Теория кодирования

Гошин Егор Вячеславович, к.т.н., доцент
кафедры суперкомпьютеров и общей информатики

Многочлены над полем K

Традиционно циклические коды представляются в форме многочленов.

Многочленом степени n над полем K называется многочлен $a_0 + a_1 x + \ldots + a_n x^n$, где коэффициенты $a_0, \ldots, a_n$ являются элементами K.

Многочлены над полем K складываются и умножаются как обычно за исключением того, что $x^i + x^i = 0$.

Пример:

$$(1 + x + x^3 + x^4) + (x + x^2 + x^3) = 1 + x^2 + x^4$$

$$(1 + x + x^3 + x^4)(x + x^2 + x^3) =$$
$$= x + 2x^2 + 2x^3 + 2x^4 + 2x^5 + 2x^6 + x^7 = x + x^7$$

Деление многочленов с остатком

$$
\begin{array}{r|l}
x + x^2 + x^6 + x^7 + x^8 & 1 + x + x^2 + x^4 \\ \cline{2-2}
x^4 + x^5 + x^6 + x^8 & x^4 + x^3 \\ \cline{1-1}
\\
x + x^2 + x^4 + x^5 + x^7 & \\
x^3 + x^4 + x^5 + x^7 & \\ \cline{1-1}
x + x^2 + x^3 &
\end{array}
$$

Многочлены и слова

Каждому многочлену $f(x) = a_0 + a_1 x + \ldots + a_{n-1} x^{n-1}$ степени не выше $n - 1$ над полем K может быть поставлено в соответствие слово $v = a_0 a_1 \ldots a_{n-1}$.

Пусть $n = 7$, тогда

$$1 + x + x^2 + x^4 \leftrightarrow (1110100)$$
$$1 + x^4 + x^5 + x^6 \leftrightarrow (1000111)$$

Так код C длины n может быть представлен в виде множества многочленов степени не выше $n - 1$.

Пример линейного кода

$$
\begin{aligned}
(0000) &\leftrightarrow 0 \\
(1010) &\leftrightarrow 1 + x^2 \\
(0101) &\leftrightarrow x + x^3 \\
(1111) &\leftrightarrow 1 + x + x^2 + x^3
\end{aligned}
$$

Циклические коды

Циклическим сдвигом $\pi(v)$ вектора v называется вектор, полученный перемещением последнего элемента вектора в его начало и сдвигом остальных на одну позицию вправо.

$$v_1 = (10110), \qquad \pi(v_1) = (01011)$$
$$v_2 = (111000), \qquad \pi(v_2) = (011100)$$

Код называется циклическим, если циклический сдвиг любого его кодового слова является кодовым словом.

Линейный, но не циклический код: $\{(000), (101), (010), (111)\}$.

Циклический, но не линейный код: $\{(110), (101), (011)\}$.

Циклический линейный код: $\{(000), (110), (101), (011)\}$

Порождающий вектор

Если некоторое слово v и его циклические сдвиги формируют множество S_v, линейная оболочка которого образует код C, говорят, что v является порождающим вектором (генератором) циклического линейного кода C.

Пример:

Линейные сдвиги $v = (1101000)$ образуют множество:

$$S = \begin{Bmatrix} (1101000), (0110100), (0011010), (0001101), \\ (1000110), (0100011), (1010001) \end{Bmatrix}$$

линейной оболочкой которого является циклический линейный код

$$C = \begin{Bmatrix} (0000000), (1101000), (0110100), (0011010), \\ (0001101), (1000110), (0100011), (1010001), \\ (1011100), (0101110), (0010111), (1001011), \\ (1100101), (1110010), (0111001), (1111111), \end{Bmatrix}$$

Циклический код в форме многочленов

Циклические коды могут быть представлены в терминах многочленов. Так результат циклического сдвига $v(x)$ может быть представлен как

$$xv(x)\;mod\;1+x^n$$

Пусть вектор равен $v=(1101)$, тогда ему соответствует многочлен $v(x)=1+x+x^3$. $\pi(v)=(1110)$ соответствует

$xv(x)=x(1+x+x^3)\;mod\;(1+x^4)=x+x^2+x^4\;mod\;(1+x^4)=1+x+x^2$.

Циклические коды в форме многочленов

Рассматривая циклические коды, будем интерпретировать коды одновременно как кодовые слова и как многочлены.

Например, все циклические сдвиги слова v длины n могут быть представлены в виде $x^i v(x)\ mod\ 1 + x^n$, где $i = 0,1, \dots n - 1$.

Пусть $v = (1101000)$, $v(x) = 1 + x + x^3$. Тогда

$$0110100 \leftrightarrow xv(x) = x + x^2 + x^4$$
$$0011010 \leftrightarrow x^2 v(x) = x^2 + x^3 + x^5$$
$$0001101 \leftrightarrow x^3 v(x) = x^3 + x^4 + x^6$$
$$1000110 \leftrightarrow x^4 v(x) = x^4 + x^5 + x^7 \equiv 1 + x^4 + x^5$$
$$0100011 \leftrightarrow x^5 v(x) = x^5 + x^6 + x^8 \equiv x + x^5 + x^6$$
$$1010001 \leftrightarrow x^6 v(x) = x^6 + x^7 + x^9 \equiv 1 + x^2 + x^6$$

Порождающий многочлен циклического кода

Определим порождающий многочлен $g(x)$ циклического линейного кода C как уникальный ненулевой многочлен наименьшей степени из C.

Для кода длины n и размерности $n - k$ степенью порождающего многочлена будет k.

При этом порождающий многочлен должен быть делителем $x^n + 1$ и быть неприводимым.

Кодирование

Кодирование заключается в умножении исходного сообщения на порождающий многочлен:

$$a(x) \to v(x) = a(x)g(x)$$

Матрица такого кода имеет вид:

$$G = \begin{bmatrix} g(x) \\ xg(x) \\ \dots \\ x^k g(x) \end{bmatrix}$$

Пример матриц циклического кода

$$g(x) = 1 + x^2 + x^3$$

Матрица имеет вид:

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

Синдромы циклического кода

Каждой одиночной ошибке соответствует одночлен $e(x) = x^i$.

Синдромом циклического кода называется остаток

$$r\big(e(x)\big) = e(x)\, mod\ g(x).$$

На самом деле, все кодовые слова должны делиться на $g(x)$ без остатка (в силу способа формирования кода). Поэтому, если принято сообщение $w(x) = v(x) + e(x) = a(x)g(x) + e(x)$, его остаток:

$$w(x)\, mod\ g(x) = \big(a(x)g(x) + e(x)\big) mod\ g(x) = e(x) mod\ g(x).$$

Таким образом, принцип работы аналогичен использованию проверочной матрицы:

$$wH = (v + e)H = vH + eH = eH.$$

Декодирование циклического кода

Очевидно, что процесс декодирования при отсутствии (или после исправления) ошибок представляет собой деление на порождающий многочлен.

Как правило, для исправления ошибок с использование циклических линейных кодов **не используются** таблицы синдромов.

«Цикличность» кодов позволяет использовать значительно более экономичное решение.

Кроме того, это решение обладает ещё дополнительным свойством: исправлять пакеты ошибок (об этом чуть позже).

Алгоритм декодирования

Для принятого слова $w(x)$:

1. Вычислить синдром $s(x) = w(x) \, mod \; g(x)$.
2. Для каждого $i \geq 0$ вычислить s_i: $s_i(x) = x^i s(x) \, mod \; g(x)$ до тех пор, пока не будет найден синдром с $wt(s_i) \leq t$. Если такой синдром найден, то вектор ошибки равен e: $e(x) = x^{n-i} s_i(x)$.
3. Если среди $i = 0, \ldots, n-1$ такой синдром не найден, то исправление ошибки невозможно.

Исправление пакетов ошибок

С точки зрения практики (по ряду причин) возникновение большого числа равномерно распределённых ошибок значительно менее вероятно, чем появление этих ошибок в форме пакета – нескольких идущих подряд ошибок.

Ограничение на характер ошибки позволяет значительно уменьшить длину кода для той же размерности. А циклические коды обеспечивают эффективный метод исправления такого типа ошибок.

Исправление пакетов ошибок

Будем называть циклическим пакетом ошибок длины не более t все возможные циклические сдвиги вектора $(xx \dots x00 \dots 0)$, где длина участка $xx \dots x$ равна t, а $x \in \{0,1\}$.

Так для кода длины 7 циклическим пакетом ошибок длины не более 3 будут называться все ошибки, соответствующие шаблонам

$$E = \{1110000, 1010000, 1100000, 1000000\}$$

и всем их циклическим сдвигам.

Алгоритм исправления пакетов ошибок

Алгоритм полностью идентичен алгоритму исправления для циклического кода за исключением этапа остановки:

Для принятого слова $w(x)$:

1. Вычислить синдром $s(x) = w(x) \; mod \; g(x)$.
2. Для каждого $i \geq 0$ вычислить s_i: $s_i(x) = x^i s(x) \; mod \; g(x)$ до тех пор, пока не будет найден синдром с $s_i \in E$, где E – множество шаблонов ошибок. Если такой синдром найден, то вектор ошибки равен e: $e(x) = x^{n-i} s_i(x)$.
3. Если среди $i = 0, \dots, n-1$ такой синдром не найден, то исправление ошибки невозможно.

Пример

$g(x) = 1 + x + x^3$.

Код (7, 4).

Кодирование:

$a = (1001),\ a(x) = 1 + x^3$,

$v(x) = a(x)g(x) = 1 + x + x^4 + x^6,\ v = (1100101)$.

Пусть возникла ошибка:

$w = (1100001)$.

Пример

Декодирование

$w = (1100001),\ w(x) = 1 + x + x^6.$

$s(x) = w(x)\, mod\ g(x) = x + x^2.$

$i = 0,\ s_0(x) = x + x^2,$

$i = 1,\ s_1(x) = x^2 + x^3\, mod\ g(x) = 1 + x + x^2,$

$i = 2,\ s_2(x) = x^3 + x^4\, mod\ g(x) = 1 + x^2,$

$i = 3,\ s_3(x) = x^2 + x^3\, mod\ g(x) = 1,$

$e(x) = x^{n-3} s_3(x) = x^4.$

$w(x) + e(x) = 1 + x + x^4 + x^6 \leftrightarrow (1100101).$

Пример

$g(x) = 1 + x + x^2 + x^3 + x^6$,

Код (15, 9).

Кодирование:

$a = (1001.0001.1),\ a(x) = 1 + x^3 + x^7 + x^8$,

$v(x) = a(x)g(x) = (1 + x^3 + x^7 + x^8)(1 + x + x^2 + x^3 + x^6) = 1 + x + x^2 + x^4 + x^5 + x^7 + x^9 + x^{11} + x^{13} + x^{14}$,

$v = (1110.1101.0101.011)$.

Пусть возникла ошибка:

$w = (1110.1101.1111.011)$.

Пример

Декодирование

$w = (1110.1101.1111.011),\ w(x) = 1 + x + x^2 + x^4 + x^5 + x^7 + x^8 + x^9 + x^{10} + x^{11} + x^{13} + x^{14}.$

$g(x) = 1 + x + x^2 + x^3 + x^6,$

$s(x) = w(x)\ mod\ g(x) = 1 + x^2 + x^3 + x^4.$

$i = 0,\ s_0(x) = 1 + x^2 + x^3 + x^4,$

$i = 1,\ s_1(x) = x + x^3 + x^4 + x^5,$

$i = 2,\ s_2(x) = x^2 + x^4 + x^5 + x^6\ mod\ g(x) = 1 + x + x^3 + x^4 + x^5,$

$i = 3,\ s_3(x) = x + x^2 + x^4 + x^5 + x^6\ mod\ g(x) = 1 + x^3 + x^4 + x^5,$

$i = 4,\ s_4(x) = x + x^4 + x^5 + x^6\ mod\ g(x) = 1 + x^2 + x^3 + x^4 + x^5,$

$i = 5,\ s_5(x) = x + x^3 + x^4 + x^5 + x^6\ mod\ g(x) = 1 + x^2 + x^4 + x^5,$

$i = 6,\ s_6(x) = x + x^3 + x^5 + x^6\ mod\ g(x) = 1 + x^2 + x^5,$

$i = 7,\ s_7(x) = x + x^3 + x^6\ mod\ g(x) = 1 + x^2,$

$1 + x^2 \in \{1, 1 + x, 1 + x^2, 1 + x + x^2\}.$

$e(x) = x^{15-7} s_7(x) = x^8(1 + x^2) = x^8 + x^{10}.$

$w(x) + e(x) = 1 + x + x^2 + x^4 + x^5 + x^7 + x^9 + x^{11} + x^{13} + x^{14} \leftrightarrow (1110.1101.0101.011).$