Защита информации

Павел Юдаев

МГТУ им. Баумана, Кафедра ИУ-9

Москва, 2014

Применение кодов, исправляющих ошибки

Угрозы и защита информации:

1) непреднамеренное искажение при передаче.

Защита - коды, исправляющие ошибки:

- Хэмминга (исправляют 1 ошибку)
- Рида-Маллера
- Рида-Соломона (испр. не менее заданного числа ошибок; компакт-диски, цифровое видео DVB-S)
- коды БЧХ (испр. не менее заданного числа ошибок, большое семейство)

Зачем нужны коды, исправляющие ошибки? (дополнительный материал)

Хотим передавать много данных с малой вероятностью искажения. Канал передачи шумный. Сигнал искажается.

Искажения (ошибки) только типа замены. Без стираний.

Демодуляция - преобразование сигнала (напр., радиосигнала) в цифровые данные.

Чем выше мощность сигнала, тем больше отношение сигнал/шум на приемнике и меньше вероятность неправильной демодуляции.

Почему **нельзя просто повысить мощность** передатчика или чувствительность приемника?
Две основные причины:

1) Пусть мощность источника сигнала = 1. На расстоянии r эта же мощность распределена по поверхности сферы площадью $4\pi r^2$. Поэтому для направленного излучателя мощность на приемнике пропорциональна $\alpha/(4\pi r^2)$.
Квадратично убывает от расстояния, линейно растет от сужения диаграммы направленности α.

2) Мощность, чувствительность и узость диаграммы направленности излучателя и приемника ограничены. Особенно у спутников.

Замечание: Без использования кодов
вероятность правильного приема бита увеличивается при уменьшении скорости передачи информации.
Метод - кратное повторение сигнала.
Скорость передачи падает в разы. Это не эффективно.

Пусть вероятность демодулировать бит неправильно мала и равна θ - вероятность ошибки.
Тогда вероятность правильно демодулировать пакет из n битов равна $\delta_{n,0} = (1-\theta)^n$. Убывает экпоненциально с ростом n.
Плохо.

Вероятность правильно демодулировать $n-t$ символов из n равна $\delta_{n,t} = C_n^t \cdot (1-\theta)^{n-t}$.
При $t = 1$ $\delta_{n,t} = n \cdot (1-\theta)^{n-t} \approx n \cdot \delta_{t,0}$
При $t = n/2$ $\delta_{n,t} \sim (1-\theta)^{n-t} \cdot 2^{n+0.5}/\sqrt{\pi n}$ асимптотически при больших n.

Т.е. с ростом t $\delta_{n,t}$ начинает стремительно увеличиваться.
Поэтому нужны коды, исправляющие много ошибок в длинных пакетах данных.

Скорость передачи информации (отношение числа информационных бит* в кодовом слове к длине кодового слова) при этом падает незначительно.

Например, код Рида-Соломона,
исправляющий t ошибочных блоков из $q^m - 1$
при длине блока в m бит,
имеет скорость передачи $\frac{q^m - 1 - 2t}{q^m - 1}$.

*Все биты кодового слова можно условно поделить на информационные и проверочные.

Пусть $q = 2$ (двоичные данные), $m = 8$ (блок: 1 байт = 8 бит). Длина кодового слова равна $(2^8 - 1) \cdot 8 = 2040$ бит = 255 байт.

Пусть $t = 10$ (код исправляет произвольное число ошибок, компактно расположенных в каких-нибудь 10 байтах из 255). Число информационных бит в кодовом слове $(2^8 - 21) \cdot 8 = 1880$, или 235 байт.

Скорость передачи $235/255 = 0.92$.

Эта конструкция нам поможет, например, если ошибки расположены компактно и вероятность ошибки в одном бите меньше $\approx 10/(255 * 2) \approx 1/50$ (очень грубая оценка).

Применение кодов Рида-Соломона:

Связь со спутниками “Вояджер”, компакт диски, цифровое телевидение и др.

In 1977, Reed-Solomon (RS) codes were notably implemented in the Voyager program in the form of concatenated codes. The first commercial application in mass-produced consumer products appeared in 1982 with the compact disc, where two interleaved RS codes are used. Today, RS codes are widely implemented in digital storage devices and digital communication standards, though they are being slowly replaced by more modern low-density parity-check (LDPC) codes or turbo codes. For example, RS codes are used in the digital video broadcasting (DVB) standard DVB-S, but LDPC codes are used in its successor DVB-S2.