

Инструкция

по применению макетной программы декодирования недвоичных кодов на основе многопорогового алгоритма коррекции ошибок в каналах с большим уровнем шума.

Данная инструкция подготовлена для обеспечения быстрого запуска и правильной интерпретации результатов работы программной версии алгоритма многопорогового декодирования (МПД) символьных данных.

Специальные недвоичные (символьные) коды с мажоритарным декодированием и собственно высокоэффективный алгоритм их декодирования, обозначаемый далее как QМПД, впервые были открыты в [1-2] и затем описаны в [3-7], а также в других работах автора.

До недавнего времени фактически единственными недвоичными кодами, которые можно было использовать в различных цифровых системах, являлись коды Рида-Соломона (РС). Однако их возможности всегда оказываются очень ограниченными вследствие небольшой длины этих кодов и довольно значительной сложности их декодирования. Именно поэтому реально до сих пор применялись коды РС длины не более 256, что соответствует символьному алфавиту такого же или даже меньшего размера. Таким образом, использование кодов РС с размером символов более 1 байта (8 бит) более чем проблематично.

Предлагаемая демопрограмма недвоичного МПД позволяет работать с кодами произвольно большой длины при использовании абсолютно любого символьного алфавита, размер которого совершенно не зависит от длины используемого кода. Коды символьного типа с мажоритарным декодированием одновременно для упрощения всех процедур кодирования/декодирования и для повышения их эффективности строятся из известных двоичных кодов так, что в символьном коде и порождающая, и проверочная матрицы состоят только из нулей и единиц. Это позволяет все процедуры кодирования сводить только к операциям сложения, вычитания и сравнения целых чисел, образующих любую простую группу по сложению. Это позволяет дополнительно значительно снизить объём вычислений в QМПД.

В целом данная демопрограмма QМПД и инструкция сохраняют идеологию демопрограммы и инструкции для двоичного свёрточного кода с кодовой скоростью $R=3/4$, уже более двух лет представленного на нашем сайте на образовательной странице. Целью предлагаемой инструкции является демонстрация высокой эффективности работы QМПД в условиях большого шума с одновременно достигаемой очень большой производительностью (скоростью декодирования). Для этого программный модуль **QMTD.exe** считывает входные данные из файла **qmtdd.c** и затем работает, осуществляя декодирование последовательности закодированных данных недвоичного систематического блочного кода. Поскольку реально недвоичные коды особенно полезны в каскадных схемах кодирования и в системах повышения достоверности хранения в сверхбольших базах данных, где обычно всегда имеет место небольшая избыточность кодирования, программа хорошо иллюстрирует возможности именно малоизбыточного кода с высокой кодовой скоростью $R=0,95$.

Специально для подчёркивания очень высокой реальной помехоустойчивости алгоритма QМПД для кода с кодовой скоростью $R \sim 1$ при большом уровне шума (когда вероятность ошибки в недвоичном канале близка к $p_0 \sim 10^{-2}$) моделируется работа алгоритма с весьма длинным кодом, длины 80'000 символов. Таким образом, длина информационной части кода равна 76'000 символов, а проверочной 4000 символов. Размер символов может меняться от 4 битов (алфавит 16 символов) до 26 битов (алфавит 2^{26} символов). Как известно, МПД алгоритмы относятся к итеративным процедурам, которые могут любое заданное число раз повторять попытки коррекции всех символов принятого

кода в некотором порядке. В программе число итераций также может варьироваться в некоторых пределах, достаточных для выявления возможностей декодера QМПД по производительности и по корректирующей способности. Кроме того, во входном массиве данных задаются объёмы моделирования (выраженные числом декодируемых информационных символов), а также вероятности ошибки в канале, полагая, что значения ошибок всегда выбираются случайно и равномерно среди всех возможных их значений, как это обычно и имеет место в q-ичных симметричных каналах.

Рассмотрим конкретные входные параметры моделируемой системы кодирования. Входной файл данных **gmtdd.c** имеет следующий вид:

```
kst=339   erl=10000
          ndkdd=1000000   iterat=5
          apstart= 0.015   apincr= 0.001   apend= 0.0011
          qs=65536
          =====
          qs=67108864 2^26
          qs=65536 2^16
          qs=256 2^8
          qs=16 2^4
```

Первый параметр **kst** должен быть целым положительным числом в диапазоне 1 – 999 . Это число определяет начальное состояние датчика случайных чисел, на основе которого генерируется поток ошибок канала;

erl – предельное количество ошибок декодера, после достижения которого прекращается процесс декодирования и происходит переход к каналу с меньшей вероятностью ошибки;

ndkdd - объём моделирования процесса исправления ошибок, выраженный числом декодированных символов. Напомним, что в языке C++ есть ограничения на представления больших целых чисел. Поэтому нежелательно назначать ndkdd большим, чем миллиард;

iterat - количество итераций (попыток повторной коррекции) декодирования;

apstart - начальная вероятность ошибки в канале;

apincr - шаг уменьшения вероятности ошибки в канале;

apend - конечная минимальная вероятность ошибки канала в заданном интервале вероятностей ошибки при моделировании работы QМПД;

qs - размер алфавита используемого кода. Минимальное значение **qs=16**. Для удобства формирования кода этот параметр следует выбирать как степень двойки. Это позволяет реализовать в данной моделирующей программе группу по сложению через простейшую маску. Под чертой комментария приведены примеры выбора возможных значений таких **qs**. Если $qs > 2^{26}$, то $qs = 2^{26}$. Разумеется, в реальных кодовых системах с применением QМПД размеры символов могут определяться архитектурой используемых процессоров, оперирующих, возможно, многими байтами одновременно. Это приводит к росту достоверности кодирования при большом уровне шума и одновременно увеличивает производительность декодеров обсуждаемого класса за счёт просто роста параллелизма в работе декодера, совершенно ни в чём не усложняя (!!!) принципов и конкретных методов работы QМПД . Все справочные данные под чертой комментария во входном файле не влияют на процесс декодирования, поскольку последний считываемый программой символ – **qs**, находится выше этой строки комментария.

Рассмотрим далее вид выходной распечатки работы демопрограммы мажоритарного многопорогового декодирования недвоичных кодов с именем **graft.c**, представленной ниже.

```

block QMTD block length N=80000 symbols R=19/20, symbol =>from 4 to 26 bits
Current time: Tue Apr 17 19:32:50
kst=339
ndkdd=1000000 nnbl=14 new ndkdd=1064000
iterat=5
pstart=0.015000 pincr=0.001000 pend=0.001100
qs=65536
=====
time work= 140 ms 0 s
*****
=>=>=> channel noise p0=0.015000
block= 1 194 errors
block= 2 15 errors
block= 3 324 errors
block= 4 5 errors
block= 5 18 errors
block= 7 312 errors
block= 8 29 errors
block= 9 188 errors
block= 10 298 errors
block= 11 30 errors
block= 12 44 errors
block= 13 3 errors
=>=>=> channel noise p0=0.015000

Ps channel Ps decoder Ns decoded err Sdec
1.500e-002 1.372e-003 1064000 1460
time decoding total work 2187 msec 2 sec delta=2 sec
decoder speed= 486000 symbols/s

*****
=>=>=> channel noise p0=0.014000
block= 3 1 errors
block= 7 12 errors
=>=>=> channel noise p0=0.014000

Ps channel Ps decoder Ns decoded err Sdec
1.400e-002 1.222e-005 1064000 13
time decoding total work 4140 msec 4 sec delta=1 sec
decoder speed= 544000 symbols/s

*****
=>=>=> channel noise p0=0.013000
=>=>=> channel noise p0=0.013000

Ps channel Ps decoder Ns decoded err Sdec
1.300e-002 0.000e+000 1064000 0
time decoding total work 6000 msec 6 sec delta=1 sec
decoder speed= 572000 symbols/s
*****
Current time: Tue Apr 17 19:32:56
===== end =====

```

Рассмотрим содержание выходной распечатки.

Сначала печатается шапка-заголовок выходных данных со справочными сведениями о программе: длина кодового блока 80000, кодовая скорость $R = 19/20$ и допустимые размеры символов кода, выраженные числом битов, которые необходимы для представления символов кода: от 4 до 26 битов для данной демопрограммы, а также дата и текущее время.

Затем полностью повторяются те входные данные программы, которые присутствовали во входном файле и были активно использованы демопрограммой в процессе работы. Здесь нужно обратить внимание на то, что после печати объёма декодируемых символов, которые были во входном массиве, идёт печать двух новых параметров **nnbl** и **new ndkdd**. Это определяется тем, что программа перед началом работы определяет количество кодовых блоков, которое необходимо декодировать, чтобы реализовать заданный объём эксперимента. Для этого заданный объём декодированных данных она увеличивает на столько символов, чтобы этот объём моделирования был кратен 76000 символам, т. е. длине информационной части кодового блока. Именно количество декодированных блоков и новое число информационных символов в них печатается дополнительно к введённому в программу объёму работы.

Далее в программе идёт параллельная с процессом моделирования печать данных о работе демопрограммы.

Сначала печатается время работы в мсек на подготовку предварительных условий моделирования: формирование массивов данных и предварительные вычисления, что занимает время ~ 140 мсек - по распечатке. Далее идёт печать стартовой вероятности ошибки в симметричном недвоичном канале и собственно данные о работе декодера с каждым принятым из канала блоком. Если блок декодируется ошибочно хотя бы в одном символе, то печатается его номер и количество оставшихся в нём ошибок. Разумеется, если общее число ошибок во всех неправильно декодированных блоках превысит величину err из входного массива, эксперимент будет завершён раньше, чем это определяется параметром **nnbl**.

После завершения эксперимента при заданном стартовом значении вероятности ошибки в канале идёт печать сведений о результатах моделирования работы QМПД. Повторяется печать вероятности p_0 канала и ниже её – строка основных статистических данных: т. е. - снова p_0 , обозначаемая **Ps channel**, вероятность ошибки QМПД на символ **Ps decoder**, объём эксперимента как число декодированных информационных символов **Ns decoded** и общее количество ошибок QМПД в декодированных символах принятого цифрового потока **err Sdec**. Далее ниже указывается полное время работы программы, а также время **delta** – время декодирования данных при текущей вероятности ошибки в канале.

Наконец, ниже указывается **важнейший параметр** - скорость работы программной демоверсии QМПД, выраженная в декодированных за секунду символах. В приведённом модельном описании в последней строке напечатано при: **channel noise p0=0.013000**

decoder speed= 572000 symbols/s .

Это значит, что при двухбайтовых символах кода, поскольку во входных данных программы указан размер символа **qs=65536**, скорость декодирования в битах оказывается равной $V = 572000 * 16 = 9,152$ Мбит/сек. Подчеркнём, что эта распечатка получена на относительно медленном портативном компьютере и длина обрабатываемых символов выбрана также весьма небольшой – 2 байта. При выборе более длинных символов кода, например, 4 - 16 байтов и в случае применения существенно более производительных уже широко доступных персональных компьютеров, как уже успешно показали наши эксперименты с различными типами ПК, производительность QМПД может достичь $20 \div 60$ Мбит/сек даже в таком программном варианте реализации, который представлен здесь. Разумеется, этот важнейший параметр, иллюстрирующий очень небольшую сложность

реализации данного алгоритма декодирования, в случае создания его даже весьма простой аппаратной реализации с некоторым набором ускоряющих его работу возможных приёмов, позволит поднять скорость декодирования до величин порядка нескольких сотен мегабитов в секунду, что несомненно удовлетворяет требованиям, предъявляемым практически ко всем средствам контроля ошибок в системах передачи данных и в системах хранения информации на дисках и в сверхбольших цифровых базах данных. Более детальные характеристики QМПД приведены в [4], где указаны также пути дальнейшего роста эффективности его работы.

Нельзя не отметить также, что реальная производительность программной версии QМПД, представленной здесь, на самом деле много выше измеряемой в процессе работы декодера, потому что данная демопрограмма в процессе моделирования алгоритма ещё успевает выполнять полный цикл работ по формированию входных массивов данных, их кодированию и даже моделирование высококачественного двоичного канала с независимыми ошибками.

После завершения распечатки данных о декодировании при текущем уровне шума, программа уменьшает вероятность ошибки в канале на величину **pincr**, после чего происходит моделирование работы QМПД при меньшем уровне шума и т.д.

Если при некоторой вероятности ошибки эксперимент завершился без ошибок декодирования, то программа прекращает работу независимо от заданной минимальной вероятности ошибки в канале.

Работа программы может быть приостановлена в любой момент нажатием на произвольную клавишу клавиатуры компьютера. Повторное нажатие клавиши позволяет продолжить эксперимент по набору статистики. Если второе нажатие будет на клавишу 'e' (латинское), программа завершит работу с распечаткой статистических данных при текущем значении уровня шума канала.

* * * * *

Приведённых данных достаточно для использования всех возможностей выкладываемой на специализированный веб-сайт ИКИ РАН www.mtdbest.iki.rssi.ru демопрограммы. Более того, её возможности позволяют получить целый ряд новых характеристик алгоритма QМПД, даже ещё не опубликованных в научной печати. Автор предоставляет безусловное право всем желающим всесторонне проанализировать возможности QМПД для малоизбыточных кодов и опубликовать свои результаты на эту интересную тему самостоятельно. Поскольку программа действительно реализует очень высокие скорости декодирования, можно указать, что набор статистики объёмом порядка 10^9 символов (до $3 \cdot 10^{10}$ битов!) даже для компьютера весьма средней производительности требует всего лишь около часа его работы (ориентировочно). В случае публикации такой статьи автор демопрограммы готов также выложить на наш веб-сайт по методам МПД эту статью (разумеется, с согласия автора будущей статьи) и обсудить дальнейшее направление совместных исследований в этой и смежных областях теории и техники кодирования.

Приглашаем всех к сотрудничеству и быстрому совместному продвижению новых методов кодирования в реальные технические системы!

ЛИТЕРАТУРА

1. Золотарёв В.В. Многопороговое декодирование в двоичных каналах. - Вопросы радиоэлектроники, серия ЭВТ, вып.12, 1984, с.73-76.

2. Золотарёв В.В. Алгоритмы коррекции символьных данных в вычислительных сетях. - В сб.: "Вопросы кибернетики", ВК-105, АН СССР, Научный совет по комплексной проблеме «Кибернетика», М., 1985, с.54-62.
3. Золотарёв В.В. Многопороговое декодирование для информационных потоков с байтовой структурой. - Мобильные системы, М., 2006, №3, с.25-27.
4. Золотарёв В.В. Обобщение алгоритма МПД на не двоичные коды. – Мобильные системы, М., 2007, №2, с.15-19.
5. Золотарёв В.В. Теория и алгоритмы многопорогового декодирования. - М., «Радио и связь», «Горячая линия – Телеком», 2006, 270 с.
6. Золотарёв В.В., Овечкин Г.В. Помехоустойчивое кодирование. Методы и алгоритмы. Справочник. М., «Горячая линия – Телеком», 2004, 126 с.
7. Специализированный веб-сайт ИКИ РАН www.mtdbest.iki.rssi.ru .