Ф. И. Соловьева

ВВЕДЕНИЕ В ТЕОРИЮ КОДИРОВАНИЯ

1 Линейные коды

Определение. *Линейным (или групповым) кодом* называется подмножество E^n, являющееся линейным подпространством (подгруппой) в E^n.

Пусть от отправителя в кодер поступило сообщение $u = (u_1, u_2, \ldots, u_k)$. Сформируем кодовое слово $x = (x_1, x_2, \ldots, x_n)$. Положим первую часть кодового слова состоящей из символов самого сообщения (называемых *информационными символами*): $x_1 = u_1, x_2 = u_2, \ldots, x_k = u_k$. Далее следуют $n - k$ символов называемых *проверочными* $x_{k+1}, \ldots, x_n$. Они выбираются таким образом, чтобы все кодовые слова удовлетворяли уравнению

$$H \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = Hx^\top = 0,$$

где H – матрица порядка $(n-k) \times n$, называемая *проверочной*. Пусть она имеет вид $H = [A_{n-k,k}|E_{n-k}]$, называемый *каноническим*, где $A_{n-k,k}$ – некоторая матрица порядка $(n-k) \times k$ из 0 и 1, E_{n-k} – единичная матрица порядка $n-k$. Все операции выполняются над полем Галуа $GF(2)$ характеристики 2.

Теорема 1. (О связи проверочной и порождающей матриц) *Если проверочная матрица линейного кода задана в каноническом виде $H = [A_{n-k,k}|E_{n-k}]$, то порождающая матрица этого кода имеет вид $G = [E_k| - A^\top_{n-k,k}]$. Верно обратное.*

Доказательство. Рассмотрим произвольное кодовое слово

$$x = (x_1, \ldots, x_k, x_{k+1}, \ldots, x_n),$$

где $x_{k+1}, \ldots, x_n$ – проверочные символы, а $x_1, \ldots, x_k$ – информационные, т.е. информационный блок имеет вид

$$u = (u_1, \ldots, u_k), \text{ где } x_1 = u_1, x_2 = u_2, \ldots, x_k = u_k,$$

что можно записать в матричном виде

$$\begin{pmatrix} x_1 \\ \vdots \\ x_k \end{pmatrix} = E_k \begin{pmatrix} u_1 \\ \vdots \\ u_k \end{pmatrix}. \tag{1}$$

Пусть

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1k} \\ a_{21} & a_{22} & \dots & a_{2k} \\ \dots & \dots & \dots & \dots \\ a_{n-k} & a_{n-k,2} & \dots & a_{n-k,k} \end{pmatrix}.$$

Тогда из определения проверочной матрицы имеем $Hx^\top = \mathbf{0}$, т.е.

$$a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{ik}x_k + x_{k+i} = 0$$

для любого $i = 1, \ldots, n-k$. Отсюда

$$x_{k+i} = -(a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{ik}x_k),\ i = 1, \ldots, n-k.$$

Таким образом,

$$\begin{pmatrix} x_{k+1} \\ \vdots \\ x_n \end{pmatrix} = -A_{n-k,k} \begin{pmatrix} x_1 \\ \vdots \\ x_k \end{pmatrix} = -A_{n-k,k} \begin{pmatrix} u_1 \\ \vdots \\ u_k \end{pmatrix}.$$

Из последнего и из (1) имеем

$$x^\top = \begin{pmatrix} E_k \\ -A_{n-k,k} \end{pmatrix} u^\top.$$

Транспонируя, получаем: $x = u\,G$, где $G = u[E_k| - A^\top_{n-k,k}]$.

2 Границы объемов кодов

Теорема 2. (Граница Хэмминга) *Для любого кода C длины n (не обязательно линейного) с кодовым расстоянием d выполняется неравенство*

$$|C| \le \frac{2^n}{\sum_{i=0}^{[(d-1)/2]} C_n^i}.$$

Доказательство. Обозначим $t = [(d-1)/2]$. Поскольку кодовое расстояние равно d, то шары

$$S_t^n(x) = \{y \mid y \in E^n, d(y,x) \le t]\}$$

радиуса t, описанные около кодовых слов x, не пересекаются, следовательно

$$|C| \times |S_t^n(x)| \le 2^n. \tag{2}$$

Вычисляя объем шара $S_t^n(x)$:

$$|S_t^n(x)| = C_n^0 + C_n^1 + \ldots + C^{[\frac{d-1}{2}]}$$

и подставляя его в (2), получаем требуемое. $\Diamond$

Определение. Код называется *совершенным* или *плотно-упакованным*, если

$$|C| = \frac{2^n}{\sum_{i=0}^{[(d-1)/2]} C_n^i},$$

т.е. имеет место плотная упаковка E^n шарами радиуса $[(d-1)/2]$.

Теорема 3. *Кодовое расстояние $[n, k, d]$-линейного кода равно минимальному из весов его ненулевых кодовых слов.*

Теорема 4. *Если H – проверочная матрица кода длины n, то код имеет кодовое расстояние d тогда и только тогда, когда любые $d-1$ столбцов матрицы H линейно независимы и найдутся d линейно зависимых столбцов.*

Доказательство. Необходимость.

Вектор x веса ω принадлежит коду тогда и только тогда, когда

$$Hx^\top = \mathbf{0}, \tag{3}$$

что эквивалентно линейной зависимости некоторых ω столбцов матрицы H. Обозначим i-ый столбец матрицы H через h_i, т.е.

$$H = [h_1, h_2, \ldots, h_n].$$

Через

$$supp(x) = \{i \mid x_i = 1\}$$

обозначим *носитель* вектора x. Отсюда и из (3) получаем

$$\sum_{i=1}^{n} h_i x_i = \mathbf{0},$$

откуда следует соотношение линейной зависимости $h_{i_1} + \ldots + h_{i_w} = \mathbf{0}$. По теореме 3 кодовое расстояние кода равно минимальному из весов его ненулевых кодовых слов. По условию теоремы код имеет кодовое расстояние d, откуда получаем линейную зависимость некоторой совокупности d столбцов матрицы H. Если существует $d-1$ линейно зависимых столбцов в матрице H, то найдется вектор веса $d-1$, принадлежащий коду C, противоречие.

Достаточность очевидна.

Непосредственным следствием теоремы 4 является следующая верхняя граница объема кода.

Теорема 5. (Граница Синглтона) *Для любого $[n, k, d]$-кода выполняется $n - k \geq d - 1$.*

Теорема 6. (Граница Синглтона для нелинейных q-значных кодов) *Для любого $(n, M, d)_q$-кода выполняется $\log_q M \leq n - d + 1$.*

Доказательство. Укорачивая $(n, M, d)_q$-код последовательно $d-1$ раз, получим код длины $n - d + 1$ с кодовым расстоянием по крайней мере 1 и мощности M.

Теорема 7. (Граница Плоткина) *При $n < 2d$ для любого (n, M, d)-кода C справедливо неравенство*

$$M \leq 2[d/(2d - n)],$$

где M –мощность кода C.

Доказательство. Вычислим двумя способами сумму

$$S = \sum_{u \in C} \sum_{v \in C, v \neq u} d(u, v)$$

для различных кодовых слов u и v из C. Поскольку при $u \neq v$ расстояние $d(u, v) \geq d$, то сумма не меньше чем $M(M-1)d$. С другой стороны, пусть A обозначает кодовую $M \times n$-матрицу, строками которой являются все кодовые слова. Предположим, что i-й столбец A содержит x_i нулей и $M - x_i$ единиц. Тогда вклад этого столбца в сумму S равен $2x_i(M - x_i)$. Суммируя по всем столбцам, получаем

$$S = \sum_{i=1}^{n} 2x_i(M - x_i).$$

При четном M максимум этого выражения достигается при $x_i = M/2$ для любого i, следовательно эта сумма не превышает $nM^2/2$, т.е. имеем:

$$M(M-1)\, d \leq nM^2/2,$$

отсюда

$$M \leq 2d/(2d - n).$$

Так как M – четно, то

$$M \leq 2[d/(2d - n)].$$

При нечетном M эта сумма не превышает $n(M^2 - 1)/2$ и следовательно

$$M \leq n/(2d - n) = 2d/(2d - n) - 1.$$

Отсюда, с учетом $[2x] \leq 2[x] + 1$ получаем

$$M \leq [2d/(2d - n)] - 1 \leq 2[d/(2d - n)].$$

Теорема 8. (Граница Варшамова-Гилберта) *Если выполняется неравенство*

$$1 + C_{n-1}^{1} + \cdots + C_{n-1}^{d-2} < 2^r,$$

то существует двоичный линейный код длины n с минимальным расстоянием по крайней мере d, имеющий не более чем r проверочных символов, т.е. $[n, k, d']$ - код, где $k \geq n - r$, $d' \geq d$.

Доказательство. Теорема будет доказана, если построим $(r \times n)$-матрицу H такую, что любые ее $d-1$ столбцов линейно независимы. Тогда применяя теорему 4, получим требуемое утверждение. В качестве первого столбца матрицы H возьмем любой ненулевой вектор длины r. Предположим, что выбрали i столбцов матрицы H так, что любые $d-1$ из них линейно независимы. Имеем не более

$$C_i^1 + \ldots + C_i^{d-2}$$

различных линейных комбинаций из этих i столбцов, содержащих $d-2$ или меньше столбцов. Если это число меньше чем $2^r - 1$ (числа всех ненулевых векторов длины r), то мы можем добавить еще один столбец, не равный ни одной из всех этих линейных комбинаций. При этом любые $d-1$ столбцов новой матрицы размера $r \times (i+1)$ по-прежнему остаются линейно независимы. Будем выполнять эту процедуру до тех пор, пока выполняется неравенство

$$C_i^1 + \ldots + C_i^{d-2} < 2^r - 1.$$

3 Код Хэмминга и его свойства

3.1 Определение кода Хэмминга

Для построения линейного кода Хэмминга с m проверками на четность, исправляющего одну ошибку, воспользуемся Теоремой 4: определим код посредством проверочной матрицы, столбцами которой являются **все** ненулевые векторы длины m. Очевидно, что любые два столбца этой матрицы линейно независимы и найдутся три линейно зависимых столбца, следовательно по теореме 4 код исправляет одну ошибку. Этот код называется *кодом Хэмминга*, далее будем его обозначать H^n.

Параметры кода Хэмминга:

$$[n = 2^m - 1, k = n - \log(n+1), d = 3],$$

$m = \log(n+1)$ (здесь и всюду далее $\log(\cdot)$ является двоичным логарифмом, если не оговорено особо).

Предложение 1. *Код Хэмминга H^n является совершенным кодом, исправляющим одну ошибку.*

Доказательство. Код H^n исправляет одну ошибку (по определению кода). По построению его мощность равна

$$|H^n| = 2^{n-m} = \frac{2^n}{n+1}.$$

Следовательно, он достигает границы Хэмминга (см. Теорему 2) и потому является совершенным.

Предложение 2. *Код Хэмминга единствен с точностью до изоморфизма.*

3.2 Примеры кодов Хэмминга длины 7

Рассмотрим три различных представления кода Хэмминга длины 7.

1) Код Хэмминга длины 7 задан в каноническом виде, см. [1], гл. 1, т. е. проверочная матрица имеет вид

$$H = \begin{pmatrix} 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{pmatrix}.$$

2) Код Хэмминга длины 7 задан в циклическом виде.

Определение. Линейный код C длины n называется *циклическим*, если для любого кодового слова $x = (x_1, x_2, \ldots, x_n)$ слово $(x_2, x_3, \ldots, x_n, x_1)$ принадлежит коду C.

Проверочная матрица кода Хэмминга длины 7 в циклическом представлении имеет вид:

$$H = \begin{pmatrix} 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 1 \end{pmatrix}.$$

3) Во многих случаях полезно определять код Хэмминга посредством проверочной матрицы, столбцы которой записаны в лексикографическом порядке возрастания двоичных представлений десятичных чисел:

$$H = \begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix}.$$

3.3 Декодирование кода Хэмминга

Код Хэмминга допускает самое простое декодирование. Рассмотрим представление кода Хэмминга H^n посредством проверочной матрицы, столбцы которой записаны в лексикографическом порядке:

$$H_m = [B(1), B(2), \ldots, B(n)],$$

здесь H_m – проверочная матрица кода H^n, $m = \log(n+1)$, $B(i)$ – двоичное представление числа i. Используя эту проверочную матрицу H_m, можно определить код Хэмминга следующим образом:

$$H^n = \{x = (x_1, \ldots, x_n) : x \in E^n, \ \sum_{i=1}^{n} B(i) \cdot x_i = \mathbf{0^m}\}.$$

Пусть в канале связи при передачи вектора x произошла одна ошибка в i-й координате и получен вектор y. Воспользовавшись тем, для любого кодового слова x кода H^n выполняется $Hx^\top = \mathbf{0^m}$, найдем синдром вектора y:

$$S = Hy^\top = Hx^\top + He_i^\top = He_i^\top = B(i),$$

он указывает столбец, номер которого i является номером ошибочной координаты вектора y. Здесь e_i – двоичный вектор длины n с единицей только в i-й координате.

Позднее будет рассмотрен пример кода Хэмминга над полем Галуа $GF(q)$ для произвольного $q = p^m$, где p – любое простое число.

Упражнения.

1. Найти число различных базисов в n-кубе E^n.
2. Найти число различных линейных двоичных кодов длины n размерности k.
3. Доказать предложение 2.

4 Способы построения новых кодов

4.1 Конструкция Плоткина

Нетрудно доказать справедливость следующего утверждения.

Предложение 3. *Для любых векторов x и y из E^n справедливо*

$$w(x+y) \geq w(x) - w(y).$$

Теорема 9. (М. Плоткин, 1960, см. [1].) *Пусть C и D – двоичные (n, M_1, d_1) и (n, M_2, d_2)-коды соответственно. Тогда множество*

$$C^{2n} = \{(x, x+y) : x \in C, y \in D\}$$

является $(n, M_1 \cdot M_2, d = min\{2d_1, d_2\})$-кодом.

Доказательство. Пусть

$$u = (x, x+y), \;\; v = (x', x'+y') -$$

произвольные кодовые слова кода C^{2n}, где $x, x' \in C$, $y, y' \in D$.
Если $y = y'$, то

$$d(u,v) = d((x,x),(x',x')) = 2d(x,x') \geq 2d_1.$$

Пусть $y \neq y'$, тогда, используя предложение 3, получим

$$d(u,v) = w(x-x') + w(x+y-x'-y') =$$

$$= w(x-x') + w((y-y') + (x-x')) \geq$$

$$\geq w(x-x') + w(y-y') - w(x-x') = w(y-y') = d_2.$$

Теорема доказана.

5 Коды Васильева

В 1959 г. С. Шапиро и Д. С. Злотник предположили, что не существует совершенных кодов, неэквивалентных коду Хэмминга. В 1962 г. Ю. Л. Васильев опроверг эту гипотезу, предложив богатый класс неэквивалентных совершенных двоичных кодов. Рассмотрим этот итеративный способ построения совершенных кодов.

Пусть $C^{(n-1)/2}$ – произвольный совершенный код длины $(n-1)/2 = 2^m - 1, m \geq 2$ и λ – произвольная функция из кода $C^{(n-1)/2}$ в множество $\{0,1\}$. Положим $|x| = x_1 + \ldots + x_{(n-1)/2} (\mathrm{mod} 2)$, где $x = (x_1, \ldots, x_{(n-1)/2}) \in E^{(n-1)/2}$.

Теорема 10. (Ю. Л. Васильев, 1962 г.) *Множество*

$$V^n = \{(x+y, |x| + \lambda(y), x) : x \in E^{(n-1)/2}, y \in C^{(n-1)/2}\} \tag{4}$$

является совершенным двоичным кодом длины n.

Доказательство. Проверим параметры построенного кода, а именно, его длину, мощность кода и кодовое расстояние.

1. Легко видеть, что n является длиной совершенного кода, т. е.

$$n = 2 \cdot (n-1)/2 + 1 = 2^{m+1} - 1.$$

2. Мощность кода

$$|V^n| = |E^{(n-1)/2}| \cdot |C^{(n-1)/2}| = 2^{(n-1)/2} \cdot 2^{(n-1)/2}/((n-1)/2+1) = 2^n/(n+1)$$

достигает границы Хэмминга.

3. Проверим, что кодовое расстояние равно 3. Рассмотрим два произвольных различных кодовых слова

$$u = (x + y, |x| + \lambda(y), x),$$
$$v = (x' + y', |x'| + \lambda(y'), x').$$

Возможны случаи.

3a. Если $y = y'$ и $x \neq x'$, то

$$d(u, v) = d((x, |x|, x), (x', |x'|, x')) \geq 3,$$

поскольку $x, x' \in E^{(n-1)/2}$ и $d(x, x') \geq 1$.

3b. Пусть $y \neq y'$ и $x = x'$. Векторы y, y' принадлежат коду $C^{(n-1)/2}$, следовательно $d(y, y') \geq 3$ и получаем

$$d(u, v) \geq d(y, y') \geq 3.$$

3c. Если $y \neq y'$ и $x \neq x'$, то для

$$d(x, x') \geq 1, 2, 3, \ldots$$

имеем

$$d(x + y, x' + y') \geq 2, 1, 0, \ldots$$

соответственно. Складывая эти расстояния, получаем

$$d(u, v) \geq 3.$$

Теорема доказана.

Код (4) будем далее называть *кодом Васильева.*

Следствие 1. *При $\lambda \equiv 0$ и $C^{(n-1)/2} = H^{(n-1)/2}$, конструкция Васильева дает код Хэмминга длины n:*

$$H^n = \{(x + y, |x|, x) : x \in E^{(n-1)/2}, y \in H^{(n-1)/2}\}.$$

Следствие 2. *Если $\lambda(y) + \lambda(y') \neq \lambda(y + y')$ для некоторых $y, y' \in C^{(n-1)/2}$, то код Васильева длины n является нелинейным.*

Поскольку функция λ произвольна, то, принимая во внимание предыдущие итеративные шаги, т. е. подставляя в (4) снова произвольный код Васильева длины $(n-1)/2$, затем произвольный код Васильева длины $(n-3)/4$ и т. д., получим следующее утверждение.

Следствие 3. *Число D_n различных кодов Васильева длины n удовлетворяет следующей нижней оценке*

$$D_n \geq 2^{2^{\frac{n+1}{2}-\log(n+1)}} \cdot 2^{2^{\frac{n+5}{4}-\log(n+1)}} \cdot 2^{2^{\frac{n+17}{8}-\log(n+1)}} \cdot \ldots$$

для достаточно больших n.

Зная нижнюю оценку числа различных кодов длины n, легко вычислить нижнюю оценку числа неэквивалентных кодов с теми же параметрами. Для этого достаточно разделить число различных кодов на $2^n \cdot n!$, где 2^n – число различных сдвигов на векторы из E^n и $n!$ – число различных подстановок на n координатах. Нетрудно из следствия 3 получить следующее утверждение.

Следствие 4. *Для числа N_n неэквивалентных кодов Васильева длины n справедливо*

$$N_n \geq 2^{2^{\frac{n+1}{2}-\log(n+1)}} \cdot 2^{2^{\frac{n+5}{4}-\log(n+1)}}$$

при достаточно больших n.

Эта оценка до 1996 года оставалась лучшей нижней оценкой числа неэквивалентных совершенных кодов, несмотря на многочисленные усилия многих исследователей.

Для $n = 7$ существует единственный совершенный код длины 7, для $n = 15$ существует 11 неэквивалентных кодов Васильева длины 15 и по крайней мере 963 неэквивалентных кода, полученных каскадной конструкцией, см. [2]. Следует отметить, что классификация совершенных кодов даже длины 15 до сих пор не найдена.

Упражнения.

1. Доказать следствия 1 и 2.

2. Доказать следствие 4, используя формулу Стирлинга

$$n^n e^{-n}\sqrt{2n\pi} \leq n! \leq n^n e^{1-n}\sqrt{2n\pi}. \tag{5}$$

6 Теорема Глаголева

В этом параграфе множество строк порождающей матрицы кода будем называть *базовым множеством*. Для доказательства теоремы Глаголева потребуется следующий несложно доказываемый факт.

Лемма 1. *Если G – линейный код с кодовым расстоянием d и если найдется такой вектор x, что $d(G, x) \geq d$, то множество $G \cup (G + x)$ является линейным кодом с кодовым расстоянием d.*

Теорема 11. (Глаголев, 1971 г.) *Для любого двоичного линейного $[n, k, d]$-кода C существует линейный код C' с теми же параметрами такой, что его базовое множество состоит из кодовых слов минимального веса d.*

Доказательство. Рассмотрим базовое множество

$$T_d \cup T_{d+1} \cup \ldots \cup T_{d+p}$$

кода C. Здесь множество T_d – максимальное линейно независимое множество кодовых слов веса d, множество T_{d+1} – множество кодовых слов веса $d+1$, которое может быть выбрано в коде C так, что $T_d \cup T_{d+1}$ – максимальное линейно независимое множество кодовых слов веса не более $d+1$. Аналогично выбираем остальные множества вплоть до T_{d+p} кодовых слов веса $d+p$ для некоторого p. Таким образом код C совпадает с линейной оболочкой множества $T_d \cup T_{d+1} \cup \ldots \cup T_{d+p}$, т.е.

$$C =< T_d \cup T_{d+1} \cup \ldots \cup T_{d+p} > .$$

Рассмотрим произвольный вектор y из T_{d+1}. Докажем, что расстояние между y и любым кодовым словом из T_d больше d. Пусть это неверно и найдется вектор $z \in T_d$ такой, что $d(y,z) = d$. Тогда $w(y+z) = d$ и в силу линейности кода C имеем $y+z \in C$ и $y+z \notin T_d$. Следовательно, получили подмножество $T_d \cup (y+z)$ в коде C, которое является линейно независимым множеством кодовых слов веса d более мощным, чем T_d, противоречие выбору множества T_d. Следовательно

$$d(T_d, y) \geq d+1. \tag{6}$$

Возьмем любой вектор y' веса d, предшествующий вектору y, т. е. $y' \prec y$, что означает, что все единичные координаты вектора y' находятся среди единичных координат кодового слова y. Используя (6), получим

$$d(T_d, y) > d(T_d, y') \geq d.$$

Рассмотрим множество $T_d \cup (T_d + y')$. Согласно лемме 1 оно является линейным кодом с кодовым расстоянием d. Далее аналогичным образом в множестве T_{d+1} найдем вектор y'' и рассмотрим множество $T_d \cup \{y', y''\}$, которое позволяет построить новый линейный код с расстоянием d и т. д., переходя от множества T_{d+1} к множеству T_{d+2} и далее до множества T_{d+p}, не более чем за k шагов построим линейный $[n,k,d]$-код C' с базовым множеством, состоящим из кодовых слов минимального веса d.

Замечание.

В 1992 г. Ю. Симонис получил аналогичный результат для q-значных линейных кодов над $GF(q)$.

Следующее утверждение вытекает из теоремы Глаголева и предложения 2 о единственности кода Хэмминга.

Следствие 5. *Для произвольного кода Хэмминга существует базовое множество, состоящее из кодовых слов веса 3.*

Упражнения.

1. Доказать, что базовое множество кода Хэмминга, состоящее из кодовых слов веса 3, может быть построено индуктивно из представления кода Хэмминга посредством конструкции Васильева.

———————————

Список литературы

[1] *Мак-Вильямс Ф.Дж., Слоэн Н.Дж.А.* Теория кодов, исправляющих ошибки: М.: Связь. 1979.

[2] F. I. Solov'eva, On perfect codes and related topics, Com^{2}Mac Lecture Note Series 13, Pohang 2004.