Серия «Математика»
2015. Т. 12. С. 49—57

Онлайн-доступ к журналу:
http://isu.ru/izvestia

ИЗВЕСТИЯ
Иркутского
государственного
университета

УДК 519.71

Об одной комбинаторной задаче для множества двоичных наборов *

К. Д. Кириченко
Иркутский государственный университет

Аннотация.

В данной работе рассматривается одна новая комбинаторная задача, возникающая в связи с изучением сложности представлений булевых функций полиномиальными нормальными формами. В теории сложности минимальная сложность представления самой сложной функции называется функцией Шеннона. Таким образом, верхняя оценка функции Шеннона гарантирует наличие представления данной сложности для любой булевой функции, что имеет существенное прикладное значение.

Обычно для нахождения верхней оценки функции Шеннона используются конструктивные алгоритмы минимизации, работающие на классе всех булевых функций. Ранее автором был разработан алгоритм минимизации полиномиальных нормальных форм булевых функций, основанный на комбинаторной технике, связанной с задачами нахождения покрытий и упаковок на множестве двоичных наборов. Полиномиальная форма для заданной булевой функции строится на основе шаблона, который описывается невырожденной матрицей над полем Z_2, где каждая строка и столбец соответствует некоторому двоичному набору. При этом для получения хорошей верхней оценки требуется, чтобы множества наборов в строках и столбцах обладали упаковками с плотностью $1 + o(1)$.

При построении матрицы шаблона естественно воспользоваться линейными кодами, исправляющими одну ошибку, в частности, может быть использован код Хэмминга. Это позволяет использовать понятия теории линейных кодов в формулировках соответствующих комбинаторных задач. Задачу, рассматриваемую в настоящей работе также можно отнести к классу задач на нахождение покрытий и упаковок. При этом на покрытие накладывается ряд дополнительных условий, вытекающих из требований к матрице. В работе приводятся некоторые из возможных покрытий, описанные на языке линейных кодов, исправляющих ошибки.

Ключевые слова: булевы функции, полиномиальные нормальные формы, линейные коды, упаковки и покрытия.

В работе рассматривается одна новая комбинаторная задача, возникшая в результате исследований одного из подходов к получению асимп-

* Работа выполнена при финансовой поддержке РФФИ, грант 13-01-00621.

тотических верхних оценок сложности булевых функций в классе полиномиальных нормальных форм. Этот подход заключается в построении полиномиальных нормальных форм специального вида, называемых шаблонами минимизации.

При исследовании шаблонов минимизации естественным образом возникает класс матриц специального вида над полем Z_2. При этом путем к получению асимптотических верхних оценок сложности является изучение алгебраических свойств этого класса матриц [1].

Данный класс матриц может быть определен следующим образом. Обозначим V_m^n — множество всех двоичных наборов длины n веса m. Назовем некоторый счетный класс множеств $X = \{X_m^n | X_m^n \subseteq V_m^n\}$ упаковываемым, если для каждого X_m^n найдется множество $P(X_m^n) \subseteq V_{m+1}^n$ такое, что для любого $\tilde{x} \in X_m^n$ найдется $\tilde{y} \in P(X_m^n)$, что $\tilde{x} < \tilde{y}$; и $|X_m^n|/|P(X_m^n)| = m + 1 - o(n)$.

Отметим, что для произвольного $\tilde{y} \in V_{m+1}^n$ имеется ровно $m + 1$ меньших наборов из V_m^n, таким образом условие $|X_m^n|/|P(X_m^n)| = m + 1 - o(n)$ означает, что упаковка наборов класса множеств X должна быть асимптотически оптимальной.

Класс матриц $A(X) = \{A_m^n\}$ определяется через класс упаковываемых множеств X следующим образом. Матрица A_m^n содержит $|V_m^n \setminus X_m^n|$ строк и $|X_{m+1}^n|$ столбцов, при этом каждой строке матрицы соответствует некоторый набор $\tilde{\alpha} \in V_m^n \setminus X_m^n$, а каждому столбцу — набор $\tilde{\beta} \in X_{m+1}^n$. Элементом матрицы, находящимся на пересечении строки $\tilde{\alpha}$ и столбца $\tilde{\beta}$, будет единица, если $\tilde{\alpha} < \tilde{\beta}$, и ноль в противном случае.

Ранее автором была доказана оценка функции Шеннона сложности булевых функций в классе ПНФ [1], которую в обозначениях данной работы можно записать

$$L(n) \leq 3 \sum_{i=0}^{n-1} |P(X_i^n)| + \sum_{i=0}^{n-1} (|V_i^n \setminus X_i^n| - rang(A_i^n))$$

Таким образом, для получения хороших верхних оценок требуется либо находить эффективные упаковки почти всех двоичных наборов, либо добиваться того, чтобы матрицы, составленные из неупакованных наборов, имели ранг близкий к количеству этих наборов.

Отметим, что задача упаковки множеств V_m^n достаточно хорошо известна как задача Турана [6]. В 1985 году Редлем было доказано, что для фиксированного m класс множеств V_m^n является упаковываемым [5], однако в данном случае нас более интересуют множества наборов с весами близкими к $\frac{n}{2}$

Дж. Н. Купер, Р. Б. Эллис и А. Б. Канг в 2001 году доказали, что асимптотическая сложность упаковки всех двоичных наборов составляет $\Theta(\frac{2^n}{n}$ [3]), таким образом асимптотика функции Шеннона сложности булевых функций в классе ПНФ равна $\Theta(\frac{2^n}{n})$ [2].

Для получения упаковываемых множеств наборов вполне естественно воспользоваться оптимальными кодами, исправляющими одну ошибку. Удобными объектами такого вида являются линейные коды и, в частности, код Хэмминга. С теорией линейных кодов, исправляющих ошибки, можно познакомиться, например, по книге [4].

Отметим, что оптимальные коды существуют только для двоичных последовательностей длины $2^k - 1$, поэтому далее будем предполагать, что в качестве n используются только такие значения.

При работе с линейными кодами используется тот факт, что множество целых чисел $\{0, 1, \ldots, 2^k - 1\}$ можно считать линейным пространством над полем Z_2, в котором каждое число представляется в виде своего вектора в двоичной записи.

Далее сумму двоичных векторов чисел a и b в поле Z_2 будем обозначать $a \oplus b$ и называть XOR-суммой.

Для упрощения изложения будем использовать следующее обозначение: если p — некоторое натуральное число, то двоичный набор у которого в позиции p находится единица, а остальные элементы — нули будем обозначать $\tilde{p}$.

Напомним, что в теории линейных кодов, исправляющих ошибки, синдромом двоичного набора называется XOR-сумма номеров позиций, в которых произошла ошибка. Здесь синдромом набора $\tilde{\sigma} = (\sigma_1, \ldots, \sigma_n)$ мы будем называть XOR-сумму всех двоичных представлений номеров, в позиции которых набор содержит единицы. Синдром набора $\tilde{\sigma}$ далее будем обозначать $s(\tilde{\sigma})$. Формально $s(\tilde{\sigma}) = \bigoplus_{i=1}^{n} \sigma_i \cdot i$.

Код Хэмминга можно определить как множество кодовых слов, синдром которых равен нулю. При этом все смежные классы кода Хэмминга также являются оптимальными кодами, исправляющими одну ошибку. В качестве вершин упаковки можно взять любой из смежных классов кода Хэмминга, однако здесь для упрощения обозначений мы будем рассматривать упаковку, построенную непосредственно на множестве кодовых слов кода Хэмминга.

Здесь будет использоваться свойство оптимальности кода Хэмминга, которое заключается в том, что каждый набор $\tilde{\alpha}$ либо входит в код, либо в код входит ровно один из соседних с $\tilde{\alpha}$ наборов.

Пусть $\tilde{\sigma}$ некоторый набор такой, что в позиции $p = s(\tilde{\sigma})$ находится ноль. Тогда очевидно, что набор $\tilde{\gamma}$, полученный в результате замены нуля в позиции p на единицу, больше чем $\tilde{\sigma}$, и кроме того, его синдром равен нулю. $s(\tilde{\gamma}) = s(\tilde{\sigma} \oplus \tilde{p}) = s(\tilde{\sigma}) \oplus p = 0$. Таким образом, $\tilde{\gamma}$ является кодовым словом кода Хэмминга.

Определим $C = \{C_m^n\}$ как класс множеств, где каждое множество состоит из всех наборов длины n веса m, у которых в позиции синдрома находится ноль. Этот класс множеств является упаковываемым, поскольку в вершинах упаковки находятся кодовые слова кода Хэмминга,

который является оптимальным. Наборы, входящие в класс C, соответствуют столбцам матриц шаблона. Строкам же соответствуют все наборы, которые не входят в C. Таким образом, их можно определить как множества наборов, синдром которых равен нулю или в позиции синдрома которых находится единица.

Пусть $\tilde{\gamma}$ кодовое слово кода Хэмминга и $||\tilde{\gamma}|| = m$. Тогда, всякий набор $\tilde{\sigma}$ такой, что $||\tilde{\sigma}|| = m+1$ и $\tilde{\sigma} > \tilde{\gamma}$ может быть получен изменением нуля в позиции p на единицу. Тогда $s(\tilde{\sigma}) = s(\tilde{\gamma}) \oplus p = 0 \oplus p = p$. Таким образом, $\sigma_{s(\tilde{\sigma})} = 1$ и все наборы большие, чем $\tilde{\gamma}$ не входят в C^n_{m+1}. Это означает, что в матрице шаблона строка, соответствующая набору $\tilde{\gamma}$ состоит из всех нулей. Поэтому такие строки матрицы можно исключить из дальнейшего рассмотрения, далее рассматриваем только те строки матрицы, которые соответствуют наборам, в позиции синдрома которых находится единица.

Обозначим множества таких наборов за $R^n_m = \{\tilde{\sigma} \subseteq V^n_m | \sigma_{s(\tilde{\sigma})} = 1\}$

Это приводит нас к задаче нахождения ранга матриц A^n_m над полем Z_2, у которых каждой строке соответствует некоторый двоичный набор $\tilde{\alpha}$ длины n, такой, что вес набора $||\tilde{\alpha}|| = m$ и $\alpha_{s(\tilde{\alpha})} = 1$, а каждому столбцу набор $\tilde{\beta}$ длины n, такой что $||\tilde{\beta}|| = m + 1$ и $\beta_{s(\tilde{\beta})} = 0$.

В связи с нахождением ранга таких матриц естественно возникает задача описания их нулевых миноров. Для поля Z_2 это эквивалентно тому, что в каждом столбце минора имеется четное число единиц. Таким образом, возникает комбинаторная задача следующего вида: для заданных натуральных чисел n, m таких, что $n = 2^k - 1$, $m < n$ описать множества наборов $Y \subset R^n_m$ такие, что для всякого $\tilde{\beta} \in C^n_{m+1}$ существует четное число наборов $\tilde{\alpha} \in Y$ меньших, чем $\tilde{\beta}$. Такие множества двоичных наборов далее будем называть четнопокрытыми.

Далее в этой работе будут описаны некоторые из конструкций четнопокрытых множеств.

Предложение 1. *Пусть $\tilde{\sigma} \in V^n_{m-1}$ и синдром $s(\tilde{\sigma}) = 0$. Тогда $Y = \{\tilde{y} \in V^n_m | \tilde{y} > \tilde{\sigma}\}$ — четнопокрытое.*

Доказательство. Докажем, что $Y \subseteq R^n_m$. Если произвольный набор $\tilde{y} \in Y$ больше набора $\tilde{\sigma}$ и их веса отличаются на единицу, то эти наборы различаются в одной позиции, номер которой обозначим p. Таким образом, $y_p = 1$. Тогда $s(\tilde{y}) = s(\tilde{\sigma} \oplus \tilde{p}) = s(\tilde{\sigma}) \oplus p = p$. Тогда $y_p = y_{s(\tilde{y})} = 1$ и $\tilde{y} \in R^n_m$.

Далее возьмем произвольный набор $\tilde{\beta} \in C^n_{m+1}$. Если $\tilde{\beta} > \tilde{\sigma}$, набор $\tilde{\beta}$ отличается от $\tilde{\sigma}$ в двух позициях, номера которых обозначим p и q. Тогда существует ровно два набора $\tilde{\sigma} \oplus \tilde{p}$ и $\tilde{\sigma} \oplus \tilde{q}$, больших чем $\tilde{\sigma}$ и меньших чем $\tilde{\beta}$. По определению оба этих набора входят в Y, то есть $\tilde{\beta}$ накрывает четное число наборов из Y.

Иначе, если наборы $\tilde{\beta}$ и $\tilde{\sigma}$ несравнимы, то не существует наборов из Y меньших чем $\tilde{\beta}$. Таким образом, предложение доказано. □

Данная конструкция является довольно простой и имеет место для почти всех весов k. Очевидно, что симметрическая разность двух четнопокрытых множеств также является четнопокрытым множеством. В связи с этим четнопокрытые множества, представимые в виде симметрической разности нескольких множеств указанного вида, будем называть тривиальными. Основная цель данной работы будет заключаться в нахождении нетривиальных конструкций четнопокрытых множеств.

Для предъявления других конструкций будет полезно ввести еще несколько обозначений. Нам будет удобно представлять двоичные наборы в виде характеристических множеств следующим образом: множество B представляет двоичный набор $\tilde{\beta}$, если B содержит те и только те натуральные числа i, что $\beta_i = 1$.

Заметим, что множество натуральных чисел с нулем меньших 2^k может рассматриваться как линейное векторное пространство над полем Z_2, где каждое число представляется своим вектором двоичной записи длины k. Таким образом, в характеристическом множестве набора каждый компонент является элементом линейного векторного пространства.

Пусть $L = \{0, l_1, l_2, \ldots, l_m\}$ некоторое подпространство линейного векторного пространства над полем Z_2, а $H = \{0, h_1, h_2, \ldots, h_s\}$ его ортогональное дополнение, причем L и H содержат более 2 векторов. Обозначим L_{h_j} смежный класс L по элементу $h_j \in H$.

При доказательстве основных утверждений будет использоваться следующая простая лемма.

Лемма. *Пусть L некоторое подпространство линейного векторного пространства над полем Z_2, состоящее более чем из двух векторов. Тогда* $\sum\limits_{a \in L} a = 0$.

Доказательство. Пусть $b_1, \ldots b_r$ — базис L. Тогда каждый элемент пространства L является одной из 2^r различных линейных комбинаций базисных векторов. При этом каждый базисный вектор будет входить ровно в 2^{r-1} линейную комбинацию. Тогда, учитывая что $2^{r-1} mod\, 2 = 0$ при $r > 1$, получим

$$\sum_{a \in L} a = \sum_{i=1}^{r} 2^{r-1} b_i = 0$$

□

Теорема. *Пусть $L = \{0, l_1, l_2, \ldots, l_k\}$ — некоторое подпространство линейного векторного пространства над полем Z_2, а*

$$H = \{0, h_1, h_2, \ldots, h_s\}$$

его ортогональное дополнение, причем $k > 1, s > 1$; выбрано произвольное множество $L_0 \subseteq L \setminus \{0\}$ и набор нечетных натуральных чисел $n_2, \ldots, n_s$, $n_i \leq k$. Тогда четнопокрытым является множество всех наборов, характеристические множества которых могут быть представлены в виде
$B(L, L_0, h_1, n_2, \ldots, n_s) = \{\tilde{\alpha} | \tilde{\alpha} = L_0 \cup L_{h_1} \cup L_{h_2}^{n_2} \cup \ldots \cup L_{h_s}^{n_s}\}$
где $L_{h_i}^{n_i}$ некоторое подмножество L_{h_i} мощности n_i.

Пример 1. Для $n = 15$, $m = 7$ множество $B(\{0,1,2,3\}, \{1\}, 12, 1, 1)$ содержит следующие наборы
{(100.1000.1000.1111),(100.1000.0100.1111),(100.1000.0010.1111),(100.1000.0001.1111),
(100.0100.1000.1111),(100.0100.0100.1111),(100.0100.0010.1111),(100.0100.0001.1111),
(100.0010.1000.1111),(100.0010.0100.1111),(100.0010.0010.1111),(100.0010.0001.1111),
(100.0001.1000.1111),(100.0001.0100.1111),(100.0001.0010.1111),(100.0001.0001.1111)}
В данном примере выбрано подпространство $\{0,1,2,3\}$, нумерация элементов в наборе идет слева направо, начиная с единицы. Таким образом, первая тройка элементов набора соответствует подпространству. Второй параметр задает подмножество $L_0 \subset L$ и означает, что в первой позиции всех наборов находится единица.

Смежными классами данного подпространства являются множества $\{4,5,6,7\}$, $\{8,9,10,11\}$, $\{12,13,14,15\}$. Третий параметр задает смежный класс $\{12,13,14,15\}$, в позициях которого во всех наборах находятся единицы. Наконец указано, что для смежных классов $\{8,9,10,11\}$, $\{12,13,14,15\}$ всеми способами выбираются подмножества из 1 элемента, и каждый способ порождает один из наборов.

Доказательство. Рассмотрим некоторый набор $\tilde{\beta}$ и его характеристическое множество B. Предполагаем, что B удовлетворяет условию теоремы.

Сначала докажем, что набор $\tilde{\beta}$ содержит 1 в позиции синдрома. Это означает, что сумма всех элементов из B принадлежит B.

Заметим, что $\sum\limits_{b \in L_{h_j}^{n_j}} b = h_j \oplus d_j$, где $d_j \in L$ поскольку

$$\sum_{b \in L_{h_j}^{n_j}} b = \sum_{i=1}^{n_j} h_j \oplus l_i = n_j \cdot h_j \oplus \sum_{i=1}^{n_j} l_i = h_j \oplus d_j$$

так как по построению n_j нечетное число.

Из доказанной леммы $\sum_{j=1}^{s} h_j = 0$, отсюда $\sum_{j=2}^{s} h_j = h_1$. Тогда

$$\sum_{b \in B} b = \sum_{b \in L_0} b \oplus \sum_{b \in L_{h_1}} b \oplus \sum_{j=2}^{s} \sum_{b \in L_{h_j}} b = d_0 \oplus 0 \oplus \sum_{j=2}^{s} (h_j \oplus d_j) = h_1 \oplus d$$

где $d \in L$. Учитывая, что $h_1 \oplus d \in L_{h_1}$ и $L_{h_1} \subset B$ получаем, что в позиции синдрома для набора соответствующего B наверняка находится единица.

Теперь рассмотрим произвольный набор $\tilde{\alpha} = \tilde{b} \oplus \tilde{p}$, полученный заменой в позиции p нуля на единицу из набора $\tilde{\beta}$.

Рассмотрим его характеристическое множество $A = B \vee \{p\}$. При этом возможны два случая: либо $p \in L$, либо $p \in L_{h_j}$, где $j \geq 2$.

В первом случае имеет место доказательство, аналогичное приведенному выше, позволяющее заключить, что в позиции синдрома $\tilde{\alpha}$ находится единица, т. е. $\tilde{\alpha}$ не принадлежит множеству стобцов.

Во втором случае в позиции синдрома набора α может быть как ноль, так и единица, однако независимо от этого множество $L_{h_j}^{n_j} \vee \{p\}$ содержит четное количество элементов, следовательно, существует четное количество способов исключить из этого множества один элемент, получая при этом наборы из B. Таким образом, теорема доказана. □

Другая конструкция с аналогичным набором $n_1, \ldots n_s$ нечетных чисел в тех же обозначениях выглядит следующим образом.

Теорема. *Пусть $L = \{0, l_1, l_2, \ldots, l_k\}$ — некоторое подпространство линейного пространства над полем Z_2, а $H = \{0, h_1, h_2, \ldots, h_s\}$ его ортогональное дополнение, причем $k > 1, s > 1$ и выбраны нечетные натуральные числа $n_1, \ldots, n_s$, $n_i \leq k$. Тогда четнопокрытым является множество всех наборов, характеристические множества которых могут быть представлены в виде*

$$B(L, n_1, \ldots, n_s) = \{\tilde{\alpha} | \tilde{\alpha} = L \setminus \{0\} \cup L_{h_1}^{n_1} \cup L_{h_2}^{n_2} \cup \ldots \cup L_{h_s}^{n_s}\}$$

такие, что $s(\tilde{\beta}) \neq 0$, где $L_{h_i}^{n_i}$ некоторое подмножество L_{h_i} мощности n_i

Доказательство. для данной конструкции очень похоже на приведенное выше. Отметим только одно различие.

$$\sum_{b \in B} b = \sum_{b \in L} b \oplus \sum_{j=1}^{s} \sum_{b \in L_{h_j}} b = 0 \oplus \sum_{j=1}^{s} (h_j \oplus d_j) = \sum_{j=1}^{s} h_j \oplus \sum_{j=1}^{s} d_j = d$$

Таким образом, сумма элементов B в данной конструкции будет принадлежать L, и из условия теоремы не равняться нулю. При этом построение гарантирует, что $L \subseteq B$, то есть в позиции синдрома $\tilde{\beta}$ находится единица. Далее доказательство полностью аналогично. □

Список литературы

1. Кириченко К. Д. Оценки сложности шаблонов минимизации полиномиальных форм булевых функций / К. Д. Кириченко // Изв. Иркут. гос. ун-та. Сер. Математика. – 2009. – Т. 2, № 2. – С. 67–76.
2. Башов М. А. О длине функций k-значной логики в классе полиномиальных нормальных форм по модулю k / М. А. Башов, С. Н. Селезнева // Дискрет. математика. – 2014. – Т. 26, № 3. – С. 3–9.
3. Мак-Вильямс Ф. Дж., Слоэн Н. Дж. А. Теория кодов, исправляющих ошибки. — М.: Связь. — 1979. — 744 с.
4. Cooper J. N. Asymmetric binary covering codes / J. N. Cooper, R. B. Ellis, A. B. Kahng // Journal of Combinatorial Theory. – 2001. – Vol. 100, N 2. – P. 232–249.
5. Rodl V. On a packing and covering problem / V. Rodl // European Journal of Combinatorics. – 1985. – Vol. 6. – P. 69–78.
6. Turan P. Reseach Problems / P. Turan // Magyar Tud. Acad. Mat. Kutato Int. Kozl. – 1961. – Vol. 6. – P. 417–423.

Кириченко Константин Дмитриевич, кандидат физико-математических наук, Иркутский государственный университет, 664011, Иркутск, ул. К. Маркса, 1, тел.: (3952)241097 (e-mail: constkir@gmail.com)

K. D. Kirichenko
On a Combinatorial Problem for the Set of Binary Vectors

Abstract. In this paper we introduce a new combinatorial problem for covering binary sets. This problem appears in connection to research of complexity of ESOP. Shannon function is called maximum from complexities of the shortest representation of each Boolean function. Hence the upper bound of the Shannon function guarantees the existence of the representation of any Boolean function with this complexity. It is important for applications.

As usual implicit algorithms of minimisation working with any Boolean function are used for defining the upper bound of Shannon function. Previously we have developed the algorithm of minimisation of Boolean functions in ESOPs which uses the combinatorial technique connected with tasks of finding covering and packing of binary sets. ESOP for given Boolean function is built by pattern which is described of non-singular matrix over the field Z_2 in that earch row and column matches any binary set. These binary sets should have the packing with density $1 + o(1)$ for getting the effective upper bound.

It is normal to use error-correcting linear codes for building a matrix of pattern. In this case Hamming code may by used. And so it lets use terms of the linear codes theory in definitions of combinatorial problems. In this paper we investigate a problem which belongs to covering and packing design. In doing so requirements to matrix impose several conditions to cover. In this work some of possible covers are introduced which have been described in terms of error-correcting linear codes.

Keywords: boolean function, ESOP, Hamming code, covering design.

References

1. Kirichenko K.D. Bounds of the minimization patterns' complexity of ESOP (in Russian). *IIGU Ser. Matematika*, 2009, vol. 2, pp. 67–76.
2. Bashov M.A., Selezneva C.N. On the length of functions of k-valued logics in modulo k ESOP (in Russian). *Diskretnaya matematika*, 2014, vol. 26, no 3, pp. 3–9.
3. Cooper J.N., Ellis R.B., Kahng A.B. Asymmetric binary covering codes. *Journal of Combinatorial Theory*, 2001, vol. 100, no 2, pp. 232–249.
4. MacWilliams F.J., Sloane N.J.A. The Theory of Error Correction Codes. *North-Holland Publishing Company*, 1977.
5. Rodl V. On a packing and covering problem. *European Journal of Combinatorics*, 1985, vol. 6, pp. 69–78.
6. Turan P. Reseach Problems. *Magyar Tud. Acad. Mat. Kutato Int. Kozl*, 1961, vol. 6, pp. 417–423.

Kirichenko Konstantin Dmitrievich, Candidate of Sciences (Physics and Mathematics), Irkutsk State University, 1, K. Marx st., Irkutsk, 664003, tel.: (3952)241097 (e-mail: `constkir@gmail.com`)