Лекция №13

Криптосистема МакЭлиса. LDPC коды.

I Криптосистема МакЭлиса

Предложена МакЭлисом в 1978г. (основано на кодах Гоппы)

Сегодня: модификация Нидеррайтера + код Гоппы.

- **KeyGen** $(pp = n, k, t)$

  1. C – случайный лин. код $[n, k, d=2t+1]$ над $\mathbb{F}_2$ с эффективным алг-мом декодирования (C – код Гоппы)
  2. $H \in \mathbb{F}_2^{n-k \times n}$ – проверочная матрица
  3. Сгенерировать $S \in \mathbb{F}_2^{n-k \times n-k}$ – случ. несингулярная матрица
  4. Сгенерировать $P \in \mathbb{F}_2^{n \times n}$ – случ. матрица перестановки
  5. $pk = H' = S \cdot H \cdot P$

     $sk = (S, H, P)$

- **Enc** $(m \in \{0,1\}^n,\ wt(m) = t,\ pk)$

  1. $c = H' \cdot m \in \mathbb{F}_2^{n-k}$

- **Dec** $(c,\ sk = (S, H, P))$

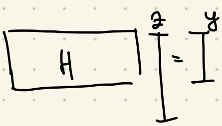
  1. $y = S^{-1} \cdot c \in \mathbb{F}_2^{n-k}$ // $c = H'm = S \cdot H \cdot P \cdot m$

     $S^{-1} c = H \cdot Pm$

  2. Используя алг-м Гаусса, найти $z \in \mathbb{F}_2^n$ т.ч.

     $H \cdot z = y = H \cdot P \cdot m$

  3. Decode(z) – алг-м декодирования для C (например, алг-м декодирования кода Гоппы)



Получим m' – результат декодирования

4. $m = P^t \cdot m' \qquad (P^t \cdot P = P \cdot P^t = Id)$

Корректность

Пусть $c = H' \cdot m$ – корректно сформированный шифр-текст

Тогда, алг-м Dec вычисляет:

1. $y = H \cdot P \cdot m$

2. $z: \quad y = H \cdot z$

$$H \cdot Pm = H \cdot z$$
$$\Downarrow$$
$$H(\underbrace{P \cdot m - z}) = 0$$
$$\Downarrow$$

$P \cdot m - z$ – кодовое слово из C

$wt(Pm) = wt(m) = t$

$\uparrow$

матрица перестановки P не меняет вес $\Rightarrow$

$\Rightarrow \Delta(z, \underset{\in C}{c}) = t \Rightarrow Decode(z)$ вернёт $P \cdot m = m'$

4. $P^t \cdot m' = \underbrace{P^t \cdot P}_{Id} \cdot m = m.$

Безопасность

основана на трудности задачи декодирования случ. лин. кода (с проверочной матрицей H')

Полагаем, что код, полученный из H' ведёт себя как случайный $[n, k]$-код.

Современные пар-ры: $n = 6960$

$k = 5413$

$t = 119$

II Коды с малой плотностью проверки на чётность.
(Low-density parity-check codes, LDPC)

802.3 Ethernet
802.11 Wireless Lan

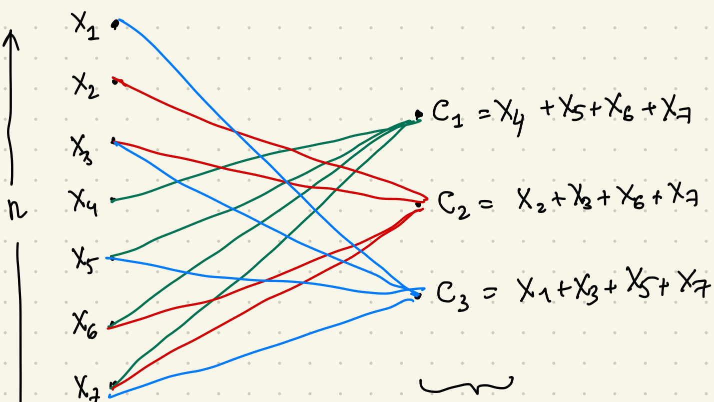
II.1. Мотивация: улучшить алг-м декодирования, а не min. расстояние.

LDPC код – код на графах

$\forall$ лин. код может быть представлен в виде двудольного графа, т.е. графа, мн-во вершин которого можно разбить на два $\not\cap$ мн-ва U, V, т.ч. рёбра графа соединяют вершины из U только с вершинами из V.

Пример $[7,4,3]_2$ – код Хэмминга

$$H = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \end{bmatrix} = \begin{bmatrix} x_4 + x_5 + x_6 + x_7 \\ x_2 + x_3 + x_6 + x_7 \\ x_1 + x_3 + x_5 + x_7 \end{bmatrix}$$



Вершины переменных

Проверочные вершины

Коды LDPC соответствуют "разреженным" (sparse) графам, т.е. графам с малым кол-вом рёбер; экв-но, проверочная матрица LDPC кода содержит число 1-ц в каждой строке $\ll n$, число 1-ц в каждом столбце $\ll n-k$.

Код LDPC называется регулярным, если его граф является регулярным, т.е. степени вершин x_i равны м/д собой и степени c_j равны м/д собой. (Код Хэмминга регулярным не является)

II.2 Жёсткое декодирование LDPC кодов (метод вероятностного итеративного декодирования)

$n = 8$

$$H = \begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}$$

полученное слово

$$y = [0\ 1\ 0\ 0\ 1\ 0\ 0\ 1]$$

Раунд 2	Раунд 1	
[0, 0, 0]	$\begin{array}{ccc} c_1 & c_2 & c_3 \end{array}$ [1, 1, 0]	$x_1 \cdot \boxed{0}$
[1, 1, 1]	$\begin{array}{ccc} c_4 & c_5 & c_6 \end{array}$ [1, 0, 1]	$x_2 \cdot \boxed{1}$
[0, 0, 0]	[1, 0, 0]	$x_3 \cdot \boxed{0}$
[0, 0, 0]	[1, 0, 0]	$x_4 \cdot \boxed{0}$
[0, 0, 0]	[0, 0, 0]	$x_5 \cdot \boxed{1}$
[0, 0, 0]	[1, 0, 0]	$x_6 \cdot \boxed{0}$
[0, 0, 0]	[0, 1, 0]	$x_7 \cdot \boxed{0}$
[1 1 1]	[1, 0, 1]	$x_8 \cdot \boxed{1}$

ошибка (!)

	Раунд 1
$\cdot\ c_1 = \overset{0}{x_1} + \overset{0}{x_4} + \overset{1}{x_5} + \overset{0}{x_6}$	$\overset{1}{x_1} + \overset{1}{x_4} + \overset{0}{x_5} + \overset{1}{x_6}$
$\cdot\ c_2 = \overset{0}{x_1} + \overset{0}{x_3} + \overset{1}{x_5} + \overset{0}{x_6}$	$\overset{1}{x_1} + \overset{1}{x_3} + \overset{0}{x_5} + \overset{1}{x_6}$
$\cdot\ c_3 = \overset{0}{x_1} + \overset{0}{x_3} + \overset{0}{x_4} + \overset{0}{x_6}$	$\overset{0}{x_1} + \overset{0}{x_3} + \overset{0}{x_4} + \overset{0}{x_6}$
$\cdot\ c_4 = \overset{1}{x_2} + \overset{0}{x_4} + \overset{0}{x_7} + \overset{1}{x_8}$	$\overset{1}{x_2} + \overset{0}{x_4} + \overset{0}{x_7} + \overset{1}{x_8}$
$\cdot\ c_5 = \overset{1}{x_2} + \overset{1}{x_5} + \overset{0}{x_7} + \overset{1}{x_8}$	$\overset{0}{x_2} + \overset{0}{x_5} + \overset{1}{x_7} + \overset{0}{x_8}$
$\cdot\ c_6 = \overset{1}{x_2} + \overset{0}{x_3} + \overset{0}{x_7} + \overset{1}{x_8}$	$\overset{1}{x_2} + \overset{0}{x_3} + \overset{0}{x_7} + \overset{1}{x_8}$

Алгоритм

I Раунд 0:

1. Вершины x_i получают значения y_i
   x_i посылают y_i смежным вершинам c_j

II Раунд i:

1. Для всех c:

   Вершина c посылает смежному x сумму всех полученных y_i mod 2 за исключением бита, полученного от самого x (экв. c посылает x полученный от x бит, если $\sum x_j = 0$, 1 иначе)

2. Для всех x:

   Вершина x посылает смежному c бит b, если x получил $\underline{b}$ от всех вершин, кроме c, иначе (если хотя бы одно значение не совпадает), x посылает полученный бит от y.

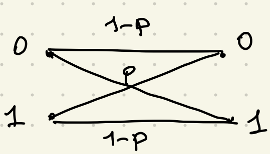
III Повторяем шаг II пока все условия не будут удовлетворены.

Замечание Можно использовать мажоритарное декодирование: вершина x решает об изменении бита $\boxed{\ }$ с помощью мажоритарного голосования

Пример $(1,0) \overset{x}{\cdot} \boxed{0} \Rightarrow Maj(1,0,0)=0$

III.3 Анализ LDPC кодов состоит в вычислении вероятности ошибки, т.е. передачи неверного бита от x к c и обратно в раунде i.

Модель канала коммуникации: бинарный симметричный канал



$p \in [0,1]$ - вероятность ошибки

Положим, исходное кодовое слово было нулевым ($x=0$), т.е. "ошибка" - передали бит 1.

Вершины X_i получат 0 с в-ю $1-p$

1 — || — p

Обозначим за p_i — вер-ть передачи от x к c "1" в раунде i

$p_0 = p$

q_i — —— || —— с к ~~x~~ "1" —— || ——

Положим, $\deg(x)=d$, $\deg(c)=e$ (граф регулярный)

1) Выразим p_{i+1} через p_i (для фикс. x и c)

$Pr[x \xrightarrow{\text{"1"}} c]$: 1. $Pr[y=1 \wedge$ хотя бы одно значение, полученное x от всех смежных вершин c, равно 1 $]$

$= p \cdot (1 - Pr[\underbrace{\text{все смежные к } x}_{\text{всего } d-1}$ вершины отправили "0"$])$

$= p \cdot (1-(1-q_i)^{d-1})$

2. $Pr[y=0 \wedge$ все значения, полученные x от всех смежных вершин, кроме c, $=$ "1"$]$

$= (1-p) \cdot (q_i)^{d-1}$

$$p_{i+1} = p\left(1-(1-q_i)^{d-1}\right) + (1-p)\cdot q_i^{d-1} \quad (1)$$

Оценим q_i для фикс. c и x

$q_i = 1 \iff \sum x_j = 1 \bmod 2$

x_j-смежные с c

и $x_j \neq x$

Лемма] X_j - независ. случ. величины $\in \{0,1\}$, т.ч. $Pr[X_j = 1] = p$

Тогда $Pr\left[\sum_{j=1}^{\ell-1} X_j \bmod 2 = 1 \right] = \dfrac{1 - (1-2p)^{e-1}}{2}$

(док-во с помощью мат. индукции)

(1) + Лемма

$$p_{i+1} = p\left(1 - \left(\frac{1 + (1-2p_i)^{e-1}}{2}\right)\right)^{d-1} + (1-p)\left(\frac{1-(1-2p_i)^{e-1}}{2}\right)^{d-1}$$