УДК 681.3

СРАВНИТЕЛЬНЫЙ АНАЛИЗ ВЕРОЯТНОСТИ ОШИБОК ДЕКОДИРОВАНИЯ КОДОВЫХ КОНСТРУКЦИЙ RIJNDAEL И БЧХ

Д.М. БИЛЬДЮК, С.Б. САЛОМАТИН

Белорусский государственный университет информатики и радиоэлектроники
П. Бровки, 6, Минск, 220013, Беларусь

Поступила в редакцию 27 декабря 2013

Рассматривается нелинейный помехоустойчивый код на базе алгоритма криптографического преобразования данных *Rijndael*. Произведен сравнительный анализ декодирования кодов *Rijndael* и БЧХ.

Ключевые слова: криптографическое преобразование данных, помехоустойчивое кодирование, алгоритм *Rijndael*, границы помехоустойчивого кодирования, нелинейный код.

Введение

Нелинейный двоичный код на основе алгоритма *Rijndael* (далее *R*-код) кодовой длины n и информационной длины k обладает как криптографическими, так и помехоустойчивыми свойствами [1–3]. Декодирование такого кода осуществляется переборными методами – либо по принципу максимального правдоподобия, либо на основе перебора возможных ошибок (например, на основе декодеров Чейза) [1]. Существование других вариантов декодирования свидетельствовало бы о криптографической уязвимости используемого в основе кодирования преобразования [4].

Недостатком алгоритма используемого в декодере максимального правдоподобия (далее ДМП) является его высокая вычислительная сложность (по сравнению с известными алгоритмами декодирования для линейных кодов) [4]. Однако ДМП осуществляет декодирование за границей корректирующей способности используемого кода, определяемой минимальным кодовым расстоянием в метрике Хэмминга (далее $d_{\min}$) [5]. Тогда актуальной является задача сравнительного анализа помехоустойчивых свойств *R*-кода со свойствами известных двоичных линейных кодов.

Среди обширного класса помехоустойчивых кодов БЧХ существуют хорошие (с точки зрения дистанционных свойств) двоичные коды [6]. Структура таких кодов позволяет определить их как коды с алгебраическим декодированием. Реализация алгебраического декодера (далее АД) может осуществляться на основе ряда алгоритмов [6], эффективность которых зависит от конкретной аппаратно-программной платформы. Тогда логичным был бы сравнительный анализ помехоустойчивых свойств *R*-кода с ДМП, БЧХ-кода с АД в пределах границы декодирования (на основе алгоритма использующего понятие полинома локаторов ошибок, например, алгоритма Берлекемпа-Месси) и БЧХ-кода с ДМП. Для оценки помехоустойчивых свойств кода принято использовать зависимость вероятности ошибки на информационный бит от отношения сигнал/шум определяемого как отношение энергии одного бита кодового слова к спектральной плотности мощности шума в пределах ширины спектра передаваемого сигнала в канале с аддитивным белым гауссовским шумом (далее АБГШ) [6]. В качестве сигнала чаще всего используют сигнал с двоичной фазовой манипуляцией (далее ДФМ). Отраженные в статье результаты экспериментов получены при использовании такого сигнала с жесткой демодуляцией и частотой дискретизации $10f_0$ (f_0 – несущая частота), 12-разрядным квантованием и с размещением десяти периодов несущей частоты в одном бите.

Схема формирования двоичного *R*-кода

Двоичный *R*-код с параметрами $(n,k,d_{\min})$ определяется как множество отображенных *k*-мерных векторов $\mathbf{a}=(a_0,a_1,\ldots,a_{k-1})\in GF(2^k)$ в другое множество *n*-мерных векторов $\mathbf{c}=(c_0,c_1,\ldots,c_{n-1})\in GF(2^n)$, $k<n$, с минимальным расстоянием Хэмминга среди всех возможных пар кодовых слов – $d_{\min}$ [1]. Отображение реализуется на основе криптографического алгоритма $Rijndael_{m,n}$ с ключом шифрования $\mathbf{s}=(s_0,s_1,\ldots,s_{m-1})\in GF(2^m)$, вектором избыточности $\mathbf{v}=(v_0,v_1,\ldots,v_{r-1})\in GF(2^r)$, $r=n-k$, и задается функцией $\varphi(\mathbf{a},\mathbf{s},\mathbf{v}):GF(2^k)\to GF(2^n)$ [1]:

$$\mathbf{c}\leftarrow\varphi(\mathbf{a},\mathbf{s},\mathbf{v}):(\mathbf{a},\mathbf{s},\mathbf{v})\to Rijndael_{m,n}(\mathbf{a}\,|\,\mathbf{v},\mathbf{s}). \qquad (1)$$

Параметры *m* и *n* (длина ключа шифрования и длина блока шифрования соответственно) в основном режиме (режиме электронной кодовой книги) криптографического преобразования *Rijndael* могут принимать фиксированные значения 128, 192 и 256 [6]. Для формирования *R*-кода произвольной длины *n* необходимо использовать режимы криптографического преобразования с обратной связью длины в один бит, вектор инициализации $\mathbf{iv}=(iv_0,iv_1,\ldots,iv_{m-1})\in GF(2^u)$, $u\in\{128,192,256\}$, можно считать частью ключа. Тогда отображение задается функцией $\psi(\mathbf{a},\mathbf{s},\mathbf{v},\mathbf{iv}):GF(2^k)\to GF(2^n)$ [1]:

$$\begin{aligned}
&\mathbf{c}\leftarrow\psi(\mathbf{a},\mathbf{s},\mathbf{v},\mathbf{iv}):(\mathbf{a},\mathbf{s},\mathbf{v},\mathbf{iv})\to(\\
&\quad\mathbf{o}\leftarrow\mathbf{iv};\ \mathbf{a}'\leftarrow\mathbf{a}\,|\,\mathbf{v};\\
&\quad\textbf{for } i \textbf{ from } 0 \textbf{ to } n-1 \textbf{ do}\\
&\qquad c_i\leftarrow\left(Rijndael_{m,u}(\mathbf{o},\mathbf{s})\right)_{u-1}\oplus a'_i;\ \mathbf{o}\leftarrow(o_1,o_2,\ldots o_{u-1},c_i);\\
&\quad\textbf{enddo};\\
&\quad\textbf{return } \mathbf{c};\ ).
\end{aligned} \qquad (2)$$

Отображения при помощи функций (1) и (2) задают помехоустойчивый код с близкими дистанционными свойствами [2].

Моделирование системы связи с помехоустойчивым кодированием и каналом связи с АБГШ

Существует теоретическая оценка верхней границы вероятности ошибки на бит P_{eb} информационного слова $\mathbf{a}$ в зависимости от E_b/N_0 (отношение энергии одного бита кодового слова E_b к спектральной плотности мощности шума в пределах ширины спектра передаваемого сигнала N_0) в канале с АБГШ для блокового группового двоичного кода $(n,k,d_{\min})$, $d_{\min}=2t+1$, t – гарантированная максимальная кратность исправляемой ошибки в кодовом слове [5]. Для определенных во введении условий данная оценка определяется неравенством [6]:

$$P_{eb}\le\frac{1}{n}\sum_{i=t+1}^{n}(t+i)C_n^i p^i(1-p)^{n-i}, \qquad (3)$$

где $C_n^i=\dfrac{n!}{i!(n-i)!}$, $p=Q\left(\sqrt{\dfrac{2E_bk}{N_0n}}\right)$, $Q(x)=\dfrac{1}{\sqrt{2\pi}}\int_x^\infty e^{-\frac{t^2}{2}}dt$.

Для некодированного ДФМ используется оценка [6]:

$$P_{eb}=Q\left(\sqrt{\frac{2E_b}{N_0}}\right). \qquad (4)$$

Оценки (3) и (4) используются для определения достоверности модели системы связи, структурная схема которой изображена на рис. 1.

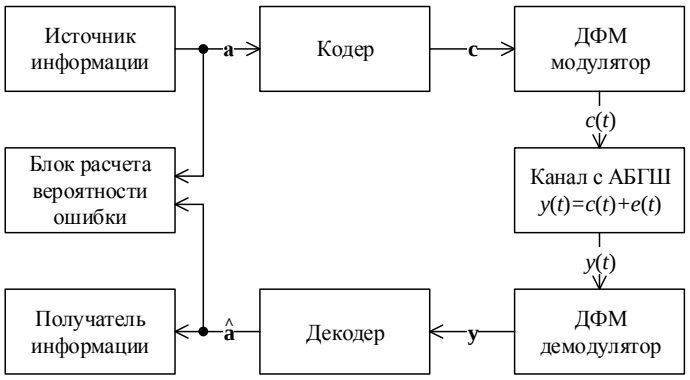


Рис. 1. Структурная схема модели системы связи с помехоустойчивым кодированием и каналом связи с АБГШ при использовании ДФМ

Источник информации формирует информационные равновероятные слова $\mathbf{a}$ в виде двоичных векторов из $GF(2^k)$. Кодер может отсутствовать (в случае системы связи с ДФМ без помехоустойчивого кодирования, далее – некодированный ДФМ-сигнал), либо реализовывать функцию избыточного кодирования, осуществляющую отображение $GF(2^k) \to GF(2^n)$ для формирования двоичного вектора кодового слова $\mathbf{c}$. В случае использования *R*-кода кодер реализует функции (1) или (2), а для кодов БЧХ может использоваться функция умножения на порождающий полином [6].

ДФМ-модулятор формирует дискретные комплексные отсчеты фазоманипулированного сигнала $c(t)$ используя кодовые слова в качестве модулирующих последовательностей. ДФМ сигнал поступает в канал с АБГШ, где формируется искаженный сигнал $y(t) = c(t) + e(t)$. Отсчеты АБГШ $e(t)$ рассчитываются исходя из ширины спектра ДФМ сигнала и в соответствии с заданным отношением E_b / N_0 [6]. ДФМ демодулятор осуществляет побитовую демодуляцию принятого сигнала $y(t)$ с жесткими (двоичными) решениями в виде вектора $\mathbf{y} \in GF(2^n)$ на выходе [6].

Декодер (также может отсутствовать в случае некодированного ДФМ-сигнала) использует внесенную в кодере избыточность для исправления ошибок, произошедших в результате искажений в канале, при этом предполагается наличие двоичного вектора ошибок $\mathbf{e} \in GF(2^n)$, такого что $\mathbf{y} = \mathbf{c} + \mathbf{e}$ [5]. В случае *R*-кода используется ДМП, а в случае БЧХ-кода может быть использован как ДМП, так и АД. На выходе декодера формируется оценка переданного информационного вектора $\hat{\mathbf{a}}$ [5].

Блок расчета вероятности ошибки P_{eb} вычисляет отношение количества ошибочных бит в потоке информационных слов к общему количеству переданных бит.

Разработка блоков кодирования и ДМП для кодов *R* и БЧХ под среду *Simulink*

Библиотека *Simulink* не содержит стандартных блоков кодирования и декодирования *R*-кода, а также ДМП для кодов БЧХ. Разработка таких блоков осуществляется реализацией программ на основе *MEX*-файлов и динамических библиотек, подключаемых к ним, на языке программирования *C* [7]. Кодер *R*-кода разработан на основе реализации функций (1) и (2), кодер БЧХ реализуется как функция $bch(a(x), g(x)) : a(x) \to c(x)$ [6]:

$$c(x) \leftarrow bch(a(x), g(x)) : (a(x), g(x)) \to a(x)g(x), \qquad (5)$$

где $g(x)$ – порождающий полином кода, $deg(a(x)) = n-k$; $a(x) = \sum_{i=0}^{k-1} a_i x^i$ – информационный полином, коэффициенты $a_i \in GF(2)$ интерпретируются как координаты информационного вектора $\mathbf{a} = (a_0, a_1, \ldots, a_{k-1}) \in GF(2^k)$; $c(x) = \sum_{i=0}^{n-1} c_i x^i$ – кодовый полином, коэффициенты $c_i \in GF(2)$ интерпретируются как координаты кодового вектора $\mathbf{c} = (c_0, c_1, \ldots, c_{n-1}) \in GF(2^n)$. В векторном виде функция (5) реализуется на основе модели регистров сдвига с конечным откликом $bch(\mathbf{a}, \mathbf{g}) : GF(2^k) \to GF(2^n)$ [6]:

$$
\begin{aligned}
&\mathbf{c} \leftarrow bch(\mathbf{a}, \mathbf{g}) : (\mathbf{a}, \mathbf{g}) \to (\\
&\quad \mathbf{a}' \leftarrow (a_0, a_1, \ldots, a_{k-1}, 0_k, 0_{k+1}, \ldots 0_{n-1});\ \mathbf{s} \leftarrow (0_0, 0_1, \ldots 0_{n-k-1});\\
&\quad \textbf{for } i \textbf{ from } 0 \textbf{ to } n-1 \textbf{ do}\\
&\qquad c_i \leftarrow a'_i\, g_{n-k} \oplus s_0 g_{n-k-1} \oplus s_1 g_{n-k-2} \oplus \ldots \oplus s_{n-k-1} g_0;\\
&\qquad \mathbf{s} \leftarrow (a'_i, s_0, s_1, \ldots, s_{n-k-2});\\
&\quad \textbf{end do};\\
&\quad \textbf{return } \mathbf{c};\ ),
\end{aligned}
\tag{6}
$$

где $\mathbf{g} = (g_0, g_1, \ldots g_{n-k})$ – вектор коэффициентов порождающего полинома $g(x) = \sum_{i=0}^{n-k} g_i x^i$.

ДМП для БЧХ- и *R*- кодов осуществляет поиск ближайшего к принятому вектору $\mathbf{y}$ среди всех $M = 2^k$ кодовых слов. Расчет расстояния в метрике Хэмминга осуществляется как оценка веса разности двух векторов, и в двоичном поле реализуется при помощи функции [6]:

$$
d \leftarrow dist(\mathbf{y}, \mathbf{c}_i) : (\mathbf{y}, \mathbf{c}_i) \to wt(\mathbf{y} \oplus \mathbf{c}_i),
\tag{7}
$$

где d – расстояние Хэмминга; wt – вес Хэмминга; $i = 0,1,\ldots M-1$.

Перебор всех возможных кодовых слов ведет к высокой вычислительной сложности ДМП с ростом параметра k. Операции вычисления расстояний, на основе функции (7), до всех кодовых слов независимы и могут быть вычислены при помощи параллельных вычислительных средств, что приведет к снижению временных затрат декодирования искаженного кодового слова. Среди таких средств существуют решения на основе технологии *CUDA* выделяющиеся, среди существующих параллельных вычислительных систем, низкой стоимостью вычислений и использующиеся как сопроцессорный ускоритель вычислений [7]. Архитектура вычислительных средств *CUDA* (далее просто – *CUDA*) относится к классу *SIMD* (*Single Instruction, Multiple Data* – одиночный поток команд, множественный поток данных) [7]. Программный код приложений для *CUDA* реализуется на языке *C* путем разработки последовательного алгоритма, копии которого исполняются во множестве потоков для множества различных данных параллельно. Алгоритм ДМП для двоичного кода $\mathbf{C} = \{\mathbf{c}_i \mid i = 0,1,\ldots M-1\}$ в метрике Хэмминга задается функцией:

$$
\begin{aligned}
&\hat{\mathbf{a}} \leftarrow DMP(\mathbf{y}, da_{\min}) : (\mathbf{y}, da_{\min}) \to (\\
&\quad \mathbf{a}_i \leftarrow binary(i); && (for\ \forall i)\\
&\quad \mathbf{c}_i \leftarrow enc(\mathbf{a}_i); && (for\ \forall i)\\
&\quad d_i \leftarrow dist(\mathbf{y}, \mathbf{c}_i); && (for\ \forall i)\\
&\quad da_i \leftarrow (d_i \mid i); && (for\ \forall i)\\
&\quad atomicMin(da_{\min}, da_i); && (for\ \forall i)\\
&\quad i_{\min} \leftarrow (d_{\min} \mid i_{\min}) \leftarrow da_{\min};\\
&\quad \hat{\mathbf{a}} \leftarrow binary(i_{\min});\\
&\quad \textbf{return } \hat{\mathbf{a}};\ ),
\end{aligned}
\tag{8}
$$

где $da_{\min}$ – общая для всех потоков память для записи номера потока i с наименьшим расстоянием d_i, $i = 0,1,\ldots M-1$; $binary(i)$ – функция преобразования числа i в двоичный вектор длины k; $enc(\mathbf{a})$ – функция кодирования информационного вектора $\mathbf{a}$, для *R*-кода реализуется при помощи функции (1) либо (2), для БЧХ-кода при помощи функции (5) либо (6); $atomicMin(x, y)$ – функция атомарного доступа к памяти хранящей переменную x и реализующей запись в эту память наименьшей из двух переменных x и y [8]. Модель системы на основе *R*-кодирования в среде Simulink с реализацией функции (8) для *CUDA* в ДМП изображена на рис. 2.

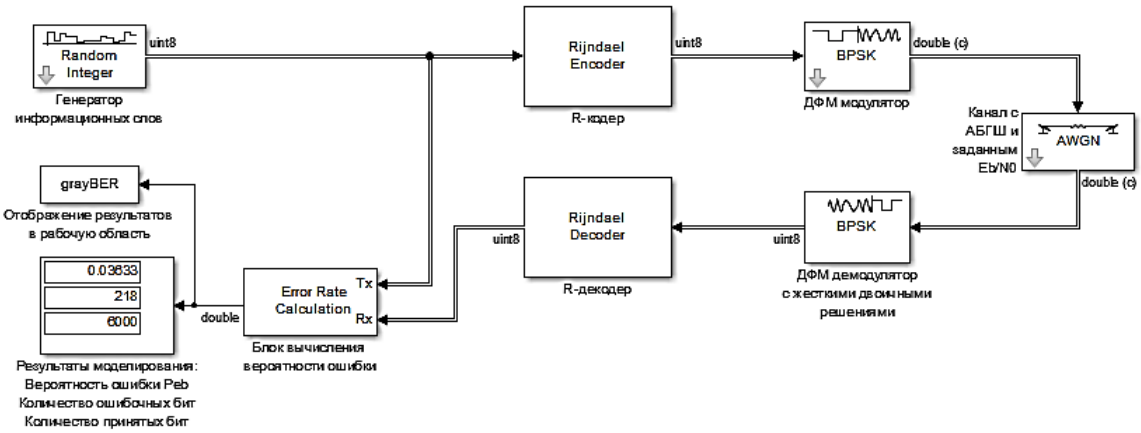


Рис. 2. Модель системы связи с *R*-кодированием и ДМП на основе *CUDA* в среде *Simulink* пакета *MatLab*

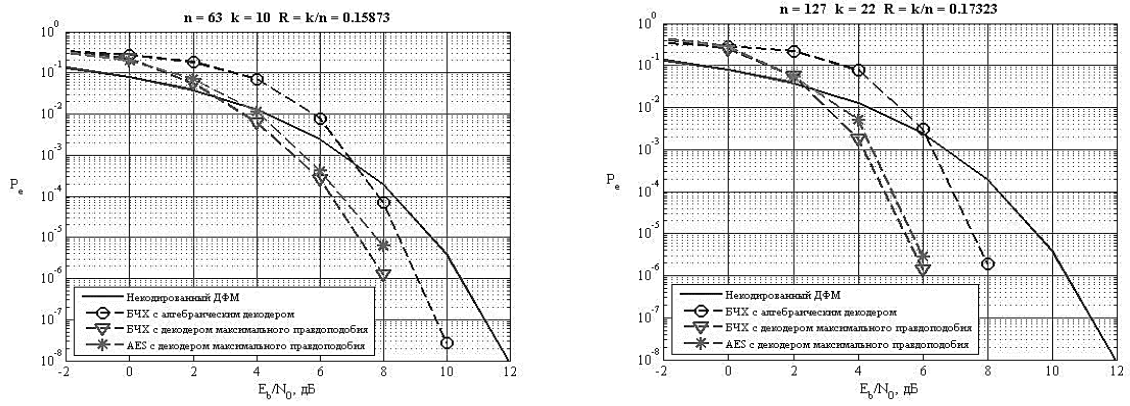
Оценка вероятности ошибки кодов *R* и БЧХ в канале с АБГШ

Сравнительный анализ эффективности корректирующих кодов *R* и БЧХ осуществляется с использованием рассмотренных выше моделей (рис. 1 и 2) на основе зависимости P_{eb} от отношения E_b / N_0. Параметры выбранных для сравнения кодов, представлены в таблице.

Параметры *R* и БЧХ кодов

n	7	15	15	15	31	31	31	31	63	63	63	63	63	127	127	127
k	4	5	7	11	6	11	16	21	7	10	16	18	24	8	15	22
$d_{\min}$ (БЧХ)	3	7	5	3	15	11	7	5	31	27	23	21	15	63	55	47
$d_{\min}$ (*R*)	2	4	2	1	8	3	1	1	19	15	9	7	2	44	32	20

Формирование БЧХ осуществлялось с использованием функции (6), а *R*-кода – с использованием функции (2), вектора избыточности $\mathbf{v}$ и инициализации $\mathbf{iv}$ использованы с координатами равными нулю. Вектор $\mathbf{s}$ – случайный. Примеры результатов моделирования с использованием параметров, приведенных в таблице представлены на рис. 3.



а *б*

Рис. 3. Примеры результатов моделирования с использованием *R* и БЧХ-кодов с параметрами *а* – (63,10) и *б* – (127,22)

Представленные результаты показывают, что:

– ДМП осуществляет декодирование за границей корректирующей способности помехоустойчивого кода;

– существуют *R*-коды с зависимостью P_{eb} от отношения E_b / N_0 близкой к зависимости БЧХ-кодов при одинаковых параметрах (n, k), в случае использования ДМП (заметим, что $d_{\min}$ *R*-кода меньше чем у БЧХ-кода – см. таблицу).

Анализ полученных данных для всех кодов таблицы показывает, что расстояние между характеристиками $P_{eb}(E_b / N_0)$ двоичного БЧХ-кода при использовании АД и ДМП зависит как от длины кода, так и от скорости (рис. 4).

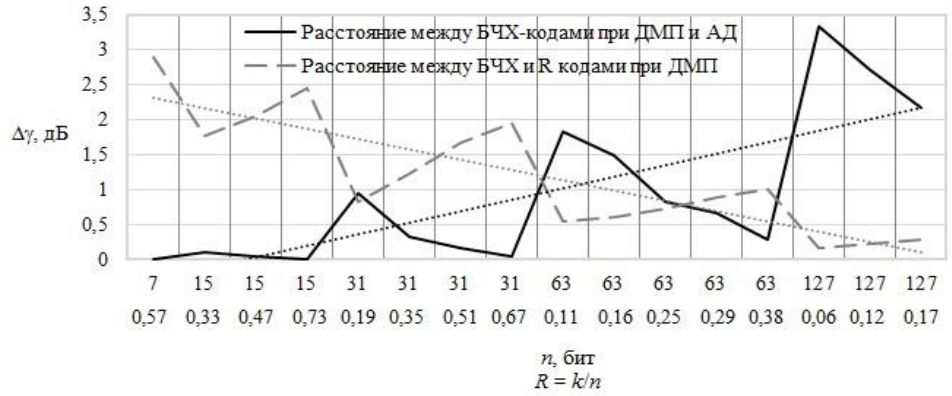


Рис. 4. Зависимость расстояния $\Delta\gamma$ по параметру отношения $\gamma = E_b / N_0$ на уровне $P_{eb} = 10^{-6}$ от длины кода n и его скорости $R = k / n$ для БЧХ-кода при ДМП и АД и *R*-кода по сравнению с БЧХ-кодом при ДМП

Из зависимостей, представленных на рис. 4, видно, что с ростом скорости кодов расстояние между характеристиками $P_{eb}(E_b / N_0)$ для БЧХ-кодов при ДМП и АД сокращается – БЧХ-код приближается к коду Хэмминга, относящийся к классу совершенных кодов, для которого характерно совпадение указанных характеристик [6]. Также увеличение длины БЧХ-кода ведет к большему количеству возможных кодов между кодом с минимально возможной скоростью и кодом Хэмминга, но всегда $\Delta\gamma$ достигает нуля при совпадении с последним. Для *R*-кода характерно отдаление от границы максимального правдоподобия БЧХ-кода с ростом скорости вследствие уменьшения комбинаторных возможностей псевдослучайного выбора кодовых слов. С ростом длины *R*-кода указанные комбинаторные возможности увеличиваются, что ведет к приближению к границе максимального правдоподобия БЧХ-кода.

Выводы

Существуют двоичные *R*-коды с границами максимального правдоподобия (по параметру вероятности ошибки на бит в зависимости от отношения сигнал шум в канале с АБГШ) близкими к аналогичными границам БЧХ-кодов, преимуществами которых являются дополнительные криптографические свойства. Указанные *R*-коды обнаруживаются малых скоростях формируемых кодовых конструкций, что более выгодно (с точки зрения быстродействия) для ДМП.

ANALYSIS OF DECODERS BIT ERROR FOR RIJNDAEL AND BCH CODES

D.M. BILDZIUK, S.B. SALOMATIN

Abstract

The nonlinear error control code on the basis of cryptographic transformation of data through Rijndael algorithm is considered. Bit error properties of decoding algorithms of Rijndael codes and BCH codes for AWGN channel are compared.

Список литературы

1. *Бильдюк Д.М., Саломатин С.Б.* // Докл. БГУИР. 2012. № 8 (70). С. 75–80.
2. *Бильдюк Д.М., Саломатин С.Б.* // Докл. БГУИР. 2012. № 7 (63). С. 105–109.
3. *Elumalai R., Reddy A.R.* // Int. J. of Scientific Research. 2011. Vol. 2, Iss 3.
4. *Фомичев В.М.* Дискретная математика и криптология. М., 2003.
5. *MacWilliams F.J., Sloane N.J.A.* The theory of error correcting codes. North-Holland, 1977.
6. *Peter Sweeney*. Federal Information Processing Standards Publication 197. USA, 2001.
7. *Farber. R.* CUDA Application Design and Development. MK, 2011.