УДК 621.396

А.А. КУЗНЕЦОВ, д-р техн. наук, А.А. СМИРНОВ, канд. техн. наук

ВСТРАИВАНИЕ ДАННЫХ В КОНТЕЙНЕРЫ-ИЗОБРАЖЕНИЯ С ИСПОЛЬЗОВАНИЕМ СЛОЖНЫХ ДИСКРЕТНЫХ СИГНАЛОВ

Введение

Одним из перспективных направлений в развитии методов стеганографической защиты информации является использование технологии прямого расширения спектра для встраивания информационных данных в контейнеры-изображения [1 – 4]. Реализация таких стеганографических систем предполагает использование больших ансамблей дискретных сигналов с улучшенными корреляционными и ансамблевыми свойствами [2 – 4]. В то же время, как показали исследования [3, 4], число ошибок в извлекаемых данных в стеганографических системах с прямым расширением спектра неприемлемо высоко, а вероятность правильного извлечения встроенных данных непосредственно зависит от статистических свойств используемого контейнера-изображения. Другими словами, применение технологии прямого расширения спектра в стеганографических целях выдвигает дополнительные требования к используемым дискретным сигналам: их коррелированность с отдельным фрагментами контейнера-изображения должна быть минимизирована.

В данной работе проводятся исследования эффективности стеганографических методов встраивания информации в неподвижные контейнеры-изображения с использованием технологии прямого расширения спектра, исследуются зависимости величины вносимых искажений и вероятности возникающих ошибок при извлечении информационных данных от пропускной способности стеганоканала. Предлагается использование адаптивно формируемых ансамблей слабокоррелированных (квазиортогональных) дискретных сигналов для встраивания информационных данных. Показано, что за счет минимизации коэффициента корреляции формируемых дискретных сигналов и используемых контейнеров удается существенно снизить вероятность ошибочного извлечения встроенных данных.

Прямое расширение спектра в стеганографии

Технология прямого расширения спектра традиционно используется для построения широкополосных помехозащищенных систем скрытной радиосвязи, реализации множественного доступа с кодовым разделением каналов (CDMA – Code Division Multiple Access), обеспечения высокой имитостойкости на уровне применяемых ансамблей дискретных сигналов [5 – 7]. В стеганографических целях прямое расширение спектра частот используется для скрытного встраивания информационных данных в цифровые контейнеры, например неподвижные изображения [1 – 4]. При этом стеганографическая система наследует все преимущества широкополосных систем связи: устойчивость к несанкционированному извлечению встроенных сообщений (аналог скрытности в системе связи), устойчивость к разрушению или модификации встроенных сообщений (аналог помехозащищенности), устойчивость к навязыванию ложных сообщений (аналог имитостойкости в системе связи) [3, 4].

Введем следующие обозначения. Пусть $\Phi = \{\Phi_0, \Phi_1, ..., \Phi_{M-1}\}$ будет множеством (ансамблем) из $M = |\Phi|$ дискретных сигналов $\Phi_i = (\varphi_{i_0}, \varphi_{i_1}, ..., \varphi_{i_{n-1}})$ длины n. Каждый дискретный сигнал представляет собой двоичную псевдослучайную последовательность (ПСП), т.е. его элементы принимают одно из значений: $\varphi_{i_z} = \begin{cases} +1 \\ -1 \end{cases}, z = 0, ..., n-1$.

Важной характеристикой ансамбля Φ является коррелированность (похожесть) его элементов, для оценки которой используют коэффициент корреляции [5 – 7]:

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

Величина $\rho(\Phi_i,\Phi_j)$ характеризует статистическую взаимосвязь ПСП из ансамбля Φ. Для ортогональных сигналов $\rho(\Phi_i,\Phi_j)=0$, для квазиортогональных $\rho(\Phi_i,\Phi_j)\approx 0$.

Основной задачей теории дискретных сигналов является разработка вычислительно эффективных методов синтеза больших ансамблей слабокоррелированных дискретных сигналов, т.е. при заданных ограничениях на максимальное значение $\rho(\Phi_i,\Phi_j)$ требуется максимизировать мощность ансамбля $M=|\Phi|$ (либо при фиксированном $M=|\Phi|$ минимизировать максимальное значение $\rho(\Phi_i,\Phi_j)$).

Для передачи данных в широкополосной системе связи с прямым расширением спектра информационный сигнал $x(t)=\begin{cases}+1\\-1\end{cases}$ модулируется посредством его умножения на расширяющий кодовый сигнал $g(t)=\Phi_i\in\Phi$:

$$y(t)=x(t)g(t). \qquad (2)$$

Предположим, что временная длительность немодулированного сигнала $x(t)$ равна T, а его частота соответственно $F(x(t))=\frac{1}{T}$. Передача модулированного сигнала $y(t)$ при той же временной длительности T приведет к расширению частотного спектра передаваемого сигнала, пропорционально числу элементов псевдослучайной последовательности, т.е. пропорционально длине n: $F(y(t))=n\frac{1}{T}=nF(x(t))$. Поскольку кодовый сигнал по своим статистическим свойствам подобен шуму, то полученный расширенный (широкополосный) сигнал $y(t)$ слабо отличим от шумов в канале связи, что и позволяет осуществить скрытую передачу. Таким образом, передаваемые сообщения приобретают вид шумоподобных последовательностей, а за счет большой мощности ансамбля Φ и прямого расширения частотного спектра обеспечивается высокая имитостойкость, помехозащищенность и скрытность цифровых каналов связи [5 – 7].

При приеме в демодуляторе полученный сигнал $y'(t)=y(t)+e(t)$ как смесь переданной последовательности $y(t)$ и произошедших в канале связи ошибок $e(t)$ умножается на синхронизированную копию расширяющего сигнала $g(t)$. Другими словами, на приемной стороне осуществляется вычисление коэффициента корреляции (1), значение которого определяет правило принятия решения:

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

Учитывая псевдослучайность Φ_i, используемых в качестве $g(t)$, вторым слагаемым в правой части равенства можно пренебречь (количество «+1» примерно равно количеству «-1»), т.е.

$$\rho(y'(t),g(t))\approx\rho(y(t),g(t))=x(t)\frac{1}{n}\sum_{z=0}^{n-1}\left(\Phi_{i_z}\right)^2=x(t), \qquad (3)$$

т.е. значение информационного сигнала на приемной стороне определяется из выражения

$$x(t)=\begin{cases}+1, при\ \rho\left(y'(t),g(t)\right)\approx+1;\\ -1, при\ \rho\left(y'(t),g(t)\right)\approx-1;\end{cases} \quad (4$$

где знак «≈» предполагает наличие ошибок $e(t)$, вызванных естественными или преднамеренными помехами в канале связи.

Использование прямого расширения спектра передаваемого сигнала обеспечивает одновременную передачу многих других информационных сигналов в той же полосе частот. Это следует из слабой взаимной коррелировнаности ПСП из применяемых ансамблей дискретных сигналов. Действительно, если на приемной стороне принята аддитивная смесь $\sum_l y_l(t$ нескольких модулированных сигналов, тогда вычисление коэффициента корреляции дает следующее:

$$\rho\left(\sum_l y_l(t),g(t)\right)=\frac{1}{n}\sum_l\sum_{z=0}^{n-1}x_l(t)\Phi_{l_z}\Phi_{i_z}. \quad (5$$

Но все последовательности из множества Ф имеют низкое значение взаимной корреляции, т.е. при $l\neq i$ имеем $\rho(\Phi_l,\Phi_i)\approx 0$ (для ортогональных сигналов имеем равенство $\rho(\Phi_l,\Phi_i)=0$). Следовательно, всеми слагаемыми при $l\neq i$ в правой части равенства (5 можно пренебречь. Отсюда, при наличии в аддитивной смеси $\sum_l y_l(t)$ дискретного сигнала $\Phi_{l=i}$, имеем выражение (3) и соответствующее правило принятия решения (4).

Рассмотренный подход к организации цифровых помехозащищенных каналов связи нашел применение и при построении стеганографических методов защиты информации. Так, в работе [2] расширение спектра прямой последовательностью использовано для создания стеганографического метода встраивания данных в неподвижные изображения.

Целью стеганографического преобразования информации является встраивание каждого отдельного блока сообщения m_i, $i=0,...,N-1$ в соответствующий блок контейнера-изображения. Встраивание информационного сообщения осуществляется следующим образом. Каждый блок информационного сообщения сопоставляется с отдельным блоком контейнера-изображения. Каждый информационный бит блока m_{i_j}, $j=0,...,n-1$ представляется в виде информационного сигнала $m_{i_j}(t)=\begin{cases}+1, m_{i_j}=1;\\ -1, m_{i_j}=0;\end{cases}$ и по аналогии с (2) модулируется расширяющим кодовым сигналом, т.е. ПСП $\Phi_j\in\Phi$.

В результате, для каждого информационного блока m_i формируется модулированный информационный сигнал

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

Полученный информационный сигнал $E_i(t)$ попиксельно суммируется с подблоком контейнера C_i: $S_i=C_i+E_i\cdot G$, где $G>0$ – коэффициент усиления расширяющего сигнала, задающий «энергию» встраиваемых бит m_{i_j}, $j=0,...,n-1$ информационной последовательно-

сти. Стеганограмма (заполненный контейнер) S формируется посредством объединения массивов данных S_i, $i = 0,..., N-1$.

При извлечении информационных данных первичный контейнер C не требуется. Операция декодирования заключается в восстановлении скрытого сообщения путем проецирования каждого блока S_i полученной стеганограммы S на все $\Phi_j \in \Phi$, $i = 0,..., N-1$. Для этого каждый блок S_i представляется в форме вектора $S_i = \left(S_{i_0}, S_{i_1}, ..., S_{i_{n-1}}\right)$, $i = 0,..., N-1$

Чтобы извлечь j-й бит сообщения из i-го блока стеганоизображения, необходимо вычислить коэффициент корреляции между Φ_j и принятым блоком S_i (представленного в виде вектора):

$$\rho\left(S_i, \Phi_j\right) = \frac{1}{n}\sum_{z=0}^{n-1} S_{i_z}\Phi_{j_z} = G\cdot\frac{1}{n}\sum_{z=0}^{n-1} E_{i_z}\Phi_{j_z} + \frac{1}{n}\sum_{z=0}^{n-1} C_{i_z}\Phi_{j_z}, \qquad (6)$$

где под C_i понимается одномерный массив, т.е. соответствующий блок контейнера, представленный в форме вектора.

Предположим, что массив C_i имеет случайную статистическую структуру, т.е. положим, что второе слагаемое в правой части выражения (6) близко к нулю и им можно пренебречь. Тогда имеем:

$$\rho\left(S_i, \Phi_j\right) \approx G\cdot E_i\cdot\Phi_j = G\cdot\sum_{l=0}^{n-1}\sum_{z=0}^{n-1} m_{i_x}(t)\cdot\Phi_{l_z}\Phi_{j_z}. \qquad (7)$$

По аналогии с (5) отметим, что все последовательности из множества Φ слабокоррелированы, т.е. при $l \neq j$ имеем $\rho\left(\Phi_l, \Phi_j\right) \approx 0$. Следовательно, всеми слагаемыми в правой части равенства (7) при $l \neq j$ можно пренебречь. Отсюда имеем:

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

Тогда, по аналогии с (5) значения $m_{i_j}(t)$ могут быть легко восстановлены с помощью выражения:

$$m_{i_j}(t) = \begin{cases} +1, при\ \rho\left(S_i, \Phi_j\right) \approx G \\ -1, при\ \rho\left(S_i, \Phi_j\right) \approx -G; \end{cases}$$

где знак «≈» предполагает наличие незначительной статитической взаимосвязи отдельных элементов множества Φ и C_i.

Таким образом, процесс встраивания информационных сообщений для скрытной передачи очень схож с технологией расширения спектра дискретных сигналов в системах связи. Поэлементное сложение модулированного сообщения $E(t)$ с контейнером-изображением $C(t)$ следует интерпретировать как наложение ошибок $e(t)$ на полезный сигнал $y(t)$ в канале связи. Задача извлечения сообщения $m(t)$ из $S(t)$ на приемной стороне стеганосистемы эквивалентна задаче детектирования $x(t)$ из смеси полезного сигнала и помехи $y'(t) = y(t) + e(t)$ в широкополосной системе связи. Другими словами, рассмотренная стеганосистема фактически является некоторой интерпретацией широкополосных систем связи:

устойчивость к несанкционированному извлечению встроенных сообщений соответствует скрытности в системе связи, устойчивость к разрушению или модификации встроенных сообщений – помехозащищенности, устойчивость к навязыванию ложных сообщений – имитостойкости в системе связи.

Таким образом, использование развитого математического аппарата теории дискретных сигналов и технологии прямого расширения спектра в стеганографических целях позволяет осуществить встраивание информационных данных в неподвижные изображения для скрытной передачи и реализовать, таким образом, стеганографическую защиту информации.

Исследование эффективности стеганосистем с прямым расширением спектра

Для оценки эффективности стеганосистем в [3, 4] введены следующие показатели:

- *пропускная способность* Q организовываемого стеганоканала – отношение объема V встраиваемой в контейнер информации к общему объему D контейнера $Q = V / D$;

- *величина вносимых искажений* I как процентное отношение среднеарифметического всех абсолютных значений данных контейнера к максимально возможному значению;

- *вероятность ошибочного извлечения* $P_{ош}$ информационных данных сообщения.

Для проведения экспериментальных исследований эффективности встраивания сообщений в неподвижные изображения с использованием прямого расширения спектра разработана программная реализация, с помощью которой получены следующие эмпирические оценки: зависимости $I(V)$ и $P^*_{ош}(V)$ при фиксированном $D = 256$; зависимости $I(G)$ и $P^*_{ош}(G)$, где $P^*_{ош} \approx P_{ош}$ – частота ошибок извлечения, G – коэффициент усиления из (6).

Исследования проводились при встраивании информационных данных в растровые данные изображения (цветовая модель R,G,B) с 8-битным кодированием каждого цвета.
Эмпирические зависимости для случая использования ансамблей ортогональных сигналов Уолша–Адамара приведены на рис. 1 – 4.

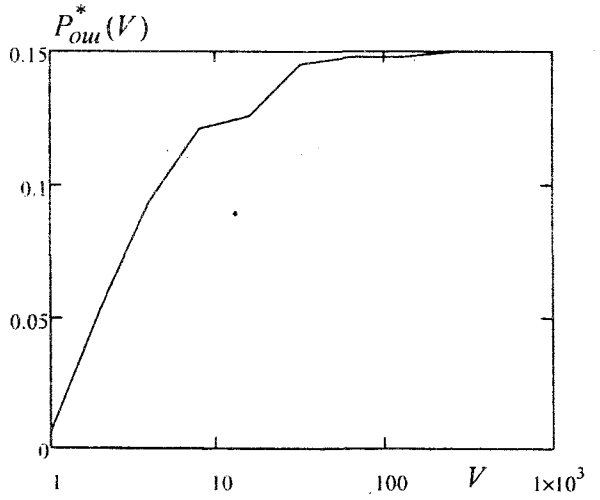

Рис. 1. Зависимость $P^*_{ош}(V)$, $D = 256$, $G = 1$

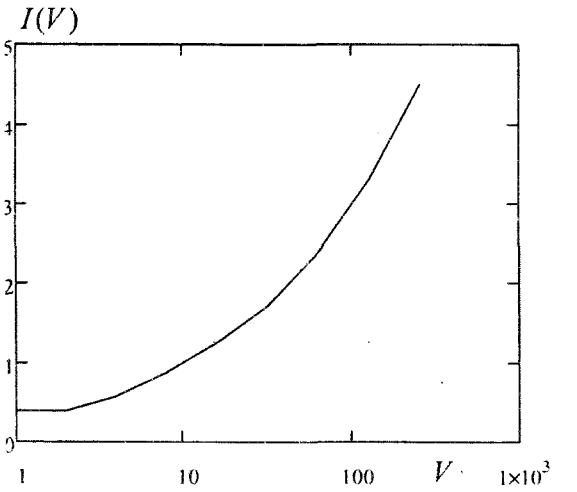

Рис. 2. Зависимость $I(V)$, $D = 256$, $G = 1$

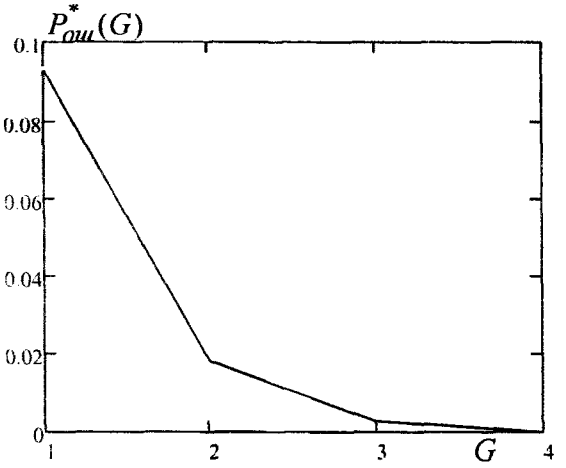

Рис. 3. Зависимость $P^*_{ош}(G)$, $D = 256$, $V = 4$

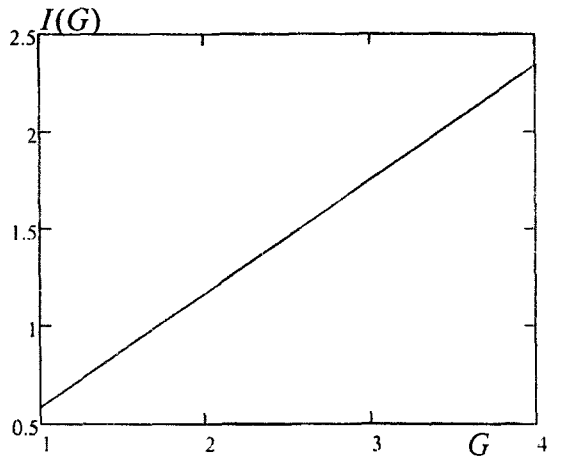

Рис. 4. Зависимость $I(G)$, $D = 256$, $V = 4$

Приведенные графики показывают, что повышение объема встраиваемых данных приводит к увеличению как вероятности ошибочного извлечения, так и к повышению доли вносимых искажений в контейнер-изображение. Приведенные зависимости также свидетельствуют, что увеличение коэффициента усиления G приводит к резкому снижению вероятности ошибочного извлечения и одновременному повышению величины вносимых искажений.

Следует отметить, что рассчитанное значение I является усредненной величиной, характеризующей долю внесенных искажений в среднем, по всем пикселям контейнера-изображения. Отдельные пиксели или группа пикселей могут быть значительно искажены, доля вносимых искажений для этих фрагментов изображения может существенно превышать рассчитанное усредненное значение V. Для примера на рис. 5, 6 приведены: контейнер S, заполненный с показателями: V = 128, G = 1 (рис. 5); исходный контейнер (рис. 6).

Рис. 5. Пример заполненного контейнера

Рис. 6. Пример пустого контейнера

Как следует из приведенных выше графиков, встраивание с такими параметрами дает усредненное значение доли вносимых искажений в пределах порога зрительной чувствительности человека. Однако, как видно из приведенных изображений (рис. 5, 6), для некоторых фрагментов искажения очень существенны. Таким образом, применение ортогональных дискретных сигналов оправдано только для стеганографических систем с малым объемом встраиваемых данных. Для повышения пропускной способности стеганоканалов при прочих равных условиях целесообразно использовать квазиортогональные дискретные сигналы.

На рис. 7, 8 приведены результаты исследования эффективности стеганосистем с использованием квазиортогональных дискретных сигналов, которые сформированы:

1) без адаптации к статистическим свойствам контейнера-изображения;
2) с адаптивным формированием последовательностей.

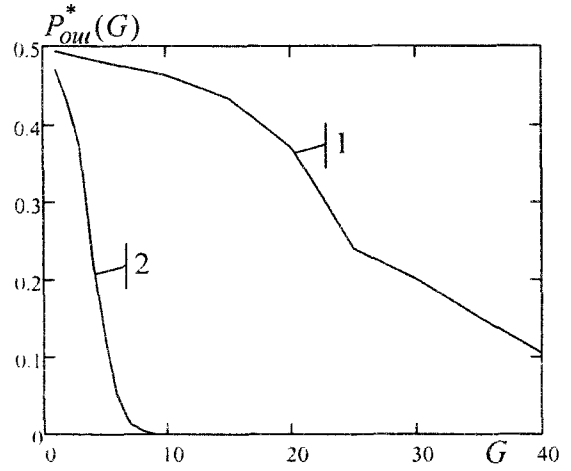

Рис. 7. Зависимость $P^*_{ош}(G)$, $D = 256$, $V = 10$

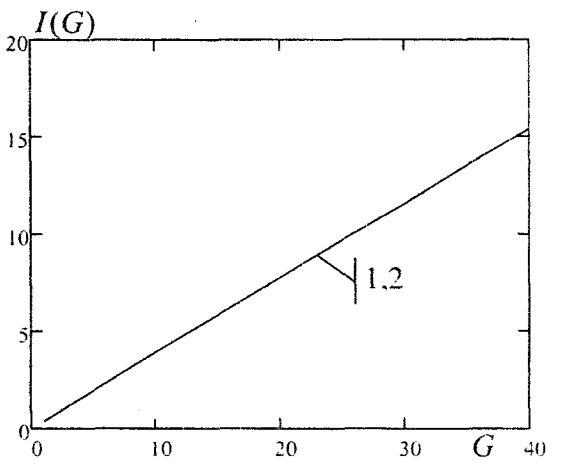

Рис. 8. Зависимость $I(G)$, $D = 256$, $V = 10$

Полученные эмпирические зависимости показывают, что повышение коэффициента усиления приводит к резкому снижению вероятности ошибочного извлечения информационных бит сообщения (см. рис.7). Однако это также ведет к увеличению вносимых искажений в контейнер-изображение (рис. 8). В то же время, как показывают проведенные исследования, по сравнению с использованием ортогональных дискретных сигналов применение квазиортогональных последовательностей приводит к меньшему искажению контейнера-изображения. Так, например, при V = 4 с использованием ортогональных дискретных сигналов при коэффициенте усиления G = 4 величина вносимых искажений составляет более 2,33 % (см. рис. 4). При большем числе вносимых бит данных (V = 10), а следовательно, и при большей пропускной способности стеганографического канала передачи данных применение квазиортогональных дискретных сигналов даже с большим значением коэффициента усиления (G = 5) приводит к меньшим искажениям контейнера, величина вносимых искажений не превосходит 2 % (см. рис. 8). Однако, как следует из представленных на рис. 7 зависимостей, использование квазиортогональных дискретных сигналов без адаптации к статистическим свойствам контейнера-изображения приводит к существенному повышению вероятности ошибочного извлечения бит сообщения (за счет сильной коррелированности ПСП с отдельными фрагментами контейнера-изображения). Так, например, для значений G < 10 величина ошибочно извлекаемых бит данных составляет почти 50 % (см. зависимость 1, рис 7). Избавиться от этого негативного фактора можно за счет адаптивного формирования квазиортогональных дискретных сигналов с учетом особенностей используемого контейнера-изображения.

В качестве примера на рис. 7, 8 цифрой 2 отмечена эмпирическая зависимость, полученная для случая использования адаптивно формируемых ансамблей квазиортогональных дискретных сигналов. Очевидно, что без увеличения величины вносимых искажений удается существенно снизить вероятность ошибочного извлечения информационных бит сообщений. Уже для значений G = 7..10 удается реализовать практически безошибочное извлечение информационных сообщений.

Таким образом, как следует из полученных результатов, применение адаптивно формируемых квазиортогональных дискретных сигналов позволяет существенно повысить пропускную способность организовываемых стеганографических каналов при сравнимых искажениях, вносимых в контейнер-изображение.

Выводы

Применение технологии прямого расширения спектра для встраивания данных в контейнеры-изображения позволяет строить эффективные стеганографические системы для

скрытной передачи информационных сообщений. Наиболее целесообразным является использование больших ансамблей квазиортогональных дискретных сигналов с адаптивным формированием (выбором) ПСП по критерию минимизации коэффициента корреляции с контейнером-изображением. Это позволяет существенно повысить пропускную способность организовываемого стеганоканала и снизить вероятность ошибочного извлечения встраиваемых данных.

Список литературы: 1. *Конахович Г. Ф., Пузыренко А. Ю.* Компьютерная стеганография. Теория и практика. – К. : МК-Пресс, 2006. – 288 с. 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. – P. 207-227. 3. *Кузнецов А. А.* Встраивание информационных данных в неподвижные изображения с использованием прямого расширения спектра / А. А. Кузнецов, А. М. Ботнов, П. А. Лаптий // Прикладная радиоэлектроника. – Т.9, N 3. – С.470-478. 4. *Kuznetsov A., Serhiienko R., Kovtun V., Botnov A.* Use of Complex Discrete Signals for Steganographic Information Security // Statistical Methods of Signal and Data Processing (SMSDP-2010): Proceedings. – Kiev : National Aviation University "NAU-Druk" Publishung House, 2010. – P. 143 – 146. 5. *Цифровые методы* в космической связи / Под ред. С. Голомба. – М. : Связь, 1969. – 272 с. 6. *Скляр Б.* Цифровая связь. Теоретические основы и практическое применение. – М. : Вильямс, 2003. – 1104 с. 7. *Горбенко И.Д., Стасев Ю.В.* Анализ производных ортогональных систем сигналов // Радиотехника. – 1989. – № 9. – С. 16 – 18.

Харьковский национальный университет радиоэлектроники

Поступила в редколлегию 10.07.2011