Efficient Coding of the Embedded Signal in Steganographic Systems with Multiple Access

Kobozeva A.A., Sokolov A.V.
Odessa National Polytechnic University,
Odessa, Ukraine

***Abstract*.** Today, steganographic systems with multiple access are of considerable importance. In such systems, the orthogonal Walsh-Hadamard transform is most often used for multiplexing and dividing channels, which leads to the need for efficient coding of the Walsh-Hadamard transform coefficients for the convenience of their subsequent embedding. The purpose of the research is to develop a theoretical basis for efficient coding of the embedded signal in steganographic systems with multiple access with an arbitrary number of users *N*, based on MC-CDMA technology. This purpose was fulfilled by forming the theoretical basis for constructing effective codes designed to encode the embedded signal in steganographic systems with multiple access. The most important results obtained are the proposed and proven relations that determine both the possible values of the Walsh-Hadamard transform coefficients, for a given value of the number of divided channels, and the probability of occurrence of the given values of the Walsh-Hadamard transform coefficients, which allow the construction of effective codes to represent the embedded signal. In the case of the number of divided channels N=4, we propose to use a constant amplitude code that provides a smaller value of the average codeword length in comparison with the Huffman code, while the constructed code has correcting capabilities. The significance of the obtained results is determined by the possibility of using the developed theoretical basis when constructing effective codes for encoding the embedded signal in steganographic systems with multiple access at an arbitrary value of the number of divided channels *N*.
***Keywords*:** steganography, multiple access, Walsh-Hadamard transform, efficient coding, C-code.
DOI: https://doi.org/10.52254/1857-0070.2021.2-50.09
UDC: 004.056

Codificare eficientă a semnalului de grup în steganosistemele cu acces multiplu

Kobozeva A.A., Sokolov A.V.
Universitatea Națională poltehnică din Odesa
Odesa, Ucraina

Rezumat. Astăzi, sistemele steganografice cu acces multiplu, care sunt capabile să asigure transmiterea simultană a diferitelor mesaje steganografice către mai mulți utilizatori autorizați într-un singur container, prezintă un interes considerabil. În astfel de sisteme, transformare ortogonală Walsh-Hadamard este cea mai des utilizată pentru multiplexarea și divizarea canalelor, ceea ce duce la necesitatea unei codificări eficiente a coeficienților de transformare Walsh-Hadamard pentru comoditatea încorporării lor ulterioare. În practică, pentru a rezolva această problemă, se folosesc cel mai des codurile Huffman, care, cu toate acestea, necesită statistici cunoscute ale alfabetului codificat, ceea ce duce la o complexitate crescândă a construcției lor cu o creștere a numărului de canale divizate. Scopul cercetării este de a dezvolta o bază teoretică pentru codificarea eficientă a unui semnal încorporat în sisteme steganografice cu acces multiplu cu un număr arbitrar de utilizatori N, pe baza tehnologiei MC-CDMA. Acest scop a fost atins prin formarea bazei teoretice pentru construirea unor coduri eficiente concepute pentru a codifica un semnal încorporat în sistemele steganografice cu acces multiplu. Cele mai importante rezultate obținute sunt relațiile propuse și dovedite care determină atât valorile posibile ale coeficienților de transformare Walsh-Hadamard, pentru o valoare dată a numărului de canale divizate, cât și probabilitatea de apariție a valorilor date ale coeficienții de transformare Walsh-Hadamard, care permit construirea de coduri eficiente care să reprezinte semnalul încorporat. Pentru a rezolva problema codării eficiente a semnalului încorporat, se propune utilizarea codurilor de amplitudine constantă înainte de a aplica transformarea ortogonală la semnalul încorporat.
***Cuvinte-cheie*:** steganografie, acces multiplu, transformare Walsh-Hadamard, codificare eficientă, cod C.

Эффективное кодирование группового сигнала в стеганосистемах с множественным доступом

Кобозева А.А., Соколов А.В.
Одесский национальный политехнический университет, Одесса, Украина

***Аннотация*.** Сегодня практический интерес представляют стеганосистемы с множественным доступом, способные обеспечить одновременную передачу различных стеганосообщений множеству

авторизованных пользователей в рамках одного контейнера. В таких системах для разделения каналов чаще всего используется ортогональное преобразование Уолша-Адамара, что приводит к необходимости предварительного кодирования группового сигнала с помощью эффективного кода. На практике для решения этой задачи чаще всего применяют коды Хаффмана, которые требуют известной статистики кодируемого алфавита, что приводит к возрастающей сложности их построения при увеличении числа разделяемых каналов. Целью исследования является разработка теоретического базиса для обеспечения эффективного кодирования группового сигнала в стеганосистемах с множественным доступом с произвольным числом абонентов N на основе технологии MC-CDMA. Поставленная цель была достигнута за счет формирования теоретического базиса для построения эффективных кодов, предназначенных для кодирования группового сигнала в стеганосистемах с множественным доступом. Наиболее важными результатами являются предложенные и доказанные соотношения, определяющие как возможные значения коэффициентов преобразования Уолша-Адамара для заданного значения числа разделяемых каналов, так и вероятности появления заданных значений коэффициентов преобразования Уолша-Адамара, что позволяет построение эффективных кодов для представления группового сигнала при любом количестве разделяемых каналов. Для решения задачи эффективного кодирования группового сигнала предложено использовать коды постоянной амплитуды перед применением ортогонального преобразования к групповому сигналу, например, в случае числа разделяемых каналов N=4 построен код постоянной амплитуды, обеспечивающий меньшее значение средней длины кодового слова по сравнению с применяемым ранее кодом Хаффмана, при этом построенный код обладает корректирующими способностями. Значимость полученных результатов состоит в возможности применения разработанного теоретического базиса при построения эффективных кодов для кодирования внедряемого сигнала в стеганосистемах с множественным доступом при произвольном значении числа разделяемых каналов N.

Ключевые слова: стеганография, множественный доступ, преобразование Уолша-Адамара, эффективное кодирование, С-код.

1. ВВЕДЕНИЕ И ПОСТАНОВКА ЗАДАЧИ

В настоящее время существенное распространение получила технология PLC (Power Line Communication), допускающая использование линий электропередач для организации высокоскоростной связи. Применение данной технологии позволяет существенно снизить затраты на организацию систем связи за счет использования существующей инфраструктуры, а также обеспечить доступ к высокоскоростным соединениям в местах, где затруднена прокладка дополнительных линий для передачи информации.

Тем не менее, в виду распространения информации в электросетях и её возможного попадания за пределы защищённого периметра применение технологии PLC сопряжено с особыми требованиями к подсистеме защиты информации. Часто в системах связи, основанных на технологии PLC, применяют криптографические методы защиты информации, тем не менее, при передаче особо чувствительной информации интерес также представляют перспективны использования цифровой стеганографии.

Современные стенографические методы не только делают невозможным прочтение информации злоумышленниками, но и скрывают от них сам факт наличия данной информации [1…4]. На сегодняшний день известно множество эффективных методов, предназначенных для внедрения цифровой информации в разнообразные контейнеры — изображения, видео, аудио, текст, программный код.

При этом, для внедрения информации используются самые разнообразные математические преобразования, начиная от классического метода LSB и его множественных модификаций [5…9], заканчивая методами внедрения информации с использованием коэффициентов дискретного косинусного преобразования [10], коэффициентов преобразования Уолша-Адамара [11] или даже с использованием генетических алгоритмов [12] и теории хаоса [13].

Отметим, однако, что большинство разработанных стеганографических методов позволяют организовать скрытый стеганографический канал лишь между парой абонентов.

При этом, для решения ряда практических задач возникает необходимость организации множественного доступа к скрытому каналу, что допускает одновременную работу множества авторизованных пользователей при обеспечении эффективного сокрытия самого факта передачи информации [14…16].

Исследования [14] показывают, что одним из эффективных способов организации множественного доступа в стеганоканале является использование технологии кодового разделения каналов MC-CDMA (Multi-Code Code Division Multiple Access). Технология MC-CDMA [17…20] предоставляет значительную гибкость распределения ресурсов стеганоканала между пользователями. Так, для некоторых, более приоритетных пользователей, может быть выделено несколько каналов связи, что приведет к кратному увеличению пропускной способности стеганоканала для данных пользователей.

Тем не менее, несмотря на высокую эффективность и перспективность использования технологии кодового разделения в стеганоканалах, данный вопрос остается достаточно малоисследованным. В литературе [14, 16] представлены лишь данные об организации частных случаев стеганосистем на основе технологии кодового разделения каналов MC-CDMA с числом абонентов $N = 4$, и только на основе кодов Хаффмана, при этом фундаментальные параметры кодирования группового сигнала, ровно как и вопросы оптимизации выбора числа каналов N с точки зрения минимизации значения среднего числа двоичных разрядов для представления элемента группового сигнала, остаются неизвестными.

Целью настоящей статьи является разработка теоретического базиса для обеспечения эффективного кодирования группового сигнала в стеганосистемах с множественным доступом с произвольным числом абонентов N на основе технологии MC-CDMA.

При этом под эффективным (статистическим) кодированием понимается применение кодов, позволяющих уменьшить избыточность [21] группового сигнала.

Для достижения данной цели необходимо решить следующие задачи:

— выбрать ортогональное преобразование, используемое для организации множественного доступа к скрытому каналу связи;

— исследовать динамику изменения средней длины кодового слова, необходимой для кодирования каждого коэффициента выбранного преобразования от количества разделяемых каналов N;

— разработать эффективный код для кодирования коэффициентов выбранного преобразования, который позволит уменьшить среднюю длину кодового слова по сравнению с используемыми до настоящего момента кодами для практически ценного значения $N = 4$ количества абонентов стеганоканала.

2. ОСНОВНЫЕ ПРИНЦИПЫ ПРИМЕНЕНИЯ ТЕХНОЛОГИИ MC-CDMA В СТЕГАНОГРАФИЧЕСКОЙ СИСТЕМЕ

Для полноты изложения материала кратко рассмотрим особенности функционирования технологии MC-CDMA для организации множественного доступа в стеганоканале.

Выберем для организации кодового разделения каналов ортогональное преобразование Уолша-Адамара. Неоспоримым преимуществом преобразования Уолша-Адамара является то, что элементы его базисных векторов принимают только значения бинарного алфавита $\{\pm 1\}$, что соответствует бинарной природе внедряемой информации. Данное обстоятельство, в совокупности с простотой правил построения матриц преобразования Уолша-Адамара, позволяет значительно упростить как программную, так и аппаратную реализацию алгоритмов внедрения и извлечения дополнительной информации.

В общем виде схема стеганосистемы с множественным доступом с использованием технологии MC-CDMA имеет вид, изображенный на рис. 1 [14].

На рис. 1 приняты следующие условные обозначения: ИС — источник сообщений, ПС — получатель сообщений.

Рассмотрим подробнее принципы функционирования устройства уплотнения каналов. В соответствии с технологией MC-CDMA биты входных данных d_i, поступающие по каждому каналу, изменяют знак одной из ортогональных функций W_i.

Исходя из выбранного вида ортогонального преобразования для построения матриц Уолша-Адамара порядка $N = 2^k$, строки (столбцы) которых представляют собой указанные функции

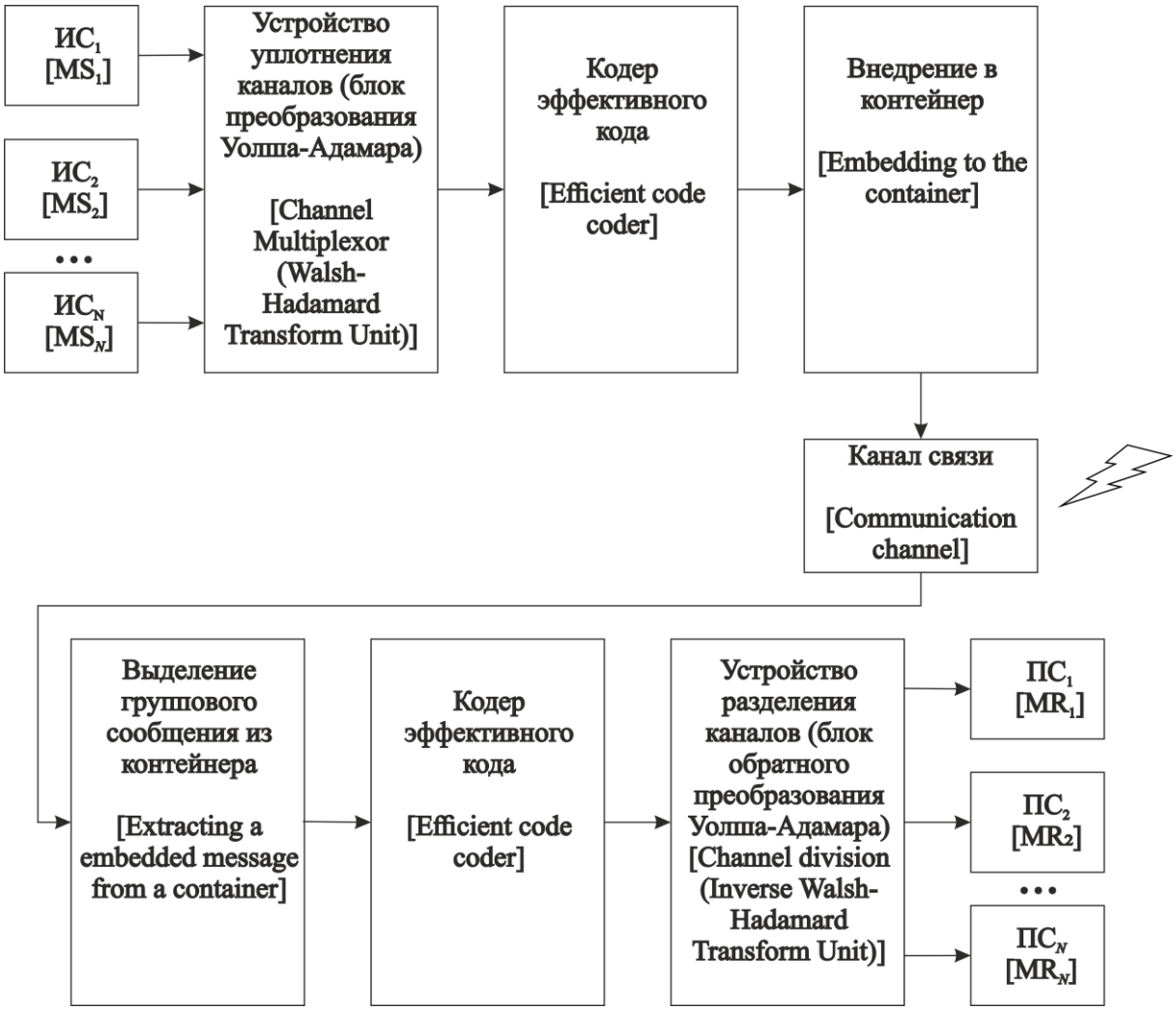

Рис. 1. Структурная схема стеганосистемы с множественным доступом с использованием технологии MC-CDMA[1].

Уолша, будем использовать конструкцию Сильвестра [22, 23], которая определяется следующей формулой

$$H_{2^k} = \begin{bmatrix} H_{2^{k-1}} & H_{2^{k-1}} \\ H_{2^{k-1}} & -H_{2^{k-1}} \end{bmatrix}, \quad (1)$$

где $H_1 = +1$.

$$S = DH. \quad (2)$$

Таким образом, групповой сигнал, который должен быть внедрен в контейнер в стеганосистемах с кодовым разделением каналов на основе технологии MC-CDMA, фактически представляет собой последовательность коэффициентов преобразования Уолша-Адамара вектора данных $D = \{d_i\}$, которая вычисляется в соответствии со следующей формулой [23]

Данное обстоятельство ведет к необходимости применения эффективного кодирования коэффициентов преобразования Уолша-Адамара для обеспечения удобства их последующего внедрения.

[1] Appendix 1

3. ПРОБЛЕМЫ ЭФФЕКТИВНОГО КОДИРОВАНИЯ ГРУППОВОГО СИГНАЛА

В работе [14] для осуществления операции эффективного кодирования коэффициентов преобразования Уолша-Адамара предложено использовать эффективные коды Хаффмана [24], требующие предварительного сбора статистики появления символов кодируемого алфавита.

Данная задача была решена в работе [14] с помощью полного перебора для значения

$N = 4$. Тем не менее отметим, что для практически ценных значений числа каналов $N > 32$, применение метода полного перебора для получения статистики появления коэффициентов преобразования Уолша-Адамара затруднено с вычислительной точки зрения, что обуславливает необходимость разработки соответствующего теоретического базиса.

Рассмотрим основные определения из теории кодирования, необходимые нам для проведения дальнейших исследований.

Определение 1 [24]. Весом Хэмминга w_t бинарной последовательности называется количество символов "–1" в ней.

Определение 2 [23]. Разбалансом Δ бинарной последовательности называется разность числа её символов "+1" и "–1"

$$\Delta = K^+ - K^-, \qquad (3)$$

где K^+ и K^- — число символов, соответственно, "+1" и "–1", содержащихся в бинарной последовательности.

На основе рассмотренных определений сформулируем утверждения, определяющие возможные значения коэффициентов преобразования Уолша-Адамара.

Утверждение 1. Множество значений $\{0, \pm 2, ..., \pm N\}$, которые может принимать разбаланс Δ двоичных последовательностей, составляет множество возможных значений коэффициентов преобразования Уолша-Адамара ω_i.

Доказательство. В соответствии с определением преобразования Уолша-Адамара (2) каждый коэффициент преобразования Уолша-Адамара кодового слова T, которое принадлежит множеству кодовых слов полного кода, является скалярным произведением исходного кодового слова T на соответствующую функцию Уолша-Адамара (столбец матрицы Уолша-Адамара). Таким образом, в соответствии с **Определением 2**, каждый коэффициент преобразования Уолша-Адамара представляет собой разбаланс поэлементного произведения исходного кодового слова T на соответствующую функцию Уолша-Адамара.

В силу того, что длина последовательности равна N, а все её элементы $t_i \in \{\pm 1\}$ то разбаланс последовательности веса w_t равен

$$\Delta = (N - w_t) - w_t = N - 2w_t. \qquad (4)$$

В виду того, что возможные значения веса $w_t \in \{0, 1, ...N\}$, то значения разбаланса принадлежат множеству $\Delta \in \{0, \pm 2, ..., \pm N\}$.

Утверждение 2. Вероятность появления коэффициента преобразования Уолша-Адамара с заданным значением ω_i определяется как

$$p(\omega_i) = p(-\omega_i) = \frac{C_N^{\frac{N-\omega_i}{2}}}{2^N}. \qquad (5)$$

Доказательство. Рассмотрим первый коэффициент преобразования Уолша-Адамара, который представляет собой разбаланс произведения последовательности T на первый столбец матрицы Уолша-Адамара (в соответствии с доказательством **Утверждения 1**), который по построению состоит из N символов "+1", т.е. $[+1+1...+1]^T$. Таким образом очевидно, что данный коэффициент представляет собой разбаланс последовательности T.

С другой стороны, количество последовательностей длины N с заданным весом w_t равно $J_{w_t} = C_N^{w_t}$.

Пусть вес последовательности T равен w_t, тогда в соответствии с (4) значение её веса может быть выражено через разбаланс как $w_t = \frac{N - \Delta}{2}$. В силу действия **Утверждения 1** возможные значения коэффициентов преобразования Уолша-Адамара равны значениям разбаланса последовательности T, и принадлежат множеству $\omega_i \in \{0, \pm 2, ..., \pm N\}$. А значит количество значений первого коэффициента преобразования Уолша-Адамара ω_i равно количеству последовательностей с заданным разбалансом $J(\omega_i) = C_N^{\frac{N-\omega_i}{2}}$.

Покажем, что указанное верно для любого коэффициента преобразования Уолша-Адамара.

Рассмотрим полный бинарный код мощности $J = 2^N$, каждое кодовое слово которого имеет длину N

$$\begin{matrix}[+1\ +1\ ...\ +1]\\ [+1\ +1\ ...\ -1]\\ \vdots\ \ \vdots\ \ \ \ \vdots\\ [-1\ -1\ ...\ -1]\end{matrix}. \qquad (6)$$

Оставшиеся коэффициенты преобразования Уолша-Адамара находятся как значения разбалансов произведений соответствующей функции Уолша (столбца матрицы Уолша-Адамара) на последовательности из полного кода

$$\begin{matrix}[h_N\ \ h_{N-1}\ ...\ \ h_1]\\ \times\ \ \ \times\ \ \ \ \ \times\\ [+1\ \ +1\ \ ...\ +1]\\ [+1\ \ +1\ \ ...\ -1]\\ \vdots\ \ \ \vdots\ \ \ \ \ \vdots\\ [-1\ \ -1\ \ ...\ -1]\end{matrix}. \qquad (7)$$

Очевидно, что данная операция (известная в теории сигналов как построение производной системы сигналов [23]), примененная к полному коду, сохранит все его кодовые слова с точностью до порядка их следования. Данное обстоятельство обусловлено отсутствием столбцов полного кода, которые являются инверсией друг друга, что ведет к тому, что любое знаковое кодирование столбцов полного кода сохраняет все его кодовые слова.

А значит, неизменным останется и количество кодовых слов полного кода, обладающих заданным значением разбаланса Δ и, соответственно, частоты появления того или иного значения коэффициентов преобразования Уолша-Адамара ω_i. В виду того, что общее количество коэффициентов преобразования Уолша-Адамара в одном векторе составляет N, то частота появления того или иного значения коэффициента преобразования Уолша-Адамара составляет $J(\omega_i) = NC_N^{\frac{N-\omega_i}{2}}$. А, стало быть, вероятность появления того или иного коэффициента преобразования Уолша-Адамара будет равна отношению частот появления данного коэффициента $J(\omega_i)$ к общему количеству коэффициентов преобразования Уолша-Адамара полного кода

$$p(\omega_i) = \frac{NC_N^{\frac{N-\omega_i}{2}}}{N2^N} = \frac{C_N^{\frac{N-\omega_i}{2}}}{2^N}.$$

Отметим также, что свойство равенства вероятностей появления равных по амплитуде, но противоположных по знаку коэффициентов преобразования Уолша-Адамара $p(\omega_i) = p(-\omega_i)$ легко доказать, раскрыв число сочетаний в числителе формулы (5)

$$p(\omega_i) = \frac{C_N^{\frac{N-\omega_i}{2}}}{2^N} = \\ = \frac{1}{2^N}\frac{N!}{\left(\frac{N-\omega_i}{2}\right)!\left(N-\frac{N-\omega_i}{2}\right)!} = , \qquad (8)$$
$$= \frac{1}{2^N}\frac{N!}{\left(\frac{N-\omega_i}{2}\right)!\left(\frac{N+\omega_i}{2}\right)!}$$

и с другой стороны

$$p(-\omega_i) = \frac{C_N^{\frac{N+\omega_i}{2}}}{2^N} = \\ = \frac{1}{2^N}\frac{N!}{\left(\frac{N+\omega_i}{2}\right)!\left(N-\frac{N+\omega_i}{2}\right)!} = . \qquad (9)$$
$$= \frac{1}{2^N}\frac{N!}{\left(\frac{N+\omega_i}{2}\right)!\left(\frac{N-\omega_i}{2}\right)!}$$

Рассмотрим пример работы **Утверждения 2**. При $N=16$ и $\omega_i = 2$. получаем, что вероятность появления данного коэффициента равна $p(2) = \frac{C_{16}^{\frac{16-2}{2}}}{2^{16}} = \frac{C_{16}^7}{2^{16}} = 0.1746$, что соответствует практически полученному результату. В соответствии с (5) вычислим все остальные вероятности появления коэффициентов преобразования Уолша-Адамара для данного значения $N=16$, для каждого из которых построим кодовые слова кода Хаффмана (табл. 1).

Анализ представленных в табл. 1 данных приводит к выводу, что средняя длина кодового слова необходимого для передачи одного коэффициента преобразования Уолша-Адамара при количестве разделяемых каналов $N=16$, составляет $l_{av} = 3.1041$.

Используя **Утверждение 2**, не составляет труда рассчитать l_{av} и для других значений

N . Тем не менее, в данном случае l_{av} зависит от особенностей кода Хаффмана.

В соответствии с теоремой Шеннона о кодировании источника [24], средняя длина кодового слова, необходимая для кодирования символа его алфавита не превышает информационную энтропию данного алфавита $l_{av} \geq H(\{\omega_i\})$, где

Таблица 1[2].

Код Хаффмана для коэффициентов преобразования Уолша-Адамара при $N = 16$

Символ [Symbol]	Вероятность появления [Probability of occurrence]	Код [Codeword]	Длина кодового слова [Codeword length]
ω_0	0.1964	11	2
ω_1	0.1746	001	3
ω_2	0.1746	000	3
ω_3	0.1222	100	3
ω_4	0.1222	011	3
ω_5	0.0667	0101	4
ω_6	0.0667	0100	4
ω_7	0.0278	10100	5
ω_8	0.0278	1011	4
ω_9	0.0085	1010100	7
ω_{10}	0.0085	101011	6
ω_{11}	0.0018	101010100	9
ω_{12}	0.0018	10101011	8
ω_{13}	0.0002	10101010100	11
ω_{14}	0.0002	1010101011	10
ω_{15}	0.0000153	101010101011	12
ω_{16}	0.0000153	101010101010	12

$$H(\{\omega_i\})_N = -\sum_{i=0}^{N} p(\omega_i) \cdot \log_2 p(\omega_i). \quad (10)$$

Таким образом, при произвольном значении N имеем следующее неравенство

$$l_{av} \geq -\sum_{i=0}^{N} \frac{C_N^{\frac{N-\omega_i}{2}}}{2^N} \cdot \log_2 \frac{C_N^{\frac{N-\omega_i}{2}}}{2^N}, \quad . \quad (11)$$
$$\omega_i = 0, \pm 2, ..., \pm N - 1$$

При заданном N правая часть выражения (11) составляет нижнюю границу l_{av} количества двоичных разрядов, необходимых для кодирования одного коэффициента преобразования Уолша-Адамара.

В табл. 2 приведены вычисленные в соответствии с (5) вероятностные характеристики коэффициентов преобразования Уолша-Адамара для других практически ценных значений числа разделяемых каналов $N = 2, 4, 8, 16, 32, 64, 128, 256, 512$, а также значения l_{av} и $H(\{\omega_i\})$.

Для наглядности, по данным табл. 2, на рис. 2 построен график зависимости средней длины кодового слова Хаффмана, необходимой для кодирования одного коэффициента преобразования Уолша-Адамара, а также информационной энтропии (11) от количества разделяемых каналов N.

Анализ данных рис. 2 свидетельствует о нарастании средней длины кодового слова

[2] Appendix 1

при росте количества разделяемых каналов N, что говорит о возрастании избыточности в представлении передаваемой информации при увеличении числа разделяемых каналов с помощью технологии MC-CDMA.

Рассмотрим последовательность элементов, которые представляют собой разности $\{\delta_i\} = H(\{\omega_i\})_N - H(\{\omega_i\})_{N/2}$ для значений $N = 4, 8, 16, 32, 64, 128, 256, 512$

Таблица 2[3].

Средняя длина кодового слова Хаффмана в зависимости от значения N

N	Количество различных ω_i [Number of different ω_i]	$p(\omega_0)$	$p(\omega_N)$	l_{av}	$H(\{\omega_i\})_N$
2	3	0.5	0.25	1.5	1.5
4	5	0.375	0.0625	2.125	2.0306
8	9	0.2734375	0.0039	2.5859	2.5442
16	17	0.196380615234375	0.000015258789063	3.1041	3.0465
32	33	0.139949934091419	0.000000000232831	3.5665	3.5470
64	65	0.099346753747967	$5.4 \cdot 10^{-20}$	4.0899	4.0471
128	129	0.070386092170015	$2.9 \cdot 10^{-39}$	4.5656	4.5471
256	257	0.049819109936140	$8.6 \cdot 10^{-78}$	5.0898	5.0471
512	513	0.035244635485839	$7.5 \cdot 10^{-155}$	5.5664	5.5471

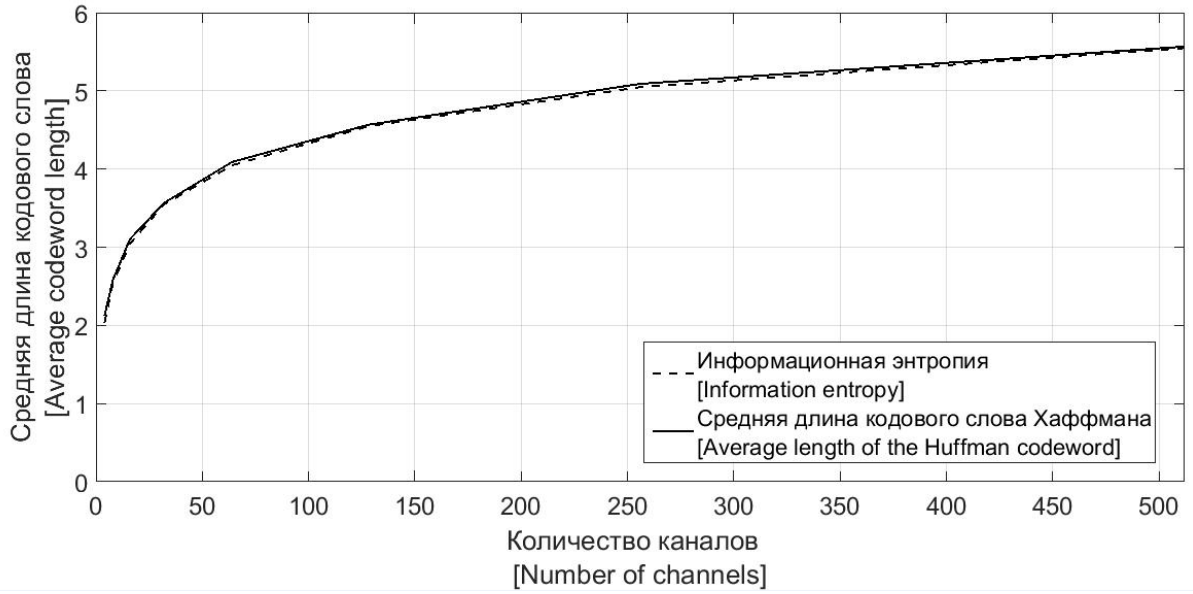

Рис. 2. График зависимости средней длины кодового слова от количества разделяемых каналов N [4].

$$\{\delta_i\} = [0.53063906,\ 0.51355844,\ 0.50235205,\ 0.50042043,\ 0.50009527,\ 0.50002286,\ 0.50000560]. \quad (12)$$

Анализ эмпирически построенной последовательности (12) позволяет утверждать, что с ростом числа разделяемых каналов N приращение в информационной энтропии коэффициентов преобразования Уолша-Адамара (а, значит, и в минимально возможном значении средней длины кодового слова l_{av}) стремится к значению $1/2$ при каждом удвоении величины N.

[3,4] Appendix 1

4. Метод эффективного кодирования коэффициентов преобразования Уолша-Адамара при $N = 4$

Проведенные в настоящей работе исследования позволили установить, что при значении количества разделяемых каналов $N = 4$ применение эффективного кода Хаффмана не является наиболее эффективным способом кодирования коэффициентов преобразования Уолша-Адамара.

В настоящей работе установлено, что для повышения эффективности кодирования коэффициентов преобразования Уолша-Адамара может быть использован специальный класс совершенных алгебраических конструкций — бент-последовательностей.

Определение 3 [25]. Бинарная последовательность $B = [b_0, b_1, \cdots, b_i, \cdots, b_{N-1}]$, где $b_i \in \{\pm 1\}$ — коэффициенты, четной длины $N = 2^{2k}$, $i = 0,1,...,N-1$; N — порядок матрицы Уолша-Адамара, называется бент-последовательностью (БП), если она имеет равномерный по модулю спектр Уолша-Адамара $W_B(\omega)$, который представим в матричной форме

$$S_B(\varpi) = BH,\ \varpi = 0,1,...,N-1, \qquad (13)$$

где H — матрица Уолша-Адамара порядка N.

Исходя из определения, каждый коэффициент преобразования Уолша-Адамара бент-последовательности $S_B(\varpi = 0), S_B(\varpi = 1),..., S_B(\varpi = N-1)$ принимает значения из множества $\left\{\pm\sqrt{N}\right\}$. Таким образом, вектор коэффициентов преобразования Уолша-Адамара бент-последовательности является, по своей сути, бинарной последовательностью, отображенной на алфавит $\left\{\pm\sqrt{N}\right\}$, т.к. каждый его элемент принимает только одно из двух возможных значений, которые отличаются знаком.

Ясно, что коэффициенты преобразования Уолша-Адамара бент-последовательности являются исключительно удобными для внедрения информации в контейнер.

Тем не менее, бент-последовательности являются крайне непредсказуемыми и сложными математическими объектами в силу своей максимально возможной нелинейности.

Сегодня в открытой литературе недоступны не только методы синтеза бент-последовательностей для произвольной длины N, но отсутствует также точная оценка количества бент-последовательностей для длин $N \geq 1024$.

Отметим, однако, что классы бент-последовательностей для длин $N \leq 64$ сегодня достаточно хорошо изучены [25, 26], в частности, разработаны эффективные методы их синтеза.

В табл. 3 представлены мощности классов бент-последовательностей для практически ценных длин, а также процентное содержание бент-последовательностей в полном коде соответствующей длины.

Анализ данных, представленных в табл. 3 позволяет сделать вывод о том, что процент

Таблица 3[5].

Мощности классов бент-последовательностей

N	Мощность полного кода [Cardinality of the complete code]	Количество БП [Number of bent-sequences]	% содержания БП в полном коде [% of the content of bent-sequences in the complete code]
4	$2^4 = 16$	8	50
16	$2^{16} = 65536$	896	1.37
64	$2^{64} = 18446744073709551616$	5425430	$2.94 \cdot 10^{-11}$
256	2^{256}	99270589265934370305785861242880	$8.57 \cdot 10^{-44}$

[5] Appendix 1

содержания бент-последовательностей в полном коде стремительно убывает с возрастанием числа N.

Тем не менее, в случае $N = 4$ мы имеем уникальную ситуацию, когда ровно половина всех последовательностей являются бент-последовательностями, т.е. удовлетворяют условиям **Определения 3**.
Данный факт может быть использован для увеличения количества внедряемой информации в контейнер в стеганосистемах с кодовым разделением каналов при $N = 4$ следующим образом: групповой сигнал может быть подвергнут преобразованию с помощью С-кода еще до нахождения его коэффициентов преобразования Уолша-Адамара.

Определение 4 [17]. С-кодом называется код, каждое кодовое слово которого обладает заданным (обычно, минимально возможным) значением пик-фактора (peak-to-average-power-ratio) κ, который определяется как

$$\kappa = \frac{P_{\max}}{P_{av}} = \frac{1}{N} \max_t \left\{ \left| S_c(t) \right|^2 \right\}, \qquad (14)$$

где $P_{\max}$ — максимальная мощность сигнала $S_c(t)$, P_{av} — средняя мощность сигнала $S_c(t)$.

Много внимания исследователей уделяется проблеме синтеза С-кодов для технологии MC-CDMA [17, 18]. Как показано в данной работе, С-коды могут успешно использоваться в стеганографических системах, основанных на технологии MC-CDMA.

Таким образом, схема стеганосистемы с разделением каналов с использованием технологии MC-CDMA (рис. 1) приобретает вид, показанный на рис. 3.

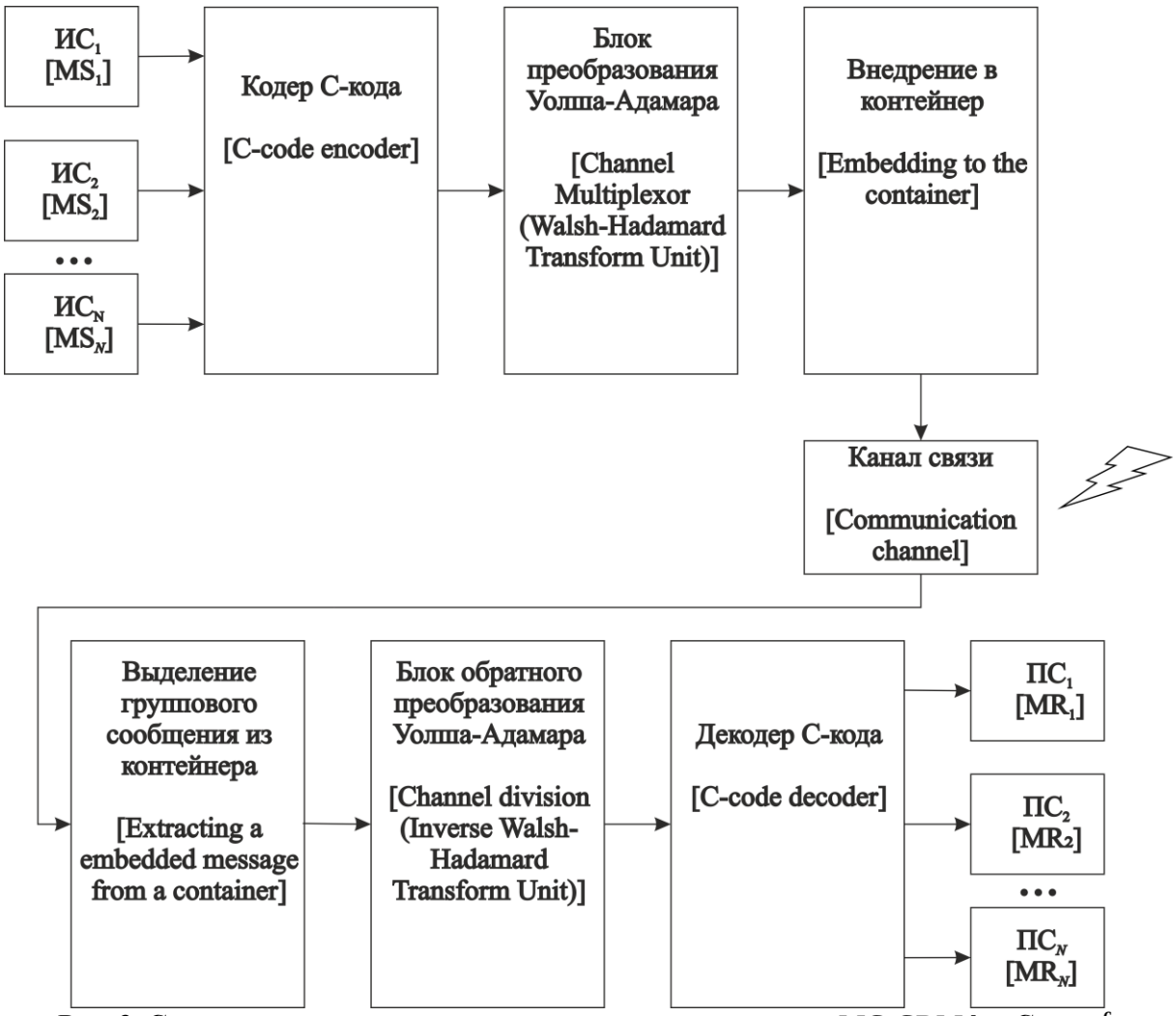

Рис. 3. Схема стеганосистемы с использованием технологии MC-CDMA и С-кода[6].

[6] Appendix 1

При этом предлагается таблица кодирования С-кодом, построенная в соответствии со следующими правилами.

Правило 1. Каждое кодовое слово, соответствующее бент-последовательности, кодируется двумя такими кодовыми словами.

Правило 2. Каждое кодовое слово, не соответствующее бент-последовательности, кодируется кодовым словом, соответствующим бент-последовательности и её инверсии.

Ясно, что при этом исходные кодовые слова могут быть закодированы кодовыми словами С-кода с точностью до перестановок исходных кодовых слов внутри класса бент-последовательностей и оставшихся последовательностей, не являющихся бент-последовательностями.

Один из вариантов кодирования изображен в табл. 4. При этом жирным шрифтом выделены кодовые слова, закодированные в соответствии с *Правилом 1*.

Таблица 4[7].

Способ кодирования кодовых слов полного кода С-кодом

Исходное кодовое слово [Original codeword]	Кодовое слово после преобразования С-кодом [Codeword after coding with C-code]	Исходное кодовое слово [Original codeword]	Кодовое слово после преобразования С-кодом [Codeword after coding with C-code]
0000	1000 1110	0001	**0001 0001**
0011	0100 1011	0010	**0010 0010**
0101	0010 1101	0100	**01000100**
0110	0001 1110	0111	**01110111**
1010	0111 1000	1000	**10001000**
1100	1011 0100	1001	**10011001**
1101	1101 0010	1011	**10111011**
1111	1110 0001	1110	**11101110**

Нетрудно видеть, что кодовое расстояние [8] такого кода равно $d_c = 2$. В виду того, что $d_c = f + 1$, где f — количество выявленных кодом ошибок, представленный в табл. 4 код может выявлять как минимум однократную ошибку, подтверждая тем самым целостность внедренной информации.

При этом средняя длина кодового слова, необходимая для кодирования в каждом кадре одного канала передачи информации составляет 2 бита вместо 2.125 бита, как это необходимо в системе с использованием кодов Хаффмана [6], т.е.:

a. по сравнению с использованием кода Хаффмана для кодирования коэффициентов преобразования Уолша-Адамара, предложенный код позволяет упаковывать коэффициенты Уолша-Адамара на 6,25% эффективнее;

b. в отличие от случая использования кода Хаффмана для кодирования коэффициентов преобразования Уолша-Адамара, предложенный код позволяет детектировать как минимум одну возникшую ошибку, подтверждая целостность внедренной информации.

ВЫВОДЫ

В настоящей статье разработан теоретический базис для эффективного кодирования группового сигнала в стеганосистемах с множественным доступом на основе технологии MC-CDMA, в рамках чего получены следующие теоретически и практически ценные результаты:

1. Выведены и доказаны соотношения, определяющие как возможные значения коэффициентов преобразования Уолша-Адамара для заданного значения числа разделяемых каналов N, так и вероятности появления заданных значений коэффициентов преобразования Уолша-Адамара. Указанные соотношения позволяют строго теоретически рассчитать вероятностные характеристики алфавита коэффициентов преобразования Уолша-Адамара для произвольного значения числа разделяемых каналов N, что исключает необходимость набора соответствующей статистики переборным методом при построении эффективных кодов.

2. Для стеганографических систем, использующих технологию MC-CDMA проведены исследования зависимости

[7] Appendix 1

средней длины кодового слова от числа разделяемых каналов N. Для практически ценных значений количества разделяемых каналов N определены средние длины кодового слова Хаффмана. Эмпирически показано, что с ростом числа разделяемых каналов N приращение в информационной энтропии коэффициентов-преобразования Уолша-Адамара стремится к значению $1/2$ при каждом удвоении величины N.

3. Для практически ценного числа разделяемых каналов $N = 4$ предложен эффективный метод кодирования коэффициентов преобразования Уолша-Адамара с помощью кодовых слов С-кода, основанного на бент-последовательностях. По сравнению с использованием кода Хаффмана для кодирования коэффициентов преобразования Уолша-Адамара, предложенный код позволяет упаковывать коэффициенты Уолша-Адамара на 6,25% эффективнее, а также детектировать как минимум одну возникшую ошибку, подтверждая целостность внедренной информации. Таким образом при использовании предложенного эффективного метода кодирования избыточность, вносимая в информацию, передаваемую по стеганоканалу, расходуется не только на уплотнение каналов, но и на добавление корректирующих свойств.

APPENDIX 1 (ПРИЛОЖЕНИЕ 1)

[1]**Fig. 1.** Block diagram of a multiple-access steganographic system using MC-CDMA technology.

In Fig. 1 and Fig. 3 the following designations are accepted: MS is the Message Sender, MR is the Message Receiver.

[2]**Table 1.** Huffman code for Walsh-Hadamard transform coefficients for $N = 16$.

[3]**Table 2.** The probabilistic and informational characteristics of the Walsh-Hadamard transform coefficients depending on the value of N.

[4]**Fig. 2.** Graph of the dependence of the average length of the codeword from the number of divided channels N.

[5]**Table 3.** Cardinalities of classes of bent-sequences.

[6]**Fig. 3.** Diagram of a steganographic system using MC-CDMA technology and C-code.

[7]**Table 4.** Method of encoding of codewords of the complete code with C-code.

Литература (References)

[1] Provos N., Honeyman P. Hide and seek: An introduction to steganography. IEEE security & privacy. 2003. Vol. 1. No. 3. P. 32-44.

[2] Morkel T., Eloff J. H. P., Olivier M. S. An overview of image steganography. ISSA. 2005. Vol. 1. No. 2. P. 322-330.

[3] Cheddad A., Condell J., Curran K., Mc.Kevitt P. Digital image steganography: Survey and analysis of current methods. Signal processing. 2010. Vol. 90, No. 3. P. 727-752.

[4] Johnson N. F., Duric Z., Jajodia S. Information hiding: Steganography and watermarking – attacks and countermeasures. Kluwer Academic Publishers, 2000. 137 p.

[5] Chandramouli R., Memon N. Analysis of LSB based image steganography techniques. Proceedings 2001 International Conference on Image Processing. IEEE, 2001. Vol. 3. P. 1019-1022.

[6] Karim S. M. M., Rahman M. S., Hossain M. I. A new approach for LSB based image steganography using secret key. International conference on computer and information technology, 2011. P. 286-291.

[7] Raja K. B., Chowdary C. R., Venugopal K. R., Patnaik L. M. A secure image steganography using LSB, DCT and compression techniques on raw images. International conference on intelligent sensing and information processing, 2005. P. 170-176.

[8] Akhtar N., Johri P., Khan S. Enhancing the security and quality of LSB based image steganography. International Conference and Computational Intelligence and Communication Networks. IEEE, 2013. P. 385-390.

[9] Qazanfari K., Safabakhsh R. A new steganography method which preserves histogram: Generalization of LSB++. Information Sciences. 2014. Vol. 277. P. 90-101.

[10] Patel H., Dave P. Steganography technique based on DCT coefficients. International Journal of Engineering Research and Applications. 2012. Vol. 2. No. 1. P. 713-717.

[11] Bhattacharyya S., Mondal S., Sanyal G. A Robust Image Steganography using Hadamard Transform. International Conference on Information Technology in Signal and Image Processing, Mumbai, 2013. P. 416-426.

[12] Wang S., Yang B., Niu X. A secure steganography method based on genetic algorithm. Journal of Information Hiding and Multimedia Signal Processing. 2010. Vol. 1. No. 1. P. 28-35.

[13] Roy R., Sarkar A., Changder S. Chaos based edge adaptive image steganography. Procedia Technology. 2013. Vol. 10. P. 138-146.

[14] Sheidaeian H., Zolfaghari B., Zobeiri M. An Efficient and Secure Approach to Multi-User Image Steganography Using CRC-Based CDMA. International Conference on Signal Acquisition and Processing. Singapore, 2011. Vol. 2. P. 1-5.

[15] Amirtharajan R., Rayappan J. B. B. Covered CDMA multi-user writing on spatially divided image. International Conference on Wireless Communication, Vehicular Technology, Information Theory and Aerospace & Electronic Systems Technology (Wireless VITAE), 2011. P. 1-5.
[16] Czvetkov K. Yu., Fedoseev V. E., Korovin V. M., Abazina E. S. Model kodera skrytogo kanala s kodovym uplotneniem s ispolzovaniem signalnykh posledovatelnostej Franka-Uolsha, Franka-Krestensona [Model of a covert channel code division multiplex coder using Frank-Walsh, Frank-Chrestenson signaling sequences]. Trudy Nauchno-issledovatelskogo instituta radio [Proceedings of the Radio Research Institute]. 2015. No. 1. P. 2-11. (In Russian).
[17] Paterson K. G. On codes with low peak-to-average power ratio for multicode CDMA. HP Laboratories Technical Report HPL-2001-115, May 2001. P. 1-16.
[18] Paterson K. G. Sequences for OFDM and Multicode CDMA: two problems in algebraic coding theory, Sequences and their applications. Seta 2001. Second Int. Conference (Bergen, Norway, May 13–17, 2001). Proc. Berlin: Springer, 2002, P. 46-71.
[19] Wada T., Yamazato T., Katayama M., Ogawa A. A constant amplitude coding for orthogonal multi-code CDMA systems. IEICE TRANSACTIONS on Fundamentals of Electronics, Communications and Computer Sciences. 1997. Vol. 80, No. 12. P. 2477-2484.
[20] Schmidt K. Quaternary Constant-Amplitude Codes for Multicode CDMA, IEEE International Symposium on Information Theory. Nice, 2007, P. 2781-2785.
[21] McEliece R. The theory of information and coding. Cambridge University Press, 2002. 400 p.
[22] Ahmed N., Rao K. R. Walsh-Hadamard transform. Orthogonal transforms for digital signal processing. Springer, Berlin, Heidelberg, 1975. P. 99-152.
[23] Mazurkov M. I. Sistemy shirokopolosnoj radiosvyazi [Broadband radio systems]. Odesa: Nauka i Tekhnika [Odesa: Science and Technology], 2010. 340 p. (In Russian).
[24] Mazurkov M. I. Osnovi teoriyi peredavannya informacziyi [Fundamentals of information transmission theory]. Odesa: Nauka i Tekhnika [Odesa: Science and Technology], 2005. 168 p. (In Ukrainian).
[25] Tokareva N. N. Bent-funkczii: rezultaty i prilozheniya. Obzor rabot [Bent-functions: results and applications. Review of works]. Priklad. diskret. matematika [Applied discrete mathematics]. 2009. No. 1(3). P. 15-37. (In Russian).
[26] Langevin P., Leander G. Counting all bent functions in dimension eight 99270589265934370305785861242880. Designs, Codes and Cryptography. 2011. Vol. 59. No. 1. P. 193-205.

Сведения об авторах.

Кобозева Алла Анатольевна. Одесский национальный политехнический университет. Кафедра кибербезопасности и программного обеспечения, заведующая кафедрой, доктор технических наук, профессор. Область научных интересов: информационная безопасность, в частности, стеганография, экспертиза целостности информационного контента.
E-mail: alla_kobozeva@ukr.net

Соколов Артем Викторович. Одесский национальный политехнический университет. Кафедра кибербезопасности и программного обеспечения, доцент, кандидат технических наук. Область научных интересов: методы защиты информации на основе совершенных алгебраических конструкций.
E-mail: radiosquid@gmail.com