# Помехоустойчивые кодеки на основе каскадных схем кодирования/декодирования для систем передачи информации с одиночными и модульными ошибками.

# Раздел №1. аналитический обзор литературы и постановка задачи;

На текущий момент существуют тенденции построения и эксплуатации распределённых информационных вычислительных систем для решения разного класса задач. Для корректного функционирования распределенных приложений необходимы устойчивые и быстрые каналы связи. Современные каналы передачи данных достаточно быстры и защищены от помех, а дополнительное применение современных помехоустойчивых кодов позволяет свести на нет вероятность неправильной передачи информации. Помехоустойчивое кодирование решает две основные задачи:

* повышение эффективности передачи данных, за счёт устранения избыточности данных в каналах, не содержащих помехи;
* повышение вероятности правильного декодирования пакета в канале с помехами с минимальной дополнительной избыточностью.

В контексте работы рассмотрим вторую задачу применения помехоустойчивых кодов. Целью научной работы является построение имитационной модели передачи информации по каналу с шумами с применением каскадного способа построения кодеков и выявление общих закономерностей и характеристик каскадных кодеков.

В ходе работы необходимо разработать программное средство помехоустойчивого кодека на основе последовательной каскадной схемы. Для разработки будет использована язык python версии 3.7.+. Программное средство должно поддерживать и позволять применить в каскаде четыре вида кодеров (Хемминга, Циклический, Свёрточный, Фонтанный), а также задать их параметры. Разработанная система должна поддерживает использование перемежителей после каждого составного кодера. В функции программного каскадного кодека входит установка разных уровней и видов шумов в канале при передаче кодового сообщения по каналу связи. Существует возможность определить контент передаваемых данных и количество выполняемых тестов для каждого уровня шума.

## Теория помехоустойчивого кодирования.

Для дальнейшего упрощения понимания контекста данной работы далее будут введены основные понятия и объекты касаемо темы помехоустойчивого кодирования.

**Канал связи –** система технических средств, предназначенная для передачи информации от источника к приёмнику.

**Теория кодирования** — одна из областей математики область действия которой распространяется на передачу данных по зашумленным каналам, а предметом является обеспечение корректности переданной информации.

**Теория кодов, контролирующих ошибки** — **кодирование**, предназначенное для защиты цифровых данных при передаче по каналам связи от появляющихся ошибок.

Теория кодирования изучает возможности модифицирования данных, чтобы

## Формы представления информации. Модель системы передачи информации.

Различают две системы передачи информации:

* дискретная;
* непрерывная.

Сигнал называется непрерывным, если его параметр в заданных пределах может принимать любые промежуточные значения. Сигнал называется дискретным, если его параметр в заданных пределах может принимать отдельные фиксированные значения.

При дискретном канале связи система связи соединяет источник информации с получателем посредством специальных устройств. В системе связи можно выделить основные функции, представленные на рисунке 1.1 ниже.

Сперва данные, которые поступают к источнику обрабатываются кодером источника. Полученные данные после этого этапа именуются кодовым словом источника. Затем данные обрабатываются кодером канала. Кодовое слово канала может быть представлено в виде битов или группы битов. Далее модулятор преобразует данные в аналоговое представление информации в соответствие с протоколом канала передачи данных. При передаче по каналу могут возникнуть шумы, из-за которых информация может подвергнуться искажению.

Демодулированная информация называется принятым словом. Декодер канала используя избыточность кодового слова источника позволяет декодировать принятое слово с некоторыми найденными ошибками.

В данной работе планируется создать имитацию дискретного канала связи с помехами.

  
  
Рисунок 1.1. Блок схема цифровой системы связи

Кодирование – преобразование сообщения в сигнал, т.е. отображение сообщений сигналами в виде определенного сочетания элементарных дискретных символов, называемых кодовыми комбинациями (кодовыми словами).

Код – правило, согласно которому каждому сообщению однозначно ставится в соответствие некоторая кодовая комбинация. Кодер – устройство, осуществляющее кодирование.

Кодер источника (КИ) – кодер, использование которого позволяет путем устранения избыточности существенно снизить среднее число символов на букву сообщения (такое кодирование называется оптимальным или эффективным). При отсутствии помех это дает выигрыш во времени передачи или в объеме ЗУ, т.е. повышает эффективность системы передачи данных.

Кодер канала (КК) – позволяет путем внесения избыточности обеспечить достоверность передачи данных при наличии помех (такое кодирование называется помехоустойчивым).

Канал – совокупность средств, предназначенных для передачи сигнала от передатчика к приемнику информации (передатчик, приемник, линия связи и т.д.). Канал связи может быть односторонний (симплексный) и двухсторонний (дуплексный).

Передатчик – служит для преобразования электрического сигнала в сигнал, пригодный для передачи по линии связи.

Модуляцией называется изменение параметров переносчика сигнала в соответствии с функцией, отображающей сообщение. Несущим сигналом может быть ток (телеграфия), гармонические низкочастотные или высокочастотные колебания (телефония и т.д.), высокочастотные импульсы (радиорелейная связь и т.д.). Модулируемые параметры называются информативными и могут быть амплитудой, частотой, фазой и т.д. Модулятор – устройство, осуществляющее модуляцию.

При передаче по каналу связи происходит ослабление и искажение передаваемого сигнала, вносимых каналом и действием помех.

Линейные искажения – определяются частотными и временными характеристиками канала. Нелинейные искажения – определяются нелинейностью звеньев канала и видом модуляции.

**Определение энтропии с помощью собственной информации**

Можно определить энтропию случайной величины, введя предварительно понятия распределения случайной величины *X*, имеющей конечное число значений: (1.1).

От основания логарифма зависит единица измерения информации и энтропии: бит, трит и другие.

 Интегральный показатель качества функционирования цифровых систем связи. Определяется как отношение количества искаженных битов данных к общему числу переданных битов.

Мера качества передачи. В общем случае выражается отрицательной степенью 10-7, например 10-7 означает 1 ошибку на 107 бит.

**Коэффициент ошибок**— отношение числа неверно принятых битов (0 вместо 1 и наоборот) к полному числу переданных битов при передаче по каналу связи. Эквивалентно понятию вероятности ошибки. В современных сетях связи характерные значения коэффициента - 1E-9 и лучше.

## Математическое выражение коэффициента битовых ошибок

Определим коэффициент битовых ошибок для реальных приёмников, которым свойственно наличие различных источников шумов. При этом будем считать, что приёмник принимает решение, какой бит (0 или 1) был передан в каждом битовом интервале путем стробирования фототока. Очевидно, что из-за наличия шумов данное решение может быть неверным, что приводит к появлению ошибочных битов. Поэтому, чтобы определить коэффициент битовых ошибок, необходимо понять, каким образом приемник принимает решение относительно переданного бита.

Обозначим через I1 и I0 фототоки, стробированные приемником в течение 1 и 0 битов, соответственно, а через s12 и s02 соответствующие шумы. Принимая, что последние имеют гауссовское распределение, проблема установления истинного значения принятого бита имеет следующую математическую формулировку. Фототок для битов 1 и 0 является выборкой гауссовской переменной со средним значением I1 и вариацией s1, а приёмник должен отслеживать этот сигнал и решать, является ли переданный бит 0 или 1. При этом существует много возможных правил принятия решения, которые могут быть реализованы в приёмнике с целью минимизации коэффициента битовых ошибок. Для значения фототока I этим оптимальным решением является наиболее вероятное значение переданного бита, которое определяется путём сравнения текущего значения фототока с пороговым значением Iп, используемым для принятия решения.

Пусть при  I ³ Iп принимается решение о том, что был передан бит 1, в противном случае – бит 0. Когда биты 1 и 0 равновероятны, что и рассматривается в дальнейшем, пороговый ток приблизительно равен:

https://studfiles.net/html/2706/752/html_5VPw5yrLaC.oMPo/img-GF4lJO.png(1.2)

Вероятность того, что I < Iп, т. е. вероятность ошибки при передаче бита 1, обозначим через Р0,1, а вероятность решения для переданного бита 1, когда I ³ Iп при переданном 0, обозначим Р1,0.

Пусть Q(х) обозначает вероятность того, что нулевая средняя вариация гауссовской переменной превышает значение х, тогда:

https://studfiles.net/html/2706/752/html_5VPw5yrLaC.oMPo/img-9RIwmH.png(1.3) https://studfiles.net/html/2706/752/html_5VPw5yrLaC.oMPo/img-aPnabl.png(1.4)https://studfiles.net/html/2706/752/html_5VPw5yrLaC.oMPo/img-x6_0us.png(1.5)

Можно показать легко увидеть, что BER определяется по формуле,

https://studfiles.net/html/2706/752/html_5VPw5yrLaC.oMPo/img-4_TeoR.png(1.6)

Очень важно отметить, что в ряде случаев эффективным является использование изменяемого в зависимости от уровня сигнала порога принятия решения, как, например, шума оптического усилителя. Многие высокоскоростные приёмники обладают такой особенностью. Однако более простые приемники имеют порог, соответствующий среднему уровню принимаемого тока, а именно (I1 + I0)/2. Такая настройка порогового   значения дает большой коэффициент битовых ошибок, определяемый выражением.

https://studfiles.net/html/2706/752/html_5VPw5yrLaC.oMPo/img-ScdEDR.png(1.7)

Выражение (1.6) можно использовать для оценки BER, когда известны как мощность полученного сигнала, соответствующего битам 0 и 1, так и статистика шумов.

Битовые ошибки являются основным источником ухудшения качества связи, проявляющегося в искажении речи в телефонных каналах, недостоверности передачи информации или снижении пропускной способности передачи данных, и характеризуются статистическими параметрами и нормами на них, которые определены соответствующей вероятностью выполнения этих норм. Последние делятся на долговременные и оперативные нормы, первые из которых определяются рекомендациями ITU-T G.821 и G.826, а вторые – М.2100, М.2110 и М.2120, при этом, согласно М.2100, качество цифрового тракта по критерию ошибок делят на три категории:

· нормальное – BER <10-6;

· пониженное – 10-6≤ BER <10-3 (предаварийное состояние);

· неприемлемое – BER ≥ 10-3 (аварийное состояние).

Так как появление ошибок является следствием совокупности всех текущих условий передачи цифровых сигналов, имеющих случайный характер, то при отсутствии данных о законе распределения ошибок его отдельные элементы могут быть определены с определенной степенью достоверности только по результатам продолжительных измерений. В то же время на практике необходимо, чтобы значения параметров ошибок для ввода в эксплуатацию и технического обслуживания систем передачи основывались на достаточно коротких интервалах времени измерения.

Для измерения коэффициента ошибок разработан ряд специальных BER анализаторов – измерителей коэффициента ошибок, включающих генераторы псевдослучайных и детерминированных последовательностей передаваемых кодированных символов, а также приемное оборудование, осуществляющее собственно измерение коэффициента ошибок. В случае посимвольного сравнения кодов измерение может быть выполнено с использованием шлейфа, т.е. путем измерения ошибок с одной оконечной станции при установке на противоположном конце шлейфа. Другой метод основан на выделении ошибок благодаря избыточности используемых кодов и используется для измерений от передающей до приемной сторон тракта или участка линии, т.е. когда выделение и фиксация ошибок производятся на ее приемном конце. Очевидно, что в первом случае требуется использование одного комплекта, а во втором – двух комплектов приборов. При этом измеренное значение коэффициента ошибок отражает качество передачи при прохождении сигнала в обоих направлениях и в каждом направлении соответственно.

**Помехоустойчивое** **кодирование** — **кодирование**, предназначенное для передачи данных по каналам с помехами, обеспечивающее исправление возможных ошибок передачи вследствие помех.

 Для обнаружения ошибок используют **коды обнаружения ошибок**, для исправления — **помехоустойчивые коды**.

Классификация помехоустойчивых кодов

В непрерывных кодах передаваемая информационная последовательность не разделяется на блоки. Избыточные элементы размещаются в определенном порядке между информационными.

Равномерные блочные коды делятся на разделимые и неразделимые. В разделимых кодах элементы информационной и проверочной частей кодовой комбинации всегда стоят на определенных местах. В неразделимых кодах деление на информационные и проверочные разряды отсутствует.

Разделимые коды, в свою очередь, делятся на систематические (линейные) и несистематические (нелинейные). Систематическими кодами называются блочные разделимые (n,k)-коды, в которых проверочные элементы представляют собой линейные комбинации информационных, несистематические коды таким свойством не обладают.

Помехоустойчивое кодирование предполагает введение в передаваемое сообщение, наряду с информационными, так называемых проверочных разрядов, формируемых в устройствах защиты от ошибок (кодерах-на передающем конце, декодерах — на приемном). Избыточность позволяет отличить разрешенную и запрещенную (искаженную за счет ошибок) комбинации при приеме, иначе одна разрешенная комбинация переходила бы в другую.

Помехоустойчивый код характеризуется тройкой чисел (n, k, d0), где n— общее число разрядов в передаваемом сообщении, включая проверочные (г), k=n-r - число информационных разрядов, d0— минимальное кодовое расстояние между разрешенными кодовыми комбинациями, определяемое как минимальное число различающихся бит в этих комбинациях. Иногда используются дополнительные показатели избыточности, производные от приведенных выше характеристик n, k:R = r/n- относительная избыточность, v = k / n - относительная скорость передачи.

Существующие помехоустойчивые коды можно разделить на ряд групп, только часть из которых используется для обнаружения ошибок в передаваемых по сети пакетах. В группе систематических (линейных) кодов общим свойством является то, что любая разрешенная комбинация может быть получена в результате линейных операций над линейно-независимыми векторами. Это способствует упрощению аппаратной и программной реализации данных кодов, повышает скорость выполнения необходимых операций.

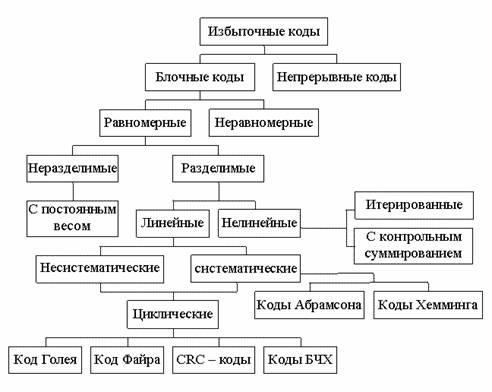
  


Рис. 1.2. Классификация помехоустойчивых кодов

Простейшими систематическими кодами являются биты четности/нечетности. Они не позволяют обнаружить ошибки четной кратности (т.е. ошибки одновременно в двух, четырех и т.д. битах) и поэтому используются при невысоких требованиях к верности принимаемых данных (или при малой вероятности ошибок в линии передачи). Примером может служить бит Parity (соответствие) в установках режимов работы последовательного порта с помощью команды MODE (MS DOS). Несмотря на ограниченные возможности обнаружения ошибок, биты четности/нечетности имеют большое значение в теории помехоустойчивого кодирования. Одни иг первых математически обоснованных и практически использовании? помехоустойчивых кодов - коды Хэмминга представляют собой простс совокупность перекрестных проверок на четность/нечетность. Циклические коды могут рассматриваться как обобщенные проверки на четность/ нечетность

Клод Шеннон сформулировал теорему для случая передачи дискретной информации по каналу связи с помехами, утверждающую, что вероятность ошибочного декодирования принимаемых сигналов может быть обеспечена сколь угодно малой путем выбора соответствующего способа кодирования сигналов. В теореме Шеннона не говорится о том, как нужно строить помехоустойчивые коды. Однако в ней указывается на принципиальную возможность кодирования, при котором может быть обеспечена сколь угодно высокая верность передачи. Это явилось стимулом к разработке помехоустойчивых кодов.

Помехоустойчивость кодирования обеспечивается за счет введения избыточности в кодовые комбинации, т.е. за счет того, что не все символы в кодовых комбинациях используются для передачи информации.

Все помехоустойчивые коды можно разделить на два основных класса: блочные и непрерывные (рекурентные или цепные).

В блочных кодах каждому сообщению (или элементу сообщения) сопоставляется кодовая комбинация (блок) из определенного количества разрядов. Блоки кодируются и декодируются отдельно друг от друга.

Блочные коды могут быть равномерными, когда длина кодовых комбинаций *п* постоянна, или неравномерными, когда *п* непостоянно.

В непрерывных кодах введение избыточности в последовательность входных символов осуществляется без разбивки ее на отдельные блоки. Процессы кодирования и декодирования в непрерывных кодах имеют также непрерывный характер.

Как блочные, так и непрерывные коды в зависимости от методов внесения избыточности подразделяются на разделимые и неразделимые. В разделимых кодах четко разграничена роль отдельных символов. Одни символы являются информационными, другие являются проверочными и служат для обнаружения и исправления ошибок. Разделимые блочные коды называются обычно *п,k-*кодами, где *п –* длина кодовых комбинаций, *k –*число информационных символов в комбинациях.

Неразделимые коды не имеют четкого разделения кодовой комбинации на информационные и проверочные символы.

Разделимые блочные коды делятся, в свою очередь, на несистематические и систематические. Несистематические разделимые коды строятся таким образом, что проверочные символы определяются как сумма подблоков длины *l,* на которые разделяется блок информационных символов.

Большинство известных разделимых кодов составляют систематические коды. У этих кодов значение проверочных символов определяется в результате проведения линейных операций над определенными информационными символами. Для случая двоичных кодов каждый проверочный символ выбирается таким, чтобы его сумма по модулю два с определенными информационными символами стала равной нулю (т.е. сумма единиц была четной). Декодирование сводится к проверке на четность определенных групп символов. В результате таких проверок дается информация о наличии ошибок, а в случае необходимости – о позиции символов, где имеются ошибки.