ՀԱՅԱՍՏԱՆԻ ԳԻՏՈՒԹՅՈՒՆՆԵՐԻ ԱԶԳԱՅԻՆ ԱԿԱԴԵՄԻԱ
НАЦИОНАЛЬНАЯ АКАДЕМИЯ НАУК АРМЕНИИ
NATIONAL ACADEMY OF SCIENCES OF ARMENIA
ДОКЛАДЫ ԶԵԿՈՒՅՑՆԵՐ REPORTS

Հատոր Том Volume 120 2020 № 1

МАТЕМАТИКА

УДК 621.391.15

В. К. Леонтьев[1], Г. Л. Мовсисян[2], Ж. Г. Маргарян[3]

Верхняя и нижняя границы мощности кода, исправляющего ошибки алгебраического канала

(Представлено академиком Г. Г. Хачатряном 25/IX 2019)

Ключевые слова: *оптимальный код, алгебраический канал связи, нижняя и верхняя границы, мощности кода.*

Введение. Основная задача теории корректирующих кодов для фиксированного канала формулируется следующим образом: построить код максимальной мощности, исправляющий ошибки этого канала.

Как показывает опыт, полное решение такой задачи является недоступным уже для классического канала с множеством слов $\{0,1\}^n$, искажениями вида 0→1, 1→0 и с ограниченным числом ошибок t. Поэтому были разработаны различные методы получения верхних и нижних границ для мощности оптимального кода, а их уточнение, наряду с эффективными конструкциями самих кодов, включая процедуры декодирования является предметом современной теории кодов, исправляющих ошибки.

§1. Каналы связи и словарные функции. Пусть $B = \{a_1, a_2, \dots, a_p\}$ – конечный алфавит и B^* – множество всех слов конечной длины над алфавитом B. Словарная функция – произвольное частичное отображение ψ следующего вида:

$$B^* \overset{\psi}{\to} B^*.$$

Таким образом, словарная функция отображает слово в слово. Примеров таких функций великое множество. Это и представление вещественных чисел в позиционных системах счисления, это и так называемые функции k-значной логики, это и машины Тьюринга (ТМ), продукции Поста и т.п. [1].

Понятие канала связи не имеет общего математического определения, хоть во многих ситуациях допускается адекватная формулировка в терми-

нах, близких к теории вероятности. Так получаются определения двоично-симметричного канала, гауссовского канала и т.д. Мы будем рассматривать канал как преобразователь информации, и если принять тезис о том, что в любом канале связи происходит преобразование одних слов в другие, то достаточно общий канал можно описать следующим образом.

Задано некоторое множество

$$\Psi = \{\psi_0, \psi_1, \dots, \psi_m\}$$

частичных словарных функций

$$B^* \overset{\psi_i}{\to} B^*, \qquad i = \overline{0, m}$$

и следующее многозначное отображение:

$$f(x) = (\psi_0(x), \psi_1(x), \dots, \psi_m(x)),$$

где $x \in W \subseteq B^*$. Содержательно это означает, что если $x \in W$, то после передачи по каналу $K(\Psi)$ это слово переходит в одно из слов $\{\psi_0(x), \psi_1(x), \dots, \psi_m(x)\} \subseteq W$.

Множество всех обратимых отображений $\{\psi_i\}$, $\psi_i(W) \subseteq W$ обозначим через T. При этом все суперпозиции $\psi_{i_1}\psi_{i_2}\dots\psi_{i_k}$ функции ψ_{i_j} из множества Ψ определены на W.

Определение [2]. *Алгебраическим каналом связи* $K(\Psi)$ *называется многозначное отображение*

$$f(x)=(\psi_0(x),\psi_1(x), \dots, \psi_m(x)), \tag{1}$$

где $\Psi \subseteq$ T, *и если*

$$\psi_i \in \Psi \to {\psi_i}^{-1} \in \Psi. \tag{2}$$

Формулу (1) следует понимать следующим образом. На вход канала подается слово v. На выходе получается ровно одно из значений $\psi_0(v), \psi_1(v), \dots, \psi_m(v)$.

Условие (2) требует, чтобы любое «преобразованное» слово могло быть возвращено к исходному виду путем тех же самых трансформаций.

Отметим, что любой аддитивный канал удовлетворяет условию (2) и является алгебраическим [3]. Однако не все матричные каналы являются алгебраическими, например матричный канал с выпадением символов [4, 5].

В дальнейшем мы всегда будем считать, что $\psi_0(x)= x$, что можно интерпретировать как возможность безошибочной передачи слова по этому каналу, $W = B^n \subseteq B^*$.

Следовательно, на множестве B^n действует группа преобразования T, которая переводит слово из B^n в слово из этого же множества. Подобные допущения упрощают многие технические детали и не влияют, как мы увидим далее, на ситуацию в целом. При этом изложенный материал становится более доступным для понимания.

После передачи некоторого слова $\boldsymbol{v}$ по имеющемуся каналу связи на выходе мы получаем слово $\boldsymbol{u}$. В классической постановке требуется вос-

становить исходное слово $\boldsymbol{v}$ по его искаженному образу $\boldsymbol{u}$ с максимально возможной достоверностью. Следуя стандартным традициям, любое подмножество V из B^n мы будем называть кодом V, который и используется для связи, т.е. передаваться по каналу могут только слова кода V. Таким образом, каждое слово $\boldsymbol{u}$, полученное на выходе канала, является образом кодового слова. И мы хотим восстановить исходное слово $v \in V$ по его образу. Здесь основное искусство состоит в правильном выборе кода V, позволяющего по любому искаженному сигналу однозначно восстановить исходное сообщение.

Определение. *Множество $V \subseteq B^n$ называется кодом, исправляющим ошибки канала $K(\Psi)$, если выполнено условие*

$$\psi_i(u) \neq \psi_j(v) \tag{3}$$

для всех $\boldsymbol{i}$ и $\boldsymbol{j}$ и для всех слов $\boldsymbol{u}, \boldsymbol{v} \in V$.

Условие (3) означает, что последствия действий канала $K(\Psi)$ на кодовые слова различны, и поэтому искажения могут быть обнаружены и исправлены.

В дальнейшем обозначим через $V(\Psi)$ код, исправляющий ошибки канала $K(\Psi)$. В терминах, введенных выше, основная задача при заданном канале состоит в построении кода $V(\Psi)$ максимальной мощности – $\overline{V}(\Psi)$.

Ясно, что мощность кода $V(\Psi)$ зависит от «структуры» и мощности множества Ψ, «порождающего» канал $K(\Psi)$, и т.д. В других терминах, можно условие (3) естественным образом увязать с понятием «**окрестности**» и сформулировать эквивалентные понятия, используемые в дальнейшем.

Окрестность k-го порядка слова v определим индуктивно:

$\Psi^k(v) = \Psi^1(x), \quad x \in \Psi^{k-1}(v), \quad$ где $\Psi^0(v) = v, \quad \Psi^1(v) = \Psi(v).$

При этом выполняются включения

$$\Psi^0(v) \subseteq \Psi^1(v) \subseteq \cdots \subseteq \Psi^k(v).$$

В терминах окрестностей условие, что код $V = \{v_0, v_1, \dots, v_N\}$ исправляет ошибки канала $K(\Psi)$, можно сформулировать следующим образом:

$$\Psi^1(v_i) \cap \Psi^1(v_j) = \emptyset, \ i \neq j. \tag{4}$$

Нетрудно проверить, что условия (3) и (4) являются эквивалентными.

Определение кода V, исправляющего ошибки канала $K(\Psi)$, в значительной мере копирует классическое определение кода, исправляющего искажения типа **0→1**, **1→0** в двоичном симметричном канале [6].

§2. Граница плотной упаковки. Сама процедура исправления ошибок на выходе канала $K(\Psi)$ в общем случае может быть представлена в виде следующей схемы, которую обычно называют **«таблицей декодирования»** $D(\Psi)$ или решающим правилом на приемном конце канала $K(\Psi)$. При выполнении условия (3) правило декодирования для $V(\Psi) =$

$\{v_0, v_1, \ldots, v_N\} \subseteq B^n$ может быть осуществлено с помощью стандартной таблицы – «таблицу декодирования»:

v_0	$v_1 \ldots$	$\ldots$	v_N
$\psi_1(v_0)$	$\psi_1(v_1)$	$\ldots$	$\psi_1(v_N)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\psi_m(v_0)$	$\psi_m(v_1)$	$\ldots$	$\psi_m(v_N)$

Первая строка этой таблицы – все кодовые слова: $v_0, v_1, \ldots, v_N$. Далее, первый столбец – все слова $\psi_0(v_0), \psi_2(v_0), \ldots, \psi_m(v_0)$. Второй столбец – $\psi_0(v_1), \psi_1(v_1), \ldots, \psi_m(v_1)$ и т.д. В силу условия (3) каждый элемент $\psi_i(v_s)$ может попасть лишь в один из столбцов таблицы декодирования.

Пусть, как и выше, $\Psi(x) = \{\psi_0(x),\ \psi_1(x), \ldots,\ \psi_m(x)\}$ – некоторое множество словарных функций, осуществляющее отображение

$$B^n \overset{\psi_i}{\to} B^n,\quad i = \overline{0, m}.$$

Утверждение 1. *Код $V \subseteq B^n$ исправляет ошибки канала $K(\Psi)$, если и только если столбцы таблицы декодирования $D(\Psi)$ не пересекаются.*

Самая общая верхняя граница мощности кода $V(\Psi)$, исправляющего ошибки канала $K(\Psi)$, состоит в следующем очевидном утверждении.

Утверждение 2. *Общее число различных элементов в таблице декодирования $D(\Psi)$ не превосходит $|B^n| = (p+1)^n$.*

Ясно, что окрестности одного порядка могут иметь разную мощность, и тогда верхняя граница для мощности кода $V(\Psi)$ имеет форму

$$|V(\Psi)| \le \frac{(p+1)^n}{\Psi^1}, \tag{5}$$

где

$$\Psi^1 = \min_{v \in B^n} |\Psi^1(v)|.$$

Действительно,

$$\bigcup_{v \in V(\Psi)} \Psi^1(v) \subseteq B^n.$$

Следовательно,

$$\sum_{v \in V(\Psi)} |\Psi^1(v)| \le (p+1)^n,$$

и поэтому

$$|V(\Psi)| \cdot \Psi^1 \leq (p+1)^n \ ,$$

что и доказывает (5). Это неравенство обычно называют границей плотной упаковки [6].

§3. Граница Варшамова – Гилберта. Утверждение 3. *Код* $V = \{v_0, v_1, \dots, v_N\}$ *исправляет ошибки канала* $K(\Psi)$ *тогда и только тогда, когда для любого* $v_i \in V$ *имеет место* $\Psi^2(v_i) \cap V = v_i$.

Доказательство. Действительно, если для канала $K(\Psi)$ и кода $V(\Psi)$ имеет место $\Psi^2(v_i) \cap V(\Psi) = \{v_i, v_j\}$, $v_i \neq v_j$, для некоторых v_i, v_j, то тогда $v_j \in \Psi^2(v_i)$. Следовательно, существуют $\psi_r, \psi_s \in \Psi$, для которых $\psi_r(v_j) = \psi_s(v_i)$. что противоречит условию (3).

Если для канала $K(\Psi)$ и кода V имеет место: $\Psi^2(v_i) \cap V = v_i$ для любого $v_i \in V$, но при этом $V \notin \{V(\Psi)\}$, т.е. существуют $\psi_r, \psi_s \in \Psi$ и $v_i, v_j \in V, v_i \neq v_j$, такие, что $\psi_r(v_i) = \psi_s(v_j)$, то тогда $\psi_s^{-1}\psi_r(v_i) = v_j$. Отсюда следует, что $v_j \in \Psi^2(v_i) \cap V$, что противоречит начальному предположению. Утверждение доказано.

Как следствие из этого утверждения используем конструкцию, которая носит название процедуры Варшамова – Гилберта [6], и получим нижнюю границу мощности кода V, исправляющего ошибки канала $K(\Psi)$.

Пусть канал $K(\Psi)$ является алгебраическим, то есть $\psi_i^{-1} \in \{\psi_0, \psi_1, \dots, \psi_m\}$ $i = \overline{0, m}$. Рассмотрим следующий алгоритм построения кода $V(\Psi)$:

1) в качестве кодового слова v_0 выберем произвольное слово из B^n;
2) рассмотрим окрестность 2-го порядка слова v_0: $\Psi^2(v_0)$;
3) в качестве следующего кодового слова v_1 выберем любое слово из $B^n \setminus \Psi^2(v_0)$;
4) продолжаем эту процедуру включая на каждом шаге в код V то слово, которое не принадлежит окрестностям второго порядка уже включенных в код V слов;
5) алгоритм заканчивает свою работу, когда все множество B^n будет покрыто окрестностями второго порядка кодовых слов.

Утверждение 4. *Код* V *исправляет ошибки канала* $K(\Psi)$.

Доказательство. Пусть не так и

$$\psi_i(v_r) = \psi_j(v_s). \qquad (6)$$

Согласно алгебраичности $K(\Psi)$ из (6) получаем

$$v_s = \psi_j^{-1}\, \psi_i(v_r) \ ,$$

и таким образом $v_s \in \Psi^2(v_r)$, что противоречит построению кода V.

Пусть

$$\Psi^2 = \max_{v \in B^n} |\Psi^2(v)|.$$

Тогда

$$\left|\overline{V}(\Psi)\right| \geq \frac{(p+1)^n}{\Psi^2}. \quad (7)$$

Действительно, в силу свойства 5) описанного выше алгоритма

$$\bigcup_{v \in V(\Psi)} \Psi^2(v) = B^n.$$

Отсюда получаем

$$\left|\overline{V}(\Psi)\right| \cdot \Psi^2 \geq |B^n|,$$

что и доказывает (7).

Таким образом, в случае алгебраического канала $K(\Psi)$ мы имеем следующие общие границы для объема кода $\overline{V}(\Psi)$, исправляющего ошибки канала $K(\Psi)$.

Утверждение 5. *Справедливы неравенства*

$$\frac{(p+1)^n}{\Psi^2} \leq \left|\overline{V}(\Psi)\right| \leq \frac{(p+1)^n}{\Psi^1}. \quad (8)$$

Верхняя граница в (8) – граница плотной упаковки, а нижняя граница – граница типа Варшамова – Гилберта. Неравенства (8) показывают, что достаточно универсальные границы для мощности кода могут быть выражены в терминах объема окрестностей 1-го и 2-го порядка.

Границы могут сильно различаться, поэтому вопрос о точности оценок (8) неоднократно формулировался в форме «какая из границ ближе к истинной – Варшамова – Гилберта или плотной упаковки». Не зная ответа на этот вопрос, мы приведем несколько примеров, которые показывают все разнообразие ситуации при построении максимальных кодов, исправляющих ошибки алгебраического канала, и могут быть интерпретированы как информация к размышлению.

Примеры. Пусть $= \{y_0,\ y_2, \ldots, y_m\}$ – некоторое подмножество B^n, «+» – сложение по mod (p+1).

1) Если $K(\Psi)$ – групповой канал (который будет рассмотрен нами в дальнейшем), то $\Psi^1(x) = \{\psi_0(x),\ \psi_1(x), \ldots,\ \psi_m(x)\}$ и $\Psi^r(x) = \Psi^1(x)$ для $r = 1,2, \ldots$

Таким образом, в групповом случае окрестности всех порядков слова $x \in B^n$ совпадают, и в обеих частях неравенства (8) достигается равенство, если при этом окрестности первого порядка всех слов из B^n имеют одинаковую мощность. Например, если в данном примере определить $\psi_i(x) = x + y_i$, то получим, что $\Psi^1 = \Psi^2$ и, следовательно, верхняя и нижняя границы в (8) совпадают.

2) Если $K(\Psi)$ – произвольный канал, где $\psi_i(x) = x + y_i$, то справедливы оценки вида (8), где основными параметрами являются мощности окрестностей 1-го и 2-го порядка. Например, если $\left|y_i\right| \leq t$, то из (8) имеем

$$\frac{(p+1)^n}{\sum_{i=0}^{2t} p^i C_n^i} \le \left|\overline{V}(\Psi)\right| \le \frac{(p+1)^n}{\sum_{i=0}^{t} p^i C_n^i}.$$

Отметим лишь следующий очевидный факт: для получения каких-то нетривиальных границ для мощности кода $V(\Psi)$ требуется наложить определенные ограничения на функции $\psi_i(x)$.

[1]ФИЦ ИУ РАН, Москва
[2]Группа Бит, Москва
[3]Ереванский государственный университет
e-mails: vkleontiev@yandex.ru, garib@hkzap.ru, j.margaryan@ysu.am

В. К. Леонтьев, Г. Л. Мовсисян, Ж. Г. Маргарян

Верхняя и нижняя границы мощности кода, исправляющего ошибки алгебраического канала

Рассматривается задача построения оптимального кода для алгебраических каналов связи. Получены универсальные границы (нижняя и верхняя) мощности оптимального кода, где основными параметрами являются мощности окрестностей первого и второго порядка.

Վ. Կ. Լեոնտին, Գ. Լ. Մովսիսյան, Ժ. Գ.Մարգարյան

Հանրահաշվական կապի գծերում սխալներ ուղղող կոդի հզորության վերին և ստորին գնահատականներ

Դիտարկվել է օպտիմալ կոդի կառուցման խնդիրը հանրահաշվական կապի գծերի համար: Ստացված են օպտիմալ կոդի հզորության ընդհանուր ստորին և վերին գնահատականներ, որտեղ հիմնական պարամետրերն առաջին և երկրորդ կարգի միջակայքերի հզորություններ են:

V. K. Leontiev, G. L. Movsissian, Zh. G. Margaryan

Upper and Lower Bounds of the Power of the Error Correction Code of an Algebraic Channel

The problem of constructing the optimal code for algebraic communication channels has been considered. The universal boundaries (lower and upper) of the power of the optimal code have been obtained, where the main parameters are the cardinalities of the vicinities of the first and second order.

Литература

1. *Мальцев А. И.* Алгоритмы и рекурсивные функции. М. Наука, Физматлит. 1986. 368 с.
2. *Леонтьев В. К., Мовсисян Г. Л.* Алгебраические каналы связи. The First International Algebra and Geometry Conference 16-20 may 2007. Yerevan, Armenia.
3. *Леонтьев В. К., Мовсисян Г. Л.* – ДНАН Армении. 2004. Т. 104. № 1. С. 23-27.
4. *Леонтьев В. К., Мовсисян Г. Л., Осипян А. А*. В кн.: Матричные каналы связи. Матер. XI междунар. семинара «Дискретная математика и ее приложение». М. Изд. МГУ. 2012. С. 415-416.
5. *Левенштейн В. И*. – ДАН СССР. 1965. Т. 163. № 4. С. 845-848.
6. *Мак-Вильямс Ф. Дж., Слоэн Н. Дж. А*. Теории кодов, исправляющих ошибки. М. Связь. 1979. 744 с.