

П. С. Поперечный, аспирант ИППМ РАН, e-mail: prpoperechny@elvees.com
Институт проблем проектирования в микроэлектронике РАН

Сравнительная оценка корректирующей способности кодов Рида—Соломона и Боуза—Чоудхури—Хоквингема

Приведены сравнительные характеристики кодов Рида—Соломона (РС) и Боуза—Чоудхури—Хоквингема (БЧХ), которые находят широкое применение, в том числе для помехоустойчивого хранения данных. Предложен способ сравнения корректирующей способности кодов исправлять пакетные ошибки. Приведено аналитическое выражение для построения передаточной функции декодеров РС и БЧХ для двоичного симметричного канала (ДСК). Показано, что при одинаковых длинах кодированных данных и скоростях кода кодек БЧХ исправляет лучше случайно распределенные ошибки, однако проигрывает кодеку РС, если ошибки сгруппированы. Также проведен синтез логической схемы двух кодеков, приведено их сравнение по занимаемой площади и задержек на критическом пути.

Ключевые слова: РС, БЧХ, корректирующая способность, ДСК, поле Галуа

Введение

Помехоустойчивые коды Рида—Соломона (РС) и Боуза—Чоудхури—Хоквингема (БЧХ) находят широкое применение в системах передачи-приема информации, хранения данных и т. д. Данные коды относятся к блочным помехоустойчивым кодам, кодирующим относительно большие блоки данных и способным исправлять множественные ошибки в этих блоках [1, 2].

Ввиду того, что данные коды с самого начала широко применялись в системах передачи-приема информации в радиосвязи, существуют различные способы оценки корректирующей способности кодов, основанные на математических моделях каналов связи [1]. Классификация каналов связи — достаточно широкая в зависимости от применения, однако все модели наиболее приближены к системам передачи информации: двоичный симметричный канал (ДСК), канал с аддитивным белым гауссовым шумом (АБГШ), канал с замиранием и т. д. Модели каналов делятся на непрерывные и дискретные, с памятью и без памяти. Например, дискретный канал без памяти — это такой канал, для которого выходной символ в данный момент статистически зависит лишь от соответствующего входного символа. Дискретные каналы с памятью — это такие каналы, в которых выход в данный момент статистически зависит как от текущего входа, так и от предыдущих входов и выходов [3].

В статье рассмотрены коды РС и БЧХ применительно к устройствам хранения информации, таким как flash-память, SSD-диск [4], для которых известные модели каналов связи не подходят ввиду разной природы ошибок.

Зачастую составить полную модель канала сложно из-за различных видов искажений [5, 6]. Во flash-памяти причинами искажений бита данных могут быть: цикл записи/стирания, парное соединение битов, ошибки сохранения данных, инверсия битов, случайно распределенные дефектные блоки, ограни-

ченный ресурс, уменьшение технологии производства и применение многоуровневой ячейки и, как следствие, уменьшение уровня напряжения на бит [7]. Ввиду множества причин искажения информации не исключено появление пакетных ошибок. При этом такие способы, как переспрос или перемеживание, могут оказаться неэффективными. В статье предложена модель канала с пакетными ошибками и показано преимущество кодов РС в этом случае.

Эквивалентные коды БЧХ и РС для применения в носителях информации

Согласно стандарту ONFI [8], flash-память делится на область хранения непосредственно информационных данных и на дополнительную область (spare-область) для хранения контрольных битов. Выберем наиболее подходящие и эквивалентные коды РС и БЧХ в части избыточности и длины кодированных данных.

Коды РС являются недвоичными, для систем хранения данных предпочтительнее, чтобы символ представлялся байтом [9]. Таким образом, в поле Галуа $GF(2^8)$ выберем для рассмотрения код РС(255, 239) с числом битов на символ $m = 8$. То есть $n = 255$ — число закодированных символов, $k = 239$ — число информационных символов. Данный код позволяет исправлять восемь символов в кодовом слове. Каждый символ является 8-битным. На рис. 1 показана структура кода и одна из

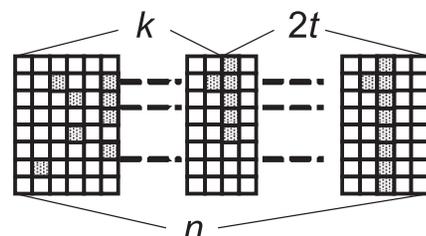


Рис. 1. Структура кода РС с возможными ошибками

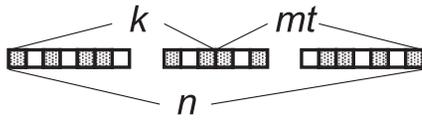


Рис. 2. Структура кода БЧХ с возможными ошибками

комбинаций ошибок (черным выделены ошибочные биты), которую может исправить данный код.

Символы могут быть испорчены частично либо полностью, любую комбинацию код РС(255, 239) способен исправить, если число испорченных символов не более $t = 8$. Если представить этот код в битах, т. е. в поле $GF(2)$, то получится код $(n \cdot m, k \cdot m) = (255 \cdot 8, 239 \cdot 8) = (2040, 1912)$.

Эквивалентный код по скорости и длине данных подходит БЧХ (2047, 1915) в поле $GF(2^{11})$, способный исправлять $t = 12$ случайных битовых ошибок. На рис. 2 представлена структура кода и возможная комбинация исправляемых ошибок.

Таким образом, из вышеприведенного следует, что код БЧХ(2047, 1915) исправляет 12 случайных битовых ошибок, в то время как код РС(255, 239) примерно с тем же числом информационных и контрольных битов исправляет лишь $t = 8$ случайных ошибок, но байтовых.

Для оценки исправляющей способности кода строят передаточную функцию — зависимость числа неисправленных ошибок на выходе из декодера от числа ошибок в канале, т. е. на входе декодера [10, 11]. Аналитическое выражение для вероятности ошибки на выходе декодера от вероятности битовой ошибки на входе приближенно выражается следующим образом [9, 10, 12, 13]:

$$p_b = \frac{1}{n} \sum_{j=t+1}^n j \binom{n}{j} p^j (1-p)^{n-j}, \quad (1)$$

где $\binom{n}{j} = \frac{n!}{j!(n-j)!}$ — число различных способов выбора

из n битов j ошибочных битов; $\binom{n}{j} p^j (1-p)^{n-j}$ —

вероятность появления j ошибок в блоке из n битов. По сути, выражение (1) показывает вероятность появления ошибок кратности $t + 1$ и выше в блоке из n битов вне зависимости от вида кодиро-

вания. Очевидно, если ошибок больше t , то заявленный кодек не справится, и ошибки появятся уже на выходе декодера. Поэтому формулу (1) для битовой ошибки можно применить и для недвоичных кодов РС, если его рассматривать как двоичный $(n \cdot m, k \cdot m) = (255 \cdot 8, 239 \cdot 8) = (2040, 1912)$. Однако для недвоичных кодов существует несколько другая формула [10], выражающая вероятность появления ошибки в символе на выходе (ошибки в символе показаны на рис. 1):

$$p_E = \frac{1}{2^{m-1}} \sum_{j=t+1}^{2^m-1} j \binom{2^m-1}{j} p_s^j (1-p_s)^{2^m-1-j}, \quad (2)$$

где $p_s = 1 - (1-p)^m$ — вероятность ошибки на символ в зависимости от битовой ошибки [9]. Так как в символе РС содержится m битов, то битовая вероятность на выходе декодера РС выражается формулой

$$p_b = \frac{p_E}{m}. \quad (3)$$

Выражение (1) можно упростить, чтобы не вычислять огромного порядка факториалы для построения зависимостей на графике. Из [10] известно:

$$\frac{1}{n} \sum_{j=2}^n j \binom{n}{j} p^j (1-p)^{n-j} = p - p(1-p)^{n-1}.$$

Потому выражение (1) можно расписать на две суммы и потом упростить:

$$\begin{aligned} p_b &= \frac{1}{n} \sum_{j=t+1}^n j \binom{n}{j} p^j (1-p)^{n-j} = \\ &= \frac{1}{n} \sum_{j=2}^n j \binom{n}{j} p^j (1-p)^{n-j} - \frac{1}{n} \sum_{j=2}^t j \binom{n}{j} p^j (1-p)^{n-j} = \\ &= p - p(1-p)^{n-1} - \frac{1}{n} \sum_{j=2}^t j \binom{n}{j} p^j (1-p)^{n-j} = \\ &= p - p(1-p)^{n-1} - \frac{1}{n} \sum_{j=2}^t j \prod_{i=1}^{j-1} \frac{(n-j+i)}{j!} p^j (1-p)^{n-j}. \end{aligned} \quad (4)$$

Для подтверждения выражений (1), (3), (4) можно собрать схему в среде моделирования MATLAB, как представлено на рис. 3, с примене-

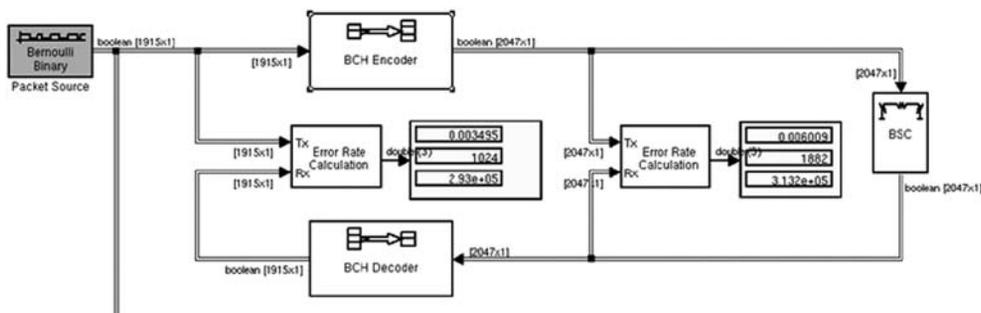


Рис. 3. Схема для оценки передаточной функции кода БЧХ

нием библиотечных кодеков БЧХ и РС с задаваемыми параметрами, генератора случайных данных и двоичного симметричного канала с задаваемой вероятностью ошибки. Аналогичную схему можно собрать для кода РС.

Изменяя в схеме вероятность появления ошибок в канале, можно снять зависимость $BER_{out}(BER_{in})$ — скорость появления ошибки на выходе декодера от скорости на входе. На рис. 4 (см. вторую сторону обложки) изображены данные эксперимента (VSP-b1 RS-b1), а также аналитические зависимости (VCH-b1-calc, RS-b1-calc), описываемые выражениями (3), (4) для кодов БЧХ и РС. Из графика видно, что БЧХ кодек для исправления случайных ошибок обладает лучшей исправляющей способностью (почти 2 порядка при вероятности ошибки на входе $2 \cdot 10^{-3}$). Однако в работах [6, 9] утверждается, что коды РС обладают наилучшей исправляющей способностью среди кодов с такой же избыточностью. Такое утверждение следует из [9], а именно: минимальное расстояние кода (d^*) удовлетворяет неравенству

$$d^* \geq d = 2t + 1 = n - k + 1.$$

Но для любого линейного кода имеет место граница Синглтона:

$$d^* \leq n - k + 1 \Rightarrow d^* = n - k + 1 \Rightarrow d^* = d.$$

Коды РС чрезвычайно эффективны для исправления пакетов ошибок [10]. Поскольку данная схема эксперимента не раскрывает возможности кода РС, необходимо составить другую модель канала.

Оценка корректирующей способности исправлять пакетные ошибки

Существуют способы избегания пакетных, т. е. зависимых, ошибок: перемеживанием данных [3, 12, 14] либо переспросом [15]. В случае работы с flash-памятью нельзя однозначно сказать, что эти способы однозначно устраняют зависимые ошибки ввиду сложной природы их возникновения (см. введение).

Согласно [3] пакетной ошибкой следует считать как минимум два подряд ошибочных бита, отделенных защитным интервалом относительно другого пакета ошибок. Для генерации пакетных ошибок предлагается использовать ДСК подобно схеме рис. 3 с отличием в дополнительном блоке. А именно, предлагается каждый ошибочный бит как бы "размазывать" по данным на задаваемую ширину — собственно ширину пакета ошибок. Ввиду того, что в нашем случае число исправляемых кодеками ошибок значительно меньше длины данных, можно считать, что ошибки в ДСК достаточно прорежены, и защитный интервал будет иметь место. Работа канала ДСК с пакетными ошибками продемонстрирована на рис. 5.

Столь небольшое усложнение канала ДСК позволяет провести ряд экспериментов с целью оценить

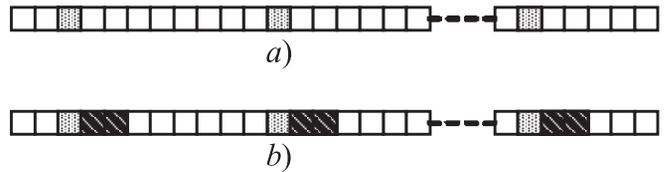


Рис. 5. Ошибки в данных после канала ДСК (а); пакетные ошибки (длина пакета — 3 бита) (б)

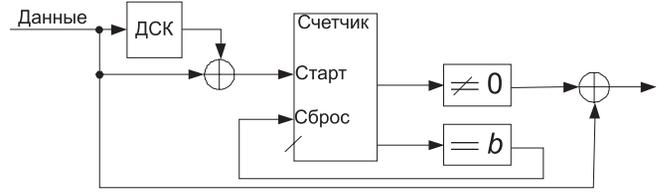


Рис. 6. Схема канала ДСК с пакетами ошибок длины b

корректирующую способность сравниваемых кодов БЧХ и РС. Предлагаемый канал является вырожденным случаем канала с конечным числом состояний [3] или канала Гильберта—Эллиота [16]. При этом вероятность возникновения пакета теперь задается вероятностью появления битовой ошибки или эту же вероятность можно интерпретировать как вероятность перехода в "плохое" состояние, при этом в "хорошем" состоянии данные не портятся вовсе, а в "плохом" портятся с вероятностью 1. В подходах, предложенных в работах [17—20], рассматривается канал Гильберта—Эллиота, однако анализ результатов оказывается весьма сложным ввиду большого числа переменных. В предлагаемом вырожденном канале добавляется лишь один переменный параметр b — длина пакета ошибок. На рис. 6 представлена схема канала ДСК с пакетными ошибками.

Вероятность битовой ошибки на выходе декодера БЧХ в этом случае будет выражаться следующим образом:

$$p_b = \frac{b}{n} \sum_{j=t/b+1}^n \binom{n}{j} p_L^j (1 - p_L)^{n-j}, \quad (5)$$

где $p_L = pb$ — вероятность появления битовых ошибок на входе декодера; p — вероятность появления пакетов ошибок (параметр ДСК); b — длина пакета ошибок.

Экспериментальные результаты

Семейство зависимостей (5) для пакетов ошибок в 1, 2, 6, 12 бит, а также экспериментальные данные приведены на рис. 7 (см. вторую сторону обложки). Очевидно, что при увеличении длины пакета ошибки кодек декодирует все хуже, устремляясь по поведению к не декодированным данным. Это объясняется тем, что для кодов БЧХ нет разницы — сгруппированы ошибки или нет, влияет лишь общее число неисправных битов.

Подобное семейство экспериментальных данных для кода РС(255, 239) представлено на рис. 8 (см. вторую сторону обложки). Видно, что в данном случае нельзя однозначно сказать, что при увеличении длины пакета кодек исправляет ошибки все хуже.

На рис. 9 (см. вторую сторону обложки) представлены сравнительные характеристики кодеков БЧХ и РС в канале ДСК с пакетными ошибками (длина пакетов 2, 8, 13). На рис. 4 было очевидно, что БЧХ для случайных ошибок выигрывает у РС почти 2 порядка, однако уже в случае если ошибки сгруппированы по два бита, кодек РС показывает преимущество. Далее преимущество увеличивается при группировании ошибок по 8 бит (два порядка при вероятности на входе 0,002). При этом характеристика БЧХ при увеличении длины пакета стремится к характеристике не декодированных данных, а РС улучшается. В предельном случае, т. е. при длине пакета ошибок ($b = 13$), превышающей исправляющую способность кода БЧХ (12 исправляемых ошибок), характеристика практически не отличается от характеристики без декодера, однако кодек РС продолжает выполнять успешное декодирование.

Результаты синтеза логической схемы на базе ПЛИС фирмы Altera (Agria II) приведены в таблице.

Заключение

Несмотря на то, что для случайных ошибок преимущества остаются за кодом БЧХ (почти два порядка при вероятности ошибки на входе $2 \cdot 10^{-3}$), предложенная модель канала ДСК с пакетными ошибками наглядно показывает преимущества кода РС (те же два порядка) в сравнении с кодом БЧХ, при прочих равных условиях, а именно скорости кода и длине данных. Даже при группировании ошибок по два бита кодек РС исправляет лучше БЧХ.

Результаты синтеза логической схемы на ПЛИС показали, что для кода РС требуется в 2,5 раза меньше логических элементов, при этом задержка на критическом пути несколько больше. Это объясняется тем, что арифметика для кода РС(255, 239) выполняется в поле Галуа $GF(2^8)$, однако для эквивалентного кода БЧХ (2047, 1912) — в поле $GF(2^{11})$, что усложняет такие элементы, как умножители. Алгоритм РС сложнее БЧХ на один дополнительный шаг (поиск значения ошибки Форни [2]),

Ресурсы ПЛИС, задержка на критическом пути для кодеков РС и БЧХ

Сравниваемые параметры	РС	БЧХ
Логические ячейки, шт.	2885	6676
Регистры, шт.	866	1132
Задержка, нс	9,8	7,5

в котором осуществляется затратная по скорости выполнения операция деления в поле Галуа.

В статье представлены как аналитические выражения для битовой ошибки на выходе декодеров БЧХ и РС, так и их упрощенная интерпретация для построения на ЭВМ. Также предложена относительно простая модель канала с пакетными ошибками, наглядно демонстрирующая преимущества одного кода в сравнении с другим. Данный способ может использоваться для проверки различных блоковых, сверточных и других кодов, без ограничения на длину данных и скорость кода.

Список литературы

1. **Блейхут Р.** Теория и практика кодов, контролируемых ошибки. М.: Мир, 1986, 576 с.
2. **Морелос-Сарагоса Р.** Искусство помехоустойчивого кодирования. Методы алгоритмы, применение. М.: Техносфера, 2005. 320 с.
3. **Галлагер Р.** Теория информации и надежная связь. М.: Советское радио, 1974. 720 с.
4. **What Types of ECC Should Be Used on Flash Memory?** Application note, Spansion, 2011, URL: [//www.spansion.com/Support/Application%20Notes/Types_of_ECC_Used_on_Flash_A.N.pdf](http://www.spansion.com/Support/Application%20Notes/Types_of_ECC_Used_on_Flash_A.N.pdf)
5. **Березюк Н. Т.** Кодирование информации. Двоичные коды. Харьков: Вища школа. 1978, 252 с.
6. **Петерсон У., Уэлдон Э.** Коды, исправляющие ошибки. М.: Мир, 1976. 580 с.
7. **Баишев А.** Твердотельные диски — надежное решение для ответственных применений // Современные технологии автоматизации. 2007. № 4. С. 68—71.
8. **Open NAND Flash Interface Specification.** URL: www.onfi.org
9. **Ратайчук И. А., Шульгин В. И.** Исследование помехоустойчивости цифрового канала связи с использованием недвоичных кодов Рида—Соломона // Радиоэлектронные и компьютерные системы. 2011. № 1.
10. **Скляр Б.** Цифровая связь. Теоретические основы и практическое применение. М.: Изд. дом "Вильямс", 2003. 1106 с.
11. **Forward error correction for high bit-rate DWDM submarine systems.** ITU-T Recommendation G.975.1. URL: <https://www.itu.int/rec/T-REC-G.975.1/en>
12. **Odenwalder J. P.** Error Control Coding Handbook. Linkabit Corporation, 1976. 213 с.
13. **Jordanova L., Laskov L., Dobrev D.** Influence of BCH and LDPC Code Parameters on the BER Characteristic of Satellite DVB Channels // Engineering, Technology&Applied Science Research. 2014. N. 1.
14. **Hafmann W., Pless V.** Fundamentals of Error-correcting Codes. Cambridge University press, 2003.
15. **Галлагер Р.** Коды с малой плотностью проверок на четность. М.: Мир, 1966.
16. **Ryan W. E., Lin S.** Channel Codes Classical. Cambridge University press, 2009.
17. **Мальцев Г. Н., Джумков В. В.** Обобщенная модель дискретного канала передачи информации с группированием ошибок // Информационно-управляющие системы. 2013. № 1.
18. **Певнев В. Я., Пуранов М. В.** Экспериментальное исследование моделей групповых ошибок в канале связи. Харьков: Изд. ХНУВД, 2006.
19. **Золотарев В. В., Овечкин Г. В., Овечкин П. В.** Исследование недвоичных многопороговых декодеров в каналах связи с пакетизируемыми ошибками. М.: Изд. института космических исследований, 2009.
20. **Фрейман В. И., Пирожков А. П.** Исследование моделей дискретных каналов с памятью в рамках лабораторного практикума, построенного на базе пакета MatLab // Электротехника, информационные технологии, системы управления. 2013. № 7.

Errr-Correcting Capability Estimating for RS and BCH Codes

This paper proposes the approach of estimating of error-correcting capability by means of changed Binary Symmetric Channel (BSC). It's known that Reed-Solomon codes are more powerful than any codes with the same code-rate and data length. But to prove it there is a special BSC with packet error. Actually it's shown that RS codes have advantages both in correcting-level capability and in apparatus resources.

Keywords: RS, BCH, correcting level capability, BSC, Galois field

References

1. **Blahut R.** *Teoriya i praktika kodov, kontrolirujushih oshibki.* Moscow: Mir, 1986, 576 p.
2. **Morelos-Saragosa R.** *Iskusstvo pomehoustojchivogo kodirovaniya. Metodi algoritmi primeneniya.* Moscow: Tehnosfera, 2005, 320 p.
3. **Gallager R.** *Theoria informacii I nadejnaja svjaz.* Moscow: Sovetskoe radio, 1974, 720 p.
4. **What** Types of ECC Should Be Used on Flash Memory? Application note, Spansion, 2011, URL: [//www.spansion.com/Support/Application%20Notes/Types_of_ECC_Used_on_Flash_AN.pdf](http://www.spansion.com/Support/Application%20Notes/Types_of_ECC_Used_on_Flash_AN.pdf)
5. **Berezjuk N. T.** *Kodirovanie informacii. Dvoichnie kodi.* Kharkov: Visha shkola, 1978, 252 p.
6. **Peterson W., Weldon A.** *Kodi ispravljajushie oshibki.* Moscow: Mir, 1976, 580 p.
7. **Baishev A.** Tverdotelnye diski — nadejnoe reshenije otvetstvennih primenenij, *Sovremennije tehnologii avtomatizacii*, 2007, no. 4, pp. 68–71.
8. **Open** NAND Flash Interface Specification. URL: www.onfi.org
9. **Ratajchuk I. A., Shulgin V. I.** Issledovanie pomehoustojchivosti cifrovogo kanala svjazi s ispolzovaniem nedvoichnih kodov Reed—Solomon, *Radioelektronnye komputernije sistemi*, 2011, no. 1.
10. **Sklar B.** *Cifirovaja svjaz. Teoriticheskie osnovy i prakticheskoe primeneniya.* Moscow: Williams, 2003, 1106 p.
11. **Forward** error correction for high bit-rate DWDM submarine systems. ITU-T Recommendation G.975.1. URL <https://www.itu.int/rec/T-REC-G.975.1/en>
12. **Odenwalder J. P.** *Error Control Coding Handbook.* Linkabit Corporation, 1976, 213 c.
13. **Jordanova L., Laskov L., Dobrev D.** Influence of BCH and LDPC Code Parameters on the BER Characteristic of Satellite DVB Channels, *Engineering, Technology & Applied Science Research*. 2014, no. 1.
14. **Hafmann W., Pless V.** *Fundamentals of Error-correcting Codes.* Cambridge University press, 2003.
15. **Gallager R.** *Kodi s maloj plotnostju proverok na chetnost.* Moscow: Mir, 1966.
16. **Ryan W. E., Lin S.** *Chahnel Codes Classical.* Cambridge University press, 2009.
17. **Malcev G. N., Djumkov V. V.** Obobshennaja model diskretnogo kanala peredachi informacii s gruppировaniem oshibok, *Informacionno-upravljajushie sistemi*, 2013, no. 1.
18. **Pevnev V. J., Curanov M. V.** *Eksperimentalnoje issledovanie modelej gruppированных oshibok v kanale svjazi.* Kharkov: Izdatelstvo KNUVD, 2006.
19. **Zolotarjev V. V., Ovechkin G. V., Ovechkin P. V.** *Issledovanie nedvoichnih mnogoporogovih dekoderov v kanalah svjazi c pajetirujushimisja oshibkami.* Moscow: Institut kosmicheskikh issledovanij. 2009.
20. **Frejman V. I., Pirojkov A. P.** Issledovanie modelej diskretnih kanalov s pamatju v ramkah laboratornogo praktikuma, postroennogo na baze paketa MatLab. Perm, *Elektrotehnika, informacionne tehnologii, sistemi upravlenija*, 2013, no. 7.

УДК 621.391

С. С. Дворников, лаборант, e-mail: dvornikov_s_s@mail.ru,

С. В. Дворников, д-р техн. наук, проф., профессор кафедры, e-mail: practicsv@yandex.ru,
Военная академия связи имени С. М. Буденного, г. Санкт-Петербург

Формирование сигналов с непрерывной фазой для передачи дискретной информации

Рассмотрены вопросы формирования сигналов каналов управления с фазовой модуляцией. Представлены результаты, характеризующие компактность спектра описываемых сигналов. Даны предложения по практическому использованию результатов. Описаны дальнейшие этапы исследований.

Ключевые слова: частотная модуляция, фазовая модуляция, канал управления, спектр, спектрально-эффективные сигналы

Введение

Ограниченность частотного ресурса стимулирует разработчиков информационно-технических систем (ИТС) к поиску спектрально-эффективных сигнальных конструкций, используемых в интересах передачи информации. При этом к основным показателям, характеризующим тот или иной мо-

дуляционный формат, следует отнести помехоустойчивость и спектральную эффективность [1]. Данные показатели тесно взаимосвязаны между собой, поэтому сложно обеспечить высокие показатели по одному из них, без снижения требований по остальным. Между тем, широко известные модуляционные форматы, используемые для передачи