ЭФФЕКТИВНОЕ НЕДВОИЧНОЕ МНОГОПОРОГОВОЕ ДЕКОДИРОВАНИЕ ПОМЕХОУСТОЙЧИВЫХ КОДОВ ДЛЯ СИСТЕМ ДИСТАНЦИОННОГО ЗОНДИРОВАНИЯ ЗЕМЛИ

Р.Р. Назиров, В.В. Золотарёв, Г.В. Овечкин, П.В. Овечкин, И.В. Чулков

Институт космических исследований РАН,
117997, Москва, ГСП=7, Профсоюзная ул., д.84/32
E-mails: zolotasd@yandex.ru, g_ovechkin@mail.ru, chulkov@iki.rssi.ru

Анализируются эффективность недвоичных многопороговых алгоритмов декодирования (qМПД) самоортогональных кодов. Представлены новые характеристики qМПД для новых кодовых схем с параллельным кодированием, обеспечивающих эффективную работу при существенно большем уровне шума в канале. Описано применение qМПД для защиты файлов от искажений.

Ключевые слова: помехоустойчивое кодирование, системы передачи данных, системы хранения данных, q-ичный симметричный канал, недвоичные коды, недвоичные самоортогональные коды, недвоичный многопороговый декодер

Введение

В системах дистанционного зондирования Земли (ДЗЗ) важной и особенно сложной является задача быстрой записи и высоконадежного хранения принятой дискретной информации в базах данных очень большого объема. Эта задача наиболее эффективно решается методами помехоустойчивого кодирования, которые применяются для исправления ошибок, возникающих при передаче данных по каналам с шумами или при их длительном хранении на различного рода носителях информации. В настоящее время в литературе наибольшее внимание уделяется методам коррекции ошибок в двоичных данных. Однако во многих системах (ДЗЗ и др.), проще обрабатывать данные, имеющие байтовую структуру. В подобных системах для защиты данных от ошибок целесообразно применение недвоичных помехоустойчивых кодов.

На сегодняшний день в теории кодирования известен ряд недвоичных кодов, различающихся корректирующей способностью, вносимой избыточностью, сложностью декодирования и многими другими важными параметрами. Среди них практическое применение в реальных системах нашли только коды Рида-Соломона (РС) [1], обладающие рядом положительных свойств. В частности коды РС характеризуются тем, что для исправления в пределах кодового слова любой комбинации из t символьных ошибок достаточно использовать лишь $2t$ проверочных символов. Для коротких кодов РС существуют эффективные алгоритмы декодирования, в полной мере использующие корректирующие возможности кода [2]. Сложность реализации наиболее простых из них пропорциональна $n \cdot \log^2 n$ [2], где n – длина кода. Под сложностью реализации здесь и далее понимается число арифметических операций, требуемых для декодирования кодового блока. Однако короткие коды РС часто не могут обеспечить требуемой в настоящее время степени защиты данных от ошибок, а для длинных кодов РС практически невозможно создать эффективные декодеры. В последнее время зарубежные специалисты стали активно развивать декодеры недвоичных низкоплотностных кодов [3–5]. Данные методы обладают очень высокой корректирующей способностью, однако сложность их реализации, особенно при больших размерах алфавита q, оказывается слишком большой для применения в реальных системах.

Гораздо лучшей эффективностью обладают недвоичные многопороговые декодеры (qМПД) [6–9], разрабатываемые в Институте космических исследований РАН и Рязанском государственном радиотехническом университете. Предложенные еще в 1984 году qМПД обладают линейной сложностью реализации и позволяют практически оптимально декодировать даже очень длинные, потенциально гораздо более эффективные коды. В результате, применение недвоичных МПД вместо кодов РС может на много порядков повысить уровень защиты информации от ошибок при одновременном существенном упрощении процесса коррекции ошибок.

Недвоичные многопороговые декодеры

Недвоичные многопороговые декодеры предназначены для декодирования недвоичных самоортогональных кодов [7, 8, 9]. qМПД, как и обычные двоичные МПД [8], обладают свойством стремления к решению оптимального декодера при линейной от длины кода сложности реализации, которая свойственна только пороговым процедурам. В отличие от кодов РС для qМПД никаких ограничений по длине кода вообще нет, поскольку длина кода n и размер алфавита q в недвоичных кодах с мажоритарным декодированием совершенно не зависят друг от друга. При этом сложность декодирования кодового блока qМПД оказывается пропорциональной $n \cdot d \cdot I$, где n – длина кода, d – кодовое расстояние (обычно $d \leq 20$), I – число итераций декодирования (обычно $I \leq 30$).

Рассмотрим характеристики недвоичных многопороговых декодеров в q-ичном симметричном канале (qСК). 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. Подобная модель, например, соответствует каналу с пакетами ошибок при использовании перемежения/деперемежения на уровне символов. Зависимости вероятности символьной ошибки P_s после декодирования от вероятности символьной ошибки P_0 в qСК для кодов с кодовой скоростью R=1/2 представлены на рис. 1. Здесь кривыми *1* и *2* показаны характеристики qМПД для кодов с длиной блока n=4000 и 60000 символов при использовании 8-ми битовых символов (размер алфавита q=256). Объем моделирования в нижних точках данных графиков составлял от $5\cdot10^{10}$ до $2\cdot10^{12}$ символов, что свидетельствует о крайней простоте метода. Для сравнения на данном рисунке кривой 4 показаны характеристики (255, 128) кодов РС для q=256. Из рис. 1 видно, что эффективность qМПД оказывается гораздо лучше эффективности кодов РС для символов такого же размера. При увеличении длины блока qМПД разница в эффективности становится еще более существенной. Характеристики qМПД при использовании двухбайтовых символов представлены на рис. 1 кривой *3*. Здесь использовался код с R=1/2 и n=32000 символов. Отметим, что очень простой для реализации qМПД декодер для двухбайтового кода длины 32000 оказывается способным обеспечить помехоустойчивость, принципиально недостижимую даже для кода РС длины 65535 двухбайтовых символов (кривая *5* на рис. 1), декодер для которого не подлежит реализации в обозримом будущем. При этом qМПД для двухбайтовых символов практически ни в чем не сложнее однобайтового, так как даже обычные микропроцессоры одинаково просто и быстро работают и с однобайтовыми символами, и с 2-х и даже с 8-байтовыми символами.

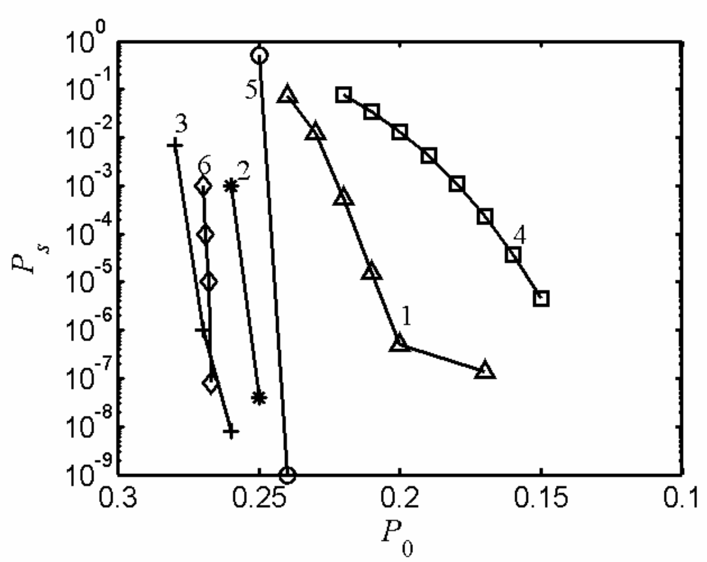

Рис. 1. Характеристики недвоичных кодов с кодовой скоростью *R*=1/2 в *q*СК

Для систем ДЗЗ большой интерес представляют малоизбыточные помехоустойчивые коды. Характеристики qМПД для недвоичных кодов с R=7/8, n=100000 символов и q=256 представлены на рис. 2 кривой *1*, а характеристики кодов РС с R=7/8 и q=256 отражены кривой *4*. Здесь также видно заметное преимущество qМПД над кодами РС. Аналогичная ситуация наблюдается и при использовании кодов с еще более высокой кодовой скоростью R=19/20. Для данной кодовой скорости при q=256 эффективность qМПД показана кривой *3*, а для кодов РС – кривой *6*. Такие же высокие характеристики обеспечивает qМПД при использовании алфавита большего объема. На рис. 2 кривой *2* представлена эффективность qМПД для кода с R=7/8 при использовании двухбайтовых символов (q=65536).

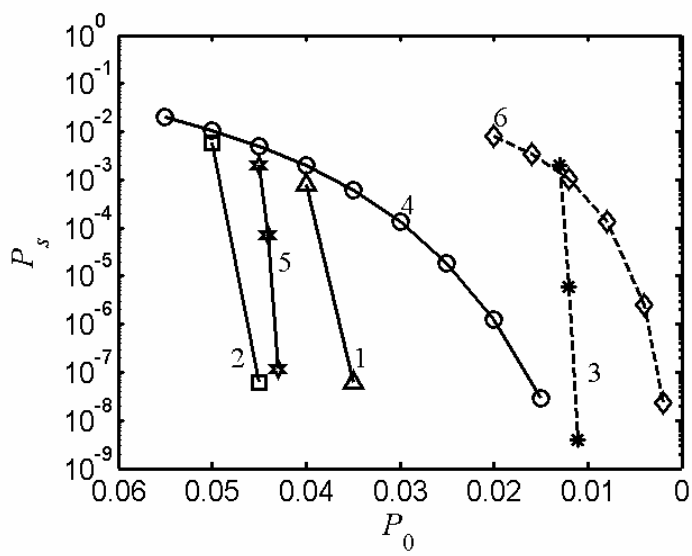

Рис. 2. Характеристики малоизбыточных недвоичных кодов в qСК

Следует заметить, 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 [8], который проявляется в том, что после первой ошибки декодирования существенно увеличивается вероятность последующих ошибок. Известно, что размножению ошибок в наименьшей степени подвержены коды для схем с параллельным кодированием [10]. В [11] показано, что оптимизируя структуру данных кодов можно еще улучшить эффективность работы qМПД. В частности, характеристики найденных в [11] кодов с q=256 и кодовыми скоростями 1/2 и 7/8 представлены на рис. 1 и 2 кривыми *6* и *5* соответственно. Видно, что данные коды обеспечивают эффективную работу при б*о*льших вероятностях ошибки в qСК, чем ранее представленные [9], при такой же сложности их декодирования. Еще большую эффективность практически без усложнения схемы коррекции ошибок можно получить при переходе к каскадным принципам кодирования. В [9] показано, что применение совместно с qМПД простейшего кода с контролем по модулю q позволяет на 1..3 порядков снизить вероятность ошибки на блок по сравнению с обычным qМПД при всего лишь 2% росте избыточности [6, 8]. При этом увеличение объема вычислений в каскадном коде составляет менее 20% по сравнению с исходным алгоритмом qМПД. А за счет применения совместно с qМПД модифицированного недвоичного расширенного кода Хэмминга можно уменьшить вероятность ошибки декодирования на 3..5 порядков [12]. Отличительной особенностью предложенных в [12] модифицированных недвоичных расширенных кодов Хэмминга от известных является то, что при кодировании и декодировании используется работа с целыми числами, а не элементами из полей Галуа. В результате данные коды можно использовать практически для любого размера символа при весьма незначительной сложности их кодирования и декодирования.

Таким образом, недвоичный аналог алгоритма МПД может обеспечить при весьма высоких уровнях шума вероятности ошибки декодирования, в ряде случаев недоступные для кодов Рида-Соломона сколько угодно большой длины. При этом сложность реализации такого алгоритма оказывается незначительной, линейно растущей с длиной кода, т.е. теоретически минимально возможной [9].

Применение qМПД в системах хранения данных

Одной из областей применения недвоичных кодов является защита данных от искажений при долговременном хранении на различных носителях информации. Для решения подобных задач в настоящее время возможно применение таких программных пакетов, как QuickPar (www.quickpar.org.uk) и ICE ECC (www.ice-graphics.com), основанных на применении кодов РС. При работе данных пакетов с большими файлами возникают сложности или с обеспечением приемлемой скорости, или надежности исправления ошибок. Применение для защиты файлов программных средств [13], использующих алгоритмы qМПД, решает перечисленные проблемы, часто предоставляя одновременно и большую корректирующую способность, и гораздо более высокое быстродействие. В частности, qМПД при программной реализации даже для длинных кодов и больших размеров алфавита обеспечивают скорость декодирования в несколько десятков Мбит/с на обычном ПК, что оказывается в десятки, сотни, а иногда и в тысячи раз быстрее других современных алгоритмов коррекции ошибок. Такие скорости, например, показывает представленная на сайте www.mtdbest.iki.rssi.ru демо-программа для qМПД, работающая даже на обычных ПК на скоростях 8…30 Мбит/с при столь больших шумах канала, при которых декодеры кодов РС вообще не работают хоть сколь-нибудь эффективно. В результате использующие qМПД программные средства в ряде случаев способны обеспечить на много порядков более высокие уровни защиты файлов от искажений, чем указанные выше программы, поддерживая при этом во много раз лучшие скорости кодирования и восстановления информации [13].

Сравнение возможностей программ для защиты файлов от искажений показало, что применение qМПД позволяет существенно повысить скорость кодирования/восстановления информации по сравнению с аналогами. Программные средства, основанные на qМПД, обеспечивали скорость кодирования/декодирования в десятки Мбайт/с, что на порядок больше скорости работы программ ICE ECC и QuickPar в тех же условиях. Особо отметим, что при этом qМПД одинаково эффективно исправляет как независимые ошибки и стирания, так и пакеты ошибок или стираний. Этого нельзя сказать о программах ICE ECC, QuickPar, которые эффективно исправляют пакеты ошибок, но не справляются даже с малым процентом независимых ошибок.

Заключение

Возможности символьных МПД алгоритмов оказываются по вероятности ошибки и по числу операций декодирования на много порядков лучше, чем возможности кодов Рида-Соломона, по праву считавшихся лучшими недвоичными кодами в течение почти полувека. Это определяется эффективным переносом идей двоичного МПД на очень просто организованные недвоичные коды сколь угодно большой длины. В результате недоступный ранее уровень помехоустойчивости, получаемый с помощью алгоритмов МПД разных типов, позволяет решать задачи обеспечения высокой надежности хранения данных без какой-либо дополнительной доработки этих алгоритмов или всего лишь при незначительной их адаптации к возможным дополнительным требованиям, возникающим при хранении данных терабайтной емкости в системах ДЗЗ. При использовании qМПД в таких системах легко обеспечить оперативный контроль за качеством хранимой информации, а также корректировку данных вследствие старения и возникающих дефектов носителя. Производительность кодеров для qМПД, которые оказываются предельно простыми однотактными узлами, может быть легко доведена до уровня 20 Гбайт/с. В результате их использование в процессе записи данных никогда не будет ограничивающим фактором для высокоскоростных систем ДЗЗ.

Работа выполнена при поддержке РФФИ (грант №08-07-00078), Института космических исследований и Рязанского государственного радиотехнического университета.

Литература

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

2. Ning C., Zhiyuan Y. Complexity analysis of Reed-Solomon decoding over GF(2^m) without using syndromes // EURASIP Journal on Wireless Communications and Networking, January 2008, n.4, pp.1-11,

3. Davey M.C., MacKay D.J.C. Low density parity check codes over GF(q) // IEEE Comm. Letters, 2(6), 1998, pp.165–167.

4. Declercq D., Fossorier M. Extended min-sum algorithm for decoding LDPC codes over GF(q) // IEEE International Symp. on Inf. Theory, 2005, pp.464–468.

5. Zhang F., Pfister H. List-Message Passing Achieves Capacity on the q-ary Symmetric Channel for Large q // In Proc. IEEE Global Telecom. Conf., Washington, Nov. 2007. pp.283–287.

6. Золотарёв В.В. Каскадные схемы МПД-декодирования для больших баз данных // Мобильные системы, 2008, №3, С.66-71.

7. Золотарёв В.В., Овечкин Г.В. Эффективное многопороговое декодирование недвоичных кодов // Радиотехника и электроника, 2010, том 55, №3, С. 324–329.

8. Золотарёв В.В. Теория и алгоритмы многопорогового декодирования. М.: Радио и связь, Горячая линия – Телеком, 2006.

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

10. Золотарёв В.В. Параллельное кодирование в каналах СПД // Вопросы кибернетики. 1986. Вып. 120.

11. Овечкин Г.В., Овечкин П.В. Оптимизация структуры недвоичных самоортогональных кодов для схем параллельного кодирования // Труды НИИР, 2009. №2. С.34–38.

12. Овечкин Г.В., Овечкин П.В. Использование недвоичного многопорогового декодера в каскадных схемах коррекции ошибок // Вестник РГРТУ, 2009. №4 (выпуск 30).

13. Овечкин П.В. Применение недвоичного многопорогового декодера для защиты файлов от искажений // В сб.: «11 Международная конференция «Цифровая обработка сигналов и ее приложения- DSPA-09», М., 2009. С.200–202.

EFFECTIVE NON-BYNARY MULTITHRESHOLD DECODING FOR REMOTE EARTH SENSYNG SYSTEMS

R.Nazirov, V.Zolotarev, G.Ovechkin, P.Ovechkin, I.Chulkov

Space Research Institute RSA,
117997 Moscow, 84/32 Profsoyuznaya str.
E-mails: zolotasd@yandex.ru, g_ovechkin@mail.ru, chulkov@iki.rssi.ru

The article deals with the non-binary multithreshold decoders (qMTD) of self-orthogonal codes in q-ary symmetrical channel. The performance of qMTD for new codes with parallel coding is presented. The codes enable to work at higher channel noise. It's discussed application of qMTD for protection of files against errors in data storage systems.

Keywords: error-correcting coding, communication systems, data storage systems, q-ary symmetrical channel, non-binary self-orthogonal codes, non-binary multithreshold decoder