Тема работы: Классификация ортогональных массивов OA(2048,14,2,7) и некоторых полностью регулярных кодов

Кротов Денис Станиславович, г.н.с. ИМ СО РАН, д.ф.-м.н.
(по совместительству преподаватель НГУ)

Аннотация

Описана классификация ортогональных массивов $\mathrm{OA}(2048, 14, 2, 7)$, или, что то же самое, полностью регулярных $\{14; 2\}$-кодов в двоичном 14-кубе (30848 классов эквивалентности). В частности, в исследуемом классе ортогональных массивов найден ровно один почти-$\mathrm{OA}(2048, 14, 2, 7{+}1)$ с точностью до эквивалентности. Как производные объекты, классифицированы также $\mathrm{OA}(1024, 13, 2, 6)$ (202917 классов) и полностью регулярные $\{12, 2; 2, 12\}$- и $\{14, 12, 2;\ 2, 12, 14\}$-коды в 13- и 14-кубах соответственно.

Работа выполнена за счет гранта РНФ 22-11-00266 «Полностью регулярные коды как решения экстремальных задач комбинаторики» 2022-2024, рук. Д. С. Кротов.

1. Постановка задачи

Введем некоторые понятия и обозначения. n-куб Q_n — это граф, множество вершин которого представляет собой множество $\{0, 1\}^n$ бинарных слов длины n, образующее векторное пространство над $\mathbb{F}_2$; два таких слова смежны в Q_n тогда и только тогда, когда они различаются ровно в одной позиции. Двоичный *ортогональный массив* $\mathrm{OA}(N, n, 2, t)$ это такое мультимножество мощности N вершин n-куба Q_n, что каждый подграф, изоморфный Q_{n-t}, содержит $N/2^t$ слов из массива.

Цель работы — классификация с точностью эквивалентности ортогональных массивов с параметрами $\mathrm{OA}(2048, 14, 2, 7)$, достигающих границу Фридмана [5]

$$N \geq 2^n\Big(1 - \frac{n}{2(t+1)}\Big) \qquad (1)$$

и, следовательно, являющихся простыми (без кратных элементов) и притом полностью регулярными кодами (см. определение ниже). Как производные объекты из основного рассматриваемого класса, классифицированы также ортогональные массивы $\mathrm{OA}(1024, 13, 2, 6)$ (202917 классов эквивалентности) и полностью регулярные $\{12, 2; 2, 12\}$- и $\{14, 12, 2; 2, 12, 14\}$-коды в 13- и 14-кубах соответственно.

Множество C вершин (код) графа G называется *полностью регулярным* с *радиусом покрытия* ρ и *массивом пересечений* $\{b_0, b_1, ..., b_{\rho-1}; c_1, ..., c_\rho\}$, если разбиение $(C = C^{(0)}, C^{(1)}, ..., C^{(\rho)})$ множества вершин по расстоянию от кода C удовлетворяет следующему условию: каждая вершина из $C^{(i)}$ имеет ровно b_i соседей в $C^{(i+1))}$ и c_i соседей в $C^{(i-1)}$ (подразумевается, что $b_\rho = c_0 = 0$ и $C^{(-1)} = C^{(\rho+1)} = \emptyset$), см., например, обзор [1].

2. Современное состояние проблемы

Ортогональные массивы представляют собой комбинаторные структуры, важные как для практических приложений, таких как планирование экспериментов или тестирование программного обеспечения, так и теоретически из-за множества связей с теорией кодирования, криптографией, теорией комбинаторных дизайнов и т.д., см., например, [6]. Классификация ортогональных массивов по заданным параметрам — проблема, привлекающая внимание многих исследователей, см., например, [2], [3], [17], [15] и библиографию там. Основным результатом настоящей работы является классификация ортогональных массивов OA$(2048, 14, 2, 7)$, мощность которых является наибольшей среди всех ортогональных массивов, когда-либо классифицированных вычислительным путем.

Двоичные ортогональные массивы с малыми параметрами, лежащими на границе Фридмана, были классифицированы в серии работ, см. Таблицу 1 (включены только целочисленные параметры с $t < n - 1$ и удовлетворяющие необходимому условию $t \le 2n/3 - 1$ [4]). Классификация OA$(1024, 12, 2, 7)$, OA$(1536, 13, 2, 7)$ и OA$(2048, 15, 2, 7)$ по существу основывалась на том, что такие массивы полностью регулярные коды (в частности, OA$(2048, 15, 2, 7)$ — 1-совершенные коды). В настоящей работе развивается техника классификации полностью регулярных кодов для характеризации всех ортогональных массивов OA$(2048, 14, 2, 7)$. В частности, используется решатель ILP (целочисленное линейное программирование) для ускорения классификации промежуточных объектов, называемых локальными кодами, класс всех OA$(2048, 14, 2, 7)$ разделяется на два подкласса в зависимости от наличия специальной подконфигурации, специфичной для рассматриваемых параметров.

3. Описание работы, включая используемые алгоритмы

Для краткости будем называть полностью регулярные коды с массивом пересечений $\{14; 2\}$ (по существу, ортогональные массивы OA$(2048, 14, 2, 7)$) $\{14; 2\}$-кодами. Классификационный подход основан на концепции локальных кодов. Будем говорить, что множество $P \subset \{0, 1\}^{14}$ является *r-локальным кодом*, если

(I) (*условие локальности*) P состоит из слов веса $\le r$;

(II) (*условие «с точностью до эквивалентности»*) $\bar{0}$ не принадлежит P;

параметры	число классов	параметры	число классов
OA$(2,3,2,1)$	1	OA$(1,2,2,0)$	1
OA$(16,6,2,3)$	1	OA$(8,5,2,2)$	1
OA$(16,7,2,3)$	1	OA$(8,6,2,2)$	1
OA$(128,9,2,5)$	2, [9]	OA$(64,8,2,4)$	3, [17]
OA$(1024,12,2,7)$	16, [12]	OA$(512,11,2,6)$	37, [12]
OA$(1536,13,2,7)$	1, [11]	OA$(768,12,2,6)$	3, [11]
OA$(2048,14,2,7)$	30848	OA$(1024,13,2,6)$	202917
OA$(2048,15,2,7)$	5983, [14]	OA$(1024,14,2,6)$	38408, [14]
OA$(8192,15,2,9)$	?	OA$(4096,14,2,8)$	?

Таблица 1: Малые параметры массивов OA$(N,n,2,t)$ и OA$(N/2,n-1,2,t-1)$, $N = 2^n(1-n/2(t+1))$, $t \le 2n/3-1$

(III) (*условие точного покрытия*) окрестность каждой вершины $\bar v$ веса меньше r удовлетворяет локальному условию из определения $\{14;2\}$-кода: если $\bar v \in P$, то $\bar v$ не имеет соседей в P; если $\bar v \notin P$, то $\bar v$ имеет ровно 2 соседей в P;

(IV) (*условие граничного неравенства*) каждое слово в $\{0,1\}^{14}$ имеет не более 2 соседей в P.

Легко видеть, что с точностью до эквивалентности (это и мотивирует название условия (II)) каждый $\{14;2\}$-код C включает в себя r-локальный код R, который определяется как набор кодовых слов кода C веса не более r (т.е. (I) автоматически выполняется). Действительно, (III) выполняется по определению полностью регулярного кода и потому, что все C-соседи вершины $\bar v$ веса $< r$ находятся в R. Мы не можем сказать то же самое о вершине v "пограничного" веса r или $r+1$; однако наш массив пересечений гарантирует, что такая вершина не может иметь более 2 соседей из кода, что дает (IV) (для другого массива пересечений можно найти другие граничные неравенства, чтобы уменьшить количество локальных кодов, см., например, [10]). Наконец, если (II) не выполняется для кода C, то оно выполняется для некоторого кода, эквивалентного коду C.

Непосредственно из определений мы имеем следующий ключевой факт.

Лемма 1. *Любой $\{14;2\}$-код, не содержащий $\bar 0$, является 14-локальным кодом. В частности, любой $\{14;2\}$-код эквивалентен 14-локальному коду. Если C есть r'-локальный код и R состоит из всех кодовых слов кода C веса не более r, где $0 \le r < r'$, то R является r-локальным кодом.*

В последнем случае мы говорим, что C является *r'-продолжением* кода R.

Если для каждого r-локального кода можно построить все его $(r+1)$-продолжения, то все локальные коды и, наконец, $\{14;2\}$-коды можно классифицировать рекурсивно. Однако оказывается, что переход от 2-локального к 3-локальному коду не завершается за разумное вычислительное время. Чтобы решить эту проблему, мы добавляем промежуточный шаг на основе $(2,3)$-кодов (детали в данном отчете опущены и могут быть найдены в публикации).

Теперь опишем общий алгоритм классификации.

- Начнем с представителей 5 классов эквивалентности $(2,2)$-локальных кодов, которые можно найти вручную.
- Для каждого из 5 представителей из предыдущего шага: построить все $(2,3)$-продолжения. На данном шаге решается соответствующая задача кратного точного покрытия, при помощи библиотеки `libexact` [8].
- Найденные $(2,3)$-продолжения классифицировать с точностью до эквивалентности, сохраняя представителей каждого класса эквивалентности. Этот шаг производится при помощи матобеспечения `nauty&traces` [13] по распознаванию изоморфности графов.
- Подтвердить вычисления с помощью теоремы об орбитах и стабилизаторах. Важный этап проверки (см. [7, §10.2]), исключающий большинство ошибок вычислений.
- Для каждого из представителей предыдущего шага построить все $(3,3)$-продолжения; классифицировать их до эквивалентности; подтвердить.
- Классифицировать найденные $(3,3)$-локальные коды с точностью до эквивалентности как 3-локальные коды.
- Для каждого из представителей $(i-1)$-локальных кодов, $i = 4, 5, \ldots, 14$, начиная с шага выше, построим все i-продолжения; классифицировать их до эквивалентности; подтвердить. Аналогично предыдущим шагам, для каждого i при помощи `libexact` [8] находим все продолжения, потом при помощи `nauty&traces` [13] выбираем только неэквивалентные решения. Исключением является классификация 4-локальных кодов, когда при помощи решателя *BOP* целочисленного линейного программирования (ILP) из библиотеки *Google OR-Tools* [16] выбираются все 3-коды, которые имеют хотя бы одно 4-продолжение, а уже затем при помощи `libexact` находятся все 4-продолжения (это привело к значительному ускорению на самом трудоёмком этапе, так как большинство 3-кодов было отбраковано гораздо быстрее, чем это делает `libexact`).

Наконец, мы находим все неэквивалентные $\{14;2\}$-коды. Перебор удалось значительно сократить за счет разбиения классификации всех $\{14;2\}$-кодов на два класса: содержащих 4-цикл (квадрат) и несодержащих (бесквадратных). В первом случае мы начинаем с 2-кодов с квадратом, во втором случае избегаем квадратов на каждом этапе.

4. Полученные результаты

Приведем сначала промежуточные результаты вычислений, а затем окончательные в виде теорем. Число классов эквивалентности локальных кодов, найденных как промежуточные результаты, показано в Таблице 2; поиск $(3,3)$-локальных кодов занял 339 $(14+59+33+37+196)$ ядро-дней, 3-локальных кодов 124 $(1+18+14+6+85)$ дни, продолжаемых 3-локальных кодов 1293 $(14+464+115+50+650)$ дни, 4-, 5- и

6-локальных кодов 50, $(7+44)$, 115, $(7+44)$, и 112 $(7+44)$ ядро-дней соответственно.

	$L_{4,4,4}$, $L_{4,8}$	$L_{5,7}$, $L_{6,6}$, L_{12} (*: бесквадратные)
$(2,3)$-локал.	$14+59$	$33+37+196$
$(3,3)$-локал.	$73762927+1586116921$	$1280242055+543652569+7755763093$
3-локал.: *все,*	$36904735+793121035$	$640150181+271854554+3877947089$
бесквадратные,		$166208491^*+71966561^*+1014622649^*$
продолжаемые	$17044+78904$	$25^*+30^*+679^*$
4-локал.	$4753786+29233429$	$9^*+0^*+117^*$
5-локал.	$15286921+16399650$	$9^*+0^*+101^*$
6-локал.	$1688762+3410955$	$9^*+0^*+101^*$

Таблица 2: Промежуточные результаты вычислений

Теорема 1 (вычислительные результаты)**.** *Имеется* 30848 *классов эквивалентности ортогональных массивов* $OA(2048,14,2,7)$*; (эквивалентно, полностью регулярные коды в* $H(14,2)$ *с фактор-матрицей* $[[0,14],[2,12]]$*); восемь из них бесквадратны. Из них* 14960 *(4 без квадратов) являются проколотыми* 1*-совершенными кодами, а остальные* 15888 *(4 без квадратов) – нет. Общее количество различных ортогональных массивов* $OA(2048,14,2,7)$ *равно* 541012580165257200 *(*267743838601839600 *проколотых* 1*-совершенных кодов).*

Теорема 2 (вычислительные результаты)**.** *Существует* 202917 *классов эквивалентности ортогональных массивов* $OA(1024,13,2,6)$ *(эквивалентно, совершенные раскраски* $H(13,2)$ *с фактор-матрицей* $[[0,1,12],[1,0,12],[2,2,9]]$*). Из них* 100473 *— это выколотые укороченные* 1*-совершенные коды.*

Теорема 3 (вычислительные результаты)**.** *В* Q_{13} *имеется* 247904 *классов эквивалентности (общее количество* 541012580165257200*) полностью регулярных кодов с массивом пересечений* $\{12,2;2,12\}$*. В* Q_{14} *имеется* 36137 *классов эквивалентности полностью регулярных кодов с массивом пересечений* $\{14,12,2;2,12,14\}$.

5. Иллюстрации, визуализация результатов

В качестве иллюстрации приведем расширенный совершенный код P' в Q_{16}, из которого укорачиванием по одной координате и затем выкалыванием по другой координате получается уникальный ортогональный массив OA$(2048,14,2,7)$, единственный из 30848, который можно назвать почти-OA$(2048,14,2,7{+}1)$. Код может быть задан подгруппой группы автоморфизмов, которая регулярным действием порождает весь код как орбиту нулевого вектора, см. Таблицу 3.

6. Благодарности и публикации

Кластер ИВЦ НГУ помог осуществить исследование, потребовавшее несколько лет процессорного времени, что затруднительно осуществить на персональных ЭВМ. Ав-

$[\,00000000\,11111111,\ \mathrm{Id}\,]$, $[\,11111111\,00000000,\ \mathrm{Id}\,]$,

$[\,00001111\,00000011, (89)(ab)(cd)(ef)\,]$, $[\,00000011\,00001111, (01)(23)(45)(67)\,]$,

$[\,11000000\,11000000, (45)(67)(cd)(ef)\,]$,

$[\,00110011\,00000101, (8a)(9b)(ce)(df)\,]$, $[\,00000101\,00110011, (02)(13)(46)(57)\,]$,

$[\,01010101\,00010001, (8c)(9d)(ae)(bf)\,]$, $[\,00010001\,01010101, (04)(15)(26)(37)\,]$,

$[\,00001100\,10100000, (23)(67)(ce)(df)\,]$, $[\,10100000\,00001100, (46)(57)(ab)(ef)\,]$.

Таблица 3: Генераторы группы автоморфизмов, регулярно действующей на расширенном 1-совершенном коде P'. Каждый автоморфизм записывается в виде $[\bar{v},\pi]$, где $\bar{v}$ — вектор сдвига, а π — координатная перестановка (координаты представлены шестнадцатеричными цифрами; перестановки также указаны стрелками); действие $[\bar{v},\pi]$ на вершинах Q_{16} равно $\bar{x} \to \bar{v} + \pi(\bar{x})$.

тор проекта благодарен ИВЦ НГУ за предоставленные вычислительные ресурсы и лично Владиславу Анатольевичу Калюжному за оперативное решение всех вопросов.

По результатам вычислительных исследований опубликована статья. Кроме того проведены смежные исследования по поиску полностью регулярных кодов с другими параметрами, в которых использовалась похожая техника, но значительно меньшее вычислительного ресурса; опубликован препринт.

- D. S. Krotov, "The classification of orthogonal arrays OA(2048,14,2,7) and some completely regular codes", Discrete Mathematics, 347:5 (2024), 113923 , 8 pp. https://doi.org/10.1016/j.disc.2024.113923 (arXiv.org preprint https://arxiv.org/abs/2311.05428).
- D. S. Krotov. On the existence of some completely regular codes in Hamming graphs. ArXiv.org preprint, 2023. https://arxiv.org/abs/2312.08360

Список литературы

1. J. Borges, J. Rifà, and V. A. Zinoviev. On completely regular codes. *Probl. Inf. Transm.*, 55(1):1–45, Jan. 2019. https://doi.org/10.1134/S0032946019010010.

2. P. Boyvalenkov, T. Marinova, and M. Stoyanova. Nonexistence of a few binary orthogonal arrays. *Discrete Appl. Math.*, 217(2):144–150, Jan. 2017. https://doi.org/10.1016/j.dam.2016.07.023.

3. D. A. Bulutoglu and K. J. Ryan. Integer programming for classifying orthogonal arrays. *Australas. J. Comb.*, 70(3):362–385, 2018.

4. D. G. Fon-Der-Flaass. A bound on correlation immunity. *Sib. Èlektron. Mat. Izv.*, 4:133–135, 2007. Online: http://mi.mathnet.ru/eng/semr149.

5. J. Friedman. On the bit extraction problem. In *Foundations of Computer Science, IEEE Annual Symposium on*, pages 314–319, Los Alamitos, CA, USA, 1992. IEEE Computer Society. https://doi.org/10.1109/SFCS.1992.267760.

6. A. S. Hedayat, N. J. A. Sloane, and J. Stufken. *Orthogonal Arrays. Theory and Applications.* Springer Series in Statistics. Springer, New York, NY, 1999. https://doi.org/10.1007/978-1-4612-1478-6.

7. P. Kaski and P. R. J. Östergård. *Classification Algorithms for Codes and Designs*, volume 15 of *Algorithms Comput. Math.* Springer, Berlin, 2006. https://doi.org/10.1007/3-540-28991-7.

8. P. Kaski and O. Pottonen. libexact user's guide, version 1.0. Technical Report 2008-1, Helsinki Institute for Information Technology HIIT, 2008.

9. D. Kirienko. On new infinite family of high order correlation immune unbalanced Boolean functions. In *Proceedings 2002 IEEE International Symposium on Information Theory, Lausanne, Switzerland, June 30 – July 5, 2002*, page 465. IEEE, 2002. https://doi.org/10.1109/ISIT.2002.1023737.

10. D. S. Krotov. Equitable $[[2,10],[6,6]]$-partitions of the 12-cube. E-print 2012.00038, arXiv.org, 2020. Available at `http://arxiv.org/abs/2012.00038`.

11. D. S. Krotov. On the OA(1536,13,2,7) and related orthogonal arrays. *Discrete Math.*, 343(2):111659/1–11, 2020. https://doi.org/10.1016/j.disc.2019.111659.

12. D. S. Krotov and K. V. Vorob'ev. On unbalanced Boolean functions with best correlation immunity. *Electr. J. Comb.*, 27(1):#P1.45(1–24), 2020. https://doi.org/10.37236/8557.

13. B. D. McKay and A. Piperno. Practical graph isomorphism, II. *J. Symb. Comput.*, 60:94–112, 2014. https://doi.org/10.1016/j.jsc.2013.09.003.

14. P. R. J. Östergård and O. Pottonen. The perfect binary one-error-correcting codes of length 15: Part I—classification. *IEEE Trans. Inf. Theory*, 55(10):4657–4660, 2009. https://doi.org/10.1109/TIT.2009.2027525.

15. S. Pang, J. Wang, D. K. J. Lin, and M.-Q. Liu. Construction of mixed orthogonal arrays with high strength. *Ann. Stat.*, 49(5):2870–2884, 2021. https://doi.org/10.1214/21-AOS2063.

16. L. Perron and V. Furnon. OR-Tools, v9.8. `https://developers.google.com/optimization/`.

17. E. D. Schoen, P. T. Eendebak, and M. V. M. Nguyen. Complete enumeration of pure-level and mixed-level orthogonal arrays. *J. Comb. Des.*, 18(2):123–140, 2010. https://doi.org/10.1002/jcd.20236.