

# ЭФФЕКТИВНЫЕ МЕТОДЫ НАХОЖДЕНИЯ КРАТЧАЙШИХ РЕШЕНИЙ СИСТЕМ ЛИНЕЙНЫХ ЛОГИЧЕСКИХ УРАВНЕНИЙ<sup>1</sup>

А.Д. Закревский

Объединенный институт проблем информатики, г. Минск

Предложены эффективные методы, существенно сокращающие объем перебора, производимого при поиске кратчайшего решения. Приведены результаты их программной реализации, дано оценивание конкурентоспособности разработанных программ и указаны области их применимости.

## ВВЕДЕНИЕ

С задачей нахождения кратчайших решений систем линейных логических уравнений (СЛЛУ) автор встретился при разработке методов синтеза логических двухуровневых AND/EXOR-схем, обладающих некоторыми преимуществами в сравнении с традиционными AND/OR-схемами [1–3 и др.]. Например, их легче тестировать [4–6], и они оказываются более компактными при реализации симметрических булевых функций, типичных для арифметики [7]. Для решения задач их проектирования оказалось возможным привлечь хорошо развитую теорию линейных векторных пространств [8, 9]. Не менее важные приложения данная задача имеет в такой бурно развивающейся области, как защита информации [10].

В абстрактном виде задача формулируется следующим образом.

Пусть задана система из  $m$  линейных логических уравнений с  $n$  переменными:

$$Ax = y, \quad \bigoplus_{j=1}^n a_i^j x_j = y_i,$$

где  $A$  – булева  $m \times n$  матрица коэффициентов,  $x = (x_1, x_2, \dots, x_n)$  – булев вектор неизвестных и  $y = (y_1, y_2, \dots, y_m)$  – булев вектор свободных членов. Решением СЛЛУ называется такой набор зна-

чений неизвестных, который обращает каждое уравнение в тождество. В таком случае покомпонентная сумма по модулю два (называемая ниже просто суммой) тех столбцов матрицы  $A$ , соответствующие которым переменные принимают в решении значение 1, будет равна столбцу  $y$ .

Положим, что СЛЛУ имеет несколько решений (тогда система называется *неопределенной*, при этом обычно  $n > m$ ). Задача заключается в выборе из них *кратчайшего решения*, т. е. такого, которое представлено значением вектора  $x$  с минимальным числом единиц.

В принципе эту задачу можно решить тривиальным алгоритмом, который перебирает сочетания «из  $n$  по  $k$ » столбцов матрицы  $A$ , в порядке возрастания  $k$  и начиная с  $k = 1$ , подсчитывает их сумму и сравнивает с вектором  $y$ . Достаточно очевидно, что как только получаемая сумма оказывается равной  $y$ , рассматриваемая комбинация столбцов может быть принята за искомое решение. К сожалению, данный алгоритм слишком трудоемок, поскольку число анализируемых комбинаций столбцов очень быстро растет с ростом числа неизвестных  $n$ .

В следующих разделах излагаются более эффективные комбинаторные методы решения данной задачи, ориентированные на компьютерную реализацию.

## ЛЕСТНИЧНЫЙ АЛГОРИТМ

Оригинальный алгоритм поиска кратчайшего решения СЛЛУ был предложен при разработке ме-

<sup>1</sup> Работа доложена на конференции «Теория и практика логического управления», посвященной памяти М.А. Гаврилова (Москва, 10 ноября 2003 г.).







**Утверждение 1.**  $\alpha(m, n, k) = C_n^k 2^{-m}$ , где  $C_n^k$  – число сочетаний из  $n$  по  $k$ .

Аналогично, обозначив через  $\beta(m, n, k)$  математическое ожидание числа корней, вес которых не превышает  $k$ , получим

$$\text{Утверждение 2. } \beta(m, n, k) = \sum_{i=0}^k C_n^i 2^{-m}.$$

Искомый вес кратчайшего корня оценим максимальным значением параметра  $k$ , при котором  $\beta(m, n, k) < 1$ .

**Утверждение 3.**  $\gamma(m, n) = k$ , где  $\beta(m, n, k) < 1 \leq \beta(m, n, k+1)$ .

Например, в табл. 1 показаны значения величин  $\alpha$  и  $\beta$  для системы из 40 уравнений с 70-ю переменными, соответствующие значениям параметра  $k$  от 7 до 13. Достаточно очевидно, что вес кратчайшего решения, скорее всего, будет равен 10.

Оценивая объем вычислений при поиске кратчайшего решения, положим, что он пропорционален числу  $N$  рассматриваемых подмножеств столбцов остатка. Очевидно, значение  $N$  определяется уровнем перебора, который мы обозначим  $\delta(m, n)$ .

$$\text{Утверждение 4. } N(m, n) = \sum_{i=0}^{\delta(m, n)} C_{n-m}^i.$$

Достаточно очевидно, что объем перебора в сильной степени зависит от мощности остатка – числа столбцов в нем, которое равно разности  $n - m$ . В случае, когда это число оказывается меньшим веса кратчайшего корня, для гарантированного нахождения последнего приходится перебирать все  $2^{n-m}$  подмножества столбцов в остатке, что приводит к удвоению времени поиска решения при каждом увеличении разности  $n - m$  на единицу. Легко подсчитать, что при увеличении разности  $n - m$  на 10 время поиска возрастает примерно в тысячу раз.

Как показывают аналитические и экспериментальные (С++, PC COMPAS Presario – процессор Intel Pentium III, 1000 МГц) оценки эффективности рассмотренного метода, область его практического

приложения ограничивается условием  $n - m < 30$ . При этом условии время поиска не превышает одного часа, в то время как с дальнейшим ростом разности  $n - m$  расходы времени резко возрастают. Например, при  $m = 625$  и  $n = 700$  данный алгоритм будет вынужден перебирать и испытывать в поисках кратчайшего решения все  $2^{75}$  комбинаций наборов столбцов из остатка, и расчеты показывают, что на это уйдет три миллиарда четыреста миллионов лет!

## РАСПОЗНАВАНИЕ КОРОТКИХ РЕШЕНИЙ

Между тем, кратчайшее решение, вес которого  $\gamma(m, n)$  можно заранее оценить по формуле утверждения 3, будет содержать в остатке в среднем  $\gamma(m, n) \cdot (n - m)/n$  элементов, откуда следует, что оно, скорее всего, будет найдено на уровне перебора  $\delta(m, n) = \lfloor \gamma(m, n) \cdot (n - m)/n \rfloor$  – окаймляющие скобки означают «целое, ближе к снизу». В этом случае его можно *распознать* по весу и прекратить поиск, отказавшись от строгого доказательства оптимальности найденного решения и сэкономив тем самым время. При  $m = 625$  и  $n = 700$  ожидаемый вес  $\gamma$  кратчайшего решения равен 220, следовательно, уровень перебора  $\delta = \lfloor 220 \cdot 75/700 \rfloor = 23$ . На этом уровне решение будет найдено за 1 800 000 лет, примерно в 1900 раз быстрее, чем при полном переборе, однако недостаточно быстро с практической точки зрения. На более существенную экономию времени можно рассчитывать, удовлетворившись некоторым приближением к  $\gamma$  сверху.

Практический интерес представляет случай существования *короткого* решения СЛЛУ со случайной матрицей  $A$  и вектором  $y$ , равным сумме некоторых  $w$  столбцов этой матрицы (при этом  $w$  заметно меньше  $\gamma$ ). Например, при  $m = 625$ ,  $n = 700$  и  $w = 105$  уровень перебора  $\delta$  будет в среднем равен 11, откуда следует, что затраты времени на нахождение данного решения методом распознавания составят 193 дня. Однако, если повезет, уровень перебора может оказаться и ниже, что приведет к большой экономии времени. Так, при проведении эксперимента на компьютере указанное короткое решение было обнаружено на уровне перебора 10, что соответствует затратам времени в 25 дней.

## ДЕКОМПОЗИЦИОННЫЙ МЕТОД НЕПЕРЕСЕКАЮЩИХСЯ ОСТАТКОВ

При больших значениях разности  $n - m$  решение рассматриваемой задачи можно существенно ускорить, применяя декомпозиционный метод не-

Таблица 1

Значения  $\alpha$  и  $\beta$  при  $m = 40$ ,  $n = 70$  и  $k = 7 \dots 13$

$k$	$\alpha$	$\beta$
7	0,001	0,001
8	0,009	0,010
9	0,059	0,069
10	0,361	0,430
11	1,968	2,398
12	9,676	12,074
13	43,170	55,244

пересекающихся остатков, понижающий уровень перебора [17–19].

Предположим, что в матрице  $A$  можно найти  $q$  таких максимальных совокупностей линейно независимых столбцов, которым соответствуют взаимно непесекающиеся остатки. Заметим, что при этом должно выполняться отношение  $n \geq q(n - m)$ . Построим множество  $Q$ , составленное из  $q$  соответствующих канонических форм, базисы которых строятся на основе указанных совокупностей (очевидно, что это число ограничено отношением  $q \leq n/(n - m)$ ), и будем искать оптимальное решение в рамках этих форм, последовательно увеличивая уровень перебора.

*Утверждение 5.* В множестве  $Q$  найдется каноническая форма, где решение с весом  $v$  можно найти на уровне перебора не выше  $\lfloor v/q \rfloor$  – натурального числа, ближнего снизу к частному  $v/q$ .

*Утверждение 6.* Кратчайшее решение системы можно найти путем последовательного рассмотрения остатков канонических форм из множества  $Q$ , ограничивая уровень перебора подмножеств в этих остатках величиной  $\lfloor (v - 1)/q \rfloor$ , где  $v$  – вес кратчайшего из решений, найденных к текущему моменту времени.

Эти утверждения образуют основу предлагаемого метода. Обратим внимание на существенное понижение уровня перебора подмножеств в этом методе, в котором реализуется перебор подмножеств во всех остатках сначала на уровне 0, затем на уровне 1, и т. д., пока этот уровень не превысит  $\lfloor (v - 1)/q \rfloor$ .

*Утверждение 7.* Число  $N_r(m, n)$  подмножеств столбцов, рассматриваемых при методе непесекающихся остатков, определяется следующей формулой:

$$N_r(m, n) = q \sum_{i=0}^p C_{n-m}^i, \text{ где } p = \lfloor v/q \rfloor.$$

Предложенный метод позволяет существенно сократить время гарантированного нахождения кратчайшего решения, в сравнении с описанным выше методом Гаусса. Для того же примера ( $m = 625$ ,  $n = 700$ ) уровень перебора понизится с 220 до 23, как и у метода распознавания, с соответствующим сокращением времени поиска. Однако, в отличие от последнего, при этом гарантируется оптимальность получаемого решения.

Обратим внимание на одно весьма существенное для дальнейших рассуждений обстоятельство. Столбцы матрицы  $A$ , составляющие кратчайшее решение, распределяются по остаткам канонических форм случайно, т. е. не равномерно, а с некоторой дисперсией, благодаря чему найдется ос-

таток с минимальным числом  $k$  вошедших туда столбцов, заметно меньшим, чем  $\lfloor v/q \rfloor$  (обозначим через  $s$  номер этого остатка). Это позволяет, объединив методы декомпозиции и распознавания, найти кратчайшее решение на уровне перебора  $k$ .

Наибольший эффект такое объединение дает в случае существования короткого решения с весом  $w$ . Например, при  $m = 625$  и  $n = 700$  число канонических форм  $q$  равно 9. При анализе на компьютере примера с данными параметрами и  $w = 105$  оказалось, что  $k = 7$  и именно таким числом элементов короткого решения обладает остаток третьей формы ( $s = 3$ ). Расчеты по методу, предложенному в работе [20], показывают, что в данном случае решение будет найдено за 7,5 ч.

### ОБОБЩЕННЫЙ ДЕКОМПОЗИЦИОННЫЙ МЕТОД

Условие не пересечения остатков канонических форм существенно для метода, гарантирующего оптимальность находимых решений, но не для метода распознавания короткого решения, где более важно понизить уровень перебора. С учетом этого соображения декомпозиционный метод был обобщен следующим образом. На базе различных выборов из матрицы  $A$  максимальных множеств линейно независимых столбцов строится несколько (обозначим их число через  $p$ ) серий канонических форм, причем в каждой из этих серий данное условие не пересечения выполняется. Таким образом, число канонических форм увеличивается в два, три или более раз, что приводит к соответствующему росту времени на их анализ. Однако с большой вероятностью этот расход времени с лихвой компенсируется снижением значения параметра  $k$ , определяющего уровень производимого перебора.

В результате возникает выигрыш, хорошо иллюстрируемый на примере той же СЛЛУ с параметрами  $m = 625$ ,  $n = 700$ , обладающей коротким решением ( $w = 105$ ). Так, в эксперименте на компьютере при  $p = 2$  (две серии,  $q = 18$ ) решение находится на уровне перебора  $k = 6$  ( $s = 12$ ) примерно за три часа. А при  $p = 3$  (три серии,  $q = 27$ ) решение находится на уровне перебора  $k = 4$  ( $s = 12$ ), всего за 78 с!

### ЭКСПЕРИМЕНТЫ

Приведенный выше пример оказался весьма выигрышным для предлагаемых методов нахождения короткого решения СЛЛУ, демонстрируя снижение уровня перебора с 10 до 4, а времени решения – с 25 дней до 78 с.



Результаты испытания методов нахождения короткого решения случайных СЛЛУ

№	$p = 0$		$p = 1$		$p = 2$		$p = 3$	
	$q-s-k$	$t_0$	$q-s-k$	$t_1$	$q-s-k$	$t_2$	$q-s-k$	$t_3$
1	1-1-12	74 100 000	9-2-8	<b>177 000</b>	18-13-6	<b>11 200</b>	27-13-5	<b>996</b>
2	1-1-12	74 100 000	9-3-7	<b>27 200</b>	18-3-7	34 900	27-3-7	42 600
3	1-1-10	2 190 000	9-3-7	<b>27 200</b>	18-12-6	<b>10 400</b>	27-12-4	<b>78</b>
4	1-1-12	74 100 000	9-4-6	<b>3 410</b>	18-4-6	4 070	27-4-6	4 730
5	1-1-11	13 400 000	9-4-7	<b>34 900</b>	18-11-5	<b>807</b>	27-13-6	11 800
6	1-1-9	320 000	9-3-7	<b>27 200</b>	18-3-7	34 900	27-3-7	42 600
7	1-1-13	373 000 000	9-3-8	<b>243 000</b>	18-17-6	<b>14 300</b>	27-20-7	175 000
8	1-1-9	320 000	9-5-7	<b>42 700</b>	18-5-7	50 400	27-5-7	58 200
9	1-1-13	373 000 000	9-3-8	<b>243 000</b>	18-10-7	<b>89 200</b>	27-3-8	398 000
10	1-1-12	74 100 000	9-2-6	<b>1 840</b>	18-2-6	2 500	27-2-6	3 160
11	1-1-9	320 000	9-1-9	917 000	18-1-9	1 590 000	27-27-7	<b>229 000</b>
12	1-1-8	41 600	9-1-8	111 000	18-18-6	<b>15 100</b>	27-27-6	22 800
13	1-1-11	13 400 000	9-6-7	<b>50 400</b>	18-6-7	58 200	27-23-5	<b>1 670</b>
14	1-1-10	2 190 000	9-7-8	<b>507 000</b>	18-15-7	<b>128 000</b>	27-7-8	662 000
15	1-1-8	41 600	9-3-6	<b>2 620</b>	18-3-6	3 280	27-3-6	3 940
16	1-1-9	320 000	9-5-8	375 000	18-11-6	<b>9 580</b>	27-16-6	14 200
17	1-1-10	2 190 000	9-2-6	<b>1 840</b>	18-2-6	2 500	27-2-6	3 160
18	1-1-15	7 270 000 000	9-2-8	<b>177 000</b>	18-15-7	<b>128 000</b>	27-11-5	<b>10 200</b>
19	1-1-14	1 720 000 000	9-4-9	<b>2 390 000</b>	18-13-6	<b>11 200</b>	27-13-6	11 800
20	1-1-15	7 270 000 000	9-8-6	<b>6 560</b>	18-8-6	7 220	27-8-6	7 880
21	1-1-10	2 190 000	9-7-6	<b>5 780</b>	18-7-6	6 430	27-7-6	7 090
22	1-1-9	320 000	9-3-7	<b>27 200</b>	18-3-7	34 900	27-22-6	<b>18 900</b>
23	1-1-12	74 100 000	9-9-7	<b>73 700</b>	18-9-7	81 500	27-20-5	<b>1 470</b>
24	1-1-4	3	9-1-4	<b>8</b>	18-1-4	15	27-1-4	25
25	1-1-12	74 100 000	9-4-5	<b>284</b>	18-4-5	335	27-4-5	388
26	1-1-7	4 740	9-1-7	11 600	18-1-7	19 400	27-1-7	27 100
27	1-1-13	373 000 000	9-8-5	<b>554</b>	18-8-5	605	27-8-5	658
28	1-1-11	13 400 000	9-4-6	<b>3 410</b>	18-4-6	4 070	27-4-6	4 730
29	1-1-11	13 400 000	9-5-9	<b>2 880 000</b>	18-18-6	<b>15 100</b>	27-18-7	159 000
30	1-1-13	373 000 000	9-4-7	<b>34 900</b>	18-4-7	42 700	27-4-7	50 400

Более объективно эффективность данных методов характеризуется табл. 2, где показаны прогнозируемые результаты решения тридцати случайных СЛЛУ с теми же параметрами ( $m = 625$ ,  $n = 700$ ,  $w = 105$ ) четырьмя методами. Значением  $p = 0$  условно отмечен метод Гаусса, значением  $p = 1$  – декомпозиционный метод на одной серии канонических форм с непересекающимися остатками, и значениями  $p = 2$  и  $p = 3$  – обобщенные декомпозиционные методы на двух и трех сериях. Результаты применения этих методов отображаются тройками значений параметров  $q - s - k$  и временем  $t$  в секундах, затраченным на нахождение решения. Например, тройка 27-12-4 означает, что в процессе решения анализировалось 27 канонических форм и процесс успешно завершился при рассмотрении остатка формы 12 на уровне перебора 4. Полужирным шрифтом отмечены моменты, в которых рассматриваемый метод выигрывает по эф-

фективности у всех предшествующих (представленных столбцами слева).

Поскольку время приведено в секундах, заметим, что час содержит 3600 с, сутки – 86 400 с, год – 31 536 000 с. Так, пример 20 решается методом Гаусса ( $p = 0$ ) за 230 лет, а методом декомпозиции ( $p = 1$ ) – за два часа (в миллион раз быстрее).

### ЗАКЛЮЧЕНИЕ

Нахождение кратчайшего решения неопределенной системы линейных логических уравнений является ярким примером трудных комбинаторных задач. Для их успешного решения необходимо не только привлекать компьютеры, но и разрабатывать эффективные комбинаторные методы, учитывающие особенности рассматриваемых задач. Такие методы, предложенные в настоящей статье, на порядки (в тысячи раз и более) могут сокращать время, затрачиваемое на поиск кратчайших решений.

## ЛИТЕРАТУРА

1. Besslich P.W. Efficient computer method for XOR logic design // IEE Proc. on Computers & Digital Techniques. – 1983. – Vol. 130. – P. 203–206.
2. Sasao T. AND-EXOR expressions and their optimization // Logic Synthesis and Optimization / Kluwer Academic Publishers. – 1993. – P. 287–312.
3. Miller J.F., Thomson P. Highly efficient exhaustive search algorithm for optimizing canonical Reed-Muller expansions of boolean functions // Int. J. Electronics. – 1994. – Vol. 76. – P. 37–56.
4. Reddy S. M. Easily testable realizations for logic functions // IEEE Trans. – 1972. – C-21. – P. 1183–1188.
5. Debany W.H., Hartmann C.R.P., Sneathen T. J. Algorithm for generating optimal tests for exclusive-OR nets // IEE Proc. – 1991. – Vol. 138. – No 2. – P. 93–96.
6. Закревский А.Д., Закревский Л.А. Диагностирование схем, построенных из элементов "сумма по модулю 2" // Докл. АН Беларуси. – 1996. – Т. 10. – № 6. – С. 7–11.
7. Suprun V.P. Fixed polarity Reed-Muller expressions of symmetric Boolean functions // Proc. IFIP WG 10.5 Workshop on Applications of the Reed-Muller Expansion in Circuit Design. – Makuhari, Chiba, Japan, – 1995. – P. 246–249.
8. Кострикин А. И., Манин Ю. И. Линейная алгебра и геометрия. – М.: Наука, 1986.
9. Ланкастер П. Теория матриц. – М.: Наука, 1978.
10. Закревский А. Д. К решению систем линейных логических уравнений с искаженными правыми частями – условия корректной постановки задачи // Информ. системы и технологии (IST'2002): Матер. I Междунар. конф., Минск, 5–8 ноября 2002 г. – Минск 2002. – Ч. 2. С. 45–50.
11. Zakrevskij A.D. Looking for shortest solutions of linear logical equations: theory and applications in logic design // 2. Workshop «Boolesche Probleme», 19/20 Sept., 1996. – Freiberg / Sachsen. P. 63–69.
12. Закревский А.Д., Торопов Н.Р. Полиномиальная реализация частичных булевых функций и систем. – Минск: Ин-т техн. кибернетики НАН Беларуси, 2001.
13. Нильсон Н. Искусственный интеллект. Методы поиска решений. – М.: Мир, 1971.
14. Торопов Н.Р. Ускорение лестничного алгоритма решения системы линейных логических уравнений // Методы логического проектирования / ОИПИ НАН Беларуси. – Минск, 2003. – Вып. 2 – С. 103–114.
15. Gauss C.F. Beitrage zur Theorie der algebraischen Gleichungen. – Gött, 1849.
16. Закревский А.Д., Василькова И.В. Быстрый алгоритм нахождения кратчайшего решения системы линейных логических уравнений // Методы логического проектирования / ОИПИ НАН Беларуси. – Минск, 2002. – Вып. 1. – С. 5–12.
17. Закревский А.Д. Декомпозиционные методы нахождения кратчайшего решения системы линейных логических уравнений // Докл. НАН Беларуси. – 2003. – Т. 47. – № 2. – С. 58–60.
18. Закревский А.Д. Оптимизация решений в линейном булевом пространстве – методы декомпозиции // Автоматика и вычислительная техника. – 2003. – № 5. – С. 28–36.
19. Zakrevskij A.D., Zakrevski L.A. Optimizing solutions in a linear Boolean space – a decomposition method // Proc. of STI '2003, Orlando, Florida, USA, July 2003. P. 276–280.
20. Закревский А.Д., Василькова И.В. Прогнозирование затрат времени на выполнение комбинаторных алгоритмов // Методы логического проектирования / ОИПИ НАН Беларуси. – Минск, 2003. – Вып. 2.

E-mail: zakr@newman.bas-net.by

☎ (10-375-17) 288-65-99



## ПРАВИЛА ОФОРМЛЕНИЯ СТАТЕЙ ДЛЯ ПУБЛИКАЦИИ В ЖУРНАЛЕ «ПРОБЛЕМЫ УПРАВЛЕНИЯ»

Статья представляется в редакцию на бумаге в 2-х экземплярах, с аннотацией и направлением организации, а также обязательно в электронном виде на дискете 3,5 дюйма или по электронной почте (не более 2 МБ). Аннотация, название статьи и фамилии авторов должны быть представлены также на английском языке. Примерный объем оригинальной статьи 12, обзорной – 18 страниц текста. Текст печатается через 2 интервала с одной стороны бумаги формата А4, страницы нумеруются. В электронной форме текст должен быть в редакторе не ниже Word 97 шрифтом № 12 Times New Roman; текст не форматируется, т.е. не имеет табуляции, колонок и т.д. Рисунки должны иметь расширение, совместимое с Word 97, или в формате CorelDraw: фотографии должны быть предельно четкими, черно-белыми, на глянцева бумаге или в электронном виде TIFF с разрешением 300 dpi. Толщина линий рисунков, представляемых в электронной форме – не менее 3 пикселей.

Все буквенные обозначения, приведенные на рисунках, необходимо пояснять в основном или подрисуночном текстах (недопустимы повторные обозначения – в подрисуночных подписях и в тексте). Нумеровать следует только те формулы, на которые есть ссылка в последующем изложении. Список литературы (только органически связанной со статьей) составляется в порядке цитирования и дается в конце статьи. Ссылки на литературу в тексте отмечаются порядковыми номерами в квадратных скобках.

В конце статьи следует указать номер контактного телефона и электронный адрес. На отдельном листе дать сведения об авторах (полностью ф.и.о., ученая степень, должность и место работы, почтовый адрес и тел.)