Санкт-Петербургский государственный политехнический университет

Институт информационных технологий и управления

Кафедра информационных и управляющих систем

Работа допущена к защите

Зав. кафедрой ИУС

\_\_\_\_\_\_\_\_И.Г. Черноруцкий

«\_\_\_\_\_»\_\_\_\_\_\_\_\_ 2015 г.

**ВЫПУСКНАЯ РАБОТА БАКАЛАВРА**

Тема: Разработка протокола канального уровня с прямым исправлением ошибок

Направление**:** 230100 – Информатика и вычислительная техника

Санкт-Петербург

2015 г.

Содержание

[Введение 2](#_Toc419585883)

[1. Анализ методов передачи данных. 5](#_Toc419585884)

[2. Исследование существующих методов исправления ошибок 6](#_Toc419585885)

[2.1 Коды Соломона Рида 6](#_Toc419585886)

[2.2. Коды Боуза—Чоудхури—Хоквингема (БЧХ-коды) 8](#_Toc419585887)

[3. Разработка архитектуры и реализация протокола передачи данных с исправлением данных. 11](#_Toc419585888)

[3.1. Модель p a 11](#_Toc419585889)

[3.3. Двоичный симметричный канал 15](#_Toc419585890)

[3.4. Алгоритм проверки выбранных кодов БЧХ 17](#_Toc419585891)

[3.5. Дейтограммный протокол передачи данных 19](#_Toc419585892)

[3.6. Протокол с задержкой 19](#_Toc419585893)

[3.7. Протокол с возвращением на n шагов 19](#_Toc419585894)

[4. Анализ результатов. 20](#_Toc419585895)

[5. Заключение 21](#_Toc419585896)

[Приложение 1. Исходный код 22](#_Toc419585897)

[Приложение 2. Полученные результаты в ходе работы 23](#_Toc419585898)

# Введение

Каналом передачи информации (каналом связи)называют совокупность устройств, обеспечивающих передачу сигналов с определёнными свойствами с одного пункта к другому. При построении системы канал, как правило, является заданным звеном, с которым источники и получатели должны быть согласованы посредством передатчиков и приёмников.

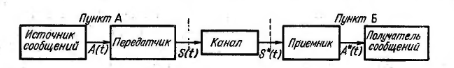


Рис. 1: система передачи информации от одного источника к одному получателю по одному каналу

A(t) – сообщение поступающее от источника.

S(t) – сигнал полученный после преобразования сообщения A(t) передатчиком.

S\*(t) – сигнал полученный на выходе канала (выходной или принимаемый).

A\*(t) – сообщение преобразованное из сигнала S\*(t) приёмником.

Каналы связи характеризуются по различным признакам:

* По используемым линиям связи:
  + Кабельные;
  + Радиорелейные;
  + Тропосферные;
  + И др.
* По полосе частот сигнала в линии:
  + Тональные;
  + Высокочастотные;
  + Коротковолновые;
  + Световые;
  + И др.
* По техническому характеру сигналов и назначению систем связи:
  + Телефонные;
  + Звукового вещания;
  + Телевизионные;
  + Телеграфные;
  + Передачи цифровой информации;
  + И др.

При математическом описании каналы различают, в первую очередь, по типу множеств которым принадлежат входные и выходные сигналы (непрерывные и дискретные), и по характеру изменения этих сигналов во времени (каналы непрерывного и дискретного времени).

Среди непрерывных каналов непрерывного времени можно выделить два наиболее типичных:

1. Непосредственно линия связи, часто с усилительным или переприёмным (ретрансляционным) оборудованием. Такой канал входит в состав всех других каналов связи.
2. Отличается от первого наличием модуляционного и демодуляционного оборудования (часто многоступенного) и полосой пропускания. Примерами таких каналов могут служить телевизионные каналы, тракты систем высокочастотного телефонирования и т.п.

Дискретный канал чаще всего дискретен и по времени. Он состоит из непрерывного канала и подключенных к нему формирователя сигналов (дискретного модулятора) и решающего устройства. Последние и обуславливают дискретность множества входных и выходных элементарных сигналов. В идеальном случае эти два устройства действуют синхронно.

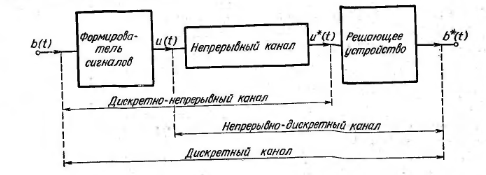


Рис.2: Дискретный канал

Дискретный канал связи предназначен для передачи дискретных сигналов (символов). При передаче по такому каналу сообщение S(t) представляется некоторой последовательностью элементарных дискретных сообщений X(t), принадлежащих конечному множеству. В результате кодирования последовательность X(t) заменяется другой последовательностью Y(t), которая ставится в соответствие E(t). Последовательность Y(t) подаётся на вход дискретного канала.

Основной характеристикой дискретного канала является вероятность того или иного изменения символа на данной позиции. Эта характеристика определяется теми изменениями, которые претерпевает символ при передаче по каналу:

* Смещение во времени;
* Отличие на некоторых позициях выходных символов от входных(аддитивные ошибки);
* Смещение номеров позиций выходной последовательности символов от позиций входной последовательности символов (ошибки синхронизации);
* Появление на некоторых ошибках символов стирания.

При действии рассмотренных факторов основная характеристика дискретного канала зависит от номера позиции, от значения, передаваемого и всех ранее переданных символов.

Поток ошибок представляет собой дискретный случайный процесс E. Каждая позиция E по определённому правилу складывается соответствующей позицией Y (сообщение на выходе).

Представляется логичным и достаточно удобным рассматривать поток ошибок дискретного канала связи как ступенчатый случайный процесс. Такой подход позволяет при исследовании каналов связи использовать многочисленные важные результаты, полученные для случайных процессов.

# Анализ методов передачи данных.

# Исследование существующих методов исправления ошибок

### **2.1 Коды Соломона Рида**

Недвоичные циклические коды, позволяющие исправлять ошибки в блоках данных. Элементами кодового вектора являются не биты, а группы битов (блоки).

Коды Рида — Соломона являются важным частным случаем БЧХ-кода, корни порождающего полинома которого лежат в том же поле, над которым строится код (m=1). Пусть \alpha — элемент поля GF(q) порядка n. Если \alpha —*примитивный* элемент, то его порядок равен q-1, то есть \alpha^{q-1}=1,\quad \alpha^i \neq 1, 0<i<q-1. Тогда нормированный полином g(x) минимальной степени над полем GF(q), корнями которого являются d-1 подряд идущих степеней \alpha^{l_0}, \alpha^{l_0+1},...,\alpha^{l_0+d-2} элемента \alpha, является порождающим полиномом кода Рида — Соломона над полем GF(q):

Обычно полагается l0 = 1. Степень многочлена g(x) равна d-1.

Длина полученного кода n, минимальное расстояние d. Код содержит   проверочных символов, где deg() обозначает степень полинома; число информационных символов k = n-r = n – d + 1.

Кодовый полином c(x) может быть получен из информационного полинома путем его перемножения на порождающий полином:

Код Рида — Соломона является одним из наиболее мощных кодов, исправляющих многократные пакеты ошибок. Применяется в каналах, где пакеты ошибок могут образовываться столь часто, что их уже нельзя исправлять с помощью кодов, исправляющих одиночные ошибки.

(q^m - 1,  q^m -2 - 2t)-код Рида — Соломона над полем \textstyle GF(q^m) с кодовым расстоянием d = 2t + 1 можно рассматривать как ((q^m - 1)m,(q^m -1 - 2t)m)-код над полем \textstyle GF(q), который может исправлять любую комбинацию ошибок, сосредоточенную в t или меньшем числе блоков из m символов. Наибольшее число блоков длины m, которые может затронуть пакет длины l_i, где l_i \leqslant mt_i - (m-1), не превосходит t_i, поэтому код, который может исправить t блоков ошибок, всегда может исправить и любую комбинацию из p пакетов общей длины l, если l+(m-1) \leqslant mt.

При несистематическом кодировании информационное слово умножается на некий неприводимый полином в поле Галуа. Полученное закодированное слово полностью отличается от исходного и для извлечения информационного слова нужно выполнить операцию декодирования и уже потом можно проверить данные на содержание ошибок. Такое кодирование требует большие затраты ресурсов только на извлечение информационных данных, при этом они могут быть без ошибок.

При систематическом кодировании к информационному блоку из k символов приписываются 2t проверочных символов, при вычислении каждого проверочного символа используются все k символов исходного блока. В этом случае нет затрат ресурсов при извлечении исходного блока, если информационное слово не содержит ошибок, но кодировщик/декодировщик должен выполнить k(n-k) операций сложения и умножения для генерации проверочных символов. Кроме того, так как все операции проводятся в поле Галуа, то сами операции кодирования/декодирования требуют много ресурсов и времени. Быстрый алгоритм декодирования, основанный на быстром преобразовании Фурье, выполняется за время порядка .

При операции кодирования информационный полином умножается на порождающий многочлен. Умножение исходного слова S длины k на неприводимый полином при систематическом кодировании можно выполнить следующим образом:

* К исходному слову приписываются 2t нулей, получается полином .
* Этот полином делится на порождающий полином G, находится остаток R,
* Этот остаток и будет корректирующим кодом Рида — Соломона, он приписывается к исходному блоку символов. Полученное кодовое слово C = Sx2r + R.

Кодировщик строится из сдвиговых регистров, сумматоров и умножителей. Сдвиговый регистр состоит из ячеек памяти, в каждой из которых находится один элемент поля Галуа.

Существует и другая процедура кодирования (более практичная и простая). Положим  - примитивный элемент поля GF(q), и пусть   - вектор информационных символов , а значит   - информационный многочлен. Тогда вектор  есть вектор кода Рида - Соломона , соответствующий информационному вектору a. Этот способ кодирования показывает, что для кода РС вообще не нужно знать порождающего многочлена и порождающей матрицы коды, достаточно знать разложение поля GF(q) по примитивному элементу α и размерность кода k (длина кода в этом случае определяется как n = q-1). Все дело в том, что за разностью n-k полностью скрывается порождающий многочлен g(x) и кодовое расстояние.

Декодировщик, работающий по авторегрессивному спектральному методу декодирования, последовательно выполняет следующие действия:

1. Вычисляет синдром ошибки.
2. Строит полином ошибки.
3. Находит корни данного полинома.
4. Определяет характер ошибки.
5. Исправляет ошибки.

### **2.2. Коды Боуза—Чоудхури—Хоквингема (БЧХ-коды)**

БЧХ-код является циклическим кодом, который задается порождающим полиномом. Для его нахождения в случае БЧХ-кода необходимо заранее определить длину кода n  и требуемое минимальное расстояние d \leqslant n. Найти порождающий полином можно следующим образом.

Пусть ~\alpha — примитивный элемент поля ~GF(q^m) (то есть \alpha^{q^m-1}=1, \alpha^i \neq 1, i< q^m-1), пусть ~\beta=\alpha^s , — элемент поля ~GF(q^m) порядка ~n, \quad s = (q^m-1) / n . Тогда нормированный полином g(x) минимальной степени над полем GF(q), корнями которого являются ~d-1 подряд идущих степеней ~\beta^{l_0}, \beta^{l_0+1},\ldots,\beta^{l_0+d-2}элемента ~\beta для некоторого целого ~l_0 (в том числе 0 и 1), является порождающим полиномом БЧХ-кода над полем ~GF(q) с длиной n и минимальный расстоянием ~d_0 \geqslant d. Поясним почему у получившегося кода будут именно такие характеристики (длина кода ~n , минимальное расстояние ~d_0). Действительно, как показано в [1] , длина БЧХ кода равна порядку элемента ~\beta, если ~d>2 и равна порядку элемента ~\beta^{l_0}, если ~d=2, тогда, так как случай ~d=2нам не интересен (такой код не может исправлять ошибки, только обнаруживать), то длина кода будет равна порядку элемента ~\beta ,то есть равна ~n. Минимальное расстояние ~d_0 может быть больше ~d, когда корнями минимальных функций(стр.83[2]) от элементов ~\beta^{l_0}, \beta^{l_0+1},\ldots,\beta^{l_0+d-2} будут элементы расширяющие последовательность, то есть элементы ~\beta^{l_0+d-1},\beta^{l_0+d},\ldots,\beta^{l_0+d_0 - 2}.

Число проверочных символов r равно степени g(x), число информационных символов k = n - r, величина d называется конструктивным расстоянием БЧХ-кода. Если n = qm - 1, то код называется примитивным, иначе не примитивным.

Так же, как и для циклического кода, кодовый полином c(x) может быть получен из информационного полинома m(x), степени не больше k-1, путём перемножения m(x) и g(x):

Для нахождения порождающего полинома необходимо выполнить несколько этапов:

* выбрать q, то есть поле GF(q), над которым будет построен код;
* выбрать длину n кода из условия:

где m,s — целые положительные числа;

* задать величину d конструктивного расстояния;

1) построить циклотомические классы элемента \beta=\alpha^s поля GF(q^m) над полем GF(q), где \alpha — примитивный элемент GF(q^m);

2) поскольку каждому такому циклотомическому классу соответствует неприводимый полином над GF(q), корнями которого являются элементы этого и только этого класса, со степенью равной количеству элементов в классе, то выбрать \beta^{l_0}, \beta^{l_0+1},\ldots, \beta^{l_0+d-2} таким образом, чтобы суммарная длина циклотомических классов была минимальна; это делается для того, чтобы при заданных характеристиках кода ~n и ~d минимизировать количество проверочных символов ~k;

3) вычислить порождающий полином g(x)=f_1(x)f_2(x)\ldots f_h(x), где f_i(x) — полином, соответствующий i-ому циклотомическому классу; или вычислить g(x), как НОК минимальных функций от элементов \beta^{l_0}, \beta^{l_0+1},\ldots,\beta^{l_0+d-2} .

При несистематическом кодировании кодовое слово получается в виде произведения информационного полинома на порождающий:

Оно может быть реализовано при помощи перемножения полиномов.

При систематическом кодировании кодовое слово формируется в виде информационного подблока и проверочного

Пусть информационное слово образует старшие степени кодового слова, тогда

Тогда из условия  , следует

Главной идеей в декодировании БЧХ кодов является использование элементов конечного поля для нумерации позиций кодового слова (или, эквивалентно, в порядке коэффициентов ассоциированного многочлена).

Основными методами декодирования кодов БЧХ:

1. Евклидов алгоритм (ЕА). Из-за высокой регулярности структуры этого алгоритма его широко используют для аппаратной реализации декодеров БЧХ и кодов Рида-Соломона.
2. Прямое решение (алгоритм Питерсона — Горенстейна — Цирлера, ПГЦ). Исторически это первый метод декодирования, найденный Питерсоном для двоичного случая (q=2), затем Горенстейном и Цирлером для общего случая. Этот алгоритм находит коэффициенты *многочлена локаторов ошибок* прямым решением соответствующей системы линейных уравнений. В действительности, так как сложность этого алгоритма растет как куб минимального расстояния ~d, прямой алгоритм может быть использован только для малых значений ~d.
3. Алгоритм Берлекемпа-Мэсси (BMA). Алгоритм поиска кратчайшего регистра сдвига с линейной обратной связью для поданной на вход бинарной последовательности. Также алгоритм позволяет найти минимальный многочлен поданной на вход линейной рекуррентной последовательности над произвольным полем.

# Разработка архитектуры и реализация протокола передачи данных с исправлением данных.

## **3.1. Модель p a**

Основным понятием, положенным в основу данной модели, является плотность ошибка порядка t. Это неслучайная функция от n и t:

.

В числитель является средним числом ошибок на блоке длинной n, содержащих t или больше ошибок. Значения плотности порядка t ограничены снизу величиной t/n, а сверху единицей, т.е. . Значения v(t,n) не убывают с ростом t; . При величине плотности v(1,n) можно судить о степени группирования ошибок, если считать, что увеличение доли ошибок высших кратностей идентично увеличению степени группирования. Для многих каналов, было установлено:

*.*

Параметр α носит название показатель группирования . При значении α равного нулю получаем модель канала с независимыми ошибками, при значении α равного единице канал с “жестким” пакетированием ошибок.

Вероятность приема блока с t ошибками равна:

Использую приближение

Но, на практике обычно применяют более простое соотношение:

Это верхняя граница вероятности . При t/n < 0,3 точные значения близки к верхней границе.

Таким образом, модель(p, a) задается соотношением:

Параметр модели p – вероятность ошибки символа, находится как и для канала ДСК[Надо вставить метод вычисления].

Параметр a вычисляется из уравнения

Данная модель учитывает факт пакетирования ошибок, что возникает в большинстве реальных каналов, имеется возможность единообразно описывать различные типа каналов. Так значения a в кабельных каналах достигает максимального значения(>0.5), а в радиоканалах минимально(~0.3 – 0.45). Но, модель имеет недостаток, заключающейся в вопросе на уровне блоков.

**3.2. Модель ОПП**

Наблюдаемое пакетирование ошибок в каналах связи при предположении о пуассоновском характере потока можно объяснить, если считать параметр  не константой, а случайной величиной или процессом. Получающийся путем рандомизации  новый случайный процесс называют обобщенным пуассоновским . Будем считать  случайной величиной, закон распределения которой известен . Тогда канал задается как поток ошибок первым способом:



Поскольку вид и параметры закона распределения для реальных каналов обычно неизвестны, указанной выше формулой воспользоваться не удается.

По экспериментальным данным относительно легко можно найти закон распределения интервалов между ошибками – функцию Пальма-Хинчина , которая полностью определяет ОПП (второй способ здания потока).

Справедлива формула:

,

где - параметр потока,

 и 

 - вероятность отсутствия ошибок за время .

Таким образом, для ОПП, зная функцию распределения интервалов между ошибками или , вычисляются вероятности , т.е. приходим к конструктивному заданию потока первым способом. Для моделирования будем считать распределение интервалов, заданное обобщенной гиперболой:

, .

Тогда для параметра потока тогда получается:

.

 и

.

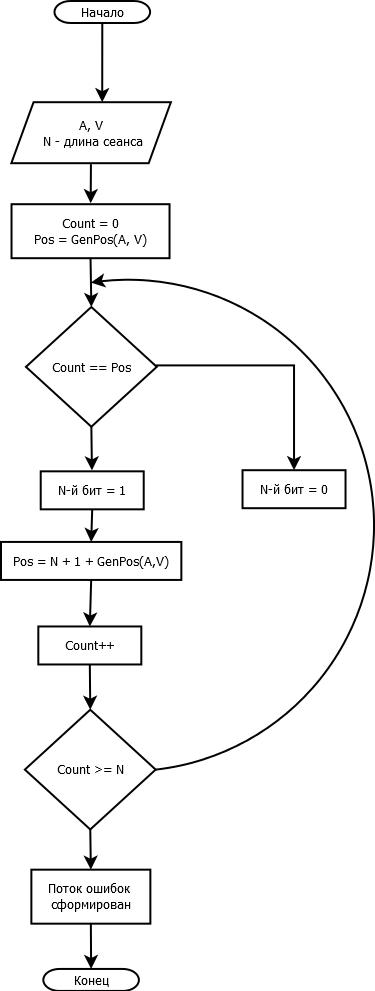
Для расчетов вероятностей наиболее удобна рекуррентная формула:



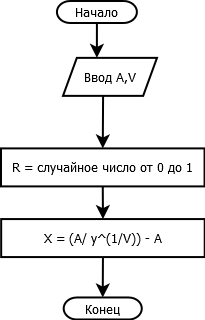
.

Неизвестные параметры  и  легко находятся, например, методом моментов, поскольку обобщенная гипербола для интервалов между ошибками приводит к гамма-распределению параметра .

К недостаткам этой модели можно отнести более трудоемкие формулы для расчета, чем у модели ДСК и тот факт, что не все каналы имеют обобщенную гиперболу в качестве закона распределения между ошибочных интервалов. На рисунках представлен алгоритм генерации потока по данной модели.

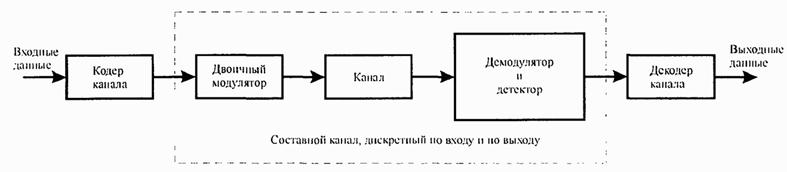


**Рис. : Алгоритм генерации потока ошибок**



**Рис. : Алгоритм функции генерации позиции ошибки в потоке битов.**

**3.3. Двоичный симметричный канал.**

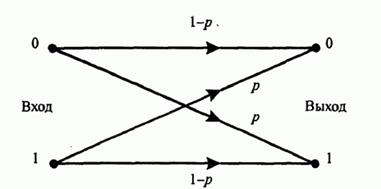


**Рис. . Составной канал, дискретный по входу и по выходу, образованный путём включения в него модулятора и демодулятора/детектора как частей канала**

Если модулятор применяет двоичные сигналы, и детектор делает жёсткие решения, то составной канал, показанный на рис., имеет на входе и выходе двоичную последовательность с дискретным временем. Такой составной канал характеризуется набором X={0,1} возможных входов, набором Y={0,1} возможных выходов и набором условных вероятностей возможных выходов при условии возможных входов. Если канальный шум и другие нарушения вызывают статистически независимые ошибки при передаче двоичной последовательности со средней вероятностью p, тогда

http://www.sernam.ru/archive/arch.php?path=../htm/book_p_net/files.book&file=p_net_114.files/image006.gif

Таким образом, каскадное соединение двоичного модулятора, канала и двоичного демодулятора и детектора в эквивалентный канал с дискретным временем, который представлен графом на рис.. Этот симметричный канал с двоичным входом и двоичным выходом обычно называют двоичным симметричным каналом (ДСК). Поскольку каждый выходной двоичный символ канала зависит только от соответствующего входного двоичного символа.



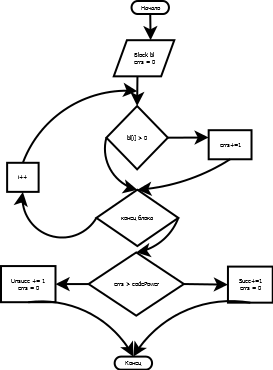
**Рис. Двоичный симметричный канал**

## **Алгоритм проверки выбранных кодов БЧХ**

Проверку выбранных кодов БЧХ будем проведем по следующему алгоритму. Так как в нашей модели данные представляется в виде нулей, ошибки представляются в виде единиц, то можем сгенерировать следующей алгоритм проверки кодов БЧХ:

1. Задаем счетчик ошибок errs = 0
2. Проверяем каждый элемент блока на 1 bl[i] > 0
3. Если получаем true увеличиваем счетчик errs на единицу
4. Проверяем достигли конца блока, если не достигли то переходим к следующему элементу в блоке
5. Достигли конца блока и делаем проверку errs > codePower, если да то увеличиваем счетчик Unsucc на единицу, если нет то счетчик Succ на единицу
6. Обнуляем счетчик errs

Для наглядности блок-схема алгоритма приведена на рисунке ниже.



**Рис. : Алгоритм проверки выбранных БЧХ кодов.**

## **3.5. Дейтограммный протокол передачи данных**

## **3.6. Протокол с задержкой**

## **3.7. Протокол с возвращением на n шагов**

# 4. Анализ результатов.

# 5. Заключение

В рамках данной работы для решения задачи разработки протокола канального уровня с прямым исправлением ошибок был осуществлен анализ существующих алгоритмов и способов помехоустойчивой передачи данных и предложена архитектура для семейства протоколов. В ходе решения поставленной задачи был разработан инструмент оценочного моделирования в виде компьютерной программы, позволяющий провести моделирование протокола при использовании различных потоков ошибок. С использованием программы был проведен анализ и для каждого потока ошибок были предложены оптимальные параметры протокола.

Так же в ходе данной работы была создана методика синтеза протокола, которую можно использовать повторно. Алгоритм синтеза описывается следующим образом:

1. Формулирование технического задания на разработку протокола (характеристики потоков ошибок; характеристики передаваемых данных; требования к скорости и надежности передачи данных).
2. Выбор необходимой архитектуры семейства протоколов. В данной работе, есть возможность выбора архитектуры протокола, что позволяет решать большое множество заданий.
3. На основе требований определить параметры протокола, которые однозначно определены.
4. Использую программу оценочного моделирования получить параметры протокола, которые невозможно определить однозначно.

# Список использованных источников информации

# Список литературы

**Р. Блейхут** Теория и практика кодов, контролирующих ошибки: Пер. с англ. [Книга]. - Москва : Мир, 1986.

**Э. Таненбаум** Компьютерные сети [Книга]. - СПб : Питер, 2003.

## **Приложение 1 Исходный код**

## **Приложение 2 Полученные результаты в ходе работы**