ДК 681.391

Р.С. Новиков, А.А. Астраханцев

Харьковский национальный университет радиоэлектроники, Харьков

АНАЛИЗ ХАРАКТЕРИСТИК ПОМЕХОУСТОЙЧИВЫХ КОДОВ

Представлен анализ характеристик, общеизвестных и используемых, алгоритмов помехоустойчивого кодирования, которые отличаются друг от друга структурой, функциональным назначением, энергетической эффективностью, алгоритмами кодирования и декодирования и многими другими параметрами.

Ключевые слова: *помехоустойчивое кодирование, вероятность декодирования, проверочный символ.*

Введение

Сегодня можно говорить о создании нового класса помехоустойчивых кодов для каналов со стиранием. Кодами из этого класса можно закодировать сообщение конечного размера (файл) потенциально-неограниченным потоком независимых пакетов. Это свойство нового класса кодов принципиально отличает его от классических блоковых или сверточных, помехоустойчивых кодов с заданной скоростью. При кодировании файла этими кодами получаем также файл кодированных данных, а не поток. Новый класс кодов также называют классом фонтанных кодов (Digital Fountain Codes).

1. Основные характеристики исследуемых кодов

1.1. Код Рида-Соломона

Код Рида-Соломона представляет собой блоковый код, в котором символы состоят из k бит. Если эти символы рассматривать как пакеты сообщения, то код может быть использован для доставки сообщений в канале со стираниями. Основным свойством кода является следующее: для доставки K информационных символов достаточно принять любые K символов из N. Или иначе: для правильного приёма сообщения из K символов в блоке из N пакетов любые из M=N–K символов могут быть стёртыми.

Оптимальность кода в указанном выше смысле достигается его жёсткой алгебраической структурой. В результате существует проблема добавления «на лету» небольшого числа проверочных символов. При переходе от M=N–K проверочных символов к большему числу M'=N'–K все проверочные символы требуется пересчитывать. Алгебраическая структура кода препятствует и неограниченному увеличению числа проверочных символов в коде. Код существует лишь при N<q=2k. Жёсткая структура кода приводит и к значительным вычислительным затратам при кодировании и декодировании. В каждый проверочный символ кода входят все K исходных символов (пакетов) сообщения. Поэтому при кодировании требуется K(N–K) операций над символами (сложений и умножений) [1].

Вероятность появления ошибки в декодированном символе, Pe, можно записать через вероятность появления ошибки в канальном символе, p, определяется по формуле:

$$Pe = \frac{1}{n} \times \sum_{j=t+1}^{n} j \times \left(\frac{n!}{j!(n-j)!}\right) \times p^{j} \times (1-p)^{n-j}, \qquad (1)$$

где m – положительное целое число, больше единицы; p – вероятность появления ошибки в канальном символе; n-число кодовых символов в кодируемом символе; t – количество ошибочных битов в символе, которые может исправить код; n – число контрольных символов [2].

1.2. LT-код (рис. 1)

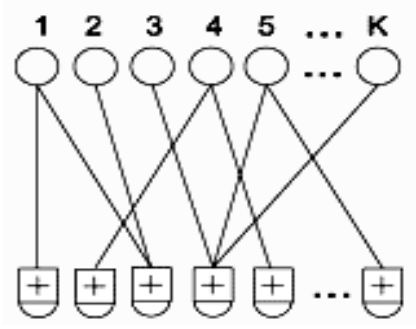

Рис. 1. Порождающий граф LT кода

Код был создан М. Лаби (Michael Luby) в 1998 г. Своё название он получил от «Laby transform» (преобразование Лаби).

В основе нахождения ρ(d) и алгоритма декодирования лежит несложная вероятностная задача. Есть сообщение из K исходных символов. Из этого множества с вероятностью 1/K производится K' случайных выборок символов. В результате формируется множество из K' кодовых символов. K'~=K ln(K/δ) при достаточно большом K, это делается для того, чтобы с вероятностью 1–δ каждый из всех K исходных символов оказался хоть один раз среди K' кодовых символов. Имея такое число кодовых символов, можно реконструировать исходное сообщение с вероятностью 1–δ. Алгоритм реконструкции

предельно быстрый и использует лишь информацию о нумерации символов.

Стоимость декодирования для кода оказывается порядка ln(K/δ) операций XOR. Эта величина при достаточно больших K и приемлемых δ оказывается намного меньше K, и поэтому код можно отнести к классу кодов с низкой плотностью порождающей матрицы [3].

1.3. Код Raptor (рис. 2)

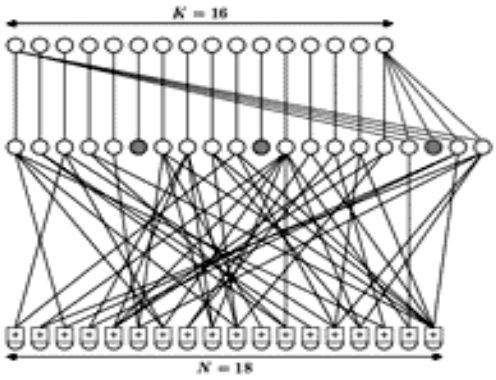

Рис. 2. Порождающий граф кода Raptor

Код представляет собой кодовую конструкцию с внутренним (по отношению к каналу) LT кодом. В качестве внешнего кода можно использоваться «почти» любой блоковый код (с фиксированной скоростью). Разработка и исследование кода принадлежит A. Shokrollahi.

Анализ конструкции показывает, что исходное сообщение может быть реконструировано с вероятностью 1–δ по K'=K(1+ε) символам, где ε – небольшое положительное число. Стоимость декодирования оказывается порядка ln(1/ε) операций XOR. Для декодирования сообщения требуется порядка Kln(1/ε) операций XOR. На сегодняшний день код является, возможно, лучшей аппроксимацией идеального фонтанного кода [4].

1.4. Коды с низкой плотностью проверок на чётность (LDPC, рис. 3)

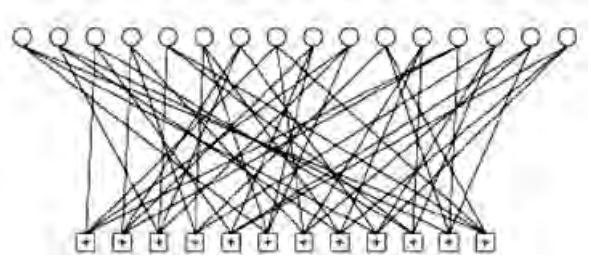

Рис. 3. Проверочный граф LDPC кода

Существует мнение, что на сегодняшний день коды с низкой плотностью проверок на чётность (Low Density Parity Code, LDPC) являются лучшими блоковыми кодами [5].

Матрица имеет размерность M×N в каждом столбце матрицы – j=количеству единиц, в каждой строке – k=количеству единиц. Для практики важны коды с очень низкой плотностью проверок на чётность, для которых j<<M и k<<N. Очень низкая плотность значительно снижает вычислительные затраты на реализацию алгоритма декодирования при больших размерах матриц. Этим кодам более 40 лет, их изобретателем является Р. Галлагер. Галлагером была предложена итеративная обменная вероятностная процедура декодирования. Применительно к обработке жёстких решений сущность этой процедуры сводится к следующему. На каждой итерации проверяются соотношения на чётность в соответствии с проверочной матрицей кода H. После первой проверки исправляется символ (ноль заменятся на единицу и наоборот), входящий в наибольшее число невыполненных проверочных соотношений на чётность.

Интерпретация обменного алгоритма декодирования особенно наглядна на проверочном графе кода. Этот граф также называют графом Таннера. Он имеет две группы проверочных узлов (Bipartite Graph). Первый набор представляет правдоподобия принятых N кодовых символов. Второй представляет правдоподобия M проверок на чётность. Отметим, что для рассмотренных кодов порождающая матрица G в корне отличается от проверочной матрицы и имеет высокую плотность единиц (High Density) [6].

Вероятность появления ошибки в декодированном символе определяется по формуле:

$$P = \int_{\sqrt{2Ec/No}}^{\infty} \frac{1}{\sqrt{2\pi}} \times e^{-x^2/2} dx, \qquad (2)$$

где Ec/No – отношение сигнал/шум в канале передачи [7].

2. Анализ вероятности декодирования помехоустойчивых кодов

Используя вышеуказанные формулы для нахождения вероятности появления ошибки в декодированном символе для кодов LDPC и Рида-Соломона, были построены графики зависимости от отношения сигнал/шум (ОСШ) и вероятности соотношения ошибки соответственно.

Исходя из рис. 4, можно сделать вывод, что LDPC код соответствует заявленным требованиям и при более “высоких” соотношениях сигнал/шум обеспечивает “низкое” значение появления ошибки в декодированном символе. С помощью одного из таких кодов удалось получить значение $Pb=10^{-5}$ при Eb /No =0.0045 дБ. LDPC-коды в современных системах передачи информации занимают нишу, аналогичную турбокодам.

Оба эти класса кодов используются в системах, где требуются повышенные скорости передачи данных при ограниченной полосе пропускания канала. К числу таких систем можно отнести, например, спутниковую связь, цифровое телевидение (в том

числе высокой четкости), а также каналы передачи в электронно-вычислительных машинах и их сетях. LDPC-кодеры могут обеспечивать поистине колоссальную скорость передачи данных (до 40 Гб/с), что обусловлено простотой их реализации

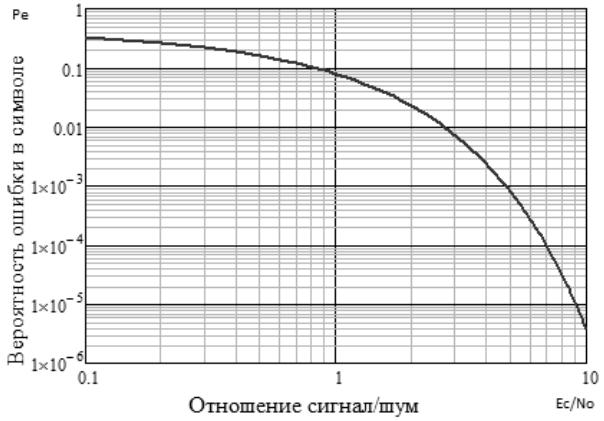

Рис. 4. Зависимость вероятности ошибки в символе от ОСШ для LDPC кодов

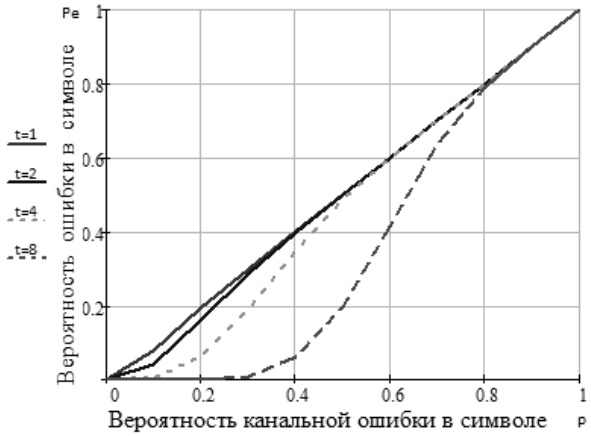

Рис. 5. Зависимость вероятности ошибки в символе Pe от вероятности канальной ошибки в символе p, для кода Рида-Соломона с характеристиками m=4, n=15

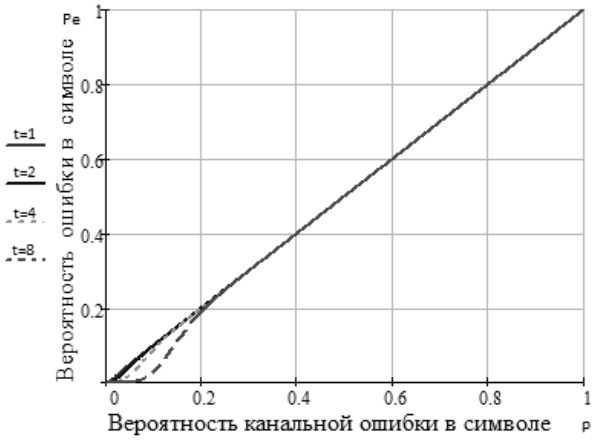

Рис. 6. Зависимость вероятности ошибки в символе Pe от вероятности канальной ошибки в символе p, для кода Рида-Соломона с характеристиками m=6, n=63

Анализируя рис.5 и 6, можно сделать выводы:

1) При увеличении числа контрольных символов вероятность “правильного” декодирования символа уменьшается.

2) При увеличении количества ошибочных битов в символе, которые может исправить код, вероятность правильного принятия символа увеличивается (при низких значениях канальной ошибки).

При передаче больших объемов данных использование кодов Рида-Соломона приобретает свою привлекательность, так как при увеличении объемов увеличивается и вероятность декодирования символа, и переданная информация с большей вероятностью дойдет до своего адресата.

Выводы

Развитие каналов связи, влекущее за собой уменьшение количества ошибок, а также все увеличивающиеся объемы передаваемой информации открывают широкие перспективы для дальнейшего внедрения и использования LDPC-кодов и кодов на их основе

Недвоичные LDPC-коды имеют существенное преимущество над кодами Рида-Соломона. Это преимущество увеличивается по мере увеличения длины кода и разрядности символа.

Все коды имеют простые алгоритмы декодирования и позволяют на практике получать результаты, близкие к предельным возможностям помехоустойчивого кодирования.

Немаловажен и рост эффективности использования кодирования с увеличением размера сообщения. Некоторое ограничение эффективности кодирования естественно наблюдается при небольших объёмах данных. Это ограничение является платой за использование статистических методов. Для многих практических приложений, однако, оно не столь существенно. В то же время, благодаря статистическому кодированию возможно решение нетривиальных сетевых задач, как, например, одновременная загрузка файла большого размера с нескольких сайтов. Можно отметить также, что алгоритмы кодирования и декодирования принципиально не зависят от размера пакета.

Можно обратить внимание и на универсальность потоковых кодов. Они могут быть использованы в любом канале со стираниями, независимо от статистики стираний. Более того, потоковые коды принципиально можно использовать и для повышения достоверности доставки сообщений в каналах со сложной помеховой обстановкой, например, в радиоканалах. Для этого физический канал с помехами следует представить моделью в виде внешнего канала со стираниями и внутреннего канала с ошибками. При этом для внутреннего канала можно использовать какой-либо хороший код для борьбы с помехами, для внешнего канала – потоковый код

для борьбы со стираниями. Такие каскадные кодовые конструкции сегодня реально рассматриваются.

Список литературы

1. Reed I.S. Polynomial codes over certain finite fields / I.S. Reed, G. Solomon // J.Soc. Industrial Appl. Math. – 1960. – Vol. 1.– P. 300.

2. Скляр Б. Цифровая связь. Теоретические основы и практическое применение / Б. Скляр. – СПб., 1999. –1105 с.

3. Luby M. LT Codes, In Proc. Of the 43rd Annual / M. Luby // IEEE Symposium on Foundations of Computer Science (FOCS). –- 2002. – P. 271-282.

4. Shokrollahi A. Raptor Codes / A. Shokrollahi. – Doctoral dissertation. –2003. – 95 p.

5. David J.C. MacKay, Information Theory, Inference, and Learning Algorithms / David J.C. MacKay. – Cambridge University Press , 2003. – 640 p.

6. Варгаузин В. Вблизи границы Шенона / В. Варгаузин // Телемультимедиа. – 2005. – №3. – С. 3-10.

7. Gallager R.G. Low Destiny Parity-Check Codes / R.G. Gallager. – Massachusetts: MIT Press, Doctoral dissertation. – 1963. – 90 p.

Поступила в редколлегию 8.10.2013

Рецензент: д-р техн. наук, проф. В.В. Поповский, Харьковский национальный университет радиоэлектроники, Харьков.

АНАЛІЗ ХАРАКТЕРИСТИК ЗАВАДОСТІЙКИХ КОДІВ

Р.С. Новиков, А.А. Астраханцев

На сьогоднішній день відомо багато різних класів завадостійких кодів, що відрізняються один від одного структурою, функціональним призначенням, енергетичною ефективністю, алгоритмами кодування і декодування і багатьма іншими параметрами. Сьогодні можна говорити про створення нового класу завадостійких кодів для каналів зі стиранням. Кодами з цього класу можна закодувати повідомлення кінцевого розміру (файл) потенційно-необмеженим потоком незалежних пакетів. Ця властивість нового класу кодів принципово відрізняє його від класичних блокових або згорткових, завадостійких кодів із заданою швидкістю. При кодуванні файлу цими кодами отримуємо також файл кодованих даних, а не потік. Новий клас кодів також називають класом фонтанних кодів (Digital Fountain Codes).

Ключові слова: *завадостійке кодування, ймовірність декодування, перевірочний символ.*

ANALYSIS OF CHARACTERISTICS ERROR-CORRECTING CODES

R.S. Novikov, A.A. Astrakhantsev

Today there are many different classes noiseimmunity codes that differ from each other structure, functionality, energy efficiency, encoding and decoding algorithms and many other parameters. Today we can speak of a new class of noiseimmunity codes for channels with erasure. Codes of this class can be coded message finite size (file) potentially unlimited stream of independent packages. This property is a new class of codes distinguishes it from classic block or convolutional, noiseimmunity codes with a given speed. When encoding file these codes are also encrypted file data rather than flow. A new class of codes is also called class Fountain Codes (Digital Fountain Codes).

Keywords*: error control codes, the probability of decoding, check digit.*