УДК 519.21

Н.Ю. КУЗНЕЦОВ

ОЦЕНКА КОЛИЧЕСТВА «ХОРОШИХ» ПЕРЕСТАНОВОК МОДИФИЦИРОВАННЫМ МЕТОДОМ УСКОРЕННОГО МОДЕЛИРОВАНИЯ

Ключевые слова: *«хорошая» перестановка, модифицированный метод ускоренного моделирования, несмещенная оценка, выборочная дисперсия, относительная погрешность.*

Задача перечисления всех полных отображений на алгебраической структуре $(G, +)$ относится к наиболее трудным проблемам дискретной математики [1, 2]. Отображение $f: G \to G$ называется полным, если $f(\cdot)$ является биекцией и отображение $h(x) = x + f(x)$ также биекция. В работе [3] исследовалась сложность проблемы нахождения количества всех полных отображений для различных структур $(G, +)$. В частности, показано, что для замкнутых структур эта проблема является *NP*-полной.

В настоящей статье основное внимание сосредоточено на сильных полных отображениях [4], когда $x + f(x)$ также перестановка. Иначе говоря, пусть $(s_0, s_1, \ldots, s_{N-1})$ — произвольная перестановка символов $(0, 1, \ldots, N-1)$. Построим новый набор $(t_0, t_1, \ldots, t_{N-1})$ согласно правилу: $t_i = i + s_i \pmod N$, $i = 0, 1, \ldots, N-1$. Если построенный набор — перестановка, то исходная перестановка $(s_0, s_1, \ldots, s_{N-1})$ называется «хорошей». Цель исследования — разработка метода, позволяющего оценивать количество «хороших» перестановок для как можно бóльших значений N.

Термин «хорошая» перестановка введен в [5]. В этой же статье дан исчерпывающий анализ основ применения «хороших» перестановок в криптографии (см. также работу [6], в которой описаны принципы использования «хороших» перестановок в роторных шифровальных системах).

Задача точного вычисления количества M_N «хороших» перестановок требует огромных вычислительных затрат, которые экспоненциально возрастают с ростом N (заметим, что $M_N = 0$ для четных N). Значения M_N для $N \le 19$ приведены в [5], для $N \le 23$ — в [7], для $N \le 25$ — в [8]. Вычисление M_N при увеличении N хотя бы на 2 составляет принципиальную проблему и зависит, в первую очередь, от развития вычислительных мощностей. Поэтому основной акцент в исследованиях M_N делается на приближенных методах расчета, в частности асимптотических и статистических методах.

Общее количество перестановок равно $N!$, следовательно, сформулированная выше детерминированная задача нахождения M_N эквивалентна вероятностной задаче оценки вероятности P_N того, что выбранная случайным образом перестановка является «хорошей», $M_N = N! P_N$. Поскольку $P_N = 0$ для четных N, то в дальнейшем рассматриваем лишь нечетные значения N.

В работе [9] доказано, что $P_N \le ae^{-cN}$ при больших N, причем $c \ge 0{,}0885$. В [10] эта оценка улучшена: $c \ge \dfrac{\ln 2}{2} \approx 0{,}3466$. В [5] выдвинута гипотеза, что

$$P_N \sim ae^{-cN} \tag{1}$$

при $N \to \infty$, причем $c \in [0{,}5, 1]$. В этой же работе предложена аппроксимация, позволяющая строить прогнозные оценки P_N при $N = 25, 35, 45, 55$ (правда, без каких-либо оценок погрешности аппроксимации). При этом делается вывод, что $c \approx 0{,}9538$.

Необходимо отметить также работу [11], в которой исследуются свойства «хороших» перестановок и доказан ряд теорем, позволяющих существенно сузить область поиска, повысив при этом эффективность создаваемых алгоритмов.

В [12] для оценки P_N предложен принципиально новый подход, основанный на использовании алгоритма ускоренного моделирования. Он позволяет с высокой точностью при относительно небольших затратах времени строить несмещенные оценки и соответствующие доверительные интервалы для P_N при весьма больших N (приведена, в частности, оценка для P_N при $N = 155$, построенная с относительной погрешностью 10 %). Полученные оценки подтверждают гипотезу (1), при этом делается вывод, что $0{,}9825 \le c \le 0{,}9883$.

Настоящая статья является продолжением работы [12]. Предлагается усовершенствованный метод ускоренного моделирования, дающий возможность при больших N более чем в 10 раз сократить время вычислений (в результате такого сокращения удалось построить оценку для P_N при $N = 205$ с относительной погрешностью 5 %). Расширение диапазона значений N и повышение точности вычислений позволяет уточнить диапазон значений константы c: $0{,}9825 \le c \le 0{,}9985$, причем есть основания полагать, что при $N > 205$ верхняя граница для c может еще возрасти. Используя тот же прием, что и в [12], а также уточненные статистические оценки для P_N, строим верхние и нижние оценки P_N в диапазоне $75 \le N \le 205$.

МОДИФИЦИРОВАННЫЙ МЕТОД УСКОРЕННОГО МОДЕЛИРОВАНИЯ

В работе [12] предложен алгоритм ускоренного моделирования, позволяющий целенаправленно размещать символы $0, 1, \ldots, N-1$ на одной из N позиций. Благодаря этому удалось резко увеличить вероятность получения «хорошей» перестановки. Именно дальнейшее увеличение данной вероятности за счет анализа последствий размещения того или иного символа на определенных позициях и является основной идеей модификации. Как и в [12], алгоритм состоит из двух этапов.

Вначале символы располагаем на позициях с номерами $m, \ldots, N-1$, затем — на позициях $0, \ldots, m-1$ (здесь $m \in \{1, \ldots, N-1\}$ — некоторый параметр). На первом этапе вероятность получения «плохой» перестановки достаточно мала, поэтому используется весьма простой («быстрый») алгоритм. По мере заполнения позиций значительно возрастает вероятность получения «плохой» перестановки. Поэтому второй этап требует более тщательного выбора символов, которые могут стоять на тех или иных позициях. Используется более сложный критерий выбора, проверка которого занимает существенно большее время.

Текущим состоянием перестановки назовем n-мерный вектор $\bar{s} = (s_0, s_1, \ldots, s_{N-1})$, где $s_i \in \{-1, 0, 1, \ldots, N-1\}$. Запись $s_i = -1$ означает, что позиция i еще не занята ни одним из символов; если $s_i = k$, то на позиции i размещен символ k. Состояние $\bar{s}$ является «хорошим», если $s_i \ne -1$ и $s_i \ne s_j, i \ne j$, для любых $i, j \in \{0, 1, \ldots, N-1\}$. Состояние $\bar{s}$ назовем «тупиковым», если из него невозможно получить «хорошую» перестановку. Обозначим:

$$\nu_i(\bar{s}) = \begin{cases} 1, \text{ если } i = s_k \text{ для некоторого } k, \\ 0 \text{ в противном случае, } i = 0, 1, ..., N-1; \end{cases}$$

$$\mu_j(\bar{s}) = \begin{cases} 1, \text{ если } j = k + s_k \,(\mathrm{mod}\, N),\ s_k \ne -1, \text{ для некоторого } k, \\ 0 \text{ в противном случае, } j = 0, 1, ..., N-1. \end{cases}$$

В приведенном далее алгоритме строится оценка $\hat{P}_{N1}$ в одной реализации для вероятности P_N.

Этап I (алгоритм телефонного диска).

1. Пусть $\hat{P}_{N1} = 1$ — начальное значение оценки, $\bar{s} = (-1, \ldots, -1)$, $\nu_i(\bar{s}) = 0$, $\mu_j(\bar{s}) = 0$, $i, j = 0, \ldots, N-1$.

2. Предположим, что r последовательно принимает значения $m, m+1, \ldots, N-1$ (номер позиции). Определим множество символов, которые можно расположить на позиции r:

$$A_r(\bar{s}) = \{i : \nu_i(\bar{s}) = 0, \mu_k(\bar{s}) = 0,\ k = r + i \ (\mathrm{mod}\, N)\}. \tag{2}$$

Обозначим $|A_r(\bar{s})|$ количество символов во множестве $A_r(\bar{s})$. Если $|A_r(\bar{s})| = 0$, то этап I (а вместе с ним и этап II) окончен и в качестве оценки имеем $\hat{P}_{N1} = 0$ (в этой реализации не удалось построить «хорошую» перестановку). Если же $|A_r(\bar{s})| > 0$, то с одной и той же вероятностью $1/|A_r(\bar{s})|$ выберем один из символов множества $A_r(\bar{s})$. Если это символ i, то полагаем

$$\hat{P}_{N1} := \hat{P}_{N1} \frac{|A_r(\bar{s})|}{N+m-r},\ s_r = i,\ \nu_i = 1,\ \mu_k = 1,$$

где $k = r + i \ (\mathrm{mod}\, N)$ (обозначение «:=» означает, что новое значение $\hat{P}_{N1}$ вычисляется как произведение прежнего значения $\hat{P}_{N1}$ на соответствующий множитель). Заметим, что $N + m - r$ — количество символов, которые еще не были распределены по позициям. Далее, увеличивая r на единицу, повторяем шаг 2 алгоритма.

По мере уменьшения незанятых позиций возрастает вероятность получения «плохой» перестановки. Поэтому на этапе II алгоритма следует более тщательно выбирать позицию и символы, которые можно расположить на данной позиции. Обозначим $\bar{y} = Z(\bar{s}, r, i)$ состояние перестановки, которое получено из состояния $\bar{s}$ $(i, r : s_r = -1, \nu_i(\bar{s}) = 0)$ путем размещения символа i на позиции r. Кроме того, обозначим S^* множество всех «тупиковых» состояний. В некоторых случаях выбор позиции и соответствующего символа может быть осуществлен однозначно. Введем преобразование $U(\bar{s})$, позволяющее осуществлять такие однозначные переходы. Пусть $\bar{s}$ — произвольное состояние. Алгоритм определения состояния $\bar{w} = U(\bar{s})$ (а также соответствующего нормирующего множителя $q(\bar{s}, \bar{w})$) формулируется следующим образом.

1. Полагаем $\bar{w} = \bar{s}$ и $q = 1$ (начальное значение нормирующего множителя).

2. Для каждой позиции r такой, что $w_r = -1$, находим множество $A_r(\bar{w})$ символов, которые могут быть размещены на этой позиции (см. (2)). Кроме того, находим множество $B_r(\bar{w})$ символов, которые могут быть получены из символов множества $A_r(\bar{w})$, если их сложить с r по $\mathrm{mod}\, N$:

$$B_r(\bar{w}) = \{k : k = i + r (\mathrm{mod}\, N),\ i \in A_r(\bar{w})\}.$$

3. Если $|A_r(\bar{w})| = 0$ для некоторого r, то алгоритм окончен. При этом получено «тупиковое» состояние $\bar{w} \in S^*$.

4. Предположим, что существует r, для которого $|A_r(\bar{w})| = 1$. Это означает, что на позиции r может быть размещен единственный символ $i \in A_r(\bar{w})$. Если n — количество незаполненных позиций (т.е. тех j, для которых $w_j = -1$), то полагаем

$$q := q\frac{1}{n},\ w_r = i,\ \nu_i = 1,\ \mu_k = 1, \tag{3}$$

где $k = r + i \ (\mathrm{mod}\, N)$. Далее возвращаемся на шаг 2 алгоритма.

5. Предположим, что существует символ i такой, что для некоторого r

$$i \in A_r(\bar{w}),\ i \notin A_l(\bar{w}), l \neq r.$$

В этом случае находим символ $k \in B_r(\bar{w})$ такой, что $k = r + i \ (\mathrm{mod}\, N)$. На позиции r может быть размещен лишь этот символ i. Далее проводим вычисления согласно (3) и возвращаемся на шаг 2 алгоритма.

6. Предположим, что существует символ k такой, что для некоторого r

$$k \in B_r(\overline{w}),\ k \notin B_l(\overline{w}), l \neq r.$$

В этом случае находим символ $i \in A_r(\overline{w})$ такой, что $k = r + i \pmod N$. На позиции r может быть размещен лишь этот символ i. Далее проводим вычисления согласно (3) и возвращаемся на шаг 2 алгоритма.

7. Если ни одно из трех приведенных выше условий (пп. 4–6) не выполнено, то алгоритм окончен. При этом полагаем $U(\bar{s}) = \overline{w}$ (текущее состояние) и $q(\bar{s}, \overline{w}) = q$.

Пусть $\bar{s}$ — состояние, полученное в результате этапа I алгоритма. Особенностью этапа II является дополнительное исследование, позволяющее исключить из множества $A_r(\bar{s})$ символы, размещение которых на позиции r неизбежно приведет за несколько ближайших шагов к попаданию в «тупиковое» состояние.

Этап II.

1. Находим новое состояние $\overline{w} = U(\bar{s})$. Если $\overline{w} \in S^*$, то $\hat{P}_{N1} = 0$ и алгоритм окончен. В противном случае полагаем $\hat{P}_{N1} := \hat{P}_{N1} q(\bar{s}, \overline{w})$.

2. Если состояние $\overline{w}$ не содержит компонент $w_i = -1$ (т.е. все позиции заполнены), то алгоритм окончен (построена «хорошая» перестановка). В противном случае выполняем следующий шаг.

3. Определяем r с минимальным значением $|A_r(\overline{w})|$ (если их несколько, то можно выбрать произвольное r, например наименьшее). Если $|A_r(\overline{w})| > 3$, то с одной и той же вероятностью $1/|A_r(\overline{w})|$ выбираем один из символов множества $|A_r(\overline{w})|$. Если это символ i, то полагаем

$$\hat{P}_{N1} := \hat{P}_{N1} \frac{|A_r(\overline{w})|}{n},\ \overline{w} := Z(\overline{w}, r, i)$$

и повторяем шаг 3 алгоритма.

4. Пусть $L = |A_r(\overline{w})| \le 3$ и $j_1, \ldots, j_L \in A_r(\overline{w})$ (рис. 1). Для каждого j_l выполняем следующие действия:

— находим состояние $\bar{y}_l = Z(\overline{w}, r, j_l)$;

— полагаем

$$I(\bar{y}_l) = \begin{cases} 0, & \text{если } \bar{y}_l \in S^*, \\ 1 & \text{в противном случае;} \end{cases}$$

— если $I(\bar{y}_l) = 1$, то находим состояние $\bar{x}_l = U(\bar{y}_l)$ и нормирующий множитель $q_l^{(1)} = q(\bar{y}_l, \bar{x}_l)$; если $\bar{x}_l \in S^*$, то полагаем $I(\bar{y}_l) = 0$ (на рисунке «тупиковые» состояния отмечены символом «–»);

— определяем номер позиции r_l с минимальным значением $|A_{r_l}(\bar{x}_l)|$ (если их несколько, то выбираем наименьшее r_l). Если $L_l = |A_{r_l}(\bar{x}_l)| > 3$ (рис. 1, случай $l = 2$), то полагаем $J(\bar{x}_l) = 0$ и повторяем действия шага 4 для j_{l+1}. Если $L_l \le 3$, то полагаем $J(\bar{x}_l) = 1$ и повторяем действия, описанные на шаге 4, если вместо $\overline{w}$ использовать $\bar{x}_l$ и вместо $j_1, \ldots, j_L$ — соответствующие элементы $k_1^{(l)}, \ldots, k_{L_l}^{(l)}$ множества $A_{r_l}(\bar{x}_l)$; при этом будут определены следующие величины: $\bar{y}_i^{(l)} = Z(\bar{x}_l, r_l, k_i^{(l)})$, $I(\bar{y}_i^{(l)})$, $\bar{x}_i^{(l)} = U(\bar{y}_i^{(l)})$ и $q_{li}^{(2)}$, $i = 1, \ldots, L_l$;

— если $\bar{x}_i^{(l)} \in S^*$, то полагаем $I(\bar{y}_i^{(l)}) = 0$ (рис. 1, случай $l = 1, i = 1$);

— если $I(\bar{y}_i^{(l)}) = 0$ для всех $i = 1, \ldots, L_l$, то полагаем $I(\bar{y}_l) = 0$ (рис. 1, случай $l = 3$).

5. Если $I(\bar{y}_l) = 0$ для всех $l = 1, \ldots, L$, то состояние $\overline{w}$ является «тупиковым». В этом случае алгоритм окончен и $\hat{P}_{N1} = 0$.

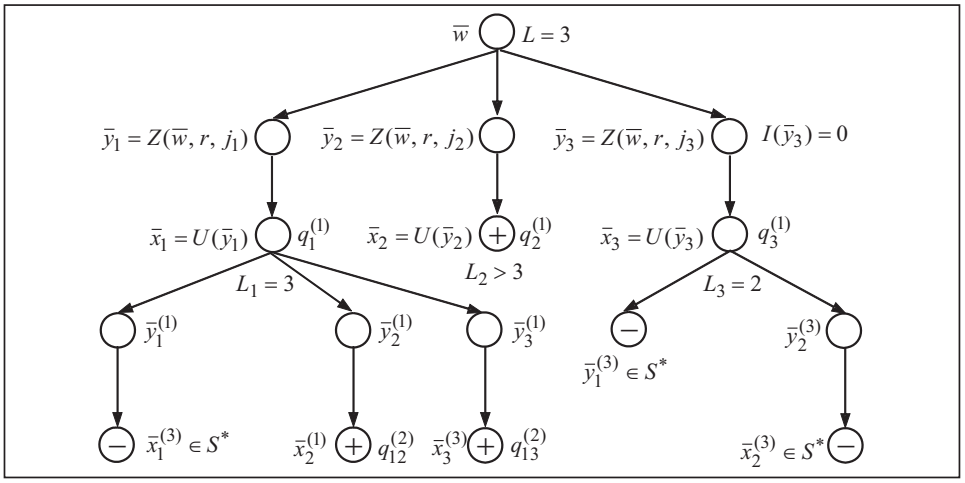

Рис. 1

6. Обозначим $n(\bar{w})$ количество незаполненных позиций при состоянии $\bar{w}$, а $C(\bar{w})$ — множество тех l, для которых $I(\bar{y}_l) = 1$, $n_1 = |C(\bar{w})|$. С одной и той же вероятностью $1/n_1$ выбираем один из элементов множества $C(\bar{w})$. Если это элемент l, то полагаем

$$\hat{P}_{N1} := \hat{P}_{N1} \frac{n_1}{n(\bar{w})} q_l^{(1)}.$$

При этом текущим состоянием становится $\bar{x}_l$.

7. Пусть $J(\bar{x}_l) = 0$ (рис. 1, случай $l = 2$). Если состояние $\bar{x}_l$ не содержит компонент со значением -1, то построена «хорошая» перестановка и алгоритм окончен. В противном случае переходим на шаг 3 алгоритма.

8. Пусть $J(\bar{x}_l) = 1$ (рис. 1, случай $l = 1$). Обозначим $n(\bar{x}_l)$ количество незаполненных позиций при состоянии $\bar{x}_l$, а $D(\bar{x}_l)$ — множество тех i, для которых $I(\bar{y}_i^{(l)}) = 1$, $n_2 = |D(\bar{x}_l)|$. С одной и той же вероятностью $1/n_2$ выбираем один из элементов множества $D(\bar{x}_l)$. Если это элемент i, то полагаем

$$\hat{P}_{N1} := \hat{P}_{N1} \frac{n_2}{n(\bar{x}_l)} q_{li}^{(2)}.$$

При этом текущим состоянием становится $\bar{x}_i^{(l)}$ (рис. 1, случай $l = 1, i = 2$; подобные состояния отмечены символом «+»). Если состояние $\bar{x}_i^{(l)}$ не содержит компонент со значением -1, то построена «хорошая» перестановка и алгоритм окончен. В противном случае полагаем $\bar{w} = \bar{x}_i^{(l)}$ и переходим на шаг 3 алгоритма.

Сформулированный алгоритм является типичным примером алгоритма направленного перебора, позволяющего существенно (на несколько десятков порядков для больших N) уменьшить дисперсию оценки. Несмещенность оценки сохраняется за счет выбора подходящих весовых множителей. При таком способе моделирования весомо возрастает доля реализаций, оканчивающихся построением «хорошей» перестановки. В отличие от алгоритма [12], приведенный алгоритм позволяет просмотреть траекторию на два шага вперед и исключить заведомо «тупиковые» состояния. За счет этого при больших N более чем в 10 раз удается повысить процент «хороших» перестановок, более чем в 50 раз понизить дисперсию оценки и, как следствие, более чем в 10 раз ускорить вычисления. Сделав необходимое число реализаций, строят оценку для P_N с требуемыми достоверностью и относительной погрешностью.

ЧИСЛЕННЫЕ РЕЗУЛЬТАТЫ

Эффективность работы алгоритма (величина, обратно пропорциональная произведению дисперсии оценки на среднее время одной реализации) существенно зависит от выбора параметра m. Численные расчеты показали, что оптимальное значение $m_{\text{opt}}^{(N)}$, максимизирующее эффективность, можно выбирать по тем же формулам, что и в [12]:

$$m_{\text{opt}}^{(3)} = 1,\ m_{\text{opt}}^{(5)} = 3,\ m_{\text{opt}}^{(7)} = 4,\ m_{\text{opt}}^{(9)} = 5,$$

$$m_{\text{opt}}^{(4k-1)} = k+3,\ m_{\text{opt}}^{(4k+1)} = k+3,\ k = 3, 4, 5, \ldots.$$

Сравним оценки, полученные методом ускоренного моделирования (МУМ) [12] и модифицированным методом ускоренного моделирования (ММУМ), предложенным выше. Все оценки, приведенные в табл. 1–5, построены при $m = m_{\text{opt}}^{(N)}$ с достоверностью 0,95 и указанной в таблицах относительной погрешностью. Вычисления производились на компьютере с процессором Pentium IV, 3.2 GHz. Использовались следующие обозначения:

- ε — относительная погрешность оценки;
- $\hat{P}_N$ — оценка, построенная для P_N с достоверностью 0,95 и относительной погрешностью ε;
- T_N — время, затраченное на построение оценки $\hat{P}_N$ с достоверностью 0,95 и относительной погрешностью ε;
- K_N — пропорция количества реализаций, в которых удалось построить «хорошие» перестановки;
- $\Delta(\hat{P}_N, \varepsilon)$ — доверительный интервал, построенный для P_N с соответствующими достоверностью и относительной погрешностью;
- $\hat{M}_N$ и $\Delta(\hat{M}_N, \varepsilon)$ — соответственно оценка и доверительный интервал, построенные для количества M_N «хороших» перестановок.

В табл. 1 для широкого диапазона значений N проведено сравнение эффективности методов МУМ [12] и ММУМ.

Таблица 1

N	ε, %	Оценка					
		МУМ			ММУМ		
		$\hat{P}_N$	T_N, с	K_N, %	$\hat{P}_N$	T_N, с	K_N, %
25	1	$2{,}70 \cdot 10^{-9}$	11,7	19,6	$2{,}68 \cdot 10^{-9}$	5,3	31,8
35	1	$2{,}02 \cdot 10^{-13}$	68,1	12,2	$2{,}01 \cdot 10^{-13}$	27,3	32,0
45	1	$1{,}32 \cdot 10^{-17}$	248,5	6,5	$1{,}33 \cdot 10^{-17}$	86,3	20,6
55	1	$8{,}08 \cdot 10^{-22}$	901,0	4,5	$8{,}11 \cdot 10^{-22}$	345,1	17,0
65	1	$4{,}75 \cdot 10^{-26}$	3 115	2,7	$4{,}72 \cdot 10^{-26}$	1 142	11,6
75	2	$2{,}66 \cdot 10^{-30}$	2 105	1,9	$2{,}71 \cdot 10^{-30}$	532,6	9,4
85	3	$1{,}44 \cdot 10^{-34}$	2 290	1,2	$1{,}48 \cdot 10^{-34}$	491,9	6,7
95	4	$7{,}76 \cdot 10^{-39}$	4 305	0,90	$8{,}08 \cdot 10^{-39}$	433,4	5,5
105	5	$4{,}06 \cdot 10^{-43}$	4 089	0,61	$4{,}12 \cdot 10^{-43}$	792,2	4,0
115	6	$2{,}07 \cdot 10^{-47}$	6 845	0,46	$1{,}99 \cdot 10^{-47}$	344,4	3,3
125	7	$1{,}12 \cdot 10^{-51}$	11 820	0,32	$1{,}08 \cdot 10^{-51}$	809,3	2,5
155	10	$1{,}36 \cdot 10^{-64}$	26 400	0,13	$1{,}35 \cdot 10^{-64}$	860,2	1,3

Прежде всего, отметим хорошее совпадение оценок по точности. За счет исключения заведомо «тупиковых» состояний удалось значительно повысить процент «хороших» перестановок (K_N увеличилось в 1,6 раза при $N = 25$ и в 10 раз при $N = 155$). При этом наблюдается заметное сокращение времени на построение оценки требуемой точности (T_N сократилось в 2,2 раза при $N = 25$ и в 30 раз при $N = 155$).

Таблица 2

N	ε, %	$\hat{P}_N$, $\Delta(\hat{P}_N, \varepsilon)$	$\hat{M}_N$, $\Delta(\hat{M}_N, \varepsilon)$	T_N, с	K_N, %
75	1	$2{,}68 \cdot 10^{-30}$ $(2{,}65 \cdot 10^{-30}, 2{,}70 \cdot 10^{-30})$	$6{,}64 \cdot 10^{79}$ $(6{,}58 \cdot 10^{79}, 6{,}71 \cdot 10^{79})$	3 080	9,4
85	1	$1{,}46 \cdot 10^{-34}$ $(1{,}44 \cdot 10^{-34}, 1{,}47 \cdot 10^{-34})$	$4{,}10 \cdot 10^{94}$ $(4{,}06 \cdot 10^{94}, 4{,}14 \cdot 10^{94})$	4 791	6,7
95	2	$7{,}87 \cdot 10^{-39}$ $(7{,}72 \cdot 10^{-39}, 8{,}03 \cdot 10^{-39})$	$8{,}13 \cdot 10^{109}$ $(7{,}97 \cdot 10^{109}, 8{,}30 \cdot 10^{109})$	2 187	5,4
105	2	$4{,}08 \cdot 10^{-43}$ $(4{,}00 \cdot 10^{-43}, 4{,}16 \cdot 10^{-43})$	$4{,}41 \cdot 10^{125}$ $(4{,}33 \cdot 10^{125}, 4{,}50 \cdot 10^{125})$	3 745	4,0
115	3	$2{,}11 \cdot 10^{-47}$ $(2{,}05 \cdot 10^{-47}, 2{,}17 \cdot 10^{-47})$	$6{,}17 \cdot 10^{141}$ $(5{,}99 \cdot 10^{141}, 6{,}36 \cdot 10^{141})$	2 940	3,3
125	3	$1{,}10 \cdot 10^{-51}$ $(1{,}07 \cdot 10^{-51}, 1{,}13 \cdot 10^{-51})$	$2{,}07 \cdot 10^{158}$ $(2{,}01 \cdot 10^{158}, 2{,}13 \cdot 10^{158})$	4 994	2,5
135	3	$5{,}52 \cdot 10^{-56}$ $(5{,}36 \cdot 10^{-56}, 5{,}69 \cdot 10^{-56})$	$1{,}49 \cdot 10^{175}$ $(1{,}44 \cdot 10^{175}, 1{,}53 \cdot 10^{175})$	5 664	2,0
145	4	$2{,}84 \cdot 10^{-60}$ $(2{,}72 \cdot 10^{-60}, 2{,}95 \cdot 10^{-60})$	$2{,}28 \cdot 10^{192}$ $(2{,}19 \cdot 10^{192}, 2{,}38 \cdot 10^{192})$	6 722	1,6
155	4	$1{,}42 \cdot 10^{-64}$ $(1{,}36 \cdot 10^{-64}, 1{,}47 \cdot 10^{-64})$	$6{,}78 \cdot 10^{209}$ $(6{,}51 \cdot 10^{209}, 7{,}05 \cdot 10^{209})$	23 068	1,3
165	4	$6{,}99 \cdot 10^{-69}$ $(6{,}71 \cdot 10^{-69}, 7{,}27 \cdot 10^{-69})$	$3{,}79 \cdot 10^{227}$ $(3{,}64 \cdot 10^{227}, 3{,}94 \cdot 10^{227})$	19 538	1,0
175	5	$3{,}44 \cdot 10^{-73}$ $(3{,}27 \cdot 10^{-73}, 3{,}61 \cdot 10^{-73})$	$3{,}87 \cdot 10^{245}$ $(3{,}67 \cdot 10^{245}, 4{,}06 \cdot 10^{245})$	18 204	0,87
185	5	$1{,}85 \cdot 10^{-77}$ $(1{,}76 \cdot 10^{-77}, 1{,}95 \cdot 10^{-77})$	$7{,}64 \cdot 10^{263}$ $(7{,}26 \cdot 10^{263}, 8{,}02 \cdot 10^{263})$	34 551	0,68
195	5	$8{,}71 \cdot 10^{-82}$ $(8{,}28 \cdot 10^{-82}, 9{,}15 \cdot 10^{-82})$	$2{,}26 \cdot 10^{282}$ $(2{,}15 \cdot 10^{282}, 2{,}37 \cdot 10^{282})$	56 915	0,57
205	5	$4{,}02 \cdot 10^{-86}$ $(3{,}82 \cdot 10^{-86}, 4{,}22 \cdot 10^{-86})$	$1{,}09 \cdot 10^{301}$ $(1{,}04 \cdot 10^{301}, 1{,}15 \cdot 10^{301})$	67 720	0,46

Повышение быстродействия алгоритма позволило значительно расширить диапазон значений N, для которых могут быть получены оценки для P_N и M_N приемлемой точности. Соответствующие оценки представлены в табл. 2 (наряду с оценками даны соответствующие доверительные интервалы).

Приведенные данные позволяют проследить рост затрат времени на получение оценок в зависимости от N и заданной относительной погрешности ε. Так, на построение оценки для $N = 205$ с относительной погрешностью 5 % требуется около 19 часов. При этом следует отметить, что «хорошую» перестановку удается строить в среднем 46 раз на каждые 10 000 реализаций (для сравнения укажем, что при использовании стандартного метода Монте-Карло было бы в среднем четыре «хороших» перестановки на 10^{86} реализаций).

ВЕРХНЯЯ И НИЖНЯЯ ОЦЕНКИ

В работе [5] выдвинута гипотеза, что при $N \to \infty$ вероятность P_N удовлетворяет соотношению (1). Большой интерес представляет оценка коэффициента c. В [12] предпринята попытка оценить c на основе статистических данных. Модифицированный метод ускоренного моделирования позволил расширить диапазон значений N, для которых могут быть построены оценки $\hat{P}_N$. Это дает возможность уточнить диапазон значений константы c. Воспользуемся тем же приемом, что и в [12]. Обозначим

$$Q_N(c) = a_N \exp\{-cN\}.$$

Параметр c подберем таким образом, чтобы при больших значениях N

$$\hat{P}_N \approx Q_N(c)$$

и при этом $\alpha_N \approx \text{const}$. Положим

$$\theta_N = \frac{a_{N+10}}{a_N} = \frac{\hat{P}_{N+10}}{\hat{P}_N} \exp\{10c\},\ N = 55, 65, \ldots, 195.$$

Тогда

$$\hat{P}_N = \hat{P}_{55} \prod_{j=1}^{k} \theta_{45+10j} \exp(-10ck),\ N = 55+10k,\ k = 1, 2, \ldots, 15. \tag{4}$$

Подберем параметр c таким образом, чтобы, начиная с некоторого N, все θ_N были как можно ближе к 1 слева. Таким значением является $c = 0,9825$ (соответствующие значения θ_N приведены в табл. 3).

Таблица 3

N	θ_N	N	θ_N	N	θ_N
55	1,0764	105	0,9557	155	0,9128
65	1,0485	115	0,9638	165	0,9106
75	1,0061	125	0,9289	175	0,9961
85	0,9996	135	0,9498	185	0,8695
95	0,9584	145	0,9219	195	0,8527

Поскольку $\theta_N \le 1$ при $N \ge 85$, из соотношения (4) получим: при $N > 75$

$$P_N \le P_{NU}^{(\text{appr})} = \hat{P}_{75} \exp\{-0,9825(N-75)\} = 268,99 \exp\{-0,9825N\}. \tag{5}$$

Аналогично строим нижнюю оценку. Подберем параметр c таким образом, чтобы, начиная с некоторого N, все θ_N были как можно ближе к 1 справа. Таким значением является $c = 0,9985$ (соответствующие значения θ_N приведены в табл. 4).

Таблица 4

N	θ_N	N	θ_N	N	θ_N
55	1,2632	105	1,1216	155	1,0712
65	1,2304	115	1,1310	165	1,0686
75	1,1807	125	1,0901	175	1,1689
85	1,1730	135	1,1146	185	1,0204
95	1,1247	145	1,0819	195	1,0006

Воспользовавшись соотношением (4), получим: при $N > 75$

$$P_N \ge P_{NL}^{(\text{appr})} = \hat{P}_{75} \exp\{-0,9985(N-75)\} = 893,08 \exp\{-0,9985N\}. \tag{6}$$

Верхние и нижние оценки (5), (6) позволяют установить границы для параметра c: $0,9825 \le c \le 0,9985$. В то же время при $c = 0,9985$ наблюдается монотонное убывание θ_N при $N \ge 175$ (табл. 4), что позволяет предположить, что верхняя граница для c может быть расширена.

В табл. 5 исследуется точность верхних и нижних оценок $P_{NL}^{(\text{appr})}$, $P_{NU}^{(\text{appr})}$.

Приведенные численные данные свидетельствуют о высокой точности как верхних, так и нижних оценок, которые позволяют прогнозировать значение P_N при $N > 75$.

Таким образом, предложенный метод является реальным инструментом, позволяющим при относительно небольших затратах времени с высокой точностью оценивать вероятность P_N выбора «хорошей» перестановки при $N \le 205$. Верхние и нижние оценки (5), (6) могут быть использованы для оценивания P_N при значительно бóльших значениях N.

Таблица 5

N	ε, %	$P_{NL}^{(appr)}$	$\hat{P}_N$, $\Delta(\hat{P}_N, \varepsilon)$	$P_{NU}^{(appr)}$
85	1	$1{,}23 \cdot 10^{-34}$	$1{,}46 \cdot 10^{-34}$ $(1{,}44 \cdot 10^{-34},\ 1{,}47 \cdot 10^{-34})$	$1{,}46 \cdot 10^{-34}$
95	2	$5{,}69 \cdot 10^{-39}$	$7{,}87 \cdot 10^{-39}$ $(7{,}72 \cdot 10^{-39},\ 8{,}03 \cdot 10^{-39})$	$7{,}87 \cdot 10^{-39}$
105	2	$2{,}62 \cdot 10^{-43}$	$4{,}08 \cdot 10^{-43}$ $(4{,}00 \cdot 10^{-43},\ 4{,}16 \cdot 10^{-43})$	$4{,}23 \cdot 10^{-43}$
115	3	$1{,}21 \cdot 10^{-47}$	$2{,}11 \cdot 10^{-47}$ $(2{,}05 \cdot 10^{-47},\ 2{,}17 \cdot 10^{-47})$	$2{,}29 \cdot 10^{-47}$
125	3	$0{,}56 \cdot 10^{-51}$	$1{,}10 \cdot 10^{-51}$ $(1{,}07 \cdot 10^{-51},\ 1{,}13 \cdot 10^{-51})$	$1{,}24 \cdot 10^{-51}$
135	3	$2{,}56 \cdot 10^{-56}$	$5{,}52 \cdot 10^{-56}$ $(5{,}36 \cdot 10^{-56},\ 5{,}69 \cdot 10^{-56})$	$6{,}70 \cdot 10^{-56}$
145	4	$1{,}18 \cdot 10^{-60}$	$2{,}84 \cdot 10^{-60}$ $(2{,}72 \cdot 10^{-60},\ 2{,}95 \cdot 10^{-60})$	$3{,}62 \cdot 10^{-60}$
155	4	$0{,}55 \cdot 10^{-64}$	$1{,}42 \cdot 10^{-64}$ $(1{,}36 \cdot 10^{-64},\ 1{,}47 \cdot 10^{-64})$	$1{,}96 \cdot 10^{-64}$
165	4	$2{,}51 \cdot 10^{-69}$	$6{,}99 \cdot 10^{-69}$ $(6{,}71 \cdot 10^{-69},\ 7{,}27 \cdot 10^{-69})$	$10{,}60 \cdot 10^{-69}$
175	5	$1{,}16 \cdot 10^{-73}$	$3{,}44 \cdot 10^{-73}$ $(3{,}27 \cdot 10^{-73},\ 3{,}61 \cdot 10^{-73})$	$5{,}73 \cdot 10^{-73}$
185	5	$0{,}53 \cdot 10^{-77}$	$1{,}85 \cdot 10^{-77}$ $(1{,}76 \cdot 10^{-77},\ 1{,}95 \cdot 10^{-77})$	$3{,}10 \cdot 10^{-77}$
195	5	$2{,}46 \cdot 10^{-82}$	$8{,}71 \cdot 10^{-82}$ $(8{,}28 \cdot 10^{-82},\ 9{,}15 \cdot 10^{-82})$	$16{,}76 \cdot 10^{-82}$
205	5	$1{,}13 \cdot 10^{-86}$	$4{,}02 \cdot 10^{-86}$ $(3{,}82 \cdot 10^{-86},\ 4{,}22 \cdot 10^{-86})$	$9{,}07 \cdot 10^{-86}$

Автор глубоко благодарен академику НАН Украины И.Н. Коваленко за привлечение внимания к указанной проблеме и критические замечания, способствовавшие улучшению статьи.

СПИСОК ЛИТЕРАТУРЫ

1. Сачков В.Н. Введение в комбинаторные методы дискретной математики. — М.: Наука, 1982. — 384 с.
2. Сачков В.Н. Цепи Маркова итерационных систем преобразований // Тр. по дискрет. математике. — 2002. — **6.** — С. 165–183.
3. Hsiang J., Hsu D.F., Shieh Y.-P. On the hardness of counting problems of complete mappings // Discrete Math. — 2004. — **277.** — P. 87–100.
4. Hsu D.F., Keedwell A.D. Generalized complete mappings, neofields, sequenceable groups and block designs. II // Pacific J. Math. — 1985. — **117.** — P. 291–312.
5. Cooper C., Gilchrist R., Kovalenko I.N., Novakovic D. Deriving the number of «good» permutations, with application to cryptography // Кибернетика и системный анализ. — 1999. — № 5. — С. 10–16.
6. Konheim R. Cryptography: a primer. — Chichester: Wiley, 1991. — 432 p.
7. Shieh Y.-P. Partition strategies for #*P*-complete problem with applications to enumerative combinatorics: Ph.D. Thesis. — Nat. Taiwan Univ., 2001.
8. http://www.research.att.com/~njas/sequences/A003111.
9. Cooper C., Kovalenko I.N. An upper bound for the number of complete mappings // Теория вероятностей и мат. статистика. — 1995. — **53.** — С. 69–75.
10. Kovalenko I.N. On an upper bound for the number of complete mappings // Кибернетика и системный анализ. — 1996. — № 1. — С. 81–85.
11. Левитская А.А. Одна комбинаторная задача в классе перестановок над кольцом Z_n вычетов по нечетному модулю n // Проблемы управления и информатики. — 1996. — № 5. — С. 99–108.
12. Кузнецов Н.Ю. Применение ускоренного моделирования к нахождению количества «хороших» перестановок // Кибернетика и системный анализ. — 2007. — № 6. — С. 80–89.

Поступила 06.11.2007