УДК 519.72 (075.8)

ПОСТРОЕНИЕ ИНЪЕКТИВНЫХ РЮКЗАЧНЫХ ВЕКТОРОВ НА ОСНОВЕ СТРУКТУРНЫХ И ЧАСТОТНЫХ СВОЙСТВ ЧИСЛОВЫХ МНОЖЕСТВ

Подколзин В. В.[1]

CONSTRUCTION OF THE INJECTIVE KNAPSACK VECTORS ON THE BASIS OF THE NUMERICAL SETS STRUCTURE AND STATISTICS PROPERTIES

Podkolzin V. V.

Properties of numbers sequences expressed through components of a knapsack vector are investigated. The search method of knapsack vector for the set of known values is defined.

Keywords: a knapsack vector, variation, injectiveness, sequence of values

Задача о рюкзаке для заданных натурального $v \in N$ и вектора $A = (a_1, a_2, \ldots, a_n)$, где $a_i \in N$, $i = 1, \ldots, n$, имеет решение в $Z_2 = \{0,\ 1\}$, если существует решение x уравнения [1]

$$Ax^T = w, x \in Z_2^n. \qquad (1)$$

Обозначим как $\mu(A)$ множество значений w, для которых уравнение (1) имеет решение. Уравнение (1) определяет отображение множества двоичных векторов Z_2^n в числовое множество $\mu(A)$, для которого вектор $A = (a_1, a_2, \ldots, a_n)$ является параметром. Вектор A — рюкзачный вектор. Из таких отображений особое значение имеют инъективные отображения, позволяющие по значению из $\mu(A)$ однозначно определить соответствующий ему x.

Рассмотрим NP-полную задачу поиска вектора A для известного множества $W \subseteq \mu(A)$. Основными методами нахождения A являются:

1) решение системы линейных уравнений относительно переменных $a_1, a_2, \ldots, a_n$;

2) решение системы нелинейных уравнений относительно переменных $a_1, a_2, \ldots, \ldots, a_n, x_1, x_2, \ldots, x_n$.

В первом случае $x = (x_1, x_2, \ldots, x_n)$ определяется последовательным перебором всех двоичных векторов длины n на основании специальных свойств W. Количество возможных решений зависит от числа известных значений и свойств рюкзачного вектора. В частности, если известно только одно значение результата инъективного отображения, то верхняя граница числа возрастающих векторов, его определяющих, равна [2]

$$\frac{n!}{[n/2]![(n+1)/2]!}.$$

Если же известны 2^n различных значений, такое решение единственное. Для рассматриваемых методов поиск рюкзачных векторов основывается только на значениях элементов множества W, но не использует свойства множества $\mu(A)$, которому они также принадлежат. Такой поиск требует большого объема вычислений.

Практическая эффективность решения NP-полной задачи определяется наличием эвристики и неформальных процедур в проблемной области ее применения. В данной работе предлагается метод решения NP-полной задачи нахождения вектора A для известного множества $W \subseteq \mu(A)$, основанный на специальных свойствах $\mu(A)$.

Определение. Вариацией вектора $A = (a_1, a_2, \ldots, a_n)$ в Z_2 назовем вектор

$$\Delta A = (\delta_1, \delta_2, \ldots, \delta_n),$$

где

$$\delta_1 = a_1, \quad \delta_i = a_i - \sum_{j=1}^{i-1} a_j, \quad i = 2, \ldots, n.$$

Определение. Последовательность

$$W_{\mu(A)} = (w_i \mid i = 0, \ldots, 2^n - 1),$$

[1]Подколзин Вадим Владиславович, старший преподаватель кафедры информационных технологий Кубанского государственного университета; e-mail: vvp_35@mail.ru

где

$$w_i = Ax_i^T, \quad x_i = (\alpha_1, \alpha_2, \ldots, \alpha_n),$$

$$i = \sum_{j=1}^{n} \alpha_j 2^{j-1},$$

называется последовательностью значений A.

Определение. Последовательность

$$\Delta W_{\mu(A)} = (\omega_1, \omega_2, \ldots, \omega_{2^n-1}),$$

где

$$\omega_i = w_i - w_{i-1}, \quad w_i, w_{i-1} \in W_{\mu(A)},$$

$$i = 1, \ldots, 2^n - 1,$$

называется вариацией

$$W_{\mu(A)} = (w_0, w_1, \ldots, w_{2^n-1}).$$

Для двух рюкзачных векторов $A = (a_1, a_2, \ldots, a_n)$, $B = (a_1, a_2, \ldots, a_n, a_{n+1})$ справедливо

$$\Delta W_{\mu(B)} = (\Delta W_{\mu(A)}, \delta_{n+1}, \Delta W_{\mu(A)}). \qquad (2)$$

Последнее соотношение следует из того, что

$$W_{\mu_(B)} = (w_0^A, w_1^A, \ldots, \ldots, w_{2^n-1}^A, a_{n+1}, w_0^A + a_{n+1}, w_1^A + a_{n+1}, \ldots, \ldots, w_{2^n-1}^A + a_{n+1}),$$

где

$$w_i^A \in W_{\mu(A)}, \quad i = 1, \ldots, 2^n - 1.$$

Рассмотрим метод поиска рюкзачного вектора по заданному множеству $\mu(A)$, использующий свойства последовательности $W_{\mu(A)}$.

Для возрастающей последовательности значений

$$W = (w_1, w_2, \ldots, w_m)$$

требуется построить инъективный рюкзачный вектор

$$A = (a_1, a_2, \ldots, a_n), \quad W \subseteq \mu(A).$$

Для любой последовательности W существует хотя бы один такой вектор, например, вектор

$$(2^0, 2^1, 2^2, 2^3, \ldots, 2^r),$$

где $r = [\log_2 w_m] + 1$.

Применим специальную схему построения вектора $\Delta A = (\delta_1, \delta_2, \ldots, \delta_n)$ (вариации A), по которому однозначно определяется рюкзачный вектор. По предположению каждый компонент последовательности W должен принадлежать $\mu(A)$, тогда все элементы W встречаются в $W_{\mu(A)}$.

Для заданной величины n построим $\Delta W_{\mu(A)}$ в общем виде. Каждой подпоследовательности $\Delta W_{\mu(A)}$ сопоставим сумму ее элементов, а каждой сумме — количество вхождений в $\Delta W_{\mu(A)}$ подпоследовательностей с такой суммой.

Например, для $A = (a_1, a_2, a_3)$ последовательность $\Delta W_{\mu(A)}$ имеет вид $(\delta_1, \delta_2, \delta_1, \delta_3, \delta_1, \delta_2, \delta_1)$. Здесь сумме $\delta_1+\delta_2+\delta_1+\delta_3$ соответствуют подпоследовательности $(\delta_1, \delta_2, \delta_1, \delta_3)$, $(\delta_2, \delta_1, \delta_3, \delta_1)$ $(\delta_1, \delta_3, \delta_1, \delta_2)$, $(\delta_3, \delta_1, \delta_2, \delta_1)$. Таких подпоследовательностей $\Delta W_{\mu(A)}$ четыре. Полный список соответствий для $A = (a_1, a_2, a_3)$ приведен в таблице.

По таблице статистики вхождений подпоследовательностей для заданного значения n определим последовательность пар $S_n = \{(x_i, s_i)\}$, где $s_i = \delta_{i_1} + \delta_{i_2} + \cdots + \delta_{i_k}$, $\delta_{i_j} \in \Delta A$ $(j = 1, \ldots, k)$, x_i — количество вхождений в $\Delta W_{\mu(A)}$ всех последовательностей длины k, образованных $\delta_{i_1}, \delta_{i_2}, \ldots, \delta_{i_k}$.

Последовательность S_n упорядочим по правилу

$$\forall (x_i, s_i) \in S_n, (x_j, s_j) \in S_n \Big(i \leqslant j \Leftrightarrow$$

$$\Leftrightarrow (x_i < x_j) \vee \Big((x_i = x_j) \wedge$$

(последовательность, образованная

$$\delta_{i_1}, \delta_{i_2}, \ldots, \delta_{i_k},$$

впервые встречается в $\Delta W_{\mu(A)}$ ранее

последовательности, образованной

$$\delta_{j_1}, \delta_{j_2}, \ldots, \delta_{j_m}) \Big) \Big).$$

Для заданной возрастающей последовательности значений W найдем

$$\Delta W = \{\omega_i | \omega_i = w_i - w_{i-1}\},$$

полагая, что $w_0 = 0$. На основе ΔW определим последовательность пар $S' = \{(k'_i, s'_i)\}$ аналогично S_n.

Поиск рюкзачного вектора определяется следующей схемой:

1. Полагаем n равным $[\log_2 |W|]$;
2. Увеличим n на 1;

Статистика вхождений подпоследовательностей для $n = 3$

Сумма элементов подпоследовательности $\Delta W_{\mu(A)}$	Количество вхождений
δ_1	4
$\delta_1 + \delta_2$	4
$\delta_1 + \delta_2 + \delta_1 + \delta_3$	4
δ_2	2
$\delta_1 + \delta_3$	2
$\delta_2 + \delta_1 + \delta_3$	2
$\delta_1 + \delta_2 + \delta_1 + \delta_3 + \delta_1$	2
$\delta_1 + \delta_2 + \delta_1 + \delta_3 + \delta_1 + \delta_2$	2
δ_3	1
$\delta_2 + \delta_1 + \delta_3$	1
$\delta_1 + \delta_2 + \delta_1 + \delta_3 + \delta_1 + \delta_2 + \delta_1$	1

3. Определим очередной вектор

$$\zeta = (s'_{j_1}, \ldots, s'_{j_n}),$$

где

$$\exists\, (x'_{j_k}, s'_{j_k}) \in S', \quad k = 1, \ldots, n,$$

$$(m < k \leqslant n) \Leftrightarrow (i_m < i_k).$$

Если указать очередной выбор невозможно, то переходим к п. 2;

4. Определим очередной вектор

$$\xi = (s_{i_1}, \ldots, s_{i_n}),$$

где

$$\exists\, (x_{i_k}, s_{i_k}) \in S_n x'_{j_k} \leqslant x_{i_k}, \quad k = 1, \ldots, n,$$

$$(m < k \leqslant n) \Leftrightarrow (i_m < i_k).$$

Если указать очередной выбор невозможно, то перейдем к п. 3;

5. Приравняем соответствующие компоненты векторов ξ и ζ и решим систему n линейных уравнений относительно элементов вектора ΔA. Если система неразрешима, перейдем к п. 4;

6. Проверим, определяет ли найденная вариация ΔA рюкзачный вектор, в котором выражаются все элементы W. Если нет, то перейдем к п. 4.

В пп. 3 и 4 перебор начинается с векторов $\xi = (s_1, \ldots, s_n)$ и $\zeta = (s'_1, \ldots, s'_n)$.

Предложенный метод поиска гарантировано найдет рюкзачный вектор минимальной размерности, так как в худшем случае $n = |W|$. Эффективность метода определяется тем фактом, что, чем раньше элемент встречается в S_n, тем больше вероятность того, что соответствующий ему элемент содержится в S'.

Пусть необходимо найти возрастающий инъективный рюкзачный вектор $A = (a_1, a_2, \ldots, a_n)$ минимальной размерности n, для которого выполняется $W = (5, 6, 8, 10, 12, 15, 21, 23) \subseteq W_{\mu(A)}$. Так как $|W| = 8$, то $n \geqslant 4$.

Построим последовательность пар

$$\begin{aligned} S' = \{&(5,2),(4,5),(3,3),(3,6),(3,7),\\ &(3,8),(3,10),(3,13),(3,15),(2,1),\\ &(2,4),(2,9),(2,11),(1,8),(1,12),\\ &(1,16),(1,17),(1,18),(1,21),(1,23)\}. \end{aligned}$$

Начальные элементы S_4 имеют вид:

$$\begin{gathered} (8,\delta_1),(8,\delta_1+\delta_2),(8,\delta_1+\delta_2+\delta_1+\delta_3),\\ (8,\delta_1+\delta_2+\delta_1+\delta_3+\delta_1+\delta_2+\delta_1+\delta_4),\\ (4,\delta_2),(4,\delta_1+\delta_3),(4,\delta_1+\delta_2+\delta_1),(4,\delta_1+\delta_2+\delta_3),\\ (4,\delta_1+\delta_2+\delta_1+\delta_3+\delta_1),(4,\delta_1+\delta_2+\delta_1+\delta_3+\delta_1+\delta_2),\\ (4,\delta_1+\delta_2+\delta_1+\delta_4),\ldots. \end{gathered}$$

В целях упрощения вычислений при построении системы уравнений проверим возможность выбора очередных компонентов векторов ξ и ζ. Последовательно сравнивая вторые компоненты каждой пары S' с соответствующими значениями S_4 получим:

1. Для $(5, 2) \in S'$ возможно соответствие первых четырех элементов S_4. Полагаем соответствие $\delta_1 = 2$ (если данное предположение окажется ложным, будем сравнивать 2 значениям вторых компонент остальных элементов S_4).

2. С учетом первого шага, так как из $\delta_1 + \delta_2 = 2$ следует $\delta_2 = 0$, что противо-

речит возрастанию отыскиваемого вектора, приравниваем $\delta_1 + \delta_2 = 5$. Получим $\delta_2 = 3$.

3. Поскольку $\delta_1 + \delta_2 + \delta_1 + \delta_3 = 3$, то $\delta_3 = -4$, следовательно искомый вектор не будет возрастающим (так как $\delta_2 = 3$, то логично предположить, что $(3, 3) \in S'$ соответствует $(4, \delta_2) \in S_4$). Проверяем следующий элемент S': $\delta_1+\delta_2+\delta_1+\delta_3 = 6$. Тогда $\delta_3 = -1$.

4. Для элемента

$$(8, \delta_1 + \delta_2 + \delta_1 + \delta_3 + \delta_1 + \delta_2 + \delta_1 + \delta_4) \in S_4,$$

последовательно проверяя элементы S', получим

1) $\delta_1 + \delta_2 + \delta_1 + \delta_3 + \delta_1 + \delta_2 + \delta_1 + \delta_4 = 2$, т.е. $\delta_4 = -11$;

2) $\delta_1 + \delta_2 + \delta_1 + \delta_3 + \delta_1 + \delta_2 + \delta_1 + \delta_4 = 5$ и $\delta_4 = -8$;

3) $\delta_1 + \delta_2 + \delta_1 + \delta_3 + \delta_1 + \delta_2 + \delta_1 + \delta_4 = 3$ и $\delta_4 = -10$;

4) $\delta_1 + \delta_2 + \delta_1 + \delta_3 + \delta_1 + \delta_2 + \delta_1 + \delta_4 = 6$ и $\delta_4 = -7$;

5) $\delta_1 + \delta_2 + \delta_1 + \delta_3 + \delta_1 + \delta_2 + \delta_1 + \delta_4 = 7$ и $\delta_4 = -6$;

6) $\delta_1 + \delta_2 + \delta_1 + \delta_3 + \delta_1 + \delta_2 + \delta_1 + \delta_4 = 8$ и $\delta_4 = -5$;

7) $\delta_1 + \delta_2 + \delta_1 + \delta_3 + \delta_1 + \delta_2 + \delta_1 + \delta_4 = 10$ и $\delta_4 = -3$.

Первые шесть вариантов противоречат возрастанию вектора A.

Проверяя последнее найденное значение, убеждаемся, что вектор $A = (2, 5, 6, 10)$ является искомым.

Следует отметить, что количество вариантов, которые необходимо проверить при поиске решений, зависит от мощности W.

Предложенный метод также применим для поиска рюкзачного вектора заданной размерности. В этом случае мощность множества W является существенным фактором, чем больше значение первой компоненты элемента S', тем меньше итераций поиска необходимо выполнить. Существенным фактором является наличие множества таких значений, которые в последовательности $W_{\mu(A)}$ образуют подпоследовательность за исключением небольшого количества случаев. В рассмотренном выше примере такими значениями являются 5, 6, 8, так как в $W_{\mu(A)}$ для найденного вектора имеется подпоследовательность 5, 7, 6, 8. Чем больше имеется таких «близких» значений, тем меньше может быть мощность множества W.

Можно предположить, что количество значений подпоследовательности W, принадлежащих отрезку $[a_n, \sum_{j=1}^{n} a_j]$, существенно влияет на сложность нахождения решения. В данном отрезке находятся значения, в разбиении которого используются все элементы рюкзачного вектора, а их статистические (в выше определенном смысле) характеристики соотносятся с характеристиками элементов последовательности S_n. Применение вышеописанного метода построения рюкзачного вектора имеет смысл только при $|W| \geqslant n$, как и всех известных неперебороных методов поиска рюкзачного вектора.

Метод нахождения инъективного вектора A уравнения (1) для известного множества $W \subseteq \mu(A)$, представленный в данной работе, позволяет существенным образом сократить перебор вариантов решений. Метод основан на гипотезе о том, что чем чаще совокупность значений вариации ΔA встречается в виде подпоследовательностей во множестве $\Delta W_{\mu(A)}$, тем выше вероятность ее появления в виде подпоследовательности в ΔW. Статистические характеристики таких совокупностей значений определяют возможность выявлять заведомо тупиковые варианты в процессе поиска решения. Использование свойств инъективного рюкзачного вектора позволило сократить количество кандидатов на значение очередного компонента вариации.

Литература

1. *Саломаа А.* Криптография с открытым ключом. М.: Мир, 1995. 318 с.
2. *Подколзин В. В., Осипян В. О.* Верхняя граница числа решений обобщенной задачи о рюкзаке на заданном входе // Актуальные проблемы безопасности информационных технологий: Материалы III Междунар. научно-практич. конф. Красноярск: Сибирский. гос. аэрокосмический ун-т, 2009. 142 с.

Ключевые слова: рюкзачный вектор, вариация, инъективность, последовательность значений

Статья поступила 26 ноября 2010 г.
Кубанский государственный университет, г. Краснодар
© Подколзин В. В., 2010