УДК 681.391

Г.В. Овечкин, П.В. Овечкин

ИСПОЛЬЗОВАНИЕ НЕДВОИЧНОГО МНОГОПОРОГОВОГО ДЕКОДЕРА В КАСКАДНЫХ СХЕМАХ КОРРЕКЦИИ ОШИБОК

Анализируются возможности недвоичных многопороговых декодеров (qМПД) самоортогональных кодов в q-ичных симметричных каналах. Представлена новая каскадная схема кодирования, состоящая из qМПД и модифицированного недвоичного кода Хэмминга. Показано, что использование предложенной каскадной схемы позволяет уменьшить вероятность ошибки декодирования в области эффективной работы qМПД на три и более порядков.

Ключевые слова: *помехоустойчивое кодирование, системы передачи данных, системы хранения данных, недвоичные коды, коды Рида-Соломона, недвоичный многопороговый декодер, каскадные схемы коррекции ошибок, коды Хэмминга.*

Введение. К современным системам передачи цифровых данных предъявляются очень жесткие требования по безошибочности передачи информации. Для обеспечения таких требований используют методы помехоустойчивого кодирования, применение которых позволяет улучшать многие важные характеристики систем передачи данных, например экономить мощность передатчика, увеличивать дальность связи, уменьшать размеры антенн и др. Немаловажную роль помехоустойчивые коды играют и в системах хранения данных, в которых необходимо обеспечивать высокую надежность долговременного хранения информации на носителе.

В настоящее время специалисты в области помехоустойчивого кодирования проявляют большой интерес к недвоичным кодам, работающим с цифровыми данными на уровне символов, например с байтами информации. Недвоичные коды применяются в каналах с группирующимися ошибками, в качестве составляющих элементов каскадных кодов, для защиты от ошибок информации на различного рода носителях (CD, DVD, Blu-ray и др.).

Анализ недвоичных корректирующих кодов и алгоритмов их декодирования показал, что широкое применение в реальных системах передачи и хранения информации из недвоичных кодов нашли только коды Рида-Соломона. Однако декодеры коротких кодов Рида-Соломона, которые и применяются на практике, не могут обеспечить высокую эффективность декодирования, а для длинных кодов Рида-Соломона сложность создания декодера оказывается слишком высокой.

Среди других методов коррекции ошибок наиболее перспективным является метод недвоичного многопорогового декодирования, предложенный В.В. Золотарёвым [1 – 4]. Отличительными особенностями данного метода являются линейная зависимость сложности реализации от длины кода и высокая эффективность декодирования. Поэтому недвоичный многопороговый декодер (*q*МПД) может применяться в высокоскоростных системах передачи и хранения больших объемов информации.

Эффективность недвоичных многопороговых декодеров. Рассмотрим эффективность *q*МПД при различных параметрах кода и размерах символа в *q*-ичном симметричном канале (*q*СК). В таком канале каждый символ искажается независимо с вероятностью P_0, причем при искажении символ с равной вероятностью переходит в один из q–1 других символов. Здесь q обозначает размер используемого алфавита.

На рисунке 1 представлены полученные с помощью компьютерного моделирования зависимости вероятности ошибки декодирования qМПД от вероятности ошибки в qСК для кодов с кодовой скоростью R=1/2. Здесь кривые «qМПД(n=4000, q=256)» и «qМПД(n=32000, q=256)» соответствуют характеристикам qМПД для кодов длиной 4000 и 32000 однобайтовых символов (q=256). При этом в процессе

исправления ошибок использовалось от 5 до 15 итераций декодирования. Для сравнения на рисунке 1 кривой «РС(n=255, q=256)» показаны характеристики (255, 128) кода Рида-Соломона [5, 6], заданного над полем GF(256). Отметим, что qМПД обеспечивает гораздо лучшую эффективность, чем коды Рида-Соломона для символов того же размера благодаря большей длине используемых кодов и хорошей сходимости решений qМПД к решению оптимального декодера. Отметим, что для достижения с помощью qМПД таких результатов требуется очень тщательно выбирать применяемые коды, основным критерием при отборе которых является степень устойчивости к эффекту размножения ошибок, и параметры декодирования [2].

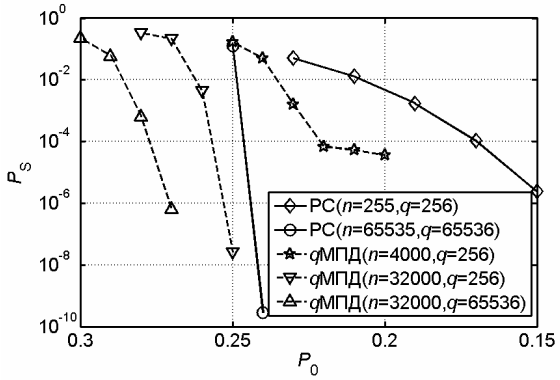

Рисунок 1 – Характеристики недвоичных декодеров в *q*СК

Существенным преимуществом *q*МПД над другими методами коррекции ошибок является то, что он позволяет работать с символами практически любого размера, обеспечивая при этом такую же высокую корректирующую способность. Характеристики *q*МПД для кода с R=1/2, n=32000 и $q=2^{16}$ (двухбайтовые символы) представлены на рисунке 1 кривой «*q*МПД(n=32000, q=65536)». Видно, что характеристики *q*МПД такого кода превосходят возможности декодера кодов Рида-Соломона длиной 65535 двухбайтовых символов [кривая «РС(n=65535, q=65536)»].

Таким образом, недвоичный аналог алгоритма МПД может обеспечить при весьма высоких уровнях шума вероятности ошибки декодирования, в ряде случаев недоступные для кодов Рида-Соломона сколь угодно большой длины. При этом сложность реализации такого алгоритма оказывается незначительной, линейно растущей с длиной кода, т.е. теоретически минимально возможной [2].

Вместе с тем, результаты проведенного исследования показали, что для *q*МПД свойственно наличие области насыщения вероятности ошибки [правая часть кривой «*q*МПД(n=4000, q=256)»], в которой скорость уменьшения вероятности ошибки декодирования существенно замедляется. Это усложняет получение очень малых вероятностей ошибки декодирования (порядка 10^{-12}), требуемых, например, в системах хранения данных. В связи с этим **актуальной является задача** дополнительного уменьшения вероятности ошибки декодирования *q*МПД.

Наиболее общим подходом к решению данной задачи является применение *q*МПД в составе каскадных схем коррекции ошибок. Причем совместно с *q*МПД следует использовать наиболее простые методы коррекции ошибок, чтобы сложность декодера каскадного кода осталась невысокой.

Каскадные схемы коррекции ошибок на базе *q*МПД. Как известно, коды Хэмминга являются одними из самых простых кодов, которые могут быть использованы в каскадных схемах кодирования. Однако применение известных недвоичных кодов Хэмминга в составе каскадных схем кодирования сопряжено с определенными трудностями. В процессе кодирования и декодирования недвоичных кодов Хэмминга все операции выполняются в расширенных полях Галуа. Существенным недостатком использования полей Галуа для кодирования и декодирования информации является то, что при большом размере символа реализация операций в таком поле оказывается слишком сложной.

Кроме того, длина недвоичного кода Хэмминга рассчитывается как

$$n = \frac{q^m - 1}{q - 1},$$

где q – размер алфавита, m – число проверочных символов. В этом случае уже при использовании всего двухбайтовых символов длина кода оказывается слишком большой (более 65000 символов). В результате декодер таких недвоичных кодов Хэмминга может исправлять только одну ошибку в блоке из 65000 символов, что приводит к нецелесообразности использования данных кодов в составе каскадных схем кодирования.

Поэтому в каскадных схемах совместно с *q*МПД предлагается использовать модифицированные недвоичные коды Хэмминга. Данные коды задаются такой же проверочной матрицей **H**, как и двоичные коды Хэмминга, но при кодировании и декодировании все операции выполняются в кольце целых чисел (по модулю q). Предлагаемые модифицированные недвоичные коды Хэмминга свободны от недостатков

известных недвоичных кодов Хэмминга. Кроме того, декодер таких модифицированных кодов Хэмминга может исправлять не только одну, но и в некоторых случаях две ошибки, что невозможно для обычных недвоичных кодов Хэмминга.

Выявим ситуации, в которых возможно исправление двух ошибок декодером модифицированных недвоичных кодов Хэмминга.

Если синдром кода содержит только нули и символы со значением a, b и c, причем $c=(a+b) \bmod q$, то с большой вероятностью в блоке было две ошибки со значениями a и b. Синдром $\mathbf{s}$ в этом случае может быть представлен как

$$\mathbf{s} = \mathbf{s}_a + \mathbf{s}_b ,$$

где $\mathbf{s}_a$ – синдром при ошибке a, содержащий только нули и элементы со значением a; $\mathbf{s}_b$ – синдром при ошибке b, содержащий только нули и элементы со значением b.

В этом случае на значение a корректируется такой символ принятого из канала сообщения, для которого столбец матрицы $\mathbf{H}$ совпадает с вектором $\mathbf{s}_a$, где все элементы a заменены на единицы; на значение b корректируется такой символ принятого из канала сообщения, для которого столбец матрицы $\mathbf{H}$ совпадает с вектором $\mathbf{s}_b$, где все элементы b заменены на единицы.

Результаты исследования показали, что декодер модифицированного недвоичного кода Хэмминга исправляет две ошибки только в том случае, если их позиции в двоичном представлении имеют вид, показанный на рисунке 2. То есть в k-м бите позиции имеют общую единицу; в i-м бите позиция первой ошибки имеет единицу, а позиция второй – ноль; в t-м бите позиция первой ошибки имеет ноль, а позиция второй – единицу.

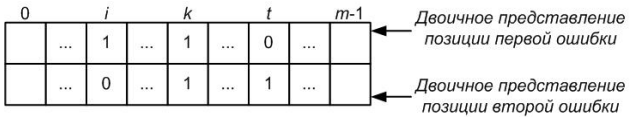

Рисунок 2 – Двоичное представление позиций ошибок

Доля блоков с двумя ошибками, исправляемых декодером модифицированных недвоичных кодов Хэмминга, может быть рассчитана как

$$1 - \frac{3^{m+1} - 3 \cdot 2^m + 1}{4^m}.$$

Например, для модифицированных недвоичных кодов Хэмминга с длиной N_h=255, то есть m=8, исправляется порядка 71% блоков с двумя ошибками.

Несколько большей корректирующей способностью обладают модифицированные недвоичные расширенные коды Хэмминга, декодер которых почти всегда исправляет конфигурацию из двух ошибок в блоке. Данные коды построены на основе двоичных расширенных кодов Хэмминга по тому же принципу, что и модифицированные недвоичные коды Хэмминга. Конфигурация из двух ошибок в блоке такого кода не исправляется только в том случае, если значения ошибок двух символов совпадают между собой или их сумма равна q.

Доля блоков с двумя ошибками, исправляемых декодером модифицированных недвоичных расширенных кодов Хэмминга, может быть рассчитана как

$$\frac{q-2}{q}.$$

Например, для модифицированных недвоичных расширенных кодов Хэмминга с q=256 исправляется порядка 99.2% блоков с двумя ошибками, а при переходе к двухбайтовым символам, то есть q=65536, исправляется порядка 99.997% блоков с двумя ошибками.

Отметим, что предложенные модифицированные недвоичные коды Хэмминга используются во внешнем каскаде каскадного кода. Кроме того, чтобы уменьшить потери в энергетике, возникающие из-за необходимости передачи дополнительных проверочных символов внешнего кода, предполагается использование модифицированного кода Хэмминга достаточно большой длины (N_h≥127).

Аналитические оценки вероятности ошибки каскадной схемы. Для предварительной оценки эффективности предложенной каскадной схемы кодирования, состоящей из qМПД и модифицированного недвоичного кода Хэмминга, получим нижнюю оценку вероятности ошибки оптимального декодирования каскадного кода при работе в q-ичном симметричном канале. Пусть при передаче по qСК каждый символ кодового слова с вероятностью P_0 искажается. После этого принятое из канала сообщение поступает на вход qМПД. Далее после многопорогового декодирования на вход декодера модифицированных кодов Хэмминга длинной N_h поступает поток символов с вероятностью ошибки P_s, которая может быть оценена по формулам, представленным в [1].

События, приводящие к ошибкам на выходе декодера модифицированного кода Хэмминга:

– в блоке из N_h символов есть два ошибочных символа, остальные символы правильные. Причем структура ошибок не должна иметь вид, показанный на рисунке 2. В этом случае после декодирования в блоке останутся две ошибки. Вероятность ошибки на выходе декодера:

$$P_{errh1} = (1-P_S)^{N_h-2} \cdot P_S^2 \cdot (N_h-1) \cdot \frac{3^{m+1} - 3 \cdot 2^m + 1}{4^m};$$

– в блоке из N_h символов есть три ошибочных символа, остальные символы правильные. В этом случае после декодирования в блоке останутся три ошибки. Вероятность ошибки на выходе декодера:

$$P_{errh2} = \frac{(1-P_S)^{N_h-3} P_S^3 (N_h-1)(N_h-2)}{2}.$$

Вероятности других типов событий несущественно влияют на результирующую вероятность ошибки.

Таким образом, нижняя оценка вероятности символьной ошибки оптимального декодирования каскадного кода определяется суммой найденных выше вероятностей:

$$P_{err} = P_{errh1} + P_{errh2}. \qquad (1)$$

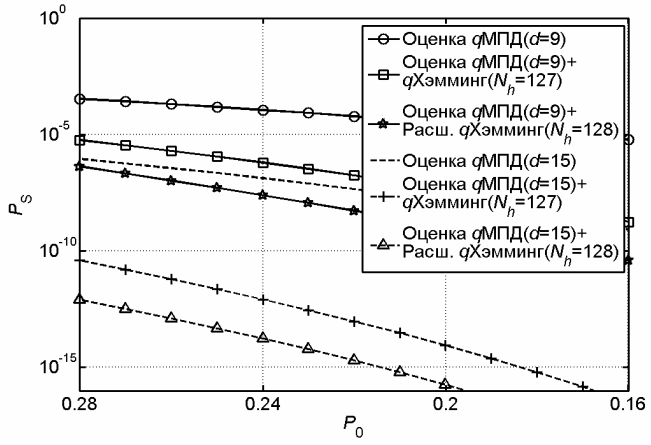

Рисунок 3 – Оценка вероятности ошибки на выходе каскадной схемы

На рисунке 3 приведены оценки эффективности работы каскадных схем кодирования в q-ичном симметричном канале, рассчитанные по формуле 1. Здесь используются qМПД для кодов с кодовой скоростью R=1/2, минимальными кодовыми расстояниями $d = 9$ и $d = 17$, размером алфавита q=256. Совместно с ними использовался модифицированный код Хэмминга с длиной блока $N_h = 127$. Из данных оценок следует, что применение предложенной схемы позволит уменьшить вероятность ошибки декодирования в области эффективной работы qМПД на три и более порядков для кода с d=9 и на пять и более порядков для кода с d=15.

Оценка эффективности каскадной схемы кодирования, состоящей из qМПД и модифицированного недвоичного расширенного кода Хэмминга, выполняется аналогично. На вход декодера модифицированного расширенного кода Хэмминга также поступает поток символов после недвоичного многопорогового декодирования с вероятностью ошибки P_s. Декодер модифицированного недвоичного расширенного кода Хэмминга с длиной блока N_h имеет минимальное кодовое расстояние d=4 и может исправлять одну и в большинстве случаев две ошибки в кодовом слове.

Рассмотрим события, приводящие к ошибкам на выходе декодера модифицированного расширенного кода Хэмминга:

– в блоке из N_h символов есть два ошибочных символа, остальные символы правильные. Значения ошибок двух символов совпадают между собой или их сумма равна q. В этом случае после декодирования в блоке останутся две ошибки. Вероятность ошибки на выходе декодера:

$$P_{exth1} = \frac{2 \cdot (1-P_S)^{N_h-2} P_S^2 (N_h-1)}{q};$$

– в блоке из N_h символов есть три ошибочных символа, остальные символы правильные. В этом случае после декодирования в блоке останутся три ошибки. Вероятность ошибки на выходе декодера:

$$P_{exth2} = \frac{(1-P_S)^{N_h-3} P_S^3 (N_h-1)(N_h-2)}{2}.$$

Вероятности других типов событий несущественно влияют на результирующую вероятность ошибки.

Таким образом, нижняя оценка вероятности символьной ошибки оптимального декодирования каскадной схемы с декодером модифицированного недвоичного расширенного кода Хэмминга определяется суммой найденных выше вероятностей:

$$P_{exterr} = P_{exth1} + P_{exth2}. \qquad (2)$$

На рисунке 3 приведены оценки эффективности работы данной каскадной схемы кодирования в q-ичном симметричном канале, рассчитанные по формуле 2. Здесь в каскаде с qМПД используется декодер модифицированного расширенного кода Хэмминга с длиной блока $N_h = 128$. Из данных оценок следует, что применение предложенной схемы позволит уменьшить вероятность ошибки декодирования в области эффективной работы qМПД на четыре и более порядков для кода с d=9 и на семь и более порядков для кода с d=15.

Экспериментальные оценки вероятности ошибки каскадной схемы. На рисунке 4 показаны экспериментальные характеристики каскадных кодов, состоящих из *q*МПД для кода с минимальным кодовым расстоянием *d*=9 и модифицированных недвоичных кодов Хэмминга с длиной блока $N_h = 127$ и модифицированных недвоичных расширенных кодов Хэмминга с длиной блока $N_h = 128$. Экспериментальная оценка эффективности каскадных схем кодирования проводилась для *q*-ичного симметричного канала и выполнялась с помощью разработанных программных средств. Анализ зависимостей показывает, что применение совместно с недвоичным многопороговым декодером в области его эффективной работы декодера модифицированного недвоичного кода Хэмминга позволяет снизить вероятность ошибки на выходе каскадной схемы на три и более порядков. При переходе к модифицированному расширенному недвоичному коду Хэмминга можно снизить вероятность ошибки еще более чем на один порядок, причем сложность итоговой каскадной схемы при этом изменится незначительно. Также отметим, что полученные оценки (1) и (2), представленные на рисунке пунктиром, оказываются достаточно точными для предварительного оценивания характеристик каскадных схем кодирования.

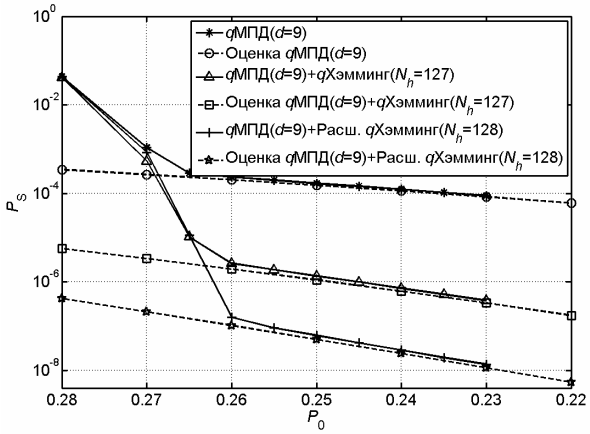

Рисунок 4 – Экспериментальная оценка эффективности использования каскадной схемы кодирования

Заключение. В статье рассмотрены возможности *q*МПД алгоритмов, которые оказываются по вероятности ошибки и по числу операций декодирования на много порядков лучше кодов Рида-Соломона, по праву считавшихся лучшими недвоичными кодами в течение почти полувека. Предложена каскадная схема на базе недвоичного самоортогонального кода и модифицированного недвоичного кода Хэмминга, использование которой позволяет уменьшить частоту появления ошибок на выходе *q*МПД в области его эффективной работы более чем на 3 порядка.

В результате недоступный ранее уровень помехоустойчивости, получаемый с помощью алгоритмов МПД разных типов, позволяет решать задачи обеспечения высокой надежности передачи и хранения данных без какой-либо дополнительной доработки этих алгоритмов или всего лишь при незначительной их адаптации к возможным дополнительным требованиям, возникающим в крупномасштабных цифровых системах.

Дополнительную информацию о *q*МПД можно найти на веб-сайте [4].

Работа выполнена при финансовой поддержке РФФИ (грант №08-07-00078).

Библиографический список

1. *Золотарев В.В., Овечкин Г.В.* Помехоустойчивое кодирование. Методы и алгоритмы: Справочник. М.: Горячая линия – Телеком, 2004. 126 с.

2. *Золотарев В.В.* Теория и алгоритмы многопорогового декодирования. М.: Радио и связь, Горячая линия – Телеком, 2006.

3. *Золотарев В.В.* Обобщение алгоритма МПД на недвоичные коды // Мобильные системы, 2007, №3, С.39–42.

4. http://www.mtdbest.iki.rssi.ru

5. *Reed I.S., Solomon G.* Polynomial codes over certain finite fields // J. Soc. Industrial Appl. Math., 1960, vol.8, PP.300–304.

6. *Sudan M.* Decoding of Reed Solomon codes beyond the error-correction bound // Journal of Complexity, 1997, vol.13, PP.180–193.