СПбГЭТУ «ЛЭТИ»
ПЕРВЫЙ ЭЛЕКТРОТЕХНИЧЕСКИЙ

А.Б. Левина

Теория информации и теория кодирования

Нелинейное кодирование

СПбГЭТУ «ЛЭТИ», 2022 г.

1 ВВЕДЕНИЕ

На данной лекции мы познакомимся со вторым направлением в области теории кодирования – нелинейными кодами, рассмотрим методы построения, матрицы Адамара и коды Адамара, робастными надежными кодами.

2. НЕЛИНЕЙНЫЕ КОДЫ

На прошлой лекции мы познакомились с линейными кодами, и обсудили, что одним из их минусов является легкая взламываемость.

Для решения этой проблемы были разработаны нелинейные коды.

Определение: Назовем (n, M, d) кодом множества из M векторов длины n такое, что любые два вектора различаются по меньшей мере в d позициях и d является наибольшим числом.

Нелинейные коды применяются в криптографических устройствах, рассмотрим их способы построения.

3. МАТРИЦА АДАМАР И КОДЫ АДАМАРА

Введем определение матрица Адамара.

Определение: Матрица Адамара H – квадратная матрица размера $\mathrm{n} \times \mathrm{n}$, составленная из чисел 1 и -1, такая, что

$$\mathrm{H_n} \cdot \mathrm{H_n^T} = \mathrm{n} \cdot \mathrm{E_n}$$

где $\mathrm{E_n}$ - единичная матрица размера n.

Пример:

$$\begin{bmatrix} 1 & 1 \\ 1 & - \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 \\ 1 & - \end{bmatrix} = 2 \cdot \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Определение: Нормализованная матрица Адамара – матрица Адамара, все элементы первой строки и первого столбца которой равны 1.

Пример:

$$\mathrm{H_4} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & - & 1 & - \\ 1 & 1 & - & - \\ 1 & - & - & 1 \end{bmatrix}$$

Теорема 1. Если существует матрица Адамара H порядка n, то n равно 1 или 2 или делится на 4

Рассмотрим методы построения матрицы Адамара:

1. Матрицы Сильвестра (порядок 2^n)
2. Матрицы Пэйли (порядок $n = p + 1$, кратный 4)
   где p - простое число

Порядок 2^n, построение через рекуррентное соотношение:

$$H_{2n} = \begin{bmatrix} H_n & H_n \\ H_n & -H_n \end{bmatrix}$$

Примеры:

$$H_1 = [1]$$

$$H_2 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

$$H_4 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & - & 1 & - \\ 1 & 1 & - & - \\ 1 & - & - & 1 \end{bmatrix}$$

$$H_8 = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & - & 1 & - & 1 & - & 1 & - \\ 1 & 1 & - & - & 1 & 1 & - & - \\ 1 & - & - & 1 & 1 & - & - & 1 \\ 1 & 1 & 1 & 1 & - & - & - & - \\ 1 & - & 1 & - & - & 1 & - & 1 \\ 1 & 1 & - & - & - & - & 1 & 1 \\ 1 & - & - & 1 & - & 1 & 1 & - \end{bmatrix}$$

Введем определение важного математического понятия, квадратичный вычет.

Определение: Квадратичные вычеты по модули p – числа $1^2, 2^2, 3^2, \dots, (p-1)^2$ приведенные по модулю p, где p – простое нечетное число.

Теорема 2. Для нахождения вычетов достаточно рассмотреть квадраты $1^2, 2^2, 3^2, \dots, ((p-1)/2)^2 \pmod p$

Итого:

- $(p-1)/2$ квадратичных вычетов
- $(p-1)/2$ квадратичных невычетов
- 0 - ни вычет, ни невычет

Пример: $p = 7$

$$1^2 = 1$$
$$2^2 = 4$$
$$3^2 = 9 \equiv 2 \pmod 7$$

Вычеты: $1, 2, 4$

Невычеты: 3, 5, 6

Введем еще одно важное математическое понятие- символ Лежандра.

Символ Лежандра

$$\chi\left(\frac{a}{p}\right) = \begin{cases} 0, \text{ a делится на p} \\ 1, \text{ a } - \text{квадратичный вычет по модулю p} \\ -1, \text{ a } - \text{квадратичный невычет по модулю p} \end{cases}$$

где a - целое число, p - простое число.

Прежде чем переходить к матрицам и кодам Адамара нам надо познакомиться еще с двумя матрицами.

Матрица Джекобстола.

Матрица Джекобстола Q - квадратная матрица размера $p \times p$, строки и столбцы которой пронумерованы числами $0,1,\dots,p-1$, а элементы $q_{ij} = \chi\left(\frac{j-i}{p}\right)$

Пример: $p = 7$ $1^2 = 1,\ 2^2 = 4,\ 3^2 = 9 \equiv 2 \pmod 7$

Вычеты: 1, 2, 4

Невычеты: 3, 5, 6

Матрица Пэйли $H_n = \begin{bmatrix} 1 & \mathbf{1}_p \\ \mathbf{1}_p^T & Q_p - E_p \end{bmatrix}$

Пример:

$$\mathbf{Q_7} = \begin{bmatrix} \mathbf{0} & \mathbf{1} & \mathbf{1} & - & \mathbf{1} & - & - \\ - & \mathbf{0} & \mathbf{1} & \mathbf{1} & - & \mathbf{1} & - \\ - & - & \mathbf{0} & \mathbf{1} & \mathbf{1} & - & \mathbf{1} \\ \mathbf{1} & - & - & \mathbf{0} & \mathbf{1} & \mathbf{1} & - \\ - & \mathbf{1} & - & - & \mathbf{0} & \mathbf{1} & \mathbf{1} \\ \mathbf{1} & - & \mathbf{1} & - & - & \mathbf{0} & \mathbf{1} \\ \mathbf{1} & \mathbf{1} & - & \mathbf{1} & - & - & \mathbf{0} \end{bmatrix}$$

Коды Адамара

Двоичная матрица Адамара A - матрица Адамара, в которой все 1 заменены на 0, а -1 на 1.

$$H_8 = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & - & 1 & 1 & - & 1 & - & - \\ 1 & - & - & 1 & 1 & - & 1 & - \\ 1 & - & - & - & 1 & 1 & - & 1 \\ 1 & 1 & - & - & - & 1 & 1 & - \\ 1 & - & 1 & - & - & - & 1 & 1 \\ 1 & 1 & - & 1 & - & - & - & 1 \\ 1 & 1 & 1 & - & 1 & - & - & - \end{bmatrix}$$

$$A_8 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{bmatrix}$$

Расстояние Хэмминга = $n/2$

1. $(n-1, n, n/2)$: отбрасывается первый символ A_n

Пример: $A_8 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{bmatrix}$ $\mathbb{A}_8 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{bmatrix}$

2. $(n,\ 2n, n/2)$: к A_n снизу добавляется инвертированная A_n^-

Пример: $A_4 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix}$ $\mathbb{B}_4 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}$

3. $(n-1,\ 2n, n/2-1)$: отбрасывается первый символ $\mathbb{B}_n$

Пример: $\mathbb{B}_4 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}$ $\mathbb{C}_4 = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

Использование матриц Адамара

- Кодирование информации (коды, исправляющие ошибки, ECC)
- Шифрование
- Рентгеновские телескопы

Пример исправления ошибки

$$H_8 = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & - & 1 & - & 1 & - & 1 & - \\ 1 & 1 & - & - & 1 & 1 & - & - \\ 1 & - & - & 1 & 1 & - & - & 1 \\ 1 & 1 & 1 & 1 & - & - & - & - \\ 1 & - & 1 & - & - & 1 & - & 1 \\ 1 & 1 & - & - & - & - & 1 & 1 \\ 1 & - & - & 1 & - & 1 & 1 & - \end{bmatrix}$$

$$w^* = [1 \quad 0 \quad 1 \quad 1 \quad 1 \quad 0 \quad 1 \quad 0]$$
$$\overline{w^*} = 2 \cdot w - h_1 = [1 \quad -1 \quad 1 \quad 1 \quad 1 \quad -1 \quad 1 \quad -1]$$
$$s = \overline{w^*} \cdot H_8 = [2 \quad 6 \quad -2 \quad 2 \quad 2 \quad -2 \quad -2 \quad 2]$$
$$w = 2 \cdot h_1 - h_2 \bmod 3 = [1 \quad 0 \quad 1 \quad 0 \quad 1 \quad 0 \quad 1 \quad 0]$$

4. НАДЕЖНЫЕ КОДЫ

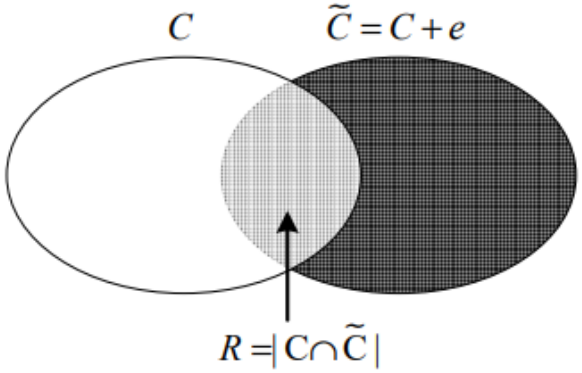
Введем определение надежного кода.

Определение: Код $C \subseteq GF(2^n)$ называется R-надежным кодом, если область пересечения кода C и любого его дополнения $C' = \{ x' \mid x' = x+e, x \in C, e \in GF(2^n), e \neq 0 \}$ ограничена сверху значением R: $R = \max_{e \in GF(2^n)} |\{ x \mid x \in C, x+e \in C \}|$, где x это кодовое слово, e это ошибка.

Пусть $M = |C|$ количеством кодовых слов в коде C. По определению R-надежного кода существует не более R кодовых слов, которые могут быть не обнаружены для любой фиксированной ошибки e. Таким образом, вероятность маскировки ошибки $Q(e)$ может быть определена как:

$$Q(e) = \frac{|\{x \mid x \in C, x + e \in C\}|}{M}$$

Равномерно надёжным называется код, мощность пересечения которого со всеми сдвигами равна R. Для линейных кодов $R = M$. Графическое изображение определения надёжного кода представлено на рисунке:



1. Надежный код не имеют необнаруживаемых ошибок.
2. Для любого R-надежного кода, наихудшая вероятность маскировки ошибки не больше, чем R/M при любой ошибке, если считать все появления кодовых слов надежного кода равновероятными.

Рассмотрим области использования надежных кодов:

- Проектирование устройств, усиленно защищённых от наиболее частых конфигураций встречаемых ошибок, сохраняя при этом возможность обнаружения любых ошибок.
- Коды с минимальным расстоянием не менее трёх могут быть использованы для коррекции ошибок.

Квадратичный систематический код (Код Кердока).

Пусть $x=(x_1, x_2,\ldots,x_{2s}, x_{2s+1})$, $s\geq 1$.

Вектор $x \in GF(2^{(2s+1)})$ принадлежит коду тогда и только тогда:

$$x_{2s+1}=x_1*x_2+x_3*x_4+\ldots+x_{2s-1}*x_{2s},$$

где * - это произведение в поле *GF(2)*.

Надежный повторяющийся код

Пусть $x=(x_1,x_2)$, $x_1, x_2 \in GF(2^r)$.

Надежный повторяющийся код *C* содержит все вектора $x \in GF(2^{2r})$, которые удовлетворяют соотношению $x_1^3=x_2$, где все вычисления происходят в поле $GF(2^r)$. Получившийся код является надежным кодом с длиной $n=2r$ и $M=2^r$.

Принцип построения

Пусть f представляет нелинейную функцию, которая отображает GF (2^k) в GF (2^r), где $k \geq r$.

Набор векторов, получающихся в результате конкатенации (y|f(y)), где y € GF (2^k) и f(y) € GF (2^r), образует надежный систематический код с параметрами $R=2^kP_f$, $n=k+r$ и $M=2^K$

Исправление ошибки

- Ошибка $e=(e_z|e_f)\neq 0$, $e_z, e_f \in GF(2^k)$, произошла на трёх необязательно последовательных тактах работы устройства и была обнаружена при декодировании.
- Обозначим элементы первого и второго искажённых слов как:

$$z_1 + e_z = \tilde{z}_1,$$
$$z_1^3 + e_f = \tilde{f}_1^3,$$
$$z_2 + e_z = \tilde{z}_2,$$
$$z_2^3 + e_f = \tilde{f}_2^3.$$

Сложив два полученных выражения и выполнив математические преобразования, получаем следующее выражение:

$$z_2^2 + z_2(\tilde{z}_1 + \tilde{z}_2) + (\tilde{z}_1^2 + \tilde{z}_2^2) + \frac{\tilde{f}_1 + \tilde{f}_2}{\tilde{z}_1 + \tilde{z}_2} = 0.$$

Повторив аналогичные действия для второго и третьего искажённых слов, получаем

$$z_2^2 + z_2(\tilde{z}_2 + \tilde{z}_3) + (\tilde{z}_2^2 + \tilde{z}_3^2) + \frac{\tilde{f}_2 + \tilde{f}_3}{\tilde{z}_2 + \tilde{z}_3} = 0.$$

$$z_2(\tilde{z}_1 + \tilde{z}_3) + (\tilde{z}_1^2 + \tilde{z}_3^2) + \frac{\tilde{f}_2 + \tilde{f}_3}{\tilde{z}_2 + \tilde{z}_3} + \frac{\tilde{f}_1 + \tilde{f}_2}{\tilde{z}_1 + \tilde{z}_2} = 0$$

$$z_1(\tilde{z}_2 + \tilde{z}_3) + (\tilde{z}_2^2 + \tilde{z}_3^2) + \frac{\tilde{f}_1 + \tilde{f}_3}{\tilde{z}_1 + \tilde{z}_3} + \frac{\tilde{f}_1 + \tilde{f}_2}{\tilde{z}_1 + \tilde{z}_2} = 0,$$
$$z_3(\tilde{z}_1 + \tilde{z}_2) + (\tilde{z}_1^2 + \tilde{z}_2^2) + \frac{\tilde{f}_2 + \tilde{f}_3}{\tilde{z}_2 + \tilde{z}_3} + \frac{\tilde{f}_1 + \tilde{f}_3}{\tilde{z}_1 + \tilde{z}_3} = 0.$$

Надежные/робастные коды представляю особый интерес в системах, где необходимо обеспечить высокий уровень безопасности.

5. ЗАКЛЮЧЕНИЕ

На лекции мы познакомились с нелинейными кодами, рассмотрели методы построения, матрицы Адамара и коды Адамара, робастными надежными кодами.