ՀԱՅԿԱԿԱՆ ՍՍՀ ԳԻՏՈՒԹՅՈՒՆՆԵՐԻ ԱԿԱԴԵՄԻԱՅԻ ԶԵԿՈՒՅՑՆԵՐ
ДОКЛАДЫ АКАДЕМИИ НАУК АРМЯНСКОЙ ССР
LXXIV 1982 4

УДК 519.21

МАТЕМАТИКА

А. Л. Нерсисян

Полиномиальный алгоритм кодирования дискретной информации с заданным уровнем искажения

(Представлено академиком АН Армянской ССР С. А. Амбарцумяном 19/XI 1981)

1°. На практике часто встречаются ситуации, когда использователь применяет информацию не в том виде, в каком она к нему поступает, а в некотором преобразованном. Это имеет место, например, когда информация порождается не в том алфавите, в котором она используется, когда информация допускает погрешности в заданных пределах или когда она является недоопределенной и при использовании должна быть доопределена и т. д.

Во всех этих случаях вместо первоначальной информации используется некоторая другая, которую будем называть „аппроксимирующей". Степень приближения аппроксимирующей информации к исходной оценивается некоторой совокупностью критериев.

В вероятностной постановке задача о кодировании информации с заданной степенью точности была впервые сформулирована К. Шенноном, и ему принадлежат основополагающие результаты в этой области ([1]). Этой проблематике посвящена монография ([2]) (см. также обзор ([3])).

В данной работе рассмотрена детерминированная постановка задачи о представлении информации с заданной точностью при дополнительном требовании эффективности процедур кодирования и декодирования. Для ее решения были использованы детерминированные методы сжатия информации, первоначально развитые О. Б. Лупановым ([4]), Э. И. Нечипоруком ([5]) и Л. А. Шоломовым ([6]) для синтеза логических схем. Частный случай этой задачи, относящийся к кодированию недоопределенной информации, рассматривался в ([7]).

2°. Пусть имеются конечные алфавиты $A=\{a_1, \ldots, a_s\}$ и $B=\{b_1, \ldots, b_t\}$. Исходная последовательность $\tilde{x}=(x_1, \ldots, x_n)$, являющаяся словом в алфавите A, должна быть воспроизведена у пользователя в виде последовательности $\tilde{y}=(y_1, \ldots, y_n)$ той же длины в алфавите B. Последовательность $\tilde{y}$ будем называть *аппроксимацией* последовательности $\tilde{x}$. Качество воспроизведения определяется некоторой совокупностью критериев $f_1, \ldots, f_N$. Каждый критерий f_α характеризуется побуквенной мерой искажения $\rho^{(\alpha)}(a_i, b_j)=a_{ij}^{(\alpha)}$, где $0\leqslant a_{ij}\leqslant\infty$ представляет собой „штраф" за то, что вместо символа a_i воспроизведено b_j. *Общее искажение* между последовательностями $\tilde{x}$ и $\tilde{y}$

$$\rho^{(d)}(\tilde{x}, \tilde{y}) = \sum_{u=1}^{n} \rho^{(d)}(x_u, y_u) = \sum_{i,j} \alpha_{ij}^{(d)} w_{ij},$$

где w_{ij} означает число позиций u, для которых $x_u = a_i$, $y_u = b_j$. По каждому критерию f_d задан допустимый уровень относительного искажения $\varepsilon^{(d)} \geqslant 0$, $d = 1, \ldots, N$, и требуется, чтобы аппроксимация $\tilde{y} = \tilde{y}(\tilde{x})$ удовлетворяла совокупности условий

$$\rho^{(d)}(\tilde{x}, \tilde{y}) \leqslant \varepsilon^{(d)} n, \quad d = 1, \ldots, N.$$

Такую аппроксимацию будем называть $\tilde{\varepsilon}$-*точной*, где $\tilde{\varepsilon} = \varepsilon^{(1)}, \ldots, \varepsilon^{(N)})$. Нетрудно видеть, что для выполнимости этих условий при малых значениях $\varepsilon^{(1)}, \ldots, \varepsilon^{(N)}$ необходимо, чтобы для каждого i существовало такое $j = j(i)$, что $\alpha_{ij(i)}^{(d)} = 0$ при всех d.

Обозначим через $m_n(k_1, \ldots, k_s)$ класс всех последовательностей длины n, содержащих k_i символов a_i $(k_1 + \ldots + k_s = n)$. Далее считаем, что $k_1 = k_1(n), \ldots, k_s = k_s(n)$. Пусть имеется некоторый *способ кодирования* $K = K_n$, который любой последовательности $\tilde{x} \in m_n(k_1, \ldots, k_s)$ ставит в соответствие двоичный набор (код) $K(\tilde{x})$, и пусть имеется *способ декодирования* $D = D_n$, такой, что $D(K(\tilde{x}))$ является словом длины n в алфавите B. Слово $\tilde{y} = D(K(\tilde{x}))$ будем считать аппроксимацией $\tilde{x}$. Обозначим через $l(\tilde{x})$ длину кодового слова $K(\tilde{x})$ и положим

$$l_n(k_1, \ldots, k_s) = \max_{\tilde{x} \in m_n(k_1, \ldots, k_s)} l(\tilde{x}).$$

Точность воспроизведения будем характеризовать величинами

$$\rho^{(d)}(\tilde{x}) = \rho^{(d)}(\tilde{x}, D(K(\tilde{x}))), \quad d = 1, \ldots, N.$$

Метод кодирования—декодирования будем называть $\tilde{\varepsilon}$-*точным*, где $\tilde{\varepsilon} = (\varepsilon^{(1)}, \ldots, \varepsilon^{(N)})$, $\varepsilon^{(1)} \geqslant 0, \ldots, \varepsilon^{(N)} \geqslant 0$, если для всех $\tilde{x} \in m_n(k_1, \ldots, k_s)$

$$\rho^{(d)}(\tilde{x}) \leqslant \varepsilon^{(d)} n, \quad d = 1, \ldots, N.$$

При заданных $\tilde{\varepsilon} = (\varepsilon^{(1)}, \ldots, \varepsilon^{(N)})$ и $P = (p_1, \ldots, p_s)$, $\sum_i p_i = 1$, определим *функцию* $\tilde{\varepsilon}$-*энтропии* следующим образом*:

$$H_{\tilde{\varepsilon}}(P) = \min_{\|p_{ij}\|} \sum_{i,j} p_{ij} \log \frac{p_{ij}}{p_i \sum_u p_{uj}},$$

где минимум берется по всевозможным $(s \times t)$-матрицам $\|p_{ij}\|$, удовлетворяющим условиям

$$\sum_j p_{ij} = p_i, \quad p_{ij} \geqslant 0, \quad \sum_{i,j} \alpha_{ij}^{(d)} p_{ij} \leqslant \varepsilon^{(d)}, \quad d = 1, \ldots, N.$$

Теорема. 1) *Пусть выполнено условие*

$$\frac{H_{\tilde{\varepsilon}}\left(\frac{k_1}{n} \cdots \frac{k_s}{n}\right) \log n}{n \log \log n} \to \infty.$$

* Здесь и далее под $\log x$ понимается $\log_2 x$.

Тогда для произвольной функции $\alpha(n)\to\infty$ существуют $\tilde{\varepsilon}$-точное кодирование K_n и декодирование D_n, для которых максимальная длина удовлетворяет оценке

$$l_n(k_1,\ldots,k_s)\leqslant n\left(H_{\tilde\varepsilon}\left(\frac{k_1}{n},\ldots,\frac{k_s}{n}\right)+\frac{\alpha(n)\log\log n}{\log n}\right).$$

2) *Если функция $\alpha(n)$ допускает вычисление со сложностью n, то кодирование K_n и декодирование D_n имеют сложность, не большую, чем $n^{1+\gamma}$, где $\gamma>0$ произвольно*.*

3) *Для любых ε-точных способов кодирования и декодирования***

$$l_n(k_1,\ldots,k_s)\geqslant nH_{\tilde\varepsilon}\left(\frac{k_1}{n},\ldots,\frac{k_s}{n}\right)-c_1\log n. \qquad (1)$$

Под *сложностью алгоритма*, фигурирующей в пункте 2, можно понимать число шагов или потребную память машины с произвольным доступом к памяти (RAM ([8])), использующей побуквенные операции, либо число элементов схем из функциональных элементов ([9]), осуществляющих кодирование и декодирование.

3°. Пусть $W=\|w_{ij}\|$ является $(s\times t)$-матрицей, удовлетворяющей условию

$$\sum_j w_{ij}=k_i,\quad i=1,\ldots,s.$$

С матрицей W будем связывать функцию

$$I(W)=nI(\|p_{ij}\|),$$

где

$$I(\|p_{ij}\|)=\sum_{i,j}p_{ij}\log\frac{p_{ij}}{\sum_u p_{uj}\sum_v p_{iv}},\quad p_{ij}=\frac{w_{ij}}{n}.$$

Положим*** $\Delta=\left[\dfrac{n}{\log^2 n}\right]$ и рассмотрим всевозможные $(s\times t)$-матрицы $W=\|w_{ij}\|$, удовлетворяющие условиям

$$\sum_{i,j}\alpha_{ij}^{(d)}w_{ij}\leqslant\varepsilon^{(d)}n,\quad d=1,\ldots,N$$

и такие, что при всех $j\neq j(i)$

$$w_{ij}=l_{ij}\Delta,$$

где l_{ij} целые числа, а при $j=j(i)$

$$w_{ij(i)}=k_i-\sum_{j\neq j(i)}w_{ij}.$$

Среди них найдем матрицу $W_0=\|w_{ij}^0\|$ с минимальным значением $I(W)$.

---

* Условие вычислимости функции $\alpha(n)$ со сложностью n необременительно, ибо оно означает, что допустима экспоненциальная сложность относительно длины двоичной записи аргумента.

** Здесь и далее буквами c с индексами будем обозначать константы.

*** Через $[z]$ (через $]z[$) обозначено наибольшее (наименьшее) целое число, не превосходящее (не меньше) z.

Лемма 1. 1) *Для матрицы W_0 выполнено неравенство**

$$I(W_0) \leqslant nH_{\tilde{\varepsilon}}\left(\frac{k_1}{n}, \ldots, \frac{k_s}{n}\right) + O\left(\frac{n}{\log n}\right).$$

2) *Вычислительная сложность нахождения матрицы W_0 ограничена величиной $n(\log n)^{2st+4}$.*

4°. Введем величины

$$q^0(j/i) = \frac{w^0_{ij}}{k_i}, \quad i=1, \ldots, s, \; j=1, \ldots, t$$

и образуем $(s \times t)$-матрицу $Q^0 = \|q(j/i)\|$.

Пусть задан целочисленный параметр $\nu \leqslant n$. Рассмотрим произвольный набор $\varkappa_1, \ldots, \varkappa_s$, такой, что $\sum\limits_i \varkappa_i = \nu$.

Обозначим

$$I\left(\frac{\varkappa_1}{\nu}, \ldots, \frac{\varkappa_s}{\nu}; \; Q^0\right) = \sum_{i,j} \frac{\varkappa_i}{\nu} q^0(j/i) \log \frac{\frac{\varkappa_i}{\nu} q^0(j/i)}{\frac{\varkappa_i}{\nu} \sum\limits_i \frac{\varkappa_i}{\nu} q^0(j/i)}.$$

Скажем, что множество $\mathfrak{n}$ последовательностей в алфавите B $\tilde{\varepsilon}$*-точно аппроксимирует* множество $\mathfrak{m}$ последовательностей в алфавите A, если для каждой последовательности из $\mathfrak{m}$ в $\mathfrak{n}$ имеется некоторая ее $\tilde{\varepsilon}$-точная аппроксимация. Обозначим через $T_{\tilde{\varepsilon}}(\varkappa_1, \ldots, \varkappa_s)$ минимальную мощность множества $\mathfrak{n}$, $\tilde{\varepsilon}$-точно аппроксимирующего $\mathfrak{m}_\nu(\varkappa_1, \ldots, \varkappa_s)$.

Лемма 2. 1) *Имеет место оценка*

$$\log T_{\tilde{\delta}}(\varkappa_1, \ldots, \varkappa_s) \leqslant \nu I\left(\frac{\varkappa_1}{\nu}, \ldots, \frac{\varkappa_s}{\nu}; Q^0\right) + O(\log \nu),$$

где $\tilde{\delta} = (\delta^{(1)}, \ldots, \delta^{(N)})$, *причем*

$$\delta^{(d)} = \sum_{i,j} \alpha^{(d)}_{ij} \omega_{ij}, \quad \omega_{ij} = [\varkappa_i q^0(j/i)], \; j \neq j(i).$$

2) *Множество $\mathfrak{n}$, $\tilde{\delta}$-точно аппроксимирующее $\mathfrak{m}_\nu(\varkappa_1, \ldots, \varkappa_s)$ и удовлетворяющее оценке пункта 1, может быть найдено со сложностью, не превосходящей $c_2^\nu + c_3 \log^3 n$.*

Эта лемма доказывается на основе градиентной процедуры, использованной в ([6]) и являющейся модификацией процедуры из ([5]).

5°. Пусть имеется последовательность $\tilde{x} \in \mathfrak{m}_n(k_1, \ldots, k_s)$ в алфавите A. Зададимся натуральным параметром ν и разобьем последовательность $\tilde{x}$ на куски длины ν

$$\tilde{x} = \tilde{\chi}_1 \ldots \tilde{\chi}_M, \; M = \left]\frac{n}{\nu}\right[. \tag{2}$$

---

* Записи $\varphi_n \leqslant \psi_n$ и $\varphi = O(\psi_n)$ означают, что $\lim\limits_{n\to\infty} \frac{\varphi_n}{\psi_n} < \infty$, а запись $\varphi_n = o(\psi_n)$ — что $\lim\limits_{n\to\infty} \frac{\varphi_n}{\psi_n} = 0$.

Лемма 3. *Замена каждого куска $\tilde{\chi}_u$ его $\tilde{\delta}_u$-точной аппроксимацией $\tilde{\eta}_u$, построенной в соответствии с леммой 2, дает последовательность $\tilde{y}=\tilde{\eta}_1 \ldots \tilde{\eta}_M$, являющуюся $\tilde{\varepsilon}$-точной аппроксимацией $\tilde{x}$.*

6°. Опишем процедуру кодирования. *Кодовое слово* $K(\tilde{x})$ для последовательности $\tilde{x} \in m_n(k_1, \ldots, k_s)$ будет состоять из трех частей

$$K(\tilde{x}) = \Lambda \Sigma \Xi,$$

которые соответственно будем называть *справочной*, *главной* и *вспомогательной* частями.

Для построения главной части Σ кодового слова $K(\tilde{x})$ последовательность $\tilde{x} \in m_n(k_1, \ldots, k_s)$ представим в виде (2), разбив ее на куски длины ν. Произведем группировку кусков с одинаковыми параметрами $\varkappa_1, \ldots, \varkappa_s$ ($\varkappa_1 + \ldots + \varkappa_s = \nu$) в один класс и все полученные классы занумеруем $m_1, \ldots, m_R$. Для каждого класса $m_v = m_\nu(\varkappa_1, \ldots, \varkappa_s)$ с помощью леммы 2 отыщем множество n_v, $\tilde{\delta}_v$-точно аппроксимирующее m_v. Мощность n_v обозначим через $\hat{T}_{\tilde{\delta}_v}(\varkappa_1, \ldots, \varkappa_s)$. Расположим все последовательности $\tilde{\eta} \in n_v$ в некотором порядке и занумеруем их двоичными последовательностями $\tilde{\pi}(\tilde{\eta})$ длины $\tau_v =]\log \hat{T}_v[$. Если i является некоторым числом, то через $\tilde{i} = (i_1 \ldots i_p)$ будем обозначать его двоичную запись минимальной длины, а через $\tilde{i}^*$ — двоичную последовательность вида

$$\tilde{i}^* = (i_1 i_1 \ldots i_p i_p 01).$$

Код $\tilde{\sigma}(\tilde{\chi}_l)$ для куска $\tilde{\chi}_l$ из класса $m_v = m_\nu(\varkappa_1, \ldots, \varkappa_s)$ будет иметь вид

$$\tilde{\sigma}(\tilde{\chi}_l) = \tilde{v}^* \tilde{\pi}(\tilde{\eta}_j),$$

где $\tilde{\eta}_j$ означает $\tilde{\delta}_v$-точную аппроксимацию $\tilde{\chi}_l$.

Последовательность приписанных друг к другу кодов кусков $\tilde{\chi}$ в образует главную часть Σ кодового слова $K(\tilde{x})$. Длину главной части обозначим через L_Σ.

Вспомогательная часть Ξ представляет собой таблицу $\tilde{\delta}$-аппроксимаций. Пусть всеми $\tilde{\delta}_v$-аппроксимациями для класса $m_v = m_\nu(\varkappa_1, \ldots, \varkappa_s)$ являются $\tilde{\eta}_{v1}, \ldots, \tilde{\eta}_{v\hat{T}_v}$, где $\hat{T}_v = \hat{T}_{\tilde{\delta}_v}(\varkappa_1, \ldots, \varkappa_s)$.
Положим

$$\tilde{\xi}(m_v) = \tilde{\eta}_{v1} \ldots \tilde{\eta}_{v\hat{T}_v},$$

где $\hat{\eta}_{v\hat{T}_v}$ получается из $\tilde{\eta}_{v\hat{T}_v}$ путем замены каждого символа $b_j \in B$ последовательности $\tilde{\eta}_{v\hat{T}_v}$ двоичной записью $\tilde{j}$ длины $]\log t[$ его номера. Вспомогательная часть будет иметь вид

$$\Xi = \tilde{\xi}(m_1) \ldots \xi(m_R),$$

где R число классов $m_\nu(\varkappa_1, \ldots, \varkappa_s)$. Длину слова $\xi(m_i)$ обозначим через l_i, а общую длину вспомогательной части через L_Ξ.

Справочная часть Λ содержит числовые параметры, необходимые для того, чтобы по главной и вспомогательной частям осуществить декодирование. Она имеет вид

$$\Lambda = \tilde{n}^* \tilde{k}_1^* \ldots \tilde{k}_s^* \tilde{L}_\Sigma^* \tilde{L}_\Sigma^* \tilde{\tau}_1^* \ldots \tilde{\tau}_n^* \tilde{l}_1^* \ldots \tilde{l}_R^*.$$

7°. Для $\tilde{\chi} \in m_\nu(x_1, \ldots, x_s)$ введем обозначение

$$\varphi(\tilde{\chi}) = \nu I\left(\frac{x_1}{\nu}, \ldots, \frac{x_s}{\nu}, Q^0\right).$$

Следующая лемма является следствием выпуклости взаимной информации, при фиксированных переходных вероятностях ([10]).

Лемма 4.

$$\varphi(\tilde{\chi}_1 \ldots \tilde{\chi}_g) \geqslant \sum_{u=1}^{g} \varphi(\tilde{\chi}_u).$$

Можно подсчитать, что длина главной части составляет

$$L_\Sigma = \sum_{i=1}^{M} (\varphi(\tilde{\chi}_i) + O(\log \nu))$$

и в соответствии с леммой 4 и леммой 1 с учетом равенства $\varphi(\tilde{x}) = I(W_0)$ оценивается величиной

$$L_\Sigma \leqslant \varphi(\tilde{x}) + O\left(\frac{n}{\nu}\log \nu\right) < nH_\varepsilon\left(\frac{k_1}{n}, \ldots, \frac{k_s}{n}\right) +$$

$$+ O\left(\frac{n}{\log n}\right) + O\left(\frac{n \log \nu}{\nu}\right).$$

Вспомогательная и справочная части дают дополнительный вклад $c_4^\nu + \nu^{c_5} \log n$.

Положив

$$\nu = \left[\frac{\log n}{\alpha(n)}\right],$$

получаем требуемую оценку длины кода. Непосредственный подсчет показывает, что сложность процедур кодирования и декодирования удовлетворяет оценке пункта 2 теоремы.

8°. Способом, приложенным в ([6]) при доказательстве нижней оценки леммы 1, можно получить неравенство

$$\log T_\varepsilon(k_1, \ldots, k_s) \geqslant nH_\varepsilon\left(\frac{k_1}{n}, \ldots, \frac{k_s}{n}\right) - c_1 \log n,$$

из которого с помощью мощностных соображений получим оценку (1).

Замечание. На основе метода кодирования, предложенного в данной работе, может быть получено доказательство теоремы К. Шеннона ([1]) о кодировании источника с заданным критерием верности, при произвольной побуквенной мере искажения, не использующее метод случайного кодирования.

Автор выражает свою признательность Л. А. Шоломову за большую помощь в процессе выполнения работы.

ВНИИ системных
исследований ГКНТ и АН СССР

Ա. Լ. ՆԵՐՍԻՍՅԱՆ

Տրված աղավաղման մակարդակով դիսկրետ ինֆորմացիայի կոդավորման ալգորիթմ

Աշխատանքում քննարկվում է n երկարությամբ բառերի տրված ճշտությամբ կոդավորման խնդիրը, միջտառային աղավաղման մի քանի չափանիշների առկայության դեպքում։

Առաջարկվում է մեթոդ, որը ապահովում է ասիմպտոտիկ նվազագույն երկարության կոդ։ Մեթոդի հաշվողական բարդությունը չի գերազանցում $n^{1+\gamma}$, *որտեղ՝* $\gamma > 0$ *կամայական հաստատուն է։*

ЛИТЕРАТУРА — ԳՐԱԿԱՆՈՒԹՅՈՒՆ

[1] *К. Шеннон*, Работы по теории информации и кибернетике, ИЛ, М. 1963. [2] *T. Berger*, Rate Distortion Theory, Prentice-Hall, Engenlood Cliffs, N. Y., 1971. [3] *Ю. М. Штарьков*, в кн.: Информационный обмен в вычислительных сетях, Наука, М., 1980. [4] *О. Б. Лупанов*, в сб.: Проблемы кибернетики, вып. 14 (1965). [5] *Э. И. Нечипорук*, ДАН СССР, т. 163, № 1 (1965). [6] *Л. А. Шоломов*, в сб.: Проблемы кибернетики, вып. 19 (1967). [7] *А. Л. Нерсисян*, ВНИИСИ, Сборник «Методы исследований сложных систем». Труды конференции аспирантов и молодых специалистов ВНИИСИ, М. 1981, ст. 29—33. [8] *А. Ахо, Дж. Хопкрофт, Дж. Ульман*, Построение и анализ вычислительных процедур, Мир, М., 1979. [9] *О. Б. Лупанов*, в сб.: Проблемы кибернетики, вып. 7 (1962). [10] *Р. Галлагер*, Теория информации и надежная связь, Советское радио, М., 1974.