Передача информации. Помехи. Помехозащитное кодирование

Александра Игоревна Кононова

МИЭТ

8 июля 2024 г. — актуальную версию можно найти на
`https://gitlab.com/illinc/otik`

Информационный канал

— совокупность устройств, объединённых линиями связи, предназначенных для передачи информации от источника информации (начального устройства канала) до её приёмника (конечного устройства канала).

- достоверность передачи информации;
- надёжность работы устройств;
- скорость передачи информации (пропускная способность, ёмкость);
- задержка сигнала во времени (латентность).

X, Y — источники сообщений: по каналу передаются сообщения из X. Из-за помех приёмником воспринимается Y.

Защита от помех

Защита от помех — добавление избыточности соответственно помехам: $|code(x)| \to |x| + \Delta$.

Если известен источник $X \ni x$ и изначальный код избыточен, его избыточность желательно вначале удалить: $|code(X)| \to I(X) + \Delta$.

Величина добавленной избыточности Δ соответствует характеристикам канала.

Пропускная способность (ёмкость) C канала

$$C = \lim_{T\to\infty} \frac{\max\limits_X (I(X,Y))}{T} \quad \left[\frac{\text{бит}}{\text{с}}\right]$$

бод — по одним источникам то же, по другим — бод $= \frac{\text{тактов}}{\text{с}}$

— максимальное количество информации, передаваемое в единицу времени.

Для канала без шума: $$C = \lim_{T\to\infty} \frac{\max\limits_X (I(X))}{T} = \lim_{T\to\infty} \frac{\log_2 N(T)}{T},$$

где $N(T)$ — число всех возможных сигналов (сообщений) за время T.

Первая теорема Шеннона (для канала без помех)

1. При любой производительности источника сообщений, меньшей пропускной способности канала: $\frac{I(X)}{T} < C$, существует способ кодирования, позволяющий передавать по каналу все сообщения, вырабатываемые источником.
2. Не существует способа кодирования, обеспечивающего передачу сообщения без их неограниченного накопления, если $\frac{I(X)}{T} > C$.

Вторая теорема Шеннона (для канала с помехами)

1. При любой производительности источника сообщений, меньшей пропускной способности канала:
$$\frac{I(X)}{T} < C$$
существует способ кодирования, позволяющий обеспечить передачу всей информации со **сколь угодно малой вероятностью ошибки.**

2. Не существует способа кодирования, обеспечивающего передачу информации со сколь угодно малой вероятностью ошибки, если
$$\frac{I(X)}{T} > C$$

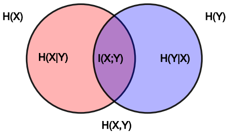
Матмодель канала

1. источник X сообщений на входе, Y — на выходе;
2. условные вероятности — статистические свойства шумов (помех):

$p(y_j|x_i) = \frac{p(x_i,y_j)}{p(x_i)}$ — вероятность того, что отправив x_i — получим y_j

$p(x_i|y_j) = \frac{p(x_i,y_j)}{p(y_j)}$ — после получения y_j, что было отправлено именно x_i

---



$H(X) = I(X)$ — энтропия X (средняя информация в X)
$H(Y) = I(Y)$ — энтропия Y (средняя информация в Y)
$I(X,Y) = I(Y,X)$ — относительная информация X и Y
$H(X,Y) = H(Y,X)$ — энтропия объединения X и Y

$H(Y|X)$ — условная энтропия Y относительно X (шум)
$H(X|Y)$ — условная энтропия X относительно Y (инф. потери)

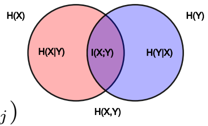
Канал без шумов: $X = Y$, $p(y|x) = \begin{cases} 1, & \text{при } y = x \\ 0, & \text{при } y \neq x \end{cases}$ $\quad I(X,Y) = I(X)$

Взаимные информация и энтропия

$$I(X,Y) = \sum_i \sum_j p(x_i, y_j) \log_2 \frac{p(x_i, y_j)}{p(x_i) \cdot p(y_j)}$$

$$H(X|Y) = M[-\log_2 p(X|Y)] = -\sum_i \sum_j p(x_i, y_j) \cdot \log_2 p(x_i|y_j) =$$

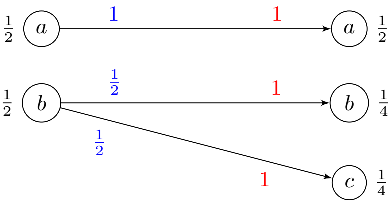
$$\left[p(x_i|y_j) = \frac{p(x_i, y_j)}{p(y_j)}\right] = -\sum_j p(y_j) \sum_i p(x_i|y_j) \cdot \log_2 p(x_i|y_j)$$



$$H(X,Y) = M[-\log_2 p(X,Y)] = -\sum_i \sum_j p(x_i, y_j) \cdot \log_2 p(x_i, y_j)$$

1. $I(X,Y) \geqslant 0, \quad I(X,Y) = 0 \Leftrightarrow X$ и Y независимы;
2. $H(X) = 0 \ \big(I(X) = 0\big) \ \Leftrightarrow \ X$ — константа;
3. $I(X,Y) = I(Y,X)$;
4. $I(X,Y) = I(X) + I(Y) - H(X,Y) = I(X) - H(X|Y) = I(Y) - H(Y|X)$
5. $I(X,Y) \leqslant I(X,X) = I(X) = H(X)$.
   Если $I(X,Y) = I(X)$, то X — функция от Y (разные y при разных x, передача без потерь).

Шум и потери



$I(X) = 1,$ $\quad I(Y) = \frac{3}{2}$

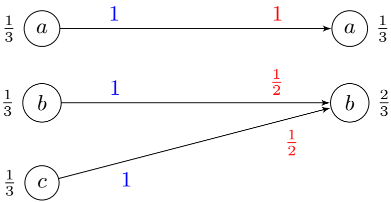
$H(X|Y) = 0$

$$H(Y|X) = -\sum_i \sum_j p(x_i, y_j) \log_2 p(x_i|y_j) =$$

$$= p(a,a) \cdot 0 + p(b,b) \cdot 1 + p(b,c) \cdot 1 =$$
$$= p(x{=}b) = \frac{1}{2}$$

$I(X,Y) = 1 = I(X)$

Есть шумы, нет потерь



$I(X) = \log_2 3,$ $\quad I(Y) = \log_2 3 - \frac{2}{3}$

$H(Y|X) = 0$

$$H(X|Y) = -\sum_i \sum_j p(x_i, y_j) \log_2 p(y_j|x_i) =$$

$$= p(a,a) \cdot 0 + p(b,b) \cdot 1 + p(c,b) \cdot 1 =$$
$$= p(y{=}b) = \frac{2}{3}$$

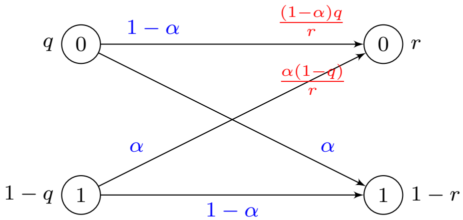
$I(X,Y) = \log_2 3 - \frac{2}{3} = I(Y)$

Есть потери, нет шума

Двоичный симметричный канал

От X к Y передаются символы 0 и 1 (k символов в единицу времени).

Каждый символ, независимо от других, с вероятностью α инвертируется.
Есть как шум, так и потери.



Пусть X производит $x_1 = 0$ и $x_2 = 1$ с вероятностями q и $1-q$,
на выходе Y — $y_1 = 0$ и $y_2 = 1$ с вероятностями $r = (1-\alpha)q + \alpha(1-q)$ и $1-r$.

$$H(Y|X) = -\sum_i p(x_i) \sum_j p(y_j|x_i) \cdot \log_2 p(y_j|x_i) = q \cdot H(Y|x{=}0) + (1-q) \cdot H(Y|x{=}1)$$

$$H(Y|x{=}0) = -\sum_{j=1}^{2} p(y_j|x{=}0) \cdot \log_2 p(y_j|x{=}0) = -(1-\alpha)\log_2(1-\alpha) - \alpha\log_2\alpha$$

$$H(Y|x{=}1) = -\sum_{j=1}^{2} p(y_j|x{=}1) \cdot \log_2 p(y_j|x{=}1) = -\alpha\log_2\alpha - (1-\alpha)\log_2(1-\alpha) = H(Y|x{=}0)$$

$$H(Y|X) = \Big(q + (1-q)\Big) \cdot H(Y|x{=}0) = H(Y|x{=}0) = -\alpha \cdot \log_2\alpha - (1-\alpha) \cdot \log_2(1-\alpha)$$

$I(Y) = -r \cdot \log_2 r - (1-r) \cdot \log_2(1-r)$

Передаваемая информация на символ $I(X,Y) = I(Y) - H(Y|X) =$
$= \Big(- r \cdot \log_2 r - (1-r) \cdot \log_2(1-r)\Big) - \Big(- \alpha \cdot \log_2 \alpha - (1-\alpha) \cdot \log_2(1-\alpha)\Big)$

Обозначим $\eta(x) = -x \cdot \log_2 x$: $I(X,Y) = \big(\eta(r) + \eta(1-r)\big) - \big(\eta(\alpha) + \eta(1-\alpha)\big)$

Макс. передаваемая информация на символ
$\max\limits_X \big(I(X,Y)\big) = \max\limits_X \Big(\big(\eta(r) + \eta(1-r)\big) - \big(\eta(\alpha) + \eta(1-\alpha)\big)\Big) =$
$= \max\limits_r \Big(\eta(r) + \eta(1-r)\Big) - \Big(\eta(\alpha) + \eta(1-\alpha)\Big) \quad = \quad 1 - \Big(\eta(\alpha) + \eta(1-\alpha)\Big)$

Пропускная способность:

$$C = k \cdot \max_X \big(I(X,Y)\big) = k \cdot \Big(1 - \big(\eta(\alpha) + \eta(1-\alpha)\big)\Big)$$

При $\alpha = 0$ или единице $C = k$; при $\alpha = 0{,}5$ получим $C = 0$.

Вероятность бессбойной передачи m битов: $p(m,0) = (1-\alpha)^m$
одиночной инверсии в блоке из m битов: $p(m,1) = m \cdot \alpha(1-\alpha)^{m-1}$
двойной инверсии: $p(m,2) = C_m^2 \cdot \alpha^2(1-\alpha)^{m-2} = \frac{m(m-1)}{2}\alpha^2(1-\alpha)^{m-2}$

При $m = 8 \cdot 16$ и $\alpha = 10^{-5}$:
$p(m,0) \approx 0{,}9987$; $p(m,1) \approx 0{,}0013$; $p(m,2) \approx 8{,}1 \cdot 10^{-7}$
$p(8m,0) \approx 0{,}99$; $p(8m,1) \approx 0{,}01$; $p(8m,2) \approx 5{,}2 \cdot 10^{-5}$

Помехозащитное кодирование

Файл **разрезается на блоки** по N байт (последний блок, если неполный, дополняется до N), каждый из которых дополняется избыточными (контрольными) данными до M байт.

Размер блока (N и M) выбирается исходя из:

- особенностей алгоритма (удобства реализации);
- свойств канала (информационных потерь);

и ни в коем случае не зависит от размера файла n.

Совместно: вначале применяются все алгоритмы сжатия, затем — защита от помех.

После декодирования необходимо восстановить исходную длину файла n!

Синдром S блока — величина, равная нулю при успешной передаче (для непротиворечивого блока) и указывающая на место ошибки при $S \neq 0$.

Простейшие помехозащитные коды

1. Обнаруживающий одиночную ошибку (здесь и далее — инверсию) в одном бите — двойное повторение каждого бита.
2. Обнаруживающий одиночную ошибку в блоке из ν бит — **контроль чётности** (добавление к каждому блоку $\nu+1$-го бита так, чтобы дополнить количество единиц до заранее выбранного для кода чётного (even) или нечетного (odd) значения).

   Двойная ошибка в блоке не будет обнаружена.
3. Исправляющий одиночную ошибку в одном бите — тройное повторение каждого бита.
4. Исправляющий одиночную ошибку в блоке из μ бит — код Хэмминга.

   Двойная ошибка в блоке будет принята за одиночную не в том месте.
5. Исправляющий одиночную ошибку и обнаруживающий двойную в блоке из $\mu+1$ бит — **код Хэмминга с дополнительным битом чётности.**

Положения кода Хэмминга

1. Информация передаётся блоками.
2. В блоке (μ битов) никогда не встретится более чем одна ошибка.
3. Ошибка — инверсия бита.

Биты блока разделяются на
- **информационные** (независимые)
- и **проверочные** (значение рассчитывается по информационным).

Общий размер блока после кодирования

$$\mu = (\nu \text{ информационных}) + (\kappa \text{ проверочных})$$

— ошибки нет;
— ошибка в i-й позиции.

$\mu+1$ указаний

$2^\kappa \geqslant \mu + 1$

κ	1	2	3	4	5	6	7	8	9	10	11
$\sup(\mu) = 2^\kappa - 1$	1	3	7	15	31	63	127	255	511	1023	2047
$\sup(\nu) = \sup(\mu) - \kappa$	0	1	4	11	26	57	120	247	502	1013	2036

Бит чётности и группы

1. Бит чётности позволяет обнаружить одиночную ошибку в группе:

$$c = \bigoplus_{b_i \in G} b_i \text{ , соответственно, } \bigoplus_{b_i \in \{c\} \cup G} b_i = c \oplus \bigoplus_{b_i \in G} b_i = 0$$

при одиночной ошибке в $\{c\} \cup G$ получим $\bigoplus_{b_i \in \{c\} \cup G} b_i = 1$.

2. Несколько пересекающихся контрольных групп позволяют уточнить положение ошибки.

---

3. Набор групп должен быть различным для каждого бита (для локализации ошибки до конкретного бита).
4. Контрольный бит не должен входить более чем в одну группу (для упрощения расчёта).
5. Каждый информационный бит должен входить как минимум в две группы (из 3 и 4).

Несистематический (наивный) код Хэмминга

- Набор контрольных групп — единицы натурального двоичного кода номера бита (с 1, чтобы каждый входил хотя бы в одну группу).

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
×		×		×		×		×		×		×		×
	×	×			×	×			×	×			×	×
			×	×	×	×					×	×	×	×
							×	×	×	×	×	×	×	×

Для 15 бит (11 инф-х + 4 кон-х)

$$\begin{aligned}
&\text{КС1:}\quad b_1 \oplus b_3 \oplus b_5 \oplus b_7 \oplus b_9 \oplus b_{11} \oplus b_{13} \oplus b_{15} = 0\\
&\text{КС2:}\quad b_2 \oplus b_3 \oplus b_6 \oplus b_7 \oplus b_{10} \oplus b_{11} \oplus b_{14} \oplus b_{15} = 0\\
&\text{КС3:}\quad b_4 \oplus b_5 \oplus b_6 \oplus b_7 \oplus b_{12} \oplus b_{13} \oplus b_{14} \oplus b_{15} = 0\\
&\text{КС4:}\quad b_8 \oplus b_9 \oplus b_{10} \oplus b_{11} \oplus b_{12} \oplus b_{13} \oplus b_{14} \oplus b_{15} = 0
\end{aligned}$$

- Биты $1, 2, 4, \dots 2^s$ — контрольные (входят только в одну группу):
  $b_1 = b_3 \oplus b_5 \oplus b_7 \oplus b_9 \oplus b_{11} \oplus b_{13} \oplus b_{15} \dots$
  $b_2 = b_3 \oplus b_6 \oplus b_7 \oplus b_{10} \oplus b_{11} \oplus b_{14} \oplus b_{15} \dots$
  $\dots$
- При наличии ошибки несошедшиеся контрольные суммы образуют натуральный двоичный код инвертированного бита $\rightarrow$ исправление.

Систематический код Хэмминга

Перестановка столбцов кода Хэмминга образует другой код Хэмминга

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
×		×		×		×		×		×		×		×
	×	×			×	×			×	×			×	×
			×	×	×	×					×	×	×	×
							×	×	×	×	×	×	×	×

Систематический код Хэмминга (простейший):

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
3	5	6	7	9	10	11	12	13	14	15	1	2	4	8
×	×		×	×		×		×		×	×			
×		×	×		×	×			×	×		×		
	×	×	×				×	×	×	×			×	
				×	×	×	×	×	×	×				×

Систематический код Хэмминга

Перестановка столбцов кода Хэмминга образует другой код Хэмминга

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
×		×		×		×		×		×		×		×
	×	×			×	×			×	×			×	×
			×	×	×	×					×	×	×	×
							×	×	×	×	×	×	×	×

Систематический код Хэмминга (Л. Бриллюэн):

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
15	7	11	13	14	3	5	9	6	10	12	1	2	4	8
×	×	×	×		×	×	×				×			
×	×	×		×	×			×	×			×		
×	×		×	×		×		×		×			×	
×		×	×	×			×		×	×				×

Коды, исправляющие одиночную ошибку и обнаруживающие двойную $\mu + 1 = 2^\kappa$

Длина блока Хэмминга $\mu = 2^\kappa - 1$ бит $\to$ один бит не используется.

$$b_0 = \bigoplus_{i=1}^{n-1} b_i \text{ — дополнительный бит чётности } \left(\bigoplus_{i=0}^{n-1} b_i = 0\right)$$

0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
	×		×		×		×		×		×		×		×
		×	×			×	×			×	×			×	×
				×	×	×	×					×	×	×	×
								×	×	×	×	×	×	×	×
×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×

Количество единиц в контрольных группах	**Общее количество единиц**	**Вывод**
Чётное во всех	Чётное	Данные верны
Чётное во всех	Нечётное	Ошибка в дополнительном контрольном разряде b_0
Нечётное в некоторых	Нечётное	Однократная ошибка в коде Хэмминга $b_1 \dots b_n$
Нечётное в некоторых	Чётное	Двойная ошибка

Систематический код Хэмминга с контролем двойной инверсии

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
15	7	11	13	14	3	5	9	6	10	12	1	2	4	8	0
×	×	×	×		×	×	×				×				
×	×	×		×	×			×	×			×			
×	×		×	×		×		×		×			×		
×		×	×	×			×		×	×				×	
×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×

Биты 1-11 — информационные, 12-16 — контрольные:

$b_{12} = k_1 = b_1 \oplus b_2 \oplus b_3 \oplus b_4 \oplus b_6 \oplus b_7 \oplus b_8$

$b_{13} = k_2 = b_1 \oplus b_2 \oplus b_3 \oplus b_5 \oplus b_6 \oplus b_9 \oplus b_{10}$

$b_{14} = k_4 = b_1 \oplus b_2 \oplus b_4 \oplus b_5 \oplus b_7 \oplus b_9 \oplus b_{11}$

$b_{15} = k_8 = b_1 \oplus b_3 \oplus b_4 \oplus b_5 \oplus b_8 \oplus b_{10} \oplus b_{11}$

$b_{16} = k_0 = b_1 \oplus b_2 \oplus b_3 \oplus b_4 \oplus b_5 \oplus b_6 \oplus b_7 \oplus b_8 \oplus b_9 \oplus b_{10} \oplus b_{11} \oplus b_{12} \oplus b_{13} \oplus b_{14} \oplus b_{15}$

Расчёт контрольных битов — битовая маска + подсчёт единиц в числе.

Расчёт позиции по синдрому — таблица.

Размер блока $N \to M$ (октетов) I

На практике: натуральное число байтов (байт обычно кратен октету):

- блок до кодирования: ν информационных битов $\sim N$ инф-х октетов [всегда $\nu \leqslant 8N$, в идеале $\nu = 8N$ — запишем $\nu \lessapprox 8N$];
- вносимая избыточность: $\kappa + 1$ контрольных битов (κ битов Хэмминга + бит общей чётности) $\sim K$ контрольных октетов [$\kappa + 1 \lessapprox 8K$];
- блок после кодирования: $\mu + 1 = \nu + \kappa + 1$ битов $\sim M = N + K$ октетов [$\mu + 1 \lessapprox 8M$]. В блоке из M октетов исправляется одна ошибка, обнаруживается двойная: $8M \gtrapprox \mu + 1 = (\nu + \kappa) + 1 \leqslant (2^\kappa - 1) + 1 = 2^\kappa$.

$\kappa = 2 \;\Rightarrow\; \nu = 1$ — учетверение информационного октета: $N = 1,\; K = 3$.

$\kappa = 3$: $\nu = 4$ инф-х бита и $\kappa + 1 = 4$ контрольных $\Rightarrow N = 1$ и $K = 1 \Rightarrow\; M = 2$: первый (информационный) октет делим на две инф-е тетрады, второй — на две контрольные тетрады (один блок алгоритма включает два блока Хэмминга с контролем двойной ошибки — в блоке из $M = 2$ октетов исправляется от одной до двух ошибок, обнаруживается от двух до четырёх).

$\kappa \geqslant 4$: выбираем $K \;\Rightarrow\; \kappa \leqslant 8K - 1 \;\Rightarrow\; M(\kappa) \;\Rightarrow\; N = M - K$.

Размер блока $N \to M$ (октетов) II

1. Пусть $K = 1$, тогда максимально возможное число битов Хэмминга $\sup(\kappa) = 7$, общее число битов $\sup(\mu + 1) = 2^7 = 128$, октетов $\sup(M) = \frac{2^\kappa}{8} = 2^{\kappa-3} = 16$. Из них инф-х $\sup(N) = \sup(M) - K = 15$.
   1. $\kappa = 7$ (полный контрольный октет): $8 < M \leqslant 16 \;\Rightarrow\; 8 \leqslant N \leqslant 15$
   2. $\kappa = 6$: $M \leqslant 8 \Rightarrow N \leqslant 7$ $(4 \leqslant N \leqslant 7)$ — 1 лишний бит: в контрольный октет включаем две копии бита общей чётности k_0.
   3. $\kappa = 5$: $M \leqslant 4 \Rightarrow N \leqslant 3$ $(2 \leqslant N \leqslant 3)$ две новые контрольные группы, либо две копии существующих ($k_0 \times 2$, $k_0 \cup k_1$, $k_1 \cup k_2$ и т. п.).
   4. $\kappa = 4$: $M \leqslant 2 \Rightarrow N = 1$ — но при $N = 1$ и $M = 2$ лучше $2 \times (\kappa = 3)$; $\kappa = 4$ без контроля двойной чётности: $N = 2$ и $K = 1$ (2 блока Хэмминга по 8 бит, $2 \times \kappa$ собираются в контрольный октет).

Размер блока $N \to M$ (октетов) III

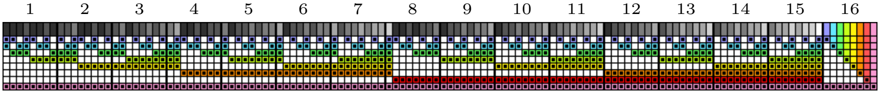
2. Пусть $K = 2$, тогда $\sup(\kappa) = 15$, $\sup(\mu + 1) = 2^{15} = 32\,768$ (бит), $\sup(M) = 2^{12} = 4096$ (октетов), и $\sup(N) = 4094$ (октета).
   Аналогично, два полных контрольных октета при $2048 < M \leqslant 4096$, то есть $2046 < N \leqslant 4094$ (в том числе $N = 2048 = 2^{11}$) инф-х октетах.
   При $15 < N \leqslant 2046$ потребуется от $\kappa = 9$ до $\kappa = 14$ контрольных бит (и блок длинный, и контрольные два октета неполные).

   $K > 2$ (и, соответственно, $N > 4094$) используется крайне редко.

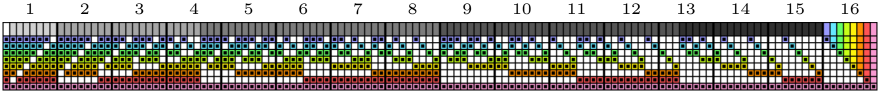
Код Хэмминга для $M = 16$ октетов

Некоторые из вариантов кода Хэмминга с контролем двойной ошибки для $N = 15$ информационных октетов и одного полного контрольного (всего $M = 16$).

Сортировка информационных битов по синдрому:



Сортировка информационных битов по количеству КС, затем по синдрому:



Яркость серого цвета вверху показывает количество контрольных сумм, в который входит информационный бит.

Полиномиальные коды

Минимальная единица передачи — **символ** (элемент некоторого поля).

Каждый символ может быть искажён при передаче независимо от других (заменой $a \to \widetilde{a}$, но без перестановок, выпадений и вставок).

Информационный полином (ν символов) степени $\nu - 1$
$a(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_{\nu-1} x^{\nu-1}$.

Порождающий полином $g(x)$ степени κ ($\kappa + 1$ символов, обычно $g_\kappa = 1$).

Кодовое слово C ($\mu = \nu + \kappa$ символов) степени $\mu - 1$ делится на $g(x)$:

- несистематический код $C(x) = a(x) \cdot g(x)$;
- **систематический код** (ν информационных и κ проверочных символов):

$C(x) = a(x) \cdot x^\kappa - r(x)$, где $r(x) = a(x) \cdot x^\kappa \bmod g(x)$, $\deg(r) < \deg(g) = \kappa$
$r(x)$ рассчитывается без деления, по табличным $x^i \bmod g(x)$, $\kappa \leqslant i < \mu$

Полученное слово $C(x) + err(x) = \widetilde{C}(x) = g(x) \cdot p(x) + r(x)$,
$r(x) = err(x) \bmod g(x)$ — **синдром**, $r(x) \neq 0$ — сбой
(но для Рида—Соломона синдромом называется другой многочлен).

Циклические коды

— циклическая перестановка символов в кодовом слове дает другое допустимое слово того же кода: $C_1 = (c_0, c_1, \dots c_{\mu-1})$, $C_2 = (c_{\mu-1}, c_0, c_1, \dots c_{\mu-2})$

Таким образом, $C_2 = x \cdot C_1 - c_{\mu-1} \cdot (x^\mu - 1) \implies c_{\mu-1} \cdot (x^\mu - 1) = x \cdot C_1 - C_2$.

Полиномиальный код циклический $\Leftrightarrow$ $x^\mu - 1$ делится на $g(x)$.

Проверочный многочлен $h(x) = \frac{x^\mu - 1}{g(x)}$ используется для извлечения информации из несистематического кода:

$C(x)h(x) = a(x)g(x)h(x) = a(x) \cdot (x^\mu - 1) = a(x) \cdot x^\mu - a(x)$

$\mu = \nu + \kappa > \deg(a) = \nu - 1$ — две разнесённых копии коэф-тов $+a$ и $-a$.

$\diamondsuit$ символы из $\mathbb{Z}_5$, $a(x) = 3x + 1$, $\mu = 4$,

$g(x) = x^2 + 4x + 3 \implies h(x) = \frac{x^4 - 1}{g(x)} = \frac{x^4 + 4}{g(x)} = x^2 + x + 3$

$C(x) = a(x)g(x) = 3x^3 + 3x^2 + 3x + 3$

$C(x)h(x) = (3x^3 + 3x^2 + 3x + 3)(x^2 + x + 3) = 3x^5 + x^4 + 2x + 4 = 3x^5 + x^4 - 3x - 1$

Полиномиальный код Хэмминга

Над $\mathrm{GF}(2)$, $g(x)$ — делитель $x^\mu - 1$ (код циклический) степени κ (причём $\mu = 2^\kappa - 1$), не имеет корней в $\mathrm{GF}(2)$ и делителей.
В $\mathrm{GF}(2)$ (то есть $\mathbb{Z}_2$) верно $(-1) = 1$, то есть сложение = вычитанию.

$\kappa = 1, \mu = 1$: мн-н $x^1 - 1$, то есть $x + 1 = 1 \cdot (x+1) \Rightarrow g(x) = 1$

$\kappa = 2, \mu = 3$: $x^3 - 1 = x^3 + 1 = (x+1)(x^2+x+1) \Rightarrow g(x) = x^2 + x + 1$, $a(x) = a_0$
сист-й и несист-й коды совпадают: $C(x) = a_0 x^2 + a_0 x + a_0 \sim (a_0, a_0, a_0)$

a_0	k_2	k_1
×	×	
×		×

синдром: $\Big(k_2(\widetilde{a}), k_1(\widetilde{a})\Big) \oplus \Big(\widetilde{k_2}, \widetilde{k_1}\Big)$

$\kappa = 3, \mu = 7$: $x^7 + 1 = (x+1)(x^3+x+1)(x^3+x^2+1) \Rightarrow \left[\begin{aligned} g(x) &= x^3 + x + 1 \\ g(x) &= x^3 + x^2 + 1 \end{aligned}\right.$

$a(x) = a_3 x^3 + a_2 x^2 + a_1 x + a_0$, пусть $g(x) = x^2 + x + 1$, тогда сист-й код:
$C(x) = a_3 x^6 + a_2 x^5 + a_1 x^4 + a_0 x^3 + (a_1 + a_2 + a_3) x^2 + (a_0 + a_1 + a_2) x + (a_0 + a_2 + a_3)$

a_3	a_2	a_1	a_0	k_3	k_2	k_1
×	×	×		×		
	×	×	×		×	
×	×		×			×

Полиномиальный код Рида-Соломона

Корни $g(x)$ Рида—Соломона лежат в том же поле, над каким и строится код
Пусть β — элемент поля $\mathrm{GF}(q)$ порядка μ (обычно — примитивный элемент).
Тогда порождающий полином кода Рида—Соломона:
$g(x) = (x - \beta^{l_0})(x - \beta^{l_0+1}) \dots (x - \beta^{l_0+\kappa-1}),$ $\deg(g) = \kappa.$
где l_0 — некоторое целое число. Обычно $l_0 = 1$.

Длина полученного кода μ, минимальное расстояние δ,
проверочных символов $\kappa = \delta - 1 = \deg(g)$,
информационных символов $\nu = \mu - \kappa = \mu - \delta + 1$.

Если β — примитивный элемент $\mathrm{GF}(q)$, то $\mu = q - 1$. Количество проверочных κ однозначно определяет $g(x)$.
Исправляется до $\kappa/2$ ошибок.

Декодирование систематического кода Рида-Соломона

- Остаток $e(x) = C(x) \mod g(x)$ — можно не вычислять.
- Синдром $S(x) : s_i = e(\beta^{i+1}) = C(\beta^{i+1})$.
- Локатор ошибки — $X_i = \beta^\ell$ для x^ℓ.
  Многочлен локаторов $L(x) = (1 - xX_1)(1 - xX_2)\dots(1 - xX_u)$
- Многочлен ошибок $W(x)$ — степень не превышает $u - 1$,
  где u — количество ошибок ($u \leqslant \kappa/2$),
  причём $L(x) \cdot S(x) = W(x) \mod x^\kappa$.
- Значения ошибок $Y_i = \frac{W(X_i^{-1})}{L'(X_i^{-1})}$ (коррекция: $C(c) = \widetilde{C}(x) + \sum Y_i \cdot x^{\ell_i}$).

Код Рида-Соломона над $\mathrm{GF}(5)$

Символы: $\mathrm{GF}(5) = \mathbb{Z}_5$ — вычеты по модулю 5, $(-1) = 4 \neq 1$, поэтому формулы частично отличаются от $\mathrm{GF}(2^s)$, где всегда $(-1) = 1$.

Максимальная длина кода $\mu = 4$ (количество ненулевых элементов поля), $\beta = 2$ — примитивный: $\beta^2 = 4, \beta^3 = 3, \beta^4 = 1$ (все ненулевые элементы).

Возможны многочлены: $g(x) = (x-2)$ ($\kappa = 1$, исправляет $\left\lfloor \frac{\kappa}{2} \right\rfloor = 0$ ошибок),
$g(x) = (x-2)(x-4)$ ($\kappa = 2$, исправляет $\left\lfloor \frac{\kappa}{2} \right\rfloor = 1$ ошибку),
$g(x) = (x-2)(x-4)(x-3)$ ($\kappa = 3$, исправляет $\left\lfloor \frac{\kappa}{2} \right\rfloor = 1$ ошибку).

$u \leqslant 1$: м-н локаторов одной ошибки $L(x) = 1 - xX_1 = 1 - x\gamma$ (производная $L'(x) = -\gamma$), м-н ошибок $W(x) = c$ нулевой степени (то есть $Y_i = \frac{c}{-\gamma}$).

При $\mu = 4$ код циклический: $\beta^\mu = 1 \Rightarrow (\beta^\ell)^\mu = 1 \Rightarrow$ все корни $g(x)$ также являются корнями $x^\mu - 1$: $x^\mu - 1 = x^4 - 1 = x^4 + 4 = g(x)(x^2 + x + 3)$.

Выбираем $g(x) = (x-2)(x-4) = x^2 + 4x + 3$, $\kappa = 2$ контрольных символа, $\nu = \mu - \kappa = 2$ информационных символа.

Систематический код Рида-Соломона (3, 1)

Сообщение: $(3, 1) \sim a(x) = 3x + 1$, — коэффициенты записываем наоборот, чтобы в систематическом коде информ-е символы располагались в начале.
$g(x) = (x - 2)(x - 4) = x^2 + 4x + 3$

Систематический код: $C(x) = a(x) \cdot x^{\kappa} - r(x)$

Вычисление остатка: $r(x) = a(x) \cdot x^{\kappa} \mod g(x) = 3x^3 + x^2 \mod g(x) =$
$= 3 \cdot \left(x^3 \mod g(x)\right) + 1 \cdot \left(x^2 \mod g(x)\right)$, где $x^{\kappa+i} \mod g(x)$ — табличные.

Здесь: $x^{\kappa+0} = x^2 = (x^2 + 4x + 3) + x + 2 \equiv x + 2$,
$x^{\kappa+1} = x^3 = x \cdot x^2 \equiv x(x + 2) = x^2 + 2x \equiv 3x + 2$.

То есть $r(x) = 3(3x + 2) + (x + 2) = (4x + 1) + (x + 2) = 3$.

$C(x) = 3x^3 + x^2 - 3 = 3x^3 + x^2 + 2 = g(x) \cdot (3x + 4)$

Код: $C(x) \sim (\underbrace{3, 1}_{\nu}, \underbrace{0, 2}_{\kappa})$ — первые (старшие) ν символов информационные.

$C(x)$ делится на $g(x)$ $\Leftrightarrow$ $C(2) = C(4) = 0$ $\Leftrightarrow$ синдром $S(x) = 0$.

Коррекция ошибок

Ошибка: $(3,1,0,\mathbf{2}) \to (3,1,0,\mathbf{0})$ $\widetilde{C}(x) = 3x^3 + x^2 = 2^3 \cdot x^3 + x^2$

Приняли $\widetilde{C}(x) = C(x) + e(x)$. Ошибка $e(x)$ неизвестна $e(x) = -2 = 3$

Найдём коэффициенты синдрома (степень $\kappa - 1 = 1$):

$s_0 = \widetilde{C}(2) = 3 \cdot 2^3 + 2^2 = 4 + 4 = 3$

$s_1 = \widetilde{C}(4) = 3 \cdot 4^3 + 4^2 = 2 + 1 = 3$

Синдром $S(x) = 3x + 3 \neq 0$ — ошибка есть, то есть $\widetilde{C}(x) \neq C(x)$.

Найдём параметры мн-в локаторов $L(x) = 1 - \gamma x$ и ошибок $W(x) = c$:

$$(3x+3)(1-\gamma x) = c \mod x^2$$

получаем систему уравнений: $\begin{cases} 3 - 3\gamma = 0 & \text{коэффициенты при } x \\ 3 \quad = c & \text{свободные члены} \end{cases}$ откуда

$\gamma = 1 = 2^0$ и $c = 3$. Решение системы — самая сложная часть.

Место ошибки: x^0 (так как $\gamma = 2^0$) — испорчен контрольный символ,

коррекция $Y_1 = \frac{3}{-1} = -3 = 2$: $C(x) = \widetilde{C}(x) + 2 \cdot x^0 = 3x^3 + x^2 + 2$.

Несистематический код Рида-Соломона

Тот же порождающий многочлен $g(x) = (x-2)(x-4) = x^2 + 4x + 3,$
то же сообщение $(3,1) \sim a(x) = 3x + 1.$

Несистематический код: $C(x) = a(x)g(x) = 3x^3 + 3x^2 + 3x + 3 \sim (3,3,3,3)$
тоже $\mu = 4$ символа, но нельзя отделить инф-е от контрольных.

Восстановление сообщения: $C(x)h(x) = a(x)g(x)h(x) = a(x)(x^\mu - 1),$
где $h(x)$ — проверочный многочлен $h(x) = \frac{x^\mu - 1}{g(x)} = x^2 + x + 3.$

$a(x)(x^\mu - 1) = (ax+b)(x^4-1) = ax^5 + bx^4 - ax - b \sim (a, b, 0, 0, -a, -b)$
$\nu + \mu - 1 = \nu + (\nu + \kappa) - 1$ степени; $2\nu + \kappa$ символов, из них κ нулей.

$C(x)h(x) = (3x^3 + 3x^2 + 3x + 3)(x^2 + x + 3) = 3x^5 + x^4 + 2x + 4$
$\sim (3,1,0,0,2,4) = (3,1,0,0,-3,-1)$

Синдром и коррекция — аналогично систематическому коду.

ДПФ Рида-Соломона

То же сообщение $(3,1) \sim a(x) = 3x + 1$ (коэффициенты записываем в обратном порядке, как и ранее, но здесь это неудобно).

$\beta^{-4} = 1, \beta^{-3} = 2, \beta^{-2} = 4, \beta^{-1} = 3,\ \beta^0 = 1, \beta^1 = 2, \beta^2 = 4, \beta^3 = 3, \beta^4 = 1$

Кодирование:
$$\begin{aligned}
c_0 &= a(\beta^0) = a(1) = 3\cdot 1 + 1 && = 4\\
c_1 &= a(\beta^1) = a(2) = 3\cdot 2 + 1 = 1 + 1 &&= 2\\
c_2 &= a(\beta^2) = a(4) = 3\cdot 4 + 1 = 2 + 1 &&= 3\\
c_3 &= a(\beta^3) = a(3) = 3\cdot 3 + 1 = 4 + 1 &&= 0
\end{aligned}$$

$(0,3,2,4) \sim C(x) = 3x^2 + 2x + 4$

Восстановление:
$$\begin{aligned}
a_0 &= \frac{C(\beta^0)}{\mu} = \frac{3\cdot 2^0 + 2\cdot 2^0 + 4}{4} = \frac{3+2+4}{4} = \frac{4}{4} = 1\\
a_1 &= \frac{C(\beta^{-1})}{\mu} = \frac{3\cdot 2^{-2} + 2\cdot 2^{-1} + 4}{4} = \frac{2+1+4}{4} = \frac{2}{4} = 3\\
a_2 &= \frac{C(\beta^{-2})}{\mu} = \frac{3\cdot 2^{-4} + 2\cdot 2^{-2} + 4}{4} = \frac{3+3+4}{4} = \frac{0}{4} = 0\\
a_3 &= \frac{C(\beta^{-3})}{\mu} = \frac{3\cdot 2^{-6} + 2\cdot 2^{-3} + 4}{4} = \frac{2+4+4}{4} = \frac{0}{4} = 0
\end{aligned}$$

Сообщение (2, 1)

Сообщение: $(2,1) \sim a(x) = 2x+1$, $\qquad g(x) = (x-2)(x-4) = x^2+4x+3$

$r(x) = 2(3x+2)+(x+2) = (x+4)+(x+2) = 2x+1.$

$C(x) = 2x^3+x^2-(2x+1) = 2x^3+x^2+3x+4 = g(x)(2x+3) \sim (2,1,3,4)$

Ошибка №1: $(2,1,\mathbf{3},4) \to (2,1,\mathbf{0},4)$ $\qquad \widetilde{C}(x) = 2x^3+x^2+4$

Синдром: $s_0 = \widetilde{C}(2) = 4$, $s_1 = \widetilde{C}(4) = 3$: $\qquad S(x) = 3x+4 \neq 0$

Из $(3x+4)(1-\gamma x) = c \mod x^2$ находим: $\gamma = \frac{3}{4} = 2$ и $4 = c$.

Место ошибки: x^1 (так как $\gamma = 2^1$) — испорчен контрольный символ, коррекция

$Y_1 = \frac{4}{-2} = -2 = 3$: $\qquad C(x) = \widetilde{C}(x)+3x = 2x^3+x^2+3x+4.$

Ошибка №2: $(\mathbf{2},1,3,4) \to (\mathbf{4},1,3,4)$ $\qquad \widetilde{C}(x) = 4x^3+x^2+3x+4$

Синдром: $s_0 = \widetilde{C}(2) = 1$, $s_1 = \widetilde{C}(4) = 3$: $\qquad S(x) = 3x+1 \neq 0$ — ошибка.

Из $(3x+1)(1-\gamma x) = c \mod x^2$ находим: $\qquad \gamma = 3 = 2^3$, $\qquad c = 1$,

коррекция $Y_1 = \frac{1}{-3} = -2 = 3$: $\qquad C(x) = \widetilde{C}(x)+3x^3 = 2x^3+x^2+3x+4.$

Тот же порождающий многочлен $g(x) = (x-2)(x-4) = x^2 + 4x + 3,$
то же сообщение $(2,1) \sim a(x) = 2x + 1.$

$\beta^{-4} = 1, \beta^{-3} = 2, \beta^{-2} = 4, \beta^{-1} = 3,\ \beta^0 = 1, \beta^1 = 2, \beta^2 = 4, \beta^3 = 3, \beta^4 = 1$

Несистематический код: $C(x) = a(x)g(x) = 2x^3 + 4x^2 + 3 \sim (2,4,0,3)$

$$C(x)h(x) = (2x^3 + 4x^2 + 3)(x^2 + x + 3) = 2x^5 + x^4 + 3x + 4$$
$$\sim (2,1,0,0,3,4) = (2,1,0,0,-2,-1)$$

ДПФ Рида—Соломона: $(2,4,0,3) \sim C(x) = 2x^3 + 4x^2 + 3$ совп. случайно

$$\begin{aligned} c_0 &= a(1) = 2\cdot 1 + 1 &&= 3 \\ c_1 &= a(2) = 2\cdot 2 + 1 = 4 + 1 &&= 0 \\ c_2 &= a(4) = 2\cdot 4 + 1 = 3 + 1 &&= 4 \\ c_3 &= a(3) = 2\cdot 3 + 1 = 1 + 1 &&= 2 \end{aligned}$$

$$\begin{aligned} a_0 &= \frac{2\cdot 2^0 + 4\cdot 2^0 + 3}{4} = \frac{2+4+3}{4} = \frac{4}{4} = 1 \\ a_1 &= \frac{2\cdot 2^{-3} + 4\cdot 2^{-2} + 3}{4} = \frac{4+1+3}{4} = \frac{3}{4} = 2 \\ a_2 &= \frac{2\cdot 2^{-6} + 4\cdot 2^{-4} + 3}{4} = \frac{3+4+3}{4} = \frac{0}{4} = 0 \\ a_3 &= \frac{2\cdot 2^{-9} + 4\cdot 2^{-6} + 3}{4} = \frac{1+1+3}{4} = \frac{0}{4} = 0 \end{aligned}$$

Спасибо за внимание!

МИЭТ
www.miet.ru

Александра Игоревна Кононова
illinc@mail.ru
gitlab.com/illinc/raspisanie