

УДК 519.725

ВОССТАНОВЛЕНИЕ КОДОВ ПО КОЭФФИЦИЕНТАМ КОРРЕЛЯЦИИ ИХ ПОДКОДОВ *)

С. В. Августинович, Е. В. Горкунов

Аннотация. Получено обобщение на произвольный случай теоремы о восстановлении двоичного кода по набору размерностей его подкодов. Предложено понятие коэффициента корреляции набора подкодов, который в настоящем рассмотрении является аналогом размерности двоичного подкода.

Ключевые слова: код, сильная изометрия, восстановление, эквивалентные коды.

Введение

Рассмотрим q -ичный куб E_q^n — множество слов длины n над алфавитом $A = \{0, 1, 2, \dots, q-1\}$. Расстояние Хэмминга $d(x, y)$ между двумя словами $x, y \in E_q^n$ определяется числом символов, в которых x и y различаются, т. е. $d(x, y) = |\{i \mid x_i \neq y_i\}|$. Вес $w(x)$ слова $x \in E_q^n$ полагается равным количеству его ненулевых символов, или $w(x) = d(x, 0)$. Куб E_q^n вместе с определённым на его элементах (вершинах) расстоянием Хэмминга образует метрическое пространство. Кодом называется произвольное подмножество C пространства E_q^n . Элементы кода называются *кодowymi словами*. Два кода *эквивалентны*, если существует изометрия пространства E_q^n , отображающая один код в другой.

В настоящей статье изучаются метрические инварианты кодов, обобщается понятие размерности двоичного кода [1] и устанавливается достаточность предложенных инвариантов для восстановления кодов с точностью до эквивалентности.

Хорошо известно [1–5], что равенство тех или иных метрических инвариантов у двух кодов далеко не всегда означает их эквивалентность. При $q = 2$ оказалось, что если между парой двоичных кодов имеется биекция, сохраняющая размерность каждого подкода, то эта биекция

*) Исследование выполнено при финансовой поддержке Российского фонда фундаментальных исследований (проект № 10-01-00424) и целевой программы СО РАН на 2012–2014 гг. (интеграционный проект № 14).

продолжается до изометрии всего пространства. Иначе говоря, набор размерностей подкодов двоичного кода задаёт этот код с точностью до эквивалентности. Здесь под *размерностью кода* понимается размерность минимальной грани куба E_2^n , содержащей этот код. В дальнейшем выяснилось [6], что полный набор размерностей подкодов избыточен, и достаточно иметь размерности подкодов лишь чётной мощности — биекция между кодами, сохраняющая такие размерности, продолжается до изометрии E_2^n , а коды оказываются эквивалентными.

Прямое обобщение подхода, применённого для двоичных кодов, при произвольном $q > 2$ не даёт ничего существенного. Имеются, например, неэквивалентные троичные коды, между которыми существует биекция, сохраняющая размерности их подкодов. Таким образом, в общем случае для q -ичных кодов необходимы более тонкие методы. Предлагаемый в этой статье подход для наглядности продемонстрируем на троичных кодах. Обобщение для произвольного целого положительного q происходит естественным образом (см. разд. 3).

1. Необходимые определения и основной результат

Пусть C — произвольный троичный код длины n . Если все кодовые слова кода C в некоторой позиции имеют один и тот же символ, то будем говорить, что эта позиция *несущественная* для кода C . Если C_1 и C_2 — непересекающиеся подкоды кода C , то через $K(C_1, C_2)$ обозначим число координатных позиций, несущественных для обоих подкодов, но в которых кодовые слова из различных кодов имеют разные значения. Формально

$$K(C_1, C_2) = |\{i \mid \exists a, b \in A, a \neq b : \forall x \in C_1 x_i = a, \forall y \in C_2 y_i = b\}|.$$

Величину $K(C_1, C_2)$ назовём *коэффициентом корреляции* кодов C_1 и C_2 . Отметим некоторые характерные равенства, которые проясняют это понятие:

- (i) $K(x, y) = d(x, y)$, где $x, y \in E_3^n$;
- (ii) $K(\{x, y\}, \emptyset) = n - d(x, y)$;
- (iii) $K(C, \emptyset) = n - \text{Dim}(C)$, где $\text{Dim}(C)$ — размерность кода C .

Таким образом, с помощью коэффициентов корреляции, в частности, можно определить расстояния между кодовыми словами.

По аналогии с [6] биекцию $I: C_1 \rightarrow C_2$ между кодами $C_1, C_2 \subset E_3^n$ назовём *сильной изометрией*, если она сохраняет коэффициент корреляции любой пары подкодов кода C_1 . Пусть M_1 — кодовая матрица кода C_1 ,

обозначим через $M_2 = I(M_1)$ кодовую матрицу кода C_2 , полученную применением отображения I к каждой строке M_1 . Основным результатом настоящей статьи является

Теорема 1. *Произвольная сильная изометрия троичных кодов продолжается до изометрии троичного куба.*

Рассмотрим код $C \subset E_3^n$ мощности m , и пусть M — его кодовая матрица. Символы алфавита A для произвольного столбца M порождают *алфавитное разбиение* множества $\{1, \dots, m\}$ номеров строк: каждое подмножество разбиения включает все номера строк, в которых элементы рассматриваемого столбца имеют одинаковые значения. Например, столбцам $(0, 1, 0, 1, 2)^T$ и $(2, 1, 2, 1, 0)^T$ соответствует разбиение $(\{1, 3\}, \{2, 4\}, \{5\})$, а также ещё пять упорядоченных разбиений, отличающихся от указанного порядком записи множеств. Для произвольного 3-разбиения P через $k(P)$ обозначим число столбцов матрицы M , для которых P — одно из шести алфавитных разбиений.

В дальнейшем будем придерживаться следующих обозначений:

$M = \{1, \dots, m\}$ — множество номеров строк матрицы M ;

$P = (P_0, P_1, P_2)$, $Q = (Q_0, Q_1, Q_2)$, $R = (R_0, R_1, R_2)$ — алфавитные разбиения M .

Для любого подмножества $S \subseteq M$ через $C(S)$ обозначим подкод кода C , образованный строками матрицы M с номерами из S . Если ясно, о каком коде идёт речь, то для $P_1, P_2 \subseteq M$ вместо $K(C(P_1), C(P_2))$ будем писать $K(P_1, P_2)$.

Несложно показать, что справедливо

Утверждение 1. *Произвольная биекция $I: C_1 \rightarrow C_2$ продолжается до изометрии всего куба тогда и только тогда, когда для любого алфавитного разбиения кодовые матрицы M_1 и $M_2 = I(M_1)$ содержат одинаковое число столбцов с таким разбиением.*

На множестве алфавитных разбиений введём частичный порядок \preceq . Положим $(P_0, P_1, P_2) \preceq (Q_0, Q_1, Q_2) \iff P_0 \subseteq Q_0, P_1 \supseteq Q_1$ и $P_2 \supseteq Q_2$. Заметим, что для произвольного кода $C \subset E_3^n$ мощности m , его фиксированной кодовой матрицы M и алфавитного разбиения $P = (P_0, P_1, P_2)$ имеет место равенство

$$K(P_1, P_2) = \sum_{Q \preceq P} k(Q). \quad (1)$$

2. Восстановление троичных кодов по коэффициентам корреляции их подкодов

Одним из этапов в доказательстве теоремы 1 является обращение формулы (1). Иными словами, ближайшая цель — имея коэффициенты корреляции подкодов, научиться вычислять количество столбцов определённого типа в кодовой матрице.

Утверждение 2. Число столбцов кодовой матрицы произвольного троичного кода, имеющих алфавитное разбиение P , равно

$$k(P) = \sum_{Q \preccurlyeq P} (-1)^{|P_0| - |Q_0|} K(Q_1, Q_2). \quad (2)$$

Доказательство проведём индукцией по мощности P_0 . Согласно (1) если $P_0 = \emptyset$, то $K(P_1, P_2) = k(P)$. Таким образом построена база индукции. Пусть утверждение верно при $|P_0| < s$, докажем его для $|P_0| = s$.

Из (1) переносом слагаемых получаем

$$k(P) = K(P_1, P_2) - \sum_{Q \prec P} k(Q).$$

Далее воспользуемся индукционным предположением, подставим (2) в последнее равенство и изменим порядок суммирования:

$$\begin{aligned} k(P) &= K(P_1, P_2) - \sum_{Q \prec P} \sum_{R \preccurlyeq Q} (-1)^{|Q_0| - |R_0|} K(R_1, R_2) \\ &= K(P_1, P_2) - \sum_{R \prec P} (-1)^{-|R_0|} K(R_1, R_2) \sum_{R \preccurlyeq Q \prec P} (-1)^{|Q_0|}. \end{aligned} \quad (3)$$

Зафиксируем разбиение $R \prec P$, и пусть $|R_0| = t$, где $0 \leq t < s$. Для того чтобы получить разбиение Q из интервала между R и P такое, что $|Q_0| = t + i$, необходимо и достаточно некоторые i элементов из множества $P_0 \setminus R_0$ присоединить к R_0 . Здесь i может принимать любые целые значения от 0 до $s - t - 1$. Таким образом,

$$\begin{aligned} \sum_{R \preccurlyeq Q \prec P} (-1)^{|Q_0|} &= \sum_{i=0}^{s-t-1} C_{s-t}^i (-1)^{t+i} \\ &= (-1)^t (1 - 1)^{s-t} - (-1)^t (-1)^{s-t} = -(-1)^s. \end{aligned}$$

Продолжив с учётом этого равенства (3), получим

$$\begin{aligned} k(P) &= K(P_1, P_2) + \sum_{R \prec P} (-1)^{s-|R_0|} K(R_1, R_2) \\ &= \sum_{R \preceq P} (-1)^{|P_0|-|R_0|} K(R_1, R_2). \end{aligned}$$

Утверждение 2 доказано.

ДОКАЗАТЕЛЬСТВО теоремы 1. Из утверждения 2 следует, что матрицы M_1 и M_2 имеют одинаковые с точностью до перестановки наборы алфавитных разбиений. Остаётся применить утверждение 1.

Следствие 1. *Сильно изометричные троичные коды эквивалентны.*

3. Замечания

В завершение приводим некоторые суждения, касающиеся обобщения результатов настоящей статьи.

1. Теорема 1 может быть обобщена для кодов над алфавитом произвольной мощности q следующим образом. Для такого кода необходимо принять в рассмотрение коэффициент корреляции его произвольных $q-1$ попарно не пересекающихся подкодов. При таком переходе к общему случаю формулировка теоремы практически не меняется.

2. Основной целью при доказательстве продолжаемости сильной изометрии в том виде, в котором оно изложено в настоящей статье, является выяснение столбцового состава кодовых матриц рассматриваемых кодов. В кодовой матрице троичного кода мощности m могут встречаться 3^m различных столбцов (при достаточной длине кода). Поэтому, вообще говоря, для установления столбцового состава кодовой матрицы необходимо вычислить 3^m величин, характеризующих число вхождений каждого m -столбца в указанную кодовую матрицу. Очевидно, что знание коэффициентов корреляции каждой пары подкодов для этой цели избыточно. Таким образом, возникает задача о поиске минимального набора коэффициентов корреляции, достаточного для восстановления кодовой матрицы и установления эквивалентности кодов.

ЛИТЕРАТУРА

1. **Августинович С. В.** О сильной изометрии бинарных кодов // Дискрет. анализ и исслед. операций. Сер. 1. — 2000. — Т. 7, № 3. — С. 3–5.
2. **Абдурахманов Ж. К.** О геометрической структуре кодов, исправляющих ошибки: Дис. ... канд. физ-мат. наук: 01.01.09. — Ташкент, 1991. — 66 с.

3. **Августинович С. В., Соловьёва Ф. И.** К метрической жёсткости двоичных кодов // Проблемы передачи информ. — 2003. — Т. 39, № 2. — С. 23–28.
4. **Красин В. Ю.** О слабых изометриях булева куба // Дискрет. анализ и исслед. операций. Сер. 1. — 2006. — Т. 13, № 4. — С. 26–32.
5. **Solov'eva F. I., Avgustinovich S. V., Honold T., Heise W.** On the extendability of code isometries // J. Geom. — 1998. — V. 61. — P. 3–16.
6. **Горкунов Е. В., Августинович С. В.** О восстановлении двоичных кодов по размерностям их подкодов // Дискрет. анализ и исслед. операций. — 2010. — Т. 17, № 5. — С. 15–21.

Августинович Сергей Владимирович,
e-mail: avgust@math.nsc.ru
Горкунов Евгений Владимирович,
e-mail: evgumin@gmail.com

Статья поступила
27 марта 2012 г.
Переработанный вариант —
17 октября 2012 г.