(Пленарный доклад на 10-й Международной конференции DSPA'08 26.03.08 Москва)

КАСКАДНЫЕ СХЕМЫ ДЕКОДИРОВАНИЯ ДЛЯ БАЗ ДАННЫХ НА ОСНОВЕ МПД

Зубарев Ю.Б.[1], Золотарёв В.В.[2]

[1]МНИТИ, [2]ИКИ РАН

Использование многопороговых декодеров (МПД) для двоичных кодов позволяет обеспечивать достаточно эффективное исправление ошибок. В частности, этот алгоритм вполне работоспособен при отношении битовой энергетики передаваемого потока данных к спектральной плотности мощности шума $E_b/N_0 < 1$ дБ при высокой производительности и в аппаратном, и в программном вариантах реализации декодера для кодовой скорости R~1/2 [1-6].

Менее известны возможности недвоичных алгоритмов МПД[1,3-9], впервые, видимо, описанных в [10,11]. Ниже рассмотрена эффективность алгоритмов МПД при обработке символьной (недвоичной) информации, закодированной мажоритарно декодируемыми кодами нового класса. Выполнено сравнение его характеристик с возможностями единственных реально используемых сейчас недвоичных кодов Рида-Соломона (РС). Описан простой метод уменьшения на несколько порядков вероятности ошибки QМПД на блок с помощью каскадирования.

В [1,3,6,8] отмечалось, что в случае больших значений основания **q**, **q**>10, недвоичного кода совершенно невозможно создать эффективные истинно оптимальные декодеры (ОД), в том числе и алгоритм Витерби, поскольку при этом их сложность во многих случаях будет иметь вид $\mathbf{q}^{\mathbf{K}}$, где **K** - длина кодирующего регистра. Если ещё учесть, что возможности декодеров для кодов Рида-Соломона, имеющих широкую область применения, очень ограничены, в первую очередь из-за их небольшой длины в реальных системах, исследование и дальнейшая разработка очень простых алгоритмов типа МПД представляются особенно полезными.

Недвоичный МПД декодер (далее: **QМПД**) устроенный согласно [1-4], при каждом изменении декодируемых символов переходит к более правдоподобному решению по сравнению с предыдущими состояниями декодера и может даже для достаточно высоком уровне шума обычно достигать оптимальных решений, хотя он и не является ОД.

Рассмотрим предельные корректирующие возможности QМПД.

Рассмотрим для кода класса СОК некоторые события, приводящие к ошибкам ОД. К искомым векторам ошибки относятся такие, что при использовании QМПД [10,11]:

- все проверочные символы, подаваемые на ПЭ, и декодируемый символ $\mathbf{i_k}$ ошибочны;
- все проверочные символы ошибочны, но два из них одинаковы, а $\mathbf{i_k}$ принят верно;
- есть один правильно принятый проверочный символ, а остальные ошибочны, как и $\mathbf{i_k}$.

Более полное перечисление различных событий, приводящих к ошибкам недвоичного ОД, и оценки их вероятностей рассматривались в [1,3,9-11]. Вероятности ошибки в первом символе недвоичного ПД, которые для МПД с несколькими итерациями декодирования можно рассматривать как верхние оценки вероятностей ошибки в QМПД, приведены в [1,10]. Простая схема QМПД дана в [3]. Поскольку QМПД на каждом шаге стремится к решению ОД, то можно ожидать, что при некотором достаточно высоком уровне шума он во многих случаях достигнет искомого оптимального решения, что обычно требует полного перебора. А QМПД при этом сохраняет линейную сложность реализации от длины кода.

Оценим возможности простейшей каскадной схемы. При использовании блоковых кодов и декодировании, близком к оптимальному, когда вероятность ошибки в канале достаточно мала, QМПД обычно совершает редкие одиночные ошибки, поскольку все ближайшие кодовые слова в кодах СОК отличаются между собой в единственном информационном символе. В подобных случаях желательно избавиться от этих одиночных ошибок простейшими средствами, поскольку исправление такой единственной ошибки приведёт и к правильному декодированию всего очередного весьма длинного блока данных. Эту задачу легко решить с помощью кодов с контролем по mod q, полного аналога кодов контроля по чётности, эффективно применимых для двоичных кодов [1,3,6]. Этот кодер ставится перед кодером для СОК.

На втором этапе декодирования в каскадной схеме после обычного QМПД декодирования предлагается реализовать корректирующие возможности кодов контроля по mod q. В каждом подблоке длины L вычисляется сумма первых (L-1) символов и если эта сумма не равна значению L-го символа, то на величину их разности изменяется самый ненадёжный символ в данном блоке. Применение этого кода позволяет исправить не одну, а достаточно большое число ошибок в полном блоке каскадного кода, оставшихся после QМПД декодера.

Очевидная нижняя оценка вероятности ошибки рассматриваемого кода с контролем по mode q равна вероятности появления двух ошибок декодирования в пределах блока длины L. Если нижнюю оценку вероятности ошибки QМПД в каждом символе кода обозначить как $Pq_{OD}(e)$, то вероятность появления двух ошибок без учёта размножения ошибок декодирования, т. е. их группирования, можно оценить снизу как $P_2=L*(L-1)*(P_{qOD}(e))^2/2$. Эта оценка соответствует наличию двух ошибок QМПД декодера в подблоке длины L . Она вполне приемлема для предварительных оценок.

Во многих случаях в реальных системах удобно работать с данными, имеющими байтовую структуру. Отметим, что кроме кодов Рида-Соломона (РС) в настоящее время вообще нет других сколько-нибудь эффективных и одновременно довольно простых методов декодирования недвоичных символьных данных, если выбранный код достаточно короткий. Сравним вероятностные характеристики кодов РС с возможностями QМПД. Подчеркнём, что для QМПД никаких ограничений по длине кода вообще нет, поскольку длины кодов и величина его основания в недвоичных кодах с мажоритарным декодированием - независимые параметры..

Ниже рассмотрены результаты моделирования работы QМПД в недвоичном симметричном канале QСК, характеристики каскадирования таких кодов с кодами контроля по mod q , а также возможности обычных декодеров кодов РС. Объём моделирования в нижних точках графиков для QМПД составлял от $5*10^{10}$ до $2*10^{12}$ битов, что свидетельствует о крайней простоте метода.

Сравнение кодов РС и QМПД при кодовой скорости R=1/2 было выполнено в [5]. Отметим только, что QМПД декодер для кода длины 32000 оказывается способным обеспечить простейшими мажоритарными методами помехоустойчивость, принципиально недостижимую даже для кода РС длины 65536 и двухбайтовых символов.

На рис.1 показаны возможности высокоскоростных QМПД и кодов РС при кодовой скорости R=7/8. Сплошными линиями представлены вероятности ошибки на символ для кодов РС.

Пунктирными линиями представлены коды с QМПД декодированием и длиной 48000 символов: b1 -байтовых (символ - 8 битов) и b2 – бвухбайтовых (символ – 16 битов). Аналогично случаю R=1/2, возможность создания кодов РС длины 4096 при R=7/8 в ближайшее время останется очень проблематичной, в то время как даже для кодов длины 48000 байтов рассматриваемые недвоичные мажоритарные декодеры остаются очень простыми.

Наконец, на рис.2 представлены аналогичные характеристики для QМПД и кодов РС при очень малой избыточности для R=0,95. Для сопоставления там же приведён также график для кода РС с n=256 и R=7/8 с рис.1. Пунктирами b1 и b2 указаны возможности двух QМПД для кодов длины n~80000 и символов размером 1 и 2 байта. Демопрограмму быстрого QМПД можно переписать с сетевого ресурса [6], со странички «Обучение». Она работает при R=0,95 и успешно декодирует цифровые потоки с весьма большой вероятностью ошибки на символ 0,01 и выше, с очень большой скоростью 5-25 Мбит/с и более на обычных ПК, снижая вероятности ошибки после декодирования на 6 ÷ 8 и более десятичных порядков.

Из сопоставления кодов РС длины n=256 при R=7/8 и R=19/20 видно, насколько последний менее эффективен, чем первый и насколько труднее обеспечивать хорошую эффективность при уменьшении избыточности. Тем не менее характеристики малоизбыточных кодов с мажоритарным декодированием на основе QМПД оказываются весьма высокими и могут существенно поднять уровень помехоустойчивости благодаря просто выбору длинных кодов.

Более того, QМПД при R=0,95, как следует из рис.2, эффективнее кода РС с R=7/8, у которого избыточность в 2,25 раз больше. Этим снимаются все вопросы о применении любых самых сложных методов декодирования для кодов РС: они малоэффективны по сравнению с QМПД. Везде для графиков b1 и b2 нижние точки соответствуют оптимальному (совпадающему с переборным!) декодированию с одиночными ошибками. Другие экспериментальные данные для QМПД можно найти в [1,3,6-9].

Возможности каскадирования QМПД с использованием кодов контроля по mod q также представлены для высокоскоростных кодов на рис.1 и рис.2.

Вероятности ошибки на блок для каскадных кодов с внутренним СОК кодом при R=7/8 и внешнего кода с L=190 представлены кривой B1 для кода с q=256 и кривой B2 для кода с q=65536 (символ - 2 байта) на рис.1.

Для каскадного кода с внутренним кодом при R=0,95 и внешнего кода с L=190 на рис.2 вероятность ошибки на блок также даётся кривыми B1 для q=256 и B2 для q=65536. Из вида кривых для каскадного кода следует, что характеристики МПД обеспечивают исправление всех одиночных и часть многократных ошибок благодаря тому, что число m подблоков длины L во внутреннем коде оказывается достаточно большим.

Во всех случаях на графиках B1 и B2 нижние точки соответствуют величине N_B^{-1}, где N_B – число декодированных блоков, поскольку в экспериментах не было ни одного случая неправильного декодирования каскадного кода в этих точках.

Как и ожидалось, применение каскадирования на много порядков снижает вероятность ошибки на блок по сравнению с обычным QМПД почти без увеличения избыточности. Рост объёма вычислений в каскадном коде составляет менее 20% по с равнению с исходным алгоритмом QМПД.

Кроме естественных областей применения простых высокоэффективных методов кодирования в сетях связи следует отметить хорошие возможности применения QМПД для кодирования информации на дисках и других носителях больших объёмов информации, в сверхбольших базах аудио- и видео- данных с намного более высоким уровнем достоверности, чем это было доступно до недавнего времени, а также при обновлении, восстановлении и использовании хранимых данных. При этом легко обеспечить и оперативный постоянный контроль за качеством хранимой информации, а также корректировку данных вследствие старения и возникающих дефектов носителя.

Принципиально новый уровень помехоустойчивости, на 4÷7 десятичных порядков более высокий, просто достигаемый с помощью QМПД, позволяет решать вышеперечисленные задачи без доработки этих алгоритмов или всего лишь при незначительной их адаптации к возможным дополнительным требованиям, возникающим в подобных масштабных цифровых системах.

Таким образом, возможность очень простого исправления ошибок в длинных недвоичных кодах при эффективности, близкой к уровню, доступному только оптимальным переборным алгоритмам, открывает принципиально новые возможности для кодирования символьной информации, т. е. основных видов данных, практически непосредственно используемых современным информационным обществом. Кодирование обеспечивает высокое контролируемое качество хранимой, передаваемой и формируемой информации. Применение простых и одновременно высокоэффективных методов кодирования может создать новые высокие стандарты информационного обеспечения всех аспектов развития цивилизации.

Дополнительная информация об МПД разных классов – на специализированном тематическом двуязычном веб-сайте ИКИ РАН **www.mtdbest.iki.rssi.ru** .

Исследования велись при финансовой поддержке РФФИ по гранту №05-07-90024в.

Литература

1. Золотарёв В.В. Теория и алгоритмы многопорогового декодирования. –Под научной редакцией члена-корреспондента РАН Ю.Б.Зубарева, Москва, Радио и связь, Горячая линия - Телеком, 2006, 270с.

2. V.V.Zolotarev, S.V.Averin, I.V.Chulkov. Optimum Decoding Characteristics Achievment on the Basis of Multithreshold Algorithms. –- 9-th ISCTA'07, July, UK, Ambleside, 2007.

3. Золотарёв В.В., Овечкин Г.В. Помехоустойчивое кодирование. Методы и алгоритмы. Справочник. "Горячая линия - Телеком", Москва, 2004, с.124.

4. V.V.Zolotarev, R.R.Nazirov, I.V.Chulkov - The Quick Almost optimal multithreshold decoders for Noisy Gaussian Channels- RCSGSO International Conference ESA in Moscow, SRI RAS, Russia, June, 2007.

5. Ю.Б.Зубарев, В.В.Золотарёв - Достижение характеристик оптимального декодирования на основе многопороговых алгоритмов. – В сб.: 9-я Международная конференция и выставка "Цифровая обработка сигналов и её применение", Доклады-1, ***Пленарный доклад***, Москва-2007, с.12-15.

6. Золотарёв В.В. Многопороговые декодеры.- Веб-сайт ИКИ РАН www.mtdbest.iki.rssi.ru .

7. V.V.Zolotarev, S.V.Averin – Non-Binary Multithreshold Decoders with Almost Optimal Performance - 9-th ISCTA’07, July, UK, Ambleside, 2007.

8. Золотарёв В.В. Многопороговое декодирование для информационных потоков с байтовой структурой. - Мобильные системы, М., 2006, №3, с.25-27.

9. Золотарёв В.В. Обобщение алгоритма МПД на недвоичные коды. - Мобильные системы, М., 2007, №2, с.36-39.

10. Золотарёв В.В. Алгоритмы кодирования символьных данных в вычислительных сетях. - В сб.: "Вопросы кибернетики", Научный совет по комплексной проблеме «Кибернетика» АН СССР, ВК-106, М.,1985, с.54-62.

11. Золотарёв В.В. Многопороговое декодирование в недвоичных каналах. - В сб.: "Вопросы радиоэлектроники", Серия ЭВТ, вып.12, М.,1984.

Характеристики кодов РС и QМПД с R=7/8

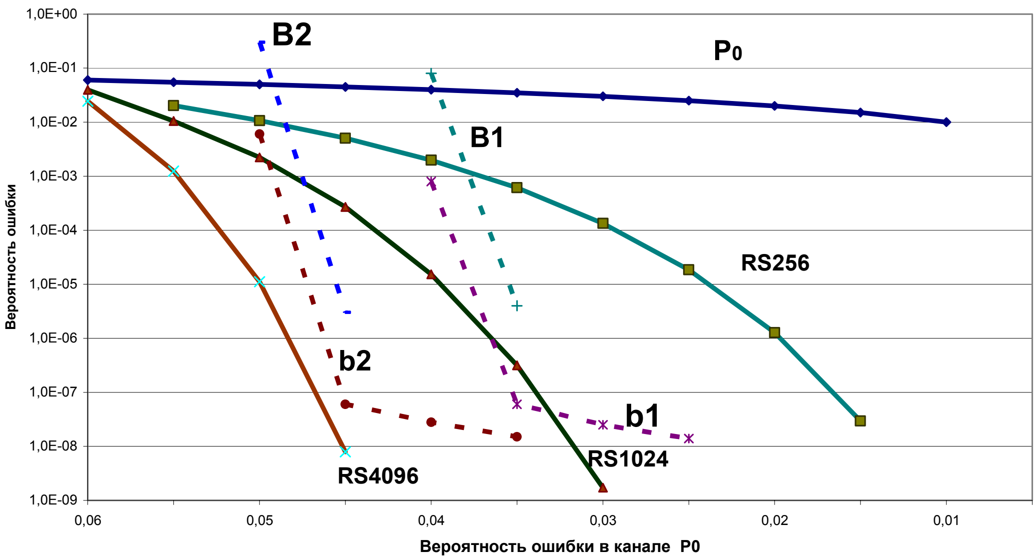

Рис.1.

Характеристики QМПД и кодов РС при R=19/20

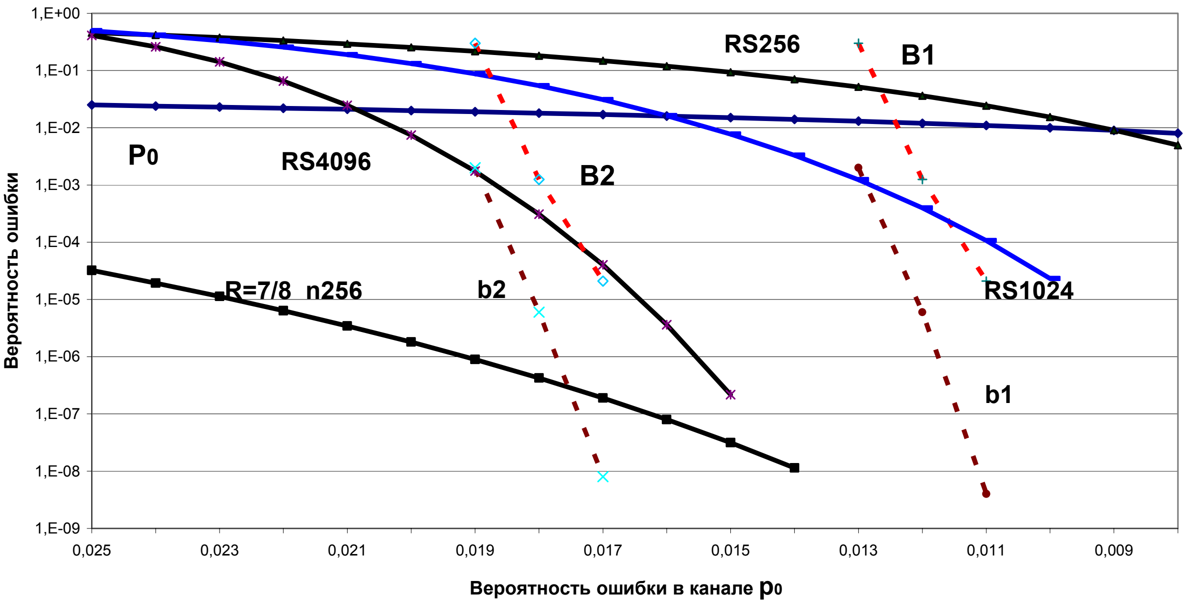

Рис.2.