Математика и ее приложения : ЖИМО. 2013. Вып. 1 (10). С. 65–70

УДК 512.543

С. И. Хашин, Ю. А. Хашина[1]

Свойства b-ранга системы булевых полиномов

Ключевые слова: булевы полиномы, автоморфизм.

При разработке алгоритмов булева сжатия информации и в некоторых разделах криптографии возникает задача эффективного нахождения множества, порождающего данную систему булевых полиномов. В работе предлагается метод ее решения эффективный, если количество булевых переменных не превосходит 32.

Keywords: boolean polynomials, automorphism.

In developing algorithms for Boolean data compression and in some sections of the cryptography, the problem of effective finding the generated set for given system of Boolean polynomials. This paper proposes a method for its effective solution if the number of Boolean variables does not exceed 32.

1. Введение

Булевы кольца и алгебры являются классическими, хорошо изученными алгебраическими объектами ([1, 3]). В последнее время они находят и новые области применения, например в криптографии с открытым ключом ([4]). В некоторых работах предлагается их использование для сжатия информации ([2, 5, 6, 7]). В работах ([6, 7]) сформулированы некоторые задачи, касающиеся свойств решений систем булевых уравнений, которые требуется решить для целей сжатия информации. В настоящей работе исследуется одна из этих задач, а именно – эффективное нахождение наименьшего порождающего множества данной системы булевых полиномов. В работе приводится полное решение этой задачи, практически применимое при количестве булевых переменных ≤ 32.

2. Определения

Определение 1. *Пусть N – натуральное число.*

а) Обозначим через U_N векторное пространство размерности N над полем $\mathbb{Z}_2$, $U_N = \mathbb{Z}_2^N$.

© Хашин С. И., Хашина Ю. А., 2013

[1]Ивановский государственный университет; E-mail: khash2@mail.ru. Работа выполнена при поддержке гранта РФФИ 13-07-00628

б) Обозначим через G_N кольцо булевых полиномов от N переменных, рассматриваемых как функции на U со значениями в $\mathbb{Z}_2$, то есть

$$G_N = \mathbb{Z}_2[x_1, \ldots, x_N]/(x_1^2 - x_1, \ldots, x_N^2 - x_N).$$

Определение 2. *Пусть $v = (v_1, \ldots, v_k)$ – произвольное, возможно пустое подмножество в $[1, 2, \ldots, N]$. Через J_v обозначим произведение*

$$J_v = x_{v_1} \cdot \cdots \cdot x_{v_k}.$$

В частности, если v – пустое подмножество, то $J_v = 1$.

Теорема 1. *Множество мономов $\{J_v, v \subset [1, 2, \ldots, N]\}$ образует $\mathbb{Z}_2$-базис в кольце G_N.*

Этот базис кольца булевых полиномов называется базисом Жегалкина.

Пример 1. В кольце G_2 базис Жегалкина состоит из 4 мономов

$$\{1,\ x_1,\ x_2,\ x_1x_2\}.$$

Таким образом, кольцо G_N является векторным пространством над $\mathbb{Z}_2$ размерности 2^N.

Теорема 2. *В кольце G_N единственным обратимым элементом является* 1.

Доказательство. Обратимая функция на U_N со значениями в $\mathbb{Z}_2$ должна всюду принимать ненулевые значения, то есть всюду равняться 1. ■

Определение 3. *Элемент f из кольца G_N называется неразложимым, если его нельзя представить нетривиальным образом в виде произведения двух других элементов кольца (тривиальными назовем представления $f = f \cdot 1 = f \cdot f$).*

Теорема 3. *Множество неразложимых элементов кольца G_N взаимнооднозначно соответствует множеству точек пространства U_N.*

Доказательство. Функция на U_N равная 0 в одной точке этого пространства и 1 во всех остальных точках будет неразложимой. Если же она имеет не менее двух корней, то ее можно представить нетривиальным образом в виде произведения двух функций. ■

Следствие 1. *Множество неразложимых элементов образует $\mathbb{Z}_2$-базис в кольце G_N.*

Определение 4. *Определим два булевых многочлена от одной переменной:* $L_0(x) = x$ *и* $L_1(x) = 1 - x$.

Определение 5. *Для каждого битового вектора* $v = \{v_1, \ldots, v_N\} \in U_N$ *обозначим через* $L_v = L_v(x_1, , x_N)$ *многочлен* $L_{v_1}(x_1) \cdot \ldots \cdot L_{v_N}(x_N)$. *Многочлены* L_v *называются многочленами Лагранжа. Множество всех многочленов* L_v *образует базис кольца* G_N, *который называется базисом Лагранжа.*

Пример 2. В кольце G_2 базис Лагранжа состоит из 4 полиномов

$$\{x_1x_2,\ (1+x_1)x_2,\ x_1(1+x_2),\ (1+x_1)(1+x_2)\}.$$

Следствие 2. *Многочлены вида* $1+L_v$ *и только они являются неразложимыи.*

Доказательство. Для каждой точки $v \in U_N$ многочлен L_v, рассматриваемый как функция $L_v : U_N \to \mathbb{Z}_2$ принимает значение 1 в точке v и значение 0 во всех остальных точках. ■

3. Гомоморфизмы и автоморфизмы булевых колец

Определение 6. *Пусть* s *– произвольное, не обязательно линейное отображение множества* U_N *в* U_M *для некоторых натуральных* M, N. *Для произвольной функции* f *на* U_M *через* s^*f *обозначим функцию на множестве* U_N, *определяемую формулой* $s^*f(v) = f(s(v))$. *Отображение* s^* *переводит сумму функций в сумму, произведение – в произведение и, следовательно, является гомоморфизомом колец*

$$s^* : G_M \to G_N.$$

Теорема 4. *Группа автоморфизмов кольца* G_N *изоморфна группе перестановок точек пространства* U_N *и состоит из* $(2^N)!$ *элементов.*

Доказательство. Пусть s – произвольная перестановка на множестве U_N. Тогда $s^* : G_N \to G_N$ – эндоморфизм кольца G_N. Так как для перестановки s существует обратная, то f^* – автоморфизм. Обратно, каждый автоморфизм переводит неразложимые элементы в неразложимые и, следовательно, задает некоторую перестановку на множестве U_N. ■

Определение 7. *Пусть* $F = (f_1, \ldots, f_s)$, $f_i \in G_N$ *– набор из* s *булевых полиномов. Будем рассматривать его как отображение* $F : U_N \to \mathbb{Z}_2^s$.

а) Назовем его объемом количество элементов в образе:

$$vol(F) = |Im(F)|.$$

б) Назовем его b-рангом наименьшее натуральное k такое, что $vol(F) = |Im(F)| \leq 2^k$.

Очевидно, что $vol(F) \leq 2^s$ и b-ранг системы F не превосходит s – количества элементов в ней.

Теорема 5. *Пусть* $F = (f_1, \ldots, f_s)$, $f_i \in G_N$ *– набор из* s *булевых полиномов и* M *– его b-ранг. Тогда существуют* M *булевых полиномов* $(h_1, \ldots, h_M)$, $h_i \in G_N$ *таких, что исходные полиномы* f_i *через них выражаются, то есть* $f_i = g_i(h_1, \ldots, h_M)$ *для некоторых булевых полиномов* $g_i \in G_M$.

Доказательство. Будем считать в кольце G_N порождающими переменными $(x_1, \ldots, x_N)$, в кольце G_M – $(y_1, \ldots, y_M)$, в кольце G_s – $(z_1, \ldots, z_s)$. Рассмотрим отображение $F : U_N \to U_s$. Согласно определению $f_i = F^*(z_i)$. Выберем в пространстве U_M произвольные $vol(F)$ различных точек. Это можно сделать, так как по определению b-ранга $vol(F) \leq 2^M$. Отображение $F : U_N \to U_s$ в этом случае можно пропустить через U_M:

$$U_N \xrightarrow{F_1} U_M \xrightarrow{F_2} U_s \tag{1}$$

для некоторой пары отображений F_1 и F_2. Переходя к булевым кольцам получим пару отображений

$$G_s \xrightarrow{F_2^*} G_M \xrightarrow{F_1^*} G_N \tag{2}$$

Введем обозначение $g_i = F_2^*(z_i) \in G_M$, $i = 1, \ldots, s$ и $h_j = F_1^*(y_j) \in G_N$, $j = 1, \ldots, M$. Согласно (1), $f_i = F_1^*(g_i)$, то есть $f_i = g_i(h_1, \ldots, h_M)$. ■

4. Алгоритм нахождения порождающих

Теорема (5) утверждает, что для произвольного набора многочленов $(f_1, \ldots, f_s)$ из G_N существуют полиномы $(h_1, \ldots, h_M)$, где M – b-ранг, через которые можно выразить исходные полиномы.

Рассмотрим более подробно алгоритм нахождения этих полиномов в случае $s >> M$ и оценим его вычислительную сложность.

Согласно теореме (5), множество U_N должно быть разбито на не более чем 2^M частей, которые будем называть сегментами, а само разбиение множества N на такие части – сегментацией. На каждом сегмента каждый из многочленов f_i должен принимать одно и тоже значение.

Так как количество сегментов не превышает 2^M, для хранения сегментации требуется массив A из 2^N элементов, каждый длиной по M битов. Учитывая, что $M \leq N$, при $N = 32$ мы получаем 2^{32} степени элементов по

32 бита каждый, то есть 16 Гбайт памяти. Это немало, но вполне доступно на современных компьютерах. Таким образом, $N = 32$ можно рассматривать как предельный размер кортежа, при котором можно реализовать рассматриваемый алгоритм.

Описание алгоритма.

В начале все точки из U_N относим к нулевому сегменту, то есть устанавливаем все элементы массива A в 0.

Обрабатываем исходные многочлены по одному, последовательно. Если очередной многочлен во всех точках какого-то из уже имеющихся сегментов принимает одно и тоже значение, то такой сегмент оставляем без изменений. Если же он принимает как значения 0, так и 1, то такой сегмент разбиваем на две части.

5. Заключение

В работах ([6, 7]) сформулирована задача поиска для данного набора булевых многочленов $(f_1, \ldots, f_s)$ из G_N наименьшего количества полиномов $(h_1, \ldots, h_M)$, через которые можно выразить исходные полиномы.

В этих работах, задача сводится к решению большой системы булевых уравнений относительно коэффициентов искомых полиномов. В такой формулировке задача получается очень сложной.

В настоящей работе предлагается другой подход к этой задаче, в результате которого задача оказывается полностью решенной в большинстве практически важных случаев. По крайней мере, пока количество булевых переменных не превосходит 32, задача оказывается посильной даже для персональных компьютеров, правда достаточно мощных. Для полного решения задачи при 32 переменных требуется оперативная память размером 16Гбайт. На сегодняшний день это большая, но вполне реальная цифра.

Для большего количества булевых переменных предложенный метод будет, фактически неработоспособным из-за чересчур большого требуемого объема памяти. В этих случаях требуется поиск более эффективных алгоритмов.

Список литературы

1. *Владимиров Д. А.* Булевы алгебры. М. : Наука, 1969. 320 с.
2. *Гришко М. Е.* Один из возможных способов разбиения файла на буферы при булевом сжатии файлов // Математика и ее приложения : журн. Иван. мат. о-ва. 2010. Вып. 1(7). С. 25–28.
3. *Логачев О. А., Сальников А. А., Ященко В. В.* Булевы функции в теории кодирования и криптологии. М. : МЦНМО, 2004. 470 с.

4. *Ростовцев А. Г.* Защита от side channel attack на основе случайных изоморфизмов. 2004. URL: http://www.ssl.stu.neva.ru/ssl/archieve/sidech1.pdf (дата обращения: 4.12.2013).

5. *Толстопятов А. А.* О возможности использования булевых уравнений для сжатия файлов // Вестник Иван. гос. ун-та. 2003. Вып. 3. С. 82–84.

6. *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.

7. *Толстопятов А. А.* Построение кодирующего уравнения при булевом сжатии файлов // Математика и ее приложения : журн. Иван. мат. о-ва. 2010. Вып. 1(7). С. 69–83.

Поступила в редакцию 12.12.2013.