На правах рукописи

Курмаев Олег Феатьевич

Нумерационное кодирование двоичных последовательностей с ограничениями на длины серий, вес, заряд

05.13.17 – Теоретические основы информатики

АВТОРЕФЕРАТ
диссертации на соискание ученой степени
кандидата физико-математических наук

Москва – 2011

Работа выполнена в *Московском государственном институте электронной техники (Национальном исследовательском университете)*.

Научный руководитель:	доктор технических наук, профессор В. А. Кустов
Официальные оппоненты:	доктор физико-математических наук, профессор Дьячков Аркадий Георгиевич кандидат физико-математических наук Лебедев Владимир Сергеевич
Ведущая организация:	Санкт-Петербургский государственный университет аэрокосмического приборостроения

Защита состоится «____»____________ 2011 г. в ____ часов на заседании диссертационного совета Д.002.077.01 при Институте проблем передачи информации им. А. А. Харкевича РАН, расположенном по адресу: 127994, г. Москва, ГСП-4, Большой Каретный переулок, д. 19, стр. 1

С диссертацией можно ознакомиться в библиотеке ИППИ РАН.

Автореферат разослан «____»____________ 2011 г.

Ученый секретарь
диссертационного совета Д.002.077.01,
доктор физико-математических наук Цитович И. И.

Общая характеристика работы

Актуальность работы. Более чем сорокалетняя история кодов с ограничениями на длины серий или канальных кодов берёт своё начало с практической задачи модуляционного кодирования для канала магнитной записи. С первых публикаций и по настоящее время, эта тема стала объектом интереса большого числа специалистов в области теории информации и кодирования. Порождаемые свойствами канала ограничения на минимальную — d и максимальную — k длины серий нулей в потоке записываемых на магнитный носитель двоичных данных привели к разработке и изучению нового класса двоичных кодов — кодов с ограничениями на длины серий или, как их принято называть в зарубежной литературе, RLL кодов.

В 1965 г. американский специалист в области теории информации У. Каутц (W. Kautz) предложил нумерационную процедуру кодирования для канала с одним ограничением (d или k). В 1970 г. Д. Тонг (D. Tang) и Л. Бал (L. Bahl) распространили идею Каутца на кодирование канала с (d, k) - ограничениями. В 1971 г. В. Ф. Бабкин, затем в 1973 г. Т. Ковер (T. Cover) предложили универсальную схему нумерационного кодирования для источников. Метод Бабкина и Ковера оказался чрезвычайно удачным и, благодаря усилиям таких учёных как К. Имминк (K. A. S. Immink), В. Д. Колесник, Ч. Шалкенс (T. Tjalkens), Ф. Браун (V. Braun), позволил дополнительно ввести стыковочные ограничения (l, r) и стал активно применяться для блокового кодирования каналов с (d, k, l, r) - ограничениями.

Этот метод дал начало многочисленным теоретическим исследованиям в области кодирования последовательностей с ограничениями. Кроме того, он позволил распространить существующие достижения, полученные для кодирования источника на кодирование канала с (d, k) - или (d, k, l, r) - ограничениями. Нумерационная техника явно показала комбинаторную природу этого канала. Она позволила получить рекуррентные уравнения и производящие функции для перечисления последовательностей с ограничениями, что в свою очередь, дало возможность изучать асимптотику и получать оценки для канала.

В инженерной практике техника нумерационного кодирования последовательностей с ограничениями, несмотря на конкуренцию альтернативных способов, позволила Иммину разработать EFM код, который лёг в основу стандартов систем компакт-диск (CD) и цифровой диск (DVD). С этого момента канальное кодирование для оптической записи базируется на нумерационном методе.

Динамичное развитие теории канального кодирования и растущие практические потребности требуют приложения всё новых усилий специалистов в этой области. Так некоторые вопросы до сих пор не имеют единого общего решения или доказательства его существования. Например, известны лишь отдельные, частные решения задачи помехоустойчивого кодирования для канала с ограничениями. При расширении класса ограничений, зачастую, не получены точные оценки, имеются лишь результаты вычислительных экспериментов. Расширение области применения кодов с ограничениями, например, на задачи многоканальной связи, также требует дополнительных исследований.

Цель работы. Цель настоящей диссертации состоит в разработке и исследовании кодовой конструкции для последовательностей с ограничениями на длины серий, вес или заряд.

Задачи исследования. Для реализации цели диссертации нужно было решить следующие задачи:

- распространить метод нумерационного кодирования на лексикографически упорядоченные двоичные последовательности с широким классом ограничений;
- выделить и рассмотреть специальные классы ограничений, а именно:
  - ограничения на длины серий и вес;
  - ограничения на длины серий нулей и накопленный заряд;
- используя технику нумерационного кодирования, определить мощность множества двоичных последовательностей с указанными выше ограничениями;
- разработать алгоритмы нумерационного кодирования и декодирования исследуемых последовательностей.

Методы исследования. В качестве научного аппарата диссертационного исследования использовались методы дискретной математики, комбинаторного анализа, теории функций комплексного переменного, теории информации и кодирования, теории сложности алгоритмов. Экспериментальные исследования проводились с помощью численного моделирования в среде “Maple”, а также с использованием специальных программ, написанных автором на ANSI C.

Научная новизна. Научная новизна диссертации заключается в том, что в ней впервые:

- предложены равновесные двоичные последовательности, у которых ограничения на длины серий нулей и длины серий единиц независимы;
- получены весовые спектры кода с ограничениями на длины серий нулей и единиц, а также получены весовой и зарядовый спектры кода с ограничениями на длины серий нулей;

- найдены рекуррентные и явные формулы для определения числа двоичных последовательностей с ограничением на длины серий нулей и заряд;

- выведены производящие функции, перечисляющие множества последовательностей с ограничениями на длины серий нулей, вес, заряд и доказано, что производящая функция для числа последовательностей с ограничениями на длины серий нулей и заряд не имеет замкнутого выражения.

Теоретическая и практическая ценность работы. Работа носит теоретический характер. Результаты исследований возможно распространить на помехоустойчивое кодирование канала, на получение оценок ёмкости и скорости кода, а также алгоритмической сложности кодирования. Отдельный интерес представляет взаимная рекурсия, полученная для определения числа двоичных последовательностей с ограничением на длины серий нулей и заряд. В частности, важно доказательство факта, что данная рекурсия не может быть выражена в замкнутой форме.

Конструкции кодов, предложенные в данной работе, позволяют получить последовательности с заданными спектральными свойствами для применения в системах массовой памяти. Практическую ценность представляют также алгоритмы нумерационного кодирования данного класса кодов.

Научные положения, выносимые на защиту. На защиту выносятся следующие положения:

- кодирование двоичных последовательностей с (d, k, l, r) - ограничениями на длины серий нулей с линейной трудоёмкостью;

- расширение класса ограничений путём введения раздельных ограничений на длины серий нулей и единиц и добавления совместных ограничений на вес или заряд;

- определение мощности множеств $|\mathscr{S}_n|$ последовательностей длины n с ограничениями на длины серий, вес или заряд, получение производящих функций для последовательностей $|\mathscr{S}_n|$;

- нумерационные алгоритмы кодирования и декодирования последовательностей с ограничениями.

Апробация работы. Результаты диссертационной работы докладывались на российских и международных конференциях: 1) 7th International Workshop on Algebraic and Combinatorial Coding Theory (ACCT-7), Bansko, Bulgaria, June 18 – 24, 2000; 2) IEEE International Symposium on Information Theory, Lausanne, Switzerland, June 6 – July 5, 2002; 3) 10th International

Workshop on Algebraic and Combinatorial Coding Theory (ACCT-10), Zvenigorod, Russia, September 3 – 9, 2006. 4) 7th East-West Design & Test Symposium (EWDTS 2009), Moscow, Russia, September 18 – 21, 2009.

Основные результаты работы неоднократно обсуждались и были одобрены на семинаре по теории кодирования Института проблем передачи информации РАН.

Публикации. По теме диссертации опубликовано семь печатных работ — [1–7] общим объёмом 2,9 п.л. Из них три работы — [2, 4, 7], общим объёмом 2,5 п.л, в реферируемом журнале из перечня ВАК.

Структура и объем диссертации. Диссертация состоит из введения, трёх глав, заключения и списка литературы из 56 наименований. Общий объём работы составляет 114 страниц. Диссертация содержит 6 рисунков и 16 таблиц общим объёмом 8 страниц.

Содержание работы

Во Введении обоснована актуальность диссертационной работы, сформулирована цель и аргументирована научная новизна исследований, показана практическая значимость полученных результатов, представлены выносимые на защиту научные положения.

В первой главе вводится определение двоичных последовательностей с ограничениями на длины серий нулей: Двоичная последовательность $\boldsymbol{x} \in \{0,1\}^n$ длины n, в которой две соседние единицы разделены серией нулей, имеющей длину не менее d и не более k, называется последовательностью с (d,k) - ограничениями на длины серий нулей. Чтобы эти ограничения не нарушались на стыках блоков при блоковом кодировании, вводятся дополнительные ограничения: l – на максимальную длину первой серии нулей и r – на максимальную длину последней серии нулей. Такая последовательность называется последовательностью с (d,k,l,r) - ограничениями на длины серий нулей. Двоичные (d,k,l,r) - последовательности, начинающиеся с единицы, будем называть последовательностями с (d,k,r) - ограничениями.

Для изложения результатов далее понадобятся следующие обозначения и определения. Обозначим через $\mathscr{S}$ множество последовательностей с (d,k,r) - ограничениями, а через $\hat{\mathscr{S}}$ — множество последовательностей с (d,k,l,r) - ограничениями. Пусть на $\hat{\mathscr{S}}$ установлен лексикографический порядок. Тогда под нумерацией будем подразумевать определение номера данной последовательности в упорядоченном списке $\hat{\mathscr{S}}_n = \{\boldsymbol{x}_0 < \boldsymbol{x}_1 < \cdots < \boldsymbol{x}_{|\hat{\mathscr{S}}_n|-1}\}$. Денумерация — это отображение целого числа N, $0 \le N \le |\hat{\mathscr{S}}_n| - 1$, задающего номер в списке, в соответствующую последовательность из $\hat{\mathscr{S}}_n$. Нумерацию и денумерацию будем называть *нумерационным кодированием*.

Общий алгоритм нумерационного кодирования был описан в работах Бабкина[1] и Ковера [2]. Он опирается на подсчёт числа последовательностей, предшествующих данной последовательности $\boldsymbol{x} = (x_1, \ldots, x_n)$ в установленном лексикографическом порядке, посредством определения числа последовательностей $W(\boldsymbol{p})$, имеющих префикс $\boldsymbol{p} = (x_1, \ldots, x_{j-1}, 0)$. Временная сложность этого алгоритма зависит от сложности вычисления весовых коэффициентов $W(\boldsymbol{p})$, но не ниже $O(n)$. При этом метод Бабкина и Ковера подразумевает произвольную сложность вычисления $W(\boldsymbol{p})$ от $O(1)$ до $O(2^n)$.

Далее в первой главе показано, что если весовой коэффициент разряда не зависит от префикса $\boldsymbol{p}$, то он не зависит и от разрядности n, а зависит только от ограничений (d, k, l, r) и позиции разряда j, отсчитываемой от разряда с наименьшим весом, т. е. сложность вычисления весовых коэффициентов есть константа. При этом нумерация и денумерация сводится к поразрядному уравновешиванию. Порядок следования разрядов в последовательности при поразрядном уравновешивании изменён на обратный, принятый при записи в позиционной системе счисления, т. е. разряду с меньшим весом соответствует меньший номер. Вес очередного разряда при этом $w_n = |\hat{\mathscr{S}}_{n-1}|$, что позволяет применять для нахождения этих величин те же рекуррентные формулы, что и для метода Ковера.

Емкостная сложность алгоритма поразрядного уравновешивания составляет $O(n)$.

Возможность использования поразрядного уравновешивания с использованием фиксированного набора весов w_n и смещений (порогов) f_n определяется следующими теоремами:

Теорема 1. *Всякой последовательности $\boldsymbol{x} = (x_n, x_{n-1}, \ldots, x_1) \in \mathscr{S}_n$ можно поставить во взаимно-однозначное соответствие целое неотрицательное число $N(\boldsymbol{x}) \in [0, |\mathscr{S}_n| - 1]$, определяемое следующим образом:*

$$N(\boldsymbol{x}) = \sum_{i=1}^{n} w_i x_i - f_n, \tag{1}$$

для (d, k, r) - последовательностей и

Теорема 2. *Всякой последовательности $\boldsymbol{x} = (x_n, x_{n-1}, \ldots, x_1) \in \hat{\mathscr{S}}_n$ можно поставить во взаимно-однозначное соответствие целое неотрицательное число $N(\boldsymbol{x}) \in [0, |\hat{\mathscr{S}}_n| - 1]$, определяемое следующим образом:*

$$N(\boldsymbol{x}) = \sum_{i=1}^{n} w_i x_i - f_n(l),$$

[1] В. Ф. Бабкин. Метод универсального кодирования источника независимых сообщений неэкспоненциальной трудоемкости // Пробл. передачи информ. 1971. Т. 7, № 1. С. 13–21.

[2] T. M. Cover. Enumerative source coding // IEEE Trans. Inf. Theory. 1973. Jan. Vol. IT-19, no. 1. P. 73–77.

для (d, k, l, r) - последовательностей.

Доказательство теорем 1 и 2 основано на следующей лемме:

Лемма 1. *Для любого* $N \in [f_i, |\mathscr{S}_i| + f_i - 1]$: $N - f_{i+k} \geq 0$, $1 \leq i \leq n$, *выполняются условия:*

$$\begin{array}{l} N - w_i - a_{i-1} \geq 0, \;\; 2 \leq i < 1 + (k+1), \\ N - w_i - f_{i-1} \geq 0, \;\; 1 + (k+1) \leq i \leq n \end{array}$$

и

$$N - w_i - f_{i+k-(d-t)} < 0, \quad 1 + d - t \leq i \leq n, \quad 0 \leq t < d,$$

где смещение f_n определяется рекуррентно как

$$f_{n-1} + f_{n-(d+1)} - f_{n-1-(k+1)}, \quad 1 + 2(k+1) \leq n,$$

вес w_n определяется как

$$w_n = \begin{cases} f_{n+k} - a_{n+k}, & 1 \leq n < 1 + (k+1), \\ f_{n+k} - f_{n-1}, & 1 + (k+1) \leq n, \end{cases}$$

а величина a_n, введённая в [1] учитывает наличие ограничения r следующим образом:

$$a_n = \begin{cases} 0, & n \bmod (k+1) < r+1, \\ 1, & n \bmod (k+1) \geq r+1. \end{cases}$$

Из теоремы 1 и уравнения (1) следует, что алгоритмы нумерации и денумерации (d, k, l, r) - последовательностей, совпадают с алгоритмами Бала–Тонга[3] нумерации и денумерации (d, k) - последовательностей.

Во второй главе вводятся двоичные последовательности с ограничениями на длины серий нулей, единиц и вес. Такие последовательности были впервые рассмотрены в 2002 году в [3]. Здесь потребовалось определение новых ограничений:

1. (d_A, k_A) - ограничения: длина любой серии нулей между сериями единиц в $\boldsymbol{x}$ не менее d_A и не более k_A;

2. (d_B, k_B) - ограничения: длина любой серии единиц между сериями нулей в $\boldsymbol{x}$ не менее d_B и не более k_B;

3. левые (d_A, k_A) - ограничения: длина серии нулей между началом последовательности и первой единицей в $\boldsymbol{x}$ не менее d_{LA} и не более k_{LA};

4. левые (d_B, k_B) - ограничения: длина серии единиц между началом последовательности и первым нулём в $\boldsymbol{x}$ не менее d_{LB} и не более k_{LB};

---

[3] D. T. Tang, L. R. Bahl. Block Codes for a Class of Constrained Noiseless Channels // Inform. and Control. 1970. Vol. 17, no. 5. P. 436–461.

5. правые (d_A, k_A) - ограничения: длина серии нулей между последней единицей и концом последовательности $\boldsymbol{x}$ не менее d_{RA} и не более k_{RA};

6. правые (d_B, k_B) - ограничения: длина серии единиц между последним нулём и концом последовательности $\boldsymbol{x}$ не менее d_{RB} и не более k_{RB}.

Для множества этих ограничений введём обозначение $\mathscr{L} = \{d_A, k_A, d_{LA}, k_{LA}, d_{RA}, k_{RA}, d_B, k_B, d_{LB}, k_{LB}, d_{RB},\ k_{RB}\}$.

Кроме того, для указанных последовательностей вводится ограничение на вес $\nu = \sum_{j=1}^{n} x_j$, что позволяет расширить приложение таких кодов как для управления спектром, так и для обнаружения и исправления ошибок.

Обозначим через A_n^ν число таких последовательностей, которые начинаются с серии нулей, для которой $d_{LA} = d_A$, $k_{LA} = k_A$. Обозначим через B_n^ν число таких последовательностей, которые начинаются с серии единиц, для которой $d_{LB} = d_B$, $k_{LB} = k_B$. Аналогичным образом, обозначим через $\hat{A}_n^\nu$ число таких последовательностей, которые начинаются с серии нулей, определяемой ограничениями d_{LA} и k_{LA}. Наконец, обозначим через $\hat{B}_n^\nu$ число таких последовательностей, которые начинаются с серии единиц, определяемой ограничениями d_{LB} и k_{LB}.

Положим, что существует единственная последовательность нулевой длины и нулевого веса. Пусть это будет последовательность, начинающаяся с серии единиц. Тогда,

$$B_0^0 = 1, \quad \hat{B}_0^0 = 1, \qquad \nu = n = 0.$$

Число последовательностей A_n^ν и B_n^ν определяется рекуррентно в соответствии с утверждениями 1 и 2, а именно:

Если был хоть один переход от серии единиц к серии нулей, или наоборот, то справедливы:

Утверждение 1. *Количество последовательностей A_n^ν может быть определено следующим образом:*
Если $\nu = 0$, то

$$A_n^0 = \begin{cases} 1, & d_{RA} \le n \le k_{RA}, \\ 0, & \textit{в противном случае}. \end{cases}$$

Если $\nu \ne 0$, то

$$A_n^\nu = \begin{cases} 0, & n < d_A, \\ \sum_{j=d_A}^{\min(n,k_A)} B_{n-j}^\nu, & d_A \le n. \end{cases}$$

и

Утверждение 2. *Количество последовательностей* B_n^ν *может быть определено следующим образом:*
Если $\nu = n$, *то*

$$B_n^n = \begin{cases} 1, & d_{RB} \le n \le k_{RB}, \\ 0, & \textit{в противном случае.} \end{cases}$$

Если $\nu \ne n$, *то*

$$B_n^\nu = \begin{cases} 0, & n < d_B, \\ \sum_{j=d_B}^{\min(n,k_B)} A_{n-j}^{\nu-j}, & d_B \le n. \end{cases}$$

Если идёт первая серия нулей, то справедливо

Утверждение 3. *Количество последовательностей* $\hat{A}_n^\nu$ *может быть определено следующим образом:*
Если $\nu = 0$, *то*

$$\hat{A}_n^0 = \begin{cases} 1, & \max(d_{LA}, d_{RA}) \le n \le \min(k_{LA}, k_{RA}), \\ 0, & \textit{в противном случае.} \end{cases}$$

Если $\nu \ne 0$, *то*

$$\hat{A}_n^\nu = \begin{cases} 0, & n < d_{LA}, \\ \sum_{j=d_{LA}}^{\min(n,k_{LA})} B_{n-j}^{\nu}, & d_{LA} \le n. \end{cases}$$

Если идёт первая серия единиц, то справедливо

Утверждение 4. *Количество последовательностей* $\hat{B}_n^\nu$ *может быть определено следующим образом:*
Если $\nu = n$, *то*

$$\hat{B}_n^n = \begin{cases} 1, & \max(d_{LB}, d_{RB}) \le n \le \min(k_{LB}, k_{RB}), \\ 0, & \textit{в противном случае.} \end{cases}$$

Если $\nu \ne n$, *то*

$$\hat{B}_n^\nu = \begin{cases} 0, & n < d_{LB}, \\ \sum_{j=d_{LB}}^{\min(n,k_{LB})} A_{n-j}^{\nu-j}, & d_{LB} \le n. \end{cases}$$

При этом полагаем, что $d_A \ge 1$, $k_A \ge 0$, $d_{LA} \ge 1$, $k_{LA} \ge 0$, $d_{RA} \ge 1$, $k_{RA} \ge 0$, $d_B \ge 1$, $k_B \ge 0$, $d_{LB} \ge 1$, $k_{LB} \ge 0$, $d_{RB} \ge 1$ и $k_{RB} \ge 0$.

Весовые коэффициенты $W(\boldsymbol{p})$ разрядов теперь будут не фиксированными, как в алгоритмах Бала–Тонга, а зависимыми от префикса текущего разряда. В данном случае подразумевается либо длина последней серии нулей в префиксе, либо длина серии единиц, непосредственно предшествующих одиночному нулю в конце префикса. Из-за этой зависимости, утверждение, определяющее весовые коэффициенты, получается громоздким и здесь опущено. Но принципиальное значение имеет следствие этого утверждения.

Следствие 1. *При любых возможных ограничениях на длины серий и вес, весовые коэффициенты разрядов последовательности* $W_n^\nu(x_1, x_2, \ldots \ldots, x_{j-1}, 0)$ *зависят только от длины последовательности* n*, её веса* ν*, элементов множества ограничений* $\mathscr{L} \setminus \{k_{LB}\}$*, позиции разряда (длины префикса)* j *и длины серии, непосредственно предшествующей разряду* x_j*.*

В третьей главе рассматриваются двоичные последовательности с ограничениями на длины серий нулей, вес ν или заряд $\sigma = \sum_{j=1}^{n}(-1)^{\nu_j}$, где $\nu_j = \sum_{i=1}^{j} x_i$. Причём акцент сделан на получение явного вида для рекуррентных уравнений, определяющих число таких последовательностей, начинающихся с 1, с заданным весом — A_n^ν или с заданным зарядом — C_n^σ.

Вначале положим, что существует единственная последовательность нулевой длины и нулевого веса, начинающаяся с 1. Тогда $A_0^0 = 1$.

Далее, указанные рекуррентные соотношения определяются следующими утверждениями:

Утверждение 5. *Количество последовательностей* A_n^ν *может быть определено следующим образом:*
Если $\nu = 1$*, то*

$$A_n^1 = \begin{cases} 1, & 1 \le n \le r + 1, \\ 0, & \textit{в противном случае.} \end{cases}$$

Если $\nu > 1$*, то*

$$A_n^\nu = \begin{cases} 0, & n < d + 1, \\ \sum_{j=d+1}^{\min(n,k+1)} A_{n-j}^{\nu-1}, & d + 1 \le n. \end{cases} \tag{2}$$

Утверждение 6. *Количество последовательностей* C_n^σ *может быть определено следующим образом:*
Если $\sigma = -n$*, то*

$$C_n^{-n} = \begin{cases} 1, & n \le r + 1, \\ 0, & \textit{в противном случае.} \end{cases}$$

Если $\sigma \neq -n$*, то*

$$C_n^\sigma = \begin{cases} 0, & n < d + 1, \\ \sum_{j=d+1}^{\min(n,k+1)} C_{n-j}^{-\sigma-j}, & d + 1 \le n. \end{cases} \tag{3}$$

При этом полагаем, что $d \ge 0$, $k \ge d$, $r \ge 0$.

Явный вид (2) определяет следующая лемма 2, которая доказывается индукцией по ν.

Лемма 2. *Число A_n^ν равновесных последовательностей может быть явно выражено как*

$$A_n^\nu = \sum_{j=0}^{\nu-1} (-1)^j \binom{\nu-1}{j} \left(\binom{n-1-(\nu-1)d-jq}{\nu-1} - \binom{n-1-(r+1)-(\nu-1)d-jq}{\nu-1} \right), \quad \nu \geq 1, \tag{4}$$

где $q = k - d + 1$.

Аналогичная лемма даёт явную формулу для вычисления C_n^σ в виде суммы 12 слагаемых с соответствующими знаками:

$$C_n^\sigma = \sum_{I=1}^{12} (-1)^{\lfloor I/2 \rfloor} C(I)_n^\sigma.$$

Обозначим эти слагаемые через $C(I)_n^\sigma$ и определим их как

$$\begin{aligned} C(I)_n^\sigma = \sum_{m=0}^{\left\lceil \frac{n+\sigma}{2(d+1)} \right\rceil} &\left(\sum_{j=0}^{m-\mu} (-1)^j \binom{m-\mu}{j} \binom{\frac{n+\sigma}{2} - p_1 - md - jq - \mu}{m-\mu} \right) \\ &\times \left(\sum_{j=0}^{m} (-1)^j \binom{m}{j} \binom{\frac{n-\sigma}{2} - p_2 - md - jq}{m} \right), \end{aligned} \tag{5}$$

где

$$p_1 = \begin{cases} 0, & I = 1,2,7,8, \\ 1, & I = 3,4, \\ -d, & I = 5,6, \\ (k+1) - d, & I = 9,10, \\ (r+1) - d, & I = 11,12, \end{cases} \qquad p_2 = \begin{cases} 0, & I = 1,3,5,7,9,11, \\ 1 + (r+1), & I = 2,4, \\ 1, & I = 6,8,10,12, \end{cases}$$

и

$$\mu = \begin{cases} 0, & I = 1,2,3,4, \\ 1, & I = 5,6,7,8,9,10,11,12. \end{cases}$$

Явные выражения (4) и (5) стали основой для получения производящих функций, перечисляющих множества dkr - последовательностей с ограничениями на вес или заряд. Существует два типа последовательностей A_n^ν и C_n^σ. Бесконечные последовательности первого типа это $A_0^\nu, A_1^\nu, \ldots$ затем $C_0^\sigma, C_2^\sigma, \ldots$ или $C_1^\sigma, C_3^\sigma, \ldots$. Конечные последовательности второго типа это $A_n^1, A_n^2, \ldots, A_n^n$ или $C_n^{-n}, C_n^{-n+2}, \ldots, C_n^n$. Колесник и Крачковский рассматривали[4] рекурсивное вычисление производящей функции первого типа для пе-

[4] V. D. Kolesnik, V. Yu. Krachkovsky. Generating Functions and Lower Bounds on Rates for Limited Error-Correcting Codes // IEEE Trans. Inf. Theory. 1991. May. Vol. IT-37, no. 3. P. 778–788.

речисления последовательностей с постоянным весом. Ли (Lee)[5] предложил для этого явный метод. В диссертационной работе получены производящие функции обеих типов.

Рассмотрим бесконечную последовательность $A_0^\nu, A_1^\nu, A_2^\nu, \ldots$. Определим производящую функцию для этой последовательности как

$$A^\nu(t) = \sum_{n=0}^{\infty} A_n^\nu t^n, \quad t \in \mathbb{R}.$$

Подставляя сюда явное выражение (4) для A_n^ν и изменяя порядок суммирования, получаем производящую функцию в замкнутой форме.

$$A^\nu(t) = \left(\frac{t^{d+1}(1-t^q)}{1-t}\right)^{\nu-1} \frac{t(1-t^{r+1})}{1-t}. \tag{6}$$

Рассмотрим конечный степенной ряд

$$A_n(y) = \sum_{\nu=1}^{n} A_n^\nu y^\nu, \quad y \in \mathbb{R}.$$

Подставляя сюда выражение (6), получаем производящую функцию двух переменных t и y. Далее потребуется n-кратное дифференцирование по t с помощью интегральной формулы Коши для производных, затем применение теоремы о полной сумме вычетов даёт

$$A_n(y) = \sum_{j=1}^{k+2} \frac{\tau_j^{r+1} - 1}{\tau_j^n \prod\limits_{\substack{m=1,\\ m\neq j}}^{k+2} (\tau_j - \tau_m)},$$

где τ_1, τ_2, $\ldots$, τ_{k+2} корни $1 - \tau - y\tau^{d+1} + y\tau^{k+2}$, причём $\tau_1 = 1$.

Аналогичным образом рассмотрим бесконечную последовательность C_0^σ, $C_1^\sigma, C_2^\sigma, \ldots$. В данном случае, определим производящую функцию как

$$C^\sigma(t) = \sum_{\substack{n=0\\ n \text{ even}}}^{\infty} C_n^\sigma t^{n/2}, \quad \text{или} \quad C^\sigma(t) = \sum_{\substack{n=1\\ n \text{ odd}}}^{\infty} C_n^\sigma t^{(n-1)/2}, \quad t \in \mathbb{R}. \tag{7}$$

Попытка получить (7) в замкнутом виде предполагает подстановку (5) с последующим изменением порядка суммирования, что ранее позволило

[5] P. Lee. Combined Error-Correcting/Modulation Recording Codes: Dr. scient. thesis // University of California. San Diego, 1988. Mar.

получить (6) и с переходом к суммированию по диагонали. Основная трудность здесь заключается в нахождении замкнутого выражения для внутренней суммы, той, что получается при перемножении рядов в (5). Эта проблема решается с помощью следующей важной леммы, которая затем потребуется ещё, по меньшей мере, дважды. Здесь будет удобно положить

$$t = \frac{x-1}{x+1}, \quad x \in \mathbb{R}. \tag{8}$$

Лемма 3. *Производящая функция*

$$g(t) = \sum_{j=0}^{\infty} \binom{\kappa+j}{\kappa} \binom{\lambda+j}{\nu} t^j, \quad \lambda \geq \nu$$

может быть выражена в интегральной форме:

$$g(t) = t^{\nu-\lambda} \frac{(x+1)^{\nu+1}}{2^{\kappa+\nu+1}} \frac{1}{2\pi i} \int\limits_{\gamma} \frac{\left(\frac{\xi-1}{\xi+1}\right)^{\lambda} (\xi+1)^{\kappa+\nu}}{(\xi-x)^{\nu+1}} d\xi,$$

где γ — некоторая замкнутая жорданова кривая, охватывающая в положительном направлении точку $x = \frac{1+t}{1-t}$, лежащую на вещественной оси.

После проведения, в процессе доказательства леммы 3, гипергеометрических преобразований удобно ввести обозначения:

$$b_d = \frac{t^{d+1}}{2(1-t)} \quad \text{и} \quad b_k = \frac{t^{k+2}}{2(1-t)}.$$

Поскольку мы ввели (8), то теперь мы установим аналогичное отображение на всей комплексной плоскости, а именно, положим $\tau = \frac{\xi-1}{\xi+1}$.

Попытка получить замкнутое выражение для (7) дала следующий результат:

$$C^{\sigma}(t) = \frac{t^{\lfloor -\sigma/2 \rfloor + p_1}}{(1-t)^{1-\mu}} J_{1,4} + \frac{2^{\mu} t^{\lfloor \sigma/2 \rfloor + p_2}}{1-t} J_{2,3} - \frac{\mu 2^{\mu-1} t^{q \max(\tilde{n},0) + \lfloor -\sigma/2 \rfloor + p_1}}{1-t^q},$$

где $J_{1,4}$ и $J_{2,3}$ обозначают следующие интегралы:

$$\begin{aligned} J_{1,4} = \frac{(-1)^{\tilde{n}}}{2\pi} \int\limits_{\xi_3}^{\xi_1} \frac{\tau^{\sigma-p_1+p_2}}{Q} \left(b_d(\xi+1)\right)^{\mu} \\ \times \left(\frac{\sqrt{t}}{\tau}\right)^{\tilde{n}q} \left(\frac{H_1}{|\widetilde{R}|} \cos(\tilde{n}\varphi) - (1-\mu\tau^q) \sin(\tilde{n}\varphi)\right) d\xi, \end{aligned} \tag{9}$$

$$
\begin{aligned}
J_{2,3} = &-\frac{(-1)^{\tilde n}}{2\pi}\int\limits_{\xi_3}^{\xi_1}\frac{\tau^{-(\sigma-p_1+p_2)}}{Q}\left(\frac{b_d(\xi-1)}{\xi-x}\right)^{\mu} \\
&\times\left(\frac{\sqrt{t}}{\tau}\right)^{-\tilde n q}\left(\frac{H_2}{|\widetilde R|}\cos\left(\tilde n\varphi\right)-\left(1-\mu\left(t/\tau\right)^{q}\right)\sin\left(\tilde n\varphi\right)\right)d\xi,
\end{aligned}
\tag{10}
$$

где, в свою очередь,

$$
\varphi = -\arccos\left(\frac{P_0}{2\sqrt{b_d b_k}(\xi^2-1)}\right)
$$

это фаза, а

$$
\tilde n = \left\lceil\frac{\sigma-p_1+p_2}{q}\right\rceil
$$

есть число членов в конечных рядах, которые образуются при переходе к упомянутому выше диагональному суммированию. Далее H_1 и H_2 определяются как

$$
\begin{aligned}
H_1 &= \mu P_1 + \tau^{\mu q}P_3, \\
H_2 &= \mu\left(t/\tau\right)^{q}P_2 + P_4,
\end{aligned}
$$

$P_0,\ldots,P_4$ это многочлены степени 2 от комплексной переменной ξ:

$$
\begin{aligned}
P_0 &= \xi - x - \left(b_d+b_k\right)\left(\xi^2-1\right), \\
P_1 &= \xi - x - \left(b_d-b_k\right)\left(\xi^2-1\right), \\
P_2 &= \xi - x - \left(b_k-b_d\right)\left(\xi^2-1\right), \\
P_3 &= \xi - x - \left(b_d\left(1-2\tau^{(1-\mu)q}\right)+b_k\right)\left(\xi^2-1\right), \\
P_4 &= \xi - x - \left(b_d+b_k\left(1-2\tau^{(\mu-1)q}\right)\right)\left(\xi^2-1\right),
\end{aligned}
$$

а знаменатели Q и $\widetilde R$ это

$$
\begin{aligned}
Q &= \xi - x - \left(b_d\left(1-\tau^{q}\right)+b_k\left(1-\tau^{-q}\right)\right)\left(\xi^2-1\right), \\
\widetilde R &= \pm\sqrt{P_0^2-4b_d b_k\left(\xi^2-1\right)^2}.
\end{aligned}
\tag{11}
$$

Пределы интегрирования есть корни выражения под радикалом в (11).

$$
\begin{aligned}
\xi_{1,2} &= \frac{1\mp\sqrt{1-4\left(\sqrt{b_d}+\sqrt{b_k}\right)^2x+4\left(\sqrt{b_d}+\sqrt{b_k}\right)^4}}{2\left(\sqrt{b_d}+\sqrt{b_k}\right)^2}, \\
\xi_{3,4} &= \frac{1\mp\sqrt{1-4\left(\sqrt{b_d}-\sqrt{b_k}\right)^2x+4\left(\sqrt{b_d}-\sqrt{b_k}\right)^4}}{2\left(\sqrt{b_d}-\sqrt{b_k}\right)^2}.
\end{aligned}
$$

Теперь можно сформулировать наиболее важный результат диссертационной работы:

Теорема 3. *Рекуррентное соотношение вида* (3) *не имеет замкнутой формы.*

Доказательство следует из того факта, что интегралы (9) и (10) суть эллиптические, что в свою очередь следует из анализа (11) в заданных пределах интегрирования.

Производящая функция последовательности $C_n^{-n}, C_n^{-n+2}, \ldots, C_n^n$ имеет замкнутую форму.

Рассмотрим конечный ряд

$$C_n(y) = \sum_{\substack{\sigma=-n \\ \sigma \text{ even}}}^{n} C_n^\sigma y^{\sigma/2} \quad \text{или} \quad C_n(y) = \sum_{\substack{\sigma=-n \\ \sigma \text{ odd}}}^{n} C_n^\sigma y^{(\sigma-1)/2},$$

где $y \in \mathbb{R}$.

Здесь также вместо C_n^σ рассматривается (5). По аналогии с (8) следует положить $y = \frac{x-1}{x+1}$. Результат есть выражение

$$\begin{aligned} C_n(y) &= (-1)^{n-p_1-p_2} y^{-\frac{n+(\sigma \bmod 2)}{2}} \\ &\times (x-1)^{n-p_2+1} (x+1)^{1-p_1-\mu} \left((1-\mu)\, J_1 + J_2\right), \end{aligned}$$

где x переопределена как

$$x = \frac{1+y}{1-y}.$$

J_1 и J_2 обозначают здесь интегралы для которых получены замкнутые формы:

$$J_1 = \frac{1}{2(-1-x)^{n-p_1-p_2+1}} - \frac{1}{2(1-x)^{n-p_1-p_2+1}}$$

и

$$J_2 = \left(x^2-1\right)^{k+1} (x+1)^{\mu q} \sum_{j=1}^{2(k+2)} \frac{h}{(\xi_j - x)^{n-p_1-p_2+1} \prod\limits_{\substack{m=1, \\ m \neq j}}^{2(k+2)} (\xi_j - \xi_m)}$$

$$\times \left(\begin{cases} \dfrac{\xi_j + 1}{(x+1)^q - (x-\xi_j)^q}, & \xi_j \neq -1, \\ \dfrac{1}{q(x-\xi_j)^{q-1}}, & \text{иначе}, \end{cases} \right)^{\mu},$$

где

$$h = \begin{cases} \dfrac{q\,(x-\xi_j)^{d+1}\left((x-\xi_j)^q - (x+\xi_j)^q\right)}{2\xi_j\,(x+\xi_j)^{k+1}}, & |\xi_j| = 1, \\ 1, & \text{иначе,} \end{cases}$$

а ξ_1, ξ_2, ..., $\xi_{2(k+2)}$ есть корни

$$\psi(\xi) = \left(x^2-1\right)^{k+1}\left(\xi^2-1\right) - (x-\xi)^{2(d+1)}\left(x^2-1\right)^q \\ + (x-\xi)^{(d+1)+(k+2)}\left((x-1)^q + (x+1)^q\right) - (x-\xi)^{2(k+2)}.$$

В Заключении обобщены полученные в диссертационной работе результаты и сделаны выводы.

Основные результаты работы

1. Известный для кодирования (d,k) - последовательностей метод Бала–Тонга с двумя наборами весовых коэффициентов, имеющий линейную вычислительную сложность, распространён на кодирование (d,k,l,r) - последовательностей при сохранении линейной трудоёмкости.

2. Введено понятие двоичных последовательностей постоянного веса с раздельными независимыми ограничениями на длины серий нулей и длины серий единиц. С помощью метода нумерации, определено число таких последовательностей, имеющих заданную длину, затем для них построены алгоритмы кодирования и декодирования.

3. Для (d,k,r) - и (d,k,l,r) - последовательностей расширен класс ограничений, а именно введены дополнительные ограничения:

   – на вес;

   – на накопленный заряд.

   С помощью техники нумерационного кодирования, были получены рекуррентные формулы для определения мощности множеств последовательностей с указанными выше ограничениями. Получены весовой и зарядовый спектры (d,k,r) - и (d,k,l,r) - кодов.

4. С учётом комбинаторной природы указанных рекуррентных соотношений, были получены явные формулы для определения мощности множеств таких последовательностей.

5. С использованием явных формул получены производящие функции, перечисляющие исследуемые множества последовательностей с ограничениями.

6. Доказано, что производящая функция, перечисляющая множества последовательностей с ограничениями на длины серий нулей и заряд не имеет замкнутого выражения.

7. Разработаны алгоритмы нумерационного кодирования и декодирования исследуемых последовательностей.

Список публикаций

1. *O. Kurmaev*. Weights Calculation for Enumerative Encoding of *dkrl* - Sequences // Proc. ACCT'2000. Bansko, Bulgaria, Jun. 18–24. 2000. P. 199–206.

2. *О. Ф. Курмаев*. Кодирование последовательностей с ограниченными длинами серий // *Пробл. передачи информ.* 2001. Т. 37, № 3. С. 34–43.

3. *O. Kurmaev*. An enumerative method for encoding and decoding constant-weight run-length limited binary sequences // IEEE Intl. Symposium on Information Theory. Lausanne, Switzerland. Jun. 30 – Jul. 5. 2002. P. 328.

4. *О. Ф. Курмаев*. Нумерационное кодирование последовательностей с ограничениями на длины серий нулей и вес // *Пробл. передачи информ.* 2002. Т. 38, № 4. С. 3–9.

5. *O. Kurmaev*. Weight and charge distribution of binary run-length limited codes // Proceedings of Tenth International Workshop on Algebraic and Combinatorial Coding Theory. Zvenigorod, Russia. 2006. Sep. 3–9. 2006. P. 174–178.

6. *O. Kurmaev*. An Algorithm for Testing Run-Length Constrained Channel Sequences // Proc. EWDTS'2009. Moscow. - Russia. Sep. 18–21. 2009. P. 391–392.

7. *О. Ф. Курмаев*. Нумерация двоичных последовательностей с ограничениями на длины серий и вес // *Пробл. передачи информ.* 2011. Т. 47, № 1. С. 78–95.