ИЗВЕСТИЯ РАН. ТЕОРИЯ И СИСТЕМЫ УПРАВЛЕНИЯ, 2023, № 5, с. 67–77

КОМПЬЮТЕРНЫЕ МЕТОДЫ

УДК 519.11, 519.171.2, 519.175.3, 519.179.1

ОБЪЕДИНЕНИЕ БАЗ И ОЦЕНКА МНОЖЕСТВА ЭКСТРЕМАЛЬНЫХ 3-ОДНОРОДНЫХ ГИПЕРГРАФОВ[1]

И. С. Берецкий[a], Е. К. Егорова[a,b], А. В. Мокряков[a,c,*], В. И. Цурков[b]

[a] *МАИ (национальный исследовательский ун-т), Москва, Россия*

[b] *ФИЦ ИУ РАН, Москва, Россия*

[c] *РГУ им. А.Н. Косыгина (Технологии. Дизайн. Искусство), Москва, Россия*

**e-mail: MokryakovAlVik@ya.ru*

Поступила в редакцию 25.04.2023 г.
После доработки 10.05.2023 г.
Принята к публикации 05.06.2023 г.

Основная проблема, связанная с гиперграфами, — это их храненение. Если гиперграф без особенностей, то для его описания часто используются разряженные матрицы. Для работы с k-однородными гиперграфами часто применяют матрицы смежности, однако они занимают большое место в памяти компьютера и в целом хранение k-мерных массивов не очень удобно. Здесь предлагается одно решение для описания и хранения экстремальных k-однородных гиперграфов. Это база — уникальная характеристика экстремального гиперграфа, которая однозначно его описывает. Кроме того, базы можно применять для поиска мощности экстремальных k-однордных гиперграфов. Нами представлены алгоритмы перечисления баз и представлена гипотеза об аналитическом виде формул, описывающих мощность множества экстремальных k-однородных гиперграфов. Для данной задачи используется операция объединения баз, также введенная здесь.

DOI: 10.31857/S0002338823050037, **EDN:** TEEFBU

Введение. Гиперграфы начиная с XXI в. активно применяются для анализа сетей [1, 2], признаков и характеристик принадлежности (таких, как теги, свойства, необязательные параметры) [3]. Графы и гиперграфы широко используются как в современных фундаментальных исследованиях, так и при решении практических задач. Однако большинство алгоритмов на гиперграфах оперирует специальными классами гиперграфов [4]. Характер ограничений, накладываемых на данные классы, позволяет точнее описывать модели анализируемых данных.

При этом нужно учитывать, что сами гиперграфы суть аморфные структуры, имеющие слабые общие черты, что привело к построению класса однородных гиперграфов. Это дает возможность применять некий общий аппарат при их исследовании. Данную структуру независимо друг от друга начали разрабатывать как топологи [5], так и исследователи в области обобщения теории графов [6–13]. Большое количество статей по раскраскам гиперграфов показывает актуальность фундаментальных исследований. С другой стороны, имеется широкий спектр работ, связанный с однородными гиперграфами [14–21], где показывается их применимость в разделах управления сетями и базами данных, шифровании и моделировании.

Главными ключевыми задачами в области комбинаторной топологии и теории гиперграфов всегда были вопросы подсчета. Здесь стоит упомянуть как давние основополагающие работы, например [22, 23], так и современные [24–26]. Сложные методы подсчета зачастую опираются на специфические структуры-представления графов и гиперграфов, которые позволяют взглянуть на них с иной стороны [27, 28].

В статье продолжается исследование базы — структуры, описанной в [28], а именно расширяется понятие базы на k-мерный случай с построением операции объединения двух баз и реализации алгоритма подсчета баз.

1. Экстремальные однородные гиперграфы и базы. Для работы с базами необходимо дать определение гиперграфам, экстремальности и самим базам.

[1] Работа выполнена при финансовой поддержке РФФИ (проект № 21-51-53019 ГФЕН_а).

О п р е д е л е н и е 1. *Гиперграф* $H(V_n, E)$ – совокупность множества V_n из n вершин и множества непустых подмножеств множества вершин E [4], e_i – элементы множества E, где $i = \overline{1,|E|}$ – *гиперребра*.

Гиперграф $H^k = H^k(V_n, E)$ называют *k-однородным* гиперграфом (uniform hypergraph (UH)), если $|e_i| = k$, $\forall e_i \in E$.

В различных отраслях математики частные случаи данных объектов встречаются под иными именами: для $k = 2$ данная конструкция соответсвует ненаправленному графу, не имеющему петель. В топологии данные структуры именуются комплексами с размерностью $k - 1$ [5, 28].

Наиболее точно структуру UH описывает k-индексная матрица смежности $X^k(H^k) = (x_{i_1 \ldots i_k})$, где индексы $i_j = \overline{1,n}$ при $j = \overline{1,k}$. Такая структура явлется универсальной, но с точки зрения хранения матрицы смежности в памяти.

Теперь рассмотрим свойство экстремальности. Его можно найти несколькими способами, например через матрицу смежности. Опишем понятия доменов нулей и единиц, которые нам полезны в дальнейшем определении.

О п р е д е л е н и е 2. *Доменом нулей* – $D_0(X_n^k)$ (*единиц* – $D_1(X_n^k)$) называется множество ячеек матрицы смежности с попарно неравными индексами, равными нулю (единице), при этом не существует рассматриваемой единичной (нулевой) ячейки, отличающейся от пары ячеек из D_0 (D_1) только одним индексом и находящейся между ними.

Теперь воспользуемся данным опеределением в задании особого экстремального класса UH.

О п р е д е л е н и е 3. Однородный гиперграф H_n^k называется *экстремальным*, если для соответствующей ему матрицы смежности $X_n^k(H_n^k)$ выполняется правило разделения доменов нулей и единиц, состоящее из следующих условий:

1) все единичные ячейки образуют домен единиц;

2) все нулевые ячейки образуют домен нулей;

3) ячейка $x_{1\ldots k} \in D_1(X_n^k)$ или $D_1(X_n^k) = \varnothing$;

4) ячейка $x_{n-k+1\ldots n} \in D_0(X_n^k)$ или $D_0(X_n^k) = \varnothing$.

В дальнейшем экстремальный UH_n^k будем обозначать через EUH_n^k.

Другими словами, правило разделения доменов нулей и единиц можно сформулировать следующим образом: домен единиц сгруппирован около ячейки с минимальными индексами, а домен нулей – около ячейки с максимальными индексами, и между ними находится граница, которая изолирует нули и единицы в своих доменах.

Особый интерес представляет описание данной границы между доменами. Одним из способов ее описания является база.

2. База и критерий экстремальности UH_n^k. Пусть $n,\ k \in \mathbb{Z},\ k \ge 1$. Для множества индексов $I_n = \{i \in \mathbb{Z} : 1 \le i \le n\}$ (номеров) введем обозначение k-индексных упорядоченных подмножеств множества I_n: $I_n^k = \{(i_1, \ldots i_k) : 1 \le i_j \le n,\ i_j < i_{j+1}\ \forall j\}$. Определим на I_n^k частичный порядок: положим $(i_j : 1 \le j \le k) \ge (m_j : 1 \le j \le k)$, если $i_j \ge m_j\ \ \forall j$, и $(i_j : 1 \le j \le k) > (m_j : 1 \le j \le k)$ при $(i_j : 1 \le j \le k) \ge (m_j : 1 \le j \le k)$ и $(i_j : 1 \le j \le k) \ne (m_j : 1 \le j \le k)$.

Теперь построим конструкцию, позволяющую алгебраическим способом описать экстремальный EUH_n^k. Для $H^k = (x_{i_1 \ldots i_k}) = H^k(V_n, E)$ положим $I_n^k(H^k) = \left\{(i_1, \ldots, i_k) \in I_n^k : x_{i_1 \ldots i_k} = 1\right\} = \{(i_1, \ldots, i_k) \in I_n^k : \{u_{i_j} : 1 \le j \le k\} \in E\}$.

Определение 4. Пусть $H^k = (x_{i_1 \ldots i_k}) = H^k(V_n, E)$ – не пустой EUH_n^k $(E \ne \varnothing)$. Подмножество индексов $\overline{I}_n^k(H^k) = \{(i_1, \ldots, i_k)\}$ из I_n^k называется базой для EUH_n^k, если выполняются следующие условия:

а) для разных элементов $(i_1, \ldots, i_k)$ и $(m_1, \ldots, m_k)$ из $\overline{I}_n^k(H^k)$ отношение порядка не определено;

б) для $\forall(i_1,\ldots,i_k)\in I_n^k(H^k)$, где $(i_1,\ldots,i_k)\not\in \overline{I}_n^k(H^k)$, существует такой $(m_1,\ldots,m_k)\in \overline{I}_n^k(H^k)$, что $(i_1,\ldots,i_k)<(m_1,\ldots,m_k)$.

Рассмотрим на следующем примере базу.

П р и м е р 1. Пусть дан $H_7^4 = H^4(V_7,E))$, где множество гиперребер $S=\{\{v_1,v_2,v_3,v_4\}$, $\{v_1,v_2,v_3,v_5\}$, $\{v_1,v_2,v_3,v_6\}$, $\{v_1,v_2,v_3,v_7\}\}$. Следовательно, $\overline{I}_7^4(H^4)=\{(1,2,3,7)\}$.

Теперь зададим понятие максимального подмножества индексов.

О п р е д е л е н и е 5. Подмножество индексов $(m_1,\ldots,m_k)$ из $I_n^k(H^k)$ (т.е. $x_{m_1\ldots m_k}=1$) называется максимальным, если $x_{i_1\ldots i_k}=0$, $\forall(i_1,\ldots,i_k)>(m_1,\ldots,m_k)$.

Это определение означает следующее: для каждой тройки из $I_n^k(H^k)$ существует максимальная тройка из этого же множества.

Т е о р е м а 1. Любой EUH $H^k = H^k(V_n,E)=(x_{i_1\ldots i_k})$ имеет единственную базу и эта база есть $\overline{I}_n^k(H^k)$.

Доказательство данной теоремы для 3-мерного случая приведено в [28]. Оно легко расширяется на k-мерный случай.

Теперь можно ввести принцип построения базы на основе максимального подмножества.

Т е о р е м а 2. Пусть $\tilde{I}_n^k=\{(i_1,\ldots,i_k)\in I_n^k$: любые пары различных подмножеств индексов не связаны отношением порядка из $I_n^k\}$. Данное множество подмножеств является максимальным. Тогда $\tilde{I}_n^k$ – база некоторого EUH$_n^k$.

П р и м е р 2. Пусть $\overline{I}_8^4(H^4)=\{(1,2,5,7),(2,3,5,6)\}$ – база некоторого экстремального 4-однородного гиперграфа $H_8^4=H^4(V_8,E)$. Тогда $E=\{\{v_1,v_2,v_3,v_4\}$, $\{v_1,v_2,v_3,v_5\}$, $\{v_1,v_2,v_3,v_6\}$, $\{v_1,v_2,v_3,v_7\}$, $\{v_1,v_2,v_4,v_5\}$, $\{v_1,v_2,v_4,v_6\}$, $\{v_1,v_2,v_4,v_7\}$, $\{v_1,v_2,v_5,v_6\}$, $\{v_1,v_2,v_5,v_7\}$, $\{v_1,v_3,v_4,v_5\}$, $\{v_1,v_3,v_4,v_6\}$, $\{v_1,v_3,v_5,v_6\}$, $\{v_2,v_3,v_4,v_5\}$, $\{v_2,v_3,v_4,v_6\}$, $\{v_2,v_3,v_5,v_6\}\}$.

Из теорем 1 и 2 вытекает следующая теорема.

Т е о р е м а 3. Однородный гиперграф H_n^k является экстремальным тогда и только тогда, когда H^k имеет базу.

Доказательство последних теорем очевидно и не вызывает затруднений.

П р и м е р 3. Пусть $n=5$ и $\tilde{I}_5=\{(1,2,5),(2,3,4)\}$. Легко видеть, что 2-комплекс $H_5^3=H^3(V_5,E)$, где $E=\{\{v_1,v_2,v_3\}$, $\{v_1,v_2,v_4\}$, $\{v_1,v_2,v_5\}$, $\{v_1,v_3,v_4\}$, $\{v_1,v_3,v_5\}$, $\{v_1,v_4,v_5\}$, $\{v_2,v_3,v_4)\}$ задается базой $\tilde{I}_5$.

3. Объединение баз. Базу можно задать также через границу, разделяющую домены D_1 и D_0.

Рассмотрим матрицу смежности, которая представляет собой k-мерный куб, касающийся начала координат внешним углом ячейки с минимальнми индексами. Оси координат соответствуют росту индексов, и без потери общности можно считать, что порядок осей такой же, как порядок индексов матрицы. Границу можно представить в виде множества точек, каждая из которых имеет локальные минимум или максимум расстояния до начала координат. При этом нужно понимать, что такая поверхность имеет ряд линий симметрии, каждая из которых делит поверхность пополам. Нас интересует область, в которой $i_1<i_2<\ldots<i_{k-1}<i_k$.

С учетом вышесказанного граница между доменами является k-мерной поверхностью, которая представляет собой объединение внешней границы (по отношению к началу координат) множества k-мерных параллелепипедов. Главная диагональ каждого параллелепипеда соединяет начало координат и один из локальных максимумов расстояния от границы до начала координат. Точки максимумов в интересующей нас области с упорядоченными значениями координат описываются подмножествами индексов, образующих элементы базы. Если базу представить такой k-мерной поверхностью, то легко понять, что будут означать операции объединения и пересечения баз.

Концепция объединения баз-поврехностей. Результат операции можно представить как границу, образованую максимумами двух границ, следующим образом: если луч, проходящий из начала координат в точку максимума, проходит через вторую границу далее исследуемой точки, то эта точка в базу не включается. После перебора всех максимальных точек нужно проверить точки на наличие порядка между ними. Особенностью построения является то, что лишними остаться могут только точки, различающиеся между собой не более чем на одну координату. Таким образом можно отсортировав точки в лексикографическом порядке (т.е. сначала по первой координате, потом при равенстве по второй, если равны обе координаты, то по третьей и т.д.), найти точки, различающиеся не более чем на одну координату и имеющие меньшее значение координаты, и убрать их из базы.

Однако с точки зрения базы, мы не храним ее как поверхность, а храним как набор максимальных подмножеств множества индексов, что накладывает свои ограничения на операцию объединения.

Первым рассмотрим вспомогательный алгоритм поиска места в списке, для добавления нового элемента базы.

Алгоритм 1. Поиск места для вставки элемента.

Входные данные: отсортированный в лексикографическом порядке список A подмножеств индексов, каждое из которых имеет длину k и последний элемент которого не превосходит n; q – количество элементов в списке A; b – подмножество индексов того же типа, что входят в A; p, j – счетчики, равные нулю.

Выходные данные: t – позиция в списке A', куда будет нужно вставить элемент b.

Шаг 1. Увеличиваем j на 1.

Шаг 2. Если $j = q$, то переходим к шагу 4.

Шаг 3. Если $A_j <= b$ в лексикографическом смысле, то переходим к шагу 1, иначе переходим к шагу 4.

Шаг 4. Устанавливаем t, равным j.

Шаг 5. Завершаем алгоритм.

Данный алгоритм будет использоваться как вспомогательный в следующем алгоритме добавления элемента в базу.

Алгоритм 2. Объединение базы с одним элементом.

Входные данные: отсортированный в лексикографическом порядке список A подмножеств индексов, каждое из которых имеет длину k и последний элемент которого не превосходит n; b – подмножество индексов того же типа, что входят в A; q – количество элементов в списке A; p, j – счетчики, равные 0; t – позиция в списке A', куда будет нужно вставить элемент b.

Выходные данные: список A' подмножеств индексов, каждое из которых имеет длину k.

Шаг 1. Устанавливаем t как результат работы алгоритма 1 с параметрами A и b.

Шаг 2. Добавляем в список A' элементы A_i, где $i = \overline{1, t-1}$.

Шаг 3. Добавляем в список A' элемент b.

Шаг 4. Добавляем в список A' элементы A_i, где $i = \overline{t, q}$.

Шаг 5. Увеличиваем q на 1.

Шаг 6. Увеличиваем p на 1.

Шаг 7. Если $p = t$, то переходим к шагу 12.

Шаг 8. Если $A_{pi} <= b_i\ \forall i$, то удаляем p-й элемент из A', иначе переходим к шагу 6.

Шаг 9. Уменьшаем t на 1.

Шаг 10. Уменьшаем q на 1.

Шаг 11. Переходим к шагу 8.

Шаг 12. Увеличиваем p на 1.

Шаг 13. Если $p > q$, то перейти к шагу 15.

Шаг 14. Если $A_{pi} >= b_i\ \forall i$, то удаляем t-й элемент из A', иначе переходим к шагу 6.

Шаг 15. Возвращаем A'.

Данный алгоритм тоже используем как вспомогательный в алгоритме объединения двух баз.

Алгоритм 3. Объединение двух баз, представленных в виде списков подмножеств.

Входные данные: отсортированные в лексикографическом порядке списки A и B подмножеств индексов, каждое из которых имеет длину k и последний элемент которого не превосходит n; q_1 и q_2 – количество элементов в списках A и B соответственно; j – счетчик, равный 0.

Выходные данные: список A' подмножеств индексов, каждое из которых имеет длину k.

Шаг 1. Устанавливаем $A' = A$.

Шаг 2. Увеличиваем j на 1.

Шаг 3. Если $j > q_2$, то завершаем алгоритм.

Шаг 4. Устанавливаем A', равным результату алгоритма 2 с параметрами A, B_j, q_1.

Результатом работы алгоритма 3 будет база-объединение баз A и B.

4. Мощность множества EUH_n^k. Отдельный интерес представлет собой оценка мощности множества всех экстремальных k-однородных гиперграфов. Интерес в данной области связан с разработкой новых алгоритмов шифрования данных, построенных на топологических принципах, в частности на связях вершин внутри ключевого экстремального гиперграфа [15–17], так как чем быстрее растет мощность множества, тем более устойчивым является ключ при простом переборе.

Через $\mathfrak{A}_n^k$ обозначим множество всех экстремальных k-однородных гиперграфов на множестве вершин V_n, через $\mathfrak{B}_n^k(p)$ – множество баз EUH_n^k, состоящих из p элементов, а через $B_n^k(p)$ – мошность множества $\mathfrak{B}_n^k(p)$, т.е. $B_n^k(p) = |\mathfrak{B}_n^k(p)|$. Запишем также сумму всех $B_n^k(p)$:

$$\mathcal{B}_n^k = \sum_p B_n^k(p) = \sum_p |\mathfrak{B}_n^k(p)| = |\mathfrak{A}_n^k|.$$

Известно, что $\mathcal{B}_n^2 = 2^{n-1}$. Для определения мощности множества $\mathfrak{A}_n^k$ был предложен следующий путь.

Так как каждой базе соответствует свой EUH_n^k, то была написана программа на языке Си, которая при заданном n создавала все возможные тройки индексов $(1 \le i, j, k, \le n)$. Таким образом задавались все базы, состоящие из одной тройки. Эти данные являются начальными для следующего рекурсивного алгоритма.

Алгоритм 4. Подсчет $B_n^k(p)$.

Входные данные: A – список баз; *deep* – уровень глубины рекурсии (изначально глубина равна 1); p – номер.

Переменные: *subsets* – массив подмножеств из k индексов; *ckn* – число сочетаний из n по k; j – счетчик (равен р); *new* – список подмножеств индексов (сначала пустой).

Шаг 1. Увеличиваем j на 1.

Шаг 2. Для каждой пары i и j-й элемент из *subsets*, где $i > j$, с помощью алгоритма 2 получаем список A'.

Шаг 3. Если длина списка A' больше *deep*, то прибавляем количество найденых элементов к $B_n^k(deep)$ и вызываем текущий алгоритм с параметрами A, $deep + 1$, j.

Шаг 4. Если $j > ckn$, то завершаем алгоритм.

По завершению всей программы получаем значения $B_n^k(p)$.

Замечание. усть $\lceil a \rceil$ – округление числа a до ближайшего целого большего a. Тогда:

1) для $k = 2$ $p \le \lceil n/2 \rceil$,

2) для $k = 3$ $p \le \lceil n(n-2)/8 \rceil$.

Таблица 1. Значения $B_n^3(p)$ для $n = \overline{3,12}$

p	n									
	3	4	5	6	7	8	9	10	11	12
1	1	4	10	20	35	56	84	120	165	220
2			5	35	140	420	1050	2310	4620	8580
3				10	140	952	4420	16060	49060	131 560
4					35	770	7755	50820	250965	1011010
5					1	216	6072	80784	691 185	4377824
6						16	2078	66778	1082492	11384930
7							288	28790	990774	18465944
8							12	6283	534309	19088128
9								620	167729	12694436
10								20	29532	5424445
11									2706	1469826
12									109	245235
13									1	23800
14										1184
15										22
16										
$\mathcal{B}_n^3$	2	5	16	66	352	2431	21760	252586	3803648	74327145

Часть из полученных $B_n^k(p)$ для случая $k = 3$ представлена в табл. 1 и 2, а для $k = 4$ – в табл. 3. Верхняя строчка задает n, а левый столбец – p. В последней строке приведены B_n^k – общее количество баз на n вершинах.

Отметим, что время расчета растет быстрее экспоненты и при $n = 17$ составляет более 100 ч на 4-ядерном процессоре с частотой 3.0 GHz. Вследствие этого особый интерес представляет аналитический вид $B_n^k(p)$.

Нам известно, что $B_n^3(1) = C_n^3$ (исходя из способа задания начальных данных). Отсюда возникло предположение, что $B_n^k(p)$ – полиномиальная функция. Для подтверждения его был применен метод конечных разностей для сеточной функции. Одним из применений данного метода является нахождение полинома, проходящего через точки сеточной функции (при его существовании). Собственно для этого он и будет использован. Далее кратко опишем его.

Для значений $x : x_0, x_0 + h, x_0 + 2h, \ldots, x_0 + nh$ нам известны значения функции $f(x)$: $y_0, y_1, y_2, \ldots, y_n$.

Тогда $y_{i+1} - y_i = \Delta y_i$ – конечные разности первого порядка, $\Delta y_{i+1} - \Delta y_i = \Delta^2 y_i$ – конечные разности второго порядка, Δ^i – конечная разность i-го порядка.

Пусть задана сеточная функция для $B_n^3(2)$. Ее значения представлены в табл. 4.

Так как $\Delta^i \neq 0$ при $i = \overline{1,6}$ и равны 0 при больших i, то ищем $B_n^3(2)$ в виде многочлена шестого порядка.

На основе этой сеточной функции была построена система из семи линейных уравнений следующего вида:

$$\begin{cases} \sum_{i=0}^{6} 5^i x_{7-i} = 5, \\ \sum_{i=0}^{6} 6^i x_{7-i} = 35, \\ \sum_{i=0}^{6} 7^i x_{7-i} = 140, \\ \sum_{i=0}^{6} 8^i x_{7-i} = 420, \\ \sum_{i=0}^{6} 9^i x_{7-i} = 1050, \\ \sum_{i=0}^{6} 10^i x_{7-i} = 2310, \\ \sum_{i=0}^{6} 11^i x_{7-i} = 4620. \end{cases} \tag{4.1}$$

Здесь в качестве неизвестных выступают коэффициенты многочлена.

Решив систему (4.1), получаем

$$B_n^3(2) = \frac{1}{720}(5n^6 - 45n^5 + 125n^4 - 75n^3 - 130n^2 + 120n) = 5C_{n+1}^6. \tag{4.2}$$

На основе уравнения (4.2) можно заметить, что количество баз связано с числом сочетаний. Это можно объяснить тем, что два элемента базы используют шесть индексов, что предполагает связь количества индексов в одном элементе (три) и количество элементов в базе (два). Кроме того, количество элементов в выборке ($n + 1$) связано с тем, что при $n = 5$ уже существует база из двух элементов, это требует существования соответствующего числа сочетаний. Однако легко убедиться, что $B_n^3(3) \neq C_{n+3}^9$. Следовательно, предполагаем наличие более сложной конфигурации для данного значения.

В целом на основе вышесказанного можно выдвинуть следующую гипотезу.

Г и п о т е з а 1. Значения $B_n^3(p)$ имеют вид

$$B_n^3(p) = \sum_{i=0}^{t(n,p)-x} a_i C_{n+t(n,p)-i}^{3p}, \tag{4.3}$$

где t – целочисленная функция, равная разнице между $3p$ и мощностью наименьшего множества вершин, на котором впервые встречается база из p элементов; x – неизвестное целое число, ограничивающее количество слагаемых в формуле (определяется эмпирически).

Для нахождения a_i был разработан специальный алгоритм 5.

А л г о р и т м 5. Нахождение коэффициентов a_i из формулы (4.3).

Входные данные: b – массив найденных значений для $B_n^3(p)$.

Ш а г 1. $a_0 = b_0$.

Ш а г 2. $a_1 = b_1 - a_0 C_{n+t(n,p)+1}^{3p}$.

Ш а г 3. $a_2 = b_2 - a_0 C_{n+t(n,p)+2}^{3p} - a_1 C_{n+t(n,p)+1}^{3p}$ и т.д., пока $a_i \neq 0$.

С помощью данного алгоритма были найдены следующие формулы для $B_n^3(p)$:

$$B_n^3(1) = C_n^3,$$

$$B_n^3(2) = 5C_{n+1}^6,$$

$$B_n^3(3) = 10C_{n+3}^9 + 40C_{n+2}^9 + 2C_{n+1}^9,$$

$$B_n^3(4) = 35C_{n+5}^{12} + 315C_{n+4}^{12} + 475C_{n+3}^{12} + 55C_{n+2}^{12},$$

Таблица 2. Значения $B_n^3(p)$ для $n = \overline{13,16}$

p	n			
	13	14	15	16
1	286	364	455	560
2	15015	25025	40040	61880
3	318604	710710	1481480	2917200
4	3488485	10650640	29439410	74923420
5	22226490	95125888	355158254	1185493232
6	88238945	543587562	2796484482	12434174318
7	227858000	2080567626	15069874468	90742384028
8	393517284	5502260799	57442352968	476744725216
9	462608192	10273381932	158592213110	1848828612124
10	373867707	13744477309	322690904368	5392594433660
11	208364244	13301782998	490018279354	12000177549114
12	79794344	9358980694	560266567338	20595468085478
13	20772294	4791740884	485082228194	27483018141586
14	3601692	1779649432	318975261710	28680686348246
15	402028	475731212	159354798682	23496815033552
16	27396	90305451	60344252884	15142647607778
17	1054	11918226	17232209478	7678979141312
18	19	1057345	3680175340	3059932450790
19		59482	580612338	955105234348
20		1894	66480772	232311475937
21		26	5386142	43696630958
22			297470	6287511734
23			10500	681758612
24			201	54552102
25			1	3123620
26				121344
27				2842
28				28
29				0
$\mathcal{B}_n^3$	1885102080	62062015500	2652584509440	147198472495020

$$B_n^3(5) = C_{n+8}^{15} + 200C_{n+7}^{15} + 2736C_{n+6}^{15} + 8992C_{n+5}^{15} + 8141C_{n+4}^{15} + 1376C_{n+3}^{15} + 26C_{n+2}^{15},$$

$$B_n^3(6) = 16C_{n+10}^{18} + 1774C_{n+9}^{18} + 30032C_{n+8}^{18} + 153544C_{n+7}^{18} + \\ + 285054C_{n+6}^{18} + 191805C_{n+5}^{18} + 38545C_{n+4}^{18} + 1725C_{n+3}^{18} + 5C_{n+2}^{18},$$

$$B_n^3(7) = 288C_{n+12}^{21} + 22454C_{n+11}^{21} + 423922C_{n+10}^{21} + 2875886C_{n+9}^{21} + 8246146C_{n+8}^{21} + \\ + 10547388C_{n+7}^{21} + 5875430C_{n+6}^{21} + 1276780C_{n+5}^{21} + 89210C_{n+4}^{21} + 1240C_{n+3}^{21}.$$

Эти выражения не противоречат гипотезе и, скорее, подтверждают ее.

Кроме того алгоритм 5 можно завершать, если данные при нужном количестве вершин еще не найдены, что позволяет получить коэффициенты не для всех слагаемых формул, но для нескольких начальных. Здесь представлены обнаруженные значения для случая $p > 7$:

Таблица 3. Значения $B_n^4(p)$ для $n = \overline{4,11}$

p	n							
	4	5	6	7	8	9	10	11
1	1	5	15	35	70	126	210	330
2			15	140	721	2709	8295	21 945
3			1	140	2464	22 008	133 420	623 744
4				35	3375	84 279	1 098 438	9 536 535
5				1	2044	171 734	5 238 452	88 689 918
6					555	199 380	15 585 508	540 661 052
7					70	135 782	30 289 620	2 269 968 192
8					4	54 003	39 581 578	6 795 045 616
9						12 026	35 399 344	14 867 483 686
10						1353	21 866 088	24 205 132 388
11						62	9 349 758	29 695 839 838
12						1	2 760 306	27 694 261 340
13							561 186	19 745 037 108
14							79 284	10 797 372 548
15							8078	4 534 416 086
16							619	1 461 616 993
17							34	360 729 056
18							1	67 837 883
19								9 646 694
20								1 026 445
21								80 774
22								4642
23								186
24								4
$\mathcal{B}_n^4$	2	6	32	352	9303	683 463	161 960 219	143 145 033 003

Таблица 4. Сеточная функция конечных разностей для $B_n^3(2)$

n	5	6	7	8	9	10	11	12
$B_n^3(2)$	5	35	140	420	1050	2310	4620	8580
Δ^1		30	105	280	630	1260	2310	3960
Δ^2			75	175	350	630	1050	1650
Δ^3				100	175	280	420	600
Δ^4					75	105	140	180
Δ^5						30	35	40
Δ^6							5	5
Δ^7								0

$$B_n^3(8) = 12C_{n+15}^{24} + 5983C_{n+14}^{24} + 380834C_{n+13}^{24} + 7587703C_{n+12}^{24} + \\ + 62307684C_{n+11}^{24} + 240698789C_{n+10}^{24} + 465642053C_{n+9}^{24} + 458320446C_{n+8}^{24} + \ldots$$

$$B_n^3(9) = 620C_{n+17}^{27} + 150369C_{n+16}^{27} + 8232384C_{n+15}^{27} + 168534426C_{n+14}^{27} + \\ + 1582063660C_{n+13}^{27} + 7589760929C_{n+12}^{27} + \ldots$$

$$B_n^3(10) = 20C_{n+20}^{30} + 28912C_{n+19}^{30} + 4518253C_{n+18}^{30} + 219352392C_{n+17}^{30} + \\ + 4544828277C_{n+14}^{30} + 47003537429C_{n+13}^{30} + \ldots$$

Полученные результаты позволяют обобщить гипотезу на общий случай.

Гипотеза 2. Значения $B_n^k(p)$ имеют следующий вид:

$$B_n^k(p) = \sum_{i=0}^{t(k,n,p)-x} a_i C_{n+t(k,n,p)-i}^{k\,p},$$

где t — целочисленная функция, равная разнице между kp и мощностью наименьшего множества вершин, на котором впервые встречается база из p элементов; x — неизвестное целое число, ограничивающее количество слагаемых в формуле (определяется эмпирически).

Для $k = 4$ поиск аналитических формул затруднен в связи с недостатком найденных экспериментальных значений. Однако для $p < 4$ коэффициенты были найдены:

$$B_n^4(1) = C_n^4,$$

$$B_n^4(2) = 15C_{n+2}^8 + 5C_{n+1}^8 + C_n^8,$$

$$B_n^4(3) = C_{n+6}^{12} + 127C_{n+5}^{12} + 722C_{n+4}^{12} + 610C_{n+3}^{12} + 183C_{n+2}^{12} + 17C_{n+1}^{12} + C_n^{12}.$$

Это также подтверждает справедливость предлагаемой гипотезы.

Заключение. Согласно результатам исследования, получено описание общего принципа объединения баз EUH_n^k, а также алгоритм объединения для конкретного представления базы в виде множества подмножеств. Были сделаны оценки, связанные с задачей перечисления EUH_n^3: найден общий вид для $\mathcal{B}_n^3(p)$, получены конкретные формулы для случаев $1 \le p \le 7$; эмпирически определена верхняя оценка максимальной размерности базы в зависимости от количества вершин EUH_n^3. Важным итогом работы является предлагаемая гипотеза аналитического вида формул для $B_n^k(p)$, которой не противоречат проведенные расчеты некоторых формул.

При продолжении исследования в данной области предполагается подтвердить гипотезу и найти аналитические формулы для $B_n^k(p)$ и общую формулу для $\mathcal{B}_n^k(p)$, в том числе для $k > 3$.

СПИСОК ЛИТЕРАТУРЫ

1. *Иванов М.В., Калашников И.В., Нуруллаев М.М.* Исследование структурных свойств сети интернет на основе метаграфовых моделей // Информатика и автоматизация. 2020. Т. 19. № 4. С. 880–905.
2. *Миков А.И., Миков А.А.* Свойства геометрических гиперграфов беспроводных компьютерных сетей // Информатизация и связь. 2020. № 4. С. 60–66. https://doi.org/10.34219/2078-8320-2020-11-4-60-66
3. *Герасименко Е.М., Дышаев Н.Н.* Персонализация в фолксономиях в виде гиперграфов на основе кластеризации тегов // Информатика, вычислительная техника и инженерное образование. 2019. № 2 (35). С. 11–15.
4. *Зыков А.А.* Гиперграфы // УМН. 1974. Т. XXIX. № 6 (180). С. 89–154.
5. *Александров П.С.* Комбинаторная топология. М.: Гостехтеориздат, 1947. 660 с.
6. *Бобу А.В., Куприянов А.Э., Райгородский А.М.* О числе ребер однородного гиперграфа с диапазоном разрешенных пересечением // Проблемы передачи информации. 2017. Т. 53. № 4. С. 16–42.
7. *Балобанов А.Е., Шабанов Д.А.* О числе независимых множеств в простых гиперграфах // Мат. заметки. 2018. Т. 103. № 1. С. 38–48. https://doi.org/10.4213/mzm11508
8. *Денисов И.О., Шабанов Д.А.* О концентрации значений чисел независимости случайных гиперграфов // Дискретная математика. 2021. Т. 33. № 4. С. 32–46. https://doi.org/10.4213/dm1676
9. *Захаров П.А., Шабанов Д.А.* О максимальном разрезе в случайном гиперграфе // Доклады Российской академии наук. Математика, информатика, процессы управления. 2021. Т. 501. № 1. С. 26–30. https://doi.org/10.31857/S2686954321060187
10. *Райгородский А.М., Черкашин Д.Д.* Экстремальные задачи в раскрасках гиперграфов // УМН. 2020. Т. 75. № 1(451). С. 95–154. https://doi.org/10.4213/rm9905
11. *Ванг Л., Егорова Е.К., Мокряков А.В.* Развитие теории гиперграфов // Изв. РАН. ТиСУ. 2018. № 1. С. 111–116.

12. *Миронов А.А.* Геометрия точек пространства R^n, реализуемых в граф // УМН. 1977. Т. XXXII. № 6 (198). С. 232–233.
13. *Костяной Д.С., Мокряков А.В., Цурков В.И.* Алгоритмы восстановления гиперграфов по заданному вектору степеней вершин // Изв. РАН. ТиСУ. 2014. № 4. С. 43–48.
14. *Бондаренко В.А., Николаев А.В.* Об одном классе гиперграфов и о вершинах релаксаций разрезного многогранника // ДАН. 2012. Т. 442. № 3. С. 300–302.
15. *Егорова Е.К., Мокряков А.В., Суворова А.А., Цурков В.И.* Алгоритм передачи многомерных данных с использованием экстремальных однородных гиперграфов // Изв. РАН. ТиСУ. 2021. № 1. С. 73–78.
16. *Каменев А.Р., Ирбитский И.С., Пашковская Е.А.* Методы подбора ключа для алгоритмов шифрования на графах // Сб. тез. работ ММНК XLVIII Гагаринские чтения – 2022. М.: Перо, 2022. С. 252.
17. *Лежинский М.В.* Концепция топологически-ориентированных хэш-функций // Сб. тез. работ ММНК XLVIII Гагаринские чтения – 2022. М.: Перо, 2022. С. 252.
18. *Mironov A.A.* Minimax under Transportation Constraints. Dordrecht: Kluwer Acad. Publ., 1999. 309 p.
19. *Миронов А.А.* Равномерные обобщенные графы // ДАН. 1996. Т. 351. 4. С. 465–468.
20. *Мокрозуб В.Г., Немтинов В.А., Мордвин А.С., Илясов А.А.* Применение n-ориентированных гиперграфов и реляционных баз данных для структурного и параметрического синтеза технических систем // Прикладная информатика. 2010. № 4 (28). С. 115–122.
21. *Суворова А.А., Берецкий И.С.* Алгоритм потокового шифрования на экстремальных k-однородных гиперграфах // Сб. тез. работ ММНК Гагаринские чтения – 2020. М.: МАИ, 2020. С. 510–511.
22. *PrГjfer H.* Neuer Beweis eines Satzes Гjber Permutationen // Arch. Math. Phys. 1918. V. 27. P. 742–744.
23. *Gilbert E.N.* Enumeration of Labelled Graphs // Canad. J. Math. 1956. V. 8. P. 405–411.
24. *Кузьмин О.В., Чернигова А.Г.* Автоматизация комбинаторного кодирования и декодирования корневых деревьев // Современные технологии. Системный анализ. Моделирование. 2015. № 1(45). С. 84–88.
25. *Мокряков А.В.* Представление гиперграфов в качестве алгебраической структуры // Изв. РАН. ТиСУ. 2011. № 5. С. 53–59.
26. *Погребной В.К., Погребной А.В.* Исследование полиномиальности метода вычисления интегрального описателя структуры графа // Изв. Томск. политехн. ун-та. 2013. Т. 323. № 5. С. 146–151.
27. *Гольцова Т.Ю., Егорова Е.К., Мокряков А.В., Цурков В.И.* Сигнатуры экстремальных 2-однородных гиперграфов // Изв. РАН. ТиСУ. 2021. Т. 6. № 6. С. 52–60.
28. *Мокряков А.В., Цурков В.И.* Восстановление 2-комплексов по целочисленному неотрицательному вектору // АиТ. 2011. № 12. С. 130–143.