Математика и её приложения: ЖИМО. 2007. Выпуск 1 (4). С. 35 – 46.

УДК 512.54

А. А. Толстопятов[1]

# Быстрый алгоритм кодирования и декодирования поля порядка при булевом сжатии файлов

**Ключевые слова:** булева алгебра, сжатие информации.

Предложен алгоритм, позволяющий по заданной последовательности чисел кратности и перестановке этих чисел вычислить номер перестановки в лексикографически упорядоченой таблице всех перестановок и наоборот, по номеру перестановки восстановить саму перестановку. Быстрота работы алгоритма обеспечивается тем, что он не требует последовательного прохождения всей таблицы перестановок.

We suggest the algorithm that calculate for any sequence of frequency rate numbers and transposition of these numbers the number of transposition in the table of all lexicographically ordered transpositions. This algorithm allows also to restore the transposition by it's number. The algorithm does not demand the transition along the whole table of transpositions, so this provides it's high speed.

Предложенный в [2] подход к сжатию дискретной информации без потерь предполагает, что код сжимаемого файла содержит три поля: 1) поле принадлежности, 2) поле кратности, 3) поле порядка. Предложенный в [3] подход к построению поля порядка имеет тот недостаток, что требует последовательного прохождения всей лексикографически упорядоченной таблицы перестановок, что делает этот алгоритм медленным. В настоящей работе рассматривается другой подход к задаче о построении кода поля порядка, основанный на том, что вся таблица перестановок разбивается на отдельные блоки, соответствующие значениями первой, второй и т. д. числам в номере перестановки, и вычисления ведутся путем последовательного прохождения блоков, а не отдельных перестановок. Это позволяет проводить построение кода поля для более длинных буферов за приемлемое время. Похожий алгоритм, но для другой задачи, изложен в работе [1].

## 1. Постановка задачи

Исходный файл разбивается на кортежи по $n$ байтов. После этого кортежи объединяются в буферы. Для каждого буфера строится код из трех полей. Поэтому при построении алгоритмов кодирования и декодирования поля порядка достаточно рассмотреть один буфер. Пусть $m$ — число кортежей в буфере, $s$ — число различных кортежей, $n_k$ — кратность $k$-го кортежа, $j = 1, \ldots, s$. Последовательность $n_1, n_2, \ldots, n_s$ назовем *типом повторов.* Пусть $j_1, j_2, \ldots, j_m$ — перестановка из $m$ кортежей, среди которых могут быть повторяющиеся. Все перестановки для данных $m$ и

[1]Ивановский государственный университет; E-mail: khash2@mail.ru. Работа выполнена при финансовой поддержке РФФИ (проект 07-07-00155).

$n_j$ образуют таблицу, которую удобно упорядочить лексикографически. Таблица содержит

$$\sigma = \frac{m!}{\prod\limits_{k=1}^{s} n_k!} \tag{1}$$

строк. Задача о построении кода поля кратности распадается на задачу о кодировании и декодировании.

1. *Задача кодирования*: по заданной перестановке $j_1, j_2, \ldots, j_m$ вычислить ее номер $N$ в указанной выше таблице.

2. *Задача декодирования*: по заданному номеру $N$ восстановить перестановку $j_1, j_2, \ldots, j_m$.

## 2. Алгоритм кодирования

Перенумеруем кортежи, входящие в буфер, числами $1, 2, \ldots, s$; при этом одинаковые кортежи, если для них число повторов больше единицы, нумеруем одинаковыми числами. Тогда лексикографически упорядоченная таблица перестановок разобьется на блоки: первый блок включает перестановки, начинающиеся с 1, второй — с 2 и т. д. до последнего, начинающегося с $s$. Длина $L_j$ блока таблицы перестановок, начинающегося с $j$, вычисляется по формуле

$$L_j = \frac{(m-1)!n_j}{\prod\limits_{k=1}^{s} n_k!}. \tag{2}$$

Зная $j$, можно найти сколько блоков и какой длины предшествует перестановке $j_1, j_2, \ldots, j_m$ и в каком блоке таблицы находится эта перестановка. После этого вместо всей таблицы перестановок берем только тот блок, в котором находится перестановка $j_1, j_2, \ldots, j_m$, и разбиваем его аналогичным образом на подблоки в зависимости от числа кортежей $n_1$. Применяя описанную выше процедуру, определим, сколько перестановок предшествует заданной в зависимости от $j_2$. Повторим эту процедуру до тех пор, пока не доберемся до подблока, состоящего из единственной заданной перестановки. Складывая длины всех предшествующих подблоков получаем номер $N$ перестановки $j_1, j_2, \ldots, j_m$.

Введем следующие обозначения: $t$ — номер шага алгоритма, выделяющего подблоки таблицы перестановок, $m_t$ — число кортежей в перестановке на $t$-м шаге с учетом их кратности, $s_t$ — число разных кортежей в перестановке на $t$-м шаге, $(n_1^t n_2^t \ldots n_{s_t}^t)$ — тип повторов на $t$-м шаге, $(j_1^t j_2^t \ldots j_{m_t}^t)$ — перестановка из $m_t$ кортежей на $t$-м шаге. Тогда длина подблока в таблице пререстановок на $t$-м шаге есть

$$L_j = \frac{(m-t)!}{\prod\limits_{k=1}^{s_t} n_k^t!} \sum_{p=1}^{j_t-1} n_p^t \,, \tag{3}$$

а номер $N$ перестановки $(j_1^t j_2^t \dots j_{m_t}^t)$ есть

$$N = \sum_{t=1}^{m} L_t. \tag{4}$$

Подставляя (3) в (4), получим

$$N = \sum_{t=1}^{m} \frac{(m-t)!}{\prod\limits_{k=1}^{s_t} n_k^t!} \sum_{p=1}^{j_t-1} n_p^t. \tag{5}$$

Для иллюстрации рассматриваемого алгоритма маломерным примером, разобьем весь алгоритм на последовательно выполняемые операции.

(1) Берем подстановку $(j_1, j_2, \dots, j_m)$.

(2) По подстановке $(j_1, j_2, \dots, j_m)$ определяем тип повторов $(n_1, n_2, \dots, n_m)$.

(3) Вводим индекс $t$, нумерующий тип повторов на $t$-м шаге вычисления номера перестановки $N$.

(4) При $t = 1$, $(n_1', n_2', \dots, n_m')$ совпадает с $(n_1, n_2, \dots, n_m)$.

(5) По типу повторов из п. (4) определим число $s$ различных кортежей в перестановке $(j_1, j_2, \dots, j_m)$.

(6) Вычисляем $\prod\limits_{k=1}^{s_1} n_k'!$.

(7) Берем $j_1$ и вычисляем $j_1 - 1$.

(8) Вычисляем $\sum\limits_{p=1}^{j_1-1} n_p'$.

(9) Вычисляем $\frac{(m-1)!}{\prod\limits_{k=1}^{s_1} n_k'!} \sum\limits_{p=1}^{j_1-1} n_p'$.

(10) Получим тип повторов $n_1^2 n_2^2 \dots n_{s_2}^2$ для $t = 2$, сокращая на единицу кратность повтора, соответствующего числу $j_1'$.

(11) Вычеркивая из перестановки $j_1' j_2' \dots j_m'$ первое число $j_1'$, получаем перестановку для $t = 2$ : $j_2 j_3 \dots j_m$, которую обозначим через $j_1^2 j_2^2 \dots j_{m_2}^2$.

(12) Берем $n_1^2 n_2^2 \dots n_{s_2}^2$ из п. (10) и $j_1^2 j_2^2 \dots j_{m_2}^2$ из п. (11) и повторим операции пп. (6)–(9) для $t = 2$.

(13) Повторим операции пп. (10)–(12) еще $m-2$ раза, пока не дойдем до $t=m$.

(14) Складывая величины из п. (9) для $t = 1, 2, \ldots, m$, т. е. пользуясь формулой (5), находим номер перестановки $N$.

## 3. Алгоритм декодирования

Задача декодирования заключается в том, чтобы по содержащемуся в поле повторов типу повторов $n_1, n_2, \ldots, n_s$, который позволяет вычислить число кортежей в буфере $m$ по формуле

$$m = \sum_{k=1}^{s} n_k, \tag{6}$$

причем эти данные задают тип лексикографически упорядоченной таблицы перестановок поля порядка, и номеру $N$ перестановки $j_1 j_2 \ldots j_m$ из этой таблицы, восстановить саму перестановку $j_1 j_2 \ldots j_m$. Поскольку алгоритм декодирования осуществляется последовательно в несколько шагов, то удобно для первого шага обозначить $N$ через $N_1$, $\sigma$ из (1) через $\sigma_1$, а тип повторов $n_1, n_2, \ldots, n_s$ — через $n'_1, n'_2, \ldots, n'_s$. Идея алгоритма декодирования заключается в том, чтобы разбив таблицу перестановок на блоки, определить в каждом из этих блоков номер $N = N_1$ на 1-м шаге. Это дает первый номер перестановки $j_1$. Потом, разбив блок, в который попал $N = N_1$ на первом шаге на подблоки в выделив по $N_1$ и $L_j$ из (2) номер $N_2$ остающейся перестановки, в этом подблоке определить $j_2$ и т. д., пока не дойдем до $j_m$. Так же, как в алгоритме кодирования, алгоритм декодирования удобно разбить на последовательно выполняемые операции.

(1) Берем из кода поля кратности тип повторов $n'_1, n'_2, \ldots, n'_s$.

(2) Берем из кода поля порядка номер перестановки $N_1$.

(3) Вычисляем полное число перестановок $\sigma_1$ согласно формуле (1).

(4) Вычисляем длины блоков $L'_j = L_j$ таблицы перестановок согласно формуле (2).

(5) Составляем таблицу следующего типа

$$\begin{array}{llll}
1. & 1 & \le N_1 & \le L'_1, \\
2. & L'_1 + 1 & \le N_1 & \le L'_1 + L'_2, \\
 & \ldots & & \\
k. & L'_{k+1} + 1 & \le N_1 & \le \displaystyle\sum_{j=1}^{k} L'_j, \\
 & \ldots & & \\
s_1. & L'_{s+1} + 1 & \le N_1 & \le \displaystyle\sum_{j=1}^{s-1} L'_j.
\end{array}$$

(6) Определяем по таблице п. (5) номер $j_1$ интервала, в который попадает номер $N_1$.

(7) Первое число перестановок есть $j_1$.

(8) Вычисляем для 2-го шага тип повторов $n_1^2, n_2^2, \ldots, n_{s_2}^2$, сокращая на единицу кратность повтора, соответствующего числу $j_2$.

(9) Вычисляем $m_2 = m_1 - 1$.

(10) По типу повторов $n_1^2, n_2^2, \ldots, n_{s_2}^2$ из п. (8) определяем число $s_2$.

(11) По $n_1^2, n_2^2, \ldots, n_{s_2}^2$ из п. (8), $m_2$ из п. (9), $s_2$ из п. (10) вычисляем полное число перестановок $\sigma_2$ на втором шаге согласно формуле (1).

(12) Вычисляем длины подблоков $L_j^2$ таблицы перестановок блока, в который попадает перестановка, первое число которой есть $j_1$, согласно формуле (2) с $n_1^2, n_2^2, \ldots, n_{s_2}^2$ из п. (8), $m_2$ из п. (9), $s_2$ из п. (10).

(13) Составляем таблицу следующего типа

$$
\begin{array}{llll}
1. & 1 & \le N_2 & \le L_1^2, \\
2. & L_1^2 + 1 & \le N_2 & \le L_1^2 + L_2^2, \\
 & \ldots & & \\
k. & L_{k+1}^2 + 1 & \le N_2 & \le \sum\limits_{j=1}^{k} L_j^2, \\
 & \ldots & & \\
s_2. & L_{s+1}^2 + 1 & \le N_2 & \le \sum\limits_{j=1}^{s-1} L_j^2.
\end{array}
$$

(14) Вычисляем число $N_2 = N_1 - \sum\limits_{j=1}^{j_1-1} L'_j$.

(15) По таблице п. (13) находим номер интервала $j_2$, в который попадает $N_2$ из п. (14).

(16) Второе число перестановок есть $j_2$.

(17) Повторим для 3-го, 4-го, $\ldots$, $m-1$-го шага операции пп. (8)–(16).

(18) Приписываем справа от вычисленной последовательности $j_1, j_2, \ldots, j_{m-1}$ последнее число $j_m$, остающееся от $m$ чисел, определенное по типу повторов из кода поля кратности и получаем искомую перестановку $j_1 j_2 \ldots j_m$.

## 4. Малоразмерный пример

Чтобы проиллюстрировать рассматриваемый алгоритм, рассмотрим малоразмерный пример, взяв перестановку 434312.

### 4.1. Кодирование

(1) $j_1 = 4,\ j_2 = 3,\ j_3 = 4,\ j_4 = 3,\ j_5 = 1,\ j_6 = 2,\ m = 6,\ s = 4.$

(2) $n_1 = 1,\ n_2 = 1,\ n_3 = 2,\ n_4 = 2.$

(3) – (4) При $t = 1$ тип повторов $n'_1 = 1,\ n'_2 = 1,\ n'_3 = 2,\ n'_4 = 2,\ m_1 = 6,$ $s_1 = 6.$

(5) $s_1 = 2.$

(6) $\prod\limits_{k=1}^{s_1} n'_k! = \prod\limits_{k=1}^{4} n'_k! = 1!1!2!2! = 4.$

(7) $j_1 = 4\ \Rightarrow j_1 - 1 = 4 - 1 = 3.$

(8) $\sum\limits_{p=1}^{j_1-1} n'_p = \sum\limits_{p=1}^{3} n'_p = n'_1 + n'_2 + n'_3 = 1 + 1 + 2 = 4.$

(9) $\frac{(m-1)!}{\prod\limits_{k=1}^{s_1} n'_k!} \sum\limits_{p=1}^{j_1-1} n'_p = \frac{(6-1)!}{4} \cdot 4 = 120.$

(10) $j_1 = 4,\ n'_4 = 2 \Rightarrow n^2_1 = 1,\ n^2_2 = 1,\ n^2_3 = 2,\ n^2_4 = 1$ — тип повторов при $t = 2.$

(11) Перестановка при $t = 2$ есть $j_2 j_3 j_4 j_5 j_6 = 34312.$

(12) $s_2 = 4.$

(13) $\prod\limits_{k=1}^{s_2} n^2_k! = \prod\limits_{k=1}^{4} n^2_k! = 1!1!2!1! = 2.$

(14) $j_2 = 3 \Rightarrow j_2 - 1 = 3 - 1 = 2.$

(15) $\sum\limits_{p=1}^{j_2-1} n^2_p = \sum\limits_{p=1}^{2} n^2_p = n^2_1 + n^2_2 = 1 + 1 = 2.$

(16) $\frac{(m-2)!}{\prod\limits_{k=1}^{s_2} n^2_k!} \sum\limits_{p=1}^{j_1-1} n^2_p = \frac{(6-2)!}{2} \cdot 2 = 24.$

(17) $j_2 = 3 \Rightarrow n^3_1 = 1,\ n^3_2 = 1,\ n^3_3 = 1,\ n^3_4 = 1$ — тип повторов при $t = 3.$

(18) $s_3 = 4.$

(19) $\prod\limits_{k=1}^{s_3} n^3_k! = \prod\limits_{k=1}^{4} n^3_k! = 1!1!1!1! = 1.$

(20) $j_3 = 4 \Rightarrow j_3 - 1 = 3.$

(21) $\sum\limits_{p=1}^{j_3-1} n_p^3 = \sum\limits_{p=1}^{3} n_p^3 = n_1^3 + n_2^3 + n_3^3 = 1 + 1 + 1 = 3.$

(22) $\frac{(m-3)!}{\prod\limits_{k=1}^{s_3} n_k^3!} \sum\limits_{p=1}^{j_3-1} n_p^3 = \frac{(6-3)!}{1} \cdot 3 = 18.$

(23) $j_3 = 4 \Rightarrow n_1^4 = 1,\ n_2^4 = 1,\ n_3^4 = 1,\ n_4^4 = 1$ — тип повторов при $t = 4$.

(24) $s_4 = 3.$

(25) $\prod\limits_{k=1}^{s_4} n_k^4! = \prod\limits_{k=1}^{4} n_k^3! = 1!1!1! = 1.$

(26) $j_4 = 3 \Rightarrow j_4 - 1 = 2.$

(27) $\sum\limits_{p=1}^{j_4-1} n_p^4 = \sum\limits_{p=1}^{2} n_p^4 = 1 + 1 = 2.$

(28) $\frac{(m-4)!}{\prod\limits_{k=1}^{s_4} n_k^4!} \sum\limits_{p=1}^{j_4-1} n_p^4 = \frac{(6-4)!}{1} \cdot 2 = 4.$

(29) $j_3 = 3 \Rightarrow n_1^5 = 1,\ n_2^5 = 1$ — тип повторов при $t = 5$.

(30) $s_5 = 2.$

(31) $\prod\limits_{k=1}^{s_5} n_k^5! = \prod\limits_{k=1}^{2} n_k^5! = 1!1! = 1.$

(32) $j_5 = 1 \Rightarrow j_5 - 1 = 0.$

(33) $\sum\limits_{p=1}^{j_5-1} n_p^5 = \sum\limits_{p=1}^{0} n_p^5 = 0.$

(34) $\frac{(m-5)!}{\prod\limits_{k=1}^{s_5} n_k^5!} \sum\limits_{p=1}^{j_5-1} n_p^5 = \frac{(6-5)!}{1} \cdot 0 = 0.$

(35) $j_5 = 1 \Rightarrow n_1^6 = 1$ — тип повторов при $t = 6$.

(36) $s_6 = 1.$

(37) $\prod\limits_{k=1}^{s_6} n_k^6! = \prod\limits_{k=1}^{2} n_k^6! = 1! = 1.$

(38) $j_6 = 2 \Rightarrow j_6 - 1 = 1.$

(39) $\sum\limits_{p=1}^{j_6-1} n_p^6 = \sum\limits_{p=1}^{1} n_p^6 = 1.$

(40) $\frac{(m-6)!}{\prod\limits_{k=1}^{s_6} n_k^6!} \sum\limits_{p=1}^{j_6-1} n_p^6 = \frac{(6-6)!}{1} \cdot 1 = 1.$

(41) Номер перестановки $N$ равен, согласно (2), сумме чисел из пп. (9), (16), (22), (28), (33), (40), т. е. $N = 120 + 24 + 18 + 4 + 0 + 1 = 167$.

### 4.2. Декодирование

(1) Тип повторов $n'_1 = 1,\ n'_2 = 1,\ n'_3 = 2,\ n'_4 = 2,\ s_1 = 4$.
$m_1 = \sum\limits_{k=1} s_1 n'_k = \sum\limits_{k=1} 4n'_k = 6.$

(2) $N = N_1 = 167.$

(3) $\sigma = \sigma_1 = \frac{m!}{\prod\limits_{k=1}^{s} n'_k!} = \frac{6!}{1!1!2!2!} = 180.$

(4) $j = 1 \Rightarrow L'_1 = \frac{(m_1-1)!n'_1}{\prod\limits_{k=1}^{s_1} n'_k!} = \frac{(6-1)!1}{1!1!2!2!} = 30.$

$j = 2 \Rightarrow L'_2 = \frac{(m_1-1)!n'_2}{\prod\limits_{k=1}^{s_1} n'_k!} = \frac{(6-1)!1}{1!1!2!2!} = 30.$

$j = 3 \Rightarrow L'_3 = \frac{(m_1-1)!n'_3}{\prod\limits_{k=1}^{s_1} n'_k!} = \frac{(6-1)!2}{1!1!2!2!} = 60.$

$j = 4 \Rightarrow L'_4 = \frac{(m_1-1)!n'_4}{\prod\limits_{k=1}^{s_1} n'_k!} = \frac{(6-1)!2}{1!1!2!2!} = 60.$

(5) 1. $1 \le N_1 \le 30.$
2. $31 \le N_1 \le 60.$
3. $61 \le N_1 \le 120.$
4. $121 \le N_1 \le 180.$

(6-7) $N_1 = N = 167,\ 4.121 \leq 167 \leq 180 \Rightarrow j_1 = 4.$

(8) $j_1 = 4 \Rightarrow n_1^2 = 1,\ n_2^2 = 1,\ n_3^2 = 2,\ n_4^2 = 1$ — тип повторов при $t = 2$.

(9) $m_2 = m_1 - 1 = 6 - 1 = 5.$

(10) $s_2 = 4.$

(11) $\sigma_2 = \frac{m_2!}{\prod\limits_{k=1}^{s_2} n_k^2!} = \frac{5!}{1!1!2!1!} = 60.$

(12) $j = 1 \Rightarrow L_1^2 = \frac{(m_2-1)!n_1^2}{\prod\limits_{k=1}^{s_2} n_k^2!} = \frac{(4)!1}{1!1!2!1!} = 12.$

$j = 2 \Rightarrow L_2^2 = \frac{(m_2-1)!n_2^2}{\prod\limits_{k=1}^{s_2} n_k^2!} = \frac{(4)!1}{1!1!2!1!} = 12.$

$j = 3 \Rightarrow L_3^2 = \frac{(m_2-1)!n_3^2}{\prod\limits_{k=1}^{s_2} n_k^2!} = \frac{(4)!2}{1!1!2!1!} = 24.$

$j = 4 \Rightarrow L_4^2 = \frac{(m_2-1)!n_4^2}{\prod\limits_{k=1}^{s_2} n_k^2!} = \frac{(4)!1}{1!1!2!1!} = 12.$

(13) 1. $1 \leq N_2 \leq 12.$

2. $13 \leq N_2 \leq 24.$

3. $25 \leq N_2 \leq 48.$

4. $49 \leq N_2 \leq 60.$

(14) $N_2 = N_1 - \sum\limits_{j=1}^{j_1-1} L_j' = N_1 - \sum\limits_{j=1}^{3} L_j' = N_1 - L_1' - L_2' - L_2' =$
$= 167 - (30 + 30 + 60) = 47.$

(15) – (16) $N_2 = 47,\ 3.\ 25 \leq 47 \leq 48 \Rightarrow j_2 = 3.$

(17) $j_2 = 3 \Rightarrow n_1^3 = 1,\ n_2^3 = 1,\ n_3^3 = 1,\ n_4^3 = 1$ — тип повторов при $t = 3$.

(18) $m_3 = m_2 - 1 = 5 - 1 = 4.$

(19) $s_3 = 4.$

(20) $\sigma_3 = \frac{m_3!}{\prod\limits_{k=1}^{s_3} n_k^3!} = \frac{4!}{1!1!1!1!} = 24.$

(21) $j = 1 \Rightarrow L_1^3 = \frac{(m_3-1)!n_1^3}{\prod\limits_{k=1}^{s_3} n_k^3!} = \frac{(3)!1}{1!1!1!1!} = 6.$

$j = 2 \Rightarrow L_2^3 = \frac{(m_3-1)!n_2^3}{\prod\limits_{k=1}^{s_3} n_k^3!} = \frac{(3)!1}{1!1!1!1!} = 6.$

$j = 3 \Rightarrow L_3^3 = \frac{(m_3-1)!n_3^3}{\prod\limits_{k=1}^{s_3} n_k^3!} = \frac{(3)!1}{1!1!1!1!} = 6.$

$j = 4 \Rightarrow L_4^3 = \frac{(m_3-1)!n_4^3}{\prod\limits_{k=1}^{s_3} n_k^3!} = \frac{(3)!1}{1!1!1!1!} = 6.$

(22) 1. $1 \le N_3 \le 6.$

2. $7 \le N_3 \le 12.$

3. $13 \le N_3 \le 18.$

4. $19 \le N_3 \le 24.$

(23) $N_3 = N_2 - \sum\limits_{j=1}^{j_2-1} L_j^2 = N_2 - \sum\limits_{j=1}^{2} L_j^2 = N_2 - L_1^2 - L_2^2 = 47 - (12 + 12) = 23.$

(24) $N_3 = 23,\ 4.\ 19 \le 23 \le 24 \Rightarrow j_3 = 4.$

(25) $j_3 = 4 \Rightarrow n_1^4 = 1,\ n_2^4 = 1,\ n_3^4 = 1,\ n_4^3 = 1$ – тип повторов при $t = 4$.

(26) $m_4 = m_3 - 1 = 4 - 1 = 3.$

(27) $s_4 = 3.$

(28) $\sigma_4 = \frac{m_4!}{\prod\limits_{k=1}^{s_4} n_k^4!} = \frac{3!}{1!1!1!} = 6.$

(29) $j = 1 \Rightarrow L_1^4 = \frac{(m_4-1)!n_1^4}{\prod\limits_{k=1}^{s_4} n_k^4!} = \frac{(2)!1}{1!1!1!} = 6.$

$j = 2 \Rightarrow L_2^4 = \frac{(m_4-1)!n_2^4}{\prod\limits_{k=1}^{s_4} n_k^4!} = \frac{(2)!1}{1!1!1!} = 6.$

$j = 3 \Rightarrow L_3^3 = \frac{(m_3-1)!n_3^3}{\prod\limits_{k=1}^{s_3} n_k^4!} = \frac{(2)!1}{1!1!1!} = 6.$

(30) 1. $1 \leq N_4 \leq 2$.
2. $3 \leq N_4 \leq 4$.
3. $5 \leq N_4 \leq 6$.

(31) $N_4 = N_3 - \sum_{j=1}^{j_3-1} L_j^3 = N_3 - \sum_{j=1}^{3} L_j^3 = N_3 - L_1^3 - L_2^3 - L_3^3 = 23 - (6+6+6) = 5.$

(32) $N_4 = 5,\ 3.\ 5 \leq N_4 \leq 6 \Rightarrow j_4 = 3.$

(33) $j_4 = 3 \Rightarrow n_1^5 = 1,\ n_2^5 = 1$ — тип повторов при $t = 5$.

(34) $m_5 = m_4 - 1 = 3 - 1 = 2.$

(35) $s_5 = 2.$

(36) $\sigma_5 = \frac{m_5!}{\prod_{k=1}^{s_5} n_k^5!} = \frac{2!}{1!1!} = 2.$

(37) $j = 1 \Rightarrow L_1^5 = \frac{(m_5-1)!n_1^5}{\prod_{k=1}^{s_5} n_k^5!} = \frac{(2)!1}{1!1!} = 2.$

$j = 2 \Rightarrow L_2^5 = \frac{(m_5-1)!n_2^5}{\prod_{k=1}^{s_5} n_k^5!} = \frac{(2)!1}{1!1!} = 2.$

(38) 1. $1 \leq N_5 \leq 1$.
2. $2 \leq N_5 \leq 2$.

(39) $N_5 = N_4 - \sum_{j=1}^{j_4-1} L_j^4 = N_4 - \sum_{j=1}^{2} L_j^4 = N_4 - L_1^4 - L_2^4 = 5 - (2+2) = 1.$

(40) $N_5 = 1,\ 1.\ 1 \leq N_5 \leq 1 \Rightarrow j_5 = 1.$

(41) Собираем вместе $j_1$ из п. (6-7), $j_2$ из п. (15-16), $j_3$ из п. (24), $j_4$ из п. (32), $j_5$ из п. (40): $j_1 = 3,\ j_2 = 3,\ j_3 = 4,\ j_5 = 1.$

(42) Из типа повторов $n_1 n_2 n_3 n_4$ из п. (1) находим, что в системе $j_t$, $t = 1, 2, 3, 4, 5$ нет $2 \Rightarrow\ j_6 = 2.$

(43) Перестановка для $N = 167$ есть $j_1 j_2 j_3 j_4 j_5 j_6$ т. е. 434312.

В заключение отметим, что длина кода поля порядка как номер в таблице всех перестановок является минимальным по длине $L_3$. Эту длину нетрудно оценить как

$$L_3 \leq [\log_2 \sigma] \tag{7}$$

с $\sigma$ из формулы (1). Если считать, что $m >> 1,\ n_k >> 1,\ k = 1, 2, \ldots, s$, и воспользоваться формулой Стирлинга

$$m! \approx (2\pi m)^{1/2} \left(\frac{m}{e}\right)^m, \tag{8}$$

то с учетом (6) и (7) получаем следующую оценку:

$$L_3 \le \left[\frac{1+2m}{2}\log_2 m - \frac{s-1}{2}\log_2 2\pi - \sum_{k=1}^{s}\left(\frac{1+2n_k}{2}\log_2 n_k\right)\right]. \tag{9}$$

Вопрос о возможности уменьшения правой части в (9) за счет изменения лексикографического порядка в таблице всех перестановок требует дополнительного исследования.

## Список использованной литературы


1. *Гоппа В. Д.* Введение в алгебраическую теорию информации. – М.: Наука, 1995. – 112 с.
2. *Толстопятов А. А.* О структуре дискретной информации и общих условиях ее сжатия // Вестник ИвГУ. – 2002. – Вып. 3. – С. 80–82.
3. *Толстопятов А. А., Хашин С. И.* Алгоритм построения поля порядка при булевом сжатии // Вестник ИвГУ. – 2004. – Вып. 3. – С. 139–143.


*Поступила в редакцию 21.11.2007.*