Математика и ее приложения : ЖИМО. 2013. Вып. 1 (10). С. 47–50

УДК 512.543

А. А. Толстопятов[1]

Арифметика кольца булевых полиномов и разбиение файлов на буферы

Ключевые слова: Поле Галуа, булево сжатие, булевы полиномы.

Показано, что если булевы полиномы разложить по базису главных идеалов – полиномам Лагранжа, то вычисление в кольце булевых полиномов сводятся к арифметическим вычислениям в декартовом произведении поля Галуа $GF(2)$. Построено арифметическое представление разбиения файла на буферы. Рассмотрен переход от одного разбиения к другому.

Key words: Galua field, boolean compress, boolean polynoms.

It is shown that if the Boolean polynomials spread out on the basis of principal ideals – Lagrange polynomials, the calculation in a ring Boolean polynomials are reduced to arithmetic calculations in the Cartesian product of Galois field $GF(2)$. Built arithmetic idea of splitting the file into the buffers. Reviewed the transition from one partition to another.

Предложенный в [1] подход к сжатию файлов основан на том, что если разбить файл на кортежи равной длины n битов, а потом объединить их в L буферов, содержащих $m_l, l = 1, \ldots, L$ кортежей и рассматривать кортежи, входящие в l-й буфер как решение уравнения

$$f_l(x_i) = 0, \tag{1}$$

где $x_i, i = 1, \ldots, n$ – булевы переменные, принимающие значения из поля Галуа $GF(2)$, то булевы полиномы образуют кольцо. Это значит, что их можно не только складывать, но и умножать. А это позволяет использовать для сжатия файла не только линейные зависимости между f_l, но и нелинейные.

В настоящей работе задача вычислений в кольце полиномов f_l сведена к задаче сложения и умножения 0 и 1, т. е. элементов $GF(2)$. Тем самым эти вычисления сведены к арифметическим операциям. Такая арифметика может быть названа арифметикой булевых полиномов.

Булевы полиномы $f_l(x_i)$ могут быть представлены элементами $GF(2)$, если выбрать базис в кольце f_l и все f_l разложить по этому базису.

Базисом, который позволяет построить арифметику булевых полиномов, является базис главных идеалов в кольце f_l – полиномы Лагранжа $L_j(x_i), j = 0, 1, \ldots, 2^n - 1$. Эти полиномы строятся из полиномов Лагранжа

© Толстопятов А. А., 2013

[1]Ивановский государственный университет; E-mail: khash2@mail.ru Работа выполнена при финансовой поддержке РФФИ (проект 13-07-00628а)

от одной булевой переменной x так:

$$L_0(x) = x + 1, \quad L_1(x) = x \ . \tag{2}$$

$$L_j(x_i) = \prod_{i=1}^{n} L_{j_i}(x_i) \ , \tag{3}$$

где

$$j = \sum_{k=1}^{n} j_k 2^{k-1} \ . \tag{4}$$

Так как для $L_0(x)$ и $L_1(x)$ справедливо, что:

$$L_0^2 = L_0; \ L_1^2 = L_1; \ L_0 L_1 = 0, \tag{5}$$

то для $L_j(x_i)$ справедливо, что:

$$L_i L_j = L_j \delta_{ij}. \tag{6}$$

причем в правой части (6) нет суммирования по j. Свойство (6) можно рассматривать как ортогональность полиномов Лагранжа. Именно (6) позволяет построить арифметику кольца булевых полиномов. Действительно, пусть f, g, h – булевы полиномы от булевых переменных x_i. Их разложения по базису полиномов Лагранжа есть:

$$f = \sum_{j=0}^{2^n-1} a_j L_j(x_i) = 1 \ , \tag{7}$$

$$g = \sum_{j=0}^{2^n-1} b_j L_j(x_i) = 1 \ , \tag{8}$$

$$h = \sum_{j=0}^{2^n-1} c_j L_j(x_i) = 1 \ , \tag{9}$$

Тогда полиномы f, g, h могут быть представлены кортежами коэффициентов $a_j \in GF(2)$:

$$\begin{array}{l} f = (a_0, a_1, \ldots, a_{2^n-1}), \\ g = (b_0, b_1, \ldots, b_{2^n-1}), \\ h = (c_0, c_1, \ldots, c_{2^n-1}), \end{array} \tag{10}$$

Операции сложения f и g вводятся как:

$$f + g = \sum_{j=0}^{2^n-1} (a_j + b_j) L_j = h = \sum_{j=0}^{2^n-1} c_j L_j \ , \tag{11}$$

а значит

$$c_j = a_j + b_j, \tag{12}$$

т. е. сложение выполняется покомпонентно.

Для умножения, в силу (6) будем иметь:

$$fg = \sum_{j=0}^{2^n-1} a_j L_j \cdot \sum_{k=0}^{2^n-1} b_k L_k = \sum_{j=0}^{2^n-1} \sum_{k=0}^{2^n-1} a_j b_k L_j L_k =$$

$$= \sum_{j=0}^{2^n-1} \sum_{k=0}^{2^n-1} a_j b_k L_j \delta_{jk} = \tag{13}$$

$$= h = \sum_{j=0}^{2^n-1} c_j L_j,$$

а значит:

$$c_j = a_j b_j, \tag{14}$$

т. е. умножение выполняется тоже покомпонентно.

Если кольцо булевых полиномов f_l обозначить через $\mathcal{K}(f)$, то (12) и (14) дают, что:

$$\mathcal{K}(f) = GF(2) \times GF(2) \times \ldots \times GF(2) = GF(2)^{2^n} \tag{15}$$

где в декартовом произведении в (15) 2^n множителей.

Построенная арифметика кольца булевых полиномов позволяет представлять арифметически не только вычисления в $\mathcal{K}(f)$, но и разбиения файла на буферы, т. е. объединение кортежей в буферы [3]. Для построения такого разбиения нужно выполнить следующие операции:

1. Разбить файл на кортежи равной длины n.

2. Объединить кортежи в L буферов. содержащих по $m_l, l = 1, \ldots, L$ кортежей.

3. В каждом буфере удалить повторяющиеся кортежи, чтобы каждый кортеж в буфере остался в единственном экземпляре.

Использование в качестве базиса полиномов Лагранжа имеет еще одно преимущество. Оно позволяет безо всяких вычислений найти коэффициенты c_{lj} в разложении [2]:

$$f_l = \sum_{j=0}^{2^n-1} c_{l_j} L_j. \tag{16}$$

А именно, существует 2^n разных кортежей длины n, которые можно рассматривать как двоичный код натуральных чисел и нуля. Их удобно перенумеровать числом $j = 0, 1, \ldots, 2^n - 1$, которым нумеруются и 2^n полиномов Лагранжа L_j. Если в l-й буфер входит кортеж с номером j, то $c_{l_j} = 0$. Если этот кортеж не входит – то $c_{l_j} = 1$. Это позволяет любое разбиение файла на L буферов представить в виде матрицы $L \times 2^n$:

кортеж	0	1	...	$2^n - 1$
1	c_{10}	c_{11}	...	$c_{1,2^n-1}$
...	...	...	...	...
L	$\boldsymbol{\sigma}_{L0}$	$\boldsymbol{\sigma}_{L1}$	...	$\boldsymbol{\sigma}_{L,2^n-1}$

Таблица 1. Арифметическое представление разбиения файла на буферы.

Таблица 1 заполняется числами c_{lj} в соответствии с сформулированными выше правилами. Поскольку таким образом можно представить любое разбиение одного и того же файла, то возникает вопрос об арифметическом представлении категории всех разбиений одного и того же файла.

Ясно, что если в разбиении файла не меняется число буферов, то Таблица 1 будут иметь одно и то же число строк. Но и изменение числа буферов в разбиении позволяет построить отображение матрицы размерности $L_1 \times 2^n$ на матрицу размерности $L_2 \times 2^n$, так как эти матрицы могут быть разложены в прямое произведение матриц $L_1 \times L_1$ и $2^n \times 2^n$ и $L_2 \times L_2$ и $2^n \times 2^n$ соответственно.

Проблемы арифметизации категории разбиений связаны с пунктом 3 в правилах построения разбиений, так как при изменении разбиения в буферах могут появляться новые кортежи.

Список литературы

1. *Толстопятов А. А.* О возможности использования булевых уравнений для сжатия файлов // Вестник Иван. гос. ун-та. 2003. Вып 3. С. 82–84.
2. *Толстопятов А. А.* // Алгоритм кодирования и декодирования поля принадлежности при булевом сжатии файлов // Математика и ее приложения : журн. Иван. мат. о-ва. 2008. Вып. 1(5). С. 53–76.
3. *Толстопятов А. А.* Алгоритм разбиения файла на буферы при булевом сжатии // Математика и ее приложения : журн. Иван. мат. о-ва. 2008. Вып. 1(5). С. 77–88.

Поступила в редакцию 26.11.2013.