

Задача о покрытии множества с интервальными весами подмножеств и жадный алгоритм ее решения

А. В. ПРОЛУБНИКОВ

Омский государственный университет им. Ф.М. Достоевского, Россия

Контактный e-mail: a.v.prolubnikov@mail.ru

Рассматривается задача о покрытии множества с интервальными весами подмножеств. Представлены варианты определения слабого и сильного оптимальных решений задачи, объединенного оптимального решения. Доказывается теорема о характеристике сильного оптимального решения. Представлена модификация жадного алгоритма нахождения приближенного решения задачи, определяемого в работе. Рассматриваются вычислительная сложность представленного алгоритма и точность получаемых приближенных решений задачи.

Ключевые слова: задача о покрытии множества, интервальная неопределенность, жадный алгоритм.

Введение

Взвешенная задача о покрытии множества имеет следующую постановку. Дано множество U мощности $|U| = m$, дан конечный набор множеств $S = \{S_1, \dots, S_n\}$ такой, что $S_i \subseteq U$ и $\bigcup_{i=1}^n S_i = U$. Набор множеств $S' = \{S_{i_1}, \dots, S_{i_k}\}$, $S_{i_j} \in S$, называется *покрытием*

множества U , если $\bigcup_{i=1}^k S_{i_j} = U$. Для $S_i \in S$ заданы *веса* $w_i = w(S_i)$, $w_i > 0$. Вес $w(S')$ набора множеств $S' = \{S_{i_1}, \dots, S_{i_k}\}$ равен сумме весов входящих в него множеств. Необходимо найти покрытие множества U минимального веса — *оптимальное покрытие*.

В рассматриваемой нами далее интервальной постановке этой задачи функция веса на множествах из S принимает интервальные значения, т. е. для каждого множества из S заданы минимальный и максимальный возможные веса. Неопределенность веса множества в приложениях может быть вызвана неопределенностью данных, например погрешностью измерений или расчетов. Так, если задача размещения может быть поставлена как задача о покрытии множества, то стоимости размещения объектов могут быть даны с интервальными неопределенностями.

Задача о покрытии множества NP -полна [1], и она часто возникает при решении других оптимизационных задач и задач анализа данных. Ее приложения включают в себя задачи размещения [2], машинного обучения [3], распределения ресурсов [4], интеллектуального анализа данных [5], задачи, возникающие при проектировании интегральных схем [6], и др.

Точные алгоритмы решения задачи о покрытии множества основаны на методах ветвей и границ [6–9], отсечения [9, 10], перебора L -классов [11] и др. Поскольку точные алгоритмы для общего случая задачи экспоненциальны по времени, разработаны полиномиальные приближенные алгоритмы решения задачи, имеющие априорные оценки точности получаемых решений. Одним из них является жадный алгоритм.

При небольшой трудоемкости, составляющей $O(nm)$, как правило, жадный алгоритм в ходе вычислительных экспериментов дает хорошее приближение оптимального решения задачи. Так, в [12] показано, что для задач из [13] (библиотеки задач, используемой для тестирования алгоритмов решения задачи о покрытии) жадный алгоритм в среднем выбирает не более чем на 7% больше множеств, чем требуется для оптимального покрытия в невзвешенном случае, и получает покрытие в среднем не более чем на 13% большего веса, чем вес оптимального покрытия во взвешенном случае задачи. Для общего случая задачи о покрытии с положительными весами верна априорная оценка точности

$$\frac{w(Gr)}{w(Opt)} \leq H(m) \leq \ln m + 1, \quad (1)$$

где $H(m) = \sum_{k=1}^m 1/k$; Gr — покрытие, получаемое с помощью жадного алгоритма; Opt — оптимальное покрытие [14]. Показано, что при условии $P \neq NP$ получение полиномиального алгоритма с оценкой точности $(1 - \varepsilon) \ln m$ при $\varepsilon > 0$ невозможно [15].

Для некоторых постановок задачи о покрытии с неопределенностями во входных данных предложены подходы к решению. В [16] задача о покрытии рассматривается как задача целочисленного линейного программирования, где правая часть ограничений представляет собой бинарную случайную величину, т. е. заданы вероятности, с которыми должны быть покрыты элементы U . В работах [17–19] неопределенными считаются коэффициенты матрицы ограничений, т. е. считается неопределенной принадлежность элементов U множествам из S . В [20] рассматривается задача нахождения робастного решения задачи о покрытии множеств с интервальными весами.

В настоящей работе рассматривается задача о покрытии множества с интервальными весами, далее ее будем называть в сокращенном виде ИЗПМ, а веса, являющиеся точками из этих интервалов, — точечными. Оптимальное решение определяется нами как сильное оптимальное, т. е. как решение, оптимальное при любых возможных точечных весах множеств, и как слабое оптимальное, т. е. как решение, оптимальное при некоторых точечных весах множеств. В работе доказывается теорема о характеристизации сильного оптимального решения ИЗПМ.

Не для всякой ИЗПМ существует сильное оптимальное решение. В этой связи аналогично тому, как это делается в работах, посвященных задачам на графах с интервальными весами ребер (задаче о кратчайшем остовном дереве в графе [21], о наиболее длинном пути в ориентированном ациклическом графе [22]), нами вводится понятие объединенного оптимального решения задачи как такого набора слабых оптимальных решений, что при любых допустимых точечных весах множеств среди покрытий из этого набора найдется покрытие, являющееся оптимальным для этого случая.

В отличие от задач, решаемых в [21, 22], точечные варианты которых полиномиально разрешимы, задача о покрытии в случае точечных весов является NP -трудной. Нами вводится понятие объединенного оптимального решения ИЗПМ и предлагается интервальная модификация жадного алгоритма для нахождения ее приближенного решения.

1. Задача о покрытии множества с интервальными весами и ее оптимальные решения

Для обозначения интервалов используется жирный шрифт. Для интервала \mathbf{a} его нижняя и верхняя границы обозначаются соответственно как $\underline{\mathbf{a}}$ и $\overline{\mathbf{a}}$. Сумма интервалов \mathbf{a} и \mathbf{b} определяется как $\mathbf{a} + \mathbf{b} = [\underline{\mathbf{a}} + \underline{\mathbf{b}}, \overline{\mathbf{a}} + \overline{\mathbf{b}}]$.

В ИЗПМ заданы U , S и вектор $\mathbf{w} = (w_1, \dots, w_n)$ интервальных весов множеств из S , $\underline{w}_i > 0$. Определим, что мы будем считать слабым и сильным оптимальными решениями такой задачи.

Определение 1. *Реализацией весов множеств называется вектор точечных весов $w \in \mathbf{w}$.*

Реализация $w \in \mathbf{w}$ задает задачу о покрытии множества с теми же U и S , что и в ИЗПМ, и точечными весами множеств, задаваемыми w . Оптимальное решение для такой задачи будем называть *оптимальным решением для реализации w* .

Определение 2. *Слабое оптимальное решение ИЗПМ — это решение, являющееся оптимальным для некоторой реализации $w \in \mathbf{w}$.*

Определение 3. *Сильное оптимальное решение ИЗПМ — это решение, являющееся оптимальным для любой реализации $w \in \mathbf{w}$.*

Замечание 1. Любое сильное оптимальное решение является и слабым оптимальным решением.

Теорема (о характеристике сильного оптимального решения). *Решение ИЗПМ является сильным оптимальным решением тогда и только тогда, когда оно оптимально в следующей ситуации: веса входящих в него множеств принимают максимально допустимый вес, а веса всех других множеств — минимально допустимый.*

Доказательство. Сильное оптимальное решение оптимально для всех $w \in \mathbf{w}$, в том числе и для указанных в условии теоремы. Докажем достаточность выполнения этого условия для того, чтобы покрытие было сильным оптимальным решением задачи.

Если покрытие Opt оптимально в ситуации, указанной в условиях теоремы, то для произвольного покрытия Cvr будем иметь

$$\sum_{S_i \in Opt \setminus Cvr} \overline{w}_i + \sum_{S_i \in Opt \cap Cvr} \overline{w}_i \leq \sum_{S_i \in Cvr \setminus Opt} \underline{w}_i + \sum_{S_i \in Opt \cap Cvr} \overline{w}_i.$$

Для произвольных весов $w_i \in \mathbf{w}_i$ множеств $S_i \in Opt \cap Cvr$ будем иметь

$$\sum_{S_i \in Opt \setminus Cvr} \overline{w}_i + \sum_{S_i \in Opt \cap Cvr} w_i \leq \sum_{S_i \in Cvr \setminus Opt} \underline{w}_i + \sum_{S_i \in Opt \cap Cvr} w_i.$$

Поскольку

$$\sum_{S_i \in Opt \setminus Cvr} w_i \leq \sum_{S_i \in Opt \setminus Cvr} \overline{w}_i \quad \text{и} \quad \sum_{S_i \in Opt \setminus Cvr} \underline{w}_i \leq \sum_{S_i \in Opt \setminus Cvr} w_i,$$

имеем

$$\sum_{S_i \in Opt \setminus Cvr} w_i + \sum_{S_i \in Opt \cap Cvr} w_i \leq \sum_{S_i \in Cvr \setminus Opt} w_i + \sum_{S_i \in Opt \cap Cvr} w_i.$$

То есть $w(Opt) \leq w(Cvr)$ при всех возможных $w \in \mathbf{w}$, и значит Opt , — сильное оптимальное решение. \square

Согласно приведенной теореме некоторое слабое оптимальное решение ИЗПМ, получаемое с помощью какого-либо алгоритма, дающего точное решение для некоторой реализации весов, может быть проверено на предмет того, является ли оно сильным оптимальным решением.

Поскольку не для всякой ИЗПМ существует сильное оптимальное решение, введем понятие объединенного оптимального решения ИЗПМ как набора оптимальных решений, содержащего оптимальные решения для всех возможных реализаций $w \in \mathbf{w}$.

Определение 4. *Объединенным оптимальным решением ИЗПМ называется такой набор слабых оптимальных решений $\text{Opt} = \{\text{Opt}_1, \dots, \text{Opt}_k\}$, что для любой реализации $w \in \mathbf{w}$ найдется $\text{Opt}_i \in \text{Opt}$ такое, что Opt_i — оптимальное решение для этой реализации.*

Поскольку нахождение оптимального решения задачи о покрытии является NP -трудной задачей в случае точечных весов, и в интервальном случае нахождение и слабого, и сильного оптимальных решений также является NP -трудной задачей. Далее построим схему алгоритма нахождения приближенного решения ИЗПМ.

2. Жадный алгоритм решения задачи о покрытии множества с интервальными весами

2.1. Приближенные решения ИЗПМ

Рассмотрим жадный алгоритм решения задачи о покрытии для случая точечных весов множеств из S . В ходе его работы множества из S выбираются в покрытие Gr соответственно минимуму отношения $w_i/|S_i|$ — *относительного веса* множества S_i .

ЖАДНЫЙ АЛГОРИТМ РЕШЕНИЯ ЗАДАЧИ О ПОКРЫТИИ МНОЖЕСТВА

- 1 $Gr \leftarrow \emptyset$.
- 2 **if** Gr — покрытие U ,
- 3 **print** Gr .
- 4 **else** выбрать $S_q \in S$ такое, что $w_q/|S_q| = \min \left\{ w_i/|S_i| \mid S_i \in S \text{ и } S_i \not\subseteq \bigcup_{j \in Gr} S_j \right\}$,
- 5 $Gr \leftarrow Gr \cup \{S_q\}$,
- 6 $\forall i : S_i \leftarrow S_i \setminus S_q$.
- 7 Перейти на шаг 2.

Приведенную схему жадного алгоритма назовем *точечным жадным алгоритмом*. Далее нами будет построена схема *интервального жадного алгоритма* для решения ИЗПМ, который работает с интервальными относительными весами множеств. В ходе работы этого алгоритма производится перебор множеств реализаций из \mathbf{w} , дающих возможные варианты получения покрытия точечным жадным алгоритмом для всех $w \in \mathbf{w}$. Везде далее *приближенным решением* задачи о покрытии с точечными весами будем называть приближенное решение, получаемое с помощью точечного жадного алгоритма.

Определение 5. *Слабое приближенное решение ИЗПМ — это решение, которое дает точечный жадный алгоритм для некоторой реализации $w \in \mathbf{w}$.*

Каждое слабое приближенное решение ИЗПМ будем рассматривать как упорядоченное множество некоторых элементов из S . Через $\mathbf{w}[Gr]$ обозначим подмножество реализаций из \mathbf{w} , при которых набор множеств Gr может быть получен точечным жадным алгоритмом при некоторой нумерации множеств из S ; $\mathbf{w}[Gr] = (\mathbf{w}[Gr]_1, \dots, \mathbf{w}[Gr]_n)$.

Введем понятие объединенного приближенного решения ИЗПМ.

Определение 6. *Объединенным приближенным решением ИЗПМ называется такой набор покрытий $Gr = \{Gr_1, \dots, Gr_k\}$, при котором для любой реализации $w \in \mathbf{w}$ имеется $Gr_i \in Gr$ такое, что Gr_i — приближенное решение для w .*

2.2. Алгоритм нахождения объединенного приближенного решения

Представленный далее алгоритм нахождения объединенного приближенного решения ИЗПМ в качестве входа получает индивидуальную ИЗПМ и производит вызов процедуры, реализующей схему рекурсии с возвратом. В ходе рекурсивных вызовов производится перебор возможных вариантов построения слабых приближенных решений индивидуальной ИЗПМ, в результате чего строится объединенное приближенное решение Gr и для каждого $Gr \in Gr$ находится $\mathbf{w}[Gr]$.

Схему точечного жадного алгоритма можно представить как последовательность итераций следующего вида.

Итерация жадного алгоритма

Имеем задачу SCP .

1) выбираем S_q такое, что

$$w_q/|S_q| = \min \left\{ w_i/|S_i| \mid S_i \in S \text{ и } S_i \not\subset \bigcup_{j \in Gr} S_j \right\};$$

2) добавляем S_q в строимое покрытие Gr : $Gr \leftarrow Gr \cup \{S_q\}$;

3) получаем задачу SCP' : $U' \leftarrow U \setminus S_q$, $S'_j \leftarrow S_j \setminus S_q$, $w' \leftarrow w$.

Таким образом, в результате выполнения одной итерации жадного алгоритма, выбирая в Gr множество $S_q \in S$ такое, что

$$w_q/|S_q| = \min \left\{ w_i/|S_i| \mid S_i \in S \text{ и } S_i \not\subset \bigcup_{j \in Gr} S_j \right\}, \quad (2)$$

совершаем переход от задачи SCP с заданными U , S , w к задаче SCP' с U' , S' , w .

Пусть далее для индивидуальной ИЗПМ SCP U_{SCP} означает множество, для которого необходимо построить покрытие, S_{SCP} — набор подмножеств U_{SCP} , которые могут быть использованы для построения покрытия, вектор \mathbf{w}_{SCP} задает интервальные веса этих множеств. Для нахождения объединенного приближенного решения ИЗПМ будем выполнять итерации следующего вида.

Итерация жадного алгоритма решения ИЗПМИмеем ИЗПМ SCP .

- 1) находим множество $Q = \{S_{i_1}, \dots, S_{i_t}\}$ ($|Q| = t$), где $S_{i_j} \in Q$, если $\exists w \in \mathbf{w}_{SCP}$ такая, что

$$w_{i_j}/|S_{i_j}| = \min \left\{ w_i/|S_i| \mid S_i \in S_{SCP} \text{ и } S_i \not\subset \bigcup_{j \in Gr} S_j \right\};$$

- 2) получаем варианты построения Gr : $Gr \leftarrow Gr \cup \{S_{i_1}\}, \dots, Gr \leftarrow Gr \cup \{S_{i_t}\}$;
 3) получаем задачи $SCP^{(i_1)}, \dots, SCP^{(i_t)}$ с найденными для них $U_{SCP^{(i_j)}}, S_{SCP^{(i_j)}}, \mathbf{w}_{SCP^{(i_j)}}, j = \overline{1, t}$.

На каждой итерации жадного алгоритма решения ИЗПМ, имея SCP и определив множество Q , получаем набор ИЗПМ $SCP^{(i_1)}, \dots, SCP^{(i_t)}$, для каждой из которых будем выполнять итерации такого же вида, получая тем самым слабые приближенные решения из Gr для всех реализаций $w \in \mathbf{w}$.

Опишем процедуры, используемые алгоритмом. Процедура ВЫБОР получает на вход задачу SCP и в качестве результата дает множество Q . $S_i \in Q$ в том случае, если существует такая реализация $w \in \mathbf{w}_{SCP}$, что на множестве S_i достигается минимум относительного веса среди множеств из S_{SCP} .

ВЫБОР (SCP) : Q ;

- 1 **for** $\forall S_i \in S_{SCP}$:
- 2 $\mathbf{v}_i \leftarrow \mathbf{w}_i/|S_i|$;
- 3 $Q \leftarrow \emptyset$;
- 4 $v \leftarrow \min\{\bar{\mathbf{v}}_i \mid S_i \in S_{SCP}, S_i \neq \emptyset\}$;
- 5 **for** $\forall S_i \in S_{SCP}$:
- 6 **if** $\underline{\mathbf{v}}_i \leq v$
- 7 $Q \leftarrow Q \cup \{S_i\}$.

Здесь и далее \mathbf{v}_i означает интервал относительных весов для интервала возможных весов \mathbf{w}_i множества S_i : $\mathbf{v}_i = \mathbf{w}_i/|S_i|$. Пусть $S = \{S_1, S_2, S_3\}$ и интервалы относительных весов множеств из S соотносятся друг с другом так, как показано на рис. 1, где $\mathbf{v}_1 = \mathbf{w}_1/|S_1|$, $\mathbf{v}_2 = \mathbf{w}_2/|S_2|$ и $\mathbf{v}_3 = \mathbf{w}_3/|S_3|$. В ходе работы процедуры ВЫБОР в Q будут выбраны множества S_1 и S_2 , тогда как S_3 в Q выбрано не будет.

Процедура ВЕС ВЫБРАННОГО МНОЖЕСТВА получает на вход задачу SCP и индекс q одного из множеств из Q . В качестве результата выдается модифицированный вес этого множества $\mathbf{w}[Gr]_q$, т. е. вес \mathbf{w}_q модифицируется так, чтобы из \mathbf{w}_{SCP} исключить несовместные с выбором S_q в Q реализации. Иначе говоря, $\mathbf{w}[Gr]_q$ — интервал,

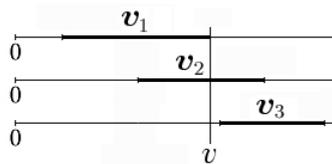


Рис. 1. Построение множества Q и модификация весов множеств из S_{SCP}

содержащий веса S_q , при которых S_q может быть выбрано в Gr точечным жадным алгоритмом, когда решается SCP при всех возможных точечных весах других множеств из S_{SCP} .

ВЕС ВЫБРАННОГО МНОЖЕСТВА $(SCP, Q, q) : \mathbf{w}[Gr]_q$;

- 1 $v \leftarrow \min\{\bar{\mathbf{w}}_i/|S_i| \mid S_i \in Q, i \neq q\}$;
- 2 **if** $\bar{\mathbf{w}}_q/|S_q| > v$
- 3 $\bar{\mathbf{w}}_q \leftarrow |S_q| \cdot v$.
- 4 $\mathbf{w}[Gr]_q \leftarrow \mathbf{w}_q$.

Для примера, изображенного на рис. 1, в случае, если в Gr выбирается S_2 , часть \mathbf{w}_2 , содержащая w_2 , большие, чем $|S_2| \cdot v$, не включается в $\mathbf{w}[Gr]_2$.

Процедура МОДИФИКАЦИЯ ЗАДАЧИ получает на вход ИЗПМ задачу SCP и индекс q множества $S_q \in Q$, выбираемого в одно из решений $Gr \in Gr$. Результатом работы процедуры является задача SCP' , для которой $U_{SCP'} = U_{SCP} \setminus S_q$, для $S_i \notin \bigcup_{j \in Gr} S_j$ полагаем $S'_i = S_i \setminus S_q$. Также производится модификация весов множеств $S_i \in S_{SCP}$ в соответствии с выбранным в Gr множеством S_q : из интервальных весов множеств исключаются те реализации, которые не совместны с таким выбором.

МОДИФИКАЦИЯ ЗАДАЧИ $(SCP, q) : SCP'$;

- 1 $S_{SCP'} \leftarrow \emptyset$;
- 2 **for** $\forall S_i \in S$
- 3 **if** $i \neq q$ и $S_i \neq \emptyset$
- 4 **if** $\underline{\mathbf{w}}_i/|S_i| < \underline{\mathbf{w}}_q/|S_q|$
- 5 $\mathbf{w}'_i \leftarrow |S_i| \cdot \underline{\mathbf{w}}_q/|S_q|$
- 6 **else** $\mathbf{w}'_i \leftarrow \mathbf{w}_i$;
- 7 $S'_i \leftarrow S_i \setminus S_q$;
- 8 $S_{SCP'} \leftarrow S_{SCP'} \cup \{S'_i\}$;
- 9 $U_{SCP'} \leftarrow U_{SCP} \setminus S_q$;
- 10 $\mathbf{w}_{SCP'} \leftarrow (\mathbf{w}'_1, \dots, \mathbf{w}'_n)$.

Отметим, что в ходе модификации ИЗПМ, производимой этой процедурой, среди множеств из $S_{SCP'}$ могут оказаться пустые. Эти множества не исключаются из $S_{SCP'}$ для того, чтобы не менялась нумерация множеств из S , заданная в исходной ИЗПМ, поданной на вход алгоритма.

В ситуации, представленной на рис. 1, при выборе в Gr множества S_2 из \mathbf{w}_1 будут исключены значения w_1 , меньшие, чем $|S_1| \cdot \underline{\mathbf{w}}_2/|S_2|$.

Общая схема алгоритма нахождения объединенного приближенного решения следующая:

НАХОЖДЕНИЕ ОБЪЕДИНЕННОГО ПРИБЛИЖЕННОГО РЕШЕНИЯ $(SCP) : Gr$;

- 1 $Gr \leftarrow \emptyset$; $Gr \leftarrow \emptyset$;
- 2 $Gr \leftarrow$ ОБЪЕДИНЕННОЕ ПРИБЛИЖЕННОЕ РЕШЕНИЕ (SCP, Gr, Gr) .

Здесь ОБЪЕДИНЕННОЕ ПРИБЛИЖЕННОЕ РЕШЕНИЕ — это рекурсивно вызываемая процедура, использующая описанные выше процедуры. При вызове этой процедуры для индивидуальной ИЗПМ SCP производится построение слабого приближенного решения Gr , включаемого в объединенное приближенное решение Gr . Gr и Gr — аргументы процедуры, изменяемые в ходе ее работы.

ОБЪЕДИНЕННОЕ ПРИБЛИЖЕННОЕ РЕШЕНИЕ (SCP, Gr, Gr);

```

1  if  $U = \emptyset$ 
2       $Gr$  — полученное слабое приближенное решение; сохранить  $\mathbf{w}[Gr]$ ;
3       $Gr \leftarrow Gr \cup \{Gr\}$ ;
4      return.
5  else
6       $Q \leftarrow$  ВЫБОР ( $SCP$ );
7      for  $\forall S_i \in Q$ :
8           $\tilde{Gr} \leftarrow Gr$ ;  $\tilde{\mathbf{w}} \leftarrow \mathbf{w}[Gr]$ ;

9           $Gr \leftarrow Gr \cup \{S_i\}$ ;
10          $\mathbf{w}[Gr]_i \leftarrow$  ВЕС ВЫБРАННОГО МНОЖЕСТВА ( $SCP, Q, i$ );
11          $SCP' \leftarrow$  МОДИФИКАЦИЯ ЗАДАЧИ ( $SCP, i$ );
12         ОБЪЕДИНЕННОЕ ПРИБЛИЖЕННОЕ РЕШЕНИЕ ( $SCP', Gr, Gr$ );

13          $Gr \leftarrow \tilde{Gr}$ ;  $\mathbf{w}[Gr] \leftarrow \tilde{\mathbf{w}}$ ;
14  print  $Gr$ .
```

Замечание 2. Представленный алгоритм является обобщением точечного жадного алгоритма. В случае, если все интервалы весов вырождены, т. е. имеем $\bar{\mathbf{w}}_i = \underline{\mathbf{w}}_i$ для всех $S_i \in S$, результатом его работы будут все покрытия, которые могут быть получены точечным жадным алгоритмом при рассмотрении всех возможных вариантов выбора множества в конструируемое покрытие, если минимум в (2) достигается на нескольких множествах из S .

2.3. Расчет вероятностей слабых приближенных решений ИЗПМ

Полагаем, что веса w_i множеств $S_i \in S$ — это случайные величины, для которых задано равномерное вероятностное распределение на интервалах \mathbf{w}_i .

Определение 7. Вероятностью слабого приближенного решения ИЗПМ назовем вероятность получения такой реализации $w \in \mathbf{w}$, при которой это решение является результатом работы точечного жадного алгоритма при весах, задаваемых w .

Через $P(Gr)$ обозначим вероятность слабого приближенного решения Gr . Если $Gr = \{S_{i_1}, \dots, S_{i_k}\}$, то, поскольку множества $S_{i_j} \in Gr$ перечисляются в порядке их включения в Gr , $P(Gr)$ может быть рассчитана как

$$P(Gr) = P(S_{i_1}) \cdot \dots \cdot P(S_{i_k}),$$

где $P(S_{i_j})$ — вероятность выбора S_{i_j} в Gr на j -й итерации точечного жадного алгоритма при последовательно выбранных ранее в Gr множествах $S_{i_1}, \dots, S_{i_{j-1}}$ и допустимых на момент выбора S_{i_j} в Gr весах множеств из S , не входящих в Gr перед выполнением j -й итерации.

Процедура ВЕРОЯТНОСТЬ ВЫБОРА получает на вход ИЗПМ SCP индекс q множества $S_q \in Q$, для которого рассчитывается вероятность получения реализации $w_q \in \mathbf{w}_q$, приводящей к выбору точечным жадным алгоритмом множества S_q в искомое слабое приближенное решение. Результат работы этой процедуры — значение вероятности $P(S_q)$ такой реализации. Для расчета $P(S_q)$ используются процедуры РАЗБИЕНИЕ и ВЕРОЯТНОСТЬ.

В ходе работы процедуры РАЗБИЕНИЕ строится разбиение P интервалов относительных весов множеств из Q , которое используется для нахождения вероятности выбора точечным жадным алгоритмом множества S_q на первой итерации решения задачи SCP при реализациях $w \in \mathbf{w}_{SCP}$.

РАЗБИЕНИЕ $(SCP, Q) : P$;

- 1 **for** $\forall S_i \in Q$:
- 2 $\mathbf{v}_i \leftarrow \mathbf{w}_i / |S_i|$;
- 3 $v_r \leftarrow \min\{\bar{\mathbf{w}}_i / |S_i| \mid S_i \in Q\}$;
- 4 **for** $\forall S_i \in Q$:
- 5 $\bar{\mathbf{v}}_i \leftarrow v_r$;
- 6 $M \leftarrow \{v_1, \dots, v_l\}$, где v_j такие, что $\exists S_i \in Q$, для которого $v_j = \mathbf{v}_i$ или $v_j = \bar{\mathbf{v}}_i$, $v_l = v_r$. M — упорядоченное по возрастанию его элементов множество.
- 7 **for** $\forall S_i \in Q$:
- 8 $M_i \leftarrow \{v_j \in M \mid v_j \in \mathbf{v}_i\}$,
 M_i — упорядоченное по возрастанию его элементов множество.
- 9 $P_i \leftarrow \{\mathbf{v}_{i1}, \dots, \mathbf{v}_{il}\}$, где $\mathbf{v}_{ik} = [v_j, v_{j+1}]$, $v_j, v_{j+1} \in M_i$.
- 10 $P \leftarrow \{P_1, \dots, P_{|Q|}\}$.

На рис. 2 проиллюстрировано построение разбиения P для набора множеств из Q , $|Q| = 3$, показаны три интервала возможных значений относительных весов для множеств из Q : \mathbf{v}_1 , \mathbf{v}_2 , \mathbf{v}_3 . В этом примере $v_r = \bar{\mathbf{v}}_1 = v_4$. Часть интервала \mathbf{v}_2 , лежащая правее значения v_4 , исключается из рассмотрения при построении разбиения $P = \{P_1, P_2, P_3\}$, поскольку реализации, принадлежащие этому интервалу при заданном \mathbf{w}_{SCP} , не совместны с выбором S_2 в Q : при любой из таких реализаций жадным алгоритмом будет выбрано либо множество S_1 с относительным весом из интервала \mathbf{v}_1 , либо S_3 с относительным весом из интервала \mathbf{v}_3 . Имеем $P_1 = \{\mathbf{v}_{11}, \mathbf{v}_{12}, \mathbf{v}_{13}\}$, где $\mathbf{v}_{11} = [v_1, v_2]$, $\mathbf{v}_{12} = [v_2, v_3]$, $\mathbf{v}_{13} = [v_3, v_4]$; $P_2 = \{\mathbf{v}_{21}\}$, где $\mathbf{v}_{21} = [v_3, v_4]$; $P_3 = \{\mathbf{v}_{31}, \mathbf{v}_{32}\}$, где $\mathbf{v}_{31} = [v_2, v_3]$, $\mathbf{v}_{32} = [v_3, v_4]$.

В ходе работы процедуры ВЕРОЯТНОСТЬ по найденному разбиению P для множества S_q рассчитывается вероятность $P(S_q)$ — вероятность выбора $S_q \in Q$ в конструируемое приближенное решение при заданных весах \mathbf{w}_{SCP} . Для случайного относительного веса $v_q = w_q / |S_q|$, соответствующего случайному весу w_q множества S_q ($w_q \in \mathbf{w}_q$, где \mathbf{w}_q — q -я компонента \mathbf{w}_{SCP}), имеем

$$P(S_q) = \sum_{\mathbf{v}_{qj} \in P_q} P(S_q \mid v_q \in \mathbf{v}_{qj}) \cdot P(v_q \in \mathbf{v}_{qj}),$$

где суммирование производится по всем интервалам из разбиения P_q интервала \mathbf{v}_q .

$$P(v_q \in \mathbf{v}_{qj}) = (\bar{\mathbf{v}}_{qj} - \underline{\mathbf{v}}_{qj}) / (\bar{\mathbf{v}}_q - \underline{\mathbf{v}}_q).$$

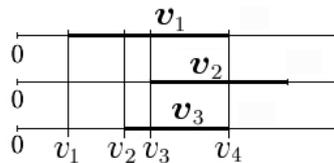


Рис. 2. Построение множества Q и модификация весов множеств из S_{SCP}

Пусть $R = \{i \mid S_i \in Q, i \neq q, \mathbf{v}_{qj} \in P_i\}$ — множество индексов из Q , не включая q , в разбиение интервалов относительных весов которых входит интервал \mathbf{v}_{qj} . Расчет вероятности $P(S_q \mid v_q \in \mathbf{v}_{qj})$ производится по формуле

$$P(S_q \mid v_q \in \mathbf{v}_{qj}) = \sum_{r \in R} \frac{1}{1 + |r|} P_r, \quad (3)$$

где r — подмножество из R , а P_r — вероятность выбора S_q при $v_i \in \mathbf{v}_{qj}$ для $i \in r$ и $v_i > \bar{\mathbf{v}}_{qj}$ для $i \notin r$. Обозначая через P_{ij} вероятность того, что $v_i \in \mathbf{v}_{qj}$, и через Q_{ij} вероятность того, что $v_i > \bar{\mathbf{v}}_{qj}$, имеем

$$P_{ij} = (\bar{\mathbf{v}}_{qj} - \mathbf{v}_{qj}) / (\bar{\mathbf{v}}_i - \mathbf{v}_i), \quad Q_{ij} = (\bar{\mathbf{v}}_i - \bar{\mathbf{v}}_{qj}) / (\bar{\mathbf{v}}_i - \mathbf{v}_i),$$

$$P_r = \prod_{i \in r} P_{ij} \prod_{i \in R \setminus r} Q_{ij},$$

где r принимает все возможные значения в (3) от $r = \emptyset$ до $r = R$. В том случае, если в (3) $r = \emptyset$ или $R \setminus r = \emptyset$, соответствующее произведение ($\prod_{i \in r} P_{ij}$ или $\prod_{i \in R \setminus r} Q_{ij}$) заменяется на 1. Множитель $1/(1 + |r|)$ в (3) обусловлен тем, что для N равномерно распределенных на одном и том же интервале случайных величин ξ_1, \dots, ξ_N для любого i имеем $P(\xi_i = \min\{\xi_1, \dots, \xi_N\}) = 1/N$. Таким образом, получаем следующие процедуры для расчета вероятности $P(S_q)$.

ВЕРОЯТНОСТЬ ВЫБОРА $(SCP, Q, q) : P(S_q)$;

- 1 $P \leftarrow$ РАЗБИЕНИЕ (SCP, Q) ;
- 2 $P(S_q) \leftarrow$ ВЕРОЯТНОСТЬ (SCP, P, Q, q) .

ВЕРОЯТНОСТЬ $(SCP, P, Q, q) : P(S_q)$;

- 1 **for** $j = 1$ **to** $|P_q|$
- 2 $P(v_q \in \mathbf{v}_{qj}) = (\bar{\mathbf{v}}_{qj} - \mathbf{v}_{qj}) / (\bar{\mathbf{v}}_q - \mathbf{v}_q)$;
- 3 $R \leftarrow \{i \in Q \mid \mathbf{v}_{qj} \in P_i\}$;
- 4 **for** $\forall i \in R$:
- 5 $P_{ij} \leftarrow (\bar{\mathbf{v}}_{qj} - \mathbf{v}_{qj}) / (\bar{\mathbf{v}}_i - \mathbf{v}_i)$;
- 6 **for** $\forall i \in R \setminus r$:
- 7 $Q_{ij} \leftarrow (\bar{\mathbf{v}}_i - \bar{\mathbf{v}}_{qj}) / (\bar{\mathbf{v}}_i - \mathbf{v}_i)$;
- 8 $P(S_q \mid v_q \in \mathbf{v}_{qj}) \leftarrow \sum_{r \in R} \left(\frac{1}{1 + |r|} \prod_{i \in r} P_{ij} \prod_{i \in R \setminus r} Q_{ij} \right)$;
- 9 $P(S_q) \leftarrow \sum_{j=1}^{|P_q|} P(S_q \mid v_q \in \mathbf{v}_{qj}) \cdot P(v_q \in \mathbf{v}_{qj})$.

В случае, если среди множеств из Q имеются множества с вырожденными интервальными весами, т.е. $\bar{\mathbf{w}}_q = \mathbf{w}_q$ для некоторых $S_q \in Q$, расчет $P(S_q)$ должен быть проведен по следующей формуле:

$$P(S_q) = \prod_{\substack{i \in Q \\ i \neq q}} [(\bar{\mathbf{v}}_i - v_q) / (\bar{\mathbf{v}}_i - \mathbf{v}_i)], \quad (4)$$

где $v_q = \bar{w}_q/|S_q| = \underline{w}_q/|S_q|$. Для других $i \in Q, i \neq q$, и таких, что \underline{w}_i — вырожденные, в произведении из (4) полагаем $(\bar{v}_i - v_q)/(\bar{v}_i - \underline{v}_i) = (\bar{v}_i - \underline{v}_i)/(\bar{v}_i - \underline{v}_i) = 1$.

Для того чтобы находить $P(Gr)$ для $Gr \in Gr$, необходимо произвести следующие модификации представленных выше процедур. На шаге 1 процедуры НАХОЖДЕНИЕ ОБЪЕДИНЕННОГО ПРИБЛИЖЕННОГО РЕШЕНИЯ производится инициализация: $P(Gr) \leftarrow 1$. Процедура ВЕРОЯТНОСТЬ ВЫБОРА вызывается перед выполнением процедуры МОДИФИКАЦИЯ ЗАДАЧИ в ходе представленной выше процедуры ОБЪЕДИНЕННОЕ ПРИБЛИЖЕННОЕ РЕШЕНИЕ. При включении в Gr множества S_i перед шагом 10 этой процедуры вычисляем текущее значение $P(Gr)$:

$$P(Gr) \leftarrow P(Gr) \cdot \text{ВЕРОЯТНОСТЬ ВЫБОРА}(SCP, Q, i).$$

Значение $P(Gr)$, рассчитанное до выбора в Gr множества S_i , должно сохраняться на шаге 8 процедуры ОБЪЕДИНЕННОЕ ПРИБЛИЖЕННОЕ РЕШЕНИЕ и восстанавливаться после выхода из нее на шаге 13 для того, чтобы рассчитывать вероятность других слабых приближенных решений, получаемых при выборе в строимое слабое приближенное решение $Gr \in Gr$ других множеств из Q . Итоговое значение должно быть сохранено перед выходом из этой процедуры на шаге 2.

2.4. Пример результата работы представленного алгоритма

Рассмотрим результат работы представленного алгоритма для нахождения объединенного приближенного решения ИЗМП со следующими входными данными:

$$U = \{1, 2, 3, 4, 5\};$$

$$S = \{S_1, S_2, S_3, S_4\}, \quad \text{где } S_1 = \{1, 2, 3\}, \quad S_2 = \{3, 4, 5\}, \quad S_3 = \{2, 5\}, \quad S_4 = \{2, 3, 4\};$$

$$\underline{w}_1 = [1, 3], \quad \underline{w}_2 = [2, 5], \quad \underline{w}_3 = [4, 6], \quad \underline{w}_4 = [1, 4].$$

Для этой задачи получаем

$$Gr = \{Gr_1, Gr_2, Gr_3, Gr_4, Gr_5, Gr_6, Gr_7\}.$$

Слабые приближенные решения $Gr_i \in Gr$ и множества реализаций $\underline{w}[Gr_i]$ представлены в табл. 1; здесь приведены также вероятности полученных слабых приближенных решений $P(Gr_i)$ (с точностью до сотых долей) и их суммарные интервальные веса, т. е. интервалы $\underline{w}(Gr_i)$ возможных весов получаемых покрытий. Вектор $\underline{w}[Gr_i]$ позволяет

Т а б л и ц а 1. Слабые приближенные решения, входящие в объединенное приближенное решение Gr

Gr_i	$\underline{w}[Gr_i]$	$P(Gr_i)$	$\underline{w}(Gr_i)$
$Gr_1 = \{S_1, S_2\}$	$[1, 3] \times [2, 5] \times [4, 6] \times [1, 4]$	0.46	$[3, 8]$
$Gr_2 = \{S_1, S_4, S_2\}$	$[1, 3] \times [2, 5] \times [4, 6] \times [1, 2.5]$	0.14	$[4, 10.5]$
$Gr_3 = \{S_1, S_4, S_3\}$	$[1, 3] \times [2, 5] \times [4, 5] \times [1, 2.5]$	0.01	$[6, 10.5]$
$Gr_4 = \{S_2, S_1\}$	$[2, 3] \times [2, 3] \times [4, 6] \times [1, 4]$	0.11	$[4, 6]$
$Gr_5 = \{S_4, S_1, S_2\}$	$[1, 3] \times [2, 5] \times [4, 6] \times [1, 3]$	0.23	$[6, 11]$
$Gr_6 = \{S_4, S_1, S_3\}$	$[1, 3] \times [2, 5] \times [4, 5] \times [1, 3]$	0.02	$[6, 11]$
$Gr_7 = \{S_4, S_2, S_1\}$	$[2, 3] \times [2, 3] \times [4, 6] \times [1, 3]$	0.02	$[5, 9]$

Т а б л и ц а 2. Модифицированная таблица слабых приближенных решений из Gr

Gr_i	$\mathbf{w}[Gr_i]$	$P(Gr_i)$	$\mathbf{w}(Gr_i)$
$Gr_1 = \{S_1, S_2\}$	$[1, 3] \times [2, 5] \times [4, 6] \times [1, 4]$	0.96	[3, 8]
$Gr_2 = \{S_1, S_4, S_3\}$	$[1, 3] \times [2, 5] \times [4, 5] \times [1, 3]$	0.03	[6, 11]

также получать наибольший возможный вес слабого приближенного решения Gr_i , его средний вес, если задано распределение на множестве реализаций \mathbf{w} , и другие характеристики решений.

Поскольку набор множеств, входящих в покрытие, получаемое точечным жадным алгоритмом, может содержать избыточные множества, а также зависит от порядка выбора в него множеств в том случае, когда минимум в (2) достигается на более чем одном множестве, в Gr могут содержаться избыточные покрытия, при удалении которых из Gr мы также получим объединенное приближенное решение, но с меньшим числом покрытий, рассматриваемых без учета порядка входящих в них множеств.

Пусть дана таблица указанного вида, содержащая слабые приближенные решения из некоторого объединенного приближенного решения Gr. Если $Gr_i, Gr_j \in Gr$ и $Gr_i \subset Gr_j$, то для любой $w \in \mathbf{w}$ будем иметь

$$w(Gr_j) = w(Gr_i) + w(Gr_j \setminus Gr_i).$$

Поскольку $Gr_j \setminus Gr_i \neq \emptyset$, а для любого $S_i \in S$ имеем $w(S_i) > 0$, для любого $w \in \mathbf{w}[Gr_j]$ имеем $w(Gr_j) > w(Gr_i)$, а следовательно, можно положить $Gr_k \leftarrow Gr_i$, Gr_k добавить в таблицу, а Gr_i и Gr_j удалить из нее. При этом полагаем

$$\mathbf{w}[Gr_k] \leftarrow \mathbf{w}[Gr_i] \cup \mathbf{w}[Gr_j], \quad P(Gr_k) \leftarrow P(Gr_i) + P(Gr_j),$$

после чего производим перенумерацию решений.

В случае, если $Gr_i, Gr_j \in Gr$ отличаются только порядком элементов, то можно положить $Gr_k \leftarrow Gr_i$, Gr_k добавить в таблицу, а Gr_i и Gr_j удалить из нее. При этом полагаем

$$\mathbf{w}[Gr_k] \leftarrow \mathbf{w}[Gr_i] \cup \mathbf{w}[Gr_j], \quad P(Gr_k) \leftarrow P(Gr_i) + P(Gr_j),$$

после чего производим перенумерацию решений.

Пример такой модификации полученного объединенного приближенного решения представлен в табл. 2. Для рассмотренной ИЗПМ при заданном равномерном распределении на интервалах весов получены слабые приближенные интервальные решения: Gr_1 с вероятностью его получения 0.96 и Gr_2 с вероятностью получения 0.03.

2.5. Вычислительная сложность и точность представленного алгоритма

Представленный алгоритм обладает в общем случае экспоненциальной вычислительной сложностью. Вычислительная сложность алгоритма определяется:

- 1) количеством вариантов ветвления, производимого в ходе рекурсивных вызовов процедуры ОБЪЕДИНЕННОЕ ПРИБЛИЖЕННОЕ РЕШЕНИЕ, т. е. значениями $|Q|$, получаемыми в результате выполнения процедуры ВЫБОР;
- 2) вычислительной сложностью выполнения процедуры ВЕРОЯТНОСТЬ, в ходе которой производится перебор подмножеств r в формуле (3), состоящих из элементов множеств $R \subset Q$, что требует $O(2^{|Q|-1})$ элементарных машинных операций.

Вычислительная сложность представленного алгоритма нахождения объединенного приближенного решения ИЗПМ может быть снижена, если поставить задачу нахождения не всего объединенного приближенного решения, а только наиболее вероятных слабых приближенных решений, входящих в него. Для этого, задавая пороговое значение P_{\min} , будем осуществлять выход из процедуры ОБЪЕДИНЕННОЕ ПРИБЛИЖЕННОЕ РЕШЕНИЕ, если в процессе работы алгоритма на какой-то его итерации получаем $P(Gr) < P_{\min}$.

Для объединенного приближенного решения Gr задачи SCP , получаемого с помощью интервального жадного алгоритма, имеем: для любого $w \in \mathbf{w}_{SCP}$ существует $Gr \in Gr$, что Gr будет построено точечным жадным алгоритмом для точечной задачи о покрытии с весами w и U_{SCP} , S_{SCP} , а значит, для него будет справедлива логарифмическая априорная оценка точности (1) для весов, задаваемых w .

Таким образом, рассмотрена задача о покрытии множества с интервальными весами. Предложены варианты определения сильного и слабого оптимального решения этой задачи. Доказана теорема о характеристизации сильного оптимального решения. Введено понятие объединенного оптимального решения задачи и приведен приближенный алгоритм его нахождения, представляющий собой модификацию жадного алгоритма решения задачи о покрытии множества с точечными весами подмножеств.

Благодарности. Автор благодарит участников семинара по интервальному анализу и его приложениям С.И. Жилина, С.И. Кумкова, Д.Ю. Надежина и С.П. Шарого за высказанные ими замечания и пожелания.

Список литературы / References

- [1] **Карп, Р.М.** Reducibility among combinatorial problems // 50 years of integer programming 1958–2008: Lectures of the Pioneers of Integer Programming. Berlin, Heidelberg: Springer-Verlag, 2010. P. 219–241.
- [2] **Farahani, R.Z., Asgari, N., Heidari, N., Hosseini, M., Goh, M.** Covering problems in facility location: A review // Computers and Industrial Engineering. 2012. No. 62(1). P. 368–407.
- [3] **Kearns, M.J.** The computational complexity of machine learning. Cambridge: MIT Press, 1990. 176 p.
- [4] **Milena, M.** Set cover with requirements and costs evolving over time // Randomization, Approximation, and Combinatorial Optimization. Algorithms and Techniques: Lecture Notes in Computer Science. Vol. 1671. Berlin, Heidelberg: Springer, 1999. P. 63–72.
- [5] **Gao, B.J., Ester, M., Cai, J.-Y., Schulte, O., Xiong, H.** The minimum consistent subset cover problem and its applications in data mining // Proc. of the 13th ACM SIGKDD Intern. Conf. on Knowledge Discovery and Data Mining, KDD '07. New York, USA, 2007. P. 310–319.
- [6] **Закревский А.Д.** Логический синтез каскадных схем. М.: Наука, 1981. 416 с.
Zakrevsky, A.D. Logical synthesis of cascade schemes. Moscow: Nauka, 1981. 416 p. (in Russ.)
- [7] **Герман О.В., Ефремов О.В.** Алгоритм решения обобщенной задачи о покрытии // Экономика и матем. методы. 1998. Т. 34, вып. 4. С. 134–140.
German, O.V., Efremov, O.V. An algorithm for the generalized set cover problem // Economics and Mathematical Methods. 1998. Vol. 34, iss. 4. P. 134–140. (in Russ.)
- [8] **Balas, E., Carrera, M.C.** A dynamic subgradient-based branch and bound procedure for set covering // Operational Research. 1996. Vol. 44(6). P. 875–890.

- [9] **Balas, E., Ho, A.** Set covering algorithms using cutting planes, heuristics, and subgradient optimization: a computational study // *Mathematical Programming*. 1980. Vol. 12. P. 37–60.
- [10] **Garfinkel, R.S., Nemhauser, G.L.** *Integer programming*. New York: Wiley, 1972. 427 p.
- [11] **Eremeev, A.V., Kolokolov, A.A., Zaozerskaya, L.A.** A hybrid algorithm for set covering problem // *Proc. of Intern. Workshop on Discrete Optimization Methods in Sheduling and Computer-aided Design*. Minsk, 2000. P. 123–129.
- [12] **Gomes, F.C., Meneses, C.N., Pardalos, P.M., Viana, G.V.R.** Experimental analysis of approximation algorithms for the vertex cover and set covering problems // *Computers and Operations Research*. 2006. Vol. 33(12). P. 3520–3534.
- [13] **Beasley, J.E.** An algorithm for set covering problem // *European Journal of Operational Research*. 1987. Vol. 31. P. 85–93.
- [14] **Chvatal, V.** A greedy heuristic for the set-covering problem // *Mathematics of Operation Research*. 1979. Vol. 4(3). P. 233–235.
- [15] **Feige, U.** A threshold of $\ln n$ for approximating set cover // *Journal of the ACM*. 1998. Vol. 45(4). P. 634–652.
- [16] **Beraldi, P., Ruszczynski, A.** The probabilistic set-covering problem // *Operations Research*. 2002. Vol. 50(6). P. 956–967.
- [17] **Hwang, M.J., Chiang, C.I., Liu, Y.H.** Solving a fuzzy set-covering problem // *Mathematical and Computer Modelling*. 2004. Vol. 40(7). P. 861–865.
- [18] **Chiang, C.I., Hwang, M.J., Liu, Y.H.** An alternative formulation for certain fuzzy set-covering problems // *Mathematical and Computer Modelling*. 2005. Vol. 42(3–4). P. 363–365.
- [19] **Fischetti, M., Monaci, M.** Cutting plane versus compact formulations for uncertain (integer) linear programs // *Mathematical Programming Computation*. 2012. Vol. 4(3). P. 239–273.
- [20] **Pereira, P., Averbakh, I.** The robust set covering problem with interval data // *Annals of Operations Research*. 2011. Vol. 207(1). P. 1–19.
- [21] **Yaman, H., Karasan, O.E., Pinar, M.C.** Longest path problem with interval data: Technical Report 9908, Department of Industrial Engineering, Bilkent University, 06533 Bilkent, Ankara, Turkey, 1999.
- [22] **Yaman, H., Karasan, O.E., Pinar, M.C.** The robust minimum spanning tree problem with interval data // *Operations Research Letters*. 2001. Vol. 29(1). P. 31–40.

*Поступила в редакцию 6 июля 2015 г.,
с доработки — 26 августа 2015 г.*

The set cover problem with interval weights and the greedy algorithm for its solution

PROLUBNIKOV, ALEXANDER V.

Omsk F.M. Dostoevsky State University, Omsk, 644077, Russia

Corresponding author: Prolubnikov, Alexander V., e-mail: a.v.prolubnikov@mail.ru

Purpose: We study the set cover problem when a set of weighted subsets of a set U is supposed to be given. Cover set of U is a collection of subsets of that set when the union of members of the set contains U . We need to choose a cover that contains sets which have minimum total weight. The specificity of our work is that we investigate the problem with interval weights of the subsets.

Methodology: We introduce definitions of optimal and approximate solutions of the problem. Realization of the weights of given subsets is a possible variant of these weights. The weak optimal solution of the problem is a solution that is optimal for some realization. The strong optimal solution is a solution that is optimal for any realization. The united optimal solution is such a set of weak solutions that it contains an optimal solution for any possible realization. The united approximate solution is such a set of covers that contains approximate solutions for every realization.

Findings: The theorem of characterization of the strong optimal solution is proved. The modification of the greedy algorithm for the problem is presented. The algorithm gives the united approximate solution as a result. For the uniform distributions of weights, we present a modification of the algorithm that could be used to find probabilities of the approximate solutions.

Originality/value: The concepts of optimal and approximate solutions are presented for the set cover problem with interval weights. The theorem of characterization of the strong optimal solution is proved. The modification of the greedy algorithm for the set cover problem with interval weights is presented.

Keywords: set cover problem, interval uncertainty, greedy algorithm.

Received 6 July 2015

Received in revised form 26 August 2015