Министерство образования и науки Российской Федерации
МОСКОВСКИЙ ФИЗИКО-ТЕХНИЧЕСКИЙ ИНСТИТУТ
(ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ)

На правах рукописи

КОБОЗЕВА ИРИНА ГЕННАДЬЕВНА

Исследование сигнально-кодовых конструкций на основе обобщенных кодов с локализацией ошибок

05.13.17 – Теоретические основы информатики

АВТОРЕФЕРАТ
диссертации на соискание учёной степени
кандидата физико-математических наук

Москва – 2013

Работа выполнена в *Федеральном государственном бюджетном учреждении науки Институте проблем передачи информации им. А.А. Харкевича Российской академии наук (ИППИ РАН).*

Научный руководитель: *доктор технических наук*,
Зяблов Виктор Васильевич

Официальные оппоненты: *доктор технических наук*
Федоренко Сергей Валентинович,
эксперт в ООО «Техкомпания Хуавэй»;
кандидат физико-математических наук
Владимиров Сергей Михайлович,
инженер программист в ООО «Одноклассники»

Ведущая организация: *Санкт-Петербургский государственный университет аэрокосмического приборостроения*

Защита состоится 16 декабря 2013 г. в 13.30 на заседании диссертационного совета *Д 212.156.04* при *Федеральном государственном автономном образовательном учреждении высшего профессионального образования «Московском физико-техническом институте (государственном университете)»* по адресу: *141700, г. Долгопрудный, Московская обл., Институтский пер., д. 9, ауд. 204.*

С диссертацией можно ознакомиться в библиотеке *МФТИ (ГУ)*.

Автореферат разослан 15 ноября 2013 г.

Ученый секретарь
диссертационного совета,
к. ф.-м. н. Стрыгин Л. В.

Общая характеристика работы

История кодирования началась в 1948 г. с публикацией знаменитой статьи К. Шеннона «Математическая теория связи». Шеннон показал, что с любым каналом передачи данных связано измеряемое в битах в секунду и называемое пропускной способностью канала число C. Если требуемая от системы связи скорость передачи информации R (измеряемая в битах в секунду) меньше C, то, используя исправляющие ошибки коды, для данного канала можно построить такую систему связи, что вероятность ошибки на выходе будет сколь угодно мала. В самом деле, понятно, что построение очень хороших каналов является сложной задачей; экономически гораздо выгоднее использовать кодирование. Шеннон, однако, не указал, как найти подходящие коды, а лишь доказал их существование.

С тех пор, течение более чем 60 лет, прошедших с момента появления кодов, исправляющих ошибки, наблюдается устойчивый рост требований к их корректирующим свойствам, и предлагаются новые, все более сложные кодовые конструкции. В последнее годы требования к качеству передаваемой информации еще более ужесточились (вероятность ошибки декодирования порядка 10^{-12} и менее). Кроме того, из-за очень высокой скорости передачи данных необходимыми условиями также являются использование методов кодирования и декодирования, требующих относительно малое количество вычислений на бит передаваемой информации, а также возможность параллельных вычислений при кодировании и декодировании. В свою очередь, при высокой кратности модуляции, обеспечивающей большую скорость передачи, более целесообразно использовать недвоичные коды, которые имеют лучшую корректирующую способность по сравнению с двоичными кодами с той же избыточностью.

Можно выделить несколько классов кодов, позволяющих построить длинный код с хорошей корректирующей способностью. Обычно это каскадные коды или МПП-коды. Среди этих кодов особое место занимает подкласс

обобщенных каскадных кодов, а именно обобщенные коды с локализацией ошибок (далее обозначены как ОЛО-коды), которые и являются основным предметом исследований в данной работе. В работе будут исследованы две разновидности ОЛО-кодов – обычные ОЛО-коды (в дальнейшем, чтобы избежать путаницы, обозначены как ОЛО-2 коды) и их новая трехмерная разновидность (далее обозначены как ОЛО-3 коды).

Цели и задачи диссертационной работы

Цель диссертационной работы состоит в исследовании методов повышения защиты от помех и увеличения пропускной способности информационных коммуникаций с использованием кодов с локализацией ошибок. Существенной частью работы является исследование свойств кодов с локализацией ошибок, включая предложенную автором усложнённую трёхмерную версию, а также разработка методов, позволяющих аналитически определять оптимальные параметры для ОЛО-2 и ОЛО-3 кодов.

Для достижения поставленных целей были решены следующие задачи:

- анализ существующих способов кодирования и декодирования обобщенными кодами с локализацией ошибок;
- разработка алгоритмов кодирования/декодирования ОЛО-3 кодов;
- разработка методов выбора оптимальных параметров для ОЛО-2 и ОЛО-3 кодов, позволяющих найти код с максимальной скоростью передачи при заданных входной и выходной вероятностях ошибки.

Научная новизна работы

- Предложена модификация обобщенных кодов с локализацией ошибок – а именно, трехмерные обобщенные коды с локализацией ошибок.
- Для трехмерных обобщенных кодов с локализацией ошибок были разработаны алгоритмы кодирования и декодирования.
- Разработан теоретический метод расчёта вероятности неправильного декодирования для ОЛО-2 и ОЛО-3 кодов, позволяющий выбрать оптималь-

ные параметры кода для известного канала передачи, а также оценить вероятность неправильного декодирования для хороших условий передачи($<10^{-12}$), что было бы крайне затратно при моделировании.

- Было проведено моделирование в среде MATLAB нескольких сигнально-кодовых конструкций на основе ОЛО-2 и ОЛО-3 кодов. Было произведено сравнение полученных результатов друг с другом и с некоторыми ныне существующими стандартами.

Теоретическая и практическая значимость работы

Работа носит, в целом, теоретический характер. Теоретическая ценность диссертации определяется теоретическими методами оценки вероятности неправильного декодирования для ОЛО-2 и ОЛО-3 кодов, позволяющими выбрать подходящие параметры для избыточности каждого из кодов-компонентов. Данный вопрос является весьма существенным, так как ОЛО-коды позиционируются как коды с малой избыточностью, для которых важно рациональное использование каждого проверочного символа. Описанные в работе методы позволяют подобрать ОЛО-код с максимальной скоростью передачи при заданных входных вероятностях ошибки и стирания в канале и выходной вероятности неверного декодирования.

На защиту выносятся следующие основные результаты и положения:

- Разработка теоретических методов расчета и оптимизации конструкций ОЛО-2 кодов на основе кодов Рида-Соломона для заданной вероятности неверного декодирования;
- Разработка конструкций ОЛО-3 кодов и оценка их параметров;
- Разработка теоретических методов расчета и оптимизации конструкций ОЛО-3 кодов на основе кодов Рида-Соломона для заданной вероятности неверного декодирования;
- Разработка алгоритмов кодирования/декодирования для рассматриваемых кодовых конструкций.

Апробация результатов работы

Основные результаты диссертации докладывались на следующих конференциях:

- XII симпозиум по проблеме избыточности в информационных системах, г. Санкт-Петербург, 2009 г.
- Конференция «Информационные технологии и системы» (ИТиС'09), 32-я конференция молодых ученых и специалистов ИППИ РАН, п. Бекасово, 2009 г.
- Конференция «Информационные технологии и системы» (ИТиС'10), 33-я конференция молодых ученых и специалистов ИППИ РАН, г. Геленджик, 2010 г.
- Конференция «Информационные технологии и системы» (ИТиС'11), 34-я конференция молодых ученых и специалистов ИППИ РАН, г. Геленджик, 2011 г.
- Конференция «Информационные технологии и системы» (ИТиС'12), 35-я конференция молодых ученых и специалистов ИППИ РАН, г. Петрозаводск, 2012 г.

Кроме того, результаты работы неоднократно докладывались на научных семинарах лаборатории №3 ИППИ РАН.

Структура и объем диссертации

Диссертация состоит из введения, обзора литературы, четырёх глав, заключения и библиографии. Общий объем диссертации 100 страниц, включая 47 рисунков. Библиография включает 37 наименований на 4 страницах.

Публикации

По теме диссертации опубликовано 3 статьи в рецензируемых отечественных и международных журналах [1- 3], 5 тезисов докладов на конфе-

ренциях [4-8]. Личный вклад соискателя в опубликованные работы является определяющим. Результаты, выносимые на защиту, получены автором самостоятельно. Все представленные в диссертации результаты получены лично автором.

Содержание работы

Диссертация состоит из введения и трех глав. Во **введении** обоснована актуальность диссертационной работы, сформулирована цель и аргументирована научная новизна исследований, показана практическая значимость полученных результатов, а также представлены выносимые на защиту научные положения.

В **первой** главе даётся краткий обзор по существующим кодам с локализацией ошибок с целью обобщить основные факты и теоретические модели для их дальнейшего использования.

Проверочная матрица внутренних кодов ОЛО-2 кода имеет вид:

$$H_B = \begin{pmatrix} H_B^{(1)} \\ H_B^{(2)} \\ \vdots \\ H_B^{(L)} \end{pmatrix} = \begin{pmatrix} Q_0^{(1)} & I^{(1)} & 0 & \dots & 0 \\ Q_0^{(2)} & Q_1^{(2)} & I^{(2)} & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ Q_0^{(L)} & Q_1^{(L)} & \cdots & Q_{L-1}^{(L)} & I^{(L)} \end{pmatrix},$$

где L -количество внешних A и внутренних B кодов, где $I^{(i)}$ – единичные матрицы, а $Q_j^{(i)}$ некоторые матрицы с элементами из $GF(q)$, $i = \overline{1,L}$, $j = \overline{0,L-1}$.

Алгоритм кодирования ОЛО-2 кодами

Вначале информационные символы записываются в левую часть матрицы E, изображенной на рис. 1, в правую часть записываются нули.

1. Информационные символы разбиваются на n_A подблоков, из которых $k_A^{(i)}$ первых подблоков имеют длину n_B, а $r_A^{(i)}$ оставшихся – длину $r_B^{(i)} = n_B - k_B^{(i)}$.

2. Каждый из $r_A^{(i)}$ подблоков длины $k_B^{(i)}$ кодируется внутренним кодом Рида-Соломона с параметрами $(n_B, k_B^{(i)}, d_B^{(i)})$. В результате получаем r_A векторов

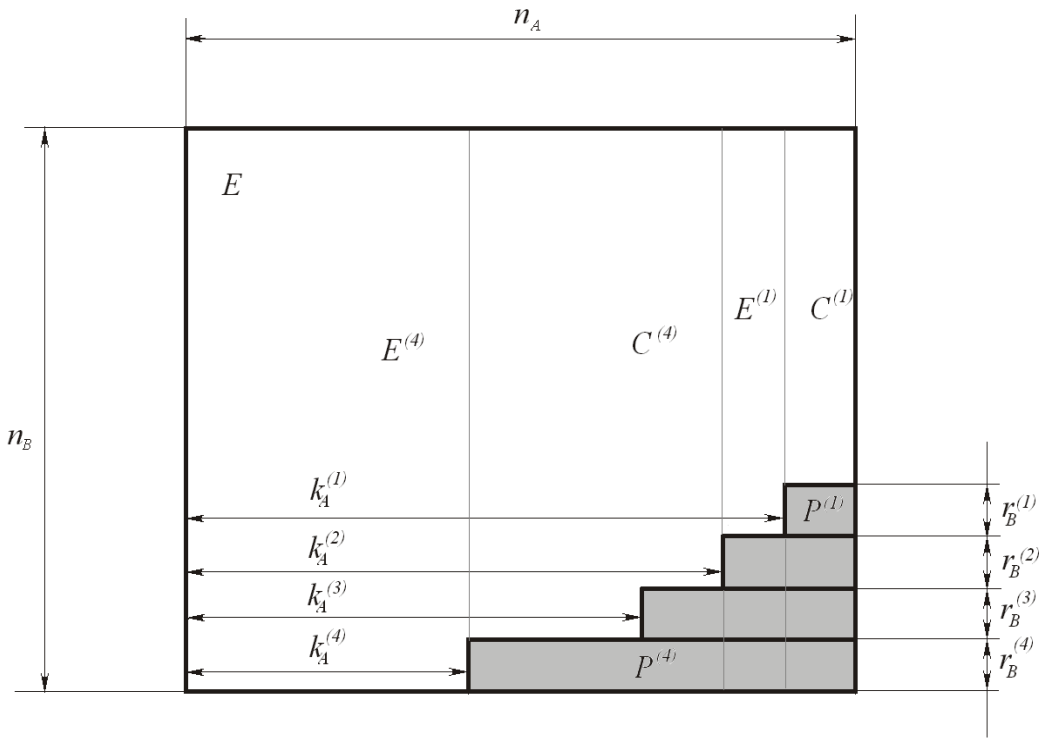

Рис. 1. Структура ОЛО-2 кода.

длины n_B, являющихся кодовыми словами внутреннего кода. На месте незаполненных ранее областей теперь находятся проверочные символы соответствующего внутреннего кода.

3. Для каждого из первых $k_A^{(i)}$ подблоков матрицы E длины n_B находим вектор длины $r_B^{(i)} = n_B - k_B^{(i)}$ по формуле $\bar{s}_\mu^{(i)} = \nu_\mu^{(i)} H_B^{(i)T}$, где $H_B^{(i)}$ – проверочная матрица кода i-ой ступени в приведённо-ступенчатой форме, а $\nu_\mu^{(i)}$ – μ-ый подблок. Представляя векторы $\bar{s}_\mu^{(i)}$, $\mu = 1,...,k_A$, как элементы $GF(2^{r_B^{(i)}})$, получаем вектор из $k_A^{(i)}$ символов над $GF(2^{r_B^{(i)}})$ и кодируем его внешним кодом Рида-Соломона с параметрами $(n_A, k_A^{(i)}, d_A^{(i)})$. В результате получаем кодовое слово внешнего кода. Затем находим проверочные символы собственно ОЛО-2 кода.

Таким образом, получаем

$$E'^{(i)} = (U^{(i)} \mid R^{(i)}) = H_B^{(i)} E = (Q_0^{(i)} ... Q_{i-1}^{(i)} \mid I_{r_B^{(i)}})(E^{(i)} \mid \frac{C^{(i)}}{P^{(i)}})$$

$$R^{(i)} = (Q_0^{(i)} ... Q_{i-1}^{(i)}) C^{(i)} + I_{r_B^{(i)}} P^{(i)} \Rightarrow P^{(i)} = R^{(i)} - (Q_0^{(i)} ... Q_{i-1}^{(i)}) C^{(i)}.$$

Алгоритм декодирования ОЛО-2 кодов.

Декодирование ОЛО-кода осуществляется в несколько этапов, количество которых равно порядку ОЛО-кода *m*. При передаче по каналу с шумом к кодовому слову ОЛО-кода *C* добавляется некоторая ошибка *E*. Таким образом, полученная комбинация E_X задается как $E_X = X \oplus E$. В начале декодирования вычисляются синдромы соответствующего данному шагу внутреннего кода, необходимые для определения символов внешнего кода $\tilde{s}^{(i)} = H_B^{(i)} E_X$, после чего происходит декодирование соответствующим внешним кодом, в результате которого получаем $\bar{s}^{(i)} = H_B^{(i)} E$.

Предполагая, что декодирование внешним кодом было успешным, находим столбцы матрицы E $\overline{b_j}$, считая что $\bar{s}_j^{(i)} = H_B^{(i)} \overline{b_j}$. На *i*-ом шаге мы комбинируем синдромы $\bar{s}_j^{(i)}$, полученные на шаге *i*, с синдромами $\bar{s}_j^{(L)}, \ldots, \bar{s}_j^{(i+1)}$, полученными на предыдущих шагах, и декодируем их *i*-ым вложенным внутренним кодом с проверочной матрицей $H_B^{L,\ldots,i}$, что позволяет нам получить корректирующую способность большую, чем у отдельно взятого внутреннего кода.

$$H_B^{L,\ldots,i} = \begin{pmatrix} H_B^{(i)} \\ H_B^{(i+1)} \\ \vdots \\ H_B^{(L)} \end{pmatrix}.$$

На каждом шаге для соответствующего внутреннего кода происходит как обнаружение, так и исправление стираний и ошибок. Таким образом, если количество обнаруженных ошибок или число стираний в кодовом слове внутреннего кода превышает корректирующую способность этого кода, то для соответствующего этому слову символа внешнего кода выносится вердикт «стирание». Если же в кодовом слове внутреннего кода, неверно исправленном на предыдущем шаге, на текущем шаге вновь обнаружены ошибки, то это слово становится таким, каким оно было на предыдущем ша-

ге. При декодировании на каждом следующем шаге используется информация о стертых символах, полученная на предыдущем шаге.

Проверочную матрицу ОЛО-2 кода можно представить в виде: $H_E = \begin{pmatrix} H_A^{(1)} \otimes H_B{}^{(1)} \\ H_A^{(2)} \otimes H_B{}^{(2)} \\ \vdots \\ H_A^{(L)} \otimes H_B{}^{(L)} \end{pmatrix}$, то есть

$$H_E^{(i)} = \left(H_A{}^{(i)} \otimes H_B{}^{(i)} \right) = \begin{pmatrix} h_{1,1}^{A(i)} H_B^{(i)} & h_{1,2}^{A(i)} H_B^{(i)} \dots h_{1,n_A}^{A(i)} H_B^{(i)} \\ h_{2,1}^{A(i)} H_B^{(i)} & h_{2,2}^{A(i)} H_B^{(i)} \dots h_{2,n_A}^{A(i)} H_B^{(i)} \\ \vdots & \vdots \quad \ddots \quad \vdots \\ h_{r_A^{(i)},1}^{A(i)} H_B^{(i)} & h_{r_A^{(i)},2}^{A(i)} H_B^{(i)} \cdots h_{r_2^{(i)},n_A}^{A(i)} H_B^{(i)} \end{pmatrix},$$

где $h_{k,j}^{A(i)}$ – соответствующие элементы матрицы $H_A^{(i)}$.

Следует отметить, что используемое выше прямое кронекеровское произведение, строго говоря, является некоммутативной операцией, поэтому обязательно следует умножать матрицы в указанном порядке.

В первой главе также проводится исследование некоторых сигнально-кодовых конструкций на основе ОЛО-2 кодов и освещается проблема выбора наиболее подходящих параметров кода при заданных условиях в канале передачи данных. Для получения оптимальных характеристик необходимо правильно выбрать порядок m (количество шагов), а также количество проверочных символов в каждом горизонтальном блоке. Структура кода выбирается таким образом, чтобы результирующее значение вероятности неверного декодирования было меньше ε. Определим, какое количество шагов для этого необходимо.

Утверждение 1.

1. *Для того, чтобы правильно выбрать порядок кода m, необходимо определить, какая избыточность $d_B^{(i)}$ для внутреннего кода на первом же шаге для нужного отношения сигнал/шум обеспечит ошибку меньше $\varepsilon / d_B^{(i)}$.*

2. *Чтобы обеспечить нужный уровень вероятности ошибки, необходимо соблюсти условие: ошибка декодирования для одного столбца имеет вид*

$$P_{err}^{B} = \sum_{t=0}^{d_B^{(i)}-1} C_n^t p_t^t \sum_{e=\left\lceil \left(d_B^{(i)}-t\right)/2 \right\rceil}^{n-t} C_{n-t}^e p_e^e (1-p_e-p_t)^{n-t-e} + \sum_{t=d_B^{(i)}}^{n} C_n^t p_t^t (1-p_t)^{n-t},$$

$$P_{\Sigma err} = 1-(1-P_{err}^{B})^{n_A} \le \varepsilon / d_B^{(i)},$$

где P_{err}^{B} *– ошибка декодирования для одного столбца, а* $P_{\Sigma err}$ *– ошибка декодирования для всего слова длины* $n_A n_B$.

Дополнительно в этой главе приводится оценка для вероятности неверного декодирования ОЛО-2 кодов на основе кодов Рида-Соломона при известных условиях в канале связи. Как известно, декодирование ОЛО-2 кода осуществляется в несколько этапов, на каждом из которых используется один внешний и один внутренний код. Обозначим через $P_{B+}^{(i)}$ вероятность того, что *i*-ый внутренний код был декодирован верно, и через $P_{A+}^{(i)}$ – вероятность правильного декодирования для *i*-ого внешнего кода.

В начале каждого шага декодирования производится обнаружение ошибок с помощью внутреннего кода. Будем считать, что обнаружение ошибок для кода Рида-Соломона происходит, только если выполняется неравенство $e+\tau \le d-1$, хотя, на самом деле, таких случаев больше. Затем производится исправление ошибок и стираний. Поскольку на каждом следующем шаге декодирование осуществляется при помощи кода Рида-Соломона, то для успешного декодирования необходимо выполнение неравенства $2e+\tau \le d-1$. Если же при исправлении ошибок и стираний при избыточности d на текущем шаге, произошла ошибка декодирования, то на следующем шаге при избыточности $d+1$ ошибки будут обнаружены, только в том случае, если $e+\tau=d$. Ниже приведены формулы, используемые для вычисления вероятности неудовлетворительного исхода при исправлении и при обнаружении ошибок для кодов Рида-Соломона:

$$P_{исправ.} = \sum_{t=0}^{d-1} C_n^t p_t^t \sum_{e=\lceil (d-t)/2 \rceil}^{n-t} C_{n-t}^e p_e^e (1-p_e-p_t)^{n-t-e} + \sum_{t=d}^{n} C_n^t p_t^t (1-p_t)^{n-t}$$

$$P_{обнаруж.} = \sum_{t=0}^{d-1} C_n^t p_t^t \sum_{e=d-t}^{n-t} C_{n-t}^e p_e^e (1-p_e-p_t)^{n-t-e} + \sum_{t=d}^{n} C_n^t p_t^t (1-p_t)^{n-t}.$$

После громоздких, но довольно несложных вычислений мы получаем, что для *i*-ого шага вероятности стирания и ошибки в символах соответствующего внешнего кода записываются как

$$P_{Be}^{(i)} = P_B(t = \{d_i - 1, .., 0\}, e \geq d_i - t),$$

$$P_{Bt}^{(i)} = P_B(e = 0, ..., n-t, t \geq d_i) + P_B(e + t = d_i - 1),$$

$$P_{B+}^{(i)} = 1 - P_{Bt}^{(i)} - P_{Be}^{(i)}.$$

Общая же формула для вероятности неверного декодирования *i+1*-го внешнего кода $P_{A-}^{(i+1)} = 1 - P_{A+}^{(i+1)}$ при условии, что все предыдущие были верно декодированы, выглядит следующим образом:

$$P_{A-}^{(i+1)} = \sum_{t=0}^{d_{i+1}^{(A)}-1} C_n^t P_{Bt}^{(i)t} \sum_{e=\lceil (d_{i+1}^A - t)/2 \rceil}^{n-t} C_{n-t}^e P_{Be}^{(i)e} P_{B+}^{(i)n-t-e} + \sum_{t=d_{i+1}^A}^{n} C_n^t P_{Bt}^{(i)t} (1 - P_{Bt}^{(i)})^{n-t}.$$

Исходя из строения кода, ОЛО-2 код будет полностью верно декодирован, если все внешние коды и последний внутренний код для всех столбцов будут декодированы правильно. В противном случае, в полученном после декодирования слове останутся искаженные столбцы. Запишем вероятности правильного и неправильного декодирования всего ОЛО-кода как

$$P_{\Sigma-} = P_{A-}^{(1)} + P_{A+}^{(1)} P_{A-}^{(2)} + P_{A+}^{(1)} P_{A+}^{(2)} P_{A-}^{(3)} + ... +$$
$$+ P_{A+}^{(1)} P_{A+}^{(2)} \cdot ... \cdot P_{A+}^{(i-1)} P_{A-}^{(i)} + P_{A+}^{(1)} \cdot ... \cdot P_{A+}^{(i-1)} P_{A+}^{(i)} (1 - P_{B+}^{(i)n_a}),$$

$$P_{\Sigma+} = P_{A+}^{(1)} P_{A+}^{(2)} \cdot ... \cdot P_{A+}^{(i-1)} P_{A+}^{(i)} P_{B+}^{(i)n_a},$$

где $P_{A-}^{(i)}$ и $P_{A+}^{(i)}$ – вероятность неверного и верного декодирования *i*-ым внешним кодом, а $(1 - P_{B+}^{(i)n_a})$ – вероятность того, что при декодировании *i*-ым внутренним кодом хотя бы в одном столбце произошла ошибка.

Необходимо также добавить, что величина, в работе называемая вероятностью неправильного декодирования, включает в себя также и вероятность отказа от декодирования, поскольку в ходе работы и ошибка, и отказ от декодирования считались неудовлетворительным результатом. На самом де-

ле, нужно отметить, что вероятность ошибки незначительна по сравнению с вероятностью отказа от декодирования.

Данные вычисления позволяют сделать довольно точную оценку сверху для ОЛО-кода с известными параметрами. Особенно важен тот факт, что полученная оценка зависит только от величин p_e и p_t, то есть, вычислив эти вероятности для канала с белым шумом, или, в более сложных случаях, получив их с помощью моделирования, мы всегда можем сделать оценку, не оглядываясь на особенности канала передачи данных. Помимо этого, мы также можем оценить вычислительную сложность для ОЛО-2 кодов, являющуюся одним из существенных факторов, влияющих на практическое применение кодов. Общеизвестно, что чем больше длина кода, тем больше операций требуется для его реализации. Для кодов Рида-Соломона длины n над полем $GF(q)$, рассматриваемых в данной работе, количество необходимых вычислений при декодировании асимптотически можно оценить как $\log(q)\cdot C\cdot n\cdot\log^2(n)$, где C – это некоторая величина, явным образом не зависящая от длины кода.

Лемма 1. *Количество операций при декодировании ОЛО-2 кода можно оценить как*

$$N=\log(q)\left(C_B n_A n_B\log^2(n_B)+C_A n_A r_B^{(i)}\log^2(n_A)+n_A n_B\sum_{i=1}^{L}r_B^{(i)}\right),$$

где C_A и C_B – некоторые коэффициенты, явным образом не зависящие от длины кода.

Доказательство. В случае ОЛО-2 кодов на каждом шаге декодирования происходит исправление ошибок и стираний при помощи внешнего кода длины n_A над полем $GF(q^{r_B^{(i)}})$, что дает нам $\log(q^{r_B^{(i)}})\cdot C_A^{(i)}\cdot n_A\cdot\log^2(n_A)$ операций, после этого, имеет место исправление и обнаружение ошибок n_A внутренними кодами длины n_B над полем $GF(q)$, что дополнительно прибавляет нам $\sum_{i=1}^{L}2C_B^{(i)}\log(q)\cdot n_A\cdot n_B\log^2(n_B)$ операций. Кроме того, для вычисления синдромов нам необходимо умножать n_A векторов на проверочные матрицы размеров

$r_B^{(i)} \times n_B$, что дает нам еще $\log(q) \cdot n_A \cdot n_B \cdot r_B^{(i)}$ операций. Таким образом, количество операций при декодировании ОЛО-2 кода можно оценить как

$$N = \sum_{i=1}^{L} 2C_B^{(i)} \log(q) n_A n_B \log^2(n_B) + \log(q) n_A n_B r_B^{(i)} + \log(q^{r_B^{(i)}}) C_A^{(i)} n_A \log^2(n_A) =$$

$$= \sum_{i=1}^{L} \log(q) n_A \left(2C_B^{(i)} n_B \log^2(n_B) + C_A^{(i)} r_B^{(i)} \log^2(n_A) + n_B \cdot r_B^{(i)} \right) =$$

$$= \log(q) \left(C_B n_A n_B \log^2(n_B) + C_A n_A r_B^{(i)} \log^2(n_A) + n_A n_B \sum_{i=1}^{L} r_B^{(i)} \right),$$

что, собственно, и требовалось доказать.

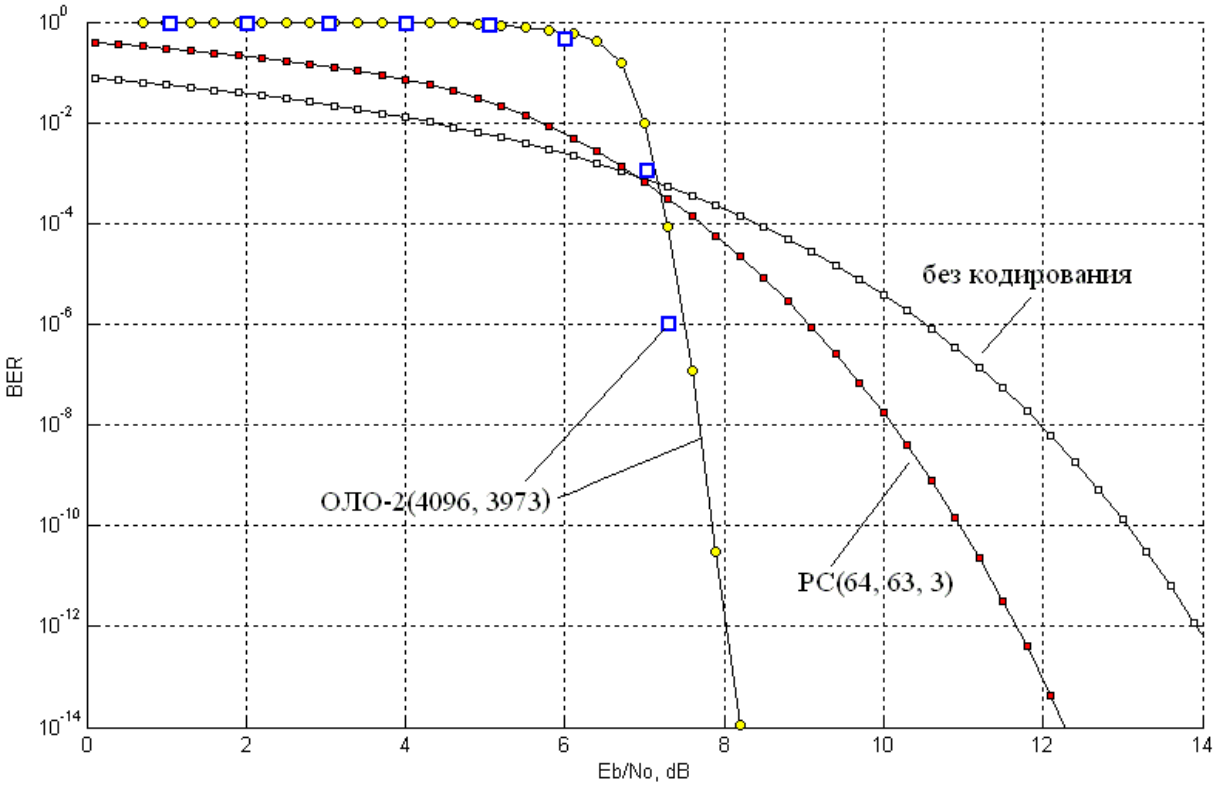

Рис. 2. Зависимость вероятности отказа от декодирования + ошибки от отношения сигнал/шум в канале для ОЛО-2 (4096, 3973), m=12 избыточностью 3% при передаче по гауссовскому каналу.

Во **второй главе** приводятся результаты моделирования в среде MATLAB для ОЛО-2 кодов. Количество проведенных испытаний для каждого значения отношения сигнал/шум равно 10^6. В результате большого количества численных экспериментов было установлено, что полученная в работе теоретическая оценка является достаточно точной оценкой сверху. Ниже, для примера, приведен график зависимости вероятности отказа от декодиро-

вания + ошибки от отношения сигнал/шум на бит в канале для ОЛО-2 (4096, 3973), m = 12 избыточностью 3% при передаче по гауссовскому каналу. Все параметры кода выбраны в соответствии с описанными выше теоретическими методами. Кружочками обозначены результаты теоретических расчетов, квадратами – результаты моделирования. Для сравнения на этом же графике также изображена теоретическая оценка зависимости вероятности отказа от декодирования + ошибки для кода Рида-Соломона с параметрами (64, 63, 3), имеющего ту же избыточность. Как мы можем убедиться, при вероятности ошибки 10^{-12} выигрыш ОЛО-2 кода по сравнению с кодами Рида-Соломона составляет более 3,5 дБ.

Третья глава посвящена разработке алгоритмов для кодирования/декодирования новой, усложненной версии ОЛО-2 кодов – трехмерных обобщенных кодов с локализацией ошибок (ОЛО-3 кодов).

Как уже говорилось выше, кодирование обобщенными кодами с локализацией ошибок осуществляется в несколько шагов. Количество этих шагов мы будем называть порядком кода L. Соответственно, каждому шагу i будут соответствовать три различных кода: внутренний код $C^{(i)}$ с параметрами $(q, n_C, k_C^{(i)}, d_C^{(i)})$, промежуточный код $B^{(i)}$ с параметрами $(q^{r_C^{(i)}}, n_B, k_B^{(i)}, d_B^{(i)})$ и внешний код $A^{(i)}$ с параметрами $(q^{r_C^{(i)} r_B^{(i)}}, n_A, k_A^{(i)}, d_A^{(i)})$, где n – длина кода, $k^{(i)}$ и $r^{(i)}$ – количество информационных и проверочных символов соответственно, а $d^{(i)}$ – кодовое расстояние. Для примера, структура кодового слова обобщенного трехмерного кода с локализацией ошибок (далее обозначенного как ОЛО-3 код) порядка 4 представлена на рис. 3.

Алгоритм кодирования ОЛО-3 кодами

Кодирование ОЛО-3 кодами, в целом, осуществляется сходным с ОЛО-2 кодами образом, разница лишь в том, что в этом случае мы имеем L наборов из 3, а не 2 кодов-компонентов – по одному внутреннему, внешнему и промежуточному на каждый этап. Вначале информационные символы запи-

сываются в информационную часть матрицы размеров $n = n_A n_B n_C$, а проверочная часть заполняется нулями. Действуя по описанной выше схеме,

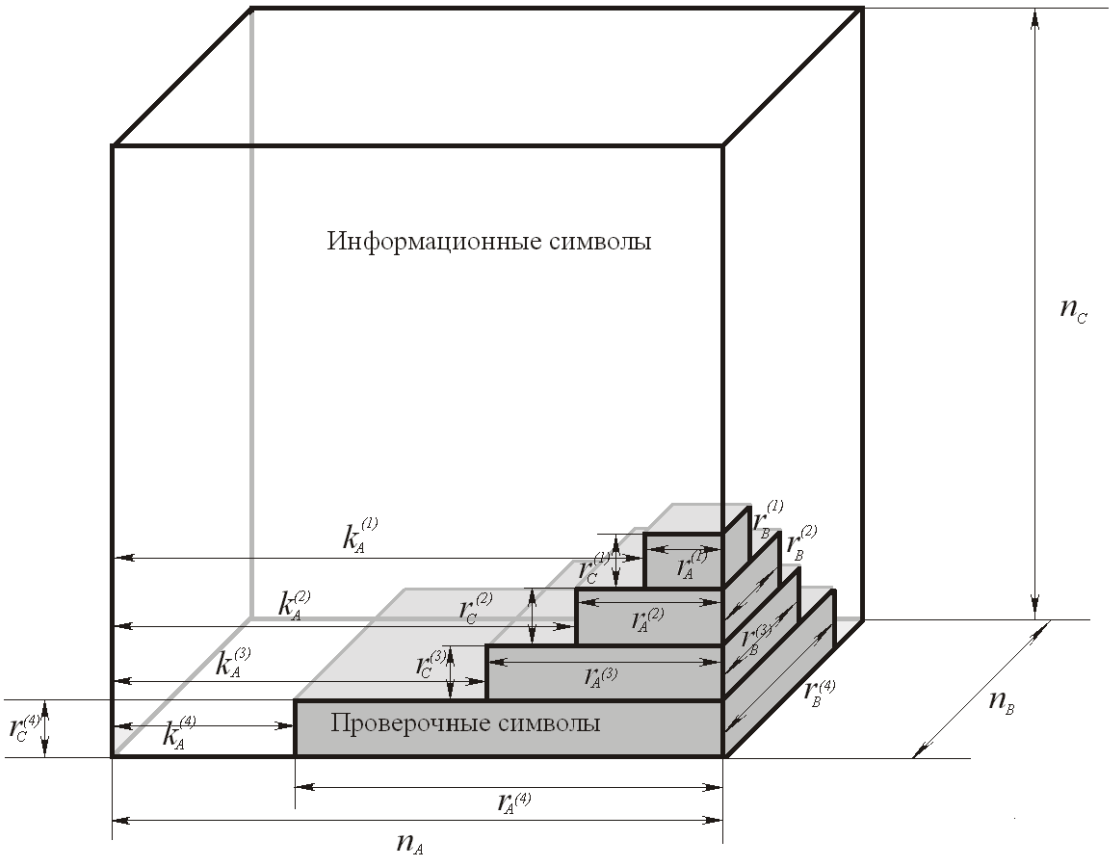

Рис. 3. Структура кодового слова ОЛО-3 кода порядка 4.

мы независимо вычисляем проверочные символы для каждого набора кодов. Умножая вертикальные столбцы матрицы *E* на проверочную матрицу $H_C^{(i)}$ соответствующего внутреннего кода, мы вычисляем синдромы ошибки, которые образуют матрицу *E'*. При этом символы, соответствующие ранее заполненной нулями части матрицы *E*, переходят в проверочные символы внутреннего кода.

Проверочная матрица внутреннего кода имеет вид $H_C^{(i)} = (Q_{C0}^{(i)} ... Q_{Ci-1}^{(i)} \mid I_{r_C^{(i)}})$, где $I_{r_C^{(i)}}$ – единичная матрица размеров $r_C^{(i)} \times r_C^{(i)}$.

Из проверочной части матрицы *E'* мы убираем все проверочные символы внутренних кодов, заменяя их нулями; оставшиеся синдромы $U_B^{(i)}$, в свою очередь, рассматриваются как символы промежуточного кода над полем

$GF(q^{r_C^{(i)}})$. Таким образом, над новым полем мы получаем двумерную матрицу E_0' размеров $n_A \times n_B$. Применяя теперь к двумерной матрице E_0' алгоритм кодирования двумерными ОЛО-2 кодами, получаем

$$E''^{(i)} = (U'^{(i)} \mid R'^{(i)}) = H_B^{(i)} \cdot E_0' = (Q_{B0}^{(i)} ... Q_{Bi-1}^{(i)} \mid I_{r_B^{(i)}})(E'^{(i)} \mid \frac{C'^{(i)}}{P^{(i)}})$$

$$R'^{(i)} = (Q_{B0}^{(i)} ... Q_{Bi-1}^{(i)})C'^{(i)} + I_{r_B^{(i)}} P^{(i)} \Rightarrow P^{(i)} = R'^{(i)} - (Q_{C0}^{(i)} ... Q_{Bi-1}^{(i)})C'^{(i)}.$$

Проверочную матрицу ОЛО-3 кода можно представить в виде:

$$H_E = \begin{pmatrix} H_E{}^{(1)} \\ H_E{}^{(2)} \\ \vdots \\ H_E{}^{(L)} \end{pmatrix} = \begin{pmatrix} H_{AB}{}^{(1)} \otimes H_C{}^{(1)} \\ H_{AB}{}^{(2)} \otimes H_C{}^{(2)} \\ \vdots \\ H_{AB}{}^{(L)} \otimes H_C{}^{(L)} \end{pmatrix}, \text{ где}$$

$H_{AB}{}^{(i)}$ – проверочная матрица каскадного кода длины $n_A n_B$.

Алгоритм декодирования ОЛО-3 кодов

И в ОЛО-2, и в ОЛО-3 кодах первым всегда декодируется внешний код. После этого для ОЛО-3 кодов последовательно осуществляется декодирование вложенных промежуточного и внутреннего кодов. Декодирование ОЛО-3 кодами (в отличие от кодирования, где все проверочные символы для каждой тройки кодов-компонентов можно вычислять параллельно) происходит последовательно в L итераций.

Рассмотрим схему декодирования более подробно. В начале каждого шага последовательно вычисляются синдромы сначала внутреннего кода, затем синдромы промежуточного кода, являющиеся соответственно символами промежуточного и внешнего кодов. После этого происходит декодирование внешним кодом, причем необходимо, чтобы его избыточности хватило для исправления всех ошибок и стираний, поскольку, если хотя бы один из внешних кодов будет декодирован неверно, то и весь ОЛО-3 код будет впоследствии неверно декодирован. После этого необходимо провести декодирование соответствующим промежуточными кодом. И уже только после это-

го происходит исправление и обнаружение ошибок соответствующим вложенным внутренним кодом с проверочной матрицей

$$H_C^{L,\dots,i} = \begin{pmatrix} H_C^{(i)} \\ H_C^{(i+1)} \\ \vdots \\ H_C^{(L)} \end{pmatrix}.$$

Таким образом, если количество обнаруженных ошибок или число стираний в кодовом слове внутреннего кода превышает корректирующую способность этого кода, то для соответствующего этому слову символа внешнего кода выносится вердикт «стирание».

Если же в кодовом слове внутреннего кода, неверно исправленном на предыдущем шаге, на текущем шаге вновь обнаружены ошибки, то это слово становится таким, каким оно было на предыдущем шаге. В результате чего, к началу каждого следующего шага вероятности появления ошибок и стираний в исходной матрице $E_X^{(i-1)} \to E_X^{(i)}$ изменяются по сравнению с предыдущим шагом. Отметим, что обнаружение происходит только внутренними кодами, и только потом, на следующей итерации, обнаруженные ошибки будут учтены.

Соответственно, при декодировании на каждом следующем шаге используется информация о стертых и ошибочных символах, полученная на предыдущем шаге. Это означает, что на следующих шагах декодирования мы можем брать меньшую избыточность внешних кодов, так как часть ошибок и стираний уже исправлена. Аналогичные действия производятся на всех последующих шагах.

Помимо этого, в третьей главе рассмотрены теоретические методы, позволяющие оценить вероятность неверного декодирования ОЛО-3 кода и подобрать оптимальные параметры при известных условиях передачи данных. Под вероятностью неверного декодирования в данном случае подразумевается сумма вероятностей ошибки и отказа от декодирования.

После вычислений получаем для *i*-ого шага вероятности стирания и ошибки в символах соответствующего промежуточного кода в виде

$$P_{Be}^{(i)} = P_C(t = \{d_i - 1, .., 0\}, e \geq d_i - t)$$

$$P_{Bt}^{(i)} = P_C(e = 0, ..., n - t, t \geq d_i) + P_C(e + t = d_i - 1),$$

где $P_C(t, e)$ – вероятность появления в столбце кодового слова t стираний и e ошибок.

Далее, вероятности стирания и ошибки в символах соответствующего внешнего кода записываются как

$$P_{eras}^{A} = P_B(e = 0..n - t, t > 0)$$

$$P_{err}^{A} = P_B(e > 0, t = 0).$$

Соответственно, общая формула для $i+1$-го внешнего кода при условии, что все предыдущие были верно декодированы, выглядит следующим образом

$$P_{Aerr}^{(i+1)} = \sum_{t=0}^{d_A^{(i+1)}-1} C_n^t P_{eras}^{At} \sum_{e=\left\lceil \left(d_A^{(i+1)} - t\right)/2 \right\rceil}^{n-t} C_{n-t}^e P_{err}^{Ae} (1 - \mathrm{P}_{err}^{A} - P_{eras}^{A})^{n-t-e} + \sum_{t=d_A^{(i+1)}}^{n} C_n^t P_{eras}^{At} (1 - P_{eras}^{A})^{n-t}.$$

ОЛО-3 код будет полностью верно декодирован, если все внешние коды, все промежуточные и последний внутренний код для всех столбцов декодированы верно, в противном случае в полученном после декодирования слове останутся искаженные столбцы. Запишем вероятности правильного и неправильного декодирования всего ОЛО-3 кода как

$$P_{\Sigma-} = P_{A-}^{(1)} + P_{A+}^{(1)}(1 - P_{B+}^{(1)n_a}) + P_{A+}^{(1)} P_{B+}^{(1)n_a} P_{A-}^{(2)} + P_{A+}^{(1)} P_{B+}^{(1)n_a} P_{A+}^{(2)} (1 - P_{B+}^{(1)n_a}) + ... +$$

$$+ P_{A+}^{(1)} P_{B+}^{(1)n_a} P_{A+}^{(2)} P_{B+}^{(2)n_a} \cdot ... \cdot P_{A+}^{(i-1)} P_{B+}^{(i-1)n_a} P_{A-}^{(i)} + P_{A+}^{(1)} P_{B+}^{(1)n_a} P_{A+}^{(2)} P_{B+}^{(2)n_a} \cdot ... \cdot$$

$$\cdot P_{A+}^{(i-1)} P_{B+}^{(i-1)n_a} P_{A+}^{(i)} (1 - P_{B+}^{(i)n_a}) + P_{A+}^{(1)} P_{B+}^{(1)n_a} P_{A+}^{(2)} P_{B+}^{(2)n_a} \cdot ... \cdot P_{A+}^{(i-1)} P_{B+}^{(i-1)n_a} P_{A+}^{(i)} P_{B+}^{(i)n_a} (1 - P_{C+}^{(i)n_a n_b})$$

$$P_{\Sigma+} = P_{A+}^{(1)} P_{B+}^{(1)n_a} P_{A+}^{(2)} P_{B+}^{(2)n_a} \cdot ... \cdot P_{A+}^{(i-1)} P_{B+}^{(i-1)n_a} P_{A+}^{(i)} P_{B+}^{(i)n_a} P_{C+}^{(i)n_a n_b},$$

где $P_{A-}^{(i)}$ и $P_{A+}^{(i)}$ – вероятность неверного и верного декодирования i-ым внешним кодом, а $(1 - P_{B+}^{(i)n_a})$ – вероятность того, что при декодировании i-ым внутренним кодом хотя бы в одном столбце произошла ошибка.

Из структуры формул для вероятностей правильного и неправильного декодирования всего ОЛО-3 кода мы можем получить следующее утверждение:

Утверждение 2.

1. *Для того, чтобы правильно выбрать порядок кода m, необходимо определить, какая избыточность $d_C^{(i)}$ для внутреннего кода на первом шаге же обеспечит ошибку меньше $\varepsilon / d_C^{(i)}$.*
2. *Чтобы обеспечить нужный уровень вероятности ошибки, необходимо соблюсти условие:*

$$P_{err}^{C} = \sum_{t=0}^{d_C^{(i)}-1} C_n^t p_t^t \sum_{e=\left\lceil \left(d_C^{(i)}-t\right)/2 \right\rceil}^{n-t} C_{n-t}^e p_e^e (1-p_e-p_t)^{n-t-e} + \sum_{t=d_C^{(i)}}^{n} C_n^t p_t^t (1-p_t)^{n-t},$$

$$P_{\Sigma err} = 1-(1-P_{err}^{C})^{n_A n_B} \le \frac{d_C^{(i)}-1}{2\varepsilon+1},$$

где P_{err}^{C} – ошибка декодирования для одного столбца, а $P_{\Sigma err}$ – ошибка декодирования для всего слова длины $n_A n_B$.

Кроме того, аналогично с двумерным случаем, мы оцениваем вычислительную сложность для ОЛО-3 кодов. Аналогично с двумерным случаем, получаем следующую лемму:

Лемма 2. *Количество операций при декодировании ОЛО-3 кода можно оценить как*

$$N = \log(q)\Big[C_C n_A n_C n_B \log^2(n_C) + C_B n_A n_B \log^2(n_B) + C_A n_A \log^2(n_A) + \sum_{i=1}^{L} (n_A n_B n_C r_C^{(i)} + n_A n_B r_B^{(i)} r_C^{(i)} + n_A r_A^{(i)} r_B^{(i)} r_C^{(i)}) \Big],$$

где C_C, C_B и C_A – некоторые коэффициенты, явным образом не зависящие от длины кода.

Доказательство. Для кодов Рида-Соломона длины *n* над полем *GF(q)* количество вычислений при декодировании можно оценить как $\log(q) \cdot C \cdot n \cdot \log^2(n)$. При декодировании ОЛО-3 кодов на каждом шаге сначала будет происходить исправление стираний и ошибок при помощи одного внешнего кода длины n_A над полем $GF(q^{r_C r_B})$, а затем при помощи n_A промежуточных кодов длины n_B над полем $GF(q^{r_C})$. Далее имеет место исправление и обнаружение ошибок $n_A n_B$ внутренними кодами длины n_C над полем *GF(q)*.

Также необходимо учесть, что на каждом шаге декодирования ОЛО-3 вначале нужно найти синдромы внутреннего и промежуточного кодов, а для их вычисления нам необходимо совершить умножение на соответствующие проверочные матрицы. Таким образом, количество операций при декодировании ОЛО-3 кода можно оценить как

$$\sum_{i=1}^{L}\Big[\log(q)\cdot\left(n_A n_C n_B 2C_C^{(i)}\log^2(n_C)+r_C{}^{(i)}n_C n_A n_B\right)+$$

$$+\log(q^{r_C{}^{(i)}})\cdot\left(C_B^{(i)}n_B n_A\log^2(n_B)+r_B{}^{(i)}n_B n_A\right)+\log(q^{r_C{}^{(i)}r_B{}^{(i)}})\cdot\left(C_A^{(i)}n_A\log^2(n_A)+r_A{}^{(i)}n_A\right)\Big]=$$

$$=\log(q)n_A\sum_{i=1}^{L}(n_C n_B 2C_C^{(i)}\log^2(n_C)+r_C{}^{(i)}n_C n_B+$$

$$+C_B^{(i)}n_B r_C{}^{(i)}\log^2(n_B)+r_B{}^{(i)}r_C{}^{(i)}n_B+C_A^{(i)}r_B{}^{(i)}r_C{}^{(i)}\log^2(n_A)+r_B{}^{(i)}r_C{}^{(i)}r_A{}^{(i)})=$$

$$=\log(q)\Big[C_C n_A n_C n_B\log^2(n_C)+C_B n_A n_B\log^2(n_B)+C_A n_A\log^2(n_A)+$$

$$\sum_{i=1}^{L}(n_A n_B n_C r_C{}^{(i)}+n_A n_B r_B{}^{(i)}r_C{}^{(i)}+n_A r_A{}^{(i)}r_B{}^{(i)}r_C{}^{(i)})\Big],$$

что, собственно, и требовалось доказать. В качестве C_C, C_B и C_A выступают некоторые коэффициенты, явным образом не зависящие от длины кода.

В **четвертой главе** приводятся результаты моделирования для ОЛО-3 кодов. Количество проведенных испытаний для каждого значения отношения сигнал/шум равно 10^6. Параметры для рассматриваемых в диссертационной работе ОЛО-3 кодов выбраны в соответствии с описанными выше теоретическими методами. Все вычисления были проделаны в среде MATLAB. Было получено, что теоретическая оценка для вероятности неверного декодирования близка к соответствующей вероятности, полученной в ходе моделирования, и является для нее довольно точной оценкой сверху.

На рис. 4, для примера, приведен график зависимости вероятности ошибки + отказа от декодирования от отношения сигнал/шум на бит в канале для ОЛО-3 кода (262144, 259492), m=14 и избыточностью 1% при передаче по каналу с аддитивным гауссовским шумом. Кружочками обозначены результаты теоретических расчетов, квадратами – результаты моделирования. Для сравнения на этом же графике также изображена теоретическая оценка

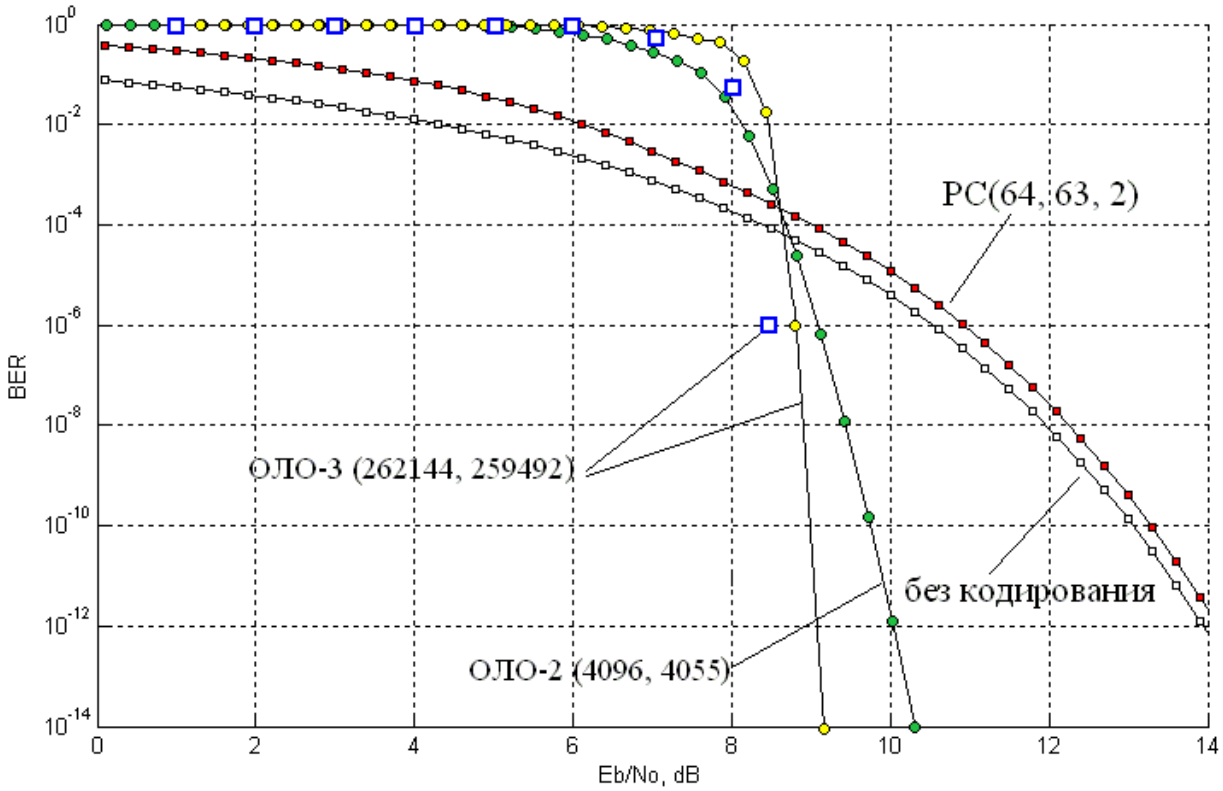

Рис. 4. Зависимость вероятности отказа от декодирования + ошибки от отношения сигнал/шум в канале для ОЛО-3 (262144, 259492), m=14 избыточностью 1% при передаче по гауссовскому каналу.

зависимости вероятности отказа от декодирования + ошибки на бит передаваемой информации для кода Рида-Соломона с параметрами (64, 63, 2), имеющего наиболее близкую избыточность; в данном случае этот код не исправляет ни одной ошибки, а лишь добавляет стирания. Кроме того, на том же графике приведена теоретическая оценка зависимости вероятности отказа от декодирования + ошибки для ОЛО-2 кода той же избыточности с параметрами (4096, 4055). Как мы можем убедиться, при уровне ошибки 10^{-12}, выигрыш ОЛО-3 кода составляет порядка 5 дБ по сравнению с кодом Рида-Соломона, и ≈1 дБ по сравнению с ОЛО-2 кодами той же избыточности.

В **заключении** диссертации сформулированы основные результаты диссертационной работы.

Основные результаты

В диссертации было проведено исследование свойств кодов с локализацией ошибок на основе кодов Рида-Соломона. Был проведен анализ суще-

ствующих способов кодирования и декодирования обобщенными кодами с локализацией ошибок.

Предложена новая разновидность ОЛО-кодов – трехмерные обобщенные коды с локализацией ошибок, для этих кодов разработаны алгоритмы кодирования/декодирования.

Также было проведена разработка методов, позволяющих выбрать оптимальные параметры для ОЛО-2 и ОЛО-3 кодов. С помощью описанных в работе методов можно выбрать код, обеспечивающий максимальную скорость передачи при заданных входной и выходной вероятностях ошибки.

В работе было проведено моделирование некоторых сигнально-кодовых конструкций на основе ОЛО-2 и ОЛО-3 кодов, и проведен анализ полученных результатов.

Основные результаты диссертации изложены в нижеследующих опубликованных работах.

Список публикаций

Статьи в рецензируемых журналах:

1. Кобозева И. Г., Зяблов В. В., “Оценка вероятности неправильного декодирования обобщенных кодов с локализацией ошибок”, “Информационно-управляющие системы”, 2013, №1, стр. 47-53
2. Кобозева И. Г., Зяблов В. В., “ Исследование сигнально-кодовых конструкций на основе трехмерных кодов с локализацией ошибок”, “Информационные процессы ”, 2013, №1, стр. 1-18.
3. Kobozeva I. G., Zyablov V. V., “Investigation of Signal- Code Structures Based on 3D Error-Locating Codes”, Journal of Communications Technology and Electronics, 2013, Vol. 58, No. 6, pp. 648–660.

Тезисы докладов на конференциях:

4. Kobozeva I., Zyablov V., “Using GEL Codes for Optical Channel”, труды XII симпозиума по проблеме избыточности в информационных системах, 2009, с. 126-129.

5. Кобозева И. Г., Зяблов В. В., “Декодирование обобщенных кодов с локализацией ошибок”, труды конференции ИТИС’09, с. 170-174.
6. Кобозева И. Г., Зяблов В. В., “Декодирование трехмерных обобщенных кодов с локализацией ошибок”, труды конференции ИТИС’10, с. 99-103.
7. Кобозева И. Г., Зяблов В. В., “Комбинаторные оценки кодового расстояния для ОЛО-кодов”, труды конференции ИТИС’11, с. 123-126.
8. Кобозева И. Г., Зяблов В. В., “Кодирование трехмерными обобщенными кодами с локализацией ошибок”, труды конференции ИТИС’12, с. 129-132.

Кобозева Ирина Геннадьевна

Исследование сигнально-кодовых конструкций на основе обобщенных кодов с локализацией ошибок

АВТОРЕФЕРАТ

Подписано в печать 15.11.2013 г.

Заказ № 9115 Тираж: 100 экз.

Печать трафаретная

Типография «11-й ФОРМАТ»

ИНН 7726330900

115230, Москва, Варшавское ш., 36

(499) 788-78-56

www.autoreferat.ru