.005$. Obviously, that for such value of carrying capacity an amplification factor cannot exceed 1..1,5 (see fig. 2). However even for such values probability of erroneous extraction is big and lies in a range of 0,1 .. 0,5. Integral dependence $I(P^*_{err})$, presented on a fig. 4, summarizes the information on a fig. 2, 3. For the fixed carrying capacity $Q=0.005$ derived an empiric curve, characterizing the dependence of size of the introduced distortions in a container-image and probabilities of erroneous extraction of informative data.

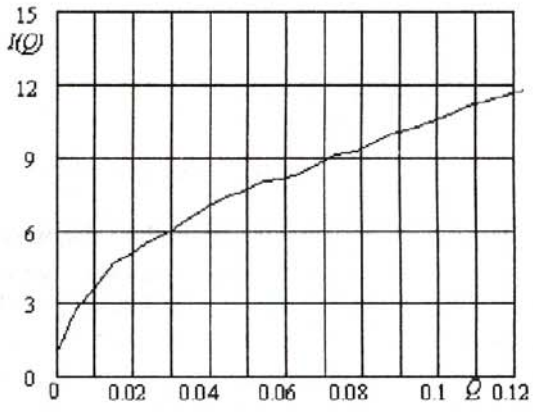

Figure 1. Dependence $I(Q)$

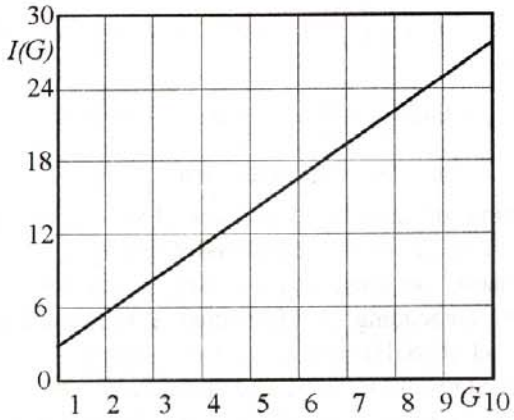

Figure 2. Dependence $I(G)$ for $Q = 0.005$

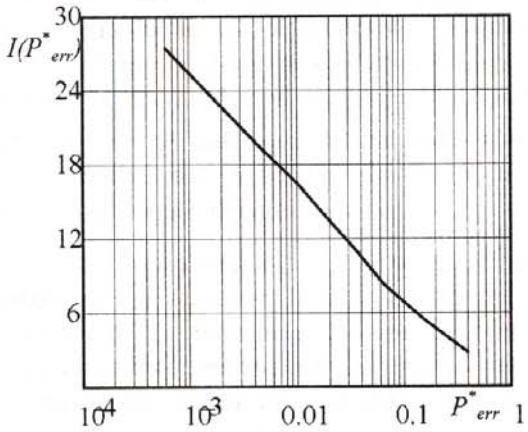

Figure 3. Dependence $P^*_{err}(G)$ for $Q = 0.005$

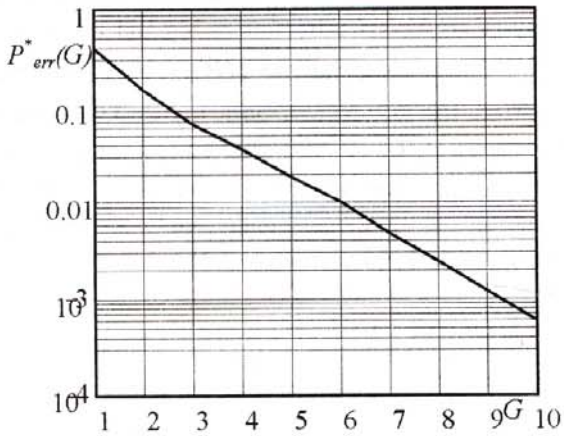

Figure 4. Dependence $I(P^*_{err})$ for $Q = 0.005$

For $Q = 0.005$ in order to obtain low distortions lying below the threshold of visual human sensitiveness ($I \le 2...3\%$) is possible only at ever-higher probability of erroneous extraction of informative data ($P_{err} \ge 0.1$). Obviously, that practical application of similar steganographic systems have to be combined with the anti-noise coding of informative data, that allow substantially reduce P_{err}.

V. Conclusions

The conducted researches show that use the direct expansion of spectrum of discrete signals in the steganographic purpoces allows to carry out secretive embedding of information messages in stable images. The task of extraction of message on the receiving side of steganographic system is equivalent to the task of finding out information from mixture of desired signal and noise (distortions) in a broadband communication network.

During researches the next lacks of the steganographic systems with expansion of spectrum of discrete signals are discovered: the probability of correct extraction of embedded data depends on the size of the introduced distortions, which, in its turn, depends on the provided carrying capacity of steganographic channel. In other words, the practical construction of steganographic system is conjugated with the search of compromise between in size of introduced distortions, by probability correctly extractions of message on a receiving side and by the provided carrying capacity. In addition it is revealed during researches that probability of correct extraction of embedded data directly depends on statistical properties of in-use container-image.

The perspective direction of further researches, by opinion of authors, is the use of large ensembles of weak-correlated (quasi-orthogonal) discrete signals for the construction of steganographic systems with direct expansion of spectrum. It allows, from one hand, without the considerable increase of the introduced distortions in a container-image to promote substantially the carrying capacity of steganographic channel. From other side, by means of the adaptive forming (choice) of discrete signals on the criterion of minimization of coefficient of correlation with a container-image it allows to reduce substantially probability of erroneous extraction of built-in data.

Acknowledgment

The authors thank Gorbenko I.D and Dolgov V.I. for their valuable comments, which helped to improve the paper.

References

[1] Kokhanovich G. F., Puzyrenko A. Y. Computer steganography. Theory and practice. – K.: «MK-Press», 2006. - 288 p., il.

[2] J. Smith, B. Comiskey, Modulation and Information hiding in Image. // Information hiding: First Int. Workshop "InfoHiding'96", Springer as Lecture Notes in Computing Science, vol 1174. 1996. – pp. 207-227.

[3] Digital methods in space communication. /Edited by S. Golomba.- M.: Sviaz, 1969. – 272 p.

[4] Skliar B. Digital communacations. Fundamentals and applications. - New Yersey: Prentice Hall PTR, 2001. – 1104 p.

[5] Gorbenko I.D., Stasiev Y.V. Analysis of derived ortogonal signal systems // Radiotekhnika. – 1989. – № 9. – pp. 16 – 18.