ДИСКРЕТНАЯ МАТЕМАТИКА И МАТЕМАТИЧЕСКАЯ КИБЕРНЕТИКА

Научная статья

УДК 519.16; 519.142
DOI: 10.18101/2304-5728-2025-3-29-37

АЛГОРИТМИЧЕСКОЕ ИССЛЕДОВАНИЕ ТРЕУГОЛЬНИКА ТИПА ПАСКАЛЯ И ПОМЕХОУСТОЙЧИВЫЕ КОДЫ

© **Кузьмин Олег Викторович**
доктор физико-математических наук, профессор,
заведующий кафедрой теории вероятностей и дискретной математики,
Иркутский государственный университет
Россия, 664003, г. Иркутск, б-р Гагарина, 20
quzminov@mail.ru

© **Терехова Анастасия Вячеславовна**
техник кафедры теории вероятностей и дискретной математики,
Иркутский государственный университет
Россия, 664003, г. Иркутск, б-р Гагарина, 20
terekhovanastena@yandex.ru

Аннотация. Изучается структура треугольника типа Паскаля в позиционной системе счисления с основанием p. С помощью рекурсивных функций на объектно-ориентированном языке программирования C++ был реализован вывод структуры треугольника типа Паскаля по модулю p. В ходе работы используется формула триномиального коэффициента, полученная с помощью двух биномиальных коэффициентов. Разработан алгоритм для выявления свойств рабочих кодовых комбинаций из таблиц типа Паскаля, использующихся для помехоустойчивого кодирования и декодирования информации с исправлением ошибок. Рассматривается граничное условие, связывающее количество кодируемых сообщений, позиционную систему и длину кодируемого сообщения. Выявлено максимальное количество ошибок, которые возможно исправить. В программе реализуются формулы избыточности кода и расстояния Хэмминга. В ходе работы анализируются возможные рабочие комбинации, полученные на основе расстояния Хэмминга.

Ключевые слова: треугольник Паскаля, биномиальные коэффициенты, p-ичная кодировка, помехоустойчивое кодирование, избыточность кода, расстояние Хэмминга, декодирование.

Для цитирования

Кузьмин О. В., Терехова А. В. Алгоритмическое исследование треугольника типа Паскаля и помехоустойчивые коды // Вестник Бурятского государственного университета. Математика, информатика. 2025. № 3. С. 29–37.

Введение

Помехоустойчивое кодирование является актуальной в настоящее время задачей теории информации. Данной задачей занимаются криптография и дискретная математика. Каналы связи — не идеальные проводники для передачи информации, достаточно часты случаи изменения кодовых комбинаций из-за влияния шума или различного рода помех. Для исправления таких ошибок в код добавляют избыточные данные, например, биты четности в коде Хэмминга. На данный момент придумано множество помехоустойчивых кодов, каждый из которых имеет свои преимущества и область применения (например, [3][4][5]).

1 Позиционная система счисления

Под *позиционной системой* счисления понимают систему счисления, в которой значение каждой цифры (базового элемента) может варьироваться в зависимости от позиции (разряда) в числе, где он находится. Количество элементов (основание системы) обозначается числом p, где $p \in \mathbb{N},\ p > 1$. Позиционную систему, в которой p элементов, будем называть p-ичной системой счисления. В такой системе используются цифры от 0 до $p-1$; если $p > 10$, то вводятся буквенные обозначения для базовых элементов.

Одним из важных и, пожалуй, интересных свойств является соответствие лексикографического порядка записей естественному порядку, но записи чисел должны быть слева дополнены ведущими незначащими нулями. Например, 101_2 и 1001_2 для сравнения этих чисел нужно дополнить нулями первое число 0101_2 и 1001_2 (видно, что уже в четвертом разряде чисел есть расхождение) $0 < 1 \implies 0101_2 < 1001_2$.

Применение p-ичной системы достаточно эффективно с точки зрения затрат памяти и ресурсов компьютера или других устройств, где каждый базовый элемент отвечает за определенное состояние единицы информации в системе. Например, использование троичных алгоритмов в ЭВМ куда более эффективно с точки зрения плотности записи информации, равной $\frac{\ln p}{p}$, принимающей максимальное значение при $p = e$, но так как $p \in \mathbb{N}$, то при p=3.

2 Треугольник Паскаля

Треугольником Паскаля называют бесконечную треугольную таблицу, элементы которой являются числами сочетаний, иначе называемыми биномиальными коэффициентами. Данная структура содержит в себе и натуральные, и треугольные числа, и числа Фибоначчи [1].

Числом сочетаний $\mathrm{C}_{\mathrm{n}}^{\mathrm{k}}$ называется количество способов выбрать неупорядоченное k-элементное подмножество из n-элементного множества.

$$C_n^k = \frac{n!}{k!(n-k)!}, \text{ где } n \geq k \geq 0 \text{ и } n, k \in \mathbb{Z}$$

```
for (int i = 0; i < n; i++) {
    Pascal_triangle[i].resize(i + 1);
    Pascal_triangle[i][0] = Pascal_triangle[i][i] = 1;
    for (int j = 1; j < i; j++) {
      Pascal_triangle[i][j] = (((Pascal_triangle[i - 1][j - 1])) +
          (Pascal_triangle[i - 1][j]))%c;
    }
}
```

Данный алгоритм [1] реализует моделирование структуры треугольника с помощью рекуррентной формулы. Здесь числа хранятся в двумерном массиве, где i-й подмассив — это i-я строка треугольника Паскаля, числа в котором рассматриваются по модулю p.

Соответствующие нули при выводе в консоль заменим на «.» для визуальной красоты.

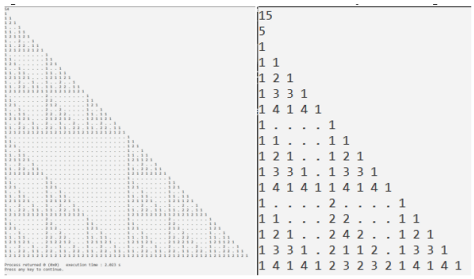

Рис. 1. Вывод структуры треугольника в консоль [1]

Элементы i-й строки треугольника представляют собой коэффициенты разложения степени бинома Ньютона $(a+b)^i$. Представление элементов треугольника Паскаля по модулю p позволяет нам заметить интересные свойства данной структуры, например, самоподобие треугольника, превращающее его во фрактал (треугольник Серпинского); мелкость разбиения зависит от p (например, [2][7]).

3 Помехоустойчивые коды

В XXI в. в связи с развитием высокоскоростного Интернета резко возросла необходимость в системах передачи и хранения информации. Главной задачей в подобных системах является помехоустойчивая передача данных. До 1948 г. считалось, что наличие шумов в канале связи — камень преткновения в безошибочной передаче данных, пока Клод Элвуд Шеннон не опубликовал основополагающую работу по теории информации, в которой была показана принципиальная возможность

обеспечения надёжной передачи данных со скоростью, не превышающей пропускной способности канала. Однако теорема, сформулированная Шенноном, лишь доказывает существование таких кодов, но не объясняет, как такие коды строить и как декодировать [8][9][10].

Если необходимо закодировать T сообщений равномерным p-ичным кодом, то нужно выбрать такую длину кодируемого сообщения n, чтобы выполнялось нестрогое неравенство $T \leq p^n$.

Данное неравенство называют *граничным условием*. В случае выполнения равенства при кодировании сообщений будут использоваться всевозможные p-ичные числа, в данном случае код не будет помехоустойчивым, ибо любая ошибка в комбинации повлечет за собой её смену на другую комбинацию, без возможности обнаружить ошибку. В случае строгого неравенства у нас есть возможность обеспечить длину строки $n > k$, где k — количество информационных элементов и $k = log_p T$, в связи с чем появляется возможность выделить избыточные $r = n - k$ элементов, которые будут необходимы для построения помехоустойчивого кода. Число n должно позволить получить T' кодовых комбинаций $T' > T$, из которых можно будет выбрать T комбинаций, которые будут называться рабочими, а остальные — нерабочими. При получении нерабочей комбинации, во-первых, мы сразу будем знать об ошибке, во-вторых, сможем найти наиболее похожую комбинацию среди имеющихся и исправить ошибку [6].

Избыточностью кода называют свойство кода, которое позволяет исправить ошибки, допущенные при передаче или приёме символов сообщения, она вычисляется по формуле:

$$g = \frac{r}{n}. \tag{1}$$

Необходимо реализовать формулу (1) в алгоритме, так как k — количество информационных элементов, значит, k должно быть целым, используя функцию ceil из библиотеки cmath, округлим значение k вверх:

```
float redundancy_bonacci(int T, int p, int n){
    float k = (ceil((log(T) / log(p))));
    float g = (n - k)/n;
    return g;
}
```

Введем несколько новых определений, необходимых для дальнейшего исследования.

Расстоянием Хэмминга d(x,y) между любой парой комбинаций x и y называют количество позиций, в которых они различаются. Больший интерес для нас представляет минимальное кодовое расстояние между последовательностями:

$$d_{\min} = \min\{d(x, y) : \forall x, y \in T, x \neq y\}. \tag{2}$$

Реализуем данную формулу в алгоритме:

```
int hamming_distance_bonacci(vector<int> a,vector<int> b){
    int d = 0;
    for (int i = 0; i < a.size(); i++){
        if (a[i]!=b[i]){
            d++;
        }
    }
    return d;
}

int d = cols + 1;
for (int i = 0; i < rows; i++){
            for (int j = 0; j < rows; j++){
                if (i!=j){
                    int a = hamming_distance_bonacci(like_pascal[i
                        ],like_pascal[j]);
                    if (d > a){d = a;}
                }
            }
}
```

Если попытаться выделить кодовую таблицу из треугольника Паскаля, то получим прямоугольник Тартальи, у которого первая строка и первый столбец называются *образующими* и состоят из единиц, что значительно уменьшит количество уникальных последовательностей.

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 3 & 6 \end{pmatrix}$$

В связи с этим рассмотрим треугольник типа Паскаля с образующими из последовательных натуральных чисел:

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 7 \\ 3 & 7 & 14 \end{pmatrix} \tag{3}$$

Кодом типа Паскаля длины p^n называется код, задаваемый таблицей (3) размерности $p^n \times p^{n+1}$, если все элементы рассматриваются по *mod p*.

Для иллюстрации метода рассмотрим подобную таблицу 3×9 при $p = 3$:

$$\begin{pmatrix} 1 & 2 & 0 \\ 2 & 1 & 1 \\ 0 & 1 & 2 \\ 1 & 2 & 1 \\ 2 & 1 & 2 \\ 0 & 1 & 0 \\ 1 & 2 & 2 \\ 2 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \tag{4}$$

В данном случае $d_{\min} = 1,\ T < p^n$. Код является помехоустойчивым, но так как $d_{\min} = 1$ при воздействии шума мы не всегда сможем при получении сигнала найти и исправить ошибку, ибо при изменении одного элемента в комбинации мы либо получим другую комбинацию из кода типа Паскаля и вовсе не обнаружим ошибки, либо получим последовательность, которой нет в нашей изначальной таблице, и ошибка будет обнаружена. Теперь появилась необходимость в коде, который всегда будет обнаруживать хотя бы b ошибок.

При $b = 1,\ d_{\min} = 1$, так как при смене одной позиции в комбинации, во-первых, ошибка будет обнаружена, так как измененной последовательности заведомо нет в таблице, во-вторых, мы всегда сможем найти комбинацию, расстояние между которой и полученной будет равно 1. При $b = 2,\ d_{\min} \geq 3$. В общем случае $d_{\min} \geq b + 1$. Заметим, что при $b = n$ (изменение всех элементов последовательности) $d_{\min} \geq n + 1$, что невозможно обеспечить в данном коде.

Чтобы обеспечить наперед заданное минимальное расстояние, нужно выбрать рабочие комбинации. Для этого необходимо рассчитать всевозможные расстояния Хэмминга с помощью алгоритма:

```
vector<vector<int>> dmin(rows, vector<int>(rows));
        int d = cols + 1;
        for (int i = 0; i < rows; i++){
            for (int j = 0; j < rows; j++){
                if (i!=j){
                    int a = hamming_distance_bonacci(like_pascal[i
                        ],like_pascal[j]);
                    dmin[i][j]=a;
                    cout<<dmin[i][j];
                    if (d > a){d = a;}
                }
                else{
                    dmin[i][j]=0;
                    cout<<dmin[i][j];
                }
            }
            cout<<endl;
        }
```

Для приведенного ранее примера (4) с помощью алгоритма все расстояния составим в виде таблицы 9×9, где в i-й строке будут даны всевозможные расстояния между i-й комбинацией и другими последовательностями.

```
1 2 0
2 1 1
0 1 2
1 2 1
2 1 2
0 1 0
1 2 2
2 1 0
0 1 1
-----
033132123
302212311
320311221
123033132
311302212
221320311
132123033
212311302
311221320
```

Основываясь на этих данных, мы можем выбрать три рабочие комбинации.

$$\begin{pmatrix} 1 & 2 & 0 \\ 2 & 1 & 1 \\ 0 & 1 & 2 \end{pmatrix}$$

При увеличении числа обрабатываемых ошибок количество рабочих комбинаций уменьшается. Данный случай необходимо рассматривать на кодах типа Паскаля при больших значениях n.

Заключение

Результаты показывают перспективу применения p-ичной кодировки треугольника типа Паскаля при решении задач кодирования информации. Кроме того, получены помехоустойчивые коды типа Паскаля при различных p. Дальнейшие исследования будут направлены на изучение других помехоустойчивых таблиц, расширение и оптимизацию уже имеющегося алгоритма.

Литература

1. Кузьмин О. В., Терехова А. В. Алгоритмические методы исследования треугольника и пирамиды Паскаля // Комбинаторные и вероятностные свойства дискретных структур: сборник научных трудов / под редакцией О. В. Кузьмина. Иркутск: Изд-во ИГУ, 2025. С. 41–47 (Дискретный анализ и информатика; вып. 11).

2. Кузьмин О. В., Старков Б. А. Фрактальные свойства бинарных матриц, построенных при помощи арифметики треугольника Паскаля, и помехоустойчивое кодирование // Современные технологии. Системный анализ. Моделирование. 2016. № 4(52). С. 138–142.

3. Кузьмин О. В., Старков Б. А. Бинарные матрицы, построенные при помощи треугольника Паскаля, и помехоустойчивое кодирование // Современные технологии. Системный анализ. Моделирование. 2016. № 1(49). С. 112–117.

4. Шавва А. И. Кодирование числовой информации // Теория и практика современной науки. 2017. № 6(24). С. 890–892.

5. Алешников С. И., Алексеенко Е. С. Математические методы защиты информации: учебное пособие. Ч. 5. Калининград: Изд-во Балт. федер. ун-та им. Иммануила Канта, 2010. 157 с.

6. Кузьмин О. В., Оркина К. П. Построение кодов, исправляющих ошибки, с помощью треугольника типа Паскаля // Вестник Бурятского университета. Сер. 13. Математика и информатика. 2006. № 3. С. 32–39.

7. Задорожный В. Н. Общая статистическая структура простейших клеточных автоматов // Омский научный вестник. 2005. № 2(31). С. 150–156.

8. Кротова Е. И. Исследование помехоустойчивости неравномерного двоичного кода по алгоритму Шеннона-Фэно // Известия ЮФУ. Технические науки. 2006. № 7. С. 152–158.

9. Использование информационной избыточности при построении сбоеустойчивых комбинационных схем / С. В. Гаврилов, С. И. Гуров, Д. В. Тельпухов, Т. Д. Жукова // Таврический вестник информатики и математики. 2018. № 2(39). С. 29–44.

10. Криксина А. А., Левенец А. В. Сравнительный анализ эффективности помехоустойчивых кодов // Информационные технологии XXI века: сборник научных трудов. Хабаровск: Изд-во Тихоокеан. гос. ун-та, 2017. С. 59–66.

Статья поступила в редакцию 30.09.2025; одобрена после рецензирования 06.10.2025; принята к публикации 08.10.2025.

ALGORITHMIC STUDY OF THE PASCAL-TYPE TRIANGLE AND ERROR-CORRECTING CODES

Oleg V. Kuzmin
Dr. Sci. (Phys. and Math.), Professor,
The Department Head of Probability Theory and Discrete Mathematics
Irkutsk State University
20 Gagarin Blvd., Irkutsk, 664003, Russia

Anastasia V. Terekhova
The Department Technician of Probability Theory and Discrete Mathematics
Irkutsk State University
20 Gagarin Blvd., Irkutsk, 664003, Russia

Abstract. The structure of a Pascal-type triangle in a positional numeral system with base p is studied. Recursive functions in the object-oriented programming language C++ are used to derive the structure of a Pascal-type triangle modulo p. A formula for the trinomial coefficient, obtained using two binomial coefficients, is used in the study. An algorithm is developed for identifying the properties of working code combinations from Pascal-type tables used for error-correcting encoding and decoding of information. A boundary condition relating the number of encoded messages, the positional system, and the length of the encoded message is considered. The maximum number of errors that can be corrected is identified. The program implements code redundancy and Hamming distance formulas. Possible working combinations obtained based on the Hamming distance are analyzed.

Keywords: Pascal's triangle, binomial coefficients, p-ary coding, error-correcting coding, code redundancy, Hamming distance, decoding.

For citation
Kuzmin O. V., Terekhova A. V. Algorithmic Study of the Pascal-Type Triangle and Error-Correcting Codes // Bulletin of Buryat State University. Mathematics, Informatics. 2025. N. 3. P. 29–37.

The article was submitted 30.09.2025; approved after reviewing 06.10.2025; accepted for publication 08.10.2025.