УДК 004.627+517.962.27

В. А. Лужецкий, д. т. н., проф.; В. М. Михалевич, д. т. н.; проф.;
А. В. Михалевич; В. А. Каплун

ОЦЕНКА КОЛИЧЕСТВА УНИКАЛЬНЫХ ЛИНЕЙНЫХ РЕКУРРЕНТНЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ ВТОРОГО ПОРЯДКА

Линейные рекуррентные последовательности, для которых каждый член, начиная с третьего, равен сумме двух предыдущих, исследованы с точки зрения возможного сжатия и шифрования данных. Получено соотношение для определения всех членов таких последовательностей, принадлежащих определенному интервалу, при условии определенных ограничений на первые два члена. Установлено, что плотность заполнения последовательности натуральных чисел указанными членами является недостаточной для получения эффекта с высокой практической ценностью.

Ключевые слова: *линейные рекуррентные последовательности, m-значные числа, плотность, сжатие.*

Введение

Существуют подходы к сжатию, основанные на использовании статистических свойств данных и на использовании словаря, а также на определенных представлениях числовых данных [1]. Однако продолжаются поиски более эффективных методов сжатия. Один из подходов в этом направлении основан на использовании определенных функциональных зависимостей данных, представленных в виде целых чисел. В работах [2, 3] предложен подход к сжатию данных, который базируется на компактном представлении больших чисел с помощью линейных однородных рекуррентных последовательностей второго порядка. Однако отсутствуют теоретические оценки количества уникальных последовательностей такого типа, при определенных ограничениях на первые члены, которые бы, в свою очередь, позволили получить оценку потенциальных возможностей подобного способа сжатия данных.

Цель данной работы заключается в исследовании плотности заполнения членами указанных последовательностей определенных диапазонов натуральных чисел.

Задачи данной работы:

– разработка методики оценки количества уникальных линейных однородных рекуррентных последовательностей второго порядка при определенных ограничениях на первые два члена;

– получение соотношения для определения количества уникальных последовательностей указанного типа;

– получение оценки для определения плотности заполнения диапазона натуральных чисел членами исследуемых последовательностей.

Основная часть

Рассмотрим обобщенные последовательности Фибоначчи [4], а именно: последовательности, для которых каждый член, начиная с третьего, равен сумме двух предыдущих:

$$u_1, u_2, \ldots, u_n = u_{n-2} + u_{n-1}, \ldots \quad n = 3, 4, \ldots, \tag{1}$$

где u_1, u_2 – положительные целые числа.

В [5] сформулировано и доказано такое свойство: любая последовательность чисел, которая определяется соотношением (1), при m>2 содержит 4 или 5 m-значных чисел. Действительно, в соответствии с приведенным доказательством [5], формулировка этого свойства должна быть следующей: для любых u_1, u_2 последовательности (1) существует такое целое число $M=M(u_1, u_2)$, что для всех $m \geq M$ в последовательности встречается только 4 или 5 m-значных чисел.

Приведенное формулирование является обобщением свойства, сформулированного в [6] относительно последовательности чисел Фибоначчи, для которой M=2.

Основываясь на замкнутом соотношении для общего члена последовательности (1)

$$u_n = u_1 \cdot \frac{\alpha^{n-2} - \beta^{n-2}}{\sqrt{5}} + u_2 \cdot \frac{\alpha^{n-1} - \beta^{n-1}}{\sqrt{5}}, \tag{2}$$

где

$$\alpha = \frac{1+\sqrt{5}}{2}, \quad \beta = \frac{1-\sqrt{5}}{2}, \tag{3}$$

можно получить формулу для вычисления количества k_m m-значных чисел последовательности (1) в зависимости от значений m, u_1, u_2:

$$k_m = \begin{cases} 5, & \{M\} \leq \left\{\dfrac{1}{\lg\alpha}\right\} \approx 0,78972, \\ 4, & \{M\} > \left\{\dfrac{1}{\lg\alpha}\right\} \approx 0,78972, \end{cases} \tag{4}$$

$$M = \frac{0,5 \cdot \lg 5 + m - \lg\left(\dfrac{u_2 - u_1}{\alpha} + u_1\right)}{\lg\alpha},$$

где дробная часть числа x обозначена $\{x\}$.

Непосредственным вычислениям в среде системы символьной математики Maple была осуществлена проверка правильности соотношения (4) для следующих диапазонов варьирования аргументов: m=8÷20, u_1=1÷300, u_2=1÷300.

Фиксированием двух параметров из трех и изменением третьего параметра в указанном диапазоне исследовалась частота появления значений k_m=4, k_m=5. На основе полученных данных построена таблица распределения возможного количества k_m m-значных чисел последовательности (1).

Таблица 1

Распределение k_m

$(k_n)_i$	4	5
p_i	0,21	0,79

Такой же результат следует из соотношения (4) и предположения о равномерном распределении мантисс чисел, определяемых соотношением

$$\left\{\left(0,5 \cdot \lg 5 + m - \lg\left(u_2 - u_1/\alpha + u_1\right)\right) \big/ \lg\alpha\right\}.$$

Очевидно, что плотность заполнения ряда натуральных чисел членами отдельной рекуррентной последовательности (1) очень мала уже для $m>4$ и быстро уменьшается с увеличением m. Вместе с тем, для оценки перспективности применения преобразований целых чисел с помощью последовательностей (1), необходимо оценить количество уникальных последовательностей при определенных ограничениях на первые два члена.

Ограничим диапазон изменения первых двух членов последовательности отрезком [1, 99]. Вычислим общее количество последовательностей (1), которые могут быть образованы при данных ограничениях. Это количество равно числу размещений с повторениями из множества N = 99 элементов по k = 2 элемента:

$$\overline{A}_N^k = N^k = 99^2 = 9801 . \qquad (5)$$

Однако не все такие последовательности уникальны. Под множеством уникальных последовательностей будем понимать такой набор последовательностей, для каждой из которых выполняется следующее условие: данная последовательность не может быть получена отбрасыванием конечного числа $p > 0$ первых членов любой другой последовательности данного множества. Приведем пример двух последовательностей типа (1):

1) 3, 2, 5, 7, 12, 19, 31, 50, 81, 131, 212, 343, 555, 898, 1453, 2351, 3804, 6155, 9959, 16114, 26073, 42187, …

2) 50, 81, 131, 212, 343, 555, 898, 1453, 2351, 3804, 6155, 9959, 16114, 26073, 42187, 68260,110447, …

Очевидно, что вторую последовательность можно получить из первой отбрасыванием ее первых семи членов.

Итак, нашей задачей является определение множества уникальных последовательностей и вычисление их количества при наличии определенных ограничений на диапазон изменения первых двух членов последовательности. Для этого предложен следующий алгоритм:

1. Создается пустое множество $M^{(2)}$, элементами которого являются упорядоченные пары чисел.

2. Для каждой из 9801 пары чисел $a_1 \in [1, 99]$, $a_2 \in [1, 99]$ последовательность (1) продолжается влево до наименьшего положительного u_1. Фиксируется полученная пара чисел $[u_1, u_2]_i$.

3. Проверяется существование текущего элемента в множестве $M^{(2)}$. Если такой элемент не существует, то текущая пара чисел вносится в указанное множество.

В результате формируется множество $M^{(2)}$ уникальных последовательностей (1), каждая из которых представлена ее первыми двумя членами u_1, u_2, удовлетворяющими условия $u_2 - u_1 \le 0$.

На рис. представлена графическая иллюстрация к расположению уникальных последовательностей для различных диапазонов возможных значений первых двух членов последовательностей (вертикальная колонка чисел – значения u_1, горизонтальная – u_2). Знаком “+” обозначены все клетки, которые являются пересечением пары чисел, порождающих уникальную последовательность. Знаком “-” – неуникальную последовательность.

На основе приведенных данных построено соотношение для вычисления количества уникальных последовательностей q_N, первые два члена которых удовлетворяют условия

$$u_2 - u_1 \le 0;\ u_1 \in [1, N],\ u_2 \in [1, N] . \qquad (6)$$

Данное соотношение имеет вид

$$q_N = N^2 - \sum_{i=1}^{\frac{N-1}{2}} (2i-1) - N + 1 - \sum_{i=1}^{\frac{N-3}{2}} 2i . \qquad (7)$$

После несложных преобразований получим соотношение для вычисления количества уникальных последовательностей q_N в зависимости от диапазона изменения первых двух членов

$$q_N = \frac{1}{2} \cdot N \cdot (N+1) . \qquad (8)$$

При $N = 99$ эффект уплотнения может быть обнаружен на 8-значных числах. Согласно (4), предположим, что количество k_m=5 m-значных чисел последовательности (1) встречается

	1	2	3	4	5	6	7	8	9
1	+	-	+	+	+	+	+	+	+
2	-	+	-	-	+	+	+	+	+
3	-	-	+	-	-	-	+	+	+
4	-	-	-	+	-	-	-	-	+
5	-	-	-	-	+	-	-	-	-
6	-	-	-	+	+	+	-	-	-
7	-	-	+	+	+	+	+	-	-
8	-	+	+	+	+	+	+	+	-
9	+	+	+	+	+	+	+	+	+

	1	2	3	4	5	6	7	8	9	10
1	+	-	+	+	+	+	+	+	+	+
2	-	+	-	-	+	+	+	+	+	+
3	-	-	+	-	-	-	+	+	+	+
4	-	-	-	+	-	-	-	-	+	+
5	-	-	-	-	+	-	-	-	-	-
6	-	-	-	-	+	+	-	-	-	-
7	-	-	-	+	+	+	+	-	-	-
8	-	-	+	+	+	+	+	+	-	-
9	-	+	+	+	+	+	+	+	+	-
10	+	+	+	+	+	+	+	+	+	+

	1	2	3	4	5	6	7	8	9	10	11
1	+	-	+	+	+	+	+	+	+	+	+
2	-	+	-	-	+	+	+	+	+	+	+
3	-	-	+	-	-	-	+	+	+	+	+
4	-	-	-	+	-	-	-	-	+	+	+
5	-	-	-	-	+	-	-	-	-	-	+
6	-	-	-	-	-	+	-	-	-	-	-
7	-	-	-	-	+	+	+	-	-	-	-
8	-	-	-	+	+	+	+	+	-	-	-
9	-	-	+	+	+	+	+	+	+	-	-
10	-	+	+	+	+	+	+	+	+	+	-
11	+	+	+	+	+	+	+	+	+	+	+

	1	2	3	4	5	6	7	8	9	10	11	12	13
1	+	-	+	+	+	+	+	+	+	+	+	+	+
2	-	+	-	-	+	+	+	+	+	+	+	+	+
3	-	-	+	-	-	-	+	+	+	+	+	+	+
4	-	-	-	+	-	-	-	-	+	+	+	+	+
5	-	-	-	-	+	-	-	-	-	-	+	+	+
6	-	-	-	-	-	+	-	-	-	-	-	-	+
7	-	-	-	-	-	-	+	-	-	-	-	-	-
8	-	-	-	-	-	+	+	+	-	-	-	-	-
9	-	-	-	-	+	+	+	+	+	-	-	-	-
10	-	-	-	+	+	+	+	+	+	+	-	-	-
11	-	-	+	+	+	+	+	+	+	+	+	-	-
12	-	+	+	+	+	+	+	+	+	+	+	+	-

Рис. Иллюстрация к закономерностям расположения уникальных последовательностей для разных диапазонов возможных значений первых двух членов последовательности (1) в $\frac{0,79}{0,21} \approx 3,76$ раз чаще, чем k_m=4. Тогда плотность заполнения диапазона натуральных чисел $10^7 \div 10^8$ -1 членами последовательности (1) можно определить как отношение всех чисел последовательности, которые попадают в заданный диапазон, к длине этого диапазона:

$$p_2 = \frac{q_{99} \cdot (5 \cdot 0,79 + 4 \cdot 0,21)}{(10^8 - 1) - (10^7 - 1)} = \\ = \frac{4950 \cdot 4,79}{9 \cdot 10^7} \approx 0,00027 \quad (9)$$

С увеличением диапазона изменения первых двух чисел последовательности возрастает и длина чисел, для которых проявляется эффект сжатия. Обобщая (8) при условии N=10^r-1 (r=2, 3, …), запишем

$$p_r = \frac{10^r \cdot (10^r - 1) \cdot 4,79}{2 \cdot (10^{2r+4} - 10^{2r+3})} \approx \\ \approx \frac{4,79}{18 \cdot 10^3} \approx 0,00027 \quad (10)$$

откуда следует, что плотность заполнения последовательности натуральных чисел числами последовательности (1) не изменяется с ростом диапазона N по указанному закону.

В соответствии с соотношением (9), количество m-значных чисел для всех последовательностей (1), которые удовлетворяют условия (6) при N=99, равна 4950*4,79≈23711. Для проверки точности полученной оценки плотности p_r по специально составленной программе в среде системы символьной математики Maple определено множество m-значных чисел для разных значений m. Результаты вычислений приведены в табл. 2.

Таблица 2

Количество n_m m-значных чисел последовательности

m	n_m	Погрешность δ, %
8	23667	0,19
9	23682	0,12
10	23712	0,004
11	23692	0,08
12	23668	0,18

Очевидно, что полученная плотность слишком мала и это обусловливает необходимость дальнейших поисков для получения результатов, имеющих большую практическую ценность.

Выводы

1. Разработана методика определения количества уникальных рекуррентных последовательностей (1), соответствующих условиям (6). Построенная для нескольких частных случаев наглядная иллюстрация к определению уникальных рекуррентных последовательностей (1) позволила получить соотношение для вычисления количества уникальных последовательностей и суммарного количества их членов, принадлежащих определенному интервалу (в зависимости от диапазона изменения первых двух членов).

2. Предложены соотношения для определения плотности заполнения диапазона натуральных чисел членами последовательности (1) в виде отношения количества всех членов уникальных последовательностей, которые попадают в заданный диапазон, к длине этого диапазона.

3. Теоретическая оценка указанной плотности не превышает $3 \cdot 10^{-4}$ и при определенных условиях не изменяется с ростом интервала варьирования первых двух членов последовательностей и соответствующего увеличения исследуемого диапазона.

4. Экспериментальная проверка с помощью специально составленной программы в среде системы символьной математики Maple подтвердила достоверность выводов для определенного диапазона значений первых двух членов последовательности.

5. Невысокое значение плотности обусловливает необходимость дальнейших поисков в выбранном направлении.

СПИСОК ЛИТЕРАТУРЫ

1. Ватолин Д. Методы сжатия данных. Устройство архиваторов, сжатие изображений и видео / Д. Ватолин, А. Ратушняк, М. Смирнов, В. Юкин. – М.: ДИАЛОГ-МИФИ, 2003. – 384 с.

2. Анисимов А. В. Обратное преобразование Фибоначчи / А. В. Анисимов, Я. П. Рындин, С. Е. Редько // Кибернетика. – 1982. – № 3. – С. 9 – 11.

3. Кшановський О. Д. Арифметичні методи ущільнення цифрової інформації / О. Д. Кшановський, С. В. Тітарчук, В. А. Лужецький // Вісник ВПІ. – 1999. – № 5. – С. 83 – 87.

4. Воробьёв Н. Н. Числа Фибоначчи / Н. Н. Воробьёв. – М.: Наука, 1978. – 144 с.

5. Лужецький В. А. Щільність заповнення ряду натуральних чисел членами лінійних рекурентних послідовностей другого порядку / В. А. Лужецький, В. М. Михалевич, О. В. Михалевич, В. А. Каплун // Вісник Вінницького політехнічного університету. – 2010. – № 4. – С. 41 – 45.

6. Алфутова Н. Б. Алгебра и теория чисел. Сборник задач для математических школ / Н. Б. Алфутова, А. В. Устинов. – М.: МЦНМО, 2002. – 264 с. – ISBN 5-94057-038-0.

Лужецкий Владимир Андреевич – д. т. н., профессор, заведующий кафедрой защиты информации, тел.: (0432)-598-386, E-mail: lva_zi@mail.ru.

Михалевич Владимир Маркусович – д. т. н., профессор, заведующий кафедрой высшей математики, тел.: (0432)-598-594, (0432)-598-591, E-mail: vmykhal@gmail.com.

Михалевич Алексей Владимирович – студент гр. 1БС-07, e-mail: mikhal.alex@gmail.com.

Каплун Валентина Аполинариевна – старший преподаватель кафедры защиты информации.

Винницкий национальный технический университет.