Учреждение образования

«БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНОЛОГИЧЕСКИЙ УНИВЕРСИТЕТ»

**Избыточное кодирование данных в информационных системах. Код Хемминга.**

Студент: Дрозд А. И.

ФИТ 3 курс 2 группа

Преподаватель:

Нистюк Ольга Александровна

Минск 2024

**Цель**: приобретение практических навыков кодирования/декодирования двоичных данных при использовании кода Хемминга.

Задачи:

1. Закрепить теоретические знания по использованию методов помехоустойчивого кодирования для повышения надежности передачи и хранения в памяти компьютера двоичных данных.
2. Разработать приложение для кодирования/декодирования двоичной информации кодом Хемминга с минимальным кодовым расстоянием 3 или 4.
3. Результаты выполнения лабораторной работы оформить в виде описания разработанного приложения, методики выполнения экспериментов с использованием приложения и результатов эксперимента.

# **Теоретические сведения**

Надежность системы – характеристика способности программного, аппаратного, аппаратно-программного средства выполнить при определенных условиях требуемые функции в течение конкретного периода времени.

Достоверность работы системы (устройства) – свойство, характеризующее истинность конечного (выходного) результата работы (выполнения программы), определяемое способностью средств контроля фиксировать правильность или ошибочность работы.

Ошибка устройства – неправильное значение сигнала (бита – в цифровом устройстве) на внешних выходах устройства или отдельного его узла, вызванное технической неисправностью, или воздействующими на него помехами (преднамеренными либо непреднамеренными), или иным способом.

Ошибка программы – проявляется в не соответствующем реальному (требуемому) промежуточном или конечном значении (результате) вследствие неправильно запрограммированного алгоритма или неправильно составленной программы.

Надежность является комплексным свойством, включающим в себя единичные свойства: безотказность, ремонтопригодность, сохраняемость, долговечность.

Простейшая структурная схема дополнена двумя интересующими нас блоками: кодером (канала), осуществляющим преобразование исходного сообщения (информационного слова) Хk (длина сообщения – k символов) в избыточное сообщение (кодовое слово) Xn длиной n символов (n > k), и декодером (канала). Схема представлена на рисунке 1.1.

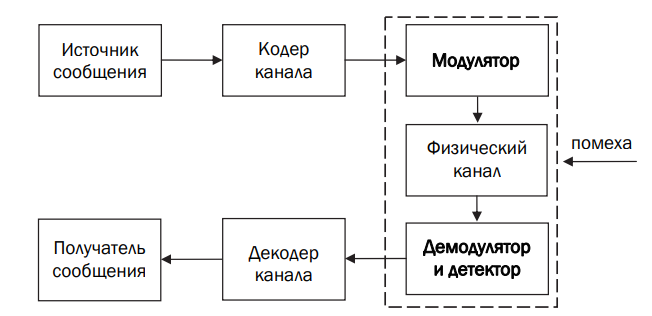


Рисунок 1.1 - Структурная схема системы передачи данных с избыточными средствами аппаратного контроля

Код Хемминга относится к классу линейных блочных кодов.

Линейные блочные коды – это класс кодов с контролем четности, которые можно описать парой чисел (п, k). Для формирования r проверочных символов (кодирования), т. е. вычисления проверочного слова Xr, используется порождающая матрица G: совокупность базисных векторов будем далее записывать в виде матрицы G размерностью k×n с единичной подматрицей (I) в первых k строках и столбцах. Формула для записи матрицы представлена на рисунке 1.2.



Рисунок 1.2 – Формула для построения матрицы G

Более точно матрица G называется порождающей матрицей линейного корректирующего кода в приведенно-ступенчатой форме. Кодовые слова являются линейными комбинациями строк матрицы G (кроме слова, состоящего из нулевых символов).

Кодирование заключается в умножении вектора сообщения Хk длиной k на порождающую матрицу по правилам матричного умножения (все операции выполняются по модулю 2). Очевидно, что при этом первые k символов кодового слова равны соответствующим символам сообщения, а последние r символов образуются как линейные комбинации первых [5, 9]. Для всякой порождающей матрицы G существует матрица Н размерности r×n, задающая базис нулевого пространства кода и удовлетворяющая равенству, представленному на рисунке 1.3.



Рисунок 1.3 – Равенство, соответствующее матрице Хемминга

В коде Хемминга с минимальным кодовым расстоянием dmin = 3 проверочная матрица Н имеет классический вид и состоит из двух подматриц: P’ размером k×r и I размером r×r соответственно. В последнем выражении I – единичная матрица порядка r (r×r). Количество r избыточных (проверочных) символов кодового слова определяется из следующей простой логической цепи рассуждений. Общее число всех возможных комбинаций 2r должно удовлетворять неравенству, представленному на рисунке 1.4.



Рисунок 1.4 – Равенство для избыточных символов

Результат умножения сообщения на выходе канала передачи (Yn) или (что равнозначно) сообщения, считываемого из памяти, на проверочную матрицу (Н) называется синдромом (вектором ошибки) S. Рисунок 1.5.



Рисунок 1.5 – Формула нахождения Синдрома

Синдром – это результат проверки четности, выполняемой над сообщением Yn для определения его принадлежности заданному набору кодовых слов. При положительном результате проверки синдром S равен 0, т. е. Yn = Хn. Если Yn содержит ошибки, которые можно исправить, то синдром имеет определенное ненулевое значение, что позволяет обнаружить и исправить конкретную ошибочную комбинацию.

Для вычисления проверочных символов, исходя из проверочной матрицы и исходного сообщения используем формулу, представленную на рисунке 1.6.

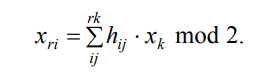


Рисунок 1.6 – Формула нахождения проверочных символов

**Практическое задание**

Изначально необходимо ввести строку с произвольным символом и переведем ее в бинарную строку. После этого, используя проверяющую матрицу, добавляем проверочные символы для исходного сообщения. На рисунке 2.1 видно исходную строку, ее бинарное представление и представление с добавленными избыточными битами:



Рисунок 2.1 – Входные данные и бинарные строки

Далее мы принимаем сообщение с 1 ошибкой и выводим проверочную матрицу для нахождения избыточных битов принятого сообщения. После нахождения избыточных битов, необходимо вычислить синдром сообщения и исправить найденную ошибку. Синдром вычисляется по формуле 1.5. Результат представлен на рисунке 2.2:

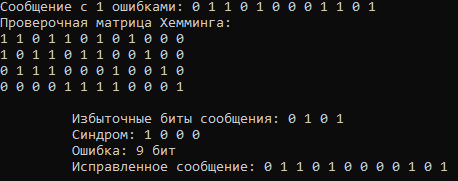


Рисунок 2.2 – Прием и исправление сообщения с 1 ошибкой

Аналогичные действия происходят и при приеме сообщения с 2 ошибками. Однако в данном случае возможно только определить ошибочные биты, без исправления сообщения. Эти биты находятся с помощью XOR-операции между двумя столбцами, результатом которой будет являться найденный синдром. Результат представлен на рисунке 2.3:

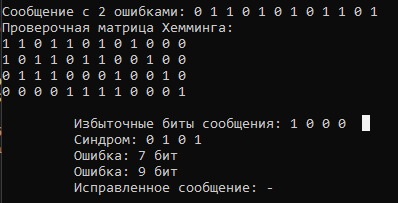


Рисунок 2.3 - Прием сообщения с 2 ошибками

**Вывод:**

Код Хемминга применяется для обнаружения и исправления ошибок при передаче данных. Он использует контрольные биты и относится к классу линейных блочных кодов. В данной лабораторной работе изучалась схема построения проверочной матрицы, нахождения синдрома и коррекции ошибок в коде Хемминга.