Труды Карельского научного центра РАН. 2022. № 4. С. 75–79
Transactions of the Karelian Research Centre RAS. 2022. No. 4. P. 75–79
DOI: 10.17076/mat1526

ОРИГИНАЛЬНЫЕ СТАТЬИ
Original articles

УДК 519.115:519.2

КОМБИНАТОРНЫЙ АНАЛИЗ СХЕМЫ СОЧЕТАНИЙ СО СЛУЧАЙНЫМ ПАРАМЕТРОМ

Н. Ю. Энатская

Московский институт электроники и математики,
Национальный исследовательский университет «Высшая школа экономики»
(ул. Таллинская, 34, Москва, Россия, 123458)

Для схемы сочетаний со случайным параметром проводится анализ по всем направлениям перечислительной комбинаторики: нахождения числа ее исходов, их перечисление, решения для них задачи нумерации, определения их вероятностного распределения и их моделирование.

Ключевые слова: выделенные ячейки; случайный параметр

Для цитирования: Энатская Н. Ю. Комбинаторный анализ схемы сочетаний со случайным параметром // Труды Карельского научного центра РАН. 2022. № 4. С. 75–79. doi: 10.17076/mat1526

N. Yu. Enatskaya. COMBINATORY ANALYSIS OF A COMBINATION SCHEME WITH A RANDOM PARAMETER

National Research University Higher School of Economics, Moscow Institute of Electronics and Mathematics (34 Tallinskaya St., 123458 Moscow, Russia)

Analysis covering the full range of enumerative combinatorics elements is carried out for a combination scheme with a random parameter: finding the number of its outcomes, enumerating them, solving the numbering problem for them, determining their probability distribution, and modeling them.

Keywords: selected cells; random parameter

For citation: Enatskaya N. Yu. Combinatory analysis of a combination scheme with a random parameter. *Trudy Karel'skogo nauchnogo tsentra RAN = Transactions of the Karelian Research Centre RAS.* 2022;4:75–79. doi: 10.17076/mat1526

Введение

В наглядной интерпретации размещения частиц по ячейкам к схеме сочетаний со случайным параметром приводит размещение r неразличимых частиц по n различимым ячейкам, вмещающим каждая не более одной частицы, при рассмотрении исходов заполнения выделенных m из n ячеек, число частиц в которых X – случайно. Далее эту схему будем называть схемой А.

Не нарушая общности анализа схемы, будем считать выделенными первые m ячеек.

1. Вспомогательные результаты

Приведем кратко результаты анализа для отдельных комбинаторных схем по перечисленным в аннотации направлениям (для каждой схемы в своих обозначениях), используемые здесь далее.

1.1. Схема сочетаний

В [1] и [2] приводятся результаты комбинаторного анализа схемы сочетаний из n элементов по r с общим числом исходов C_n^r.

Бесповторное перечисление всех C_n^r различимых исходов схемы производится случайным процессом размещения частиц, начиная с нуля частиц с добавлением на каждом шаге еще одной частицы до заданного числа частиц r, независимо по всем ячейкам, правее последней занятой, с нумерацией исходов подряд, перечисляемых группами по возрастанию номеров первых занятых ячеек, а в каждой группе – по возрастанию номеров последних занятых ячеек. Номера элементов в каждом исходе схемы $R = (n_1 n_2 \dots n_r)$ упорядочены по возрастанию, и значение n_m, $(m = \overline{1,r})$ определяет величина l_m – порядковый номер n_m по возрастанию среди значений от $(n_{m-1}{+}1)$ до n.

Прямая задача нумерации (ЗН) решается теоремой:

Теорема 1. *Дан номер N исхода схемы сочетаний. Тогда в его искомом виде $R = (n_1 n_2 \dots n_r)$ значение l_m, $(m = \overline{1,r})$ вычисляется при $l_0 = 1$, $n_0 = 0$, $C_0^0 = 1$ по формуле*

$$l_m = \min j_m :$$

$$\left\{ N \leqslant \sum_{i=1}^{m-1} \sum_{s_i = n_{i-1}+I(i-1)}^{l_{m-1}-1} C_{n-s_i}^{r-i} + \sum_{s_m = n_{m-1}}^{j_m} C_{n-s_m}^{r-m} \right\},$$

где $I(Z) = 0$ при $Z > 0$ и $I(Z) = 1$ при $Z = 0$.

Обратная ЗН решается теоремой:

Теорема 2. *Дан исход $R = (n_1 n_2 \dots n_r)$. Тогда его номер N находится по формуле*

$$N = \sum_{i=1}^{r} \sum_{s_i = n_{i-1}+1}^{n_i - 1} C_{n-s_i}^{r-i} + 1,$$

где $n_0 = 0$.

1.2. Обобщенная схема последовательных действий (ОПД)

В [1] приводятся результаты комбинаторного анализа схемы ОПД.

В схеме проводится k последовательных действий, i-е из которых $(i = \overline{1,k})$ на i-м шаге совершается $N^{(i)}$ способами. Тогда число исходов этих k действий складывается из $N^{(k-1)}$ пучков размерами $\bar{d}^{(i)} = (d_1^{(i)}, d_2^{(i)}, \dots, d_{N^{(k-1)}}^{(k)})$, т. е. общее число $N = N^{(k)}$ исходов схемы получается из рекуррентного соотношения при $i = k$ и $N = N^{(0)} = 1$

$$N^{(i)} = \sum_{l=1}^{N^{(i-1)}} d_l^{(i)}.$$

Вид исхода после совершения i действий будет формироваться из принятых видов исходов последовательных действий, которые будем соответственно обозначать через R_{ij_i}, где i – номер действия, а j_i – номер исхода в результате его совершения.

Задача нумерации решается для нашей схемы при решенной ЗН для каждого из k действий и известной пучковой структуре графа перечисления исходов нашей схемы, т. е. с известными числами исходов (размерами пучков) при каждом действии на каждой итерации. Под траекторией T исхода схемы понимается последовательность исходов, ведущая в графе их перечисления от начала к исследуемому на последней итерации.

Прямая и обратная ЗН решены следующими теоремами:

Теорема 3. *Пусть совершается k действий и задан номер исхода $N_*^{(k)}$. Тогда его вид, определяемый номерами исходов траектории T в содержащих их пучках $\{j_i\}$ от первой до k-й итераций, вычисляется по рекуррентной формуле для j_i $(i = \overline{1,k})$*

$$j_i = N_*^{(i)} - \sum_{l=1}^{N_*^{(i-1)}-1} d_l^{(i)},$$

где все пучковые структуры действий $\bar{d}^{(i)}$ заданы и

$$N_*^{(k-1)} = \delta + \max t : \left(\sum_{l=1}^{t} d_l^{(k)} = A_k \leqslant N_*^{(k)} \right),$$

где $\delta = 0$ при $A_k = N_*^{(k)}$ и $\delta = 1$ при $A_k < N_*^{(k)}$; заменяя k на i, доходим по рекурренте до первого шага.

По решенной ЗН для всех действий находим из $\{j_i\}$ виды их исходов, из которых получаем искомый вид исхода $R_*^{(k)}$.

Теорема 4. *Пусть совершается k действий и задан вид исхода $R_*^{(k)} = \{j_1, \dots, j_k\}$. Тогда его номер $N_*^{(k)}$ определяется по рекуррентной формуле при $i = k$, $i = \overline{1,k}$*

$$N_*^{(i)} = \sum_{l=1}^{N_*^{(i-1)}-1} d_l^{(i)} + j_i,$$

начиная с $i = 1$ при $N_*^{(1)} = j_1$.

2. Вид исходов схемы А, их перечисление и число

В условиях схемы каждый ее исход будет иметь вид 1) возрастающего набора номеров непустых ячеек среди выделенных m или 2) m-размерной последовательности нулей и единиц, где нули стоят на местах пустых, а единицы – непустых ячеек среди первых m ячеек.

Процедуру прямого нумерованного перечисления исходов схемы представим в виде двух этапов – итераций: перебора всех возможных растущих количеств X попавших в первые m ячеек частиц и перечисления исходов ранее изученной (см. [1, 2]) при каждом исходе первого этапа схемы сочетаний в порядке установленного там перечисления. Наглядно этот процесс перечисления представляется графом по методу графов (МГ, см. [3]). Структура графа представляет собой реализацию двух описанных итераций, состоящих соответственно на первой итерации из одного пучка размером диапазона возможных значений случайного числа X частиц в выделенных ячейках, а на второй (итоговой) – из каждого исхода пучка первой итерации образуется пучок размера числа исходов схемы сочетаний с фиксированным значением случайного числа X, от которого зависит размер пучка. Таким образом, наша схема представляет собой обобщенную схему двух последовательных действий, приведенную в п. 1.2 по результатам, используемым далее.

Теорема 5. *Число N разных исходов схемы А определяется по формуле*

$$N = \sum_{k=t}^{T} C_m^k, \tag{1}$$

где для возможных значений числа k частиц, попадающих в выделенные ячейки, выполняется соотношение $t = max(0, r - n + m) \leqslant k \leqslant min(r, m) = T$.

Доказательство. Доказательство формулы (1) для числа N следует из вышеописанной структуры графа перечисления исходов схемы А как сумма размеров пучков ее итоговой итерации в диапазоне $k \in (t, T)$ возможных значений с. в. X, а бесповторность перечисления исходов схемы А обеспечивается разными количествами размещаемых в выделенных ячейках частиц. □

3. Решение ЗН

Будем решать ЗН для схемы А в прямой и обратной постановках, определяя соответственно вид ее исхода R^* по ее номеру при перечислении N^*, и наоборот. В общем виде ЗН в обеих постановках решена в схеме ОПД (см. в п. 1.1 или в [1]). Выпишем в форме пошаговых АЛГОРИТМОВ решения ЗН при заданных параметрах схемы (n, m, r), откуда из п. 2 известны крайние значения (t, T) диапазона изменения целочисленной с. в. X.

Прямая ЗН. Дан номер исхода схемы N^*, требуется найти его вид R^*.

Шаги решения:

1) определяем номер j пучка итоговой итерации, содержащий искомый исход схемы по формуле теоремы 3

$$j = \delta + max\ s : \left\{ \sum_{k=t}^{s} C_m^k = A_s \leqslant N^* \right\}, \tag{2}$$

где $\delta = 0$ при $A_s = N^*$ и $\delta = 1$ при $A_s < N^*$;

2) находим номер h искомого исхода в пучках итоговой итерации по формуле

$$h = N^* - A_s; \tag{3}$$

3) по результату теоремы 1 решения прямой ЗН для схемы сочетаний (см. в п. 1.1 или [1] и [2]), с соответствующим номером h и значением числа k частиц X, в пучке с номером, равным $j = k - t + 1$, откуда при $k = j + t - 1$ получаем искомый вид исхода схемы R^*.

Обратная ЗН. Дан вид исхода схемы R^*, требуется найти его номер N^*.

Шаги решения:

1) из данного вида исхода R^* определяем число k непустых ячеек среди выделенных в схеме и номер пучка $j = k - t + 1$, содержащего искомый исход;

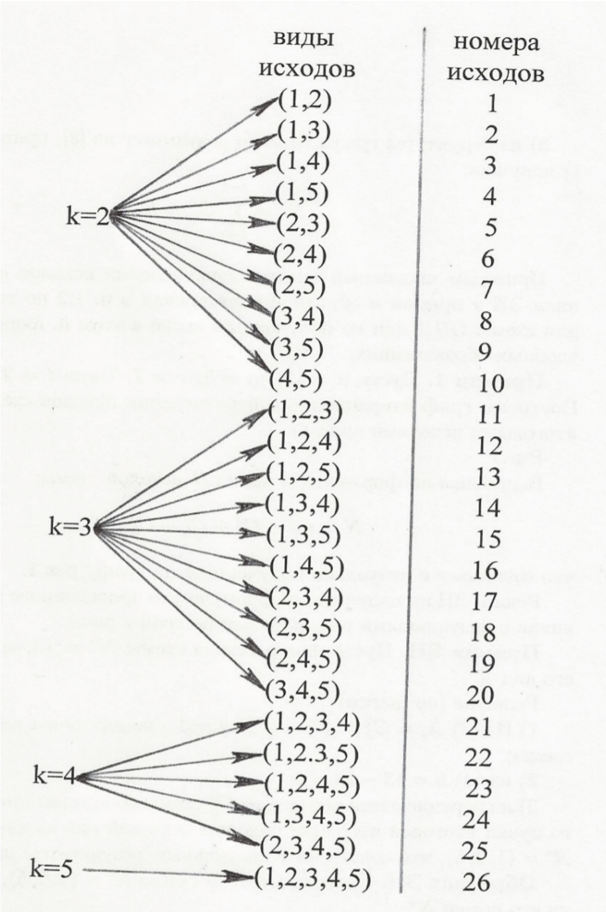
2) по результату решения обратной ЗН для схемы сочетаний п. 1.1 с параметрами чисел ячеек m и частиц k по виду R^* находим его номер h, совпадающий с номером искомого исхода в итоговом пучке;

3) по теореме 2 из п. 1.1 получаем

$$N^* = \sum_{k=t}^{j-1} C_m^k + h. \tag{4}$$

Приведем численный пример перечисления исходов схемы с решением ЗН в прямой и обратной постановках в п. 1.2 по теоремам 3 и 4 для схемы ОПД или по полученным выше в этом пункте формулам в более удобных обозначениях.

Пример 1. Пусть $n = 10$, $m = 5$, $r = 7$. Тогда $t = 2 \leqslant k \leqslant 5 = T$. Построим граф итерационного перечисления исходов схемы на рисунке с итоговыми исходами вида 1).



Граф перечисления исходов схемы в примере 1
The graph of outcomes enumeration of the scheme in example 1

Вычислим по формуле (1) число N исходов схемы:

$$N = C_5^2 + C_5^3 + C_5^4 + C_5^5 = 26,$$

что совпадает с визуально полученным по графу на рисунке.

Решим ЗН по полученным формулам ее приведенных в АЛГОРИТМАХ решений, сравнивая с визуальными результатами по графу на рисунке.

Прямая ЗН. Пусть номер исхода схемы $N^* = 13$, требуется найти его вид R^*.

Решение (по шагам):

1) из (2) $A_s = C_5^2 = 10$; $\delta = 1$; $j = 2$ – номер пучка итогового исхода схемы;

2) из (3) $h = 13 - 10 = 3$;

3) из перечисления исходов из C_5^3 схемы сочетаний при $j = 2$ – второго пучка итоговой итерации с $k = 3$ находим искомый вид исхода нашей схемы $R^* = (1, 2, 5)$, что совпадает с визуальным результатом на рисунке.

Обратная ЗН. Пусть вид исхода схемы $R^* = (1, 2, 5)$, требуется найти его номер N^*.

Решение (по шагам):

1) из данного вида исхода $R^* = (1, 2, 5)$ следует, что $k = 3$, и находим номер пучка искомого исхода $j = 2$;

2) из результата решения обратной ЗН для схемы сочетаний по данному виду исхода $R^* = (1, 2, 5)$ из результата решения обратной ЗН для схемы сочетаний с параметрами $m = 5$ ячеек, $k = 3$ частиц получаем его номер $h = 3$ в пучке;

3) из (4) $N^* = \sum_{k=t}^{j-1} C_m^k + h = C_5^3 + 3 = 13$, что совпадает с визуальным результатом на рисунке.

4. Вероятностное распределение исходов схемы

Теорема 6. *Пусть процесс бесповторного перечисления исходов схемы сочетаний приводит к равновероятным исходам. Тогда $p_i(k)$ – вероятность i-го исхода ($i = \overline{1, N}$) нашей схемы – в этом случае определяется формулой*

$$p_i(k) = \frac{C_{n-m}^{r-k}}{C_n^r}, \qquad (5)$$

где $k = \overline{t, T}$, $(k-t+1)$ – номер пучка этого i-го итогового исхода перечисления всех исходов схемы с совпадающими внутри каждого пучка вероятностями, а их номера в общем списке исходов, обозначенные индексом i, меняют значения в пределах пучка размером C_m^k.

Доказательство. В условиях теоремы вероятностное распределение случайной величины X имеет гипергеометрическое распределение $P(X{=}k){=}C_m^k C_{n-m}^{r-k}/C_n^r$, где $t{=}max(0, r{-}n{+}m) \leqslant k \leqslant min(r, m) = T$, которое по описанной выше структуре графа перечисления исходов нашей схемы является распределением исходов единственного пучка первой итерации. На второй итоговой итерации мы имеем $(T-t+1)$ пучков исходов соответствующих схем сочетаний с параметром k – числа частиц от t до T включительно, каждый со своими совпадающими в пучке вероятностями итерационных переходов, обратными размеру пучка C_m^k.

Отсюда следует, что вероятностное распределение на множестве исходов нашей схемы представляет собой группы совпадающих вероятностей в пределах исходов каждого итогового пучка, вычисляемых по траекториям исходов схемы как произведения соответствующей гипергеометрической вероятности исхода первой итерации с $k = \overline{t, T}$ частицами в выделенных m ячейках и вероятности $1/C_m^k$ исхода схемы сочетаний с параметром числа частиц равным k. Тогда

$$p_i(k) = \frac{C_m^k C_{n-m}^{r-k}}{C_n^r} \frac{1}{C_m^k} = \frac{C_{n-m}^{r-k}}{C_n^r},$$

где $k = \overline{t, T}$, $(k - t + 1)$ – номер пучка этого i-го итогового исхода перечисления всех исходов схемы с совпадающими внутри каждого пучка вероятностями, а их номера в общем списке исходов, обозначенные индексом i, меняют значения в пределах пучка размером C_m^k, что и утверждалось. □

Замечание. При неравновероятном распределении исходов схемы сочетаний, определяющем вероятностное распределение исходов первой итерации, вероятность каждого исхода нашей схемы вычисляется как вероятность ведущей к нему траектории.

Покажем на примере расчет вероятностей на множестве исходов нашей схемы в случае равновероятных исходов схемы сочетаний.

Пример 2. Будем находить вероятностное распределение исходов схемы в условиях примера 1, т. е. при $n = 10$, $m = 5$, $r = 7$. Тогда $t = 2 \leqslant k \leqslant 5 = T$.

Общее число исходов схемы $N = 26$, полученное в примере 1, в порядке их нумерации разделим на группы исходов совпадающих итоговых вероятностей в соответствии с размерами пучков второй итерации их перечисления, т. е. размерами: $C_5^2 = 10$; $C_5^3 = 10$; $C_5^4 = 5$; $C_5^5 = 1$, соответствующие значениям $k = \overline{2, 5}$ с. в. X. Тогда при каждом значении k по формуле (5) будем получать подряд совпадающие итоговые вероятности исходов нашей схемы в количестве C_5^k:

$p_1(2) = \ldots = p_{10}(2) = 1/120$;
$p_{11}(3) = \ldots = p_{20}(3) = 5/120$;
$p_{21}(4) = \ldots = p_{25}(4) = 10/120$;
$p_{26}(5) = 10/120$.

Проверка на распределение

$$10 \cdot \frac{1}{120} + 10 \cdot \frac{5}{120} + 5 \cdot \frac{10}{120} + 1 \cdot \frac{10}{120} = 1.$$

5. Моделирование исходов схемы

При изученном вероятностном распределении на множестве исходов схемы будем методом маркировки (см. [4]) разыгрывать его случайный номер от 1 до N. Тогда по результату прямой ЗН по разыгранному номеру исхода будем определять его вид, который считаем смоделированным исходом нашей схемы.

Литература

1. *Энатская Н. Ю.* Анализ комбинаторных схем в доасимптотической области изменения параметров // Труды Карельского научного центра РАН. 2018. № 7. С. 117–133. doi: 10.17076/mat750

2. *Энатская Н. Ю.* Комбинаторный анализ схемы сочетаний // Промышленные АСУ и контроллеры. 2015. № 8. С. 33–38.

3. *Энатская Н. Ю., Хакимуллин Е. Р.* Метод графов для решения задач перечислительной комбинаторики // Приборы и системы. Управление, контроль, диагностика. 2014. № 8. С. 15–21.

4. *Энатская Н. Ю., Хакимуллин Е. Р.* Стохастическое моделирование. М.: МИЭМ, 2012. 185 с.

References

1. *Enatskaya N. Yu.* Analysis of combinatorial schemes in the pre-asymptotic region of parameter change. *Trudy Karel'skogo nauchnogo tsentra RAN = Transactions of the Karelian Research Centre RAS.* 2018;7:117–133. doi: 10.17076/mat750 (In Russ.)

2. *Enatskaya N. Yu.* Combinatorial analysis of a combination scheme. *Promyshlennye ASU i kontrollery = Industrial ACS and Controllers.* 2015;8:33–38. (In Russ.)

3. *Enatskaya N. Yu., Khakimullin E. R.* Graphs method for solving enumerative combinatorics. *Pribory i sistemy. Upravlenie, kontrol, diagnostika = Instruments and Systems: Monitoring, Control, and Diagnostics.* 2014;8:15–21. (In Russ.)

4. *Enatskaya N. Yu., Khakimullin E. R.* Stochastic modelling. Moscow: MIEM; 2012. 185 p. (In Russ.)

Поступила в редакцию / received: 05.12.2021; принята к публикации / accepted: 28.04.2022.
Автор заявляет об отсутствии конфликта интересов / The author declares no conflict of interest.

СВЕДЕНИЯ ОБ АВТОРЕ:

Энатская Наталия Юрьевна
канд. физ.-мат. наук, доцент Департамента прикладной математики

e-mail: nat1943@mail.ru

CONTRIBUTOR:

Enatskaya Natalia
Cand. Sci. (Phys.-Math.), Associate Professor