

ISSN 2949-5598

ДИСКРЕТНЫЙ АНАЛИЗ И ИССЛЕДОВАНИЕ ОПЕРАЦИЙ

Том 32 № 1 2025

Новосибирск
Издательство Института математики

РЕДАКЦИОННАЯ КОЛЛЕГИЯ

Главный редактор **В. Л. Береснев**
Зам. главного редактора **А. А. Евдокимов**
Ответственный секретарь **Ю. В. Шамардин**

С. В. Августинovich	М. Я. Ковалёв	А. В. Пяткин
Г. П. Агибалов	А. В. Кононов	А. А. Сапоженко
В. Б. Алексеев	А. В. Косточка	М. Свириденко
О. В. Бородин	В. В. Кочергин	Б. Я. Рябко
В. А. Васильев	Ю. А. Кочетов	Н. Н. Токарева
Э. Х. Гимади	В. К. Леонтьев	Ю. А. Флеров
А. Ю. Григорьев	Б. М.-Т. Лин	Ф. В. Фомин
С. Демпе	В. В. Лозин	М. Ю. Хачай
А. И. Ерзин	П. Пардалос	Я. М. Шафранский

Учредители Сибирское отделение РАН
журнала Институт математики им. С. Л. Соболева СО РАН

Журнал включён в базу данных Russian Science Citation Index (RSCI) на платформе Web of Science. Переводы статей на английский язык публикуются в *Journal of Applied and Industrial Mathematics* и доступны по ссылке www.springer.com/mathematics/journal/11754.

СИБИРСКОЕ ОТДЕЛЕНИЕ РОССИЙСКОЙ АКАДЕМИИ НАУК
ИНСТИТУТ МАТЕМАТИКИ им. С. Л. СОБОЛЕВА СО РАН

ДИСКРЕТНЫЙ АНАЛИЗ И ИССЛЕДОВАНИЕ ОПЕРАЦИЙ

Выпускается с 1994 г.	Научный журнал	4 номера в год
Том 32, № 1 (163)	Январь–март 2025	

СОДЕРЖАНИЕ

Борисова И. А. Вычислительная сложность задачи выбора типичных представителей в конечном множестве точек метрического пространства	5
Глебов А. Н., Добрынин А. А. Универсальные циклы, порождающие все графы коалиционных разбиений циклов	16
Гусев Д. А., Свиридова О. А., Шидловский И. Г., Бродецкий Г. Л. Оптимизация стратегии поставок заказов при управлении запасами по многим критериям в условиях неопределённости	28
Зырянов А. О., Лавлинский С. М., Панин А. А., Плясунов А. В. Государственно-частное партнёрство в инфраструктурных проектах сырьевой территории: модель на основе консорциума недропользователей	48
Моторин К. О., Пяткин А. В. Об одной задаче оптимизации размещения товаров на складе	75
Ратушный А. В., Кочетов Ю. А. Двухстадийный алгоритм для динамической задачи упаковки в контейнеры с группами размещения	99
Талецкий Д. С. О числе вечного доминирования планарных графов диаметра 2	122
Юськов А. Д. Метод декомпозиции для задачи размещения антенн на стадионе	145

НОВОСИБИРСК
ИЗДАТЕЛЬСТВО ИНСТИТУТА МАТЕМАТИКИ

В журнале публикуются оригинальные научные статьи и обзоры теоретической и прикладной направленности по следующим разделам дискретного анализа, исследования операций и информатики:

- дискретная оптимизация
- комбинаторика
- контроль и надёжность дискретных устройств
- математические модели и методы принятия решений
- математическое программирование
- модели экономики
- моделирование процессов управления
- построение и анализ алгоритмов
- синтез и сложность управляющих систем
- теория автоматов
- теория графов
- теория игр и её приложения
- теория кодирования
- теория расписаний и размещений

Адрес редакции:

Институт математики
им. С. Л. Соболева СО РАН,
пр. Академика Коптюга, 4,
630090 Новосибирск, Россия
Телефон: +7 (383) 329–75–79
E-mail: discopr@math.nsc.ru

© Сибирское отделение РАН, 2025

© Институт математики им. С. Л. Соболева СО РАН, 2025

SIBERIAN BRANCH OF THE RUSSIAN ACADEMY OF SCIENCES
SOBOLEV INSTITUTE OF MATHEMATICS

DISKRETNYYI ANALIZ I ISSLEDOVANIE OPERATSII

/DISCRETE ANALYSIS AND OPERATIONS RESEARCH/

Published since 1994	Scientific journal	4 issues per year
Vol. 32, No. 1 (163)	January–March, 2025	

CONTENTS

I. A. Borisova. <i>Computational complexity of the choice problem for typical representatives of a finite point set in a metric space</i>	5
A. N. Glebov and A. A. Dobrynin. <i>Universal cycles that generate all graphs of coalition partitions in cycles.</i>	16
D. A. Gusev, O. A. Sviridova, I. G. Shidlovskii, and G. L. Brodetskiy. <i>Optimization of inventory management strategies for order deliveries using multicriteria decision making under conditions of uncertainty.</i>	28
A. O. Zyryanov, S. M. Lavlinskii, A. A. Panin, and A. V. Plyasunov. <i>Public-private partnership concerning infrastructure projects in a resource region: A model based on a consortium of subsoil users.</i>	48
K. O. Motorin and A. V. Pyatkin. <i>On one optimization problem for warehouse goods placement.</i>	75
A. V. Ratushnyi and Yu. A. Kochetov. <i>A two-stage algorithm for the dynamic bin packing problem with placement groups</i>	99
D. S. Taletskii. <i>On the eternal domination number of planar graphs with diameter 2.</i>	122
A. D. Yuskov. <i>A decomposition approach to a stadium antenna deployment problem</i>	145

NOVOSIBIRSK
SOBOLEV INSTITUTE PRESS

In this journal we publish original research papers and survey papers of both theoretical and practical importance on the following topics of discrete analysis, operations research and informatics:

- discrete optimization
- combinatorics
- control and reliability of discrete devices
- decision making models and methods
- mathematical programming
- economic models
- management modeling
- design and analysis of algorithms
- synthesis and complexity of control systems
- automata theory
- graph theory
- game theory and its applications
- coding theory
- theory of scheduling and facility location

Editorial office address:

Sobolev Institute of Mathematics,

4 Acad. Koptug Avenue,

630090 Novosibirsk, Russia

Phone: +7 (383) 329-75-79

E-mail: discopr@math.nsc.ru

© Siberian Branch of RAS, 2025

© Sobolev Institute of Mathematics SB RAS, 2025

ВЫЧИСЛИТЕЛЬНАЯ СЛОЖНОСТЬ ЗАДАЧИ ВЫБОРА ТИПИЧНЫХ ПРЕДСТАВИТЕЛЕЙ В КОНЕЧНОМ МНОЖЕСТВЕ ТОЧЕК МЕТРИЧЕСКОГО ПРОСТРАНСТВА

И. А. Борисова^{1,2}

¹ Институт математики им. С. Л. Соболева,
пр. Акад. Коптюга, 4, 630090 Новосибирск, Россия

² Новосибирский гос. университет,
ул. Пирогова, 2, 630090 Новосибирск, Россия

E-mail: biamia@mail.ru

Аннотация. Исследуется сложностной статус одной экстремальной задачи выбора подмножества p наиболее типичных точек в заданном конечном множестве точек метрического пространства. При этом требуется, чтобы выбранное подмножество наилучшим образом описывало исходное множество с точки зрения некоторого геометрического критерия. Рассматриваемая задача является формализацией одной прикладной проблемы из анализа данных, заключающейся в отыскании подмножества типичных представителей выборки с опорой на функцию конкурентного сходства. В статье доказывается, что рассматриваемая задача NP-трудна. Для этого к ней полиномиально сводится одна из NP-трудных задач — задача о трёхмерном сочетании. Ил. 1, библиогр. 14.

Ключевые слова: NP-трудная задача, выбор типичных представителей, функция конкурентного сходства, задача о трёхмерном сочетании, машинное обучение.

Введение

Несмотря на то, что современное общество всё чаще требует от науки сосредоточить усилия на решении актуальных прикладных задач, подобная прагматизация не должна ставить крест на фундаментальных исследованиях, а потому одним из важных вызовов для современной прикладной математики становится формализация и детальное изучение задач, порождаемых практикой.

В данной работе предлагается формальная постановка и устанавливается сложностной статус одной такой задачи, решаемой в процессе

интеллектуального анализа данных, задачи выбора типичных объектов из неклассифицированной выборки, максимально сохраняющих информацию о ней, с точки зрения некоторого формального критерия. Содержательно эта задача близка к задаче кластеризации, в которой множество объектов выборки разбивается на кластеры, состоящие из объектов, одновременно максимально похожих на центральные элементы этих кластеров и непохожих на центральные элементы соседних кластеров.

1. Формулировка задачи, её интерпретация и истоки

Задача выбора типичных объектов из неклассифицированной выборки в метрическом пространстве возникает в распознавании образов и машинном обучении при анализе выборок, описание которых необходимо «сжать», переходя от всего множества объектов к кластерам объектов, представленных наиболее типичными их представителями. Так, при анализе профилей клиентской активности сотовые операторы формируют кластеры клиентов со схожими потребностями и создают новые тарифы для типичных клиентов из каждого кластера. Принципы, в соответствии с которыми осуществляется отбор типичных представителей при машинном обучении, могут меняться в зависимости от того, как формализуется понятие типичного объекта.

В рассматриваемой постановке типичными считаются те объекты выборки, на которые максимально похожи объекты из того же кластера и непохожи объекты других кластеров. В качестве меры похожести-непохожести для объектов, представленных в виде точек многомерного метрического пространства, используется функция конкурентного сходства, или FRiS-функция (function of rival similarity) [1], которая позволяет вычислять величину конкурентного сходства на основе расстояний между объектами.

Для множества $\{f_i\}_{i=1}^n$ конечной мощности n через $\min'_{i \in \overline{1, n}} f_i$ обозначим его второй минимальный элемент, т. е. если $f_{i_1} \leq f_{i_2} \leq \dots \leq f_{i_n}$, то $\min'_{i \in \overline{1, n}} f_i = f_{i_2}$ для $n \geq 2$ и $\min'_{i \in \overline{1, n}} f_i = f_{i_1}$ для $n = 1$. В этих обозначениях имеем следующую постановку.

Задача 1. Даны конечное множество точек Λ в метрическом пространстве размерности d и натуральное p . Требуется найти непустое подмножество $S_p \subseteq \Lambda$, которое максимизирует целевую функцию

$$F(S_p) = \sum_{z \in \Lambda} \frac{\min'_{s_2 \in S_p} r(z, s_2) - \min_{s_1 \in S_p} r(z, s_1)}{\min'_{s_2 \in S_p} r(z, s_2) + \min_{s_1 \in S_p} r(z, s_1)}, \quad (1)$$

где $r(a, b)$ — расстояние между точками a и b , при ограничении $|S_p| = p$.

Исследования сложностного статуса различных задач машинного обучения, использующих в процессе анализа данных функцию конкурентного сходства, ведутся уже несколько лет [2–4].

Наиболее близкая постановка данной задачи встречается в работе [4], в которой принцип выбора подмножества типичных объектов также основан на использовании функции конкурентного сходства, однако в ней автор рассматривает редуцированный критерий, из которого исключён компонент, оценивающий непохожесть объектов на типичных представителей чужих кластеров.

Задача выбора подмножества типичных объектов из конечного множества близка к таким известным задачам кластеризации, как k -means (или k -MSSC) [5, 6] и k -median [7] в их дискретной постановке [8, 9], когда центры кластеров отыскиваются не среди всех точек метрического пространства, а только среди тех, которые включены в исходное множество точек. Похожие постановки есть в [10, 11], где в процессе кластеризации отыскиваются наиболее типичные элементы, в то время как элементы-выбросы в формируемые кластеры не включаются. Несколько вариантов постановки задачи выбора подмножества точек из заданного множества для различных геометрических критериев представлены в [12]. В [2, 13] также рассматривается задача поиска типичных представителей. В этой задаче, в отличие от всех вышеперечисленных, которые можно интерпретировать как поиск центров неизвестных кластеров, разбиение множества точек на два кластера задаётся изначально, причём каждый кластер может описываться несколькими точками, вместо одного центра.

2. Описание математической модели

Теперь поподробнее рассмотрим задачу анализа данных, моделируемую в настоящей статье. Предположим, что для каждой тройки объектов $x, y, z \in A$ задан способ вычислять конкурентное сходство $f(z, x, y)$ таким образом, что чем больше z похож на объект x и непохож на объект y , тем величина $f(z, x, y)$ больше. Тогда типичными с точки зрения конкурентного сходства считаются объекты, у которых сходство с объектами своего кластера как можно выше, а с объектами чужого кластера — как можно ниже.

Если разбиение на кластеры $\{A_i\}_{i=1}^p$ задано изначально, так что

$$A = \bigcup_{i=1}^p A_i, \quad \text{где } A_i \cap A_j = \emptyset \text{ для } i \neq j,$$

то задача выбора подмножества p типичных объектов из этой выборки заключается в отыскании такого набора элементов $B = \{b_i\}_{i=1}^p$, $b_i \in A_i$,

которые обеспечат максимальное сходство всех объектов выборки с типичными объектами своего кластера при минимальном сходстве с типичными объектами другого кластера:

$$F_c(B) = \sum_{i=1}^p \sum_{z \in A_i} \min_{\substack{y \in B, \\ y \neq b_i}} f(z, b_i, y) \rightarrow \max_{\substack{B = \{b_i\}_{i=1}^p, \\ b_i \in A_i}}.$$

Если же разбиение на кластеры не задано, то из требования максимального сходства объектов с типичным представителем своего кластера следует, что своим считается тот кластер, сходство с типичным объектом которого максимально. Тогда задача записывается следующим образом:

$$F_{nc}(B) = \sum_{z \in A} \max_{b \in B} \min_{\substack{y \in B, \\ y \neq b}} f(z, b, y) \rightarrow \max_{\substack{B \subset A, \\ |B|=p}}. \quad (2)$$

В большинстве случаев при решении прикладных задач каждый объект из множества A рассматривается как точка в пространстве описывающих характеристик, на котором определена некоторая метрика r . Это позволяет вычислять расстояние $r(x, y)$ между парой объектов $x, y \in A$. В рассматриваемой постановке в качестве величины конкурентного сходства $f(z, x, y)$ используется FRiS-функция [1], определяемая по следующей формуле:

$$\text{FRiS}(z, x, y) = \frac{r(z, y) - r(z, x)}{r(z, y) + r(z, x)}.$$

Если подставить формулу для вычисления FRiS-функции в максимизируемый функционал задачи (2), то его слагаемые преобразуются следующим образом:

$$F(B) = \sum_{z \in A_i} \max_{b \in B} \min_{\substack{y \in B, \\ y \neq b}} \frac{r(z, y) - r(z, b)}{r(z, y) + r(z, b)} = \sum_{i=1}^p \sum_{z \in A_i} \frac{\min_{\substack{y \in B, \\ y \neq y_z}} r(z, y) - \min_{b \in B} r(z, b)}{\min_{\substack{y \in B, \\ y \neq y_z}} r(z, y) + \min_{b \in B} r(z, b)},$$

где $y_z = \arg \min_{b \in B} r(z, b)$. Для получения критерия (1) из задачи 1 осталось заметить, что

$$\min_{\substack{y \in B, \\ y \neq y_z}} r(z, y) = \min'_{y \in B} r(z, y).$$

3. Вычислительная сложность задачи выбора подмножества типичных объектов

Для установления сложностного статуса задачи 1 покажем, что к ней сводится задача о трёхмерном сочетании (3D-matching) в следующей постановке.

Задача 3DM (о трёхмерном сочетании). Даны три конечных множества $X = \{x_i\}_{i=1}^p$, $Y = \{y_i\}_{i=1}^p$, $Z = \{z_i\}_{i=1}^p$, а также множество троек $T = \{t_i\}_{i=1}^M \subset X \times Y \times Z$. Требуется ответить, существует ли такое подмножество троек $S \subseteq T$, $|S| = p$, которое удовлетворяет следующему условию: для любой пары различных троек $s_1 = (x_1, y_1, z_1)$, $s_2 = (x_2, y_2, z_2) \in S$ выполняются неравенства $x_1 \neq x_2$, $y_1 \neq y_2$ и $z_1 \neq z_2$.

Задача о трёхмерном сочетании является одной из NP-полных задач из списка Карпа [14].

Теорема 1. Задача о трёхмерном сочетании полиномиально сводится к задаче 1.

ДОКАЗАТЕЛЬСТВО. Рассмотрим задачу 1 для случая, когда исходное множество точек Λ в метрическом пространстве состоит из следующих групп: множеств $\tilde{X} = \{\tilde{x}_i\}_{i=1}^p$, $\tilde{Y} = \{\tilde{y}_i\}_{i=1}^p$ и $\tilde{Z} = \{\tilde{z}_i\}_{i=1}^p$, множества $\tilde{T} = \{\tilde{t}_i\}_{i=1}^M$ и набора множеств $B_1, B_2, \dots, B_M$ таких, что $|B_1| = |B_2| = \dots = |B_M| = N > M + 3$. При этом расстояния между точками задаются следующим образом:

- $r(\tilde{t}_i, \tilde{x}_j) = r(\tilde{t}_i, \tilde{y}_k) = r(\tilde{t}_i, \tilde{z}_l) = 0,5$, если $\tilde{t}_i = (\tilde{x}_j, \tilde{y}_k, \tilde{z}_l) \in \tilde{T}$;
- $r(\tilde{t}_i, b) = 0,5$ для $\tilde{t}_i \in \tilde{T}$ и всех $b \in B_i$;
- $r(x, y) = 1$ для любой другой пары точек $x, y \in \Lambda$.

В этом случае матрица попарных расстояний удовлетворяет аксиомам метрического пространства, в том числе и аксиоме треугольника.

На рис. 1 проиллюстрирован внешний вид полученного множества точек Λ . Ребро между двумя точками означает, что расстояние между ними равно 0,5.

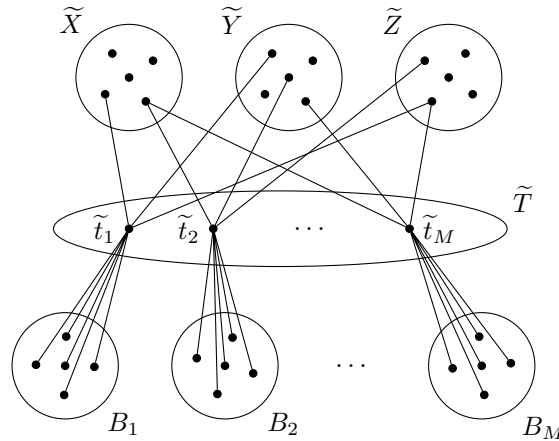


Рис. 1. Структура множества Λ из теоремы 1

Покажем, что только $S \subseteq \tilde{T}$ может быть решением задачи 1. Докажем это для тривиального случая, когда $p = 1$. При этом по определению $\min'_{s_2 \in S} r(z, s_2) = 1$. Если $S_1 = \{b\}$, где $b \in B_i$, то расстояние от b только до одного t_i равно 0,5, поэтому

$$F(\{b\}) = \sum_{z \in \Lambda} \frac{1 - r(z, b)}{1 + r(z, b)} = \frac{1}{3} + 1 = 1\frac{1}{3}.$$

Если $S_1 = \{x\}$, где $x \in \tilde{X}$, то только расстояния от x до l троек из множества $\tilde{T}$, содержащих x , равны 0,5, поэтому

$$F(\{x\}) = \sum_{z \in \Lambda} \frac{1 - r(z, x)}{1 + r(z, x)} = l \cdot \frac{1}{3} + 1 \leq M \cdot \frac{1}{3} + 1.$$

Если $S_1 = \{y\}$, где $y \in \tilde{Y}$, то по аналогии с предыдущим случаем

$$F(\{y\}) \leq M \cdot \frac{1}{3} + 1.$$

Если $S_1 = \{z\}$, где $z \in \tilde{Z}$, то по аналогии с предыдущим случаем

$$F(\{z\}) \leq M \cdot \frac{1}{3} + 1.$$

Если $S_1 = \{t\}$, где $t = (x, y, z) \in \tilde{T}$, то только расстояния от t до N точек связанного с ним множества B_i и до x, y, z равны 0,5, поэтому

$$F(\{t\}) = \sum_{z \in \Lambda} \frac{1 - r(z, t)}{1 + r(z, t)} = (N + 3) \cdot \frac{1}{3} + 1 > (M + 3 + 3) \cdot \frac{1}{3} + 1 = M \cdot \frac{1}{3} + 3.$$

Таким образом, максимум функционала достигается на множестве столпов S_1 , состоящем из элемента $t \in \tilde{T}$.

Теперь докажем, что при дальнейшем увеличении числа элементов в S по-прежнему максимальное значение функционала качества будет достигаться на элементах из $\tilde{T}$.

Пусть уже имеется множество столпов $S_{k-1} \subset \tilde{T}$, являющееся решением задачи 1 для $p = k - 1$. Оценим, как меняется качество при добавлении в S_{k-1} элементов из разных подмножеств Λ . Для $z \in \Lambda$ введём обозначение

$$F(z, S) = \frac{\min'_{s_2 \in S} r(z, s_2) - \min_{s_1 \in S} r(z, s_1)}{\min'_{s_2 \in S} r(z, s_2) + \min_{s_1 \in S} r(z, s_1)}.$$

Если $S_k = S_{k-1} \cup \{b\}$, где $b \in B_i$, то расстояние от b до t_i равно 0,5 и добавление b в S_k повлияет только на значения $F(b, S_k)$ и $F(t_i, S_k)$, поэтому $F(S_k) \leq F(S_{k-1}) + 1 + \frac{1}{3}$.

Если $S_k = S_{k-1} \cup \{x\}$, где $x \in \widetilde{X}$, то расстояния от x до l троек из множества $\widetilde{T}$, содержащих x , равны 0,5 и добавление x повлияет на значение $F(x, S_k)$ и l значений вида $F(t_i, S_k)$, поэтому

$$F(S_k) \leq F(S_{k-1}) + 1 + l \cdot \frac{1}{3} \leq F(S_{k-1}) + 1 + M \cdot \frac{1}{3}.$$

Аналогичные неравенства выполняются для $y \in \widetilde{Y}$ и $z \in \widetilde{Z}$.

Если $S_k = S_{k-1} \cup \{t\}$, где $t = (x, y, z) \in \widetilde{T}$, то расстояния от t до N элементов связанного с ним множества B_i и до x, y, z равны 0,5, и добавление t повлияет на значение $F(x, S_k)$ и N значений вида $F(b, S_k)$ для $b \in B_i$. При этом для $b \in B_i$ оно увеличится с 0 до $\frac{1}{3}$, для t — с 0 до 1. Значения $F(x, S_k)$, $F(y, S_k)$, $F(z, S_k)$ могут либо увеличиться с 0 до $\frac{1}{3}$, либо остаться равными 0, либо уменьшиться с $\frac{1}{3}$ до 0 в зависимости от того, были ли в наборе S_{k-1} тройки, содержащие x, y, z соответственно. В результате имеем

$$F(S_k) \geq F(S_{k-1}) + 1 + N \cdot \frac{1}{3} - 3 \cdot \frac{1}{3} > F(S_{k-1}) + M \cdot \frac{1}{3} + 1.$$

Таким образом, доказано, что множество столпов S_k может быть образовано только элементами $t \in \widetilde{T}$ при любом $k \leq M$.

Осталось доказать, что множество S_p , являющееся решением задачи 1, за линейное время может быть преобразовано в решение задачи 3DM S путём замены каждого $\tilde{t}_i \in \widetilde{T}$ на $t_i \in T$ с тем же индексом.

Предположим, что множество S , полученное таким образом, не является решением задачи 3DM, т. е. в нём существует пара различных троек, пересекающихся минимум по одному элементу. При этом существует множество S^* , которое является решением задачи 3DM и обеспечивает условие непересекаемости всех троек. Рассмотрим образ этого множества S_p^* , полученный переходом от множества объектов из T к множеству объектов из $\widetilde{T}$ с теми же индексами.

Сравним качество $F(S_p)$ и $F(S_p^*)$. Так как каждый из p элементов множеств $\widetilde{X}$, $\widetilde{Y}$ и $\widetilde{Z}$ имеет ровно один элемент из S_p^* на расстоянии 0,5, значение FRIS-функции для них на этом множестве типичных представителей равно $\frac{1}{3}$. Аналогично для N каждого из p множеств B_i , связанных с теми t_i , которые попали в S_p^* . Для самих $t_i \in S_p^*$ значения FRIS-функции равны 1. В результате имеем

$$F(S_p^*) = 3 \cdot p \cdot \frac{1}{3} + N \cdot p \cdot \frac{1}{3} + p = 2p + \frac{1}{3}Np.$$

С другой стороны, если в множестве столпов присутствует пара образов пересекающихся троек, то величина FRIS-функции для элемента из $\widetilde{X}$, $\widetilde{Y}$ или $\widetilde{Z}$, по которому они пересекаются, равна 0, так как два ближайших

к нему столпа находятся на одинаковом расстоянии 0,5. Тогда

$$F(S_p) \leq (3p - 1) \cdot \frac{1}{3} + Np \cdot \frac{1}{3} + p < F(S_p^*),$$

а следовательно, S_p решением задачи 1 быть не может.

Таким образом, если $F(S_p) = 2p + \frac{1}{3}Np$, то задача 3DM имеет решение, и получить его можно, включив в S образы всех элементов из S_p . Если же $F(S_p) < 2p + \frac{1}{3}Np$, то решения задачи 3DM не существует. Теорема 1 доказана.

Следствие 1. *Задача 1 выбора подмножества из множества точек в метрическом пространстве NP-трудна.*

ДОКАЗАТЕЛЬСТВО. Следствие справедливо в силу того, что к рассматриваемой задаче за полиномиальное время сводится задача о трёхмерном сочетании, которая NP-трудна. Следствие 1 доказано.

Заключение

В работе доказано, что известная NP-трудная в сильном смысле задача о трёхмерном сочетании сводится к задаче выбора из конечного множества точек метрического пространства подмножества p типичных представителей, наилучшим образом описывающих кластеры этого множества с точки зрения некоторого геометрического критерия. Тем самым доказано, что указанная задача выбора NP-трудна, и в рамках парадигмы $P \neq NP$ невозможно предложить точный полиномиальный алгоритм решения этой задачи.

Финансирование работы

Исследование выполнено при финансовой поддержке Российского фонда фундаментальных исследований (проект № 17-01-00710). Дополнительных грантов на проведение или руководство этим исследованием получено не было.

Конфликт интересов

Автор заявляет, что у неё нет конфликта интересов.

Литература

1. Zagoruiko N. G., Borisova I. A., Dyubanov V. V., Kutnenko O. A. Methods of recognition based on the function of rival similarity // Pattern Recognit. Image Anal. 2008. V. 18, No. 1. P. 1–6.
2. Борисова И. А. Вычислительная сложность задачи выбора типичных представителей в 2-разбиении конечного множества точек метрического пространства // Дискрет. анализ и исслед. операций. 2020. Т. 27, № 2. С. 5–16.

3. Кутненко О. А., Плясунов А. В. NP-трудность некоторой задачи цензурирования данных // Дискрет. анализ и исслед. операций. 2021. Т. 28, № 2. С. 60–73.
4. Кутненко О. А. Вычислительная сложность двух задач когнитивного анализа данных // Дискрет. анализ и исслед. операций. 2022. Т. 29, № 1. С. 18–32.
5. Aloise D., Deshpande A., Hansen P., Popat P. NP-hardness of Euclidean sum-of-squares clustering // Machine Learn. 2009. V. 75, No. 2. P. 245–248.
6. Dasgupta S. The hardness of k -means clustering. Tech. Rep. CS2007-0890. San Diego, CA: Univ. California, 2008. 6 p.
7. Papadimitriou C. H. Worst-case and probabilistic analysis of a geometric location problem // SIAM J. Comput. 1981. V. 10, No. 3. P. 542–557.
8. Har-Peled S., Mazumdar S. On coresets for k -means and k -median clustering and their applications // Proc. 36th Annu. ACM Symp. Theory of Computing (Chicago, USA, June 13–16, 2004). New York: ACM, 2004. P. 291–300.
9. Kaufman L., Rousseeuw P. J. Clustering by means of medoids // Statistical data analysis based on the L_1 -norm and related methods. Amsterdam: North Holland, 1987. P. 405–416.
10. Кельманов А. В., Пяткин А. В., Хандеев В. И. О сложности некоторых максиминных задач кластеризации // Тр. Ин-та математики и механики. 2018. Т. 24, № 4. С. 189–198.
11. Кельманов А. В., Пяткин А. В. NP-трудность некоторых евклидовых задач разбиения конечного множества точек // Журн. вычисл. математики и мат. физики. 2018. Т. 58, № 5. С. 852–856.
12. Aggarwal H., Imai N., Katoh N., Suri S. Finding k points with minimum diameter and related problems // J. Algorithms. 1991. V. 12, No. 1. P. 38–56.
13. Zukhba A. V. NP-completeness of the problem of prototype selection in the nearest neighbor method // Pattern Recognit. Image Anal. 2010. V. 20, No. 4. P. 484–494.
14. Karp R. M. Reducibility among combinatorial problems // Complexity of computer computations. Proc. Symp. (New York, USA, March 20–22, 1972). New York: Plenum Press, 1972. P. 85–103.

Борисова Ирина Артёмовна

Статья поступила
3 сентября 2024 г.
После доработки —
16 сентября 2024 г.
Принята к публикации
22 сентября 2024 г.

COMPUTATIONAL COMPLEXITY OF THE CHOICE PROBLEM FOR TYPICAL REPRESENTATIVES OF A FINITE POINT SET IN A METRIC SPACE

I. A. Borisova^{1,2}

¹ Sobolev Institute of Mathematics,
4 Acad. Koptuyug Avenue, 630090 Novosibirsk, Russia

² Novosibirsk State University,
2 Pirogov Street, 630090 Novosibirsk, Russia

E-mail: biamia@mail.ru

Abstract. We analyze the complexity of one extremal problem of choosing a subset of p points in a given finite set in a metric space. The chosen subset of points is required to describe given clusters in the best way from the point of view of some geometric criterion. This problem is a formalization of one applied problem from data mining that consists in finding a subset of typical representatives of a dataset based on the rival similarity function of. We prove that the problem under consideration is NP-hard by polynomially reducing the well-known NP-hard 3D-Matching problem to this one. Illustr. 1, bibliogr. 14.

Keywords: NP-hard problem, choice of typical representatives, FRiS-function, 3D-matching problem, machine learning.

References

1. N. G. Zagoruiko, I. A. Borisova, V. V. Dyubanov, and O. A. Kutnenko, Methods of recognition based on the function of rival similarity, *Pattern Recognit. Image Anal.* **18** (1), 1–6 (2008).
2. I. A. Borisova, Computational complexity of the problem of choosing typical representatives in a 2-clustering of a finite set of points in a metric space, *Diskretn. Anal. Issled. Oper.* **27** (2), 5–16 (2020) [Russian] [*J. Appl. Ind. Math.* **14** (2), 242–248 (2020)].
3. O. A. Kutnenko and A. V. Plyasunov, NP-hardness of some data cleaning problem, *Diskretn. Anal. Issled. Oper.* **28** (2), 60–73 (2021) [Russian] [*J. Appl. Ind. Math.* **15** (2), 285–291 (2021)].

4. **O. A. Kutnenko**, Computational complexity of two problems of cognitive data analysis, *Diskretn. Anal. Issled. Oper.* **29** (1), 18–32 (2022) [Russian] [*J. Appl. Ind. Math.* **16** (1), 89–97 (2022)].
5. **D. Aloise, A. Deshpande, P. Hansen, and P. Popat**, NP-hardness of Euclidean sum-of-squares clustering, *Machine Learn.* **75** (2), 245–248 (2009).
6. **S. Dasgupta**, The hardness of k -means clustering, *Tech. Rep. CS2007-0890* (Univ. California, San Diego, CA, 2008).
7. **C. H. Papadimitriou**, Worst-case and probabilistic analysis of a geometric location problem, *SIAM J. Comput.* **10** (3), 542–557 (1981).
8. **S. Har-Peled and S. Mazumdar**, On coresets for k -means and k -median clustering and their applications, in *Proc. 36th Annu. ACM Symp. Theory of Computing, Chicago, USA, June 13–16, 2004* (ACM, New York, 2004), pp. 291–300.
9. **L. Kaufman and P. J. Rousseeuw**, Clustering by means of medoids, in *Statistical Data Analysis Based on the L_1 -Norm and Related Methods* (North Holland, Amsterdam, 1987), pp. 405–416.
10. **A. V. Kel'manov, A. V. Pyatkin, and V. I. Khandeev**, On the complexity of some max–min clustering problems, *Tr. Inst. Mat. Mekh.* **24** (4), 189–198 (2018) [Russian] [*Proc. Steklov Inst. Math.* **309** (Suppl. 1), S65–S73 (2020)].
11. **A. V. Kel'manov and A. V. Pyatkin**, NP-hardness of some Euclidean problems of partitioning a finite set of points, *Zh. Vychisl. Mat. Mat. Fiz.* **58** (5), 852–856 (2018) [Russian] [*Comput. Math. Math. Phys.* **58** (5), 822–826 (2018)].
12. **H. Aggarwal, N. Imai, N. Katoh, and S. Suri**, Finding k points with minimum diameter and related problems, *J. Algorithms* **12** (1), 38–56 (1991).
13. **A. V. Zukhba**, NP-completeness of the problem of prototype selection in the nearest neighbor method, *Pattern Recognit. Image Anal.* **20** (4), 484–494 (2010).
14. **R. M. Karp**, Reducibility among combinatorial problems, in *Complexity of Computer Computations* (Proc. Symp., New York, USA, March 20–22, 1972) (Plenum Press, New York, 1972), pp. 85–103.

Irina A. Borisova

Received September 3, 2024

Revised September 16, 2024

Accepted September 22, 2024

УНИВЕРСАЛЬНЫЕ ЦИКЛЫ, ПОРОЖДАЮЩИЕ ВСЕ ГРАФЫ КОАЛИЦИОННЫХ РАЗБИЕНИЙ ЦИКЛОВ

А. Н. Глебов^а, А. А. Добрынин^б

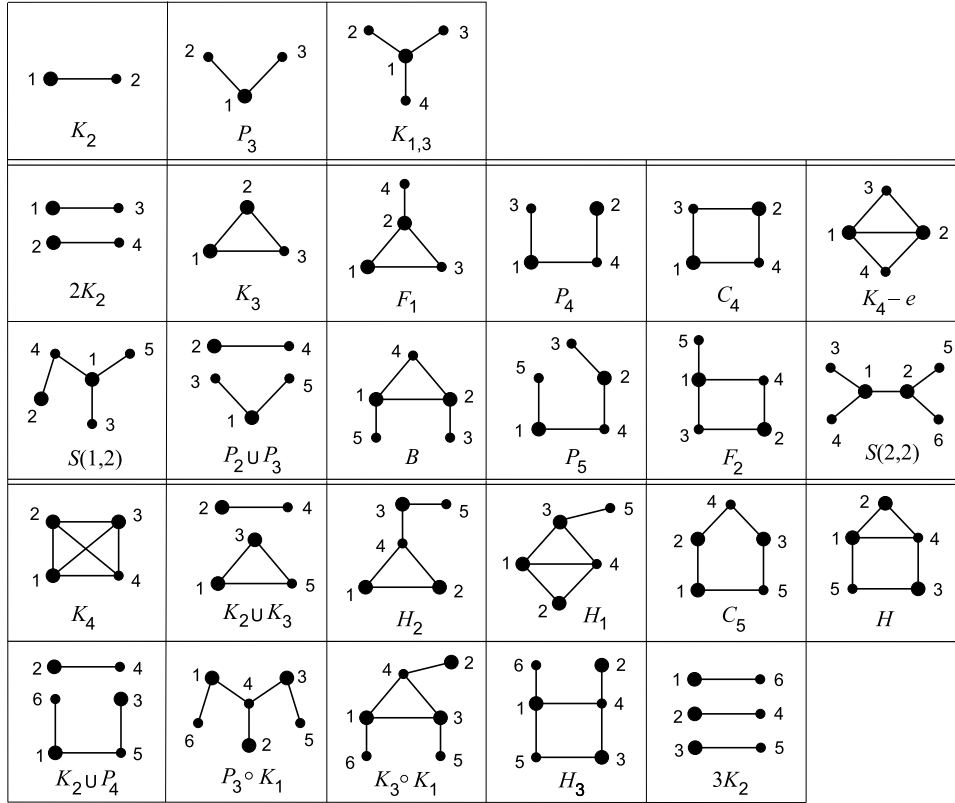
Институт математики им. С. Л. Соболева,
пр. Акад. Коптюга, 4, 630090 Новосибирск, Россия
E-mail: ^аangle@math.nsc.ru, ^бdobr@math.nsc.ru

Аннотация. *Коалицией* в графе G называется пара непересекающихся и недоминирующих подмножеств его вершин $V_1, V_2 \subset V(G)$ таких, что объединение $V_1 \cup V_2$ является доминирующим множеством. В коалиционном разбиении $\pi(G) = \{V_1, V_2, \dots, V_k\}$ каждое недоминирующее множество V_i входит в некоторую коалицию с другим множеством этого разбиения, а если V_i доминирующее, то оно одновершинное. Коалиционное разбиение вершин графа G порождает граф коалиций $CG(G, \pi)$, в котором вершины соответствуют множествам разбиения и две вершины смежны, если соответствующие множества образуют коалицию. Известно, что в совокупности все простые циклы порядка более трёх порождают 26 графов коалиций с числом вершин не более шести. Универсальный цикл порождает все такие графы. В работе показано, что только циклы C_{3k} при $k \geq 5$ универсальны. Табл. 4, ил. 1, библиогр. 25.

Ключевые слова: граф, доминирующее множество, коалиционное разбиение, граф коалиций.

Введение

Рассматриваются обыкновенные графы G с множеством вершин $V(G)$. Множество вершин $S \subseteq V(G)$ *доминирующее*, если каждая вершина $v \in V(G) \setminus S$ смежна хотя бы с одной вершиной из S . *Числом доминирования* $\gamma(G)$ графа G называется мощность его наименьшего доминирующего множества. Теория доминирования является большой и хорошо изученной областью теории графов, имеющей многочисленные приложения. Подробные сведения и библиографию по доминированию в графах можно найти в книгах [1–6]. В 2020 г. в работе [7] введено понятие


 Рис. 1. Все графы коалиций, порождаемые циклом C_n при $n \geq 4$

коалиции в графах, отражающее способность нескольких недоминирующих подмножеств вершин графа образовывать доминирующие множества. В работе рассматриваются коалиции, образованные парами подмножеств вершин графа.

Пусть V_1 и V_2 — два непересекающихся подмножества вершин графа G . Эти подмножества образуют *коалицию* в G , если ни одно из них не является доминирующим, но их объединение $V_1 \cup V_2$ будет доминирующим. Разбиение $\pi(G) = \{V_1, V_2, \dots, V_k\}$ вершин графа G называется *коалиционным*, если для каждого недоминирующего множества из π существует другое множество из π , образующее с ним коалицию, а каждое доминирующее множество из π одновершинное. Максимальное число подмножеств в коалиционном разбиении графа G называется его *коалиционным числом* и обозначается через $C(G)$. Математическое понятие коалиции возникло как модель описания организации групп людей для достижения своих целей. Пусть, например, парламент состоит

Таблица 1

Число попарно неизоморфных графов коалиций,
порождаемых циклом C_n , $8 \leq n \leq 18$

Цикл	C_8	C_9	C_{10}	C_{11}	C_{12}	C_{13}	C_{14}	C_{15}	C_{16}	C_{17}	C_{18}
Число графов	17	19	17	17	25	17	17	26	17	17	26

Таблица 2

Число графов коалиций, порождаемых циклом C_n , $n \in \{9, 12, 15, 18\}$

$CG(C_n)$	C_9	C_{12}	C_{15}	C_{18}	$CG(C_n)$	C_9	C_{12}	C_{15}	C_{18}
K_2	54	711	7 740	75 927	$S(2, 2)$	9	216	3 645	52 074
P_3	873	32 400	1 007 625	28 366 308	$K_2 \cup P_4$	—	114	3 960	91 089
$K_{1,3}$	450	29 668	1 494 235	63 580 626	$P_2 \cup P_3$	—	59	865	8 844
K_3	573	17 856	448 285	10 236 134	$K_2 \cup K_3$	—	—	16	180
$2K_2$	243	19 890	1 232 520	62 902 353	C_5	18	216	2 040	16 182
P_4	1 278	65 070	2 572 200	88 237 314	P_5	18	312	3 690	37 806
F_1	639	27 138	813 665	20 963 358	H_2	—	48	480	3 222
C_4	171	5 166	124 590	2 693 505	H_1	—	12	90	450
$K_4 - e$	198	4 116	70 995	1 113 054	H	9	123	1 540	16 849
$S(1, 2)$	180	8 388	307 020	9 656 142	$3K_2$	9	120	1 095	8 397
B	342	12 084	359 550	9 696 120	$P_3 \circ K_1$	9	84	615	3 897
K_4	72	1 950	38 445	667 629	$K_3 \circ K_1$	—	4	35	198
F_2	27	972	21 525	386 118	H_3	—	12	90	450

из нескольких партий, каждая из которых не в состоянии добиться желаемого результата (провести какой-либо закон и т. п.), в то время как коалиция двух или нескольких партий даёт возможность сделать это [8]. Коалиции в графах, основанные на доминировании с дополнительными свойствами (тотальное, связное, сильное, независимое доминирование и др.), изучались в [9–16].

Картину образования коалиций в графе G можно характеризовать графами коалиций, которые строятся по разбиениям множества его вершин [17]. Если π — некоторое разбиение $V(G)$, то оно рассматривается в качестве множества вершин графа коалиций $CG(G, \pi)$, при этом две вершины смежны в $CG(G, \pi)$, если соответствующие множества разбиения π образуют коалицию в G .

Вопросы нахождения коалиционных чисел и описания графов коалиций для различных классов графов изучались в [18–25]. В частности, доказано, что простые цепи в совокупности порождают 18 графов коалиций, три из которых существуют только для цепей малой длины,

Таблица 3

 Число графов коалиций, порождаемых циклом C_n , $8 \leq n \leq 18$, $3 \nmid n$

$CG(C_n)$	C_8	C_{10}	C_{11}	C_{13}	C_{14}	C_{16}	C_{17}
K_2	20	130	308	1 599	3 542	16 716	35 768
P_3	24	2 940	9 878	103 220	324 632	3 090 976	9 399 742
$K_{1,3}$	96	1 760	7 623	111 176	414 946	5 285 496	18 477 708
K_3	84	1 910	5 951	53 235	155 526	1 281 880	3 634 141
$2K_2$	38	1 025	4 565	80 184	317 891	4 643 930	17 221 340
P_4	88	4 760	17 930	224 952	767 760	8 449 216	27 452 688
F_1	204	2 490	8 525	87 256	268 926	2 439 404	7 180 069
C_4	48	560	1 716	15 210	43 820	350 336	975 256
$K_4 - e$	40	650	1 287	11 440	25 928	182 576	434 690
$S(1, 2)$	40	425	2 585	25 064	102 816	946 824	3 288 752
B	48	920	3 168	34 229	112 742	1 038 080	3 243 277
K_4	48	260	1 232	5 967	21 497	110 480	333 744
F_2	16	150	451	3 081	9 289	57 752	163 829
$S(2, 2)$	8	10	154	364	2 198	7 232	28 458
$K_2 \cup P_4$	4	15	110	546	2 324	13 416	44 217
$P_2 \cup P_3$	7	50	77	364	672	2 787	5 474
$K_2 \cup K_3$	8	3	44	26	210	168	935

а цепь P_{10} является наименьшей цепью, порождающей все остальные 15 графов коалиций, т. е. является универсальной цепью. Простые циклы C_n с числом вершин $n \geq 4$ порождают в точности 26 графов коалиций, порядок которых не превосходит шести [18, 19]. На рис. 1 приводятся диаграммы этих графов, обозначенных K_2 , P_3 , P_4 , P_5 , $2K_2$, K_3 , C_4 , C_5 , $K_{1,3}$, F_1 и F_2 , $K_4 - e$, K_4 , $P_2 \cup P_3$, $K_2 \cup K_3$, H , $S(1, 2)$, $S(2, 2)$, B , H_1 , H_2 , H_3 , $3K_2$, $K_2 \cup P_4$, $P_3 \circ K_1$ и $K_3 \circ K_1$. Простой цикл называется *универсальным*, если он порождает все указанные графы. Вопрос о существовании универсальных циклов сформулирован в [19]. В табл. 1–3 приводятся полученные нами с помощью компьютерных вычислений данные о числе графов коалиций для цикла C_n с числом вершин $8 \leq n \leq 18$, которые показывают, что циклы C_{15} и C_{18} универсальны. В настоящей работе докажем, что только цикл C_{3k} универсальный, где $k \geq 5$.

1. Основной результат

Докажем сначала, что простые циклы с числом вершин, не кратным трём, не могут быть универсальными.

Утверждение 1. Циклы C_{3k-2} и C_{3k-1} не универсальны для любого $k \geq 4$.

ДОКАЗАТЕЛЬСТВО. Покажем, что несвязный граф $3K_2$ не может быть графом, порождаемым коалиционными разбиениями вершин простых циклов C_{3k-2} и C_{3k-1} . Пусть, напротив, $3K_2$ является графом коалиций некоторого коалиционного разбиения такого цикла $\pi = \{V_1, V_2, \dots, V_6\}$. Без потери общности предположим, что множества V_{2i-1} и V_{2i} для каждого $i \in \{1, 2, 3\}$ образуют коалиции. Известно, что для чисел доминирования циклов выполняется равенство $\gamma(C_{3k-2}) = \gamma(C_{3k-1}) = k$. Тогда для всех i мощность каждой коалиции $|V_{2i-1} \cup V_{2i}| \geq k$ и, следовательно, $\sum_{i=1}^3 |V_{2i-1} \cup V_{2i}| = \sum_{i=1}^6 |V_i| \geq 3k$. В силу того, что π является разбиением множества вершин цикла, имеем $\sum_{i=1}^6 |V_i| \in \{3k-2, 3k-1\}$. Из полученного противоречия следует, что множество графов коалиций циклов C_{3k-2} и C_{3k-1} не может содержать граф $3K_2$. Утверждение 1 доказано.

Для универсальности цикла с числом вершин не менее 15 достаточно, чтобы его порядок делился на 3.

Утверждение 2. Цикл C_{3k} универсальный для любого $k \geq 5$.

ДОКАЗАТЕЛЬСТВО. Пусть вершины циклов имеют последовательную нумерацию, начиная с вершины 1. Коалиционные разбиения вершин цикла C_{3k} , порождающие графы на рис. 1, будут получены добавлением новых вершин в множества разбиений цикла C_{15} из табл. 4. Все графы коалиций на рис. 1 имеют наименьшие вершинные покрытия мощности не более 3, которые изображены большими вершинами. Тогда любая коалиция будет содержать множество, соответствующее хотя бы одной вершине из вершинного покрытия графа коалиций. Обозначим вершины такого покрытия через v_1, v_2 и v_3 (вершины могут совпадать), а соответствующие им множества коалиционного разбиения через V_1, V_2 и V_3 соответственно. Для получения коалиционных разбиений цикла C_{15+3t} , $t \geq 1$, к множествам V_1, V_2 и V_3 цикла C_{15} будем добавлять $3t$ новых вершин, сохраняя тем самым графы коалиций. Рассмотрим три случая в зависимости от мощности вершинного покрытия. Для более удобной проверки доминирования при изображении фрагмента цикла вместо номеров его вершин используем номера соответствующих множеств коалиционных разбиений.

СЛУЧАЙ 1. Три графа коалиций имеют вершинное покрытие мощности 1 (вершина $v_1 = 1$ в графах верхнего ряда на рис. 1), т. е. каждая коалиция будет включать множество V_1 . Согласно табл. 4 смежные вершины 1 и 2 цикла C_{15} принадлежат множеству V_1 в коалиционных разбиениях этих графов. Тогда все вставляемые в цикл C_{15} новые вершины

Таблица 4

 Разбиения вершин цикла C_{15} , порождающие все графы коалиций

$CG(C_{15})$	Разбиение множества $V(C_{15}) = V_1 \cup V_2 \cup \dots$	Рёбра $CG(C_{15})$
K_2	$\{1, 2, \dots, 12\}, \{13, 14, 15\}$	1-2
P_3	$\{1, 2, \dots, 12\}, \{13, 15\}, \{14\}$	1-2; 1-3
$K_{1,3}$	$\{1, 2, \dots, 12\}, \{13\}, \{14\}, \{15\}$	1-2; 1-3; 1-4
K_3	$\{1, 2, 5, 7, 9, 11\}, \{3, 13, 14\}, \{4, 6, 8, 10, 12, 15\}$	1-2,3; 2-3
$2K_2$	$\{1, 2, 5, 10, 11\}, \{3, 12\}, \{4, 8, 13, 14\}, \{6, 7, 9, 15\}$	1-3; 2-4
P_4	$\{1, 2, 5, 10, 11, 14\}, \{3, 12, 15\}, \{4, 8\}, \{6, 7, 9, 13\}$	1-3,4; 2-4
F_1	$\{1, 2, 5, 8, 10, 11\}, \{3, 12, 15\}, \{4, 14\}, \{6, 7, 9, 13\}$	1-2,3,4; 2-4
C_4	$\{1, 2, 5, 8, 11\}, \{3, 9, 15\}, \{4, 7, 12, 14\}, \{6, 10, 13\}$	1-3,4; 2-3,4
$K_4 - e$	$\{1, 2, 5, 8, 11\}, \{3, 9, 12, 15\}, \{4, 7, 14\}, \{6, 10, 13\}$	1-2,3,4; 2-3,4
K_4	$\{1, 4, 7, 10\}, \{2, 8, 11, 14\}, \{3, 6, 13\}, \{5, 9, 12, 15\}$	1-2,3,4; 2-3,4; 3-4
$P_2 \cup P_3$	$\{1, 2, 5, 8, 11\}, \{3, 15\}, \{4, 14\}, \{6, 9, 12\}, \{7, 10, 13\}$	1-3,5; 2-4
$S(1, 2)$	$\{1, 2, 5, 8, 11\}, \{3\}, \{4, 14\}, \{6, 9, 12, 15\}, \{7, 10, 13\}$	1-3,4,5; 2-4
P_5	$\{1, 2, 5, 8, 13\}, \{3, 9, 14, 15\}, \{4, 7, 12\}, \{6, 11\}, \{10\}$	1-4,5; 2-3,4
B	$\{1, 2, 5, 8\}, \{3, 9, 12, 15\}, \{4, 7\}, \{6, 11, 14\}, \{10, 13\}$	1-2,4,5; 2-3,4
F_2	$\{1, 2, 5, 8, 14\}, \{3, 15\}, \{4, 7, 10, 13\}, \{6, 9, 12\}, \{11\}$	1-3,4,5; 2-3,4
C_5	$\{1, 4, 13\}, \{2, 7, 10\}, \{3, 8, 15\}, \{5, 11, 14\}, \{6, 9, 12\}$	1-2,5; 2-4; 3-4,5
$K_2 \cup K_3$	$\{1, 4, 7, 10\}, \{2, 14\}, \{3, 13\}, \{5, 8, 11\}, \{6, 9, 12, 15\}$	1-3,5; 2-4; 3-5
H_2	$\{1, 4, 7, 10\}, \{2, 13\}, \{3, 15\}, \{5, 8, 11, 14\}, \{6, 9, 12\}$	1-2,4; 2-4; 3-4,5
H_1	$\{1, 4, 7, 10\}, \{2, 13\}, \{3, 12, 15\}, \{5, 8, 11, 14\}, \{6, 9\}$	1-2,3,4; 2-4; 3-4,5
H	$\{1, 4, 7, 13\}, \{2, 10\}, \{3, 15\}, \{5, 8, 11, 14\}, \{6, 9, 12\}$	1-2,4,5; 2-4; 3-4,5
$S(2, 2)$	$\{1, 2, 5, 8, 13, 14\}, \{3, 9, 12, 15\}, \{4, 7\}, \{6\}, \{10\}, \{11\}$	1-2,5,6; 2-3,4
$3K_2$	$\{1, 4, 7, 10\}, \{2, 14\}, \{3, 15\}, \{5, 8, 11\}, \{6, 9, 12\}, \{13\}$	1-6; 2-4; 3-5
$K_2 \cup P_4$	$\{1, 4, 7, 10\}, \{2, 14\}, \{3\}, \{5, 8, 11\}, \{6, 9, 12, 15\}, \{13\}$	1-5,6; 2-4; 3-5
$P_3 \circ K_1$	$\{1, 4, 7, 10\}, \{2\}, \{3, 15\}, \{5, 8, 11, 14\}, \{6, 9, 12\}, \{13\}$	1-4,6; 2-4; 3-4,5
$K_3 \circ K_1$	$\{1, 4, 7, 10\}, \{2\}, \{3, 12, 15\}, \{5, 8, 11, 14\}, \{6, 9\}, \{13\}$	1-3,4,6; 2-4; 3-4,5
H_3	$\{1, 4, 7, 13\}, \{2\}, \{3, 15\}, \{5, 8, 11, 14\}, \{6, 9, 12\}, \{10\}$	1-4,5,6; 2-4; 3-4,5

добавим только к множеству V_1 . Фрагмент $\dots - 1 - 1 - \dots$ (соответствующий вершинам 1 и 2) цикла C_{15} после добавления t троек вершин перейдёт во фрагмент цикла C_{15+3t} как показано ниже:

$$\dots - 1 - 1 - \dots \rightarrow \dots - 1 - (1 - 1 - 1)^t - 1 - \dots$$

Очевидно, что пополненное множество V_1 будет образовывать в цикле C_{15+3t} такие же коалиции, как и исходное V_1 в графе C_{15} .

СЛУЧАЙ 2. Вершинное покрытие мощности 2 имеют 12 графов коалиций (вершины $v_1 = 1$ и $v_2 = 2$ в графах второго и третьего рядов

на рис. 1). Каждая коалиция будет содержать хотя бы одно из множеств V_1 или V_2 . Смежные вершины 1 и 2 цикла C_{15} принадлежат множеству V_1 в коалиционных разбиениях всех этих графов (см. табл. 4). Тогда новые вершины, добавляемые в множества V_1 и V_2 графа C_{15} , можно разместить в цикле C_{15+3t} следующим образом:

$$\dots - 1 - 1 - \dots \rightarrow \dots - 1 - (1 - 2 - 1)^t - 1 - \dots$$

Так как все вершины указанного справа фрагмента цикла C_{15+3t} доминируются вершинами пополненных множеств V_1 и V_2 , все графы коалиций, порождаемые циклами C_{15} и C_{15+3t} , будут совпадать.

СЛУЧАЙ 3. Вершинное покрытие мощности 3 имеют 11 графов коалиций (вершины $v_1 = 1$, $v_2 = 2$ и $v_3 = 3$ в графах из двух последних рядов на рис. 1). Каждая коалиция будет содержать хотя бы одно из множеств V_1 , V_2 или V_3 . Как видно из табл. 4, смежные вершины 1 и 2 цикла C_{15} принадлежат множествам V_1 и V_2 в коалиционных разбиениях этих графов. Тем самым новые вершины, добавляемые в множества V_1 , V_2 и V_3 , можно расположить в цикле C_{15+3t} следующим способом:

$$\dots - 1 - 2 - \dots \rightarrow \dots - 1 - (2 - 3 - 1)^t - 2 - \dots$$

Легко проверить, что все вершины фрагмента C_{15+3t} при любом $t \geq 1$ доминируются вершинами пополненных множеств V_1 , V_2 и V_3 , а это влечёт совпадение графов коалиций для C_{15} и C_{15+3t} .

Таким образом, для любого $t \geq 0$ к циклу C_{15} всегда можно добавить $3t$ вершин с сохранением коалиционных разбиений из табл. 4, что обеспечивает порождение циклом C_{15+3t} всех графов коалиций, т. е. универсальность цикла C_{15+3t} . Утверждение 2 доказано.

Финансирование работы

Исследование выполнено в рамках государственного задания Института математики им. С. Л. Соболева (проект № FWNF-2022-0017).

Конфликт интересов

Авторы заявляют, что у них нет конфликта интересов.

Литература

1. Haynes T. W., Hedetniemi S. T., Slater P. J. Fundamentals of domination in graphs. New York: Marcel Dekker, 1998. 464 p. (Pure Appl. Math.; V. 208). DOI: 10.1201/9781482246582.
2. Du D.-Z., Wan P.-J. Connected dominating set: Theory and applications. New York: Springer, 2012. 206 p. (Springer Optim. Appl.; V. 77). DOI: 10.1007/978-1-4614-5242-3.

3. **Henning M. A., Yeo A.** Total domination in graphs. New York: Springer, 2013. 178 p. DOI: 10.1007/978-1-4614-6525-6.
4. Topics in domination in graphs. Cham: Springer, 2020. 546 p. (Dev. Math.; V. 64). DOI: 10.1007/978-3-030-51117-3.
5. Structures of domination in graphs. Cham: Springer, 2021. 536 p. (Dev. Math.; V. 66). DOI: 10.1007/978-3-030-58892-2.
6. Domination in graphs: Core concepts. Cham: Springer, 2023. 644 p. DOI: 10.1007/978-3-031-09496-5.
7. **Haynes T. W., Hedetniemi J. T., Hedetniemi S. T., McRae A. A., Mohan R.** Introduction to coalitions in graphs // AKCE Int. J. Graphs Comb. 2020. V. 17, No. 2. P. 653–659. DOI: 10.1080/09728600.2020.1832874.
8. **Fisher St. D., Hobolt S. B.** Coalition government and electoral accountability // Electoral Stud. 2010. V. 29, No. 3. P. 358–369. DOI: 10.1016/j.electstud.2010.03.003.
9. **Alikhani S., Bakhshesh D., Golmohammadi H.** Total coalitions in graphs // Quaest. Math. 2024. V. 47, No. 11. P. 2283–2294. DOI: 10.2989/16073606.2024.2365365.
10. **Alikhani S., Bakhshesh D., Golmohammadi H., Konstantinova E. V.** Connected coalitions in graphs // Discuss. Math. Graph Theory. 2024. V. 44, No. 4. P. 1551–1566. DOI: 10.7151/dmgt.2509.
11. **Alikhani S., Bakhshesh D., Golmohammadi H., Klavžar S.** On independent coalition in graphs and independent coalition graphs // Discuss. Math. Graph Theory. 2025. V. 45, No. 2. P. 533–544. DOI: 10.7151/dmgt.2543.
12. **Golmohammadi H., Alikhani S., Ghanbari N., Takhonov I. I., Abaturon A.** Strong coalitions in graphs. Ithaca, NY: Cornell Univ., 2024. 12 p. (Cornell Univ. Libr. e-Print Archive; arXiv:2404.11575). DOI: 10.48550/arXiv.2404.11575.
13. **Samadzadeh M. R., Mojdeh D. A.** Independent coalition in graphs: Existence and characterization // Ars Math. Contemp. 2024. V. 24, No. 3. Article ID P3.09. 17 p. DOI: 10.26493/1855-3974.3113.6f7.
14. **Brešar B., Henning M. A., Klavžar S., Rall D. F.** Domination games played on graphs. Cham: Springer, 2021. 122 p. DOI: 10.1007/978-3-030-69087-8.
15. **Bakhshesh D., Henning M. A.** The minmin coalition number in graphs // Aequat. Math. 2024. V. 99. P. 223–236. DOI: 10.1007/s00010-024-01045-5.
16. **Bednarz P., Pirga M.** On proper 2-dominating sets in graphs // Symmetry. 2024. V. 16, No. 3. Article ID 296. 10 p. DOI: 10.3390/sym16030296.
17. **Haynes T. W., Hedetniemi J. T., Hedetniemi S. T., McRae A. A., Mohan R.** Coalition graphs // Commun. Comb. Optim. 2023. V. 8, No. 2. P. 423–430. DOI: 10.22049/CCO.2022.27916.1394.
18. **Bakhshesh D., Henning M. A., Pradhan D.** On the coalition number of trees // Bull. Malays. Math. Sci. Soc. 2023. V. 46. Article ID 95. 14 p. DOI: 10.1007/s40840-023-01492-4.

19. Haynes T. W., Hedetniemi J. T., Hedetniemi S. T., McRae A. A., Mohan R. Coalition graphs of paths, cycles, and trees // Discuss. Math. Graph Theory. 2023. V. 43, No. 4. P. 931–946. DOI: 10.7151/dmgt.2416.
20. Haynes T. W., Hedetniemi J. T., Hedetniemi S. T., McRae A. A., Mohan R. Upper bounds on the coalition number // Australas. J. Comb. 2021. V. 80, No. 3. P. 442–453.
21. Alikhani S., Golmohammadi H., Konstantinova E. V. Coalition of cubic graphs of order at most 10 // Commun. Comb. Optim. 2024. V. 9, No. 3. P. 437–450. DOI: 10.22049/cco.2023.28328.1507.
22. Dobrynin A. A., Golmohammadi H. On cubic graphs having the maximum coalition number // Сиб. электрон. мат. изв. 2024. Т. 21, № 1. С. 356–362. DOI: 10.33048/semi.2024.21.027.
23. Golmohammadi H. Total coalitions of cubic graphs of order at most 10 // Commun. Comb. Optim. 2024. V. 10, No. 3. P. 601–615. DOI: 10.22049/cco.2024.29015.1813.
24. Haynes T. W., Hedetniemi J. T., Hedetniemi S. T., McRae A. A., Mohan R. Self-coalition graphs // Opuscula Math. 2023. V. 43, No. 2. P. 173–183. DOI: 10.7494/opmath.2023.43.2.173.
25. Bakhshesh D., Henning M. A., Pradhan D. Singleton coalition graph chains // Comput. Appl. Math. 2024. V. 43. Article ID 85. 22 p. DOI: 10.1007/s40314-023-02588-0.

Глебов Алексей Николаевич
Добрынин Андрей Алексеевич

Статья поступила
11 июля 2024 г.
После доработки —
16 августа 2024 г.
Принята к публикации
22 сентября 2024 г.

UNIVERSAL CYCLES THAT GENERATE ALL GRAPHS OF COALITION PARTITIONS IN CYCLES

A. N. Glebov^a and A. A. Dobrynin^b

Sobolev Institute of Mathematics,
4 Acad. Koptuyug Avenue, 630090 Novosibirsk, Russia
E-mail: ^aangle@math.nsc.ru, ^bdobr@math.nsc.ru

Abstract. A *coalition* in a graph G is a pair of disjoint nondominating subsets of its vertices $V_1, V_2 \subset V(G)$ such that $V_1 \cup V_2$ is a dominating set. In the coalition partition $\pi(G) = \{V_1, V_2, \dots, V_k\}$, every nondominating set V_i is included in some coalition and if V_i is dominating, then it is a single-vertex set. A coalition partition of vertices of a graph G generates a coalition graph $CG(G, \pi)$ whose vertices correspond to the partition sets, while two vertices are adjacent if the corresponding sets form a coalition. It is well known that all simple cycles of order greater than three generate in total 26 coalition graphs of order at most six. A universal cycle generates all such graphs. It is shown that only the cycles C_{3k} , $k \geq 5$, are universal. Tab. 4, illustr. 1, bibliogr. 25.

Keywords: graph, dominating set, coalition partition, coalition graph.

References

1. T. W. Haynes, S. T. Hedetniemi, and P. J. Slater, *Fundamentals of Domination in Graphs* (Marcel Dekker, New York, 1998) (Pure Appl. Math., Vol. 208), DOI: 10.1201/9781482246582.
2. D.-Z. Du and P.-J. Wan, *Connected Dominating Set: Theory and Applications* (Springer, New York, 2012) (Springer Optim. Appl., Vol. 77), DOI: 10.1007/978-1-4614-5242-3.
3. M. A. Henning and A. Yeo, *Total Domination in Graphs* (Springer, New York, 2013), DOI: 10.1007/978-1-4614-6525-6.
4. *Topics in Domination in Graphs* (Springer, Cham, 2020) (Dev. Math., Vol. 64), DOI: 10.1007/978-3-030-51117-3.
5. *Structures of Domination in Graphs* (Springer, Cham, 2021) (Dev. Math., Vol. 66), DOI: 10.1007/978-3-030-58892-2.

English transl.: *Journal of Applied and Industrial Mathematics* **19** (1), 33–39 (2025), DOI: 10.1134/S199047892501003X.

6. *Domination in Graphs: Core Concepts* (Springer, Cham, 2023), DOI: 10.1007/978-3-031-09496-5.
7. **T. W. Haynes, J. T. Hedetniemi, S. T. Hedetniemi, A. A. McRae,** and **R. Mohan**, Introduction to coalitions in graphs, *AKCE Int. J. Graphs Comb.* **17** (2), 653–659 (2020), DOI: 10.1080/09728600.2020.1832874.
8. **St. D. Fisher** and **S. B. Hobolt**, Coalition government and electoral accountability, *Electoral Stud.* **29** (3), 358–369 (2010), DOI: 10.1016/j.electstud.2010.03.003.
9. **S. Alikhani, D. Bakhshesh,** and **H. Golmohammadi**, Total coalitions in graphs, *Quaest. Math.* **47** (11), 2283–2294 (2024), DOI: 10.2989/16073606.2024.2365365.
10. **S. Alikhani, D. Bakhshesh, H. Golmohammadi,** and **E. V. Konstantinova**, Connected coalitions in graphs, *Discuss. Math. Graph Theory.* **44** (4), 1551–1566 (2024), DOI: 10.7151/dmgt.2509.
11. **S. Alikhani, D. Bakhshesh, H. Golmohammadi,** and **S. Klavžar**, On independent coalition in graphs and independent coalition graphs, *Discuss. Math. Graph Theory* **45** (2), 533–544 (2025), DOI: 10.7151/dmgt.2543.
12. **H. Golmohammadi, S. Alikhani, N. Ghanbari, I. I. Takhonov,** and **A. Abaturov**, Strong coalitions in graphs (Cornell Univ., Ithaca, NY, 2024) (Cornell Univ. Libr. e-Print Archive, arXiv:2404.11575), DOI: 10.48550/arXiv.2404.11575.
13. **M. R. Samadzadeh** and **D. A. Mojdeh**, Independent coalition in graphs: Existence and characterization, *Ars Math. Contemp.* **24** (3), ID P3.09 (2024), DOI: 10.26493/1855-3974.3113.6f7.
14. **B. Brešar, M. A. Henning, S. Klavžar,** and **D. F. Rall**, *Domination Games Played on Graphs* (Springer, Cham, 2021), DOI: 10.1007/978-3-030-69087-8.
15. **D. Bakhshesh** and **M. A. Henning**, The minmin coalition number in graphs, *Aequat. Math.* **99**, 223–236 (2024), DOI: 10.1007/s00010-024-01045-5.
16. **P. Bednarz** and **M. Pirga**, On proper 2-dominating sets in graphs, *Symmetry* **16** (3), ID 296 (2024), DOI: 10.3390/sym16030296.
17. **T. W. Haynes, J. T. Hedetniemi, S. T. Hedetniemi, A. A. McRae,** and **R. Mohan**, Coalition graphs, *Commun. Comb. Optim.* **8** (2), 423–430 (2023), DOI: 10.22049/CCO.2022.27916.1394.
18. **D. Bakhshesh, M. A. Henning,** and **D. Pradhan**, On the coalition number of trees, *Bull. Malays. Math. Sci. Soc.* **46**, ID 95 (2023), DOI: 10.1007/s40840-023-01492-4.
19. **T. W. Haynes, J. T. Hedetniemi, S. T. Hedetniemi, A. A. McRae,** and **R. Mohan**, Coalition graphs of paths, cycles, and trees, *Discuss. Math. Graph Theory.* **43** (4), 931–946 (2023), DOI: 10.7151/dmgt.2416.
20. **T. W. Haynes, J. T. Hedetniemi, S. T. Hedetniemi, A. A. McRae,** and **R. Mohan**, Upper bounds on the coalition number, *Australas. J. Comb.* **80** (3), 442–453 (2021).

-
21. **S. Alikhani, H. Golmohammadi, and E. V. Konstantinova**, Coalition of cubic graphs of order at most 10, *Commun. Comb. Optim.* **9** (3), 437–450 (2024), DOI: 10.22049/cco.2023.28328.1507.
 22. **A. A. Dobrynin and H. Golmohammadi**, On cubic graphs having the maximum coalition number, *Sib. Elektron. Mat. Izv.* **21** (1), 356–362 (2024), DOI: 10.33048/semi.2024.21.027.
 23. **H. Golmohammadi**, Total coalitions of cubic graphs of order at most 10, *Commun. Comb. Optim.* **10** (3), 601–615 (2024), DOI: 10.22049/cco.2024.29015.1813.
 24. **T. W. Haynes, J. T. Hedetniemi, S. T. Hedetniemi, A. A. McRae, and R. Mohan**, Self-coalition graphs, *Opuscula Math.* **43** (2), 173–183 (2023), DOI: 10.7494/opmath.2023.43.2.173.
 25. **D. Bakhshesh, M. A. Henning, and D. Pradhan**, Singleton coalition graph chains, *Comput. Appl. Math.* **43**, ID 85 (2024), DOI: 10.1007/s40314-023-02588-0.

Aleksey N. Glebov
Andrey A. Dobrynin

Received July 11, 2024
Revised August 16, 2024
Accepted September 22, 2024

ОПТИМИЗАЦИЯ СТРАТЕГИИ ПОСТАВОК ЗАКАЗОВ ПРИ
УПРАВЛЕНИИ ЗАПАСАМИ ПО МНОГИМ КРИТЕРИЯМ
В УСЛОВИЯХ НЕОПРЕДЕЛЁННОСТИ

Д. А. Гусев^{1, a}, О. А. Свиридова^{1, b},
И. Г. Шидловский^{1, c}, Г. Л. Бродецкий^{2, d}

¹ ФГБУ ВО «РЭУ им. Г. В. Плеханова»,
Стремянный пер., д. 36, 115054 Москва, Россия

² Высшая школа бизнеса, НИУ «Высшая школа экономики»,
ул. Шаболовка, д. 26, стр. 1, 119049 Москва, Россия

E-mail: ^a gussev79@mail.ru, ^b sviridova.oa@rea.ru,
^c shdlvsk-ivan@yandex.ru, ^d gbrodetskiy@hse.ru

Аннотация. В задачах управления запасами возникает необходимость оптимизации новых модифицированных моделей для выбора наилучших решений по обеспечению поставок заказов. Особенностью таких моделей является учёт различных неопределённостей и концепции временной ценности денег. При этом зачастую соответствующие решения на практике отличаются многокритериальностью, что обусловлено особенностями цепей поставок и также учитывается в постановках подобных задач.

Авторами разработан подход к оптимизации задач указанного типа на основе синтеза процедур многокритериальной оптимизации и процедур выбора в условиях неопределённости. Практика показала, что указанный подход требует специальной модификации, которая будет представлена в этой статье. Предлагаемая модификация реализуется с целью помочь менеджерам избегать нежелательных ситуаций с выбором альтернатив, т. е. ситуаций, связанных с феноменами, приводящими к выбору альтернатив, которые могут оказаться неадекватными предпочтениям лица, принимающего решения.

Соответствующая модификация обусловит реализацию специальной особенности в формате процедур оптимизации. Исходные показатели задаваемых частных критериев при такой модификации предлагается представлять на основе так называемых обобщённых данных, позволяющих избавиться от фактора их размерности.

В этой статье представлена численная иллюстрация на примере разработки стратегии управления запасами с необходимостью выбора логистического посредника в условиях неопределённости спроса и возможных задержек поставок. При этом в качестве критерия выбора используется взвешенная сумма оценок частных критериев, а для учёта фактора неопределённости используется критерий Гурвица. Табл. 7, библиогр. 20.

Ключевые слова: управление запасами, выбор в условиях неопределённости, неопределённость спроса, задержка поставок, многокритериальная оптимизация, обобщённый критерий выбора, феномен неадекватного выбора.

Введение

В современных условиях происходит очевидное нарастание неопределённости для ряда данных, которые надо использовать в формате процедур многокритериальной оптимизации в задачах управления запасами. С одной стороны, учёт указанного фактора приводит к совершенствованию алгоритмов оптимизации запасов. С другой стороны, это оказывает дополнительно дестабилизирующее влияние на функционирование цепей поставок [1]. Как видим, всё более актуальными становятся исследования возможностей учёта фактора неопределённости, в особенности применительно к задачам многокритериального выбора наилучших альтернатив при транспортном обеспечении поставок в области управления запасами.

Отметим, что исследуемые в этой статье решения представлены конечным числом альтернатив. Атрибуты рассматриваемых альтернатив обусловлены практическими особенностями задач транспортного обеспечения поставок. Учёт фактора неопределённости реализован на основе формирования полной группы дискретных событий (как это предложено в [2, 3]), связанных с параметрами стратегии управления запасами и особенностями транспортного обеспечения требуемых поставок.

1. Особенности учёта фактора неопределённости при многокритериальном выборе на дискретном множестве решений

Чтобы организовать оптимальный выбор альтернативы, причём с учётом именно фактора неопределённости, отметим, что в теории уже разработан специальный арсенал подходов и методов оптимизации [4–10]. Он представлен своими критериями выбора. Кстати, формат соответствующих процедур выбора позволяет учитывать индивидуальное отношение конкретного ЛПР (лица, принимающего решения) к особенностям анализируемого фактора неопределённости.

Как отмечено выше, для указанных областей выбора наилучшего решения об организации поставок при управлении запасами особо актуальными являются методы многокритериальной оптимизации. При этом для таких процедур оптимизации часто используют методологию принятия решений по многим критериям (multiple criteria decision making, MCDM [11]). Сегодня, как уже отмечалось, при оптимизации таких решений требуются методы, позволяющие одновременно учитывать и фактор неопределённости. Особенности таких методов рассматриваются в этой работе. В формате представленного исследования ограничимся ситуациями, когда в качестве критерия выбора (для оптимизации по многим критериям на дискретном множестве решений) предполагается использовать именно критерии выбора прямого типа [12–16].

Возможности синтеза процедур многокритериального выбора с процедурами принятия решений в условиях неопределённости для отмеченных выше приложений (для задач управления запасами) рассмотрены в [17]. Отметим, что на практике разработанные в [17] процедуры синтеза могут потребовать специальной модификации. А именно, здесь рассматривается модификация, которая соотносится с требованием исключать так называемые феномены неадекватного выбора применительно к задачам оптимизации запасов [18]. Напомним, что при реализации процедур многокритериального выбора могут возникнуть нежелательные ситуации. В частности, реализация необходимых процедур может привести к нежелательным результатам (т. е. результатам, воспринимаемым ЛПР в качестве сомнительных). Такие ситуации будем называть *феноменами неадекватного выбора*. Здесь будут рассмотрены только указанные нежелательные ситуации, которые относятся именно к так называемым феноменам несогласованности порядка значений частных критериев [18]. Речь идёт о ситуациях, когда показатели некоторых частных критериев (в формате задач многокритериальной оптимизации) могут существенно превосходить показатели других частных критериев.

Заметим, что другой тип реализованной специальной модификации для указанных задач разработан ранее для ситуации, когда воздействие внешних факторов надо соотносить только с конкретным заданным набором анализируемых частных критериев (а не всей группы таких критериев, как это формализовано в [17]). Соответствующая модификация представлена в [19], где предложены специфика требуемого подхода к процедурам оптимизации и соответствующий алгоритм вычислительных процедур для случая, когда частные критерии представлены в виде двух групп таких критериев. В одной группе — частные критерии, которые подвергаются влиянию фактора неопределённости, а в другой — частные критерии, которые такому влиянию не подвергаются. В дополнение отметим, что предложенная в настоящей статье модификация процедур

многокритериальной оптимизации для задач рассматриваемого типа может применяться и с учётом специфики указанных двух групп частных критериев.

2. Возможности идентификации воздействия феноменов неадекватного выбора и подходы к их устранению

Как описано в [18], на дискретном множестве решений интересующая нас задача многокритериального выбора может быть формализована в табличном виде. При этом в таблице по строкам представляются альтернативы X_i , $i = 1, \dots, m$, а по столбцам — частные критерии C_k , $k = 1, \dots, N$. В ячейках таблицы приводятся известные априори заданные показатели a_{ik} (для каждой альтернативы X_i по частному критерию C_k). При этом также формируется дополнительный специальный столбец с показателями критерия выбора F_i , причём для каждой альтернативы X_i . В частности, если для задач типа $C_k \rightarrow \max$ используется критерий выбора на основе взвешенной суммы оценок заданных частных критериев, то указанные показатели F_i будут определяться равенствами вида $F_i = \sum_{k=1}^N w_k a_{ik}$, где w_k — заданный вес частного критерия C_k , причём наилучшая альтернатива должна максимизировать такой показатель.

Обратим также внимание на следующее. При реализации процедур выбора может оказаться, что проявятся нежелательные ситуации. Речь идёт о ситуациях, когда некоторые частные критерии участвуют в выборе лишь формально, а их воздействие на результат выбора сомнительно. Различные типы феноменов неадекватного выбора (актуальных для цепей поставок) описаны в [18]. Здесь с учётом реализованной модификации имеет смысл отметить именно возможность воздействия феномена неадекватного выбора, который будет обусловлен несогласованностью порядка задаваемых значений для оценок частных критериев (поскольку воздействие такого феномена надо будет учитывать далее в численном примере). Воздействие этого феномена проявляется следующим образом. Числовые значения показателей некоторых частных критериев могут весьма существенно превосходить заданные числовые значения показателей остальных частных критериев. В таком случае формально все частные критерии участвуют в вычислениях, но реально некоторые из них могут не оказывать влияния на выбор. В указанной ситуации вряд ли ЛПР будет считать найденный выбор адекватным. Воздействие такого феномена можно обнаружить специальными процедурами, которые представлены далее.

Интересующее нас воздействие возникает, если при удалении некоторых столбцов таблицы, отражающих оценки соответствующих частных критериев, не наблюдается никаких изменений в оптимизации решения

по конкретному критерию выбора. Такие частные критерии, удаление которых не приводит к изменению наилучшего выбора, не оказывают влияния на выбираемое решение. Другими словами, они участвуют в процедурах оптимизации только формально.

В частности, при использовании критерия выбора по методу взвешенной суммы оценок частных критериев, как уже отмечалось, целевая (критериальная) функция имеет вид

$$F_i = \sum_{k=1}^N w_k a_{ik} \rightarrow \max.$$

В этом случае воздействие указанного феномена может быть обнаружено на основе следующих процедур.

Допустим, что некоторое слагаемое $w_k a_{ik}$ в рассматриваемой целевой функции существенно больше остальных в формате рассматриваемой альтернативы. Естественнo предположить, что именно частный критерий C_k будет обеспечивать выбор альтернативы. При этом некоторые другие частные критерии (с меньшими слагаемыми в целевой функции) могут не оказывать влияния на результаты сравнения альтернатив. Проверить это можно на основе расчёта рассматриваемой целевой функции, удалив из анализа один частный критерий с малым слагаемым. Если выбор альтернативы не изменится, то предположение подтвердится. Иначе убеждаемся, что исключённый частный критерий оказывает влияние на ранжирование альтернатив.

Понятно, что полную картину можно получить, если последовательно удалять по одному частному критерию (именно те, для которых есть подозрение в отсутствии влияния на ранжирование альтернатив) и выполнять указанные выше процедуры.

Один из способов, позволяющих устранить воздействие такого феномена, состоит в переходе к так называемым обобщённым данным для рассматриваемой задачи оптимизации [4, 18]. Уточним особенности такого перехода при максимизации задаваемых частных критериев.

Предварительно потребуется обратиться к координатам утопической точки (УТ). Рассмотрим координаты $УТ_k = \max_{i=1, \dots, m} \{a_{ik}\}$. При их использовании уже модифицированные обобщённые показатели частных критериев (с учётом их весов) примут вид

$$\frac{w_k a_{ik}}{УТ_k}.$$

Итак, далее в формате процедур реализуемой многокритериальной оптимизации предполагается использование именно обобщённых критериев выбора [4, 18].

3. Модифицированный подход к многокритериальному выбору альтернативы в условиях неопределённости

Как уже отмечалось, алгоритм выбора альтернативы по многим критериям в условиях неопределённости представлен в [17]. Приведём здесь краткий вариант этого алгоритма с модификацией, направленной на переход к обобщённым данным. Отличие от исходного алгоритма заключается в том, что теперь добавлен специальный шаг 4.

Шаг 1. Атрибуты задачи многокритериального выбора. Задаётся перечень альтернатив $X_i, i = 1, \dots, m$, а также перечень частных критериев $C_k, k = 1, \dots, N$, с весами w_k . Априори предполагается, что в задаче требуется максимизировать показатели всех частных критериев. Задаётся многокритериальный критерий выбора. В численном примере, представленном в разд. 4, в качестве такового выступает критерий взвешенной суммы показателей частных критериев.

Шаг 2. Формирование полной группы случайных событий; задание критерия выбора в условиях неопределённости. Для рассматриваемой ситуации формализуется полная группа возможных случайных событий, оказывающих влияние на экономический результат. Пусть для такой группы выделены события $\theta_j, j = 1, \dots, n$. На этом шаге также задаётся критерий выбора в условиях неопределённости. В численной иллюстрации в качестве такового выбран критерий Гурвица.

Шаг 3. Расчёт показателей частных критериев для событий полной группы. Для каждой альтернативы X_i по частному критерию C_k при реализации события θ_j определяется значение конечного результата a_{ijk} . Совокупность таких значений для каждой пары i, j представляется в виде вектора $(a_{ij1}; \dots; a_{ijk}; \dots; a_{ijN})$.

Шаг 4. Переход к обобщённым данным. Для каждой альтернативы X_i и полученного вектора $(a_{ij1}; \dots; a_{ijk}; \dots; a_{ijN})$ реализуется переход к вектору обобщённых данных $V_{ij} = (V_{ij1}; \dots; V_{ijk}; \dots; V_{ijN})$, представляющему эту альтернативу: $V_{ijk} = \frac{w_k a_{ijk}}{UT_k}$, где $UT_k = \max_{i=1, \dots, m} \{a_{ijk}\}$.

Шаг 5. Формирование матрицы полезностей. С учётом атрибутов процедур выбора в условиях неопределённости на этом шаге требуется формализовать соответствующую матрицу полезностей $B = (b_{ij})$. Напомним, что элемент b_{ij} должен представлять полезность соответствующего результата для альтернативы X_i при внешнем воздействии θ_j . При этом если в качестве показателя многокритериального критерия выбора, как в числовом примере, используется взвешенная сумма обобщённых оценок частных критериев, то $b_{ij} = V_{ij1} + \dots + V_{ijN}$.

Шаг 6. ВЫБОР НАИЛУЧШЕЙ АЛЬТЕРНАТИВЫ. Определяются показатели заданного критерия выбора в условиях неопределённости, которые образуют дополнительный столбец матрицы полезностей. Если используется критерий Гурвица, как в числовом примере, то такие показатели находят по формуле: $c \min_j \{b_{ij}\} + (1 - c) \max_j \{b_{ij}\}$. Напомним, что в случае применения критерия Гурвица необходимо задать показатель c , отражающий задаваемый ЛПР уровень пессимизма ($0 \leq c \leq 1$).

Далее при помощи численного примера проиллюстрируем предложенный алгоритм применительно к задаче выбора наилучшей стратегии при управлении запасами. В приложении А отмечены особенности реализуемой здесь оптимизации с учётом фактора временной ценности денег. Это позволит сократить объём представленного примера (в частности, за счёт возможности сослаться на атрибуты аналогичных расчётов, приведённых в [17]).

4. Численный пример выбора наилучшей стратегии управления запасами при многих критериях с учётом фактора неопределённости

Пусть рассматривается возможность распространения новой продукции парфюмерии из дружественной страны в Московскую область. Предполагается, что при определении стратегии управления запасами необходимо учитывать неопределённость спроса и неопределённость применительно к возможным задержкам, а также выбрать логистического посредника. Для упрощения изложения и последующих иллюстраций далее принимаем, что выбирается один из двух возможных таких посредников (в общем случае можно рассматривать задачу такого типа при любом соотношении долей участия в сделках с различными посредниками).

Учёт фактора неопределённости спроса предполагает, что возможны три сценария: D_1 — низкий спрос, D_2 — средний спрос, D_3 — высокий спрос (их вероятности неизвестны). Сценарии по возможным задержкам поставок товара зависят от реализуемого сценария по спросу. Для упрощения задержки поставок будем учитывать только применительно к выбору второго посредника, априори полагая, что у первого посредника задержки не случаются. В случае задержки по вине второго посредника годовые потери из-за них задаются экспертно в зависимости от реализованного сценария по спросу ($S_2(D_1)$ — есть задержки при спросе D_1 ; $S_2(D_2)$ — есть задержки при спросе D_2 ; $S_2(D_3)$ — есть задержки при спросе D_3).

Представим реализацию предложенного выше алгоритма выбора наилучшего решения. В соответствии с первым шагом алгоритма требуется задать перечень анализируемых альтернатив. Естественно, что они

обусловлены с одной стороны возможностью выбора ориентации на конкретный спрос (по указанным выше сценариям в условиях неопределённости). С другой стороны, альтернативы будут также обусловлены необходимостью выбора одного из двух возможных посредников. Далее рассмотрим шесть альтернативных решений применительно к указанной задаче оптимизации. Они обозначаются следующим образом: $X_1(D_1; \Pi_1)$, $X_2(D_1; \Pi_2)$, $X_3(D_2; \Pi_1)$, $X_4(D_2; \Pi_2)$, $X_5(D_3; \Pi_1)$, $X_6(D_3; \Pi_2)$. Здесь обозначения в скобках указывают на ориентацию применительно к одному из трёх сценариев по спросу, а также применительно к выбору одного из двух посредников (Π_1 и Π_2).

Многокритериальный выбор реализуется по следующим пяти частным критериям. Критерии и фигурирующие в них нежелательные величины заданы ЛПР:

- C_1 — разность между нежелательной величиной (17 000 у. е.) и фактическим значением годовых издержек на доставку товара от места производства до склада в Московской области; в стоимость доставки включены как затраты на транспортировку, так и таможенное оформление груза;
- C_2 — разность между нежелательной величиной (2400 у. е.) и фактическим значением годовых издержек хранения товара на складе в Московской области;
- C_3 — разность между нежелательной величиной (90 000 у. е.) и фактическим значением годовых потерь, связанных с задержками поставок по вине логистического посредника;
- C_4 — разность между нежелательной величиной (30 000 у. е.) и фактическим значением годовых потерь, связанных с ошибками прогнозирования спроса;
- C_5 — разность между нежелательной величиной (500 у. е.) и фактическим значением годовых транзакционных издержек.

По построению алгоритма показатели всех частных критериев должны быть определены таким образом, чтобы их требовалось максимизировать. Соответственно используемые при оптимизации показатели издержек формализованы так, что необходимо максимизировать разность между заданной экспертом величиной издержек, допустимой в крайнем случае, и их фактическим значением. Веса частных критериев выбраны следующим образом: $w_1 = 0,3$, $w_2 = 0,1$, $w_3 = 0,2$, $w_4 = 0,2$, $w_5 = 0,2$. Наконец, в качестве критерия многокритериального выбора применяется метод взвешенной суммы оценок частных критериев.

На втором шаге алгоритма для учёта фактора неопределённости задаём 6 событий, образующих полную группу. Речь идёт о следующих событиях полной группы: $\theta_1 \rightarrow (D_1S_1)$, $\theta_2 \rightarrow (D_2S_1)$, $\theta_3 \rightarrow (D_3S_1)$, $\theta_4 \rightarrow (D_1S_2)$, $\theta_5 \rightarrow (D_2S_2)$, $\theta_6 \rightarrow (D_3S_2)$. В частности, отметим, что

Таблица 1

**Исходные данные для рассматриваемой задачи
многокритериального выбора в условиях неопределённости**

Показатель	Обозначение	Значение	Ед. изм.
Низкий спрос на продукцию (первый сценарий)	D_1	1600	шт.
Средний спрос на продукцию (второй сценарий)	D_2	2000	шт.
Высокий спрос на продукцию (третий сценарий)	D_3	2300	шт.
Цена закупки	C_{Π}	380	у. е./шт.
Выручка от реализации единицы товара	C_s	430	у. е./шт.
Прибыль от реализации единицы товара	P_{Π}	50	у. е./шт.
Транзакционные издержки посредника I	—	100	у. е.
Транзакционные издержки посредника II	—	400	у. е.
Требуемые отчисления от прибыли на единицу товара	L_{Π}	10	у. е./шт.
Стоимость доставки посредником I	C_{01}	2500	у. е./партия
Стоимость доставки посредником II	C_{02}	2200	у. е./партия
Стоимость хранения единицы товара за год посредником I	C_{h1}	12	у. е./шт.
Стоимость хранения единицы товара за год посредником II	C_{h2}	11	у. е./шт.
Потери из-за задержек у второго посредника при спросе D_1	$S_2(D_1)$	50 000	у. е.
Потери из-за задержек у второго посредника при спросе D_2	$S_2(D_2)$	70 000	у. е.
Потери из-за задержек у второго посредника при спросе D_3	$S_2(D_3)$	80 000	у. е.

например событие $\theta_1 \rightarrow (D_1 S_1)$ соотносится с ситуацией, когда наступает случайное событие θ_1 . Кроме того, на втором шаге надо указать задаваемый ЛПР критерий выбора в условиях неопределённости. Как уже отмечалось, в качестве такого критерия выбора задан критерий Гурвица. При этом для полноты изложения результаты такого выбора рассматриваются при трёх различных уровнях пессимизма $c \in \{0,05; 0,5; 0,85\}$.

Таблица 2

Показатели альтернатив по частным критериям C_1 – C_5 применительно к ситуации с внешним воздействием по событию θ_1

Альтернатива	C_1	C_2	C_3	C_4	C_5
$X_1(D_1; \Pi_1)$	4 607,64	463,32	90 000	30 000	400
$X_2(D_1; \Pi_2)$	5 234,36	754,53	90 000	30 000	100
$X_3(D_2; \Pi_1)$	5 257,99	356,06	90 000	19 600	400
$X_4(D_2; \Pi_2)$	5 855,29	662,85	90 000	19 800	100
$X_5(D_3; \Pi_1)$	5 649,73	285,51	90 000	11 800	400
$X_6(D_3; \Pi_2)$	6 228,95	602,59	90 000	12 150	100
УТ	6 228,95	754,53	90 000	30 000	400

На третьем шаге алгоритма требуется рассчитать показатели частных критериев для всех альтернатив применительно к заданным событиям полной группы. Для этого понадобятся исходные данные, формализованные в табл. 1. Дополнительно отметим, что специальные показатели C_{Π} , C_S , P_{Π} , L_{Π} , а также C_{01} , C_{02} , C_{h1} , C_{h2} формализуются в приложении А.

Далее для краткости расчёты будут приведены только применительно к ситуации с внешним воздействием по событию θ_1 (см. табл. 2, где также представлены координаты соответствующей УТ). Возможные комментарии, в частности, для альтернативы $X_1(D_1; \Pi_1)$ даны в приложении Б. При этом по спросу реализуется сценарий D_1 , а по задержкам у второго посредника — сценарий S_1 .

Таблица 3

Обобщённые показатели альтернатив по частным критериям K_1 – K_5 применительно к ситуации с внешним воздействием по событию θ_1

Альтернатива	K_1	K_2	K_3	K_4	K_5	$\sum K_i$
$X_1(D_1; \Pi_1)$	0,222	0,061	0,200	0,200	0,200	0,883
$X_2(D_1; \Pi_2)$	0,252	0,100	0,200	0,200	0,050	0,802
$X_3(D_2; \Pi_1)$	0,253	0,047	0,200	0,131	0,200	0,831
$X_4(D_2; \Pi_2)$	0,282	0,088	0,200	0,132	0,050	0,752
$X_5(D_3; \Pi_1)$	0,272	0,038	0,200	0,079	0,200	0,789
$X_6(D_3; \Pi_2)$	0,300	0,080	0,200	0,081	0,050	0,711

Обратим внимание, что согласно табл. 2 частные критерии C_2 и C_5 по всем альтернативам имеют показатели на порядок меньше, чем показатели других частных критериев. В такой ситуации возможно проявление феномена неадекватного выбора, который будет обусловлен несогласованностью порядка значений оценок частных критериев. Чтобы это проверить, следует сопоставить финальные результаты наилучшего выбора как с учётом второго и пятого частных критериев, так и без них (в частности, когда число частных критериев сократится до трёх), что будет сделано в табл. 6, 7. Предварительно следует реализовать шаг 4 модифицированного алгоритма. Тем самым требуется выполнить переход к обобщённым показателям всех частных критериев.

Далее для реализации шага 4 алгоритма в табл. 3 представлены обобщённые показатели частных критериев. При этом частные критерии C_k после реализации этой процедуры будем обозначать через K_k .

В соответствии с четвёртым шагом алгоритма в табл. 3 реализован переход к обобщённым данным. При этом, например, для альтернативы $X_1(D_1; \Pi_1)$ её обобщённый показатель по частному критерию K_1 составит $\frac{w_{1111}}{\Sigma T_1} = \frac{0,34607,64}{6228,95} = 0,222$.

Соответствующий столбец для матрицы полезностей в случае реализации события θ_1 (что потребуется на следующем шаге алгоритма) получен в табл. 3 при сложении найденных обобщённых показателей частных критериев по строкам.

В соответствии с пятым шагом алгоритма в табл. 4 сформирована матрица полезностей, в которой её первый столбец для случая реализации события θ_1 получен на основе табл. 3. Расчёты по остальным событиям, как отмечалось выше, опускаются.

Таблица 4

Матрица полезностей в условиях неопределённости

Альтернатива	Обобщённый экономический результат при возможных внешних воздействиях (в формате полной группы случайных событий)					
	θ_1 (D_1S_1)	θ_2 (D_2S_1)	θ_3 (D_3S_1)	θ_4 (D_1S_2)	θ_5 (D_2S_2)	θ_6 (D_3S_2)
$X_1(D_1; \Pi_1)$	0,883	0,683	0,314	0,883	0,683	0,314
$X_2(D_1; \Pi_2)$	0,802	0,638	0,381	0,802	0,638	0,381
$X_3(D_2; \Pi_1)$	0,831	0,844	0,591	0,731	0,744	0,491
$X_4(D_2; \Pi_2)$	0,752	0,798	0,652	0,651	0,698	0,552
$X_5(D_3; \Pi_1)$	0,789	0,824	0,773	0,639	0,674	0,623
$X_6(D_3; \Pi_2)$	0,711	0,778	0,830	0,559	0,628	0,680

Таблица 5

Результаты выбора по критерию Гурвица
при различных уровнях пессимизма c

Альтернатива	$c = 0,05$	$c = 0,5$	$c = 0,85$
$X_1(D_1; \Pi_1)$	0,855	0,599	0,399
$X_2(D_1; \Pi_2)$	0,781	0,591	0,444
$X_3(D_2; \Pi_1)$	0,827	0,668	0,544
$X_4(D_2; \Pi_2)$	0,786	0,675	0,589
$X_5(D_3; \Pi_1)$	0,814	0,724	0,653
$X_6(D_3; \Pi_2)$	0,816	0,694	0,599

В соответствии с заключительным шагом алгоритма в табл. 5 представлены результаты выбора по критерию Гурвица при различном возможном уровне пессимизма ЛПР, реализованные по данным табл. 4.

Например, при $c = 0,05$ для альтернативы $X_1(D_1; \Pi_1)$ показатель критерия Гурвица составит $0,05 \cdot 0,314 + 0,95 \cdot 0,883 = 0,855$. Таким образом, при $c = 0,05$ альтернатива $X_1(D_1; \Pi_1)$ становится наилучшей, а при $c = 0,5$ и при $c = 0,85$ наилучшей становится альтернатива $X_5(D_3; \Pi_1)$.

5. Оценка факта проявления феномена неадекватного выбора

Для оценки факта проявления феномена неадекватного выбора сравним итоговые результаты наилучшего выбора в ситуациях с использованием различных наборов частных критериев, представленные в табл. 6.

Из табл. 6 видно, что для критерия Гурвица использование исходного алгоритма оптимизации (без перехода к обобщённым данным) при задействовании всех частных критериев даёт те же результаты, что и без учёта второго или пятого частных критериев или без них вместе взятых. Причём такое подведение исходного алгоритма наблюдается при любом

Таблица 6

Сравнение результатов наилучшего выбора
для исходного алгоритма

Атрибуты алгоритма	$c = 0,05$	$c = 0,5$	$c = 0,85$
С учётом всех критериев	$X_2(D_1; \Pi_2)$	$X_1(D_1; \Pi_1)$	$X_1(D_1; \Pi_1)$
Без 2-го частного критерия	$X_2(D_1; \Pi_2)$	$X_1(D_1; \Pi_1)$	$X_1(D_1; \Pi_1)$
Без 5-го частного критерия	$X_2(D_1; \Pi_2)$	$X_1(D_1; \Pi_1)$	$X_1(D_1; \Pi_1)$
Без 2-го и 5-го частных критериев	$X_2(D_1; \Pi_2)$	$X_1(D_1; \Pi_1)$	$X_1(D_1; \Pi_1)$

уровне пессимизма ЛПР. Таким образом, проявляется феномен неадекватного выбора в указанном выше понимании: второй и пятый частные критерии не оказывают влияния на наилучший выбор.

6. Возможности устранения воздействия феномена неадекватного выбора

Чтобы убедиться в устранении феномена неадекватного выбора, сравним итоговые результаты наилучшего выбора в ситуациях с использованием различных наборов частных критериев при реализации модифицированного алгоритма, представленные в табл. 7.

Как видно из табл. 7, удаление только пятого частного критерия (в формате процедур оптимизации) меняет наилучший выбор. Это значит, что теперь (при использовании обобщённых данных) пятый частный критерий будет оказывать влияние на выбор. Другими словами, можно говорить об устранении проявления феномена неадекватного выбора применительно к указанному частному критерию.

В то же время, как видно из табл. 7, удаление только второго частного критерия (в формате процедур оптимизации) не меняет наилучший выбор. Стоит отметить, что в данной ситуации (рассматриваемой числовой иллюстрации) предложенная модификация не обеспечила устранение указанного феномена. Тем не менее, в более общих ситуациях (см. ниже) использование указанной модификации позволяет избавиться от воздействия рассматриваемого нежелательного феномена.

Действительно, как видно из табл. 7, для обобщённого критерия Гурвица в случае одновременного удаления и второго, и пятого частных критериев при использовании предложенного модифицированного алгоритма происходит изменение наилучшего выбора. Это значит, что при использовании обобщённых данных указанные второй и пятый частные критерии оказывают влияние на выбор. Как видим, можно говорить об устранении проявления феномена неадекватного выбора.

Таблица 7

Сравнение результатов наилучшего выбора для модифицированного алгоритма

Атрибуты алгоритма	$c = 0,05$	$c = 0,5$	$c = 0,85$
С учётом всех критериев	$X_1(D_1; \Pi_1)$	$X_5(D_3; \Pi_1)$	$X_5(D_3; \Pi_1)$
Без 2-го частного критерия	$X_1(D_1; \Pi_1)$	$X_5(D_3; \Pi_1)$	$X_5(D_3; \Pi_1)$
Без 5-го частного критерия	$X_6(D_3; \Pi_2)$	$X_6(D_3; \Pi_2)$	$X_6(D_3; \Pi_2)$
Без 2-го и 5-го частных критериев	$X_6(D_3; \Pi_2)$	$X_6(D_3; \Pi_2)$	$X_6(D_3; \Pi_2)$

Наконец, отметим следующее. Предложенная модификация не может во всех случаях гарантировать исключение воздействия рассматриваемого нежелательного феномена.

Заключение

В статье предложена модификация моделей для выбора наилучших решений по обеспечению поставок заказов при управлении запасами. Речь идёт о моделях многокритериальной оптимизации, применяемых в условиях неопределённости на дискретном множестве решений, которое представляется множеством анализируемых альтернатив.

Цель такой модификации состоит в устранении феномена неадекватного выбора, обусловленного несогласованностью порядка значений имеющих оценок для частных критериев. Особенность предложенного изменения процедуры оптимизации заключается в том, что заданные показатели частных критериев представляются в виде обобщённых данных.

В статье дана численная иллюстрация представленного модифицированного алгоритма. Она соотносится с примером оптимизации стратегии управления запасами, когда в формате модели требуется учитывать фактор неопределённости и концепцию временной ценности денег. Кроме того, в модели дополнительно обеспечивается возможность выбора логистического посредника. При этом оптимизация решения о поставках заказов реализуется в условиях неопределённости спроса и возможных задержек поставок заказов.

Приведённый пример позволяет сделать следующие выводы. Использование обобщённых критериев выбора в формате указанных задач действительно может обеспечить устранение влияния феномена неадекватного выбора, обусловленного несогласованностью порядка значений оценок частных критериев. При этом в формате приведённой численной иллюстрации впервые показано, что предложенная модификация не во всех случаях будет гарантировать исключение воздействия рассматриваемого нежелательного феномена.

Финансирование работы

Исследование выполнено за счёт гранта Российского научного фонда (проект № 24–21–00339, rscf.ru/project/24-21-00339/). Дополнительных грантов на проведение или руководство этим исследованием получено не было.

Конфликт интересов

Авторы заявляют, что у них нет конфликта интересов.

Литература

1. Бродецкий Г. Л., Герами В. Д., Гусев Д. А., Колик А. В. Трансформация цепей поставок в ситуации глобального кризиса // Анализ и прогноз. Журн. ИМЭМО РАН. 2023. № 2. С. 14–23.
2. Мушик Э., Мюллер П. Методы принятия технических решений. М.: Мир, 1990. 206 с.
3. Muschick E., Müller P. H. Entscheidungspraxis: Ziele, Verfahren, Konsequenzen. Berlin: VEB Verl. Technik, 1987. 144 S. [German].
4. Бродецкий Г. Л. Системный анализ в логистике. Выбор в условиях неопределённости. М.: Изд. центр «Академия», 2018. 336 с.
5. Подиновский В. В. Теория важности критериев в многокритериальных задачах принятия решений при неопределённости. I. Исходные положения // Информ. технологии моделир. и управления. 2010. № 5. С. 599–607.
6. Косоруков О. А., Свиридова О. А. Учёт неопределённости спроса при оптимизации системы управления запасами // Логистика. 2012. № 6. С. 12–13.
7. Ben Abdelaziz F., Lang P., Nadeau R. Dominance and efficiency in multi-criteria decision under uncertainty // Theory Decis. 1999. V. 47, No. 3. P. 191–212.
8. Zhukovin V. E., Burshtein F. V., Korelov E. S. A decision making model with vector fuzzy preference relation // Fuzzy Sets Syst. 1987. V. 24, No. 1–2. P. 71–79.
9. Gilboa I. Theory of decision under uncertainty. Cambridge: Camb. Univ. Press, 2009. 214 p.
10. Baydaş M., Pamučar D. Determining objective characteristics of MCDM methods under uncertainty: An exploration study with financial data // Mathematics. 2022. V. 10, No. 7. Paper ID 1115. 25 p.
11. Tzeng G. H., Huang J. J. Multiple attribute decision making: Methods and applications. Boca Raton: CRC Press, 2011. 352 p.
12. Шикин Е. В., Чхартишвили А. Г. Математические методы и модели в управлении. М.: Дело, 2000. 440 с.
13. Ларичев О. И. Теория и методы принятия решений. М.: Университетская лавка; Логос, 2008. 392 с.
14. Ногин В. Д. Принятие решений в многокритериальной среде. Количественный подход. М.: Физматлит, 2004. 176 с.
15. Подиновский В. В. Идеи и методы теории важности критериев в многокритериальных задачах принятия решений. М.: Наука, 2019. 103 с.
16. Thakkar J. J. Multi-criteria decision making. Singapore: Springer, 2021. 390 p. (Stud. Syst. Decis. Control; V. 336).
17. Brodetskiy G. L., Gusev D. A., Shidlovskii I. G. Multi-criteria optimisation under the conditions of uncertainty in logistics and supply chain management // Int. J. Logist. Syst. Manag. 2021. V. 39, No. 2. P. 207–227.
18. Бродецкий Г. Л., Гусев Д. А., Шидловский И. Г. Оптимизация решений по многим критериям в исследованиях логистики. М.: ИНФРА-М, 2020. 284 с.

19. Бродецкий Г. Л., Гусев Д. А., Свиридова О. А., Шидловский И. Г. Возможности процедур оптимизации в условиях неопределённости при их синтезе с выбором решения по многим критериям // Мягкие измерения и вычисления. 2023. № 5. С. 22–36.
20. Brodetskiy G. L. The inventory optimisation taking into account time value of money and order payment deferrals // Int. J. Logist. Syst. Manag. 2017. V. 28, No. 4. P. 486–506.

Приложение А

Для определения оптимальной стратегии управления запасами требуется вспомогательный показатель, характеризующий рентабельность работы рассматриваемой цепи поставок. Это обусловлено тем, что процедуры оптимизации будут реализованы с учётом концепции временной ценности денег (ВЦД) [20].

Оптимальный размер заказа $q_{\text{опт}}$ вычисляется с учётом издержек хранения по фактически занимаемым местам на складе:

$$q_{\text{опт}} = \sqrt{\frac{2C_0D}{C_h + rC_{\Pi}}}, \quad (1)$$

где r — показатель рентабельности работы цепи поставок, который для указанной модели можно найти по формуле [20]

$$r = \frac{(P_{\Pi} - L_{\Pi})\sqrt{\frac{DC_h}{2C_0}} - C_h}{C_{\Pi} + \sqrt{\frac{C_0C_h}{2D}}}, \quad (2)$$

в которой участвуют следующие параметры:

- C_0 — расходы на поставку одной партии товара, причём только такие, которые не зависят от количества поставляемого товара (в числовом примере такой показатель используется в виде C_{01} для первого посредника и в виде C_{02} для второго посредника);
- D — объём годового потребления товара;
- C_h — годовые издержки хранения единицы товара (в числовом примере такой показатель используется в виде C_{h1} для первого посредника и в виде C_{h2} для второго посредника);
- C_{Π} — стоимость единицы товара (в этот показатель принято также включать расходы на поставку, которые зависят от размера заказа);
- P_{Π} — прибыль от реализации единицы товара;
- L_{Π} — требуемые отчисления от прибыли на единицу товара (нужны в модели, чтобы дать менеджерам возможность учитывать специальные расходы бизнеса, пропорциональные обороту товара: например, выплаты сотрудникам и/или страховым организациям, отчисления на хеджирование рисков и т. д.).

Приложение Б

В табл. 2 для альтернативы $X_1(D_1; \Pi_1)$ показатель частного критерия C_1 получен так:

$$C_1 = 17\,000 - C_{01} \frac{D_1}{q_1} = 17\,000 - 2500 \cdot \frac{1600}{323} = 4607,64.$$

При этом размер заказа (q_1) рассчитан по формуле (1), а рентабельность r получена по формуле (2). Показатель частного критерия C_2 получен по формуле

$$C_2 = 2400 - C_{h1} \frac{q_1}{2} = 2400 - 12 \cdot \frac{323}{2} = 463,32,$$

а показатель частного критерия C_3 равен $90\,000 - 0 = 90\,000$, поскольку у первого посредника нет задержек. Аналогично показатель частного критерия C_4 равен $30\,000 - 0 = 30\,000$, так как при реализации события θ_1 не наблюдается ошибок в прогнозировании. Действительно, размер заказа q_1 рассчитывается при ожидании спроса D_1 , а событие θ_1 предусматривает реализацию спроса D_1 . Показатель частного критерия C_5 равен $500 - 100 = 400$.

Также отметим, что если фактически реализованный спрос больше, чем предельно допустимый в модели применительно к сценарию стратегии поставок, то вычисляется упущенная прибыль в размере недопоставленной продукции. Если же фактически реализованный спрос оказывается меньше, то рассчитываются потери в размере половины упущенной прибыли и половины годового хранения избыточного количества поставленного товара.

Гусев Денис Александрович
Свиридова Ольга Александровна
Шидловский Иван Геннадьевич
Бродецкий Геннадий Леонидович

Статья поступила
5 июля 2024 г.
После доработки —
18 августа 2024 г.
Принята к публикации
22 сентября 2024 г.

OPTIMIZATION OF INVENTORY MANAGEMENT STRATEGIES
FOR ORDER DELIVERIES USING MULTICRITERIA DECISION
MAKING UNDER CONDITIONS OF UNCERTAINTY

*D. A. Gusev^{1, a}, O. A. Sviridova^{1, b},
I. G. Shidlovskii^{1, c}, and G. L. Brodetskiy^{2, d}*

¹ Plekhanov Russian University of Economics,
Stremyanny Lane, 36, 115054 Moscow, Russia

² Graduate School of Business, HSE University
Shabolovka Street, 26, Bld. 1, 119049 Moscow, Russia

E-mail: ^agussev79@mail.ru, ^bsviridova.oa@rea.ru,
^cshdlvsk-ivan@yandex.ru, ^dgbrodetskiy@hse.ru

Abstract. In today's inventory management, there is a need to optimize new models in order to select the best solution to ensure order fulfillment. These models must take into account various uncertainties, as well as the concept of time value of money. At the same time, it is important to note that in practice, appropriate solutions are often multicriteria due to the complex nature of supply chains.

To address this, the authors have developed a method for optimizing such tasks based on a combination of multicriteria optimization procedures and decision making under uncertainty. However, this approach is shown to require further refinement in practice. The proposed refinement aims to assist managers in avoiding undesirable outcomes related to alternative selection. These are situations related to phenomena that lead to the selection of alternatives that may not be optimal for the preferences of the decision maker.

The corresponding adjustment would result in introduction of a specific feature in the form of optimization procedures. In such an adjustment, it is proposed to present initial indicators of given specific criteria on the basis of so-called aggregated data, which eliminates the factor of dimensionality. A numerical example is provided using the development of a stock management strategy as an example, considering the need to

select a logistics intermediary and in conditions of uncertainty regarding demand and potential delays in delivery. In this scenario, a weighted average of estimates of specific criteria is used as the selection criterion and the Hurwitz criterion is employed to account for uncertainty. Tab. 7, bibliogr. 20.

Keywords: inventory management, choice under uncertainty, demand uncertainty, delivery delay, multicriteria optimization, generalized selection criteria, phenomena of inadequate choice.

References

1. G. L. Brodetskiy, V. D. Gerami, D. A. Gusev, and A. V. Kolik, Supply chain transformation amid a global crisis, *Anal. Progn., Zh. IMEMO RAN*, No. 2, 14–23 (2023) [Russian].
2. E. Muschick and P. H. Müller, *Method for Making Technical Decisions* (Mir, Moscow, 1990) [Russian].
3. E. Muschick and P. H. Müller, *Entscheidungspraxis: Ziele, Verfahren, Konsequenzen* (VEB Verl. Technik, Berlin, 1987) [German].
4. G. L. Brodetskiy, *Systems Analysis in Logistics. Choice Under Uncertainty* (Izd. Tsentr “Akademiya”, Moscow, 2018) [Russian].
5. V. V. Podinovski, Theory of criteria importance in multicriteria decision making problems under uncertainty. I. Initial provisions, *Inf. Tekhnol. Model. Upr.*, No. 5, 599–607 (2010) [Russian].
6. O. A. Kosorukov and O. A. Sviridova, Taking into account the demand uncertainty while optimizing the inventory management, *Logistika*, No. 6, 12–13 (2012) [Russian].
7. F. Ben Abdelaziz, P. Lang, and R. Nadeau, Dominance and efficiency in multicriteria decision under uncertainty, *Theory Decis.* **47** (3), 191–212 (1999).
8. V. E. Zhukovin, F. V. Burshtein, and E. S. Korelov, A decision making model with vector fuzzy preference relation, *Fuzzy Sets Syst.* **24** (1–2), 71–79 (1987).
9. I. Gilboa, *Theory of Decision Under Uncertainty* (Camb. Univ. Press, Cambridge, 2009).
10. M. Baydaş and D. Pamučar, Determining objective characteristics of MCDM methods under uncertainty: An exploration study with financial data, *Mathematics* **10** (7), ID 1115 (2022).
11. G. H. Tzeng and J. J. Huang, *Multiple Attribute Decision Making: Methods and Applications* (CRC Press, Boca Raton, 2011).
12. E. V. Shikin and A. G. Chkhartishvili, *Mathematical Methods and Models in Management* (Delo, Moscow, 2000) [Russian].
13. O. I. Larichev, *Theory and Methods of Decision Making* (Univ. Lavka; Logos, Moscow, 2008) [Russian].
14. V. D. Nogin, *Decision Making in a Multicriteria Environment. A Quantitative Approach* (Fizmatlit, Moscow, 2004) [Russian].

15. **V. V. Podinovski**, *Ideas and Methods of the Theory of Criteria Importance in Multicriteria Decision-Making Problems* (Nauka, Moscow, 2019) [Russian].
16. **J. J. Thakkar**, *Multi-Criteria Decision Making* (Springer, Singapore, 2021) (Stud. Syst. Decis. Control, Vol. 336).
17. **G. L. Brodetskiy, D. A. Gusev, and I. G. Shidlovskii**, Multi-criteria optimisation under the conditions of uncertainty in logistics and supply chain management, *Int. J. Logist. Syst. Manag.* **39** (2), 207–227 (2021).
18. **G. L. Brodetskiy, D. A. Gusev, and I. G. Shidlovskii**, *Optimizing Multi-Criteria Decisions in Logistics Research* (INFRA-M, Moscow, 2020) [Russian].
19. **G. L. Brodetskiy, D. A. Gusev, O. A. Sviridova, and I. G. Shidlovskii**, Possibilities of optimization procedures under uncertainty and their synthesis with solution choice by many criteria, *Myagkie Izmer. Vychisl.*, No. 5, 22–36 (2023) [Russian].
20. **G. L. Brodetskiy**, The inventory optimisation taking into account time value of money and order payment deferrals, *Int. J. Logist. Syst. Manag.* **28** (4), 486–506 (2017).

Denis A. Gusev

Olga A. Sviridova

Ivan G. Shidlovskii

Gennady L. Brodetskiy

Received July 5, 2024

Revised August 18, 2024

Accepted September 22, 2024

ГОСУДАРСТВЕННО-ЧАСТНОЕ ПАРТНЁРСТВО
В ИНФРАСТРУКТУРНЫХ ПРОЕКТАХ СЫРЬЕВОЙ
ТЕРРИТОРИИ: МОДЕЛЬ НА ОСНОВЕ
КОНСОРЦИУМА НЕДРОПОЛЬЗОВАТЕЛЕЙ

*А. О. Зырянов^a, С. М. Лавлинский^b,
А. А. Панин^c, А. В. Плясунов^d*

Институт математики им. С. Л. Соболева,
пр. Акад. Коптюга, 4, 630090 Новосибирск, Россия

E-mail: ^aalexander.zyryanov44@gmail.com, ^blavlin@math.nsc.ru,
^caapanin1988@gmail.com, ^dapljas@math.nsc.ru

Аннотация. Предложена модель для разработки планов производственно-инфраструктурного строительства, использующая специальный механизм государственно-частного партнёрства. В его основе — формирование набора кластеров месторождений, в каждом из которых государством создаётся консорциум частных инвесторов, заинтересованных в освоении месторождений кластера. Функционал такого объединения — совместное строительство необходимой для добычи полезных ископаемых транспортной и энергетической инфраструктуры. Механизм достижения компромисса интересов государства и частных инвесторов использует модель Штакельберга, в которой лидером выступает государство. Оно определяет квоты компенсаций затрат консорциумов на реализацию инфраструктурных проектов. Роль последователя отведена системе консорциумов, формирующей программу инфраструктурного строительства. Решение соответствующей двухуровневой задачи математического программирования позволяет сформировать адресный план развития, составляющими которого являются программа инфраструктурного строительства в консорциумах, список запускаемых производственных проектов, раздел инфраструктурных затрат и компенсационных трансфертов для частных инвесторов. Показано, что задача государства Σ_2^P -трудна и принадлежит классу $\Sigma_2^P O$, если переменные, определяющие график трансфертов, принимают только целые значения. Для решения двухуровневой задачи на основе метаэвристик разработан новый стохастический приближённый гибридный алгоритм, использующий идеи покоординатного спуска.

Проведены численные эксперименты для комплекса месторождений Забайкальского края. Результаты имеют практическое значение для формирования механизма компенсаций в долевом строительстве производственной инфраструктуры. Ил. 8, библиогр. 20.

Ключевые слова: двухуровневая задача математического программирования, модель Штакельберга, метаэвристика, покоординатный спуск, полиномиальная иерархия, аппроксимационная иерархия, государственно-частное партнёрство, производственно-инфраструктурный проект, консорциум недропользователей.

Введение

Проблема формирования программы недропользования в ресурсном регионе с недостаточным уровнем развития транспортной и энергетической инфраструктуры актуальна для большинства российских ресурсных территорий. Здесь инвестор не может запустить проекты освоения месторождений из-за отсутствия дорог, ЛЭП, а государство не готово строить инфраструктурные объекты без полной уверенности в их достаточной загрузке. В большинстве сибирских и дальневосточных регионов такое положение дел сдерживает развитие региональных экономик и стимулирует разработку механизмов достижения компромисса интересов участвующих сторон.

В контексте стратегического планирования ключевая роль в процессе принятия решения отведена согласованию перечня и сроков запуска проектов освоения месторождений минерального сырья с планами создания необходимой, но отсутствующей на сегодняшний день производственной инфраструктуры. Такую инфраструктуру предполагается строить в рамках коллаборации, общие контуры которой укладываются в классическую схему государственно-частного партнёрства (ГЧП) [1–3], но сам механизм ещё предстоит определить. В сегодняшних реалиях управленческий корпус минерально-сырьевого сектора ставит во главу угла политические аргументы и использует, в основном, простейшие оценки эффективности принимаемых решений на основе анализа текущих рыночных цен и затратных характеристик технологических проектов.

Опыт реализации самых крупных проектов такого рода [4–7] негативно оценивается экспертным сообществом и говорит о том, что программа освоения недр должна не только определить перечень мер, способствующих территориальному развитию, но и ответить на следующие ключевые вопросы, стоящие перед частным инвестором, потенциально заинтересованным в освоении природных богатств территории. Как помочь инвестору в освоении месторождений труднодоступных районов без какой-либо производственной инфраструктуры? Как оптимально построить институт стимулирования частных инвестиций, обеспечивающий

комплементарный эффект различных мер инвестиционной политики государства на территории? Как в организационном плане гармонизировать цели и объединить возможности недропользователей?

Основное внимание в статье направлено на эти проблемы. Дополнив классическую модель ГЧП механизмом консорциума, авторы в существенной степени расширяют ранее исследованные модели партнёрства [8–10] за счёт дополнения их горизонтальными связями частных инвесторов и эффектами от консолидации ресурсов. Принципиально важно, что механизм партнёрства и программа освоения приобретают в этих условиях максимально конкретный и адресный характер, который уже может быть использован напрямую в практике управления.

1. Математическая модель

Для большинства российских ресурсных регионов весь набор месторождений может быть представлен в виде системы непересекающихся территориальных кластеров, границы которых обусловлены рельефом местности и водоразделами рек. В основе предлагаемой модели лежит представление о том, что необходимая для запуска проектов недропользования производственная инфраструктура строится системой консорциумов, территориально привязанных к этим кластерам. Для каждого консорциума государство создаёт управляющую компанию (УК). Функционал УК — организация и координация долевого финансирования процесса создания необходимых инфраструктурных объектов частными инвесторами — членами консорциума, а также генерация графика компенсационных бюджетных выплат. Управляющая компания берёт на себя полный перечень государственных функций в основных процедурах согласования интересов: организация экологического контроля, проведение всесторонней предпроектной экспертизы, осуществление мониторинга хода реализации проекта и пр. Итогом деятельности управляющих компаний является развёрнутая комплексная программа освоения минерально-сырьевой базы территории. Она определяет, какие из инфраструктурных объектов будут построены, перечень запускаемых проектов освоения месторождений в каждом консорциуме (кластере), графики затрат на создание необходимой производственной инфраструктуры и их бюджетного возмещения для каждого из инвесторов.

Такая схема жизнеспособна и интересна для бизнеса лишь при условии достижения частным инвестором необходимого уровня рентабельности в совокупном проекте освоения месторождения и строительства части необходимой инфраструктуры. Для государства она привлекательна, во-первых, с точки зрения динамики затрат на развитие территории: в отличие от сегодняшней практики, бюджетные затраты на компенсации могут быть отнесены на период, когда уже начинают поступать налоги.

Во-вторых, государство получает инструмент регулирования для распределения природно-ресурсной ренты, существенно более гибкий, чем рентные налоги, назначая для высокорентабельных месторождений большую долю инфраструктурных затрат в консорциуме. Дело за малым — определить параметры такого партнёрства (программы освоения), обеспечивающие экономическую заинтересованность потенциальных недропользователей и максимально возможный объём налоговых платежей.

1.1. Модель консорциума. Как может быть описан процесс деятельности управляющей компании консорциума в части координации процессов добычи, долевого строительства инфраструктуры и согласования интересов в процедурах взаимных расчётов?

Управляющая компания консорциума, как субъект управления проектом ГЧП, координирует процесс нахождения компромисса интересов участников, ориентируясь на совокупность целей и особенности экономической среды. Важную роль здесь играют транзакционные издержки (ТИ) [11–13], существенно влияющие на процессы, связанные с образованием и деятельностью консорциума. По своей природе такой специализированный консорциум должен решать сложные, оптимизационные по своей сути задачи. Это обстоятельство существенно отличает его от сегодняшних дирекций вышеупомянутых мегапроектов и программ, инструментарий принятия управленческих решений которых не в полной мере соответствует уровню поставленных перед ними задач.

Транзакционные издержки включают в себя не только расходы по заключению контрактов, но и затраты, появляющиеся в процессе взаимодействия экономических агентов. Для случая масштабных сырьевых проектов транзакционные издержки в неосвоенных районах могут оказаться достаточно большими. Управляющая компания, как представитель собственника (государства), несёт ТИ в части контроля, мониторинга, затрат на создание институциональной инфраструктуры. В рамках консорциума транзакционные издержки инвестора связаны с разработкой оценки воздействия на окружающую среду, издержками на содержание служб, взаимодействующих с контролирующими органами, и т. п.

Практика экономического анализа различает транзакционные издержки до (ТИ-до) и транзакционные издержки после (ТИ-после) старта проекта [11]. Это связано с различиями в подходах к управлению природными ресурсами до и после начала реализации проектных мероприятий. Особенности институциональной среды также влияют на процесс формирования издержек, определяя то, каким образом эти издержки распределяются между правительством и бизнесом.

Традиционный подход к исчислению транзакционных издержек позволяет предположить, что в рамках консорциума общая величина ТИ

управляющей компании и частных инвесторов пропорциональна объёму капиталовложений проекта. Практика показывает, что этот факт наблюдается как у инфраструктурного проекта, так и у проекта освоения месторождения.

Та же практика позволяет уловить характер динамики ТИ — рост ТИ-до на протяжении T_1 лет к моменту старта и постепенный спад ТИ-после до некоторого уровня, который определяется условиями контракта. Кроме того, мы будем фиксировать долю ТИ, включаемых в себестоимость. Вкупе все сделанные предположения дают возможность алгоритмически определить величину транзакционных издержек и представить в агрегированном виде характер их изменения во времени.

1.2. Модель формирования ГЧП. Введём следующие обозначения:

- $T = \{-T_1, \dots, 0, 1, \dots, T_2\}$ — временной горизонт, T_2 — длительность периода планирования;
- T_0 — временной лаг, определяющий начальный момент компенсационных выплат из бюджета;
- I — множество производственных (добычных) проектов;
- J — множество инфраструктурных проектов;
- K — множество консорциумов.

Предполагается, что у каждого инвестора одно месторождение, которое он может осваивать.

Для производственного проекта $i \in I$ определим

- DBP_i^t — график налоговых платежей, возникающих в процессе реализации проекта;
- CFP_i^t — кэшфло проекта (разница доходов и расходов);
- MTCR_i^t — транзакционные издержки управляющей компании консорциума, членом которого является инвестор, на всём временном горизонте проекта: в периоды подготовки (ТИ-до, $t = -T_1, \dots, 0$) и реализации (ТИ-после, $t = 1, \dots, T_2$);
- ITCR_i^t — транзакционные издержки инвестора на всём временном горизонте проекта: в периоды подготовки (ТИ-до, $t = -T_1, \dots, 0$) и реализации (ТИ-после, $t = 1, \dots, T_2$).

Для инфраструктурного проекта $j \in J$ определим

- ZI_j^t — график капитальных вложений по проекту;
- MTCI_j^t — транзакционные издержки управляющей компании консорциума, реализующего проект;
- ITCI_j^t — совокупные транзакционные издержки инвесторов, участвующих совместно в проекте;
- VDI_j^t — бюджетные доходы, порождённые мультипликативным воздействием проекта на экономику территории.

На этапе подготовки проектов ($t = -T_1, \dots, 0$) параметры модели DBP_i^t , CFP_i^t , ZI_j^t и VDI_j^t предполагаются нулевыми.

Взаимосвязь проектов $i \in I$ и $j \in J$:

- μ_{ij} — булева переменная, фиксирующую необходимость реализации инфраструктурного проекта $j \in J$ для запуска производственного проекта $i \in I$,

$$\mu_{ij} = \begin{cases} 1, & \text{если реализация проекта } i \in I \text{ невозможна без реализации проекта } j \in J, \\ 0 & \text{иначе.} \end{cases}$$

Наборы производственных и инфраструктурных проектов разбиты на консорциумы непересекающимся образом и соответствуют исходной кластеризации территории.

Распределение проектов и инвесторов по консорциумам:

$$\alpha_{kj} = \begin{cases} 1, & \text{если инфраструктурный проект } j \\ & \text{принадлежит консорциуму } k, \\ 0 & \text{иначе;} \end{cases}$$

$$\beta_{ki} = \begin{cases} 1, & \text{если производственный проект } i \\ & \text{принадлежит консорциуму } k, \\ 0 & \text{иначе.} \end{cases}$$

Бюджетные ограничения и дисконты:

- BudG^t , DG — бюджет и дисконт государства;
- BudI_i^t , DI — бюджет и дисконт инвестора.

Целочисленные переменные модели:

$$z_i = \begin{cases} 1, & \text{если производственный проект } i \text{ запускается инвестором этого проекта,} \\ 0 & \text{иначе;} \end{cases}$$

$$c_j = \begin{cases} 1, & \text{если инфраструктурный проект } j \text{ реализуется некоторым консорциумом,} \\ 0 & \text{иначе.} \end{cases}$$

Вещественные переменные:

- $\overline{W}_k^t$ — предложенный государством трансферт на компенсацию инфраструктурных затрат недропользователей и накладные расходы УК консорциума k ;
- W_k^t — реализованный государством трансферт на компенсацию инфраструктурных затрат недропользователей и накладные расходы УК консорциума k ;
- R_i^t — график компенсационных выплат частному инвестору i ;

- D_{ij} — доля инвестора i в совокупных затратах на строительство объекта инфраструктуры j ;
- $\Theta > 0$ — некоторое большое число.

Модель формирования ГЧП формулируется как следующая задача двухуровневого математического программирования.

Задача государства ($\widetilde{\text{PG}}$). Максимизировать дисконтированный поток бюджетных доходов

$$\sum_{t \in T} \left(\sum_{i \in I} \text{DBP}_i^t z_i + \sum_{j \in J} \text{VDI}_j^t c_j - \sum_{k \in K} W_k^t \right) / (1 + \text{DG})^t \rightarrow \max_{\overline{W}, z, c, W, R, D} \quad (1)$$

при условиях

$$\overline{W}_k^t \geq 0, \quad t \in T, k \in K, \quad (2)$$

$$\sum_{k \in K} \overline{W}_k^t \leq \text{BudG}^t, \quad t \in T, \quad (3)$$

$$(z, c, W, D, R) \in \mathcal{F}^*(\overline{W}), \quad (4)$$

где $\mathcal{F}^*$ — множество оптимальных решений для задачи консорциумов:

$$\begin{aligned} & \gamma \sum_{t \in T} \sum_{i \in I} \left((\text{CFP}_i^t - \text{ITCP}_i^t) z_i - \sum_{j \in J} (\text{ZI}_j^t + \text{ITCI}_j^t) D_{ij} + R_i^t \right) / (1 + \text{DI})^t + \\ & + (1 - \gamma) \sum_{t \in T} \left(\sum_{i \in I} \text{DBP}_i^t z_i + \sum_{j \in J} \text{VDI}_j^t c_j - \sum_{k \in K} W_k^t \right) / (1 + \text{DG})^t \rightarrow \\ & \rightarrow \max_{z, c, W, R, D} \end{aligned} \quad (5)$$

при условиях

$$W_k^t \leq \overline{W}_k^t, \quad t \in T, k \in K, \quad (6)$$

$$\sum_{t \in T} \left(\sum_{i \in I} \text{DBP}_i^t z_i \beta_{ki} + \sum_{j \in J} \text{VDI}_j^t c_j \alpha_{kj} - W_k^t \right) / (1 + \text{DG})^t \geq 0, \quad k \in K, \quad (7)$$

$$R_i^t = 0, \quad -T_1 \leq t \leq T_0, \quad i \in I, \quad (8)$$

$$R_i^t \geq 0, \quad T_0 + 1 \leq t \leq T_2, \quad i \in I, \quad (9)$$

$$R_i^t \leq \Theta z_i, \quad T_0 + 1 \leq t \leq T_2, \quad i \in I, \quad (10)$$

$$R_i^t \leq \Theta \sum_{j \in J} D_{ij} \text{ZI}_j^t, \quad T_0 + 1 \leq t \leq T_2, \quad i \in I, \quad (11)$$

$$\sum_{i \in I} R_i^t \beta_{ki} + \sum_{j \in J} \text{MTCI}_j^t c_j \alpha_{kj} + \sum_{i \in I} \text{MTCP}_i^t z_i \beta_{ki} \leq W_k^t, \quad k \in K, t \in T, \quad (12)$$

$$0 \leq D_{ij} \leq \mu_{ij} z_i, \quad i \in I, j \in J, \quad (13)$$

$$\sum_{i \in I} D_{ij} \beta_{ki} = \alpha_{kj} c_j, \quad k \in K, j \in J, \quad (14)$$

$$\sum_{t \in T} \sum_{i \in I} ((\text{CFP}_i^t - \text{ITCP}_i^t) z_i - (\text{ZI}_j^t + \text{ITCI}_j^t) D_{ij} + R_i^t) / (1 + \text{DI})^t \geq 0, \quad i \in I, \quad (15)$$

$$c_j \geq \mu_{ij} z_i, \quad i \in I, j \in J, \quad (16)$$

$$\sum_{1 \leq t \leq \omega} \left(-(\text{CFP}_i^t - \text{ITCP}_i^t) z_i + \sum_{j \in J} (\text{ZI}_j^t + \text{ITCI}_j^t) D_{ij} - R_i^t \right) \leq \sum_{1 \leq t \leq \omega} \text{BudI}_i^t, \quad \omega \in T, i \in I. \quad (17)$$

В сформулированной модели управляющие компании стремятся достичь компромисса интересов партнёров. Распределение инфраструктурных затрат и компенсаций государства происходит так, чтобы, с одной стороны, сгладить стремление инвесторов к получению максимально возможной ренты, а с другой, принять во внимание тот факт, что государство является собственником недр (5). С помощью параметра γ в модели фиксируется уровень значимости интересов инвесторов. Соотношения (13)–(15) определяют, каким образом затраты на инфраструктуру (с учётом трансакционных издержек) распределяются внутри консорциума. Инфраструктурные ограничения добычи минерального сырья определяются условием (16): производственный проект не реализуется при отсутствии необходимой производственной инфраструктуры. Бюджетные ограничения (2) и (17) имеют различную степень «жёсткости»: недропользователь может задействовать в следующем году финансовые средства, не израсходованные в текущем году.

Компенсации инвесторам начинают поступать с года $T_0 + 1$ ((8), (9)). Ограничение на бюджет управляющей компании (12) учитывает, что она несёт трансакционные издержки, возникающие со стороны государства. Поток выплат недропользователям должен не только гарантировать бюджету баланс доходов всех видов и трансфертов, направляемых консорциумам (7), но и обеспечить рентабельное возмещение затрат инвестора на инфраструктуру (15).

Необходимым условием образования консорциума является прирост целевой функции государства (7) в результате формирования соответствующей организационной структуры. В такой ситуации единый консорциум, включающий весь набор месторождений, проигрывает совокупности нескольких консорциумов, сформированных в соответствии с кластеризацией территории.

Представленная модель генерирует структуру партнёрства и соответствующую ей программу развития территории $\{c_j, z_i, R_i^t, D_{ij}\}$, в которой определён не только перечень запускаемых проектов, но и сформированы детали раздела затрат в долевом строительстве и стратегия компенсационных выплат.

2. Вопросы сложности

В данной статье рассматривается новая двухуровневая модель взаимодействия государства и частных инвесторов, использующая специальный механизм государственно-частного партнёрства с консорциумом. Структурные особенности такой модели не позволяют воспользоваться оценками вычислительной сложности ранее разработанных двухуровневых моделей ГЧП [8–10]. Оценка сложностного статуса рассматриваемой модели — практический путь уточнения класса эффективных алгоритмов для задач большой размерности. Полученные далее результаты объясняют, почему в работе авторы не пошли по пути разработки для задачи государства точных или приближённых полиномиальных алгоритмов с оценками отклонения полученного решения от оптимального.

Свяжем с задачей государства $\widehat{PG}$ стандартную задачу распознавания $D(\widehat{PG})$, в которой входными данными являются исходные данные задачи государства и произвольное рациональное число U . В задаче $D(\widehat{PG})$ мы должны решить, существует ли допустимое решение со значением целевой функции, большим или равным U . Задача $D(\widehat{PG})$ принадлежит классу Σ_2^P , если существует недетерминированная оракульная машина Тьюринга, распознающая задачу $D(\widehat{PG})$ за полиномиальное время, используя в качестве оракула некоторый язык из класса NP. Класс Σ_2^P относится ко второму уровню полиномиальной иерархии [14]. Задача оптимизации принадлежит классу NPO, если её стандартная задача решения принадлежит классу NP [14]. Аналогично класс Σ_2^PO содержит оптимизационные задачи, для которых соответствующая стандартная задача распознавания принадлежит классу Σ_2^P [14]. Класс NPO относится к первому уровню иерархии аппроксимации, а класс Σ_2^PO — ко второму уровню этой иерархии.

Покажем, что задача государства Σ_2^P -трудна для случая $\gamma = 1$, сведя Σ_2^P -полную задачу о сумме подмножеств на интервале (subset-sum-interval problem) к рассматриваемой задаче [15].

Задача о сумме подмножеств на интервале. Заданы целые положительные числа $q_1, \dots, q_l$, R и $r \leq l$. Спрашивается, найдётся ли целое число S такое, что $R \leq S < R + 2^r$ и не существует подмножеств $I \subseteq \{1, \dots, l\}$, удовлетворяющих равенству $\sum_{i \in I} q_i = S$.

Теорема 1. Задача $\widetilde{\text{PG}} \Sigma_2^P$ -трудна.

ДОКАЗАТЕЛЬСТВО. Сведём указанную задачу к задаче $\widetilde{\text{PG}}$, построив следующий вход задачи государства. Рассмотрим один консорциум (положим $\text{NC} = 1$), $\text{NP} = l + \lceil \log_2(R + 2^r) \rceil + 1$ производственных проектов и один общий для всех этих проектов инфраструктурный проект (положим $\text{NI} = 1$). Так как $\text{NI} = \text{NC} = 1$, имеем $\alpha_{11} = 1$, $\mu_{i1} = 1$ и $\beta_{1i} = 1$ для всех производственных проектов. Горизонт планирования составляет 3 года ($T_1 = T_0 = 0$, $T_2 = 3$). Положим $\text{DG} = \text{DI} = 0$.

Для первых l производственных проектов имеем

$$\text{CFP}_i^1 = 0, \quad \text{CFP}_i^2 = -2q_i, \quad \text{CFP}_i^3 = q_i.$$

Для проектов i , $l + 1 \leq i \leq l + \lceil \log_2(R + 2^r) \rceil$, положим

$$\text{CFP}_i^1 = 0, \quad \text{CFP}_{l+1}^2 = -2^{i-l}, \quad \text{CFP}_i^3 = 2^{i-l-1} - \varepsilon,$$

где $\varepsilon = 1/(\lceil \log_2(R + 2^r) \rceil + 1)$. Кроме того,

$$\text{CFP}_{\text{NP}}^1 = \text{CFP}_{\text{NP}}^2 = 0, \quad \text{CFP}_{\text{NP}}^3 = \Delta,$$

где $\Delta > 0$ — достаточно большое число, например $\Delta = 2(R + 2^r) + \sum_i q_i$.

Доходы государства от всех проектов в первые два года и для первых l проектов в третьем году равны 0. В графике налоговых платежей для проектов i , $l < i \leq \text{NP}$, положим $\text{DBP}_i^3 = \Delta$.

Для инфраструктурного проекта определим следующий график капитальных вложений:

$$\text{ZI}_1^1 = \sum_i q_i, \quad \text{ZI}_1^2 = \text{ZI}_1^3 = 0.$$

Бюджет государства по годам равен

$$\text{BudG}^1 = 2^r - 1, \quad \text{BudG}^2 = 2(R + 2^r - 1) + 1, \quad \text{BudG}^3 = 0.$$

Инвесторы имеют ненулевой бюджет только в первый год:

$$\begin{aligned} \text{BudI}_i^1 &= q_i, \quad i \leq l, \\ \text{BudI}_i^1 &= 2^{(i-l-1)}, \quad l < i < \text{NP}, \\ \text{BudI}_{\text{NP}}^1 &= \sum_i q_i - R - 2^r + 1. \end{aligned}$$

Все остальные параметры производственных (ITCP_i^t , MTCP_i^t) и инфраструктурных (VDI_j^t , MTCI_j^t , ITCI_j^t) проектов равны 0.

Структура данных задачи такова, что государство и консорциум будут стремиться реализовать инфраструктурный и последний инвестиционный проекты для получения дохода Δ . Однако инвестору NP для инфраструктурного проекта не хватает $R + 2^r - 1$ бюджетных средств.

Государство в первый год может дать $\overline{W}_1^1 \leq 2^r - 1$ бюджетных средств. Таким образом формируется дефицит бюджета $S = R + 2^r - 1 - \overline{W}_1^1$, который консорциум будет стремиться покрыть за счёт бюджетов других инвесторов, обеспечивая реализацию их инвестиционных проектов, которые не выгодны ни инвесторам, ни консорциуму. Это означает, что необходимо минимизировать остаток от бюджетов привлекаемых инвесторов.

Учитывая, что $R \leq S \leq R + 2^r - 1$, такой минимум легко найти, так как бюджеты инвесторов с номерами от $l + 1$ по $l + \lceil \log_2(R + 2^r) \rceil$ покрывают любое число в данном интервале, а следовательно, искомый минимум равен 0. При уменьшении доходов этих инвесторов в последний год на ε они становятся наименее привлекательны для консорциума. Если консорциум может покрыть величину S числами q_i , то он вообще не будет привлекать этих выгодных для государства клиентов. В противном случае привлечёт ровно одного, что даст государству дополнительный доход в Δ . Кроме того, бюджет государства во второй год позволяет не только компенсировать затраты инвесторов в первый год, но и выйти на неотрицательную прибыль для инвесторов с номерами от l до $l' = l + \lceil \log_2(R + 2^r) \rceil$, поэтому

$$\overline{W}_1^2 = 2S + \varepsilon \sum_{i=l+1}^{l'} z_i.$$

Таким образом, государство будет так выбирать $\overline{W}_1^1$, чтобы S не покрывалось числами q_i , если это возможно. Теорема 1 доказана.

Переменные $\overline{W}_j^t$ и W_j^t в задаче государства $\widetilde{\text{PG}}$ непрерывны. Это обстоятельство является причиной, которая не позволяет связать данную задачу с классом $\Sigma_2^P \text{O}$. Однако если предположить, что эти переменные целочисленны, то выполняется

Теорема 2. *Задача $\widetilde{\text{PG}}$ принадлежит классу $\Sigma_2^P \text{O}$.*

ДОКАЗАТЕЛЬСТВО может быть легко получено модификацией доказательства теоремы 3 из работы [9].

Теорема 3. *Если $\text{NPO} \neq \Sigma_2^P \text{O}$, то задача государства не может лежать в классах PO , FPTAS , PTAS , APX , Log-APX , Poly-APX , Exp-APX .*

ДОКАЗАТЕЛЬСТВО следует из определения аппроксимационной иерархии [16].

Полученные результаты оценки сложностного статуса рассматриваемой модели делают бессмысленными попытки построить для задачи государства точные или приближённые полиномиальные алгоритмы с теми

или иными оценками уклонения полученного решения от оптимального. Именно поэтому для решения задачи строится стохастический приближённый гибридный алгоритм, использующий локальный поиск и идеи метода покоординатного спуска [17].

Алгоритм вначале выбирает стартовое решение $\{W_k^t\}$, для получения которого решается одноуровневая задача с целевой функцией (1) и ограничениями верхнего и нижнего уровней (2), (3), (6)–(17). Стартовое решение становится текущим рекордом, который в ходе работы алгоритма будет улучшаться. Критерием остановки алгоритма является проведение I итераций подряд без улучшения рекорда. Локальный поиск идёт по значениям переменных текущего рекорда $\{\overline{W}_j^t\}$. На каждой итерации случайно выбирается заданная доля переменных, чьи значения поочерёдно варьируются. Для каждой из выбранных переменных выбирается N' случайных значений из интервала от 0 до той доли годового бюджета государства, которая остаётся свободной, используя равномерное распределение. Это означает, что любая выбранная переменная W_j^t определяет для текущего рекорда $\{\overline{W}_j^t\}$ окрестность из N' соседей. Для каждого пробного значения при условии, что выполнены ограничения (2), решается задача нижнего уровня и вычисляется значение целевой функции государства на найденном допустимом решении. Если эта величина больше текущего рекорда, то обновляем текущее значение переменной и рекорд. Иначе оставляем всё без изменений.

Алгоритм

ШАГ 1. Выбрать стартовое решение государства $\{W_k^t\}$; текущее лучшее решение $\{\overline{W}_k^t\}$ положить равным $\{W_k^t\}$; за рекорд принять значение целевой функции государства при $\{\overline{W}_k^t\}$, $k \in K$, $t \in T$.

ШАГ 2. Присвоить $i := 0$.

ШАГ 2.1. Если $i < I$, то присвоить $i := i + 1$, $k' := 1$, $t' := 1$, иначе перейти на шаг 3.

ШАГ 2.2. С вероятностью p перейти на шаг 2.2.1, иначе перейти на шаг 2.3.

ШАГ 2.2.1. Присвоить $N := 1$.

ШАГ 2.2.2. Выбрать значение переменной $\widehat{W}_{k'}^{t'}$ из интервала от 0 до $\text{BudG}^{t'} - \sum_{k \in K \setminus \{k'\}} \overline{W}_k^{t'}$, используя равномерное распределение.

ШАГ 2.2.3. Если при $\widehat{W}_k^t$ значение целевой функции государства выше рекорда, то обновить рекорд и присвоить $\overline{W}_k^t := \widehat{W}_k^t$, $i := 0$.

ШАГ 2.2.4. Если $N < N'$, то присвоить $N := N + 1$ и перейти на шаг 2.2.2.

ШАГ 2.3. Если $t' < T_2$, то присвоить $t' := t' + 1$ и перейти на шаг 2.2, иначе если $k' < \text{НС}$, то присвоить $k' := k' + 1$, $t' := 1$ и перейти на шаг 2.2, иначе перейти на шаг 2.1.

ШАГ 3. Вернуть наилучшее найденное решение $\{\overline{W}_k^t\}$ и рекордное значение целевой функции лидера.

Данный приближённый стохастический алгоритм отличается от классического метода покоординатного спуска следующими чертами.

1. На каждой итерации используется только часть переменных, которые выбираются случайным образом.

2. В отличие от классического метода, в котором рассматривается только два значения для каждой переменной (одно получено сдвигом с заданной длиной шага в направлении соответствующего координатного вектора, а другое получено сдвигом с тем же шагом в противоположном направлении), в предлагаемом алгоритме для любой выбранной переменной, которая рассматривается как равномерно распределённая на интервале от 0 до годового бюджета государства случайная величина, порождается N' значений.

3. Численный эксперимент

В качестве модельного полигона в экспериментах рассматриваются пять попарно не пересекающихся кластеров месторождений полиметаллов, 20 производственных и 10 инфраструктурных проектов, прообразами которых являются инфраструктурные проекты из Программы 2030 Забайкальского края: автодороги, линии электропередач, железная дорога Нарын — Лугокан. Полигон сконструирован таким образом, что запуск всех инфраструктурных проектов открывает возможность реализации всех производственных проектов.

Для двухуровневой задачи планирования база данных описывает модельный полигон и формируется на основе прогнозных имитационных моделей реализации проектов всех видов — добычи полезных ископаемых, строительства энергетической инфраструктуры, создания транспортных коммуникаций, предусмотренных Стратегией развития 2030 Забайкальского края [9].

Свойства равновесия по Штакельбергу для построенного полигона изучаются в ходе численного эксперимента, анализирующего, каким образом меняются решения сформулированной двухуровневой задачи математического программирования при вариации основных параметров модели. При этом нас будет интересовать, прежде всего, зависимость

экономической эффективности, полученной согласно модели программы развития, от ключевых факторов процесса принятия решений частным инвестором — качества инвестиционного климата, проводимой государством инвестиционной политики (либеральной или консервативной), величины трансакционных издержек, отражающих уровень развития институциональной инфраструктуры в региональном минерально-сырьевом секторе.

Дисконт инвестора в модели является индикатором качества инвестиционного климата. В расчётах для него используется диапазон от $DI = 0,07$ (благоприятный инвестиционный климат) до $DI = 0,15$ (тяжёлые условия для инвестирования), соответствующего экспертной оценке современных инвестиционных условий в Забайкальском крае.

Дисконт DG отражает основные свойства государственной инвестиционной политики и варьируется от 0,01 (либеральная инвестиционная политика, ориентированная на долгосрочные мультипликативные эффекты) до 0,07 (консервативная инвестиционная политика собственника недр с коротким горизонтом оценки последствий принимаемых решений). Либеральное государство понимает, что отдача от инвестирования бюджетных средств в инфраструктуру (в виде компенсаций) реализуется на долгосрочном горизонте.

Параметр γ , определяющий степень учёта интересов частного инвестора в модели, оказывает заметное влияние только на функционал системы консорциумов. Прочие характеристики программы в меньшей степени зависят от γ и существенно меняются лишь в диапазоне $0,8 \leq \gamma \leq 1$, зависящем, вообще говоря, от особенностей региона.

Для проведения численных экспериментов зафиксируем четыре основных сценария выбора государством экономической политики в зависимости от уровня благоприятности инвестиционного климата в регионе. Первые два сценария предполагают благоприятный инвестиционный климат ($DI = 0,07$) и проведение государством либеральной (консервативной) бюджетной политики, формально описываемой дисконтом $DG = 0,01$ ($DG = 0,07$). Третий и четвёртый сценарии реализуют неблагоприятные инвестиционные условия ($DI = 0,15$). Государство в третьем сценарии в процессе принятия управленческого решения будет использовать либеральный $DG = 0,01$, а четвёртый сценарий фиксирует переход государства к консервативной политике с $DG = 0,07$.

Для каждого из сценариев будем изучать поведение отдельных показателей эффективности программы освоения в зависимости от уровня трансакционных издержек, предполагая, что соответствующий диапазон ограничен пятью процентами суммарных инвестиционных вложений в проекте. Такое предположение вполне соответствует практике природно-ресурсного комплекса — уровень развития институтов в этой сфере

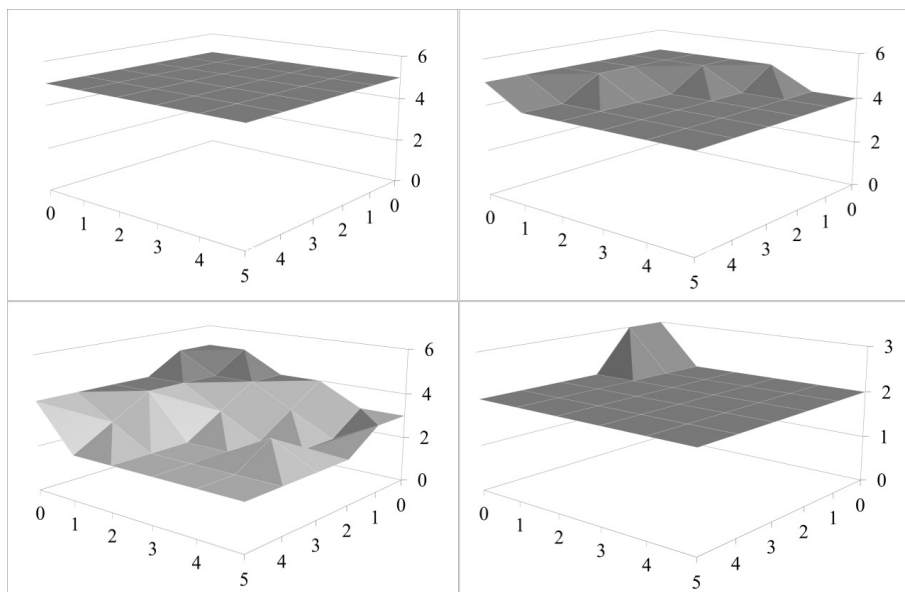


Рис. 1. Число образованных консорциумов

невысок и приводит к значительным затратам в процессе согласования интересов.

Какие показатели эффективности интересуют нас в первую очередь? Ключевой показатель для государства — бюджетный поток, включающий прямые налоговые доходы от освоения недр и ожидаемые мультипликативные эффекты от построенной в программе производственной инфраструктуры. В нём учтены понесённые затраты на формирование и содержание консорциумов, а также компенсации инфраструктурных затрат недропользователей. Именно таким образом построен целевой функционал задачи верхнего уровня, позволяющий ранжировать генерируемые моделью программы освоения недр в зависимости от выбранной бюджетной политики, особенностей сложившегося инвестиционного климата в регионе и уровня развития институтов.

Кроме бюджетного потока, нас интересует величина и структура затрат государства — в условиях бюджетного дефицита такая характеристика программы развития решающим образом влияет на процесс принятия решения. Помимо прямых характеристик затрат государства особую роль здесь играют оценки эффективности политики с точки зрения баланса стоимости новых инфраструктурных фондов и произведённых бюджетных расходов. Естественно ожидать положительности такого баланса, поскольку в противном случае государству проще построить ту же инфраструктуру за свой счёт.

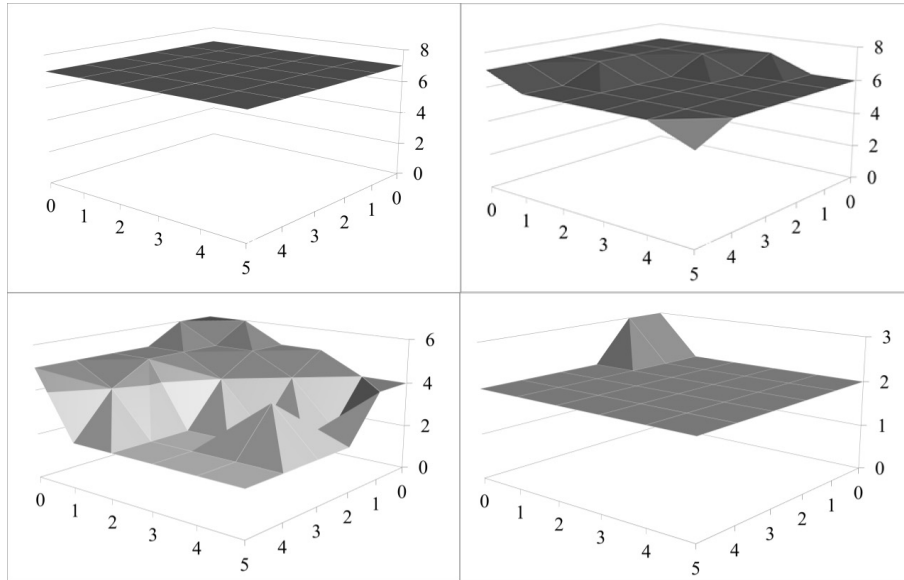


Рис. 2. Число реализованных инфраструктурных проектов

И наконец, привлекательность программы для частных инвесторов — полученный ими чистый дисконтированный доход — определяет степень достижения компромисса интересов партнёров и возможность практической реализации сгенерированной программы. Выбор месторождений для освоения и инфраструктурных проектов для реализации в сгенерированной модели программы позволяет сформировать перечисленные показатели эффективности партнёрства. Количественные характеристики производственной и инфраструктурной частей программы определяются тем набором консорциумов, для которых государство посчитало целесообразным образовать управляющие компании.

Каким образом интенсивность процесса формирования консорциумов зависит от инвестиционного климата, политики государства и уровня транзакционных издержек?

Результаты экспериментов представлены на рис. 1–8. Каждый из них состоит из четырёх фрагментов и устроен следующим образом: левый верхний квадрант содержит данные для сценария 1, правый верхний — для сценария 2, информация по третьему и четвёртому сценариям расположена в левом и правом нижних квадрантах соответственно. По осям отложены значения ТИ-до (левая на рис. 1–5, правая на рис. 6–8) и ТИ-после (правая на рис. 1–5, левая на рис. 6–8), выраженные в процентах от объёма капвложений.

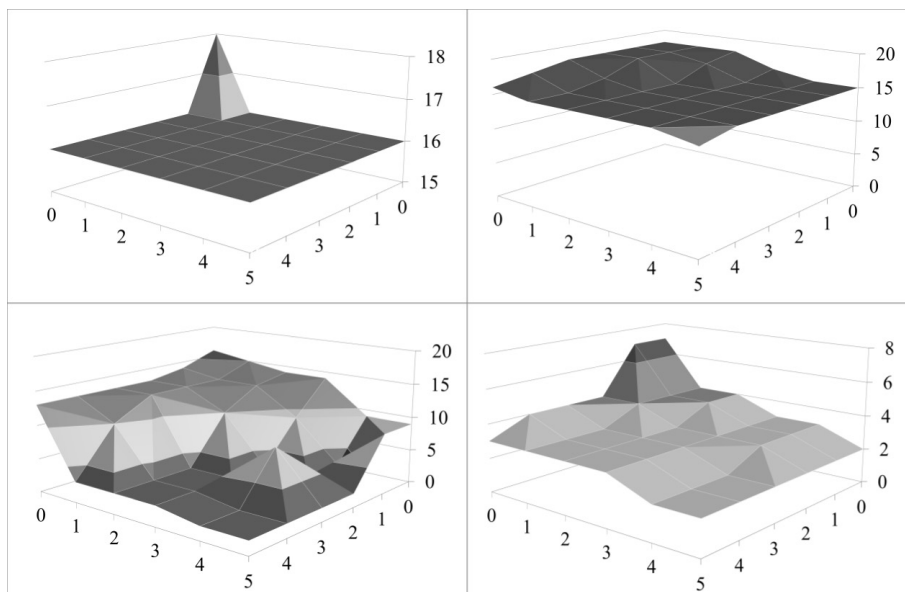


Рис. 3. Число реализованных проектов освоения месторождений

На рис. 1 видим, что в условиях первого сценария при любом уровне ТИ государство запускает все пять консорциумов, которые строят семь из десяти объектов производственной инфраструктуры (рис. 2). Это открывает возможность реализации, как минимум, шестнадцати из двадцати проектов освоения, а для нулевых ТИ их число доходит до 18 (рис. 3). Переход государства к консервативной политике в тех же благоприятных инвестиционных условиях влечёт за собой сокращение числа запускаемых консорциумов, производственных проектов и фронта инфраструктурного строительства при росте ТИ.

В неблагоприятном инвестиционном климате (сценарии 3, 4) все пять консорциумов формируются лишь при самых малых ТИ. Их рост приводит к существенному сокращению числа образованных УК в сценарии 3 и минимальному уровню программы освоения, сформированной консервативным государством на основе двух консорциумов практически при любом уровне ТИ (рис. 1). Число реализованных инфраструктурных и производственных проектов ведёт себя примерно так же (рис. 2, 3).

Зависящий от дисконтов партнёров и уровня ТИ выбор конкретного перечня осваиваемых месторождений и проектов транспортной и энергетической инфраструктуры для реализации в программе развития определяет поведение показателей эффективности программы в условиях разных сценариев. Функционал государства в благоприятном инвестиционном климате линейным образом зависит от ТИ, большую устойчивость

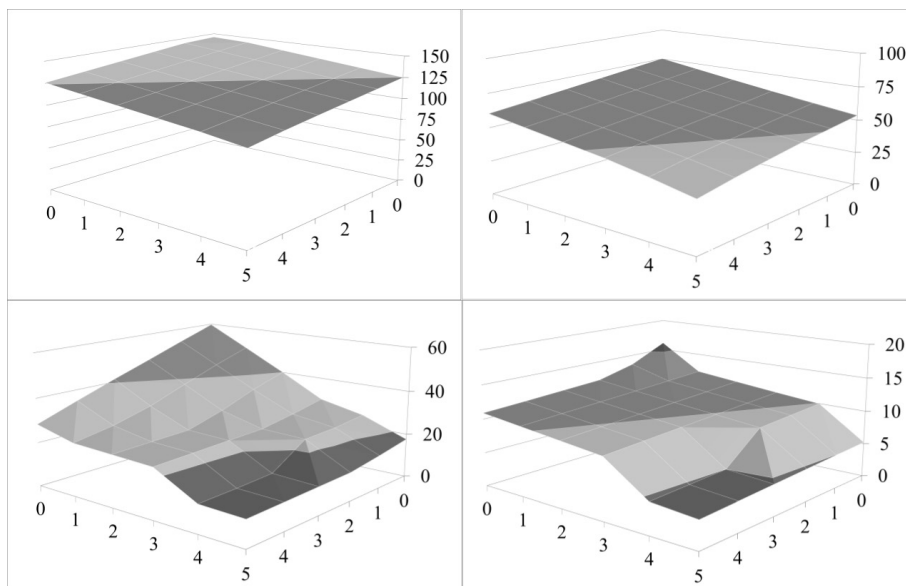


Рис. 4. Значение целевой функции государства (млрд руб.)

этому показателю придаёт либеральный характер государственной политики (рис. 4). Ухудшение инвестиционных условий для бизнеса в сценариях 3 и 4 существенно усугубляет негативное воздействие высоких ТИ,

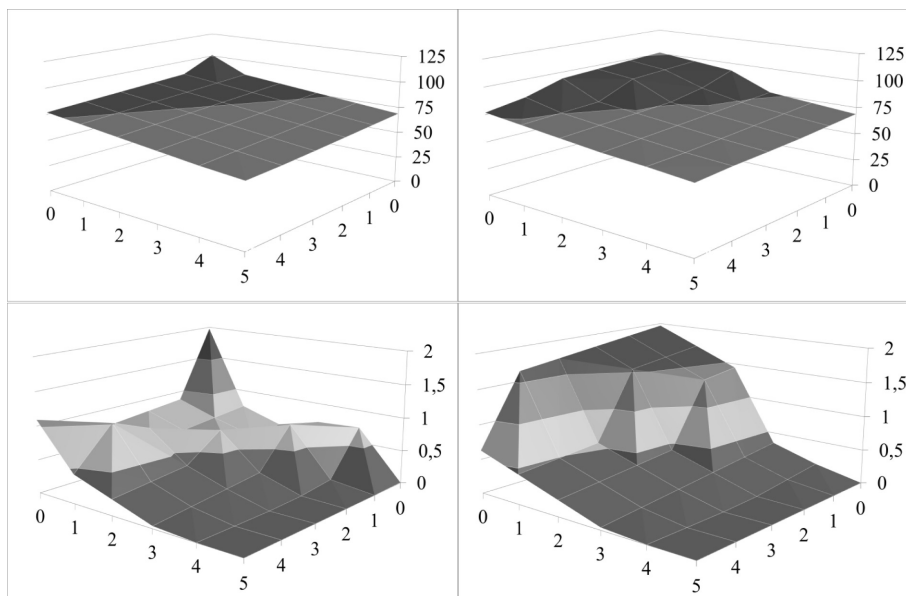


Рис. 5. Чистый дисконтированный доход инвесторов (млрд руб.)

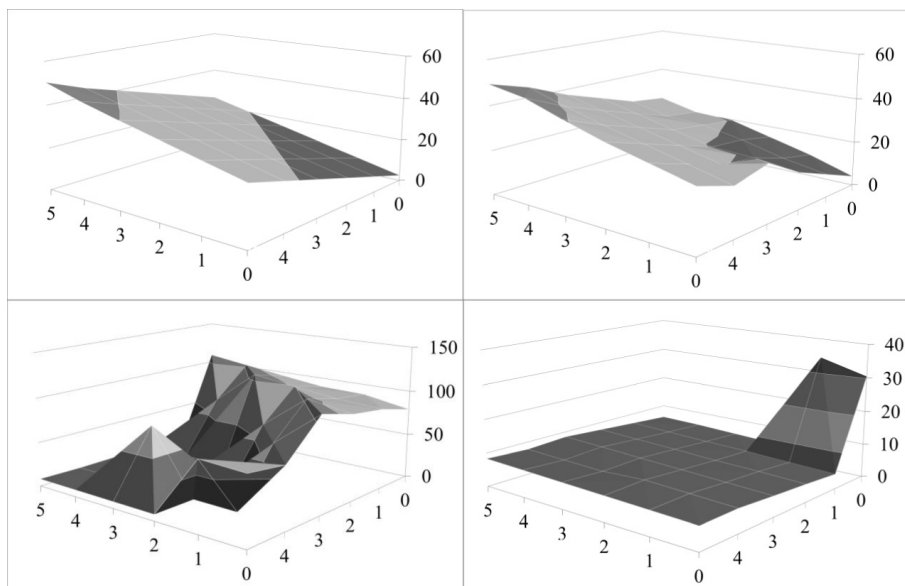


Рис. 6. Общие затраты государства на консорциумы и компенсации (млрд руб.)

особенно осязаемое при консервативной политике государства. Примерно так же рост ТИ влияет на чистый дисконтированный доход инвесторов: бизнесу нужен благоприятный инвестиционный климат, либеральное государство и развитые институты в недропользовании (рис. 5).

На рис. 6–7 представлена зависимость величины и структуры затрат государства от сценарных условий и уровня ТИ. В условиях благоприятного инвестиционного климата общие затраты государства практически линейно растут с ростом ТИ, при этом всё больше расходов требует функционирование УК, а удельный вес компенсационных выплат в государственных тратах стремительно падает (рис. 7).

Ухудшение общих инвестиционных условий для бизнеса в сценариях 3 и 4 приводит к существенно более сложной картине зависимости общих затрат государства от параметров модели. Общий вид соответствующих поверхностей на рис. 6 вполне соответствует характеру изменений интенсивности инфраструктурного строительства на рис. 2, но, в отличие от сценариев с благоприятным инвестиционным климатом, удельный вес расходов государства на компенсации превышает отметку в 70% для всех значений ТИ.

Как соотносятся стоимость новых, построенных в консорциумах инфраструктурных фондов и произведённые государством бюджетные расходы?

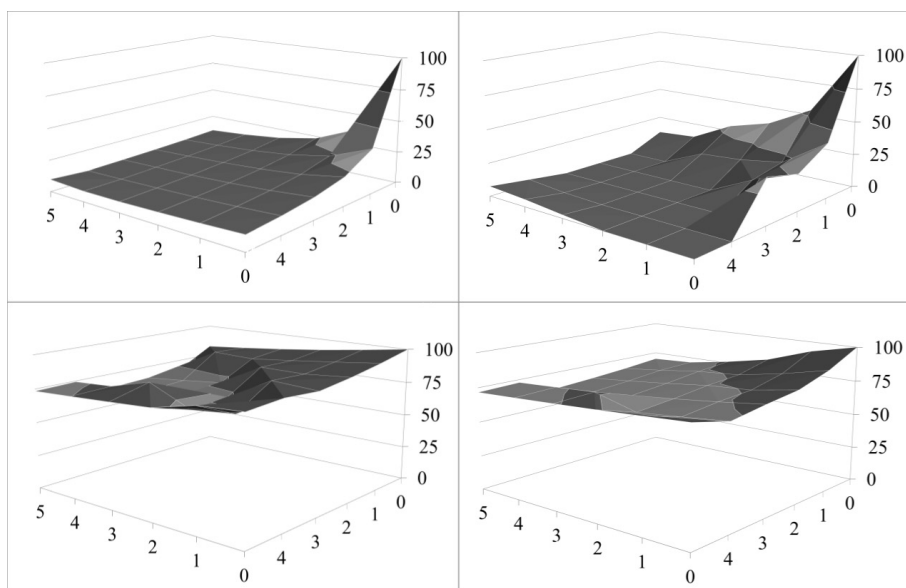


Рис. 7. Удельный вес затрат на компенсации в общих затратах государства (%)

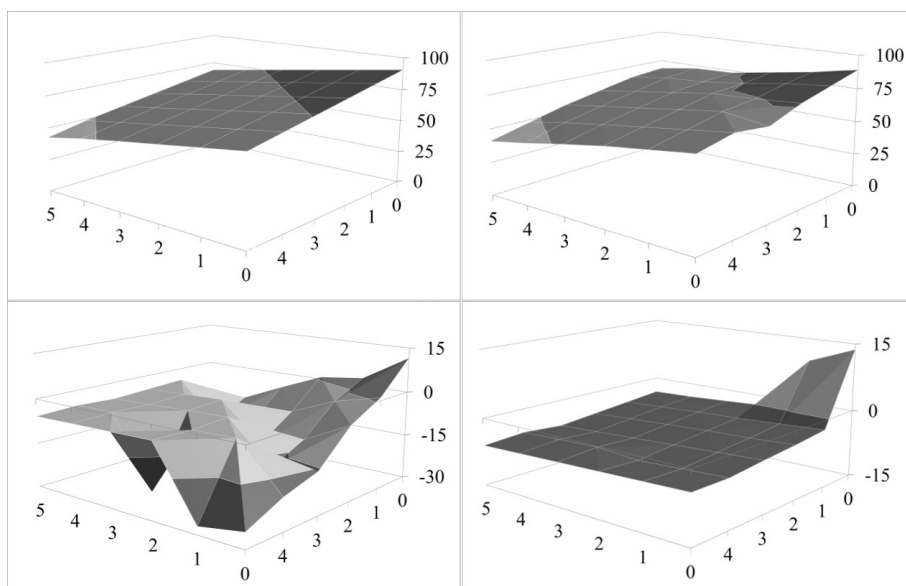


Рис. 8. Баланс стоимости новых инфраструктурных фондов и затрат государства (млрд руб.)

На рис. 8 представлена зависимость баланса стоимости созданных в программе инфраструктурных фондов и затрат государства при различных уровнях ТИ. Первые два сценария обеспечивают существенный выигрыш государства, сформировавшего программу освоения на основе консорциумов недропользователей, по сравнению с вариантом прямого финансирования строительства той же инфраструктуры за счёт бюджета. В регионе с неблагоприятным инвестиционным климатом картина существенно меняется: положительный баланс формируется только для самого малого уровня ТИ и не зависит от уровня либеральности политики государства. В этих условиях государству проще напрямую построить ту же инфраструктуру за свой счёт.

Сегодняшняя ситуация в минерально-сырьевом комплексе соответствует сценарию 4, а уровень ТИ в силу слабой договороспособности партнёров и отсутствия инструментов балансировки механизма долевого строительства чрезвычайно высок.

4. Обсуждение полученных результатов

Представленная модель (1)–(17) может служить основой для разработки инструментария, позволяющего поддержать процесс формирования программы освоения минерально-сырьевой базы в региональных администрациях на основе описанного механизма ГЧП с консорциумом.

В работе получены оценки сложности исходной задачи и установлено:

1) задача государства Σ_2^P -трудна и принадлежит классу Σ_2^PO , если переменные $\bar{W}_k^t$ и W_k^t , определяющие компенсацию инфраструктурных затрат недропользователей и накладные расходы управляющих компаний консорциумов, принимают только целые значения;

2) если $NPO \neq \Sigma_2^PO$, то задача государства не может лежать в классах PO , $FPTAS$, $PTAS$, APX , $Log-APX$, $Poly-APX$, $Exp-APX$.

Из этих результатов следует, что для задачи государства не существует детерминированных точных и приближённых с относительными оценками уклонения от оптимума алгоритмов, имеющих трудоёмкость, ограниченную полиномом от длины исходных данных. Эти факты стали причиной разработки метода решения задачи, использующего стохастический приближённый гибридный алгоритм, основанный на идеях покоординатного спуска. Продемонстрированная применимость метода для случая большой размерности позволяет использовать его в реальном инструментарии решения исходной задачи — задачи формирования программы освоения минерально-сырьевой базы ресурсного региона на основе консорциума недропользователей.

Апробация созданного модельного инструментария на примере Забайкальского края показала возможность решения задач, возникающих

в практике управления ресурсным регионом с недостаточно развитой производственной инфраструктурой. Заложенный в технологию численного эксперимента процесс изучения свойств равновесия по Штакельбергу позволил связать эффективность сформированной в модели программы освоения с качеством инвестиционного климата, либеральностью проводимой государством инвестиционной политики и уровнем развития институциональной инфраструктуры. Такого рода содержательный анализ представляет собой наиболее сложный раздел долгосрочного стратегического планирования и чрезвычайно востребован практикой.

Полученные в работе содержательные результаты представляют существенный интерес для экономистов-практиков и могут быть распространены на целый ряд сибирских и дальневосточных регионов. Предложенный инструментарий в виде модели Штакельберга, методов решения соответствующей двухуровневой задачи и технологии изучения свойств равновесных решений образует основу для создания практической методологии формирования инвестиционной политики государства, стимулирующей приход частного инвестора на основе гармонизации целей и объединения ресурсов отдельных недропользователей.

Финансирование работы

Исследование выполнено за счёт гранта Российского научного фонда (проект № 23–28–00849). Дополнительных грантов на проведение или руководство этим исследованием получено не было.

Конфликт интересов

Авторы заявляют, что у них нет конфликта интересов.

Литература

1. Dao N. H., Marisetty V. B., Shi J., Tan M. Institutional quality, investment efficiency, and the choice of public-private partnerships // Account. Finance. 2020. V. 60. P. 1801–1834. DOI: 10.1111/acfi.12514.
2. Grimsey D., Lewis M. K. Are public private partnerships value for money? // Account. Forum. 2005. V. 29, No. 4. P. 345–378.
3. Karnes Y. Are there optimal forms of public-private partnerships? // J. Behav. Appl. Manag. 2020. V. 20, No. 2. P. 51–68. DOI: 10.21818/001c.14178.
4. Малов В. Ю., Тарасова О. В., Ионова В. Д. Нижнее Приангарье как «связующее звено» между севером и югом Сибири // ЭКО. 2018. № 6. С. 50–74. DOI: 10.30680/EC00131-7652-2018-6-50-74.
5. Glazyrina I. P., Faleychik L. M., Faleychik A. A. Investments and the growth potential of the quality of life in the Russian Far East // Журн. Сиб. фед. ун-та. Сер. Гуманитар. науки. 2022. Т. 15, № 7. С. 921–929.

6. Самаруха В. И., Краснова Т. Г., Трусова С. В. Модель создания и реализации межрегионального проекта «Енисейская Сибирь» // Изв. Байкал. гос. ун-та. 2019. Т. 29, № 2. С. 324–331.
7. Bryukhanova E. A., Efimov V. S., Shishatsky N. G. Research on the issues of economic growth centres' establishment in the south of the Angara and Yenisei macroregion // Журн. Сиб. фед. ун-та. Сер. Гуманитар. науки. 2020. Т. 13, № 11. С. 1736–1745. DOI: 10.17516/1997-1370-0679.
8. Лавлинский С. М., Панин А. А., Плясунов А. В. Двухуровневая модель планирования государственно-частного партнёрства // Автоматика и телемеханика. 2015. № 11. С. 89–103. DOI: 10.1134/S0005117915110077.
9. Лавлинский С. М., Панин А. А., Плясунов А. В. Сравнение моделей планирования государственно-частного партнёрства // Дискрет. анализ и исслед. операций. 2016. Т. 23, № 3. С. 35–60. DOI: 10.17377/daio.2016.23.527.
10. Lavlinskii S. M., Panin A. A., Plyasunov A. V. Public-private partnership model with a consortium // Commun. Comput. Inf. Sci. 2023. V. 1881. P. 231–242. DOI: 10.1007/978-3-031-43257-6-18.
11. Marshall G. R. Transaction costs, collective action and adaptation in managing complex social-ecological systems // Ecol. Econ. 2013. V. 88. P. 185–194. DOI: 10.1016/j.ecolecon.2012.12.030.
12. McCann L., Colby B., Easter K. W., Kasterine A., Kuperan K. V. Transaction cost measurement for evaluating environmental policies // Ecol. Econ. 2005. V. 52, No. 4. P. 527–542. DOI: 10.1016/j.ecolecon.2004.08.002.
13. Ostrom E. A general framework for analyzing sustainability of social-ecological systems // Science. 2009. V. 325, No. 5939. P. 419–422. DOI: 10.1126/science.1172133.
14. Kochetov Yu. A., Plyasunov A. V., Panin A. A. Bilevel discrete optimisation: Computational complexity and applications // The Palgrave handbook of operations research. Cham: Palgrave Macmillan, 2022. P. 3–42. DOI: 10.1007/978-3-030-96935-6_1.
15. Caprara A., Carvalho M., Lodi A., Woeginger G. J. A study on the computational complexity of the bilevel knapsack problem // SIAM J. Optim. 2014. V. 24, No. 2. P. 823–838.
16. Панин А. А., Пащенко М. Г., Плясунов А. В. Двухуровневые модели конкурентного размещения производства и ценообразования // Автоматика и телемеханика. 2014. № 4. С. 153–169.
17. Васильев Ф. П. Численные методы решения экстремальных задач. М.: Глав. ред. физ.-мат. лит-ры, 1988. 552 с.
18. Alekseeva E. V., Kochetov Yu. A., Talbi E. G. A matheuristic for the discrete bilevel problem with multiple objectives at the lower level // Int. Trans. Oper. Res. 2017. V. 24, No. 5. P. 959–998. DOI: 10.1111/itor.12268.
19. Dempe S., Khamisov O., Kochetov Yu. A. A special three-level optimization problem // J. Global Optim. 2020. V. 76, No. 3. P. 519–531. DOI: 10.1007/s10898-019-00822-w.

- 20. Davydov I. A., Kochetov Yu. A., Dempe S.** Local search approach for the competitive facility location problem in mobile networks // Int. J. Artif. Intel. 2018. V. 16, No. 1. P. 130–143.

*Зырянов Александр Олегович
Лавлинский Сергей Михайлович
Панин Артём Александрович
Плясунов Александр Владимирович*

Статья поступила
22 августа 2024 г.
После доработки —
15 сентября 2024 г.
Принята к публикации
22 сентября 2024 г.

PUBLIC-PRIVATE PARTNERSHIP CONCERNING
INFRASTRUCTURE PROJECTS IN A RESOURCE REGION:
A MODEL BASED ON A CONSORTIUM OF SUBSOIL USERS

A. O. Zyryanov^a, S. M. Lavlinskii^b,
A. A. Panin^c, and A. V. Plyasunov^d

Sobolev Institute of Mathematics,
4 Acad. Koptug Avenue, 630090 Novosibirsk, Russia
E-mail: ^aalexander.zyryanov44@gmail.com, ^blavlin@math.nsc.ru,
^caapanin1988@gmail.com, ^dapljas@math.nsc.ru

Abstract. We propose a model for the formation of a resource region development program using a special mechanism of public-private partnership. It is based on the clustering of fields and the creation of a system of consortia of private investors jointly implementing projects for the construction of the necessary production infrastructure. The mechanism for achieving a compromise between the interests of the government and private investors uses the Stackelberg model, in which the government is the leader. It determines quotas of compensation of consortia's costs for the implementation of infrastructure projects. The role of the follower is assigned to the system of consortia, which forms the program of infrastructure construction. The solution of the corresponding bilevel mathematical programming problem allows us to form a targeted development plan. Its components are lists of infrastructure projects to be implemented in consortia, as well as schedules of costs for shared construction of infrastructure and their compensation from the budget for private investors. It is shown that the government problem is Σ_2^P -hard and belongs to the class $\Sigma_2^P\text{O}$ if the variables defining the transfer schedule take only integer values. A new stochastic approximate hybrid algorithm is developed to solve the two-level problem based on metaheuristics using the ideas of coordinate descent. The results of calculations on real data for the Zabaikalsky Krai allow us to formulate a set of practical recommendations on the formation of the mechanism of

shared construction and the policy of compensation payments. Illustr. 8, bibliogr. 20.

Keywords: bilevel problem of mathematical programming, Stackelberg model, metaheuristic, coordinate descent, polynomial hierarchy, approximation hierarchy, public-private partnership, production and infrastructure project, consortium of subsoil users.

References

1. N. H. Dao, V. B. Marisetty, J. Shi, and M. Tan, Institutional quality, investment efficiency, and the choice of public-private partnerships, *Account. Finance* **60**, 1801–1834 (2020), DOI: 10.1111/acfi.12514.
2. D. Grimsey and M. K. Lewis, Are public private partnerships value for money?, *Account. Forum* **29** (4), 345–378 (2005).
3. Y. Karnes, Are there optimal forms of public-private partnerships?, *J. Behav. Appl. Manag.* **20** (2), 51–68 (2020), DOI: 10.21818/001c.14178.
4. V. Yu. Malov, O. V. Tarasova, and V. D. Ionova, Lower Angara region as a link between the north and south of Siberia, *EKO*, No. 6, 50–74 (2018), DOI: 10.30680/EC00131-7652-2018-6-50-74 [Russian].
5. I. P. Glazyrina, L. M. Faleychik, and A. A. Faleychik, Investments and the growth potential of the quality of life in the Russian Far East, *Zh. Sib. Fed. Univ., Ser. Gumanit. Nauki* **15** (7), 921–929 (2022), DOI: 10.17516/1997-1370-0897.
6. V. I. Samarukha, T. G. Krasnova, and S. V. Trusova, The model for creation and implementation of the interregional project “Yenisei Siberia”, *Izv. Baikal. Gos. Univ.* **29** (2), 324–331 (2019) [Russian].
7. E. A. Bryukhanova, V. S. Efimov, and N. G. Shishatsky, Research on the issues of economic growth centres’ establishment in the south of the Angara and Yenisei macro-region, *Zh. Sib. Fed. Univ., Ser. Gumanit. Nauki* **13** (11), 1736–1745 (2020), DOI: 10.17516/1997-1370-0679.
8. S. M. Lavlinskii, A. A. Panin, and A. V. Plyasunov, A bilevel planning model for public-private partnership, *Avtom. Telemekh.*, No. 11, 89–103 (2015) [Russian] [*Autom. Remote Control* **76** (11), 1976–1987 (2015), DOI: 10.1134/S0005117915110077].
9. S. M. Lavlinskii, A. A. Panin, and A. V. Plyasunov, Comparison of models of planning public-private partnership, *Diskretn. Anal. Issled. Oper.* **23** (3), 35–60 (2016) [Russian], DOI: 10.17377/daio.2016.23.527 [*J. Appl. Ind. Math.* **10** (3), 356–369 (2016)].
10. S. M. Lavlinskii, A. A. Panin, and A. V. Plyasunov, Public-private partnership model with a consortium, *Commun. Comput. Inf. Sci.* **1881**, 231–242 (2023), DOI: 10.1007/978-3-031-43257-6-18.
11. G. R. Marshall, Transaction costs, collective action and adaptation in managing complex social–ecological systems, *Ecol. Econ.* **88**, 185–194 (2013), DOI: 10.1016/j.ecolecon.2012.12.030.

12. **L. McCann, B. Colby, K. W. Easter, A. Kasterine, and K. V. Kupe-
ran**, Transaction cost measurement for evaluating environmental policies, *Ecol.
Econ.* **52** (4), 527–542 (2005), DOI: 10.1016/j.ecolecon.2004.08.002.
13. **E. Ostrom**, A general framework for analyzing sustainability of social-
ecological systems, *Science* **325** (5939), 419–422 (2009), DOI: 10.1126/
science.1172133.
14. **Yu. A. Kochetov, A. V. Plyasunov, and A. A. Panin**, Bilevel discrete op-
timisation: Computational complexity and applications, in *The Palgrave Hand-
book of Operations Research* (Palgrave Macmillan, Cham, 2022), pp. 3–42, DOI:
10.1007/978-3-030-96935-6_1.
15. **A. Caprara, M. Carvalho, A. Lodi, and G. J. Woeginger**, A study on the
computational complexity of the bilevel knapsack problem, *SIAM J. Optim.*
24 (2), 823–838 (2014).
16. **A. A. Panin, M. G. Pashchenko, and A. V. Plyasunov**, Bilevel competi-
tive facility location and pricing problems, *Avtom. Telemekh.*, No. 4, 153–169
(2014) [Russian] [*Autom. Remote Control* **75** (4), 715–727 (2014)].
17. **F. P. Vasilyev**, *Numerical Methods for Solving Extremal Problems* (Glav. Red.
Fiz.-Mat. Lit., Moscow, 1988) [Russian].
18. **E. V. Alekseeva, Yu. A. Kochetov, and E. G. Talbi**, A matheuristic for
the discrete bilevel problem with multiple objectives at the lower level, *Int.
Trans. Oper. Res.* **24** (5), 959–998 (2017), DOI: 10.1111/itor.12268.
19. **S. Dempe, O. Khamisov, and Yu. A. Kochetov**, A special three-level op-
timization problem, *J. Global Optim.* **76** (3), 519–531 (2020), DOI: 10.1007/
s10898-019-00822-w.
20. **I. A. Davydov, Yu. A. Kochetov, and S. Dempe**, Local search approach
for the competitive facility location problem in mobile networks, *Int. J. Artif.
Intel.* **16** (1), 130–143 (2018).

Aleksandr O. Zyryanov
Sergey M. Lavlinskii
Artyom A. Panin
Aleksandr V. Plyasunov

Received August 22, 2024
Revised September 15, 2024
Accepted September 22, 2024

ОБ ОДНОЙ ЗАДАЧЕ ОПТИМИЗАЦИИ РАЗМЕЩЕНИЯ ТОВАРОВ НА СКЛАДЕ

К. О. Моторин^{1, a}, А. В. Пяткин^{2, b}

¹ Новосибирский гос. университет,
ул. Пирогова, 2, 630090 Новосибирск, Россия

² Институт математики им. С. Л. Соболева,
пр. Акад. Коптюга, 4, 630090 Новосибирск, Россия
E-mail: ^a k.motorin@ng.su.ru, ^b artem@math.nsc.ru

Аннотация. Рассматривается задача поиска расположения товаров на складе, при котором минимизируется суммарное время составления заказов из заданного списка. Доказано, что задача NP-трудна даже в простейшем частном случае. Построена математическая модель ЦЛП для этой задачи. Предложено два эвристических алгоритма её решения, работа которых проанализирована на случайно сгенерированных примерах. Табл. 3, ил. 9, библиогр. 17.

Ключевые слова: оптимизация склада, размещение товаров, NP-трудная задача, генетический алгоритм, поиск с запретами.

Введение

Объектом исследования в этой статье является следующая задача оптимизации размещения товаров на складе. Имеется склад прямоугольной формы с расположенными в нём стеллажами, на которых хранятся товары. Между стеллажами есть проходы, движение по каждому из которых в целях безопасности может осуществляться только в одном направлении. Дан список заказов, которые нужно составить с помощью сборщиков заказов. Каждый заказ состоит из нескольких товаров, и один сборщик собирает только один заказ. Чтобы составить заказ, сборщик должен двигаться в строго заданных направлениях обхода склада мимо нужных стеллажей, с которых он собирает входящие в заказ товары. Если на пути основного обхода склада есть стеллажи, на которых нет нужных товаров, то путь иногда можно сократить, не подходя к этим стеллажам (если при этом не нарушается общее правило одностороннего

движения по проходам). Каждый вид товара лежит ровно в одной ячейке, и гарантируется, что товаров хватит на сборку всех заказов. Требуется найти такое расположение товаров на складе, при котором суммарное время перемещения сборщиков будет минимальным.

Эта задача возникла из дискуссии на индустриальной секции конференции MOTOR-2023 (Екатеринбург, Россия, 2–8 июля 2023 г.). Насколько известно авторам, ранее в такой постановке задача не рассматривалась.

Отметим, что задачам оптимизации работы склада посвящено очень много публикаций (см., например, обзор [1]). Можно выделить несколько направлений в зависимости от того, что оптимизируется:

- структура склада [2–4];
- взаимодействие склада с внешними предприятиями [5];
- обход склада при составлении заказов [6, 7];
- размещение товаров на складе [4, 8, 9].

Рассматриваемая задача лежит на стыке двух последних типов. Однако в задачах оптимизации обхода склада при составлении заказов размещение товаров обычно фиксировано, и требуется найти оптимальный маршрут движения по складу (при этом требования односторонности движения нет). В задачах же размещения товаров в расчёт принимается не заданный список заказов, а некие общие статистические соображения (например, складывать рядом наиболее часто используемые товары или размещать рядом товары из близких групп).

С математической точки зрения наиболее близкой к рассматриваемой задаче является квадратичная задача о назначениях (КЗОН) [10], в которой требуется так расположить предприятия, чтобы суммарное произведение объёмов грузопотока на расстояние между каждой парой предприятий было бы наименьшим. В отличие от обычной задачи о назначениях, КЗОН NP-трудна, так как её частным случаем является задача о гамильтоновом цикле. И хотя в рассматриваемой задаче зависимость целевой функции от способа размещения значительно более сложная, некоторые подходы к решению КЗОН оказались полезными и для исследуемой задачи.

Статья имеет следующую структуру. В разд. 1 приводится подробная постановка задачи, строится её математическая модель и доказывается, что задача NP-трудна даже в простейшем частном случае. В следующих двух разделах описываются эвристические алгоритмы, предложенные для решения задачи (гибридный генетический алгоритм представлен в разд. 2, а итеративный алгоритм поиска с запретами — в разд. 3). В разд. 4 анализируются полученные результаты. В заключении даются финальные комментарии.

1. Постановка задачи и её сложность

В этом разделе даются содержательная и математическая постановки задачи, а также исследуется её сложность.

1.1. Содержательная постановка. Имеется склад прямоугольной формы, составленный из стеллажей и проходов между ними. Также задан список заказов, каждый из которых состоит из нескольких видов товаров. Для каждого стеллажа известна его вместимость (число ячеек). В каждую ячейку можно положить лишь один вид товара; при этом считается, что количество товара данного вида в ячейке достаточно для удовлетворения (сборки) всех заказов. На складе задан односторонний порядок движения, позволяющий обойти все стеллажи. Сборщик заказа не обязан обходить весь склад: если на каких-то стеллажах нет товаров из собираемого заказа, он иногда имеет возможность сократить маршрут (но при этом он не имеет права двигаться против заданного направления обхода склада). За один обход сборщик собирает только один заказ. Требуется так разместить товары на складе, чтобы суммарный пройденный путь при составлении всех заказов был минимальный.

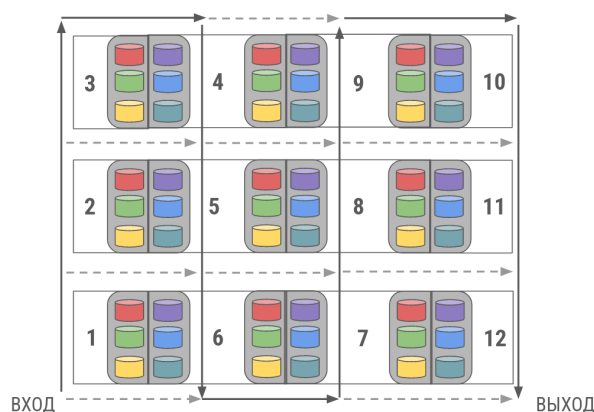


Рис. 1. Пример склада и нумерация точек размещения

На рис. 1 представлен пример такого склада. Чёрными стрелками обозначено направление обхода всего склада, в котором двигаются сборщики заказов. Пунктирные стрелки показывают места, где путь можно сократить, если на стеллажах основного пути нет нужных товаров для сбора текущего заказа. Каждый стеллаж разделён на две части, доступ к которым можно получить только с разных сторон.

Отметим, что при такой постановке задачи не возникает проблемы маршрутизации. Действительно, зная распределение товаров по ячейкам

стеллажей, можно однозначно восстановить кратчайший путь каждого из сборщика заказов. Для этого достаточно пронумеровать точки между стеллажами (будем называть их *точками размещения*) в порядке обхода склада (рис. 1) и задать матрицу кратчайших расстояний между ними. Тогда для вычисления кратчайшего пути обхода для сборщика заказов нужно сложить расстояния между точками размещения, которые нужно посетить для сборки заказа, в порядке возрастания, добавив расстояния от входа до первого стеллажа в пути и от последнего до выхода. Отметим, что в терминах точек размещения не имеет значения, в какой именно ячейке примыкающих к ней стеллажей находится нужный товар. Это значительно упрощает генерацию и формулировку примеров задачи для решения её с помощью алгоритмов.

1.2. Математическая модель. Формализуем приведённую выше содержательную задачу в виде следующей модели смешанного целочисленного линейного программирования.

Параметры задачи:

- N — число точек размещения;
- M — число видов товаров;
- K — число заказов;
- d_{ij} — кратчайшее расстояние от точки размещения i до точки размещения j ;

- b_j — вместимость точки размещения j ;

- $o_i^k = \begin{cases} 1, & \text{если в заказе } k \text{ есть товар } i, \\ 0 & \text{иначе.} \end{cases}$

Переменные задачи:

- $x_{ij} = \begin{cases} 1, & \text{если товар } i \text{ расположен в точке размещения } j, \\ 0 & \text{иначе;} \end{cases}$

- $p_j^k = \begin{cases} 1, & \text{если точка } j \text{ должна быть посещена в заказе } k, \\ 0 & \text{иначе;} \end{cases}$

- f_{ij}^k — вспомогательная переменная, представляющая результат произведения $p_i^k p_j^k$;

- $l_{ij}^k = \begin{cases} 1, & \text{если для заказа } k \text{ точка } i \text{ посещается после точки } j, \\ 0 & \text{иначе.} \end{cases}$

Индексы:

- $M_{\text{idx}} = \{0, \dots, M - 1\}$ — индексы товаров;
- $N_{\text{idx}} = \{1, \dots, N - 2\}$ — индексы точек размещения; точки размещения 0 и $N - 1$ соответствуют входу и выходу;
- $K_{\text{idx}} = \{0, \dots, K - 1\}$ — индексы заказов.

Целевая функция

$$\sum_{k=0}^{K-1} \sum_{i=0}^{N-2} \sum_{j=i+1}^{N-1} d_{ij} l_{ij}^k \rightarrow \min. \quad (1)$$

Ограничения:

$$\sum_{j=1}^{N-2} x_{ij} = 1, \quad i \in M_{\text{idx}}, \quad (2)$$

$$\sum_{i=0}^{M-1} x_{ij} \leq b_j, \quad j \in N_{\text{idx}}, \quad (3)$$

$$p_j^k \geq x_{ij} o_i^k, \quad i \in M_{\text{idx}}, j \in N_{\text{idx}}, k \in K_{\text{idx}}, \quad (4)$$

$$p_0^k = 1, \quad k \in K_{\text{idx}}, \quad (5)$$

$$p_{N-1}^k = 1, \quad k \in K_{\text{idx}}, \quad (6)$$

$$f_{ij}^k \leq p_i^k, \quad i, j \in N_{\text{idx}}, k \in K_{\text{idx}}, \quad (7)$$

$$f_{ij}^k \leq p_j^k, \quad i, j \in N_{\text{idx}}, k \in K_{\text{idx}}, \quad (8)$$

$$p_i^k + p_j^k - f_{ij}^k \leq 1, \quad i, j \in N_{\text{idx}}, k \in K_{\text{idx}}, \quad (9)$$

$$l_{ij}^k \leq f_{ij}^k, \quad i, j \in N_{\text{idx}}, i < j, k \in K_{\text{idx}}, \quad (10)$$

$$l_{ij}^k \leq 1 - p_l^k, \quad i, j, l \in N_{\text{idx}}, i < l < j, k \in K_{\text{idx}}, \quad (11)$$

$$l_{ij}^k \geq f_{ij}^k - \sum_{l=i+1}^{j-1} p_l^k, \quad i, j \in N_{\text{idx}}, i < j - 1, \quad (12)$$

$$l_{jj+1}^k \geq f_{jj+1}^k, \quad j \in \{0, \dots, N-2\}, k \in K_{\text{idx}}. \quad (13)$$

Целевая функция (1) определяет суммарную длину всех маршрутов сборки заказов. Ограничение (2) гарантирует, что один вид товара расположен ровно в одной точке размещения. Неравенства (3) задают максимальную вместимость для каждой точки размещения. Ограничения (4)–(6) нужны для определения вспомогательной переменной, которая показывает, нужно ли посещать точку размещения для конкретного заказа. В модели предполагается, что вход на склад и выход со склада — это точки размещения 0 и $N-1$ соответственно, с нулевой вместимостью, которые обязательно нужно посетить во всех заказах. Неравенства (7)–(9) задают переменную f_{ij}^k , которая нужна для индикации, нужно ли посещать обе точки размещения i, j в заказе k . Вспомогательная переменная f_{ij}^k является результатом линеаризации произведения $p_i^k p_j^k$. Ограничения (10)–(13) задают переменную l_{ij}^k , участвующую в целевой

функции. Эта переменная равна единице, если в заказе k для сборки заказа требуется посетить подряд точки размещения i и j (т. е. во всех промежуточных точках из интервала (i, j) нет нужных товаров для k -го заказа). Тогда для подсчёта суммарного пути сборщиков заказов требуется просуммировать все переменные l_{ij}^k , умноженные на соответствующие кратчайшие длины между точками размещения.

1.3. Сложность задачи. Покажем, что задача NP-трудна и в простейшем частном случае, когда на складе есть всего два стеллажа и каждый заказ содержит лишь два вида товара. Нам потребуется следующая известная NP-трудная задача [11, 12].

Задача о бисекции. Дан граф $G = (V, E)$ на $2n$ вершинах. Найти в G минимальную бисекцию, т. е. такое разбиение $V = V_1 \cup V_2$, что $|V_1| = |V_2| = n$ и число рёбер, соединяющих V_1 и V_2 , минимально.

Теорема 1. Задача о размещении товаров на складе в случае, когда имеется два стеллажа и все заказы состоят из двух видов товаров, NP-трудна.

ДОКАЗАТЕЛЬСТВО. Рассмотрим произвольный пример задачи о бисекции (граф $G = (V, E)$ на $2n$ вершинах) и построим по нему следующий пример задачи о размещении товаров на складе. Сопоставим каждой вершине графа G один вид товара, а каждому ребру — заказ, товары в котором соответствуют инцидентным этому ребру вершинам. Рассмотрим склад с двумя стеллажами длины L , вместимость каждого из которых равна n , причём размещать товары можно только с внешних сторон стеллажа. Также между внутренними сторонами стеллажей есть проход, позволяющий вернуться от конца первого стеллажа к началу второго. Ширина стеллажей считается незначительной. Схема такого склада представлена на рис. 2. Склад имеет три точки размещения, вторая из которых имеет вместимость 0. Очевидно, что каждое распределение товаров по стеллажам взаимно однозначно соответствует некоторой бисекции графа.

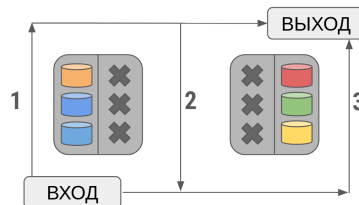


Рис. 2. Схема склада из теоремы 1

Нетрудно видеть, что если оба товара в заказе лежат на одном стеллаже, то маршрут сборщика имеет длину L , а если на разных, то $— 3L$, поэтому для минимизации пройденного пути нужно расположить товары так, чтобы максимизировать число заказов, для сборки которых требуется посетить только один стеллаж. Тогда в соответствующей бисекции графа число рёбер, концы которых лежат в разных множествах, будет минимально. И наоборот, если известна минимальная бисекция графа, то предметы, соответствующие вершинам из одного множества, нужно расположить на одном стеллаже. В таком случае число заказов, для которых потребуется посетить два стеллажа, будет минимальным.

Таким образом, по любому примеру задачи о минимальной бисекции можно построить эквивалентный пример исследуемой задачи о размещении товаров на складе с двумя стеллажами и заказами из двух видов товаров. Из этого следует, что задача о размещении товаров на складе NP-трудна в таком частном случае. Теорема 1 доказана.

2. Гибридный генетический алгоритм

В качестве первого алгоритма для исследуемой проблемы был выбран гибридный генетический алгоритм.

2.1. Структура алгоритма. Генетический алгоритм является вероятностным алгоритмом поиска, основанным на биологических принципах живой природы. Идея алгоритма заключается в генерации начальной популяции — набора решений исследуемой задачи, называемых индивидами, и последующем улучшении этой популяции. На каждой

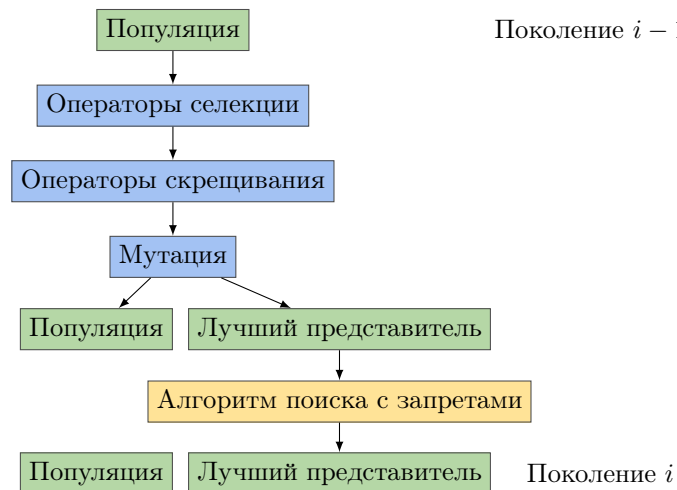


Рис. 3. Схема гибридного генетического алгоритма

итерации алгоритма представители популяции изменяются с помощью операций селекции, скрещивания и мутации, образуя новое поколение, в котором может найтись решение с лучшим значением целевой функции. Однако, в процессе работы было экспериментально обнаружено, что применение генетического алгоритма само по себе не даёт желаемых результатов, поэтому в качестве дополнительной эвристики выбран алгоритм поиска с запретами, который на каждой итерации улучшает рекордного индивидуума в популяции. Общая схема алгоритма представлена на рис. 3.

Каждое решение оптимизационной задачи (индивид популяции) кодируется M -мерным вектором, i -я координата которого указывает номер точки размещения, в которой расположен i -й вид товара. Тогда для допустимости решения необходимо и достаточно, чтобы каждое значение j встречалось не более чем в b_j координатах вектора. В алгоритме используется два оператора скрещивания и один оператор мутации.

1. Оператор равномерного скрещивания. Сначала потомок наследует одинаковые признаки от обоих родителей, затем остальные признаки наследуются равномерно. Если добавление признака нарушает допустимость решения, то текущая позиция пропускается. Далее все пропущенные позиции заполняются оставшимися свободными точками размещения случайным образом. Применение такого оператора требует $O(M)$ операций. На рис. 4 изображено схематичное представление данного оператора (во всех примерах ниже предполагается, что все $b_j = 2$).

2. Оператор двухточечного скрещивания. Сначала случайным образом выбираются две точки скрещивания — два целых числа $i, j \in [1, M]$, причём $i \leq j$. Затем потомок наследует интервал $[i, j]$ от первого родителя, а оставшиеся координаты наследуются от второго родителя, если

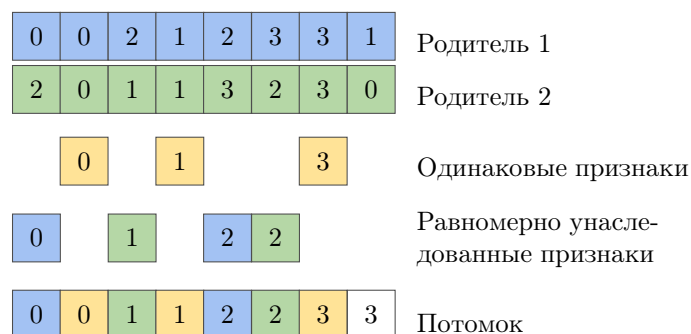


Рис. 4. Оператор равномерного скрещивания



Рис. 5. Оператор двухточечного скрещивания

это не нарушает условия допустимости решения. Далее те ячейки решения, для которых значения остались неопределёнными, заполняются случайным образом из оставшихся допустимых точек размещения. Этот оператор также требует $O(M)$ операций. Схема оператора двухточечного скрещивания представлена на рис. 5.

Отметим, что также рассматривался оператор множественного скрещивания, в котором выбирается несколько родителей и координаты вектора заполняются в случайно выбранном порядке по принципу «выбираем наиболее часто встречающееся значение этой координаты среди родителей». Известно, что такой оператор скрещивания показывает хорошие результаты для КЗОН [13]. Однако в настоящей работе было решено отказаться от его использования, так как на больших примерах время работы алгоритма значительно росло при отсутствии заметного улучшения значения целевой функции.

3. В качестве оператора мутации была использована многократная последовательная замена значений двух случайных координат исходного решения друг другом. На каждом шаге случайным образом выбираются два целых числа $i, j \in [1, M]$, происходит обмен значений этих координат текущего вектора и вычисление целевой функции. Результирующим решением считается мутировавший вектор с минимальным значением целевой функции (при этом целевая функция исходного решения не учитывается). На рис. 6 представлена схема данного оператора.

Отметим, что число шагов оператора мутации регулируется специальным параметром — степенью мутации μ , который варьируется между минимальным и максимальным значениями $\mu_{\min}$ и $\mu_{\max}$. На первой итерации алгоритма этот параметр равен минимальному значению, чтобы дать возможность алгоритму найти лучшее решение в текущей области



Рис. 6. Оператор мутации

локального минимума. На каждой следующей итерации алгоритма значение параметра увеличивается на 1, каждый раз всё больше изменяя представителей популяции. Таким образом осуществляется поиск решения в отдалённых областях пространства решений. Степень мутации возвращается к минимальному значению, если на предыдущей итерации алгоритма либо значение параметра было максимальным, либо произошло обновление лучшего найденного решения. Похожая идея использовалась в [14] при разработке итерационного алгоритма локального поиска для КЗОН.

2.2. Локальное улучшение. На каждой итерации генетического алгоритма рекордный представитель популяции (в смысле значения целевой функции) улучшается с помощью алгоритма поиска с запретами. Такой алгоритм был впервые предложен в [15] как модифицированная версия алгоритма локального поиска, в которой имеется список запретов, не позволяющий алгоритму выбирать некоторые решения из рассматриваемой окрестности. Такой механизм позволяет не заикливаться в областях локального минимума. Алгоритм поиска с запретами зарекомендовал себя во многих задачах дискретной оптимизации и имеет множество модификаций [16].

В качестве окрестности, в которой происходит поиск нового решения, выбирается множество всех решений, получающихся переменой различных точек размещения ровно двух товаров, т. е.

$$N(x_1, \dots, x_M) = \{(x'_1, \dots, x'_M) \mid \exists i, j: x_i \neq x_j, \\ x'_i = x_j, x'_j = x_i, x'_k = x_k \text{ при } k \neq i, j\}.$$

Эта окрестность имеет мощность $(M^2 - M)/2$. Будем использовать обозначение $x' = \text{swar}(x, i, j)$, если $x'_i = x_j, x'_j = x_i$ и $x'_k = x_k$ для всех $k \neq i, j$. Для реализации механизма списка запретов используется матрица T размера $M \times N$, элемент $T_{i,j}$ которой показывает, до какой итерации алгоритма запрещено назначение элемента i в точке размещения j . После выбора нового текущего решения из окрестности новые назначения двух изменённых элементов решения попадают под запрет до итерации $\text{iter} + h$, где iter — номер текущей итерации, а h — время запрета. Значение параметра h выбирается случайным образом из отрезка $[h_{\min}, h_{\max}]$ для большей стабильности алгоритма.

Так как время просмотра всей окрестности равно $O(M^2)$, на каждой итерации просматривается не вся окрестность, а только её часть. Для этого в алгоритме задаётся специальный параметр p — вероятность, с которой очередное решение из окрестности будет взято в рассмотрение. Такой подход уменьшает трудоёмкость одной итерации алгоритма и может повысить вероятность нахождения глобального минимума.

Приведём формальное описание всего алгоритма. Отметим, что процедура поиска с запретами вынесена в отдельный алгоритм, поскольку она будет использована и в следующем разделе.

Алгоритм 1. Гибридный генетический алгоритм.

ВХОД: число товаров M , размер популяции P , минимальная и максимальная степени мутации $\mu_{\min}$ и $\mu_{\max}$, минимальное и максимальное время запрета $h_{\min}$ и $h_{\max}$.

ШАГ 1. Выбрать начальную популяцию $S = \{s_1, \dots, s_P\}$ и найти в ней рекордное решение s^* с минимальным значением целевой функции $\text{obj}(s^*)$. Задать $\mu = \mu_{\min}$. Положить $g = 0$. Пока не выполнен критерий остановки, выполнять шаги 2–7.

ШАГ 2. Применить операторы скрещивания.

ШАГ 3. Применить оператор мутации с числом шагов μ . Найти в популяции индивида s' с минимальным значением целевой функции.

ШАГ 4. Применить алгоритм поиска с запретами (Алгоритм 2) с входом $M, s', h_{\min}$ и $h_{\max}$ и получить улучшенное решение x^* .

ШАГ 5. Положить $s'' = x^*$.

ШАГ 6. Если $\mu = \mu_{\max}$, то положить $\mu = \mu + 1$, иначе $\mu = \mu_{\min}$.

ШАГ 7. Если $\text{obj}(s'') < \text{obj}(s^*)$, то положить $s^* = s''$ и $\mu = \mu_{\min}$.

ВЫХОД: решение s^* .

Алгоритм 2. Локальный поиск с запретами.

ВХОД: число товаров M , начальное решение s' , минимальное и максимальное время запрета $h_{\min}$ и $h_{\max}$.

ШАГ 1. Задать текущее решение $x = s'$, лучшее решение $x^* = s'$, матрицу запретов $T = 0$, вероятность выбора элемента окрестности $p \in (0, 1)$ и номер итерации $\text{iter} = 0$. Пока не выполнен критерий останова, выполнять шаги 2–8.

ШАГ 2. Выбрать $h = \text{rand}(h_{\min}, h_{\max})$. Положить $i_{\text{best}} = M + 1$, $j_{\text{best}} = M + 1$, $R = \infty$.

ШАГ 3. Для всех $i = 1, \dots, M - 1$ и $j = i + 1, \dots, M$, если $x_i \neq x_j$ и $\text{rand}(0, 1) < p$, то выполнить шаги 4–6.

ШАГ 4. Задать решение $x' = \text{swap}(x, i, j)$ и вычислить $\text{obj}(x')$.

ШАГ 5. Если $(\text{obj}(x') < \text{obj}(x^*))$ или $(\text{obj}(x') < R$ и $T_{i,x(j)} \leq \text{iter}$ и $T_{j,x(i)} \leq \text{iter})$, то положить $i_{\text{best}} = i$, $j_{\text{best}} = j$, $R = \text{obj}(x')$.

ШАГ 6. Если $\text{obj}(x') < \text{obj}(x^*)$, то положить $x^* = x'$.

ШАГ 7. Если $i_{\text{best}} \leq M$, то положить $x = \text{swap}(x, i_{\text{best}}, j_{\text{best}})$, а также $T_{i_{\text{best}},x(j_{\text{best}})} = \text{iter} + h$ и $T_{j_{\text{best}},x(i_{\text{best}})} = \text{iter} + h$.

ШАГ 8. Положить $\text{iter} = \text{iter} + 1$.

ВЫХОД: решение x^* и номер последней итерации iter .

Отметим также технический момент, позволяющий ускорить пересчёт значения целевой функции при перестановке двух элементов. Во-первых, путь сбора заказа изменится только у тех сборщиков, у которых в заказе имеется только один тип товара из перестановки. Во-вторых, значение целевой функции для заказа изменится только при условии, что на точке размещения, где находится товар из перестановки, нет других товаров из заказа, либо в текущем маршруте для сборщика нет точки размещения, на место которой был назначен товар из перестановки, поэтому хранение информации о текущих путях обхода склада для каждого заказа организовано в виде списка, в элементе которого указан номер точки размещения для посещения и число типов товаров из заказа на этой точке. Тогда для пересчёта значений целевой функции нужно знать старое и новое расположения элемента решения, а также предыдущее значение целевой функции. Всего для подсчёта нового значения целевой функции потребуется $O(KL)$ операций, где K — это число всех заказов, а L — максимальная длина одного заказа.

3. Итеративный алгоритм поиска с запретами

Так как модель смешанного целочисленного линейного программирования для исследуемой задачи позволяет найти решение с помощью специальных программ-решателей только для примеров малой размерности, требуются дополнительные алгоритмы, с которыми можно сравнивать полученные результаты. В этом разделе представлен второй разработанный алгоритм для исследуемой проблемы, а именно алгоритм итеративного поиска с запретами. В нём используются три различные модификации алгоритма 2 в качестве процедуры локального улучшения на каждой итерации. Алгоритм заключается в итеративном применении процедур интенсификации и диверсификации, которые используются для борьбы с проблемами заикливания алгоритма в областях локальных минимумов. Первый вариант такого поиска был предложен в [17]. Фаза интенсификации нацелена на поиск решения в специальном «многообещающем» регионе пространства решений задачи, который находится рядом с текущим найденным локальным минимумом, в то время как фаза диверсификации позволяет выйти из текущей области локального минимума в сторону новых неизученных регионов. Отметим, что аналогичный подход был успешно применён к КЗОН, где с помощью него удалось найти новые лучшие решения для примеров, часто используемых в тестировании алгоритмов для данной задачи [14].

Алгоритм состоит из следующих этапов (см. схему на рис. 7).

1. Выбрать случайное решение в качестве текущего.
2. Текущее решение улучшается с помощью модификации алгоритма 2, в которой вычисляется матрица частоты назначения товаров в точки размещения.
3. На фазе интенсификации происходит поиск нового улучшенного решения в «хороших» областях пространства решения вблизи текущего значения локального минимума.

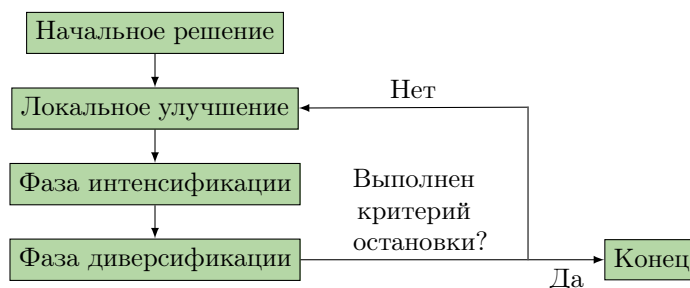


Рис. 7. Схема алгоритма итеративного локального поиска

4. После фазы интенсификации применяется фаза диверсификации для выхода из текущей области локального минимума. Если критерий остановки не выполнен, то алгоритм переходит на этап 2.

Опишем подробнее фазы интенсификации и диверсификации.

Для реализации фазы интенсификации алгоритм поиска с запретами на втором этапе заполняет матрицу F размера $M \times N$, значение $F_{i,j}$ которой показывает, как часто вид товара под номером i находился в точке размещения j . В фазе интенсификации в рекордном решении фиксируется часть компонент, которые чаще всего находились на местах лучшего решения. Доля фиксированных компонент решения задаётся специальным параметром P_I . После этого на рекордном решении запускается модификация алгоритма поиска с запретами, в которой перестановки, содержащие зафиксированные элементы решения, запрещены. Это позволяет сократить размерность задачи, тем самым ускорив работу алгоритма поиска с запретами, а также сосредоточиться на поиске лучшего решения вблизи текущего найденного локального минимума. Параметр $rate$ задаёт порог, по которому отбираются самые частые назначения в лучшем решении алгоритма. Этот порог уменьшается до тех пор, пока число фиксированных элементов решения не достигнет MP_I .

В фазе диверсификации используется похожая стратегия. Сначала задаётся матрица доступных назначений A размера $M \times N$ (т. е. i -й вид товара можно разместить в j -й точке тогда и только тогда, когда $A_{i,j} = 1$). Назначения выбираются из вариантов, которые находились в текущих решениях реже всего. Специальный параметр P_D регулирует долю возможных назначений типов товаров на точки размещения, которые будут разрешены в матрице A . Далее запускается модификация алгоритма 2, в которой разрешены только назначения из матрицы A , а число итераций выбирается небольшим. Это позволяет продолжить поиск лучшего решения в ранее не рассмотренных областях. Матрица A заполняется следующим образом. Выбирается начальное значение порога $rate$, которое ограничивает частоту назначения типов товаров. Если доля назначений товара i на место j от общего числа итераций меньше этого порога, то такое назначение становится разрешённым. Параметр $rate$ увеличивается на малое значение до тех пор, пока число разрешённых назначений не достигнет NMP_D .

Приведём формальное описание итеративного алгоритма поиска с запретами. Шаги 4–8 соответствуют фазе интенсификации, а шаги 9–13 — фазе диверсификации.

Алгоритм 3. Итеративный алгоритм поиска с запретами.

ВХОД: число товаров M и точек назначения N , минимальное и максимальное время запрета $h_{\min}$ и $h_{\max}$, доли фиксированных элементов

решения фазы интенсификации $P_I \in (0, 1)$, доля допустимых назначений в фазе диверсификации $P_D \in (0, 1)$, начальное решение x_0 .

ШАГ 1. Задать текущее решение $x = x_0$ и рекордное решение $x^* = x_0$.

ШАГ 2. Пока не выполнен критерий остановки, применять шаги 3–13.

ШАГ 3. Задать матрицу $F = 0$ и применить модификацию алгоритма 2 с входом $M, x, h_{\min}$ и $h_{\max}$, в которой на шаге 8 выполняется дополнительное действие: для всех $i = 1, \dots, M$ положить $F_{i,x(i)} = F_{i,x(i)} + 1$. На выходе получим рекордное решение x^* , а также число итераций I .

ШАГ 4. Задать вектор фиксированных товаров $\text{fix} = 0$ и их число $M_{\text{fix}} = 0$, а также порог $\text{rate} = 0,95$.

ШАГ 5. Пока $M_{\text{fix}} < MP_I$ выполнять шаги 6–7.

ШАГ 6. Для всех $i = 1, \dots, M$ если $F_{i,x^*(i)} > I \cdot \text{rate}$ и $\text{fix}[i] = 0$, то положить $\text{fix}[i] = 1$ и $M_{\text{fix}} = M_{\text{fix}} + 1$.

ШАГ 7. Положить $\text{rate} = \text{rate} - 0,05$.

ШАГ 8. Применить следующую модификацию алгоритма 2: на вход подаются $M, x, h_{\min}$ и $h_{\max}$, а также вектор fix ; при этом на шаге 3 дополнительно проверяются условия $\text{fix}[i] = 0$ и $\text{fix}[j] = 0$.

ШАГ 9. Задать матрицу доступных назначений $A = 0$ и их число $N_A = 0$, а также порог $\text{rate} = 0,05$.

ШАГ 10. Пока $N_A < NMP_D$ выполнять шаги 11–12.

ШАГ 11. Для всех $i = 1, \dots, M$ и $j = 1, \dots, N$, если $F_{i,x^*(i)} < I \cdot \text{rate}$ и $A_{i,j} = 0$, то положить $A_{i,j} = 1$ и $N_A = N_A + 1$.

ШАГ 12. Положить $\text{rate} = \text{rate} + 0,05$.

ШАГ 13. Применить следующую модификацию алгоритма 2: на вход подаются $M, x, h_{\min}$ и $h_{\max}$, а также матрица A ; при этом на шаге 3 дополнительно проверяются условия $A_{i,x^*(j)} = 1$ и $A_{j,x^*(i)} = 1$.

ВЫХОД: решение x^* .

4. Результаты

В этом разделе представлены описание способа генерации примеров, на которых производилось тестирование алгоритмов, выбранные параметры каждого из алгоритмов и полученные численные результаты.

4.1. Тестовые примеры. Так как исследуемая задача ранее не изучалась, общедоступных библиотек примеров для тестирования алгоритмов её решения нет, поэтому авторами реализована программа, которая генерирует примеры разной размерности для задачи. На вход программе подаётся число видов товаров, размеры стеллажей (число ячеек для

хранения), число стеллажей по горизонтали и по вертикали, число заказов, а также максимальный размер одного заказа. Товары в одном заказе заполняются случайным образом из всех доступных видов товаров для данного примера. Генератор примеров формирует представление задачи в терминах точек размещения, задаёт их вместимость, а также матрицу кратчайших расстояний между ними. Тестовые примеры генерировались в соответствии с несколькими возможностями изменения конфигурации: изменение размеров одного стеллажа, изменение числа стеллажей, изменение размера и числа заказов. Пример представляет собой три файла: информация о кратчайших расстояниях между точками размещения, входом и выходом, вместимость каждой из точки размещения, набор заказов для сборки.

Названия примеров закодированы в виде $a \times b_c_d$, где a — число стеллажей по вертикали, b — число стеллажей по горизонтали, c — вместимость одного стеллажа, d — число заказов. Диапазон числа товаров в одном заказе регулировался одинаковым образом для всех примеров в пределах до 15% от M , поэтому не задан в названии.

Описанные алгоритмы реализованы на языке программирования C. Сравнительные эксперименты проводились на компьютере с процессором AMD Ryzen 5 5600U 2,3 ГГц и 16 ГБ ОЗУ под управлением операционной системы Linux.

В ходе численных экспериментов сделана попытка решения исследуемой задачи с помощью программы-оптимизатора, для которой построена модель смешанного целочисленного линейного программирования. Так как число переменных и ограничений даже для примеров малой размерности не позволили запустить оптимизацию с помощью таких пакетов, как CPLEX и GUROBI с пробной лицензией, проведён эксперимент с использованием оптимизатора с открытым исходным кодом CBC. На примере, состоящем из 12 точек размещения вместимости 1 или 2 и пяти заказов размера от 5 до 9, оптимизатору удалось найти оптимальное решение за время, примерно равное 4 ч. Разработанным алгоритмам для нахождения решения с такой же целевой функцией потребовалось меньше секунды. В связи с этим сравнение результатов работы алгоритмов с оптимизаторами не проводилось.

4.2. Гибридный генетический алгоритм. Рассмотрено несколько вариантов гибридного генетического алгоритма, отличающихся выбором стратегии селекции (отбора особей для скрещивания).

1. Стратегия, в которой новое поколение формируется с помощью скрещивания «элитного» представителя (индивида с наименьшим значением целевой функции) с остальными членами популяции. В этом случае в новую популяцию всегда наследуются признаки элитного родителя, а для разнообразия популяции используется оператор мутации.

2. Стратегия, использующая для выбора индивидуумов метод турнирного отбора, который работает следующим образом. Каждый раз из популяции выбирается случайным образом заданное число особей (в нашем случае 2), и лучшая из них записывается в промежуточный массив родителей. Процесс повторяется, пока число родителей в массиве не станет равным числу особей в популяции. После этого особи для скрещивания выбираются из массива случайным образом. Преимущество данного метода селекции заключается в том, что он не требует дополнительного пересчёта значений целевой функции и упорядочивания индивидуумов.

3. Стратегия, использующая гибрид двух вышеописанных подходов: сначала формировался один набор индивидуумов первым методом, а затем этот набор принимался как популяция и изменялся вторым методом селекции. Это используется для большей диверсификации алгоритма и большей разности между представителями популяции.

Также был запрограммирован стандартный генетический алгоритм без использования алгоритма поиска с запретами.

На рис. 8 представлено сравнение работы разных подходов к выбору стратегии алгоритма. Как можно заметить, вторая стратегия (использующая только турнирный отбор) показала себя лучшим образом. Стандартный генетический алгоритм с разными стратегиями селекции

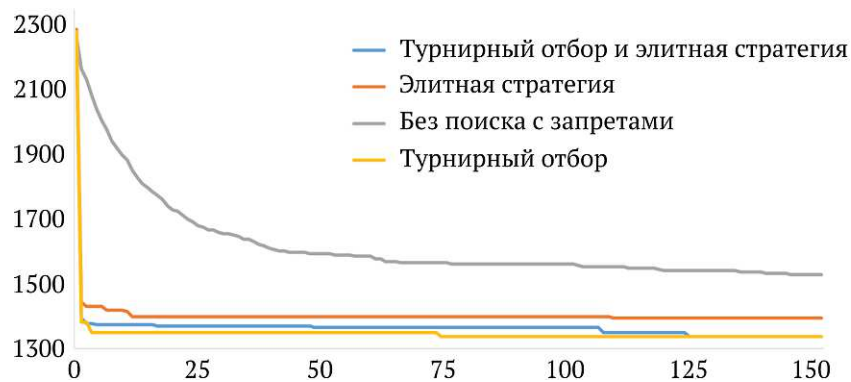


Рис. 8. Сравнение разных вариантов генетического алгоритма: по горизонтали — число итераций, по вертикали — значение целевой функции

не смог достичь значения, близкого к лучшему найденному даже на большем числе итераций.

Таблица 1

Параметры гибридного генетического алгоритма

Параметр	Значение
Размер популяции, G	M
Число особей в турнире, n_{tourn}	2
Число итераций без улучшения для всего алгоритма, I_{all}	$[0, 2M]$
Минимальное время запрета, h_{min}	$[0, 1M]$
Максимальное время запрета, h_{max}	$[0, 3M]$
Минимальная степень мутации, μ_{min}	$[0, 1M]$
Максимальная степень мутации, μ_{max}	$[0, 8M]$
Параметр рандомизированной окрестности для TS, p	0,25
Число итераций без улучшения для TS, I_{TS}	M

В табл. 1 представлены подобранные параметры гибридного генетического алгоритма. Диапазон времени запрета выбирался из промежутков разной длины от $0,1M$ до $0,5M$ по лучшему среднему результату за пять запусков алгоритма поиска с запретами на примере $5 \times 5_{-6}_{-500}$. Диапазон степени мутации выбирался из промежутков $(0,1M; 0,2M)$, $\dots$, $(0,9M; M)$ с шагом $0,1M$ при таких же условиях на запусках гибридного генетического алгоритма. В качестве критерия остановки было использовано число итераций без улучшения значения целевой функции.

4.3. Итеративный алгоритм поиска с запретами. В табл. 2 представлены подобранные параметры итеративного алгоритма поиска с запретами. Параметр P_D подбирался из диапазона $[0,05; 0,3]$ с шагом 0,05 при фиксированных остальных параметрах на одном из сгенерированных примеров. Выбор диапазона обоснован тем, что для диверсификации требуется использовать самые редкие возможные назначения, а при увеличении доли допустимых назначений туда начнут попадать уже использованные варианты изменения решения. Критерием остановки в фазе диверсификации являлось достижение итерации с номером I_D . Параметр P_I подбирался из диапазона $[0,1; 0,5]$ с шагом 0,1 при таких же условиях. В качестве критерия остановки выступало максимально допустимое число итераций без улучшения рекорда.

На рис. 9 представлен график работы итеративного алгоритма поиска с запретами на примере для 25 стеллажей с вместимостью 20 и 100 заказами. Из графика видно, что сначала алгоритм работает как стандартный алгоритм локального поиска. С течением времени он начинает

Таблица 2

Параметры итеративного алгоритма поиска с запретами

Параметр	Значение
Минимальное время запрета, $h_{\min}$	$[0,1M]$
Максимальное время запрета, $h_{\max}$	$[0,3M]$
Параметр фазы диверсификации, P_D	0,15
Параметр фазы интенсификации, P_I	0,3
Параметр рандомизированной окрестности, p_{TS}	0,25
Число итераций без улучшения для всего алгоритма, I_{all}	10
Число итераций без улучшения для TS, I_{TS}	M
Число итераций на фазе диверсификации, I_D	$[0,1M]$

попадать в локальные оптимумы, но из некоторых из них удаётся выйти и получить значения с меньшей целевой функцией. Например, перед фазой интенсификации, отмеченной зелёным цветом, алгоритм поиска с запретами заиклился. Однако в фазе интенсификации благодаря уменьшению размерности задачи удаётся получить решения с лучшими значениями целевой функции. На этапе диверсификации значение целевой функции резко возрастает, но затем алгоритм поиска с запретами запускается снова, и ему удаётся обновить рекордное решение.

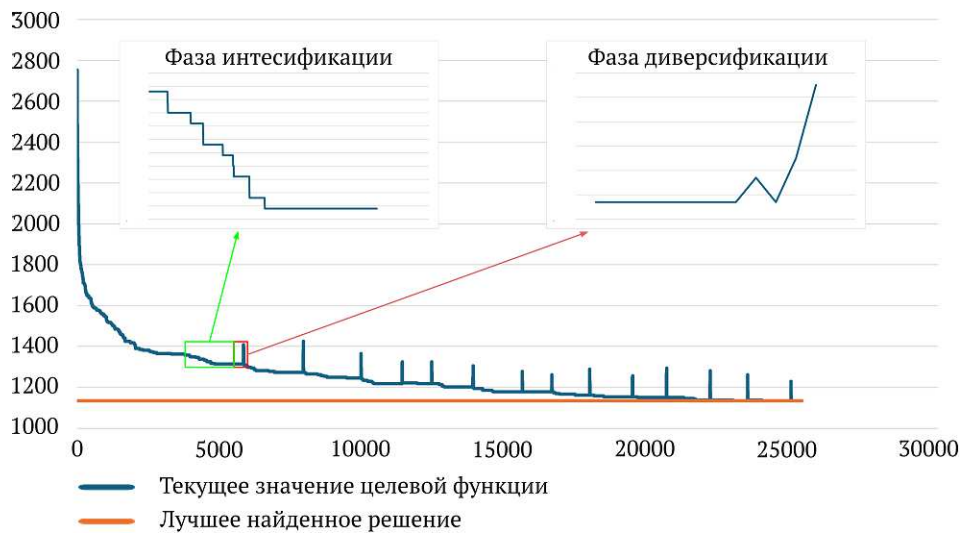


Рис. 9. Поведение алгоритма 3 на примере 5x5_20_100:
по горизонтали — номер итерации, по вертикали — значение целевой функции

Таблица 3

Численные результаты

Пример	GA + TS				ITS			
	мин.	сред.	макс.	t , с	мин.	сред.	макс.	t , с
3x3_10_10	250	250	250	4,3	250	250	250	1,7
3x3_10_50	1 006	1 006	1 006	24	1 006	1 006	1 006	7
3x3_10_100	2 592	2 592	2 592	43	2 592	2 592	2 592	12
3x3_10_500	12 324	12 334	12 348	84	12 324	12 331	12 360	32
3x3_50_10	142	142	142	58	142	142	142	38
3x3_50_50	710	713	714	512	712	716	724	134
3x3_50_100	1 540	1 540	1 540	2 868	1 540	1 540	1 540	809
3x3_50_500	10 634	10 636	10 642	6 588	10 636	10 648	10 658	3 192
5x5_6_10	218	218	218	28	218	218	218	6
5x5_6_100	4 040	4 040	4 040	412	4 040	4 040	4 040	97
5x5_6_500	24 456	24 460	24 464	1 628	24 456	24 459	24 472	649
5x5_6_1000	51 376	51 404	51 416	4 745	51 382	51 412	51 416	1 593
5x5_20_10	250	250	250	12	250	250	250	4,6
5x5_20_100	1 134	1 134	1 134	68	1 134	1 136	1 142	21
5x5_20_500	20 996	20 999	21 006	14 456	20 996	21 027	21 058	4 632
5x5_50_1000	38 876	38 877	38 880	25 573	38 880	38 880	38 880	8 923

4.4. Сравнение двух алгоритмов. В табл. 3 представлено сравнение работы разработанных алгоритмов на сгенерированных примерах. Приведены результаты пяти запусков каждого алгоритма на примерах разной размерности. В столбцах таблицы находятся минимальное, максимальное и среднее значения целевой функции, а также время работы алгоритмов в секундах.

Для примеров малой размерности можно сделать предположение, что алгоритмы получили значения глобального оптимума, так как они совпадают на всех запусках. Также стоит отметить, что на некоторых примерах большей размерности гибридный генетический алгоритм нашёл решение лучше, чем итеративный алгоритм поиска с запретами, однако для этого алгоритма потребовалось значительно больше времени.

Заключение

В представленной работе исследована задача оптимизации расположения товаров на складе, состоящем из стеллажей, путь обхода между которыми строго задан. Доказано, что задача NP-трудна даже в частном случае с двумя стеллажами и двумя типами товаров в каждом заказе. Для формализации задачи и исследования её с помощью программ-

оптимизаторов была построена модель смешанного целочисленного линейного программирования. Для решения задачи разработано два эвристических алгоритма: гибридный генетический алгоритм и итеративный алгоритм поиска с запретами. Разработан генератор примеров исследуемой задачи, на которых была протестирована и проанализирована работа предложенных алгоритмов.

Финансирование работы

Исследование выполнено в рамках гос. задания Института математики им. С. Л. Соболева СО РАН (проект № FWNF-2022-0019). Дополнительных грантов на проведение или руководство этим исследованием получено не было.

Конфликт интересов

Авторы заявляют, что у них нет конфликта интересов.

Литература

1. **Karasek J.** An overview of warehouse optimization // J. Adv. Telecommun. Electrotech. Signals Syst. 2013. V. 2, No. 3. P. 111–117.
2. **Caron F., Marchet G., Perego A.** Optimal layout in low-level picker-to-part systems // Int. J. Prod. Res. 2000. V. 38, No. 1. P. 101–117.
3. **Gu J., Goetschalckx M., McGinnis L. F.** Research on warehouse design and performance evaluation: A comprehensive review // Eur. J. Oper. Res. 2010. V. 203, No. 3. P. 539–549.
4. **De Koster R., Le-Duc T., Roodbergen K. J.** Design and control of warehouse order picking: A literature review // Eur. J. Oper. Res. 2007. V. 182, No. 2. P. 481–501.
5. **Ратушный А. А., Кочетов Ю. А.** Матэвристика для минимизации времени ожидания трейлеров при неточных временах прибытия // Дискрет. анализ и исслед. операций. 2022. Т. 29, № 3. С. 85–101.
6. **Ratliff H. D., Rosenthal A. S.** Order-picking in a rectangular warehouse: A solvable case of the traveling salesman problem // Oper. Res. 1983. V. 31, No. 3. P. 507–521.
7. **Roodbergen K. J., de Koster R.** Routing order pickers in a warehouse with a middle aisle // Eur. J. Oper. Res. 2001. V. 133, No. 1. P. 32–43.
8. **Brynzer H., Johansson M. I.** Storage location assignment: Using the product structure to reduce order picking times // Int. J. Prod. Econ. 1996. V. 46, No. 1. P. 595–603.
9. **Mantel R. J., Schuur P. C., Heragu S. S.** Order oriented slotting: A new assignment strategy for warehouses // Eur. J. Ind. Eng. 2007. V. 1, No. 3. P. 301–316.
10. **Koopmans T. C., Beckmann M.** Assignment problems and the location of economic activities // Econometrica. 1957. V. 25, No. 1. P. 53–76.

11. **Bui T. N., Chaudhuri S., Leighton F. T., Sipser M.** Graph bisection algorithms with good average case behavior // *Combinatorica*. 1987. V. 7, No. 2. P. 171–191.
12. **Гэри М., Джонсон Д.** Вычислительные машины и труднорешаемые задачи. М.: Мир, 1982. 416 с.
13. **Misevicius A., Kilda B.** Comparison of crossover operators for the quadratic assignment problem // *Inf. Technol. Control*. 2005. V. 34, No. 2. P. 109–119.
14. **Misevicius A.** An implementation of the iterated tabu search algorithm for the quadratic assignment problem // *OR Spectrum*. 2012. V. 34. P. 665–690.
15. **Glover F.** Future paths for integer programming and links to artificial intelligence // *Comput. Oper. Res.* 1986. V. 13, No. 5. P. 533–549.
16. **Glover F., Laguna M.** Tabu search // *Modern heuristic techniques for combinatorial problems*. New York: Halstead Press, 1993. P. 70–150.
17. **Lourenço H. R., Martin O. C., Stützle T.** Iterated local search // *Handbook of metaheuristics*. New York: Kluwer Acad. Publ., 2003. P. 321–353.

Моторин Кирилл Олегович
Пяткин Артём Валерьевич

Статья поступила
15 июля 2024 г.
После доработки —
5 августа 2024 г.
Принята к публикации
22 сентября 2024 г.

ON ONE OPTIMIZATION PROBLEM
FOR WAREHOUSE GOODS PLACEMENT

K. O. Motorin^{1, a} and A. V. Pyatkin^{2, b}

¹ Novosibirsk State University,

2 Pirogov Street, 630090 Novosibirsk, Russia

² Sobolev Institute of Mathematics,

4 Acad. Koptuyug Avenue, 630090 Novosibirsk, Russia

E-mail: ^ak.motorin@ng.su.ru, ^bartem@math.nsc.ru

Abstract. A warehouse goods placement problem is considered, where the aim is to minimize the total time of fulfillment of orders from a given list. NP-hardness of this problem even in the simplest special case is proved. An ILP model is suggested for this problem. Two heuristic algorithms are developed for solving this problem; their effectiveness is analyzed using randomly generated instances. Tab. 3, illustr. 9, bibliogr. 17.

Keywords: warehouse optimization, goods placement, NP-hard problem, genetic algorithm, tabu search.

References

1. **J. Karasek**, An overview of warehouse optimization, *J. Adv. Telecommun. Electrotech. Signals Syst.* **2** (3), 111–117 (2013).
2. **F. Caron, G. Marchet, and A. Perego**, Optimal layout in low-level picker-to-part systems, *Int. J. Prod. Res.* **38** (1), 101–117 (2000).
3. **J. Gu, M. Goetschalckx, and L. F. McGinnis**, Research on warehouse design and performance evaluation: A comprehensive review, *Eur. J. Oper. Res.* **203** (3), 539–549 (2010).
4. **R. de Koster, T. Le-Duc, and K. J. Roodbergen**, Design and control of warehouse order picking: A literature review, *Eur. J. Oper. Res.* **182** (2), 481–501 (2007).
5. **A. A. Ratushnyi and Yu. A. Kochetov**, Matheuristics for waiting time minimization for trailers with uncertain arrival times, *Diskretn. Anal. Issled. Oper.* **29** (3), 85–101 (2022) [Russian] [*J. Appl. Ind. Math.* **16** (3), 540–549 (2022)].

English transl.: *Journal of Applied and Industrial Mathematics* **19** (1), 77–91 (2025), DOI: 10.1134/S1990478925010089.

6. **H. D. Ratliff** and **A. S. Rosenthal**, Order-picking in a rectangular warehouse: A solvable case of the traveling salesman problem, *Oper. Res.* **31** (3), 507–521 (1983).
7. **K. J. Roodbergen** and **R. de Koster**, Routing order pickers in a warehouse with a middle aisle, *Eur. J. Oper. Res.* **133** (1), 32–43 (2001).
8. **H. Brynzer** and **M. I. Johansson**, Storage location assignment: Using the product structure to reduce order picking times, *Int. J. Prod. Econ.* **46** (1), 595–603 (1996).
9. **R. J. Mantel**, **P. C. Schuur**, and **S. S. Heragu**, Order oriented slotting: A new assignment strategy for warehouses, *Eur. J. Ind. Eng.* **1** (3), 301–316 (2007).
10. **T. C. Koopmans** and **M. Beckmann**, Assignment problems and the location of economic activities, *Econometrica* **25** (1), 53–76 (1957).
11. **T. N. Bui**, **S. Chaudhuri**, **F. T. Leighton**, and **M. Sipser**, Graph bisection algorithms with good average case behavior, *Combinatorica* **7** (2), 171–191 (1987).
12. **M. R. Garey** and **D. S. Johnson**, *Computers and Intractability. A Guide to the Theory of NP-Completeness* (Freeman, San Francisco, 1979; Mir, Moscow, 1982 [Russian]).
13. **A. Misevicius** and **B. Kilda**, Comparison of crossover operators for the quadratic assignment problem, *Inf. Technol. Control.* **34** (2), 109–119 (2005).
14. **A. Misevicius**, An implementation of the iterated tabu search algorithm for the quadratic assignment problem, *OR Spectrum* **34**, 665–690 (2012).
15. **F. Glover**, Future paths for integer programming and links to artificial intelligence, *Comput. Oper. Res.* **13** (5), 533–549 (1986).
16. **F. Glover** and **M. Laguna**, Tabu search, in *Modern Heuristic Techniques for Combinatorial Problems* (Halstead Press, New York, 1993), pp. 70–150.
17. **H. R. Lourenço**, **O. C. Martin**, and **T. Stützle**, Iterated local search, in *Handbook of Metaheuristics* (Kluwer Acad. Publ., New York, 2003), pp. 321–353.

Kirill O. Motorin
Artem V. Pyatkin

Received July 15, 2024
Revised August 5, 2024
Accepted September 22, 2024

ДВУХСТАДИЙНЫЙ АЛГОРИТМ ДЛЯ ДИНАМИЧЕСКОЙ ЗАДАЧИ УПАКОВКИ В КОНТЕЙНЕРЫ С ГРУППАМИ РАЗМЕЩЕНИЯ

А. В. Ратушный^а, Ю. А. Кочетов^б

Институт математики им. С. Л. Соболева,
пр. Акад. Коптюга, 4, 630090 Новосибирск, Россия
E-mail: ^аalexeyratushny@gmail.com, ^бjkochet@math.nsc.ru

Аннотация. Рассматривается новая задача динамической упаковки в контейнеры, актуальная для облачных вычислений. Для каждого предмета (виртуальной машины) известны время создания, удаления и требуемые ресурсы. Контейнеры (серверы) имеют NUMA-архитектуру и определённые правила при размещении машин. Серверы собраны в стойки, а некоторые машины образуют группы. Каждая группа разделена на партии. Машины из разных партий нельзя размещать на одной стойке для обеспечения отказоустойчивости системы. Требуется упаковать все машины в минимальное число стоек на заданном горизонте планирования. Для решения задачи разработан двухстадийный алгоритм: построение начального решения, в котором нарушается часть ограничений, и итеративное улучшение с помощью локального поиска, направленное на устранение нарушений. Применяя предложенный подход на открытых тестовых примерах, удалось достичь среднего отклонения от нижней границы в 3,8%. Табл. 3, ил. 4, библиогр. 27.

Ключевые слова: задача упаковки в контейнеры, виртуальная машина, конфликт, группа размещения.

Введение

Развитие облачных вычислений в последние десятилетия породило множество актуальных задач в области информационных технологий. Одним из ключевых направлений является оптимизация использования серверных ресурсов с целью повышения энергоэффективности, что включает в себя решение задачи динамической упаковки в контейнеры. В этой задаче упаковки фиксируется некоторый горизонт планирования, в течение которого необходимо оптимально распределить виртуальные машины (VM) по серверам. Все серверы идентичны и обладают NUMA-

архитектурой [1, 2], т. е. состоят из нескольких NUMA-узлов, каждый из которых характеризуется определённым объёмом оперативной памяти (RAM) и числом процессорных ядер (CPU). Каждой виртуальной машине требуются серверные ресурсы, количество которых определяется её типом. Важно уточнить, что число типов ВМ обычно значительно меньше общего числа виртуальных машин. Тип ВМ также определяет её размер — малая или большая. Малая виртуальная машина занимает ресурсы только одного NUMA-узла, тогда как большая ВМ равномерно распределяется между двумя различными NUMA-узлами. Временной интервал использования ресурсов каждой ВМ задаётся во входных данных задачи. В дополнение в задаче имеются стойки и группы размещения виртуальных машин с партициями [3], что является основным отличием от задач, рассмотренных в [4, 5]. Стойка представляет собой ограниченный набор серверов, причём каждый сервер принадлежит только одной стойке. Часть виртуальных машин объединена в группы, разделенные на непересекающиеся партиции (подмножества ВМ), находящиеся в конфликте друг с другом. Виртуальные машины из разных партиций одной группы не могут располагаться на одной стойке одновременно, а конфликт между двумя такими ВМ возникает только при пересечении их интервалов существования.

Описанная задача NP-трудна в сильном смысле, поскольку является обобщением задачи упаковки в контейнеры. Однако если рассматривать задачу разрешения конфликтов в группах ВМ с партициями изолированно, то задача раскраски графа конфликтов в этом случае становится полиномиально разрешимой. Это обусловлено возможностью раскрасить каждую партицию в отдельный цвет, что приводит к оптимальному решению. Данная особенность существенно отличает рассматриваемую задачу от сценария, где любые две ВМ могут конфликтовать, а также от задачи с конфликтами между типами ВМ, исследованной в [5]. Таким образом, наличие партиций вводит специфические структурные особенности, которые могут быть учтены при разработке алгоритмических подходов к решению данной задачи. Похожая задача рассмотрена в статье [6]. В отличие от постановки, представленной ниже, в которой рассматриваются конфликты, в [6] акцент делается на ограничениях, касающихся совместного размещения виртуальных машин из одной группы. Авторы приводят МIP модель и формулируют несколько теорем, касающихся сложности данной задачи. Оказывается, что даже при сильных предположениях, таких как наличие только одного момента времени или одной группы, задача остаётся NP-трудной.

Задача оптимизации ресурсов остаётся актуальной, причём продолжают возникать её новые вариации. Так, например, в [7] представлен широкий спектр сценариев и задач, возникающих в силу необходимости

разработки оптимальных расписаний для виртуальных машин с учётом технических особенностей современных вычислительных кластеров. Авторы работы [14] уделяют внимание важности обеспечения отказоустойчивости серверов в дата-центрах, чтобы уменьшить потери производительности и обеспечить непрерывность работы системы при возникновении сбоев. Параллельно с развитием постановок задач происходит эволюция методов их решения. В работах [9, 13] проводится анализ классических алгоритмов, таких как First Fit, Best Fit и Next Fit. В дополнение к этому авторы предлагают новые алгоритмы Hybrid First Fit и онлайн Move-to-front, построенные на основе вышеупомянутых классических подходов. Исследование [8] посвящено изучению многокритериальной постановки задачи оптимизации ресурсов. Авторы предлагают метаэвристический подход, интегрирующий жадные эвристики в адаптивную эволюционную схему. Предложенный метод демонстрирует высокую эффективность при тестировании на различных бенчмарках, показывая результаты, сопоставимые с другими современными метаэвристическими алгоритмами.

Другим важным направлением исследований является задача оптимизации энергетических, а не вычислительных ресурсов. Например, в [10] авторы детально анализируют проблемы эффективного использования энергетических ресурсов в дата-центрах и предлагают новые алгоритмы динамической упаковки: Dynamic Energy Efficient Best-Fit Decreasing и Energy Mitigation Switcher. Первый из них нацелен на оптимизацию утилизации ресурсов облачной системы, а второй выступает в качестве переключателя между различными подходами. Проведённое моделирование показывает, что эти алгоритмы могут значительно снизить общее энергопотребление и уменьшить время выполнения задач. Кроме того, в [10] применение технологий виртуализации предлагается как эффективный инструмент оптимизации распределения ресурсов. В [11] также изучается задача понижения энергопотребления, но основное внимание акцентируется на возможности динамического перераспределения виртуальных машин, что может привести к значительной экономии. Аналогично в [12] авторы рассматривают вариацию этой задачи с возможностью переупаковки виртуальных машин и предлагают алгоритм, отличительная особенность которого состоит в независимости числа переупаковок от общего числа виртуальных машин. Потенциально такой подход может повысить масштабируемость и эффективность управления ресурсами в крупных дата-центрах.

Кроме того, в последние годы приобретают популярность подходы, основанные на методах машинного обучения, в том числе это отражается на развитии программных библиотек (см., например, [15, 16]). В исследованиях [17, 18] изучаются подходы, основанные на методах обучения

с подкреплением. Авторы проводят сравнительный анализ эффективности алгоритмов машинного обучения и традиционных методов, демонстрируя возрастающую конкурентоспособность первых, что подчёркивает их потенциал в решении практических задач оптимизации. В [19] представлен обзор ограничений, присущих традиционным методам машинного обучения при решении задачи упаковки в контейнеры. Авторы показывают, как их подход может преодолеть эти ограничения благодаря обширному набору обучающих данных и точной разметке. В [20] предлагается метод решения задачи о рюкзаке, основанный на применении нейронных сетей. Приводится подробное описание архитектуры сети, принимающей на вход все доступные параметры и выдающей оптимальный набор предметов. Результаты демонстрируют потенциал данного подхода в контексте решения реальных задач. В [21] поднимается вопрос планирования в сфере распределённых вычислений. Авторы представляют прототип онлайн-планировщика, основанного на гибридном подходе, включающем в себя методы машинного обучения и динамического программирования.

Для решения рассматриваемой задачи предлагается двухстадийный подход, основанный на итеративном исправлении решения. На первом шаге строится некоторое исходное решение с помощью простого алгоритма, а затем ломается таким образом, чтобы нарушалась часть ограничений, связанных с конфликтами. На втором шаге для каждой группы определяются все конфликты, которые были нарушены, и запускается процедура их исправления. Каждая виртуальная машина с недопустимым расположением меняется местами с некоторой другой виртуальной машиной таким образом, чтобы уменьшить общее число нарушенных ограничений в текущей рассматриваемой группе. Если избавиться от всех конфликтов в группе не получается, то часть этой группы подвергается повторному размещению. Основной научный вклад данной работы заключается в исследовании новой NP-трудной задачи. Разработана математическая модель, которая учитывает различные аспекты и позволяет формально описать данную задачу. Кроме того, представлены алгоритмы получения верхних оценок. Проведённые вычислительные эксперименты на открытых данных продемонстрировали высокую эффективность предложенных методов, что подтверждает их практическую применимость для решения рассматриваемой задачи. Аналогичная задача рассмотрена в [22], где авторы предложили стохастический жадный алгоритм. Сравнение с данным подходом производится в разделе с вычислительными экспериментами.

Далее в разд. 1 приводятся необходимые обозначения, детальная постановка задачи и математическая модель. Разд. 2 содержит алгоритмы построения верхних оценок. Разд. 3 описывает используемые данные,

а также результаты численных экспериментов. В заключении кратко подводятся итоги исследования.

1. Математическая модель

Для постановки задачи и построения её математической модели введём следующие обозначения:

- T — множество моментов времени или горизонт планирования;
- $R = \{\text{RAM, CPU}\}$ — множество типов ресурсов;
- F — множество стоек;
- S_f — множество серверов, принадлежащих стойке $f \in F$;
- $S = \bigcup_{f \in F} S_f$ — множество всех серверов;
- N — множество NUMA-узлов.

Стойки представляют собой наборы серверов, причём вместимость каждой стойки ограничена и, как правило, не превышает 10–20 единиц. Каждый сервер должен быть ассоциирован ровно с одной стойкой, иначе говоря, принадлежать ей. Хотя число серверов в стойках может варьироваться, основная цель задачи заключается в минимизации общего числа используемых стоек. Следовательно, в рамках данной постановки задачи можно всегда использовать максимальное заполнение каждой стойки серверами. Все серверы идентичны, но важно отметить, что различные NUMA-узлы одного и того же сервера могут отличаться по количеству доступных ресурсов.

Для работы с группами и виртуальными машинами введём следующие обозначения:

- M^S — множество малых виртуальных машин;
- M^L — множество больших виртуальных машин;
- $M = M^S \cup M^L$ — множество всех виртуальных машин;
- G — множество групп виртуальных машин;
- $M_g \subset M$ — множество виртуальных машин из группы $g \in G$;
- P_g — множество партиций в группе $g \in G$;
- $M_{gp} \subset M$ — множество ВМ из партиции $p \in P_g$ группы $g \in G$;
- K — множество типов виртуальных машин.

Группа представляет собой множество виртуальных машин: например, это могут быть виртуальные машины одного клиента. Каждая группа разделена дополнительно на подмножества, называемые партициями. Число партиций в группе обычно невелико и не превышает 10. Ключевое ограничение заключается в том, что любые две виртуальные машины из различных партиций одной группы не могут быть размещены на одной стойке (рис. 1). Таким образом, партиции можно мыслить как частный случай конфликтов между ВМ. Однако структура графа таких конфликтов (где вершины — виртуальные машины, а рёбра — конфликты между

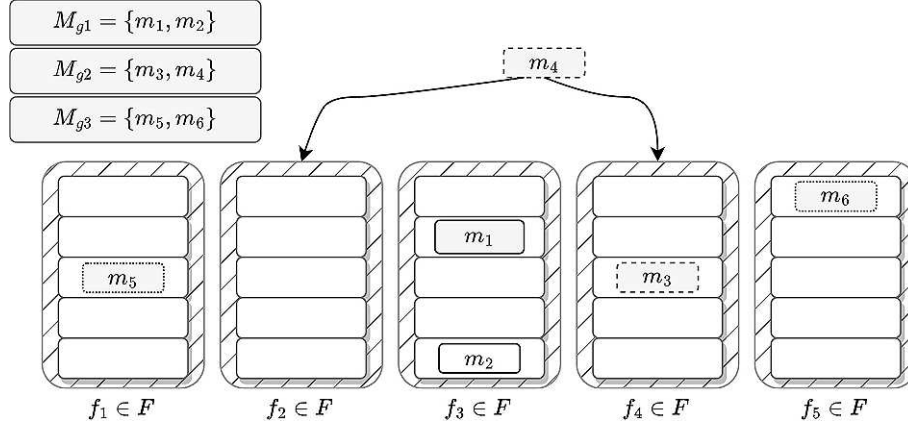


Рис. 1. Пример упаковки группы с тремя партициями

ними) достаточно специфична. Граф разбивается на несколько компонент связности, соответствующих группам. Каждую компоненту связности можно эффективно раскрасить в минимальное число цветов, просто присваивая отдельный цвет каждой партиции. Тип определяет количество памяти RAM и число ядер CPU, которые ВМ занимает на сервере. Множества типов виртуальных машин у большинства компаний, предоставляющих услуги облачных вычислений, имеют значительное пересечение. Пример такого множества типов можно найти в [24]. Также следует отметить, что мощность множества типов ВМ, как правило, намного меньше мощности множества виртуальных машин.

Следующие обозначения вводятся для описания характеристик виртуальных машин и серверов:

- C_{nr} — количество ресурса $r \in R$ на NUMA-узле $n \in N$;
- d_{mr} — количество ресурса $r \in R$, необходимого виртуальной машине для одного NUMA-узла;
- d_{kr} — количество ресурса $r \in R$, необходимого типу $k \in K$ для одного NUMA-узла;
- α_m — время создания машины $m \in M$;
- ω_m — время удаления машины $m \in M$.

Каждая виртуальная машина $m \in M^S$ ($m \in M^L$) занимает d_{mr} ($2d_{mr}$) ресурса $r \in R$ на сервере во все моменты времени $\alpha_m \leq t < \omega_m$. Для любого ресурса $r \in R$ справедливо равенство $d_{mr} = d_{kr}$, если виртуальной машине $m \in M$ приписан тип $k \in K$. Во входных данных гарантируется, что $d_{mr} \leq C_{nr}$ для любых $n \in N$, $r \in R$, $m \in M$. Важно уточнить, что если между двумя ВМ m_1 и m_2 есть конфликт, но они не пересекаются по времени, то данный конфликт можно не рассматривать.

С учётом всего вышесказанного множество конфликтов D можно задать следующим образом. Пара ВМ (m_1, m_2) , обозначающая конфликт, принадлежит D , если $m_1 \in M_{gp_1}$ и $m_2 \in M_{gp_2}$, где p_1 и p_2 — различные партии $(p_1 \neq p_2)$ некоторой группы g , и существует момент времени $t \in T$ такой, что $\alpha_{m_1} \leq t < \omega_{m_1}$ и $\alpha_{m_2} \leq t < \omega_{m_2}$.

Определим также следующие множества ВМ, функционирующих в заданный момент времени $t \in T$:

- $M_t^S = \{m \in M^S \mid \alpha_m \leq t < \omega_m\}$ — множество малых ВМ;
- $M_t^L = \{m \in M^L \mid \alpha_m \leq t < \omega_m\}$ — множество больших ВМ;
- $M_t = M_t^S \cup M_t^L$ — множество всех ВМ.

Для описания математической модели понадобятся следующие булевы переменные. Для произвольных $f \in F$, $s \in S_f$, $n \in N$ определим

$$x_{fsm} = \begin{cases} 1, & \text{если малая или первая часть большой ВМ } m \in M \\ & \text{расположена на узле } n \text{ сервера } s \text{ в стойке } f, \\ 0 & \text{иначе;} \end{cases}$$

$$y_{fsm} = \begin{cases} 1, & \text{если вторая часть большой ВМ } m \in M^L \text{ распо-} \\ & \text{ложена на узле } n \text{ сервера } s \text{ в стойке } f, \\ 0 & \text{иначе;} \end{cases}$$

$$z_f = \begin{cases} 1, & \text{если стойка } f \text{ была активна хотя бы в один} \\ & \text{момент времени } t \in T, \\ 0 & \text{иначе.} \end{cases}$$

В указанных обозначениях математическую модель рассматриваемой задачи можно записать следующим образом:

$$\min \sum_{f \in F} z_f, \quad (1)$$

$$\sum_{f \in F} \sum_{s \in S_f} \sum_{n \in N} x_{fsm} = 1, \quad m \in M, \quad (2)$$

$$x_{fsm} + y_{fsm} \leq 1, \quad f \in F, s \in S_f, n \in N, m \in M^L, \quad (3)$$

$$\sum_{n \in N} x_{fsm} = \sum_{n \in N} y_{fsm}, \quad f \in F, s \in S_f, m \in M^L, \quad (4)$$

$$\sum_{m \in M_t} d_{mr} x_{fsm} + \sum_{m \in M_t^L} d_{mr} y_{fsm} \leq C_{nr},$$

$$f \in F, s \in S_f, n \in N, t \in T, r \in R, \quad (5)$$

$$\sum_{s \in S_f} \sum_{n \in N} x_{fsm} \leq z_f, \quad f \in F, m \in M, \quad (6)$$

$$\sum_{s \in S_f} x_{f s n m_1} + \sum_{s \in S_f} x_{f s n m_2} \leq 1, \quad (m_1, m_2) \in D, f \in F, n \in N, \quad (7)$$

$$x_{f s n m}, y_{f s n m}, z_f \in \{0, 1\}. \quad (8)$$

Целевая функция (1) выражает минимизацию числа задействованных стоек. Ограничения (2) гарантируют, что все виртуальные машины упакованы, а (3)–(4) отвечают за правильную упаковку больших ВМ. Неравенства (5) ограничивают количество ресурсов на каждом сервере, а (6) необходимы для подсчёта числа стоек. Считаем, что стойка была активна, если на ней обслуживалась хотя бы одна виртуальная машина. Неравенства (7) задают конфликты между виртуальными машинами.

Потребность в решении задачи для достаточно больших примеров (более 10^4 виртуальных машин) не позволяет использовать математическую модель (1)–(8), так как она требует значительных вычислительных ресурсов и времени. В качестве альтернативы предлагается использовать хорошо зарекомендовавшие себя эвристические подходы [5, 25]. Описание рассматриваемого алгоритма приводится в разд. 2.

2. Верхние оценки

При решении задачи в онлайн-постановке существует необходимость упаковывать виртуальные машины в порядке их создания, учитывая конфликты непосредственно на этапе размещения. Динамическая постановка обладает большей свободой действий, что позволяет разрабатывать более гибкие и эффективные эвристические методы. Предлагаемый подход основан на двухэтапном процессе оптимизации. На первом этапе формируется начальное решение задачи, игнорирующее часть ограничений, связанных с конфликтами. На втором этапе применяется процедура коррекции, направленная на устранение конфликтов. Разделение поиска решения на два этапа позволило достичь большей гибкости и качества итогового размещения.

На первом этапе применяется рандомизированная версия алгоритма First Fit [23] для размещения групп виртуальных машин (RFFG). Алгоритм последовательно упаковывает группы ВМ в случайном порядке одну за другой и не приступает к упаковке следующей группы, пока не закончит с текущей. При этом упаковка каждой группы производится таким образом, чтобы не возникло конфликтов между партициями одной группы. Если в примере имеются ВМ, не принадлежащие какой-либо группе, то они упаковываются последними. Случайный порядок групп позволяет достичь заметно лучшего результата, сохраняя при этом производительность алгоритма. С помощью такого подхода можно генерировать несколько потенциальных решений, из которых затем выбирается наилучшее. Важно отметить, что порядок ВМ внутри каждой группы

также случаен, что дополнительно увеличивает вариативность решений. Далее для создания возможности дальнейшего улучшения распаковывается некоторый процент $\widehat{U}$ стоек, после чего неразмещённые ВМ перепакуются рандомизированным алгоритмом First Fit (RFF), при этом игнорируются ограничения на конфликты.

Определим схожесть виртуальных машин с помощью коэффициентов Жаккара $J(m_1, m_2)$ [26], в частности, его адаптации для прямоугольников IoU (intersection over union). В нашем случае ширина прямоугольника (виртуальной машины m) равняется $\omega_m - \alpha_m$, а высота определяется относительным значением требуемого ресурса $d_{mr} / \sum_{n \in N} C_{nr}$ ($2d_{mr} / \sum_{n \in N} C_{nr}$ в случае большой ВМ). Тем самым коэффициент $J(m_1, m_2, r)$ зависит не только от выбранных виртуальных машин m_1 и m_2 , но и от рассматриваемого ресурса $r \in R$. Для того чтобы избавиться от данной неоднозначности, будем рассматривать адаптацию этого коэффициента:

$$J(m_1, m_2) = \min_{r \in R} J(m_1, m_2, r).$$

На втором шаге имеющееся решение итеративно перестраивается так, чтобы последовательно избавляться от нарушений ограничений. Группы виртуальных машин просматриваются одна за другой в случайном порядке. Для каждой ВМ m анализируется возможность перестановки с другими ВМ $\widehat{m}$, исключая из рассмотрения те, которые удовлетворяют некоторым правилам: находятся на той же стойке f , что и m ; принадлежат уже просмотренным группам; входят в ту же группу и партицию, что и m ; или имеют низкую степень схожести с m ($J(m, \widehat{m}) < \widehat{J}$). Эти критерии отсева позволяют сфокусироваться на наиболее перспективных перестановках, которые могут уменьшить число нарушений ограничений без риска добавления новых конфликтов. Перестановка двух виртуальных машин не увеличивает целевую функцию, однако такой подход не всегда может свести к нулю число нарушенных ограничений. Это приводит к необходимости дополнительного шага. Если на некоторой итерации (полном просмотре всех виртуальных машин в группе) не удалось полностью избавиться от нарушений ограничений, то часть данной группы полностью перепаковывается. В таком случае на каждой стойке остаётся только та партиция из группы, которая имеет наибольшее число виртуальных машин, а все остальные ВМ удаляются и упаковываются заново с помощью RFF. Однако в этот раз алгоритм упаковывает так, чтобы не нарушить никаких ограничений, т. е. для каждой ВМ определяется некоторый набор запрещённых стоек, куда упаковка не производится. При необходимости создаются новые стойки и серверы, так что алгоритм всегда возвращает допустимое решение. Схему алгоритма можно представить следующим образом.

Алгоритм RALR

ШАГ 1. С помощью алгоритма RFFG строится начальное допустимое решение, не содержащее нарушений по ограничениям (2)–(8).

ШАГ 2. Случайный набор из $\hat{U}\%$ стоек распаковывается, а освободившиеся ВМ упаковываются заново с помощью RFF с возможным нарушением ограничений на конфликты между партициями.

ШАГ 3. Для очередной группы и для каждой ВМ в ней запускается локальный поиск в пространстве перестановок местами с другими ВМ таким образом, чтобы уменьшить число нарушенных ограничений. Каждая группа просматривается полностью несколько раз до тех пор, пока удаётся устранить нарушения. Если удалось устранить все нарушения, то переходим к следующей группе, иначе переходим на шаг 4.

ШАГ 4. Если не удалось устранить все нарушения для рассматриваемой группы, то она частично перепаковывается алгоритмом RFF, который исправляет все ограничения, при необходимости создавая новые сервера и стойки.

Пример поведения целевой функции и числа нарушенных ограничений во время шагов 3–4 можно найти на рис. 2. Здесь обе кривые нормализованы к интервалу $[0, 1]$. Кривая отклонения целевой функции показывает, как менялось её значение относительно решения, полученного на втором шаге. Резкие скачки на обеих кривых соответствуют переупаковкам групп, что, как можно видеть, не всегда приводит к ухудшению решения.

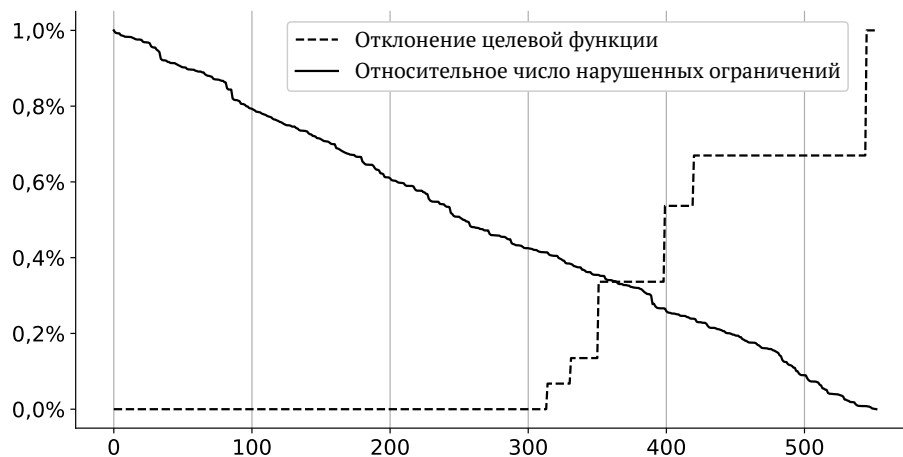


Рис. 2. Зависимость целевой функции и числа нарушенных ограничений от номера итерации

Следует отметить, что шаги 2–4 иногда способны ухудшить изначальное решение. В случае, если такие ухудшения действительно наблюдаются, то решение, полученное методом RFFG на первом шаге возвращается в качестве окончательного. Однако результаты численных экспериментов показывают, что такая ситуация возникает достаточно редко.

Сложность шагов 1, 2 и 4 имеет порядок $O(|M||N|T|\bar{S}|)$, что сопоставимо со сложностью алгоритма RFF, который учитывает динамику и NUMA-архитектуру. Параметр $\bar{S}$ ограничивает число серверов сверху: например, можно использовать $\bar{S} = |M|$. Шаг 3 наиболее трудоёмкий в алгоритме RALR. Для каждой пары виртуальных машин необходимо проверить допустимость обмена их местами, что требует $O(|M|^2|T||N|)$ операций. Количество полных просмотров одной группы ограничено числом возможных конфликтов порядка $O(|M|^2)$. На практике число итераций ограничивается сверху. В проведённых вычислительных экспериментах граница равна 20, хотя алгоритму обычно требуется не более 10 итераций для достижения локального минимума. Несмотря на сложность, благодаря различным оптимизациям и ограничениям размера окрестностей локального поиска шаг 3 выполняется достаточно быстро для получения допустимого решения в обозримое время. Например, поиск решения заметно ускоряется за счёт значения порога схожести двух виртуальных машин $\hat{J}$. Многие пары виртуальных машин (m_1, m_2) имеют достаточно низкий коэффициент $J(m_1, m_2)$, например из-за сильного разброса по времени. Попытка переставить их местами скорее всего закончится неудачей. Даже небольшое значение $\hat{J}$ позволяет эффективно исключить

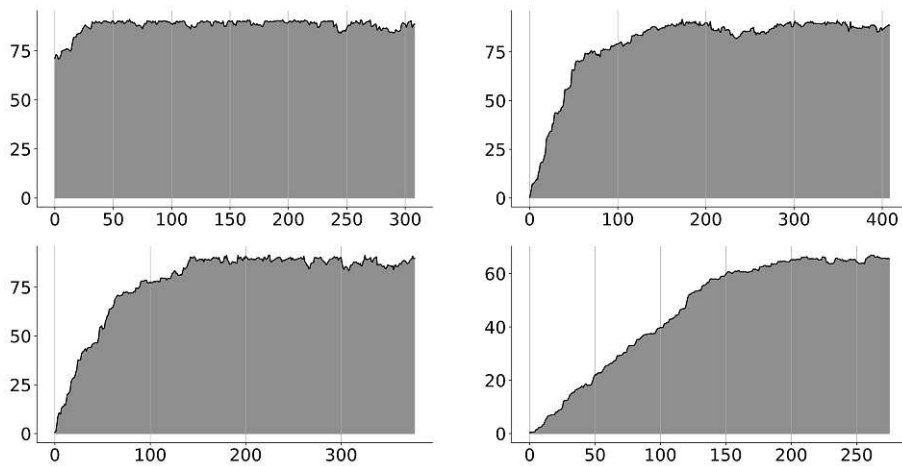


Рис. 3. Нижняя оценка по моментам времени для различных примеров

из рассмотрения большую часть неудачных пар и значительно уменьшить размер окрестности.

3. Численные эксперименты

Для проведения численных экспериментов применён случайный генератор, предоставленный одной из компаний, занимающихся облачными вычислениями. Генератор имитирует поведение реальной нагрузки на серверы. На каждом шаге он создаёт запросы на размещение виртуальных машин последовательно, стараясь, чтобы загрузка кластера (ограниченного набора стоек) находилась в определённом интервале. Новые группы размещения и типы виртуальных машин генерируются в соответствии с частотами, полученными в результате анализа истории работы одного из кластеров той же компании. Примеры суммарной нагрузки по числу стоек можно увидеть на рис. 3. По оси абсцисс отложено время, а по оси ординат — минимально необходимое число стоек, вычисленное по формуле

$$\max_{r \in R} \frac{\sum_{m \in M_t^S} d_{mr} + 2 \sum_{m \in M_t^L} d_{mr}}{\sum_{n \in N} C_{nr}}.$$

Все рассмотренные примеры находятся в открытом доступе [27].

Анализ примеров и различные статистики приведены в табл. 1. Все примеры поделены на несколько семейств. Обозначения LPO, DMP и MP в названии семейства показывают тип примеров, где LPO — примеры только с большими группами партиций, MP — примеры с различными группами партиций (с большими и малыми), а DMP — обобщение примеров MP с добавлением виртуальных машин, которые не принадлежат никакой группе, т. е. не имеют конфликтов ни с какими другими ВМ. Буквенные обозначения «l», «m» и «s» соответствуют размеру примера, т. е. числу виртуальных машин и количеству моментов времени, а обозначения «2» и «4» показывают число NUMA-узлов на серверах. Максимальный размер стойки во всех примерах равен 10 серверам. В среднем во всех примерах по 5 партиций в группе. Каждое семейство состоит из 25 примеров, а суммарно рассмотрено 450 примеров. В табл. 1 приводятся средние значения по всем примерам каждого семейства.

Для сравнения выбраны следующие алгоритмы:

- RALR — алгоритм, описанный в разд. 2, с параметрами $\hat{J} = 0,25$ и $\hat{U} = 66\%$;
- INIT + NMSLI — стохастический алгоритм, приведённый в [22]; этот подход основан на алгоритме First Fit и методе бисекции;

Таблица 1

Анализ примеров

Семейство примеров	Среднее число ВМ	Среднее число моментов времени	Среднее число групп с партициями	Среднее число ВМ в группе	Среднее число ВМ без конфликтов
LPO_l_2	64 703,12	338,28	605,92	106,85	0,00
LPO_l_4	68 349,76	280,40	671,84	101,80	0,00
LPO_m_2	28 825,24	153,72	272,20	105,89	0,00
LPO_m_4	27 328,32	115,04	271,76	100,59	0,00
LPO_s_2	13 090,92	70,28	124,56	105,20	0,00
LPO_s_4	12 781,92	55,24	124,92	102,20	0,00
DMP_l_2	38 074,36	379,68	354,36	69,60	13 452,12
DMP_l_4	42 577,24	311,80	425,72	63,27	15 668,76
DMP_m_2	32 460,80	343,88	273,60	69,51	13 452,12
DMP_m_4	32 767,40	276,00	271,36	62,99	15 668,76
DMP_s_2	22 066,76	274,48	124,88	68,95	13 452,12
DMP_s_4	23 220,84	235,72	121,88	62,06	15 668,76
MP_l_2	44 820,88	308,88	641,40	69,96	0,00
MP_l_4	49 286,20	217,60	737,04	66,91	0,00
MP_m_2	19 213,04	136,16	273,84	70,08	0,00
MP_m_4	18 679,64	84,28	278,64	66,98	0,00
MP_s_2	8 838,44	63,32	126,16	69,98	0,00
MP_s_4	8 213,20	37,24	121,92	67,46	0,00

• RFFG — первый шаг алгоритма RALR; метод является вариацией алгоритма First Fit, в которой группы упаковываются одна за другой в случайном порядке таким образом, чтобы не нарушались ограничения на конфликты, а виртуальные машины вне групп упаковываются так же в случайном порядке в конце алгоритма;

• RAR — упрощённый вариант алгоритма RALR, в котором отсутствует самый трудоёмкий шаг 3;

• RFF — рандомизированный First Fit алгоритм, который упаковывает виртуальные машины в случайном порядке, не учитывает конфликты и нецелесообразно, чтобы оценить дополнительную сложность примеров, вызванную наличием групп.

Все алгоритмы реализованы на языке программирования Python, вычисления проводились на компьютере с процессором AMD EPYC 7502 32-core. В каждом примере вычислено относительное отклонение по формуле $gap = (UB - LB) / UB$, где UB — это результат некоторого алгоритма,

Таблица 2

Сравнение алгоритмов

Семейство примеров	RALR		INIT + MSLI		RFFG		RAR	
	сред.	стд.	сред.	стд.	сред.	стд.	сред.	стд.
LPO_l_2	5,14	0,87	4,36	0,89	12,39	0,68	13,64	1,09
LPO_l_4	6,12	1,33	7,57	0,56	9,93	0,49	13,49	0,80
LPO_m_2	1,49	2,21	1,55	1,51	6,85	1,91	9,98	1,85
LPO_m_4	4,52	2,37	8,25	1,89	8,04	1,36	12,68	1,72
LPO_s_2	3,27	3,09	3,90	2,85	6,24	3,22	9,63	2,69
LPO_s_4	6,27	2,72	11,95	3,34	9,32	2,75	14,52	2,73
DMP_l_2	3,02	1,10	2,11	0,87	7,34	1,48	8,26	1,41
DMP_l_4	3,06	0,46	3,31	0,61	4,36	1,00	4,56	1,94
DMP_m_2	1,94	1,21	1,42	1,03	6,42	1,82	7,33	1,93
DMP_m_4	2,85	0,87	3,01	0,91	4,35	1,11	4,40	3,33
DMP_s_2	1,43	1,13	1,19	1,16	5,21	2,08	6,03	1,93
DMP_s_4	3,10	1,19	3,11	1,24	4,40	1,54	4,95	1,34
MP_l_2	3,96	0,78	3,13	0,63	9,79	0,73	11,16	0,80
MP_l_4	5,81	1,00	6,58	0,64	7,64	0,57	10,64	0,70
MP_m_2	2,20	1,21	2,02	1,49	7,00	2,22	9,93	2,67
MP_m_4	5,45	1,74	7,91	1,26	7,10	1,58	12,22	3,93
MP_s_2	3,05	2,48	3,85	3,60	5,94	3,79	11,05	3,88
MP_s_4	6,41	3,65	11,26	3,87	7,58	2,69	12,23	2,64
Среднее	3,84	1,63	4,80	1,57	7,22	1,72	9,81	2,08

а LB — нижняя оценка, полученная с помощью генерации столбцов для задачи упаковки в контейнеры с одним моментом времени без конфликтов [5]. В табл. 2 для каждого алгоритма и семейства примеров представлено по два значения в процентах: среднее значение гар (сред.) по семейству примеров и стандартное отклонение гар (стд.), которое иллюстрирует стабильность работы алгоритма на данном подмножестве примеров. Так как во всех алгоритмах присутствует фактор случайности, они запускались до достижения установленного ограничения по времени. Вне зависимости от семейства на каждый пример размера l выделялось 3 часа, на размер m — 2 часа, а на размер s — всего 1 час. Исключение составил алгоритм INIT + MSLI, параметры запуска которого описаны более подробно в [22].

Анализ результатов показывает, что алгоритм RALR демонстрирует значительно лучшие результаты по сравнению с RFFG и RAR на всех представленных наборах данных, даже несмотря на то, что RFFG и RAR более эффективны и успевали обработать больше запусков за отведённое

время. Наблюдаемая разница в отклонениях подчёркивает значимость наличия всех этапов, реализованных в RALR. Тем не менее, несмотря на свою простоту, алгоритм RFF смог продемонстрировать среднее отклонение, равное 2,51%, что заметно ниже, чем у других подходов. Нужно отметить, что RFF не учитывает наличие конфликтов, и выявленная разница, вероятно, обусловлена дополнительной сложностью примеров, связанной с наличием групп и партиций. Кроме того, при анализе результатов работы RFF замечено, что на примерах с четырьмя NUMA-узлами он имеет более высокие значения гар, чем на примерах с двумя NUMA-узлами. По всей видимости, примеры с четырьмя NUMA-узлами оказываются несколько сложнее, и для них обеспечение высокой утилизации ресурсов серверов вызывает больше сложностей. Важно, что такое наблюдение сделано именно для алгоритма RFF, поскольку этот подход не учитывает некоторые дополнительные ограничения и имеет более простую структуру. Дополнительно алгоритму RALR удалось улучшить результаты алгоритма INIT + MSLI на части примеров. В среднем получилось уменьшить гар более чем на 1%. Однако большая часть данной разницы связана с семействами примеров LPO_s_4, LPO_m_4 и MP_s_4, которые оказались для INIT + MSLI достаточно сложными, даже в сравнении с RFFG.

Таблица 3

Влияние параметра $\widehat{U}$ на отклонение гар

Семейство примеров	Значение $\widehat{U}$ (%)						
	10	25	33	50	66	75	90
	Значение гар (%)						
DMP_l_2	6,62	5,30	4,16	3,46	3,02	4,19	4,62
DMP_l_4	4,29	3,65	3,62	3,32	3,06	3,69	3,73
DMP_m_2	4,34	3,47	2,82	2,39	1,94	2,52	2,80
DMP_m_4	4,33	3,47	3,29	3,12	2,85	3,59	3,68
DMP_s_2	3,71	2,97	2,41	2,23	1,43	1,80	1,89
DMP_s_4	4,25	3,82	3,57	3,40	3,10	3,42	3,48
Среднее	4,59	3,78	3,31	3,02	2,57	3,20	3,37

В ходе исследования вызвал интерес также анализ влияния параметра $\widehat{U}$ на эффективность работы алгоритма RALR. Результаты данного анализа представлены в табл. 3 и на рис. 4. При относительно малых значениях параметра $\widehat{U}$ результаты RALR схожи с результатами RFFG. Это означает, что при такой малой переупаковке алгоритму RALR не удастся эффективно исправлять нарушения, и он часто возвращает решение, полученное на первом шаге. При увеличении значения параметра $\widehat{U}$

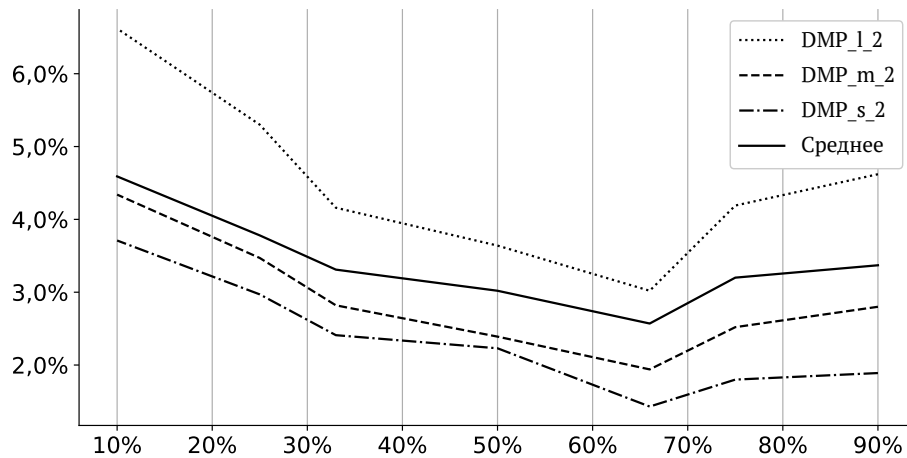


Рис. 4. Зависимость отклонения гар от значения параметра $\hat{U}$

до слишком больших величин качество алгоритма RALR также остаётся низким. В данном случае разрушается слишком большая часть решения. Упаковка с помощью RFF значительной части ВМ, по всей видимости, оказывается недостаточно хорошей для того, чтобы её можно было эффективно исправить. Это приводит к слишком частому использованию процедуры переупаковки групп с шага 4, что дополнительно снижает качество.

Заключение

В представленной работе рассмотрена динамическая задача упаковки в контейнеры с учётом групп размещения, которые накладывают ограничения на расположение виртуальных машин. Представлена математическая модель задачи, продемонстрирована работа нескольких эвристических методов, а также проведено их сравнение на большом объёме открытых примеров. Кроме того, выполнен анализ одного из алгоритмов для более глубокого понимания его поведения.

Предложенный метод в дальнейшем может быть обобщён на другие модификации задачи с учётом развития облачных сервисов и появления новых возможностей для клиентов. Например, можно исследовать дополнительные ограничения на группы размещения, такие как конфликты на уровне серверов вместо стоек, или требования по близости размещения виртуальных машин внутри одной группы [6]. Также перспективным направлением является улучшение нижних оценок, так как текущий подход не учитывает влияние групп размещений.

Финансирование работы

Исследование выполнено в рамках гос. задания Института математики им. С. Л. Соболева (проект № FWNF-2022-0019). Дополнительных грантов на проведение или руководство этим исследованием получено не было.

Конфликт интересов

Авторы заявляют, что у них нет конфликта интересов.

Литература

1. **Lameter C.** NUMA (Non-uniform memory access): An overview // Queue. 2013. V. 11. P. 40–51. DOI: 10.1145/2508834.2513149.
2. **Li P., Wu X., Luo Y., Wang L.-M., Yadav N., Pepin M., Morgan J.** NUMAware: Accelerate VM-to-VM I/O performance in NUMA servers for NFV applications // IEEE Conf. Network Function Virtualization and Software Defined Networks (Palo Alto, CA, USA, Nov. 7–9, 2016). Piscataway: IEEE, 2016. URL: nfvsdn2016.ieee-nfvsdn.org/program/demonstrations/index.html (accessed: 20.08.2024).
3. Placement groups for your Amazon EC2 instances // Amazon elastic compute cloud: User guide. Seattle: Amazon Web Serv., 2024. URL: docs.aws.amazon.com/AWSEC2/latest/UserGuide/placement-groups.html (accessed: 20.08.2024).
4. **Ratushnyi A. V., Kochetov Yu. A.** A column generation based heuristic for a temporal bin packing problem // Mathematical optimization theory and operations research. Proc. 20th Int. Conf. MOTOR 2021 (Irkutsk, Russia, July 5–10, 2021). Cham: Springer, 2021. P. 96–110. (Lect. Notes Comput. Sci.; V. 12755). DOI: 10.1007/978-3-030-77876-7_7.
5. **Ratushnyi A. V.** A pattern-based heuristic for a temporal bin packing problem with conflicts // Mathematical optimization theory and operations research. Proc. 22nd Int. Conf. MOTOR 2023 (Yekaterinburg, Russia, July 2–8, 2023). Cham: Springer, 2023. P. 161–175. (Commun. Comput. Inf. Sci.; V. 1881). DOI: 10.1007/978-3-031-43257-6_13.
6. **Borisovsky P. A., Ereemeev A. V., Panin A. A., Sakhno M. A.** Temporal bin packing problems with placement constraints: MIP-models and complexity // Mathematical optimization theory and operations research. Proc. 23rd Int. Conf. MOTOR 2024 (Omsk, Russia, June 30–July 6, 2024). Cham: Springer, 2024. P. 157–169. (Lect. Notes Comput. Sci.; V. 14766). DOI: 10.1007/978-3-031-62792-7_11.
7. **Grushin D. A., Kuzyurin N. N.** On effective scheduling in computing clusters // Program. Comput. Softw. 2019. V. 45. P. 398–404. DOI: doi.org/10.1134/S0361768819070077.
8. **Spencer K. Y., Tsvetkov P. V., Jarrell J. J.** A greedy memetic algorithm for a multiobjective dynamic bin packing problem for storing cooling objects // J. Heuristics. 2019. V. 25. P. 1–45. DOI: 10.1007/s10732-018-9382-0.

9. **Li Y., Tang X., Cai W.** Dynamic bin packing for on-demand cloud resource allocation // *IEEE Trans. Parallel Distrib. Syst.* 2016. V. 27, No. 1. P. 157–170. DOI: 10.1109/TPDS.2015.2393868.
10. **Gupta N., Gupta K., Gupta D.** [et al.]. Enhanced virtualization-based dynamic bin-packing optimized energy management solution for heterogeneous clouds // *Math. Probl. Eng.* 2022. V. 2022, No. 1. Article ID 8734198. 11 p. DOI: 10.1155/2022/8734198.
11. **Beloglazov A., Buyya R.** Energy efficient allocation of virtual machines in cloud data centers // 10th IEEE/ACM Int. Conf. Cluster, Cloud and Grid Computing (Melbourne, Australia, May 17–20, 2010). Piscataway: IEEE, 2014. P. 577–578. DOI: 10.1109/CCGRID.2010.45.
12. **Berndt S., Jansen K., Klein K.-M.** Fully dynamic bin packing revisited // *Math. Program.* 2020. V. 179. P. 109–155. DOI: 10.1007/s10107-018-1325-x.
13. **Murhekar A., Arbour D., Mai T., Rao A.** Brief announcement: Dynamic vector bin packing for online resource allocation in the cloud // *Proc. 35th ACM Symp. Parallelism in Algorithms and Architectures* (Orlando, FL, USA, June 17–19, 2023). New York: ACM, 2023. P. 307–310. DOI: 10.1145/3558481.3591314.
14. **Daudjee K., Kamali S., López-Ortiz A.** On the online fault-tolerant server consolidation problem // *Proc. 26th ACM Symp. Parallelism in Algorithms and Architectures* (Prague, Czech Rep., June 23–25, 2014). New York: ACM, 2014. P. 12–21. DOI: 10.1145/2612669.2612686.
15. **Hubbs C., Perez H., Sarwar O., Sahinidis N., Grossmann I., Wasick J.** OR-Gym: A reinforcement learning library for operations research problem. Ithaca, NY: Cornell Univ., 2020. 29 p. (Cornell Univ. Libr. e-Print Archive; arXiv:2008.06319). DOI: 10.48550/arXiv.2008.06319.
16. **Wan C., Li T., Wang J.** RLOR: A flexible framework of deep reinforcement learning for operation research. Ithaca, NY: Cornell Univ., 2023. 21 p. (Cornell Univ. Libr. e-Print Archive; arXiv:2303.13117). DOI: 10.48550/arXiv.2303.13117.
17. **Balaji B., Bell-Masterson J., Bilgin E.** [et al.]. ORL: Reinforcement learning benchmarks for online stochastic optimization problems. Ithaca, NY: Cornell Univ., 2019. 12 p. (Cornell Univ. Libr. e-Print Archive; arXiv:1911.10641). DOI: 10.48550/arXiv.1911.10641.
18. **Yang T., Luo F., Fuentes J., Ding W., Gu C.** A flexible reinforced bin packing framework with automatic slack selection // *Math. Probl. Eng.* 2021. V. 2021. Article ID 6653586. 15 p. DOI: 10.1155/2021/6653586.
19. **Mao F., Blanco E., Fu M.** [et al.]. Small boxes big data: A deep learning approach to optimize variable sized bin packing // *Proc. IEEE 3rd Int. Conf. Big Data Computing Service and Applications* (San Francisco, CA, USA, Apr. 6–9, 2017). Piscataway: IEEE, 2017. P. 80–89. DOI: 10.1109/BigDataService.2017.18.
20. **Nomer H. A., Alnowibet K. A., Elsayed A., Mohamed A. W.** Neural knapsack: A neural network based solver for the knapsack problem // *IEEE Access.* 2020. V. 8. P. 224200–224210. DOI: 10.1109/ACCESS.2020.3044005.

21. **Toporkov V. V., Yemelyanov D. M., Bulkhak A. N.** Machine learning-based online scheduling in distributed computing // Parallel processing and applied mathematics. Rev. Sel. Pap. 14th Int. Conf. PPAM 2022 (Gdansk, Poland, Sept. 11–14, 2022). Pt. II. Cham: Springer, 2023. P. 248–259. (Lect. Notes Comput. Sci.; V. 13827). DOI: 10.1007/978-3-031-30445-3_21.
22. **Turnaev A. M., Panin A. A.** Stochastic greedy algorithms for a temporal bin packing problem with placement groups // Mathematical optimization theory and operations research. Proc. 23rd Int. Conf. MOTOR 2024 (Omsk, Russia, June 30 – July 6, 2024). Cham: Springer, 2024. P. 199–211. (Lect. Notes Comput. Sci.; V. 14766). DOI: 10.1007/978-3-031-62792-7_14.
23. **Johnson D. S.** Fast algorithms for bin packing // J. Comput. Syst. Sci. 1974. V. 8. P. 272–314. DOI: 10.1016/S0022-0000(74)80026-7.
24. Amazon EC2 instance types // Amazon elastic compute cloud documentation. Seattle: Amazon Web Serv., 2024. URL: docs.aws.amazon.com/ec2/latest/instancetypes/instance-types.html (accessed: 20.08.2024).
25. **Munien C., Ezugwu A.** Metaheuristic algorithms for one-dimensional bin-packing problems: A survey of recent advances and applications // J. Intell. Syst. 2021. V. 30. P. 636–663. DOI: 10.1515/jisys-2020-0117.
26. **Kosub S.** A note on the triangle inequality for the Jaccard distance // Pattern Recognit. Lett. 2019. V. 120. P. 36–38. DOI: 10.1016/j.patrec.2018.12.007.
27. Temporal bin packing problem with placement groups // Discrete location problems. Benchmark library. Novosibirsk: Sobolev Inst. Math., 2024. URL: old.math.nsc.ru/AP/benchmarks/Temporal%20Bin%20Packing/binpack.html (accessed: 20.08.2024).

Ратушный Алексей Владленович
Кочетов Юрий Андреевич

Статья поступила
11 сентября 2024 г.
После доработки —
17 сентября 2024 г.
Принята к публикации
22 сентября 2024 г.

A TWO-STAGE ALGORITHM FOR THE DYNAMIC BIN PACKING PROBLEM WITH PLACEMENT GROUPS

A. V. Ratushnyi^a and Yu. A. Kochetov^b

Sobolev Institute of Mathematics,
4 Acad. Koptuyug Avenue, 630090 Novosibirsk, Russia
E-mail: ^aalexeyratushny@gmail.com, ^bjkochet@math.nsc.ru

Abstract. A new dynamic bin packing problem relevant to cloud computing is considered. The creation time, deletion time, and required resources are known for each item (virtual machine). The containers (servers) have a NUMA architecture and specific rules for placing the machines. The servers are grouped into racks, and some machines form groups. Each group is divided into partitions. Machines from different partitions cannot be placed on the same rack to ensure system fault tolerance. The objective is to pack all the machines into the minimum number of racks over a given planning horizon. A two-stage algorithm is developed to solve the problem: an initial solution is constructed, where some constraints may be violated, followed by iterative improvement using local search aimed at eliminating the violations. Using the proposed approach, an average deviation of 3.8% from the lower bound was achieved on open test cases. Tab. 3, illustr. 4, bibliogr. 27.

Keywords: bin packing problem, virtual machine, conflict, placement group.

References

1. C. Lameter, NUMA (Non-uniform memory access): An overview, *Queue* **11**, 40–51 (2013), DOI: 10.1145/2508834.2513149.
2. P. Li, X. Wu, Y. Luo, L.-M. Wang, N. Yadav, M. Pepin, and J. Morgan, NUMAware: Accelerate VM-to-VM I/O performance in NUMA servers for NFV applications, in *IEEE Conf. Network Function Virtualization and Software Defined Networks, Palo Alto, CA, USA, Nov. 7–9, 2016* (IEEE, Piscataway, 2016), URL: nfvsdn2016.ieee-nfvsdn.org/program/demonstrations/index.html (accessed: 20.08.2024).

English transl.: *Journal of Applied and Industrial Mathematics* **19** (1), 92–103 (2025), DOI: 10.1134/S1990478925010090.

3. Placement groups for your Amazon EC2 instances, in *Amazon Elastic Compute Cloud: User Guide* (Amazon Web Serv., Seattle, 2024), URL: docs.aws.amazon.com/AWSEC2/latest/UserGuide/placement-groups.html (accessed: 20.08.2024).
4. **A. V. Ratushnyi** and **Yu. A. Kochetov**, A column generation based heuristic for a temporal bin packing problem, in *Mathematical Optimization Theory and Operations Research* (Proc. 20th Int. Conf. MOTOR 2021, Irkutsk, Russia, July 5–10, 2021) (Springer, Cham, 2021), pp. 96–110 (Lect. Notes Comput. Sci., Vol. 12755), DOI: 10.1007/978-3-030-77876-7_7.
5. **A. V. Ratushnyi**, A pattern-based heuristic for a temporal bin packing problem with conflicts, in *Mathematical Optimization Theory and Operations Research* (Proc. 22nd Int. Conf. MOTOR 2023, Yekaterinburg, Russia, July 2–8, 2023) (Springer, Cham, 2023), pp. 161–175 (Commun. Comput. Inf. Sci., Vol. 1881), DOI: 10.1007/978-3-031-43257-6_13.
6. **P. A. Borisovsky**, **A. V. Ereemeev**, **A. A. Panin**, and **M. A. Sakhno**, Temporal bin packing problems with placement constraints: MIP-models and complexity, in *Mathematical Optimization Theory and Operations Research* (Proc. 23rd Int. Conf. MOTOR 2024, Omsk, Russia, June 30–July 6, 2024) (Springer, Cham, 2024), pp. 157–169 (Lect. Notes Comput. Sci., Vol. 14766), DOI: 10.1007/978-3-031-62792-7_11.
7. **D. A. Grushin** and **N. N. Kuzyurin**, On effective scheduling in computing clusters, *Program. Comput. Softw.* **45**, 398–404 (2019), DOI: doi.org/10.1134/S0361768819070077.
8. **K. Y. Spencer**, **P. V. Tsvetkov**, and **J. J. Jarrell**, A greedy memetic algorithm for a multiobjective dynamic bin packing problem for storing cooling objects, *J. Heuristics* **25**, 1–45 (2019), DOI: 10.1007/s10732-018-9382-0.
9. **Y. Li**, **X. Tang**, and **W. Cai**, Dynamic bin packing for on-demand cloud resource allocation, *IEEE Trans. Parallel Distrib. Syst.* **27** (1), 157–170 (2016), DOI: 10.1109/TPDS.2015.2393868.
10. **N. Gupta**, **K. Gupta**, **D. Gupta**, [et al.], Enhanced virtualization-based dynamic bin-packing optimized energy management solution for heterogeneous clouds, *Math. Probl. Eng.* **2022** (1), ID 8734198 (2022), DOI: 10.1155/2022/8734198.
11. **A. Beloglazov** and **R. Buyya**, Energy efficient allocation of virtual machines in cloud data centers, in *10th IEEE/ACM Int. Conf. Cluster, Cloud and Grid Computing, Melbourne, Australia, May 17–20, 2010* (IEEE, Piscataway, 2014), pp. 577–578, DOI: 10.1109/CCGRID.2010.45.
12. **S. Berndt**, **K. Jansen**, and **K.-M. Klein**, Fully dynamic bin packing revisited, *Math. Program.* **179**, 109–155 (2020), DOI: 10.1007/s10107-018-1325-x.
13. **A. Murhekar**, **D. Arbour**, **T. Mai**, and **A. Rao**, Brief announcement: Dynamic vector bin packing for online resource allocation in the cloud, in *Proc. 35th ACM Symp. Parallelism in Algorithms and Architectures, Orlando, FL, USA, June 17–19, 2023* (ACM, New York, 2023), pp. 307–310, DOI: 10.1145/3558481.3591314.

14. **K. Daudjee, S. Kamali, and A. López-Ortiz**, On the online fault-tolerant server consolidation problem, in *Proc. 26th ACM Symp. Parallelism in Algorithms and Architectures, Prague, Czech Rep., June 23–25, 2014* (ACM, New York, 2014), pp. 12–21, DOI: 10.1145/2612669.2612686.
15. **C. Hubbs, H. Perez, O. Sarwar, N. Sahinidis, I. Grossmann, and J. Wassick**, OR-Gym: A reinforcement learning library for operations research problem (Cornell Univ., Ithaca, NY, 2020) (Cornell Univ. Libr. e-Print Archive; arXiv:2008.06319), DOI: 10.48550/arXiv.2008.06319.
16. **C. Wan, T. Li, and J. Wang**, RLOR: A flexible framework of deep reinforcement learning for operation research (Cornell Univ., Ithaca, NY, 2023) (Cornell Univ. Libr. e-Print Archive; arXiv:2303.13117), DOI: 10.48550/arXiv.2303.13117.
17. **B. Balaji, J. Bell-Masterson, E. Bilgin, [et al.]**, ORL: Reinforcement learning benchmarks for online stochastic optimization problems (Cornell Univ., Ithaca, NY, 2019) (Cornell Univ. Libr. e-Print Archive; arXiv:1911.10641), DOI: 10.48550/arXiv.1911.10641.
18. **T. Yang, F. Luo, J. Fuentes, W. Ding, and C. Gu**, A flexible reinforced bin packing framework with automatic slack selection, *Math. Probl. Eng.* **2021**, ID 6653586 (2021), DOI: 10.1155/2021/6653586.
19. **F. Mao, E. Blanco, M. Fu, [et al.]**, Small boxes big data: A deep learning approach to optimize variable sized bin packing, in *Proc. IEEE 3rd Int. Conf. Big Data Computing Service and Applications, San Francisco, CA, USA, Apr. 6–9, 2017* (IEEE, Piscataway, 2017), pp. 80–89, DOI: 10.1109/BigDataService.2017.18.
20. **H. A. Nomer, K. A. Alnowibet, A. Elsayed, and A. W. Mohamed**, Neural knapsack: A neural network based solver for the knapsack problem, *IEEE Access* **8**, 224200–224210 (2020), DOI: 10.1109/ACCESS.2020.3044005.
21. **V. V. Toporkov, D. M. Yemelyanov, and A. N. Bulkhak**, Machine learning-based online scheduling in distributed computing, in *Parallel Processing and Applied Mathematics* (Rev. Sel. Pap. 14th Int. Conf. PPAM 2022, Gdansk, Poland, Sept. 11–14, 2022) (Springer, Pt. II. Cham, 2023), pp. 248–259 (Lect. Notes Comput. Sci., Vol. 13827), DOI: 10.1007/978-3-031-30445-3_21.
22. **A. M. Turnaev and A. A. Panin**, Stochastic greedy algorithms for a temporal bin packing problem with placement groups, in *Mathematical Optimization Theory and Operations Research* (Proc. 23rd Int. Conf. MOTOR 2024, Omsk, Russia, June 30–July 6, 2024) (Springer, Cham, 2024), pp. 199–211 (Lect. Notes Comput. Sci., Vol. 14766), DOI: 10.1007/978-3-031-62792-7_14.
23. **D. S. Johnson**, Fast algorithms for bin packing, *J. Comput. Syst. Sci.* **8**, 272–314 (1974), DOI: 10.1016/S0022-0000(74)80026-7.
24. Amazon EC2 instance types, in *Amazon Elastic Compute Cloud Documentation* (Amazon Web Serv., Seattle, 2024), URL: docs.aws.amazon.com/ec2/latest/instancetypes/instance-types.html (accessed: 20.08.2024).
25. **C. Munien and A. Ezugwu**, Metaheuristic algorithms for one-dimensional bin-packing problems: A survey of recent advances and applications, *J. Intell. Syst.* **30**, 636–663 (2021), DOI: 10.1515/jisys-2020-0117.

-
- 26. S. Kosub**, A note on the triangle inequality for the Jaccard distance, *Pattern Recognit. Lett.* **120**, 36–38 (2019), DOI: [10.1016/j.patrec.2018.12.007](https://doi.org/10.1016/j.patrec.2018.12.007).
- 27.** Temporal bin packing problem with placement groups, in *Discrete Location Problems. Benchmark Library* (Sobolev Inst. Math., Novosibirsk, 2024), URL: old.math.nsc.ru/AP/benchmarks/Temporal%20Bin%20Packing/binpack.html (accessed: 20.08.2024).

Aleksey V. Ratushnyi

Yury A. Kochetov

Received September 11, 2024

Revised September 17, 2024

Accepted September 22, 2024

О ЧИСЛЕ ВЕЧНОГО ДОМИНИРОВАНИЯ ПЛАНАРНЫХ ГРАФОВ ДИАМЕТРА 2

Д. С. Талецкий

Национальный исследовательский университет «Высшая школа экономики»,
ул. Большая Печёрская, 25/12, 603155 Нижний Новгород, Россия
E-mail: dmitailmail@gmail.com

Аннотация. Вечным доминирующим множеством графа называется доминирующее множество D , на котором располагается первоначально мобильная охрана (не более одного охранника может находиться в каждой вершине). Для любой бесконечной последовательности атак на вершины графа множество D может быть модифицировано путём передвижения охранника со смежной вершины в атакуемую вершину (предполагается, что атакуемая вершина не была занята охранником во время атаки). Конфигурация охранников должна после каждой атаки и движения охранника образовывать доминирующее множество. Числом вечного доминирования графа называется мощность его наименьшего вечного доминирующего множества. Доказано, что число вечного доминирования каждого планарного графа диаметра 2 равно числу его кликового покрытия. Ил. 5, библиогр. 10.

Ключевые слова: доминирующее множество, вечное доминирующее множество, число вечного доминирования, планарный граф.

Введение

Множество D вершин графа называется *доминирующим*, если каждая вершина не из D смежна хотя бы с одной вершиной из D . Доминирующее множество D_0 называется *вечным доминирующим* в графе G , если для любого набора вершин $v_1, \dots, v_k \in V(G)$ найдутся доминирующие множества $D_1, \dots, D_k$ и вершины $u_0 \in D_0, u_1 \in D_1, \dots, u_{k-1} \in D_{k-1}$ такие, что $D_{i+1} = (D_i \setminus \{u_i\}) \cup \{v_{i+1}\}$ при всех $i \in \overline{0, k-1}$. Числом вечного доминирования $\gamma^\infty(G)$ графа G называется мощность его наименьшего вечного доминирующего множества.

Элементы вечного доминирующего множества обычно представляются в виде охранников, расположенных в вершинах графа (не более одного

охранника в каждой вершине) и способных отразить последовательность атак на незанятые вершины сколь угодно большой длины. Предполагается, что после каждой атаки ровно один охранник перемещается по ребру в соседнюю атакованную вершину, а остальные охранники остаются на своих местах. Отметим, что существует и другая модель вечного доминирования (известная как « t -вечное доминирование»), в которой после каждой атаки на одну из вершин все охранники могут переместиться по ребру в соседнюю вершину. Рассмотрение этой модели выходит за рамки данной работы.

Впервые вечные доминирующие множества были рассмотрены в [1] в 2004 г. В [2] получены значения чисел вечного доминирования для графов из некоторых классов. В [3] доказано неравенство $\gamma^\infty(G) \leq \binom{\alpha(G)}{2}$ (здесь через $\alpha(G)$ обозначается число независимости графа G), а в [4] приведены примеры графов, для которых $\gamma^\infty(G) = \binom{\alpha(G)}{2}$. В работе [5] доказано неравенство $\gamma^\infty(G \boxtimes H) \geq \alpha(G)\gamma^\infty(H)$, где через $G \boxtimes H$ обозначается сильное произведение графов G и H . Статья [6] и книга [7] содержат подробный обзор результатов, связанных с вечным доминированием в графах.

Числом доминирования $\gamma(G)$ называется мощность наименьшего доминирующего множества графа G , а числом кликового покрытия $\theta(G)$ — наименьшее число клик, на которые можно разбить множество вершин этого графа. Как известно, число кликового покрытия графа равно хроматическому числу его дополнения. В работе [1] впервые упоминается тривиальное соотношение

$$\gamma(G) \leq \alpha(G) \leq \gamma^\infty(G) \leq \theta(G).$$

В самом деле, для отражения последовательности атак, состоящей из вершин некоторого наибольшего независимого множества G , требуется $\alpha(G)$ охранников, откуда $\alpha(G) \leq \gamma^\infty(G)$. С другой стороны, вершины графа G можно разбить на $\theta(G)$ клик, для защиты каждой из которых достаточно одного охранника, откуда $\gamma^\infty(G) \leq \theta(G)$.

Класс графов, для которых $\gamma^\infty(G) = \theta(G)$, оказался довольно обширным и уже долгое время привлекает внимание исследователей. Будем обозначать этот класс через $\mathcal{MD}$ (от англ. maximum demand). Известно, что все последовательно-параллельные (а значит, и все внешнепланарные) графы входят в $\mathcal{MD}$ [8] и существуют графы без треугольников, не входящие в $\mathcal{MD}$ [2]. Поскольку для любого совершенного графа G верно $\alpha(G) = \theta(G)$, все такие графы входят в $\mathcal{MD}$. В [9] доказано, что класс $\mathcal{MD}$ содержит бесконечное число графов H таких, что прямое произведение $H \square K_2$ не входит в $\mathcal{MD}$. В недавней работе [10] найдены минимальные по числу вершин графы G_1 и G_2 , не входящие в $\mathcal{MD}$ (рис. 1).

Эти графы содержат по 10 вершин и непланарные, причём

$$3 = \gamma^\infty(G_1) = \gamma^\infty(G_2) < \theta(G_1) = \theta(G_2) = 4.$$

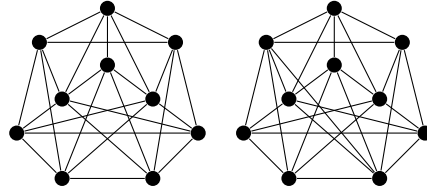


Рис. 1. Графы G_1 и G_2

Гипотеза о принадлежности классу $\mathcal{MD}$ всех планарных графов на сегодняшний день остаётся открытой. В настоящей работе доказано, что этому классу принадлежат все планарные графы диаметра 2.

1. Терминология и обозначения

Как обычно, через $V(G)$, $E(G)$ и $v(G)$ обозначаются множество вершин, множество рёбер и число вершин графа G соответственно. Через $N(v)$, $N[v]$, $\deg(v)$ и $d(u, v)$ обозначаются открытая окрестность, замкнутая окрестность, степень вершины v и длина кратчайшего пути между вершинами u и v соответственно (предполагается, что рассматриваемый граф ясен из контекста). Диаметр $d(G)$ связного графа G определяется равенством $d(G) = \max_{u, v \in V(G)} d(u, v)$. Простой цикл с k вершинами называ-

ется k -циклом. Граф называется (вершинно) k -связным, если он имеет более k вершин и остаётся связным после удаления из него любых $k - 1$ вершин. Пусть $u \in S \subset V(G)$, тогда S -приватной окрестностью вершины u называется множество $pn[u, S]$ вершин $w \in V(G)$ таких, что $N[w] \cap S = \{u\}$.

Граф называется планарным, если существует его укладка на плоскости без пересечений рёбер не по вершинам; изображение планарного графа на плоскости без пересечений рёбер называется плоским графом. Будем предполагать, что все рассматриваемые планарные графы уложены на плоскости. Через $\mathcal{PL}_2$ будем обозначать класс планарных графов диаметра 2.

Пусть плоский граф G содержит цикл C . Через $\text{int}(C)$ (соответственно $\text{ext}(C)$) обозначим множество вершин, находящихся в части плоскости, ограниченной C (соответственно в части плоскости, не ограниченной C). Назовём цикл разделяющим, если оба этих множества непусты. Заметим, что если $d(G) = 2$, то каждая вершина из $\text{int}(C)$ имеет общего соседа в C с каждой вершиной из $\text{ext}(C)$. Будем говорить, что разделяющий цикл C

отделяет множество $A \subset V(G)$, если одно из множеств $\text{int}(C)$ и $\text{ext}(C)$ совпадает с A . Если $A = \{a\}$, то будем говорить, что цикл C отделяет вершину a .

Напомним, что *числом кликового покрытия* $\theta(G)$ графа G называется наименьшее число клик, на которое можно разбить множество $V(G)$. *Наименьшим кликовым покрытием* (НКП) $\mathcal{C}$ графа G называется произвольное разбиение множества $V(G)$ на $\theta(G)$ клик. Будем называть НКП графа u -отделимым, если оно содержит клику, состоящую из единственной вершины $u \in V(G)$.

Назовём *критическим* наименьший по числу вершин граф $G \in \mathcal{PL}_2$ такой, что $\gamma^\infty(G) < \theta(G)$. Цель дальнейших рассуждений — доказать, что критических графов не существует.

2. Предварительные результаты

Следующий факт, установленный в [10], будет часто использоваться в работе для сокращения перебора.

Лемма 1. Пусть граф G планарен. Тогда

- 1) если $v(G) < 12$, то $G \in \mathcal{MD}$;
- 2) если $v(G) < 14$ и G вершинно 3-связен, то $G \in \mathcal{MD}$.

Докажем три простых свойства критических графов, которые также будут применяться в дальнейших рассуждениях.

Лемма 2. Если граф $G \in \mathcal{PL}_2$ критический, то для любой вершины $w \in V(G)$ существует w -отделимое НКП G .

ДОКАЗАТЕЛЬСТВО. Рассмотрим граф $G - w$, полученный из G удалением вершины w . Хорошо известно, что $\gamma^\infty(G - w) \leq \gamma^\infty(G)$ (действительно, если некоторое множество охранников может отразить любую последовательность атак на вершины множества $V(G)$, то оно может отразить и любую последовательность атак на вершины множества $V(G) \setminus \{w\}$). Поскольку G критический, $\gamma^\infty(G - w) = \theta(G - w) < \theta(G)$. Тогда для любого НКП $\mathcal{C}$ графа $G - w$ семейство $\mathcal{C} \cup \{\{w\}\}$ является w -отделимым НКП графа G . Лемма 2 доказана.

Лемма 3. Если граф $G \in \mathcal{PL}_2$ содержит вершины x, y такие, что $N[x] \subseteq N[y]$, то G не критический.

ДОКАЗАТЕЛЬСТВО. Предположим, что G критический. Рассмотрим некоторое y -отделимое НКП $\mathcal{C}$ графа G и его клику $Cl_x \in \mathcal{C}$, содержащую вершину x . Тогда множество $Cl_x \cup \{y\}$ тоже является кликой, что невозможно в силу минимальности $\mathcal{C}$; противоречие. Лемма 3 доказана.

Лемма 4. Если граф $G \in \mathcal{PL}_2$ не содержит разделяющих 3-циклов и для некоторой вершины $w \in V(G)$ верно $\deg(w) \geq v(G) - 4$, то G не критический.

ДОКАЗАТЕЛЬСТВО. Предположим, что G критический, и рассмотрим некоторое его w -отделимое НКП $\mathcal{C}$. Поскольку G планарен и не содержит разделяющих 3-циклов, все клики $\mathcal{C}$ содержат не более 3 вершин. Тогда из условия $v(G) \geq 12$ следует, что $\theta(G - w) \geq 4$ и $\theta(G) \geq 5$. Так как $\mathcal{C}$ содержит хотя бы 4 клики, не содержащие w , хотя бы в одной из них все вершины смежны с w , что противоречит минимальности $\mathcal{C}$. Лемма 4 доказана.

3. Разделяющие 3-циклы

Лемма 5. Если в графе $G \in \mathcal{PL}_2$ содержится разделяющий 3-цикл $C = u_1u_2u_3$, то G не критический.

ДОКАЗАТЕЛЬСТВО. Напомним, что из условия $d(G) = 2$ следует, что каждая вершина множества $\text{int}(C) \cup \text{ext}(C)$ смежна хотя бы с одной вершиной цикла C . Рассмотрим 3 случая.

СЛУЧАЙ 1. Найдётся вершина $v_1 \in \text{int}(C) \cup \text{ext}(C)$, смежная ровно с одной вершиной C (можем считать без ограничения общности, что $u_1v_1 \in E(G)$ и $v_1 \in \text{int}(C)$). Тогда для любой вершины $x \in \text{ext}(C)$ верно $xu_1 \in E(G)$, откуда $N[x] \subseteq N[u_1]$; противоречие по лемме 3.

СЛУЧАЙ 2. Найдётся вершина w , смежная со всеми вершинами C (без ограничения общности считаем, что $w \in \text{int}(C)$). Если $\deg(w) = 3$, то $N[w] \subseteq N[u_1]$; противоречие по лемме 3. Иначе хотя бы один из 3-циклов u_1u_2w , u_2u_3w , u_1u_3w (считаем, что u_1u_2w) разделяющий и найдётся вершина $y \in \text{int}(u_1u_2w)$. Из предыдущего случая вытекает, что каждая вершина множества $\text{int}(u_1u_2w)$ смежна с обеими вершинами u_1 и u_2 . Таким образом, $N[y] \subseteq N[u_1]$; противоречие по лемме 3.

СЛУЧАЙ 3. Каждая вершина множества $\text{int}(C) \cup \text{ext}(C)$ смежна ровно с двумя вершинами цикла C . В силу симметрии можем считать, что $\text{int}(C) \geq 4$ и найдутся вершины $v_1, v_2 \in \text{int}(C)$, смежные с u_1 и u_2 . Тогда либо $v_1 \in \text{int}(u_1u_2v_2)$, либо $v_2 \in \text{int}(u_1u_2v_1)$ (пусть, например, $v_1 \in \text{int}(u_1u_2v_2)$). Каждая вершина множества $N(v_1) \setminus \{u_1, u_2\}$ смежна с обеими вершинами u_1 и u_2 , откуда $N[v_1] \subseteq N[u_1]$; противоречие по лемме 3. Лемма 5 доказана.

4. Разделяющие 4-циклы

Во всех утверждениях этого пункта рассматривается граф $G \in \mathcal{PL}_2$ и его разделяющий 4-цикл $C = u_1u_2u_3u_4$. Предполагается, что G не содержит разделяющих 3-циклов. Через $\text{int}^*(C)$ обозначим подмножество

вершин $\text{int}(C)$, смежных с единственной парой соседних вершин C . Будем обозначать через $u_{a,b}$ вершину $\text{int}^*(C)$, смежную с вершинами u_a и u_b цикла C . Заметим, что в силу отсутствия разделяющих 3-циклов в G каждая такая вершина единственна, откуда $|\text{int}^*(C)| \leq 4$.

Лемма 6. *Если каждая вершина множества $\text{int}(C) \cup \text{ext}(C)$ смежна хотя бы с двумя вершинами C , то G не критический.*

ДОКАЗАТЕЛЬСТВО. Предположим, что G критический. По лемме 1 имеем $v(G) \geq 12$. Тогда без ограничения общности можем считать, что $|\text{int}(C)| \geq 4$. Рассмотрим 4 случая.

СЛУЧАЙ 1. Найдётся вершина $x \in \text{int}(C)$, смежная со всеми вершинами C . Поскольку $|\text{int}(C)| \geq 4$, хотя бы один из циклов u_1u_2x , u_2u_3x , u_3u_4x , u_4u_1x разделяющий; противоречие по предыдущей лемме.

СЛУЧАЙ 2. Найдётся вершина $x \in \text{int}(C)$, смежная ровно с тремя вершинами C . В силу симметрии считаем, что $xu_4 \notin E(G)$. Рассмотрим два подслучая.

СЛУЧАЙ 2.1. Все вершины множества $\text{int}(C) \setminus \{x\}$ смежны с u_1 . Рассмотрим некоторое u_1 -отделимое НКП $\mathcal{C}$ графа G . Поскольку G планарен и $|\text{int}(C)| \geq 4$, найдётся клика $Cl \in \mathcal{C}$, содержащая хотя бы одну вершину из $\text{int}(C)$ и не содержащая вершину u_3 . Тогда $Cl \cup \{u_1\}$ тоже клика, что противоречит минимальности $\mathcal{C}$.

СЛУЧАЙ 2.2. Найдётся вершина $w \in \text{int}(C) \setminus \{x\}$, не смежная с u_1 . Тогда w смежна с u_3 и u_4 . Без ограничения общности считаем, что найдётся также вершина $w' \in \text{int}(C)$, не смежная с u_3 , но смежная с u_1 и u_4 (если это не так, то все вершины $\text{int}(C) \setminus \{x\}$ смежны с u_3 ; тогда переименуем вершины и применим рассуждения случая 2.1). Поскольку G не содержит разделяющих 3-циклов и $|\text{int}(G)| \geq 4$, найдётся также вершина $w'' \in \text{int}(C)$, смежная с u_1 и u_3 , но не смежная с u_4 . Поскольку $\deg(u_1) \leq v(G) - 5$ по лемме 4, в $\text{ext}(C)$ найдутся хотя бы две вершины, не смежные с u_1 , которые в силу равенства $d(G) = 2$ обе смежны с u_3 и u_4 (чтобы иметь общего соседа с w'' и w' соответственно), но тогда G содержит 3-цикл; противоречие по лемме 5.

СЛУЧАЙ 3. Все вершины множества $\text{int}(C)$ смежны ровно с двумя вершинами C . Поскольку G планарен, можем считать без ограничения общности, что никакая вершина из $\text{int}(G)$ не смежна одновременно с вершинами u_2 и u_4 . Рассмотрим несколько случаев в зависимости от мощности множества $\text{int}^*(C)$.

СЛУЧАЙ 3.1: $|\text{int}^*(C)| = 4$.

СЛУЧАЙ 3.1.1: $\text{int}(C) = \text{int}^*(C)$. Тогда $v(G) = 12$ и $|\text{ext}(C)| = 4$. Предполагаем, что каждая вершина $\text{ext}(C)$ смежна ровно с двумя вершинами C (иначе применим рассуждения случаев 1 и 2 этой леммы).

Тогда каждая вершина $\text{ext}(C)$ смежна с двумя противоположными вершинами C , а поскольку G планарен, то либо все вершины $\text{ext}(C)$ смежны с u_1 и u_3 , либо все они смежны с u_2 и u_4 , откуда $\max\{\deg(u_1), \deg(u_2)\} = v(G) - 4$; противоречие по лемме 4.

СЛУЧАЙ 3.1.2: $\text{int}(C) \neq \text{int}^*(C)$. Тогда найдётся некоторая вершина $x \in \text{int}(C)$, смежная с u_1 и u_3 . По лемме 4 множество $\text{ext}(C)$ содержит вершину y , не смежную с u_1 , и вершину z , не смежную с u_3 (эти вершины различны, поскольку они должны иметь общего соседа с x). Так как $d(G) = 2$, то $u_2y, u_2z, u_4y, u_4z \in E(G)$. Тогда из отсутствия в G разделяющих 3-циклов следует, что никакая вершина множества $\text{ext}(C) \setminus \{y, z\}$ не может быть смежна с вершинами u_1 и u_3 , откуда $|\text{ext}(C)| = 2$ и $|\text{int}(C)| \geq 6$. Таким образом, найдётся вершина $x' \in \text{int}(C) \setminus \text{int}^*(C)$, отличная от x и смежная с u_1 и u_3 . Тогда вершина x не является общим соседом вершин $u_{1,2}$ и $u_{3,4}$. Так как $u_{1,2} \in \text{int}(u_1u_2u_3x)$ и $u_{3,4} \in \text{int}(u_3u_4u_1x)$, то $d(u_{1,2}, u_{3,4}) > 2$; противоречие.

СЛУЧАЙ 3.2: $|\text{int}^*(C)| = 3$. Считаем без ограничения общности, что $\text{int}^*(C) = \{u_{1,2}, u_{2,3}, u_{1,4}\}$. Тогда множество $\text{int}(C) \setminus \text{int}^*(C)$ содержит хотя бы одну вершину w , смежную с u_1 и u_3 . По лемме 4 $\deg(u_1) \leq v(G) - 5$. Тогда найдутся вершины $y, z \in \text{ext}(G)$, не смежные с u_1 . Поскольку $d(G) = 2$, вершины y, z смежны с u_2, u_3, u_4 , чтобы иметь общих соседей с вершинами $u_{1,2}, w, u_{1,4}$ соответственно, что противоречит планарности G .

СЛУЧАЙ 3.3: $|\text{int}^*(C)| = 2$. Считаем без ограничения общности, что $u_{1,2} \in \text{int}^*(C)$. Поскольку $\text{int}(G)$ содержит хотя бы одну вершину, смежную с u_1 и u_3 , каждая вершина $\text{ext}(C)$, не смежная с u_1 , должна быть смежна с вершинами u_2 и u_3 . Так как G не содержит разделяющих 3-циклов, каждое из множеств $\text{ext}(C)$ и $\text{int}(C)$ содержит не более одной вершины, не смежной с u_1 , откуда $\deg(u_1) \geq v(G) - 4$; противоречие по лемме 4.

СЛУЧАЙ 3.4: $|\text{int}^*(C)| \leq 1$. Без ограничения общности считаем, что $\text{int}^*(C)$ не содержит вершин, не смежных с u_1 . Тогда $\text{int}(C)$ тоже не содержит таких вершин, и можем применить рассуждения случая 2.1. Лемма 6 доказана.

Лемма 7. Если $\min\{\text{int}(G), \text{ext}(G)\} \geq 2$, то G не критический.

ДОКАЗАТЕЛЬСТВО. Пусть G критический. Рассмотрим 3 случая.

СЛУЧАЙ 1: $\min\{\text{int}(G), \text{ext}(G)\} \geq 3$. По лемме 6 без ограничения общности можем считать, что найдётся вершина $x \in \text{ext}(G)$, смежная с единственной вершиной C (например с u_1). Тогда каждая вершина множества $\text{int}(G)$ смежна с u_1 (и, следовательно, смежна также с u_3 по лемме 3). Поскольку G планарен и $\text{int}(G) \geq 3$, найдутся несмежные

вершины $y, z \in \text{int}(C)$. Рассмотрим u_1 -отделимое НКП $\mathcal{C}$ графа G . Обозначим через $Cl_y, Cl_z \in \mathcal{C}$ клики такие, что $y \in Cl_y$ и $z \in Cl_z$. Если $u_3 \notin Cl_y$, то множество $Cl_y \cup \{u_1\}$ является кликой, иначе множество $Cl_z \cup \{u_1\}$ является кликой; противоречие.

При рассмотрении оставшихся случаев можем считать без ограничения общности, что $\text{ext}(G) = \{x, y\}$ и G содержит вершину, смежную с единственной вершиной C (например с u_1). Заметим, что такая вершина принадлежит $\text{int}(C)$, поскольку в противном случае все вершины $\text{int}(C)$ смежны с u_1 , откуда $\deg(u_1) \geq v(G) - 4$, что невозможно по лемме 4. Тогда обе вершины x и y смежны с u_1 и, следовательно, с u_3 . Так как G не содержит разделяющих 3-циклов, то также можем считать, что $xu_4, yu_2 \notin E(G)$.

СЛУЧАЙ 2: $\min\{\deg(x), \deg(y)\} < 4$. Считаем, что $\deg(x) < 4$. Если $xy \in E(G)$ (соответственно $xu_2 \in E(G)$), то $N[x] \subseteq N[y]$ (соответственно $N[x] \subseteq N[u_2]$); противоречие по лемме 3. Если $xy \notin E(G)$ и $\deg(y) = 3$, то $yu_4 \in E(G)$ и $N[y] \subseteq N[u_4]$; снова противоречие по лемме 3. Если же $\deg(x) = \deg(y) = 2$, то рассмотрим некоторое u_1 -отделимое НКП $\mathcal{C}$ графа G . Найдётся клика $Cl \in \mathcal{C}$, содержащая одну из вершин x, y и не содержащая вершину u_3 . Тогда Cl состоит из единственной вершины, смежной с u_1 , и множество $Cl \cup \{u_1\}$ — тоже клика; противоречие.

СЛУЧАЙ 3: $\deg(x) = \deg(y) = 4$. Таким образом, $N(x) = \{u_1, u_2, u_3, y\}$ и $N(y) = \{u_1, u_3, u_4, x\}$.

СЛУЧАЙ 3.1. Найдётся вершина $w \in \text{int}(G)$, смежная с двумя противоположными вершинами C . Если $wu_1, wu_3 \in E(G)$ (соответственно $wu_2, wu_4 \in E(G)$), то $\max\{\text{int}(u_1u_2u_3w), \text{int}(u_3u_4u_1w)\} \geq 3$ (соответственно $\max\{\text{int}(u_2u_3u_4w), \text{int}(u_4u_1u_2w)\} \geq 3$) по лемме 1. Переименуем вершины графа и применим рассуждения случая 1 этой леммы.

При рассмотрении случаев 3.2 и 3.3 предполагаем, что каждая вершина множества $\text{int}(G) \setminus \text{int}^*(G)$ смежна ровно с одной вершиной C , тогда $\deg(u_2), \deg(u_4) \leq 5$.

СЛУЧАЙ 3.2: $\min\{\deg(u_2), \deg(u_4)\} \leq 4$. В силу симметрии можно считать, что $\deg(u_2) \leq 4$ и у вершин u_2 и u_3 отсутствует общий сосед в $\text{int}(C)$. Тогда либо $\deg(u_2) = 3$, либо $\deg(u_2) = 4$ и вершины u_1 и u_2 имеют общего соседа в $\text{int}(C)$. Рассмотрим u_1 -отделимое МСР графа G и его клики Cl_y и Cl_{u_2} , содержащие вершины y и u_2 соответственно (эти клики не могут совпадать, поскольку $yu_2 \notin E(G)$). Если $u_3 \in Cl_{u_2}$, то множество $Cl_y \cup \{u_1\}$ — клика, иначе множество $Cl_{u_2} \cup \{u_1\}$ — клика; противоречие.

СЛУЧАЙ 3.3: $\deg(u_2) = \deg(u_4) = 5$ и, таким образом, $|\text{int}^*(C)| = 4$. Если $u_{1,2}u_{1,4} \in E(G)$ или $u_{2,3}u_{3,4} \in E(G)$, то из отсутствия разделяющих

3-циклов в G следует, что $\max\{\deg(u_1), \deg(u_3)\} = v(G) - 4$; противоречие. Таким образом, $u_{1,2}u_{1,4}, u_{2,3}u_{3,4} \notin E(G)$, откуда $u_{1,2}u_{3,4}, u_{2,3}u_{1,4} \notin E(G)$ (если, например, $u_{1,2}u_{3,4} \in E(G)$, то $d(u_{2,3}, u_{1,4}) > 2$). Тогда вершины $u_{2,3}$ и $u_{1,4}$ имеют общего соседа $w \in \text{int}(C) \setminus \text{int}^*(C)$, который также единственно возможный общий сосед вершин $u_{1,2}$ и $u_{3,4}$. В силу симметрии можно считать, что $wu_1 \in E(G)$; тогда 3-циклы $u_1u_{1,2}w$ и $u_1u_{1,4}w$ не разделяющие. Поскольку $\deg(u_2) = \deg(u_4) = 5$, то 4-циклы $u_2u_{2,3}wu_{1,2}$ и $u_4u_{1,4}wu_{3,4}$ также не разделяющие. Если $|\text{int}(wu_{2,3}u_{3,4})| \geq 3$, то применим рассуждения случая 1 этой леммы. Если же $|\text{int}(wu_{2,3}u_{3,4})| \leq 2$, то рассмотрим 3 варианта.

СЛУЧАЙ 3.3.1: $\text{int}(u_3u_{2,3}wu_{3,4}) = \emptyset$. Тогда $v(G) = 11$; противоречие по лемме 1.

СЛУЧАЙ 3.3.2. Цикл $u_3u_{2,3}wu_{3,4}$ отделяет множество $\{x', y'\} \subset V(G)$. Из случая 2 этой леммы следует, что $\deg(x') = \deg(y') = 4$, при этом $x'u_3, y'u_3 \in E(G)$, откуда $x'w, y'w \in E(G)$. Легко проверить, что граф G 3-связный, при этом $v(G) = 13$; противоречие по лемме 1.

СЛУЧАЙ 3.3.3. Цикл $u_3u_{2,3}wu_{3,4}$ отделяет вершину $z \in V(G)$. Тогда $zu_3 \in E(G)$, откуда $zw \in E(G)$ по лемме 3. Если $\deg(z) = 4$, то граф G 3-связный; противоречие по лемме 1. Если $\deg(z) = 3$, то имеем либо $N[z] \subseteq N[u_{2,3}]$, либо $N[z] \subseteq N[u_{3,4}]$; противоречие по лемме 3.

Пусть $\deg(z) = 2$. Поскольку $\text{int}(u_1u_2u_{2,3}w) = u_{1,2}$, а по лемме 3 включение $N[u_{1,2}] \subseteq N[u_1]$ невозможно, то $u_{1,2}u_{2,3} \in E(G)$. Аналогично $u_{1,4}u_{3,4} \in E(G)$. Покажем, что $\theta(G) = 5$. Неравенство $\theta(G) \geq 5$ вытекает из того факта, что G не содержит разделяющих 3-циклов (и, следовательно, подграфов K_4), а вершина z не принадлежит ни одному 3-циклу. С другой стороны, множество $V(G)$ легко разбить на 5 клик (рис. 2), откуда и следует требуемое равенство. Поскольку множество $\{x, u_{2,3}, u_{3,4}, z\}$ независимое, то $4 = \alpha(G) \leq \gamma^\infty(G)$.

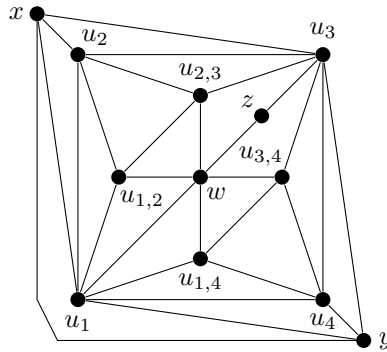


Рис. 2. Структура графа в случае 3.3.3

Предположим, что $\gamma^\infty(G) = 4$. Тогда любое независимое множество G мощности 4 будет вечным доминирующим и хотя бы один из охранников множества $\{x, u_{2,3}, u_{3,4}, z\}$ может отразить атаку на вершину u_3 таким образом, чтобы полученное множество тоже было вечным доминирующим. Легко проверить, что из множеств $\{u_3, u_{2,3}, u_{3,4}, z\}$, $\{x, u_3, u_{3,4}, z\}$, $\{x, u_{2,3}, u_3, z\}$ и $\{x, u_{2,3}, u_{3,4}, u_3\}$ доминирующее только последнее. Значит, оно способно отразить атаку на вершину u_2 таким образом, чтобы хотя бы одно из множеств $\{x, u_{2,3}, u_{3,4}, u_2\}$, $\{u_2, u_{2,3}, u_{3,4}, u_3\}$ и $\{x, u_2, u_{3,4}, u_3\}$ тоже было вечным доминирующим. Однако первое из этих трёх множеств не доминирующее, а второе (соответственно третье) перестаёт быть доминирующим после отражения атаки любым возможным способом на вершину y (соответственно вершину u_4). Полученное противоречие означает, что $\gamma^\infty(G) > 4$ и G не критический. Лемма 7 доказана.

Лемма 8. *Если C отделяет вершину x степени не более 3, то G не критический.*

ДОКАЗАТЕЛЬСТВО. Предположим, что G критический. Без ограничения общности можем считать, что $\text{ext}(G) = \{x\}$ и $xu_1 \in E(G)$. Тогда $xu_3 \in E(G)$ по лемме 3. Если $\deg(x) = 3$, то либо $N[x] \subseteq N[u_2]$, либо $N[x] \subseteq N[u_4]$; противоречие по лемме 3. Пусть $\deg(x) = 2$. По лемме 6 найдётся вершина $v_1 \in \text{int}(C)$, смежная с единственной вершиной C (например с u_1). Рассмотрим 2 случая.

СЛУЧАЙ 1. Найдётся вершина $w \in \text{int}(C)$, смежная как минимум с тремя вершинами C . Так как G не содержит разделяющих 3-циклов и $|\text{int}(C)| \geq 4$, то w смежна ровно с 3 вершинами C , и случай $wu_3 \notin E(G)$ невозможен. Если $wu_2 \notin E(G)$ (соответственно $wu_4 \notin E(G)$), то имеем $N[u_4] \subseteq N[w]$ (соответственно $N[u_2] \subseteq N[w]$); противоречие по лемме 3. Если же $wu_1 \notin E(G)$, то каждая вершина $\text{int}(C) \setminus \{w\}$ смежна с u_1 (чтобы иметь общего соседа с x), откуда $\deg(u_1) = v(G) - 3$; противоречие по лемме 4.

СЛУЧАЙ 2. Каждая вершина множества $\text{int}(C)$ смежна не более, чем с двумя вершинами цикла C . Поскольку каждая вершина $\text{int}(C)$, смежная с одной из вершин u_2 или u_4 , смежна также с одной из вершин u_1 или u_3 (чтобы иметь общего соседа с x), то $\deg(u_2), \deg(u_4) \leq 4$.

СЛУЧАЙ 2.1: $\min\{\deg(u_2), \deg(u_4)\} \leq 3$ (пусть $\deg(u_2) \leq 3$). Считаем без ограничения общности, что вершины u_2 и u_3 не имеют общего соседа в $\text{int}(C)$. Рассмотрим некоторое u_1 -отделимое НКП $\mathcal{C}$ графа G . Так как $\mathcal{C}$ минимально, то $\{x\} \notin \mathcal{C}$; тогда $\{u_3, x\} \in \mathcal{C}$. Рассмотрим клику $Cl \in \mathcal{C}$, содержащую вершину u_2 . Тогда либо $Cl = \{u_2\}$, либо $Cl = \{u_2, u_{1,2}\}$. Таким образом, множество $Cl \cup \{u_1\}$ тоже клика; противоречие.

СЛУЧАЙ 2.2: $\deg(u_2) = \deg(u_4) = 4$. Тогда $|\text{int}^*(C)| = 4$. Заметим, что $u_{1,2}u_{1,4} \notin E(G)$ (в противном случае цикл $u_1u_{1,2}u_{1,4}$ разделяющий). Кроме того, $u_{2,3}u_{3,4} \notin E(G)$, поскольку $\deg(u_1) \leq v(G) - 5$. Следовательно, пары вершин $u_{1,2}, u_{3,4}$ и $u_{2,3}, u_{1,4}$ не имеют общих соседей в $\text{int}^*(C)$, но тогда все вершины $\text{int}^*(C)$ смежны с некоторой вершиной $w \notin \text{int}^*(C)$. Так как $\deg(u_2) = \deg(u_4) = 4$, циклы $u_2u_{2,3}wu_{1,2}$ и $u_4u_{1,4}wu_{3,4}$ не разделяющие. Поскольку G не содержит разделяющих 3-циклов, то $wu_3 \in E(G)$. Тогда цикл $u_3u_{2,3}wu_{2,3}$ также не разделяющий. Из лемм 6 и 7 получаем, что $\text{int}(u_1u_{1,2}wu_{1,4}) = v_1$, откуда $v(G) = 11$; противоречие по лемме 1. Лемма 8 доказана.

Лемма 9. *Граф G не критический.*

ДОКАЗАТЕЛЬСТВО. Пусть G критический. В силу лемм 6–8 можем считать без ограничения общности, что $\text{ext}(C) = \{x\}$ и $\deg(x) = 4$. Заметим, что из леммы 3 следует неравенство $\deg(u_i) \geq 4$ при всех $i \in \overline{1,4}$.

СЛУЧАЙ 1. Найдётся вершина $w \in \text{int}(C)$, смежная с противоположными вершинами цикла C (например с u_1 и u_3). По леммам 6–8 имеем

$$|\text{int}(u_1u_2u_3w)| + |\text{int}(u_1wu_3u_4)| \leq 2,$$

откуда $v(G) \leq 8$; противоречие по лемме 1.

При рассмотрении оставшихся случаев предполагаем, что каждая вершина множества $\text{int}(C) \setminus \text{int}^*(C)$ смежна ровно с одной вершиной C .

СЛУЧАЙ 2: $|\text{int}^*(C)| = 4$. Заметим, что $u_{1,2}u_{3,4}, u_{2,3}u_{1,4} \notin E(G)$, иначе $v(G) < 12$ по леммам 6–8. Возможны следующие случаи.

СЛУЧАЙ 2.1. Вершины $u_{2,3}$ и $u_{1,4}$ имеют общего соседа в $\text{int}^*(C)$ (без ограничения общности считаем, что это $u_{3,4}$). Если $u_{1,2}u_{2,3} \in E(G)$ или $u_{1,2}u_{1,4} \in E(G)$, то из отсутствия в G разделяющих 3-циклов следует, что $\max\{\deg(u_1), \deg(u_2)\} \geq v(G) - 4$; противоречие. Иначе найдётся общий сосед $w \notin \text{int}^*(C)$ вершин $u_{1,2}$ и $u_{3,4}$. В силу симметрии можем положить, что $wu_1 \in E(G)$. Поскольку $v(G) \geq 12$, в графе G найдутся вершины $v_2, v'_2 \in \text{int}(u_2u_{1,2}vu_{3,4}u_{2,3}) \cap pn[u_2, C]$. Обе эти вершины смежны с $u_{3,4}$, чтобы иметь общего соседа с u_4 . Тогда

$$\max\{|\text{int}(u_2u_3u_{3,4}v_2)|, |\text{int}(u_2u_3u_{3,4}v'_2)|\} \geq 2;$$

противоречие по леммам 6–8.

СЛУЧАЙ 2.2. Вершины $u_{2,3}$ и $u_{1,4}$ имеют общего соседа $w \notin \text{int}^*(C)$. Предполагаем, что w также является общим соседом вершин $u_{1,2}$ и $u_{3,4}$ (иначе эти вершины имеют общего соседа в $\text{int}^*(C)$, а тогда применим рассуждения случая 2.1). В силу симметрии положим, что $u_1w \in E(G)$. По леммам 6–8 имеем

$$\max\{|\text{int}(u_2u_{2,3}wu_{1,2})|, |\text{int}(u_3u_{2,3}wu_{3,4})|, |\text{int}(u_4u_{2,3}wu_{3,4})|\} \leq 1.$$

Таким образом, $v(G) \leq 13$. При этом если $v(G) \in \{12, 13\}$, то каждый из циклов $u_2u_{1,2}wu_{2,3}$, $u_3u_{2,3}wu_{3,4}$ и $u_4u_{3,4}wu_{1,4}$ либо не разделяющий, либо отделяет единственную вершину степени 4 (иное невозможно по леммам 6–8). Тогда, как нетрудно проверить, граф G 3-связен (рис. 3); противоречие по лемме 1.

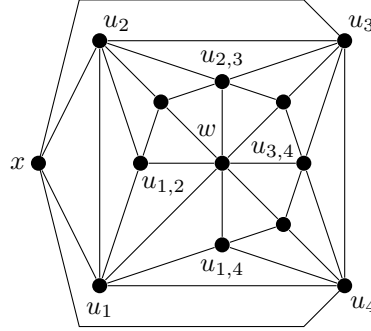


Рис. 3. Структура графа в случае 2.2 при $v(G) = 13$

СЛУЧАЙ 3: $|\text{int}^*(C)| = 3$. Без ограничения общности считаем, что $\text{int}^*(C) = \{u_{1,2}, u_{2,3}, u_{1,4}\}$.

СЛУЧАЙ 3.1. Вершина $u_{1,2}$ является общим соседом вершин $u_{2,3}$ и $u_{1,4}$. Из условия $v(G) \geq 12$ следует, что $|\text{int}(u_3u_4u_{1,4}u_{1,2}u_{2,3})| \geq 4$. При этом каждая вершина $\text{int}(u_3u_4u_{1,4}u_{1,2}u_{2,3})$ не принадлежит $\text{int}^*(G)$ и, таким образом, смежна ровно с одной из вершин u_3 и u_4 . По лемме 1 имеем $\deg(u_{1,2}) \leq v(G) - 5$, поэтому найдётся вершина $w \in \text{int}(u_3u_4u_{1,4}u_{1,2}u_{2,3})$, не смежная с $u_{1,2}$. В силу симметрии считаем, что $wu_3 \in E(G)$. Тогда из условия $d(u_1, w) \leq 2$ следует, что $wu_{1,4} \in E(G)$. По лемме 8 получаем, что $\text{int}(u_3u_4u_{1,4}w) = \emptyset$, так что $\deg(u_4) = 4$ и $\deg(u_3) = v(G) - 4$; противоречие по лемме 1.

СЛУЧАЙ 3.2. Вершина $u_{1,2}$ не является общим соседом вершин $u_{2,3}$ и $u_{1,4}$. Поскольку $v(G) \geq 12$, вершины $u_{2,3}$ и $u_{1,4}$ не смежны, так что они имеют общего соседа $w \in \text{int}(C) \setminus \text{int}^*(C)$, смежного ровно с одной из вершин C . Рассмотрим два подслучая.

СЛУЧАЙ 3.2.1: $wu_1 \in E(G)$ или $wu_2 \in E(G)$ (в силу симметрии считаем, что $wu_1 \in E(G)$). По леммам 6–8 имеем $\text{int}(u_1u_2u_{2,3}w) = \{u_{1,2}\}$ и $u_{1,2}w, u_{1,2}u_{2,3} \in E(G)$. Поскольку $v(G) \geq 12$, то $|\text{int}(u_3u_4u_{1,4}wu_{2,3})| \geq 3$. Далее возможны две конфигурации.

СЛУЧАЙ 3.2.1.1. Найдётся вершина $w' \in \text{int}(u_3u_4u_{1,4}wu_{2,3})$, смежная с вершиной u_4 , а также с $u_{2,3}$ (чтобы иметь общего соседа с u_2). Тогда $\text{int}(u_3u_4w'u_{2,3}) = \emptyset$ и найдётся вершина $w'' \in \text{int}(u_4u_{1,4}wu_{2,3}w')$, смежная

с вершинами u_4 и $u_{2,3}$. Таким образом, $w' \in \text{int}(u_3u_4w''u_{2,3})$ и $w'u_3 \notin E(G)$; противоречие по леммам 6–8.

СЛУЧАЙ 3.2.1.2. Все вершины $\text{int}(u_3u_4u_{1,4}wu_{2,3})$ смежны с вершиной u_3 . Если все эти вершины также смежны с w , то найдётся вершина $w' \in \text{int}(u_3u_4u_{1,4}wu_{2,3})$ такая, что $|\text{int}(wu_{2,3}u_3w')| > 1$; противоречие. Иначе найдётся не смежная с w вершина $w'' \in \text{int}(u_3u_4u_{1,4}wu_{2,3})$, которая смежна с вершинами $u_{2,3}$ и $u_{1,4}$ (чтобы иметь общих соседей с вершинами $u_{1,2}$ и u_1 соответственно). Тогда по леммам 6–8 получаем, что $\text{int}(u_3u_4u_{1,4}w'') = \emptyset$, откуда $|\text{int}(u_3u_4u_{1,4}wu_{2,3})| = 1$; противоречие.

СЛУЧАЙ 3.2.2: $wu_3 \in E(G)$ или $wu_4 \in E(G)$ (в силу симметрии считаем, что $wu_3 \in E(G)$). Тогда цикл $u_3u_4u_{1,4}w$ не разделяющий, откуда $\deg(u_3) = 5$ и $\deg(u_4) = 4$. Поскольку $\deg(u_1) \leq v(G) - 5$ по лемме 4, то найдётся вершина $v_2 \in pn[u_2, C]$. Из условия $d(v_2, u_4) \leq 2$ следует, что $v_2u_{1,4} \in E(G)$. Тогда по леммам 6–8 имеем $\text{int}(u_1u_2v_2u_{1,4}) = \{u_{1,2}\}$ и $u_{1,2}u_{1,4} \in E(G)$, откуда $\deg(u_1) = 5$ и $\deg(u_2) = v(G) - 4$; противоречие по лемме 1.

СЛУЧАЙ 4: $|\text{int}^*(C)| = 2$. В силу симметрии считаем, что $u_{1,2} \in E(G)$. Тогда либо $u_{1,4} \in E(G)$, либо $u_{3,4} \in E(G)$.

СЛУЧАЙ 4.1: $\text{int}^*(C) = \{u_{1,2}, u_{3,4}\}$. Если $u_{1,2}u_{3,4} \in E(G)$, то

$$|\text{int}(u_2u_3u_{3,4}u_{1,2})| + |\text{int}(u_1u_4u_{3,4}u_{1,2})| \leq 2,$$

откуда $v(G) \leq 9$; противоречие по лемме 1. Если же $u_{1,2}u_{3,4} \notin E(G)$, то некоторая вершина $w \in \text{int}(C) \setminus \text{int}^*(C)$ является их общим соседом. В силу симметрии считаем, что $u_1w \in E(G)$. Тогда по леммам 6–8 имеем $\text{int}(u_1wu_{3,4}u_4) = \emptyset$ и $\deg(u_4) = 4$. Если $\deg(u_2) = 4$, то $\deg(u_3) = v(G) - 4$; противоречие. Иначе в G найдётся вершина $v_2 \in pn[u_2, C]$. Так как $d(v_2, u_4) \leq 2$, то $v_2u_{3,4} \in E(G)$. Тогда по леммам 6–8 получаем, что $\text{int}(u_2u_3u_{3,4}v_2) = \emptyset$, откуда $\deg(u_3) = 4$ и $\deg(u_2) = v(G) - 4$; противоречие по лемме 1.

СЛУЧАЙ 4.2: $\text{int}^*(C) = \{u_{1,2}, u_{1,4}\}$. Так как $\deg(u_3) \geq 4$, найдётся $v_3 \in pn[u_3, C]$.

СЛУЧАЙ 4.2.1. Если $v_3u_{1,2}, v_3u_{1,4} \in E(G)$, имеем

$$|\text{int}(u_1u_{1,2}v_3u_{1,4})| + |\text{int}(u_2u_3v_3u_{1,2})| + |\text{int}(u_3u_4u_{1,4}v_3)| \leq 3,$$

откуда $v(G) < 12$; противоречие по лемме 1.

СЛУЧАЙ 4.2.2: $v_3u_{1,2} \in E(G)$ и $v_3u_{1,4} \notin E(G)$ (в силу симметрии случай $v_3u_{1,2} \notin E(G)$ и $v_3u_{1,4} \in E(G)$ рассматривается аналогично). В силу лемм 6–8 получаем, что $\text{int}(u_2u_3v_3u_{1,2}) = \emptyset$, откуда $\deg(u_2) = 4$. Если $\deg(u_4) \geq 5$, то найдётся вершина $v_4 \in pn[u_4, C]$, а из условия $d(u_2, v_4) \leq 2$ следует, что $u_{1,2}v_4 \in E(G)$. Тогда $\text{int}(u_1u_{1,2}v_4u_4) = \{u_{1,4}\}$ из лемм 6–8 и $u_{1,2}u_{1,4} \in E(G)$. Поскольку $\deg(u_{1,2}) \leq v(G) - 5$, найдётся вершина

$w \in \text{int}(u_3 u_4 v_4 u_{1,2} v_3)$, не смежная с $u_{1,2}$, но тогда, как нетрудно видеть, $\max\{d(w, u_1), d(w, u_2)\} > 2$; противоречие.

Пусть теперь $\deg(u_4) = 4$. Поскольку включение $N[u_{1,4}] \subseteq N[u_1]$ невозможно по лемме 3, то найдётся вершина $v'_3 \in pn[u_3, C]$, смежная с $u_{1,4}$ (из случая 4.2.1 вытекает, что v'_3 отлична от v_3 и не смежна с $u_{1,2}$). Из условия $\deg(u_1) \leq v(G) - 5$ следует, что в графе G имеется вершина $v''_3 \in pn[u_3, C]$, отличная от v_3 и v'_3 , при этом $v''_3 \in \text{int}(u_1 u_{1,2} v_3 u_3 v'_3 u_{1,4})$. Кроме того, $u_{1,2} v''_3, u_{1,4} v''_3 \notin E(G)$ (если, например, $u_{1,2} v''_3 \in E(G)$, то получим $v_3 \in \text{int}(u_2 u_3 v''_3 u_{1,2})$ и $u_2 v_3 \notin E(G)$, что противоречит леммам 6–8). Тогда u_1 и v''_3 имеют общего соседа $v_1 \in pn[u_1, C]$, при этом вершина v_1 является единственным возможным общим соседом пар вершин $v_3, u_{1,4}$ и $v'_3, u_{1,2}$ (рис. 4). По леммам 6–8 получаем, что $\text{int}(u_3 v_3 v_1 v'_3) = \{v''_3\}$, откуда $v(G) = 11$; противоречие по лемме 1.

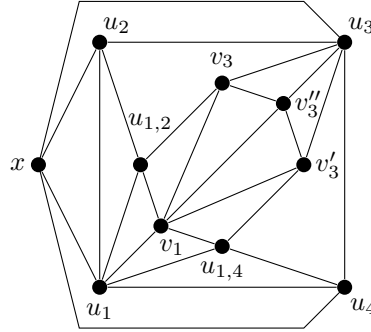


Рис. 4. Структура графа в случае 4.2.2

СЛУЧАЙ 4.2.3. Если $v_3 u_{1,2}, v_3 u_{1,4} \notin E(G)$, то имеется общий сосед $v_1 \in pn[u_1, C]$ вершин u_1 и v_3 . Поскольку включение $N[u_{1,2}] \subseteq N[u_1]$ невозможно по лемме 3, найдётся вершина $w \in N(u_{1,2}) \setminus N(u_1)$, причём $w \in \text{int}(u_1 u_2 u_3 v_3 v_1)$. Тогда u_3 является единственным возможным соседом вершин w и u_4 . Следовательно, можем применить рассуждения случая 4.2.2, рассматривая вершину w вместо v_3 .

СЛУЧАЙ 5: $|\text{int}^*(C)| = 1$. В силу симметрии можем считать, что $\text{int}^*(C) = \{u_{1,2}\}$. Поскольку $\deg(u_3), \deg(u_4) \geq 4$, найдутся вершины $v_3 \in pn[u_3, C]$ и $v_4 \in pn[u_4, C]$. Предположим, что $v_4 u_{1,2} \notin E(G)$ (случай $v_3 u_{1,2} \notin E(G)$ рассматривается аналогично). Тогда у вершин v_4 и u_2 найдётся общий сосед $v_2 \in pn[u_2, C]$, откуда $d(u_1, v_3) > 2$; противоречие.

Пусть теперь $v_3 u_{1,2}, v_4 u_{1,2} \in E(G)$. Поскольку $\deg(u_{1,2}) \leq v(G) - 5$, найдётся вершина $w' \in \text{int}(C)$, не смежная с $u_{1,2}$. По леммам 6–8 имеем $\text{int}(u_2 u_3 v_3 u_{1,2}) = \text{int}(u_1 u_4 v_4 u_{1,2}) = \emptyset$, так что $w' \in \text{int}(v_3 u_3 u_4 v_4 u_{1,2})$. Тогда либо $w' u_3 \in E(G)$ и $d(w', u_1) > 2$, либо $w' u_4 \in E(G)$ и $d(w', u_2) > 2$; противоречие.

СЛУЧАЙ 6: $\text{int}^*(C) = \emptyset$. Напомним, что $\deg(u_i) \geq 4$, а значит, найдётся вершина $v_i \in pn[u_i, C]$ для любого $i \in \overline{1, 4}$. Поскольку $d(G) = 2$, вершина v_1 смежна с некоторой вершиной $v'_3 \in pn[u_3, C]$ (которая может совпадать с v_3). Тогда $d(u_2, v_4) > 2$; противоречие. Лемма 9 доказана.

5. Разделяющие 5-циклы

Лемма 10. Если в графе $G \in \mathcal{PL}_2$ содержится разделяющий 5-цикл $C = u_1u_2u_3u_4u_5$, то G не критический.

ДОКАЗАТЕЛЬСТВО. Пусть G критический. Из лемм 5–9 следует, что G не содержит разделяющих $\{3, 4\}$ -циклов, поэтому C является порождённым 5-циклом. Назовём вершину множества $\text{int}(C) \cup \text{ext}(C)$ *хордальной*, если она смежна с парой несмежных вершин C . Обозначим через c_{int} и c_{ext} число хордальных вершин в множествах $\text{int}(C)$ и $\text{ext}(C)$ соответственно. Без ограничения общности считаем, что $c_{\text{int}} \geq c_{\text{ext}}$, а если $c_{\text{int}} = c_{\text{ext}}$, то $|\text{int}(C)| \geq 4$, так как $v(G) \geq 12$. Рассмотрим 6 случаев в зависимости от значений величин c_{int} и c_{ext} .

СЛУЧАЙ 1: $c_{\text{int}} \geq 3$. Рассмотрим две хордальные вершины x и y подграфа $\text{int}(C)$. В силу симметрии можно считать, что $u_1x, u_3x \in E(G)$. Если $u_1y, u_3y \in E(G)$, то хотя бы один из 4-циклов $u_1u_2u_3x$ и $u_1u_2u_3y$ разделяющий; противоречие. Иначе, поскольку G планарен, вершина y смежна ровно с одной из вершин u_1 и u_3 . Если, например, $u_3y \in E(G)$, то $u_5y \in E(G)$. Так как $c_{\text{int}} \geq 3$, то $\text{int}(C)$ содержит хордальную вершину z , отличную от x и y , а это невозможно, поскольку G планарен и не содержит разделяющих 4-циклов; противоречие.

При рассмотрении случаев 2–4 можем считать без ограничения общности, что найдутся хордальные вершины $x, y \in \text{int}(C)$ такие, что $xu_1, xu_3, yu_3, yu_5 \in E(G)$. Обозначим через C' цикл $u_1xu_3yu_5$.

СЛУЧАЙ 2: $c_{\text{int}} = c_{\text{ext}} = 2$. Так как циклы $u_1u_2u_3x$ и $u_3u_4u_5y$ не разделяющие, то $|\text{int}(C')| \geq 2$. Обозначим через a и b хордальные вершины из $\text{ext}(C)$. Поскольку G планарен и не содержит разделяющих $\{3, 4\}$ -циклов, вершины a и b имеют единственного общего соседа u_m среди вершин C . Если $m \in \{1, 3, 5\}$, то, как нетрудно проверить, хотя бы одна из вершин a, b образует разделяющий 4-цикл с одной из вершин x, y ; противоречие. Пусть теперь $m \in \{2, 4\}$. В силу симметрии можем считать, что $m = 2$ и $u_5a, u_4b \in E(G)$. Поскольку $u_5b \notin E(G)$, каждая вершина из $\text{int}(C')$ смежна с u_3 (чтобы иметь общего соседа с b) и с одной из вершин u_1 и u_5 (чтобы иметь общего соседа с a). Тогда каждая вершина $\text{int}(C')$ хордальная, откуда $c_{\text{int}} \geq 4$; противоречие.

СЛУЧАЙ 3: $c_{\text{int}} = 2, c_{\text{ext}} = 1$. Обозначим через z единственную хордальную вершину $\text{ext}(C)$.

СЛУЧАЙ 3.1: $u_3z \in E(G)$. Так как G не содержит разделяющих 4-циклов, то $zu_1, zu_5 \notin E(G)$. Поскольку z хордальная, то $zu_2, zu_4 \in E(G)$. Если множество $\text{int}(C')$ непусто, то каждая его вершина смежна с u_3 (чтобы иметь общего соседа с z) и при этом не смежна с u_1 и u_5 . Тогда для каждой вершины $w \in \text{int}(C')$ верно включение $N[w] \subseteq N[u_3]$; противоречие.

Таким образом, $\text{int}(C) = \{x, y\}$ и $|\text{ext}(C) \setminus \{z\}| \geq 4$. Поскольку 3-циклы u_2u_3z и u_3u_4z не разделяющие, вершины множества $\text{ext}(C) \setminus \{z\}$ не смежны с u_3 . При этом все они не хордальные и должны иметь общего соседа с каждой из вершин x и y . Следовательно, все вершины $\text{ext}(C) \setminus \{z\}$ смежны с u_1 и u_5 , но тогда G содержит разделяющий 3-цикл; противоречие.

СЛУЧАЙ 3.2: $u_3z \notin E(G)$ и $u_2z, u_4z \in E(G)$, при этом хотя бы одно из рёбер u_1z и u_5z принадлежит $E(G)$ (в силу симметрии считаем, что $u_1z \in E(G)$). Так как ни один из циклов u_1u_2z , $u_1u_5u_4z$, $u_2u_3u_4z$ не разделяющий, то $\text{ext}(C) = \{z\}$, так что $|\text{int}(C')| \geq 4$. Если $u_5z \notin E(G)$, то все вершины $\text{int}(C')$ смежны с u_1 , поэтому $\deg(u_1) = v(G) - 4$; противоречие. Если же $u_5z \in E(G)$, то считаем в силу симметрии, что $\text{int}(C')$ содержит хотя бы две вершины, смежные с u_1 . Каждая из этих вершин смежна либо с y , либо с u_5 , но тогда G содержит разделяющий $\{3, 4\}$ -цикл; противоречие.

СЛУЧАЙ 3.3: $u_2z, u_4z \in E(G)$ и $u_1z, u_3z, u_5z \notin E(G)$. Тогда множество $\text{int}(C')$ пусто, поскольку никакая его вершина не может иметь общего соседа с z , откуда $|\text{ext}(C)| \geq 5$. Поскольку цикл $u_2u_3u_4z$ не является разделяющим и $xu_5, yu_1 \notin E(G)$, каждая вершина подграфа $\text{ext}(C) \setminus \{z\}$ смежна хотя бы с одной из вершин u_1, u_2 и хотя бы с одной из вершин u_4, u_5 (чтобы иметь общих соседей с вершинами x и y). Так как вершины множества $\text{ext}(C) \setminus \{z\}$ не хордальные, каждая из них смежна с вершинами u_1 и u_5 . Значит, граф G содержит разделяющий 3-цикл; противоречие.

СЛУЧАЙ 3.4: $u_3z \notin E(G)$ и хотя бы одно из рёбер u_4z и u_2z не входит в $E(G)$ (в силу симметрии считаем, что $u_4z \notin E(G)$). Тогда $u_2z, u_5z \in E(G)$. Если $\text{ext}(C) = \{z\}$, то $|\text{int}(C')| \geq 4$ и, применяя рассуждения случая 3.2, нетрудно проверить, что G содержит разделяющий $\{3, 4\}$ -цикл, что невозможно. В противном случае найдётся вершина $a \in \text{ext}(C) \setminus \{z\}$. Поскольку цикл $u_1u_2zu_5$ не разделяющий, то $au_1 \notin E(G)$. Тогда $au_3 \in E(G)$, так как вершина a не хордальная и должна иметь общего соседа с обеими вершинами x и y . Аналогично каждая вершина множества $\text{ext}(C) \setminus \{z, a\}$ не хордальная и смежна с u_3 . Из условия $\deg(u_3) < v(G) - 4$ следует, что найдётся вершина $b \in \text{int}(C')$, не смежная с u_3 . Так как $au_1, au_5 \notin E(G)$, то $d(a, b) > 2$; противоречие.

СЛУЧАЙ 4: $c_{\text{int}} = 2$, $c_{\text{ext}} = 0$. Если ни одна вершина множества $\text{ext}(C)$ не смежна с u_3 , то все они смежны с u_1 и u_5 (чтобы иметь общего соседа с обеими вершинами x и y). Поскольку граф G не содержит разделяющих 3-циклов, $\text{ext}(C)$ состоит из единственной вершины z и $N[z] \subseteq N[u_1]$; противоречие по лемме 4. Если же найдётся вершина $w \in \text{ext}(C)$, смежная с u_3 , то все вершины $\text{int}(C')$ смежны с u_3 , чтобы иметь общего соседа с w . По лемме 4 $\text{ext}(C)$ содержит хотя бы две вершины, не смежные с u_3 . Эти вершины смежны с вершинами u_1 и u_5 , но тогда G содержит разделяющий 3-цикл; противоречие.

СЛУЧАЙ 5: $c_{\text{int}} = 1$. Обозначим через x единственную хордальную вершину $\text{int}(C)$.

СЛУЧАЙ 5.1: $\text{int}(C) = \{x\}$. Так как $|\text{ext}(C)| \geq |\text{int}(C)|$, предполагаем, что $c_{\text{ext}} = 0$. Без ограничения общности можем считать, что найдётся вершина $v_1 \in \text{ext}(C)$, для которой $u_1v_1 \in E(G)$ и $v_1u_3, v_1u_4, v_1u_5 \notin E(G)$. Тогда найдётся вершина $v_4 \in \text{ext}(G)$ такая, что $v_4u_4, v_4v_1 \in E(G)$. Поскольку вершина v_4 не хордальная, то $u_1v_4, u_2v_4 \notin E(G)$.

СЛУЧАЙ 5.1.1: $v_1u_2, v_4u_3, v_4u_5 \notin E(G)$. Так как $d(v_1, u_3) \leq 2$, найдётся вершина $v_3 \in \text{ext}(C)$ такая, что $v_1v_3, u_3v_3 \in E(G)$. Если $u_2v_3 \notin E(G)$, то $d(v_4, u_2) > 2$. Если же $u_2v_3 \in E(G)$, то $v_3u_4 \notin E(G)$ и $d(v_3, u_5) > 2$; противоречие.

СЛУЧАЙ 5.1.2: $v_1u_2 \in E(G)$ и $v_4u_3, v_4u_5 \notin E(G)$. Нетрудно видеть, что любая вершина множества $\text{ext}(C) \setminus \{v_1, v_4\}$ смежна с u_4 , иначе она не имеет общего соседа хотя бы с одной из вершин u_3 и u_5 . При этом вершина u_4 является единственно возможным общим соседом v_4 и x . Таким образом, $\deg(u_4) = v(G) - 4$; противоречие по лемме 4.

СЛУЧАЙ 5.1.3: $v_1u_2 \in E(G)$ и либо $v_4u_3 \in E(G)$, либо $v_4u_5 \in E(G)$ (в силу симметрии достаточно рассмотреть случай $v_4u_3 \in E(G)$). Положим $C'' = u_1v_1v_4u_4u_5$ и рассмотрим $w \in \text{int}(C'')$. Если $wu_5 \in E(G)$, то w смежна хотя бы с одной из вершин u_1, v_1 (чтобы иметь общего соседа с вершиной u_2), и хотя бы с одной из вершин u_4, v_4 (чтобы иметь общего соседа с u_3). Поскольку G планарен, не содержит разделяющих $\{3, 4\}$ -циклов, а вершина w не хордальная, то $\text{int}(C'') = \{w\}$ и $v(G) < 12$; противоречие по лемме 1.

Таким образом, каждая вершина $\text{int}(C'')$ смежна ровно с одной из вершин u_1 и u_4 и не смежна с u_5 . В силу симметрии можно считать, что $|pn[u_1, C]| \geq |pn[u_4, C]|$. Тогда найдутся вершины $w_1, w_2 \in pn[u_1, C]$, которые должны иметь общего соседа с u_3 . Тем самым $w_1v_4, w_2v_4 \in E(G)$ и хотя бы один из 4-циклов $u_1v_1v_4w_1$ и $u_1v_1v_4w_2$ разделяющий; противоречие.

При рассмотрении случаев 5.2 и 5.3 можем считать без ограничения общности, что хордальная вершина $x \in \text{int}(C)$ смежна с u_1 и u_3 . Обозначим через C''' цикл $u_1xu_3u_4u_5$. Так как $\text{int}(u_1u_2u_3x) = \emptyset$, то $\text{int}(C''') \neq \emptyset$ (иное уже разобрано в случае 5.1).

СЛУЧАЙ 5.2. Каждая вершина множества $\text{int}(C''')$ смежна с x .

СЛУЧАЙ 5.2.1: $|\text{int}(C''')| \geq 3$. Либо в $\text{int}(C''')$ найдутся хотя бы две вершины, смежные хотя бы с одной из вершин множества $\{u_1, u_5\}$, либо найдутся хотя бы две вершины, смежные хотя бы с одной из вершин множества $\{u_3, u_4\}$. Следовательно, G содержит разделяющий $\{3, 4\}$ -цикл; противоречие.

При рассмотрении случаев 5.2.2 и 5.2.3 предполагаем, что $|\text{int}(C)| \leq 3$ и тем самым $c_{\text{ext}} = 0$.

СЛУЧАЙ 5.2.2: $\text{int}(C''') = \{y, z\}$. Поскольку G не содержит разделяющих $\{3, 4\}$ -циклов, считаем без ограничения общности, что $yu_1, yu_5, zu_3, zu_4 \notin E(G)$. Рассмотрим некоторую вершину $w \in \text{ext}(C)$, которая по предположению не хордальная. Так как $xu_4, xu_5 \notin E(G)$, хотя бы с одной из вершин x, y, z вершина w не имеет общего соседа; противоречие.

СЛУЧАЙ 5.2.3: $\text{int}(C''') = \{y\}$. Если $yu_1 \in E(G)$ (соответственно $yu_3 \in E(G)$), то $N[y] \subseteq N[u_1]$ (соответственно $N[y] \subseteq N[u_3]$), так как вершина y не хордальная; противоречие. Если же $yu_1, yu_3 \notin E(G)$, то каждая вершина из $\text{ext}(C)$ смежна хотя бы с одной из вершин u_4 и u_5 , а также хотя бы с одной из вершин u_1, u_2, u_3 (чтобы иметь общего соседа с x и y). Поскольку G не содержит разделяющих 3-циклов, а $\text{ext}(C)$ не содержит хордальных вершин, то $|\text{ext}(C)| \leq 2$ и $v(G) < 12$; противоречие.

СЛУЧАЙ 5.3. Найдётся вершина $y \in \text{int}(C''')$, не смежная с x . Поскольку $d(u_2, y) \leq 2$, вершина y смежна ровно с одной из вершин u_1 и u_3 (в силу симметрии считаем, что $yu_1 \in E(G)$). Так как y не хордальная, имеем $yu_4 \notin E(G)$ и найдётся вершина $z \in \text{int}(C''')$ такая, что $yz, u_3z \in E(G)$. Поскольку вершины y и z не хордальные, получаем $yu_4, zu_5 \notin E(G)$. Тогда все вершины множества $\text{ext}(C)$ смежны с двумя несмежными вершинами C (так как они должны иметь общих соседей с каждой из вершин x, y, z). Таким образом, $\text{ext}(C)$ состоит из единственной хордальной вершины (обозначим её через w) и $|\text{int}(C''')| \geq 5$. Рассмотрим 4 варианта в зависимости от того, смежна ли вершина w с вершинами u_1 и u_3 .

СЛУЧАЙ 5.3.1. Если $wu_1, wu_3 \in E(G)$, то граф G содержит разделяющий 4-цикл wu_1xu_3 ; противоречие.

СЛУЧАЙ 5.3.2. Если $wu_3 \in E(G)$ и $wu_1 \notin E(G)$, то каждая вершина $\text{int}(u_1xu_3zy)$ смежна с u_3 (чтобы иметь общего соседа с w), а u_5 является общим соседом вершин y и w . Поскольку $\deg(u_3) \leq v(G) - 5$, найдётся

вершина $a \in \text{int}(u_3u_4u_5yz)$ такая, что $au_3 \notin E(G)$, но тогда $d(a, u_2) > 2$; противоречие.

СЛУЧАЙ 5.3.3. Если $wu_1 \in E(G)$ и $wu_3 \notin E(G)$, то каждая вершина $\text{int}(u_1xu_3zy)$ смежна с u_1 , а u_4 является общим соседом вершин z и w . Поскольку $\deg(u_1) \leq v(G) - 5$, найдётся вершина $a' \in \text{int}(u_1u_5u_4zy)$ такая, что $a'u_1 \notin E(G)$, но тогда $d(a', u_2) > 2$; противоречие.

СЛУЧАЙ 5.3.4. Если $wu_1, wu_3 \notin E(G)$, то вершины u_2, u_4, u_5 являются единственными возможными общими соседями вершины w в парах с вершинами x, y, z соответственно. Поскольку $v(G) \geq 12$ и G не содержит разделяющих $\{3, 4\}$ -циклов, множество $\text{int}(u_1xu_3zy)$ непусто, но никакая его вершина не может иметь общего соседа с вершиной w ; противоречие.

СЛУЧАЙ 6: $c_{\text{int}} = c_{\text{ext}} = 0$. В силу симметрии считаем, что $|\text{int}(C)| \geq 4$ и найдётся вершина $v_1 \in \text{int}(C) \cap N(u_1)$ такая, что $v_1u_5 \notin E(G)$.

СЛУЧАЙ 6.1. Если $v_1u_2 \notin E(G)$, то каждая вершина $w \in \text{ext}(C)$ смежна с u_1 и не является хордальной, откуда $N[w] \subseteq N[u_1]$; противоречие по лемме 3.

СЛУЧАЙ 6.2. Если $v_1u_2 \in E(G)$, рассмотрим общего соседа v_4 вершин v_1 и u_4 . Так как каждая вершина $\text{ext}(C)$ должна иметь общего соседа с v_1 и v_4 , то либо $v_4u_3 \in E(G)$, либо $v_4u_5 \in E(G)$ (в силу симметрии считаем, что $v_4u_3 \in E(G)$). Тогда каждая вершина $\text{ext}(C)$ смежна с вершинами u_2 и u_3 , а поскольку G не содержит разделяющих 3-циклов, то такая вершина единственна; обозначим её через w . Тем самым $N[w] \subseteq N[u_2]$; противоречие по лемме 3. Лемма 10 доказана.

6. Основной результат

Теорема 1. *Каждый планарный граф диаметра 2 принадлежит $\mathcal{MD}$.*

ДОКАЗАТЕЛЬСТВО. Предположим, что это не так. Тогда найдётся критический граф G , который по леммам 5–10 не содержит разделяющих $\{3, 4, 5\}$ -циклов. Рассмотрим вершину $u \in V(G)$ максимальной степени и положим $N(u) = \{u_1, \dots, u_s\}$ (поскольку $v(G) \geq 12$ и $d(G) = 2$, то $s \geq 4$). Предполагаем, что вершины $u_1, \dots, u_s$ расположены на плоскости по часовой стрелке (тогда из отсутствия 3-циклов в G следует, что если $u_iu_j \in E(G)$, то $|i - j| \in \{1, s - 1\}$). Выберем произвольную вершину $w \in N(u_1) \setminus N(u)$ (она существует по лемме 3). Рассмотрим три случая.

СЛУЧАЙ 1. Вершины w и u_4 не смежны и имеют общего соседа u_k . Если $k = 2$ (соответственно $k = 3$), то цикл wu_2u_4 (соответственно wu_1wu_3) разделяющий; противоречие. Если же $k > 4$, то $k = s = 5$ (в противном случае цикл $wu_1wu_ku_4$ разделяющий), при этом $u_5u_2, u_5u_3 \notin E(G)$.

Если w смежна с u_2 (соответственно с u_3), то цикл uu_2wu_5 (соответственно uu_3wu_5) разделяющий. Иначе вершины u_3 и w имеют общего соседа $x \notin N[u]$, а цикл uu_1wxu_3 разделяющий; противоречие.

СЛУЧАЙ 2. Несмежные вершины w, u_4 имеют общего соседа $x \notin N(u)$. Поскольку цикл uu_1wxu_4 не разделяющий, то $s = 4$. Легко проверить, что если хотя бы одна из вершин u_2, u_3 смежна с одной из вершин w, x , то G содержит $\{4, 5\}$ -разделяющий цикл. В противном случае у вершин u_3 и w найдётся общий сосед $y \notin N[u] \cup \{x\}$, и цикл uu_1wyu_3 будет разделяющим; противоречие.

СЛУЧАЙ 3. Вершины w и u_4 смежны. Можем считать, что все вершины множества $N(u_1) \setminus N(u)$ смежны с u_4 (если это не так, то переименуем вершины и применим рассуждения одного из предыдущих случаев). Так как G не содержит разделяющих 4-циклов, то $s = 4$ и в силу симметрии можно считать, что все вершины множества $N(u_4) \setminus N(u)$ смежны с u_1 . Поскольку цикл uu_1wu_4 не разделяющий, то

$$N(u_1) \setminus N(u) = N(u_4) \setminus N(u) = \{w\}.$$

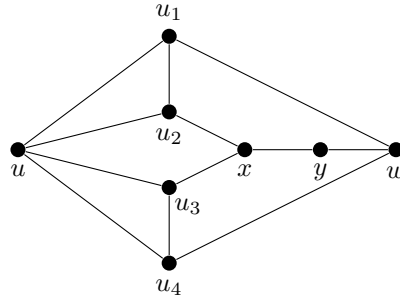


Рис. 5. Структура графа в случае 3

Если $\min\{\deg(u_1), \deg(u_4)\} = 2$, то каждая вершина из $V(G) \setminus N[u]$ смежна с w (чтобы иметь общего соседа с u_1 и u_4). Отсюда следует, что $\deg(w) \geq v(G) - 4$; противоречие по лемме 4. Тем самым $\deg(u_1) = \deg(u_4) = 3$ и $u_1u_2, u_3u_4 \in E(G)$, при этом $u_2w, u_3w \notin E(G)$, иначе G содержит разделяющий 4-цикл. Поскольку $\deg(w) \leq v(G) - 5$, найдётся вершина $x \notin N[u]$ такая, что $xw \notin E(G)$. Тогда u_2 (соответственно u_3) является единственно возможным соседом вершин x и u_1 (соответственно x и u_4). Кроме того, вершины x и w имеют общего соседа $y \in V(G) \setminus N[u]$. Поскольку циклы $uu_1wu_4, uu_2xu_3, u_1u_2xuw$ и u_4u_3xuw не разделяющие (рис. 5), то $v(G) = 8$; противоречие по лемме 1. Теорема 1 доказана.

Финансирование работы

Исследование выполнено в рамках Программы фундаментальных исследований Национального исследовательского университета «Высшая школа экономики». Дополнительных грантов на проведение или руководство этим исследованием получено не было.

Конфликт интересов

Автор заявляет, что у него нет конфликта интересов.

Литература

1. Burger A. P., Cockayne E. J., Grundlingh W. R., Mynhardt C. M., van Vuuren J., Winterbach W. Infinite order domination in graphs // J. Comb. Math. Comb. Comput. 2004. V. 50. P. 179–194.
2. Goddard W., Hedetniemi S. M., Hedetniemi S. T. Eternal security in graphs // J. Comb. Math. Comb. Comput. 2005. V. 52. P. 169–180.
3. Klostermeyer W. F., MacGillivray G. Eternal security in graphs of fixed independence number // J. Comb. Math. Comb. Comput. 2007. V. 63. P. 97–101.
4. Goldwasser J. L., Klostermeyer W. F. Tight bounds for eternal dominating sets in graphs // Discrete Math. 2008. V. 308, No. 12. P. 2589–2593.
5. Driscoll K., Klostermeyer W. F., Krop E., Magnant C., Taylor P. On eternal domination and Vizing-type inequalities // AKCE Int. J. Graphs Comb. 2020. V. 17, No. 1. P. 1–5.
6. Klostermeyer W., Mynhardt C. Protecting a graph with mobile guards // Appl. Anal. Discrete Math. 2016. V. 10. P. 1–29.
7. Topics in domination in graphs. Cham: Springer, 2020. 546 p. (Dev. Math.; V. 64).
8. Anderson M., Barrientos C., Brigham R., Carrington J., Vitray R., Yellen J. Maximum-demand graphs for eternal security // J. Comb. Math. Comb. Comput. 2007. V. 61. P. 111–128.
9. Krim-Yee A., Seamone B., Virgile V. Eternal domination on prisms of graphs // Discrete Appl. Math. 2019. V. 283. P. 734–736.
10. MacGillivray G., Mynhardt C. M., Virgile V. Eternal domination and clique covering // Electron. J. Graph Theory Appl. 2022. V. 10, No. 2. P. 603–624.

Талецкий Дмитрий Сергеевич

Статья поступила

26 июня 2024 г.

После доработки —

14 августа 2024 г.

Принята к публикации

22 сентября 2024 г.

ON THE ETERNAL DOMINATION NUMBER
OF PLANAR GRAPHS WITH DIAMETER 2

D. S. Taletskii

National Research University “Higher School of Economics”,
25/12 Bolshaya Pechyorskaya Street, 603155 Nizhny Novgorod, Russia
E-mail: dmitailmail@gmail.com

Abstract. An *eternal dominating set* in a graph is a dominating set D on which mobile guards are initially located (at most one guard is allowed on any vertex). For any infinite sequence of attacks occurring sequentially at vertices, the set D can be modified by moving the guard from an adjacent vertex to the attacked vertex, provided the attacked vertex has no guard on it at the time it is attacked. The configuration of guards after each attack must induce a dominating set. The *eternal domination number* of a graph is the cardinality of its minimum eternal dominating set. We prove that the eternal domination number of any planar graph of diameter 2 is equal to its clique covering number. Illustr. 5, bibliogr. 10.

Keywords: dominating set, eternal domination set, eternal domination number, planar graph.

References

1. A. P. Burger, E. J. Cockayne, W. R. Grundlingh, C. M. Mynhardt, J. van Vuuren, and W. Winterbach, Infinite order domination in graphs, *J. Comb. Math. Comb. Comput.* **50**, 179–194 (2004).
2. W. Goddard, S. M. Hedetniemi, and S. T. Hedetniemi, Eternal security in graphs, *J. Comb. Math. Comb. Comput.* **52**, 169–180 (2005).
3. W. F. Klostermeyer and G. MacGillivray, Eternal security in graphs of fixed independence number, *J. Comb. Math. Comb. Comput.* **63**, 97–101 (2007).
4. J. L. Goldwasser and W. F. Klostermeyer, Tight bounds for eternal dominating sets in graphs, *Discrete Math.* **308** (12), 2589–2593 (2008).

5. **K. Driscoll, W. F. Klostermeyer, E. Krop, C. Magnant, and P. Taylor**, On eternal domination and Vizing-type inequalities, *AKCE Int. J. Graphs Comb.* **17** (1), 1–5 (2020).
6. **W. Klostermeyer and C. Mynhardt**, Protecting a graph with mobile guards, *Appl. Anal. Discrete Math.* **10**, 1–29 (2016).
7. *Topics in Domination in Graphs* (Springer, Cham, 2020) (Dev. Math., Vol. 64).
8. **M. Anderson, C. Barrientos, R. Brigham, J. Carrington, R. Vitray, and J. Yellen**, Maximum-demand graphs for eternal security, *J. Comb. Math. Comb. Comput.* **61**, 111–128 (2007).
9. **A. Krim-Yee, B. Seamone, and V. Virgile**, Eternal domination on prisms of graphs, *Discrete Appl. Math.* **283**, 734–736 (2019).
10. **G. MacGillivray, C. M. Mynhardt, and V. Virgile**, Eternal domination and clique covering, *Electron. J. Graph Theory Appl.* **10** (2), 603–624 (2022).

Dmitry S. Taletskii

Received June 26, 2024

Revised August 14, 2024

Accepted September 22, 2024

МЕТОД ДЕКОМПОЗИЦИИ ДЛЯ ЗАДАЧИ РАЗМЕЩЕНИЯ АНТЕНН НА СТАДИОНЕ

А. Д. Юськов

Новосибирский гос. университет,
ул. Пирогова, 2, 630090 Новосибирск, Россия
E-mail: a.yuskov@g.nsu.ru

Аннотация. Рассматривается задача размещения антенн на стадионе. Стадион разделён на секторы, и на каждый сектор назначается некоторое число антенн связи. Пользователи должны получать сигнал определённого качества от антенн, назначенных на их сектор. Задача состоит в том, чтобы выбрать места размещения антенн, их типы, углы направления и назначения на сектора так, чтобы максимизировать три критерия качества решения: среднее отношение сигнал/помехи (signal to interference ratio, SIR), количество клиентов с хорошим качеством сигнала и согласованность назначения. Вычисление качества сигнала производится при помощи имитационной модели. В работе представлена трехэтапная эвристическая схема для решения задачи. Она использует конструктивную эвристику, процедуру локального улучшения и эвристику на основе декомпозиции задачи по секторам с использованием модели целочисленного линейного программирования. Проведены численные эксперименты на тестовых примерах с 94 антеннами 7 типов, 19 секторами и 4426 пользователями. На данных примерах удалось за 2 ч улучшить предоставленные базовые решения и получить решения, сравнимые с запуском метаэвристического пакета на 24 ч. Табл. 2, ил. 3, библиогр. 17.

Ключевые слова: оптимизация «чёрного ящика», метаэвристика, беспроводная сеть, качество сигнала, SINR.

Введение

В век информационных технологий качество услуг беспроводной связи должно отвечать высоким стандартам, поэтому существует необходимость в разработке методов оптимизации сетей связи. Услуги беспроводной сети в местах массового скопления людей, таких как стадионы [1] или лекционные аудитории [2], предоставляются с помощью специально

организованной системы антенн. Чаще всего моделирование распространения сигнала производится при помощи эмпирических формул, таких как [3] или [4].

При оптимизации размещения антенн в литературе обычно встречается два полярных подхода: один из них предполагает полное знание законов распространения сигнала, а другой — использование общих метаэвристических методов, которые оптимизируют целевую функцию как «чёрный ящик». В первом случае, принимая во внимание некоторые упрощения, иногда удаётся построить модель, которая позволяет найти оптимальное решение [5]. Однако у этого подхода есть несколько недостатков. Его применимость зависит от конкретной конфигурации устройств и геометрии местности и требует полного знания данных характеристик. Он также требует некоторого упрощения формул и, как следствие, результаты могут расходиться с действительностью. Кроме того, может потребоваться большое количество вычислительного времени для полного решения.

Второй подход не требует точного знания формул распространения сигнала: функции могут быть выражены при помощи компьютерной программы, в том числе использующей экспериментальные данные и машинное обучение [6]. Кроме того, при смене внутренней модели, например вследствие получения новых данных, он не требует изменения алгоритма. Однако, так как алгоритм оптимизации не использует никаких свойств задачи, его эффективность на больших размерностях может не удовлетворять потребности. Популярным классом алгоритмов в этой сфере являются генетические алгоритмы [7]. Их достоинством в том числе является то, что они способны решать многокритериальные задачи [8, 9].

В этой работе применяется смешанный подход, при котором часть данных задачи считается известной и используется алгоритмами, а другая часть получается от имитационной модели как от «чёрного ящика». Рассматривается задача оптимизации размещения антенн на стадионе. Весь стадион разделён на зоны обслуживания, называемые секторами. Каждый сектор содержит некоторое число мест для посетителей, которые будем называть «ячейками». В любом таком месте посетители должны получать сигнал хорошего качества, для чего на каждый сектор назначается некоторое число антенн. Одна антенна назначается ровно на один сектор, при этом антенны одного сектора действуют совместно, усиливая сигнал друг друга, а антенны, назначенные на разные секторы, мешают друг другу. Сила сигнала в каждой ячейке вычисляется при помощи имитационной модели. Задача заключается в том, чтобы определить расположение антенн, углы их направления и назначение

на секторы так, чтобы максимизировать общее качество сигнала, определяемое несколькими метриками. Такими метриками являются число клиентов с хорошим качеством сигнала, среднее значение отношения сигнал/помехи-шум (signal to interference + noise ratio, SINR) [10] и согласованность с назначением. Согласованность определяется как число ячеек, которые получают сигнал от антенн, назначенных именно на их сектор. Равномерное распределение клиентов по секторам важно для балансирования нагрузки на сеть. Таким образом, данная задача является задачей смешанной целочисленной нелинейной многокритериальной оптимизации с неявно заданными функциями [11–14].

Для её решения в настоящей статье разработана схема из трёх процедур, применяемых последовательно. На первом этапе конструктивная эвристика создаёт начальное решение. К этому решению применяется процедура локального улучшения. Затем используется итеративная эвристика, основанная на декомпозиции задачи по секторам. Эвристика оптимизирует каждый сектор в отдельности, собирает итоговое решение с помощью модели целочисленного линейного программирования и вновь оптимизирует секторы в новом окружении. Полученные решения уточняются при помощи дополнительного запуска процедуры локального улучшения. Эти процедуры на каждом из этапов используют взвешенную свёртку критериев. Варьируя весовые коэффициенты свёртки можно получать различные решения на Парето-границе. Такая схема является развитием подхода, представленного в предыдущей статье [15]. Для оценки эффективности алгоритмического подхода использованы два реальных тестовых примера, в которых также имеются базовые решения. Сравнение производилось с базовыми решениями, а также с предыдущей версией подхода. Первый пример содержит 4426 ячеек, 17 секторов, 7 типов антенн и 27 антенн изначально. Второй пример содержит 3551 ячейку, 19 секторов, 4 типа антенн и изначально 47 антенн. Разрешается ставить не более чем в два раза больше антенн, чем в базовом решении. Эксперименты показали, что разработанная схема улучшила качество покрытия на 10 %, а SINR — по меньшей мере на 12 % без ухудшения согласованности.

Остальная часть статьи организована следующим образом. В разд. 1 представлены обозначения, формулировка задачи и модель смешанного целочисленного нелинейного программирования (MINLP). В разд. 2 приведена схема декомпозиции, являющаяся основным вкладом данной статьи. В разд. 3 приводится полная схема метода для нахождения не доминируемых по Парето решений. Разд. 4 описывает тестовые примеры и показывает эффективность алгоритмов. В заключении приводятся выводы касательно данного подхода.

1. Постановка задачи

Рассматривается следующая структура задачи. На стадионе имеется несколько уровней, каждый из которых разбит на секторы, а сектор состоит из набора ячеек. Введём множество уровней F , множество секторов S и множество ячеек C . Через S_f обозначим множество секторов на уровне $f \in F$, а через C_s — множество ячеек в секторе $s \in S$. Также введём множество I доступных мест для размещения антенн и множество T типов антенн. Наша задача — определить размещение антенн $X = \{x_j\}$, $x_j = \langle i_j, t_j, p_j, s_j \rangle$, где $i_j \in I$ — слот для размещения антенны j , $t_j \in T$ — её тип, $p_j \in [0, 360] \times [0, 90]$ — её ориентация в пространстве, а $s_j \in S$ — присвоенный ей сектор.

Для вычисления мощности $R_{cs}(X)$ принимаемого сигнала (reference signal received power, RSRP) в ячейке $c \in C$ от антенн, назначенных на сектор $s \in S$, используется имитационная модель. С помощью этой модели для каждой ячейки $c \in C$ рассчитывается значение SINR $Q_c(X)$ по формуле

$$Q_c(X) = 10 \log \frac{\max_{s \in S} R_{cs}(X)}{\sum_{s \in S} R_{cs}(X) - \max_{s \in S} R_{cs}(X)}. \quad (1)$$

Эта формула не учитывает внешний шум, поэтому величину $Q_c(X)$ можно также назвать отношением сигнал/помехи (signal to interference ratio, SIR).

Качество решения измеряется следующими метриками. *Качество покрытия* $\mathcal{Q}$ — это число ячеек, где SINR не меньше некоторого порогового значения Q_0 :

$$\mathcal{Q}(X) = |\{c \in C \mid Q_c(X) \geq Q_0\}|. \quad (2)$$

Среднее значение SINR $\mathcal{S}$ вычисляется простым способом:

$$\mathcal{S}(X) = \frac{1}{|C|} \sum_{c \in C} Q_c(C, X). \quad (3)$$

Наконец, *согласованность* $\mathcal{C}$ — это число ячеек, каждая из которых принимает наилучший сигнал от антенн, назначенных на её сектор. Сектор ячейки $c \in C$ обозначим через s_c . Тогда согласованность $\mathcal{C}$ вычисляется следующим образом:

$$\mathcal{C}(X) = |\{c \in C \mid R_{cs_c}(X) \geq R_{cs}(X) \text{ для всех } s \in S\}|. \quad (4)$$

Задача состоит в максимизации данных метрик качества

$$Q(X) = (\mathcal{Q}(X), \mathcal{S}(X), \mathcal{C}(X)) \rightarrow \max_X \quad (5)$$

так, чтобы решение удовлетворяло следующим ограничениям.

Как минимум U_0 ячеек должны получать RSRP не меньше порогового значения R_0 :

$$|\{c \in C \mid R_{cs}(X) \geq R_0\}| \geq U_0. \quad (6)$$

В общей сложности можно использовать не более N_0 антенн, т. е.

$$|X| \leq N_0. \quad (7)$$

На каждом уровне можно использовать не более двух типов антенн:

$$|\{t \in T \mid \exists j \in J \text{ такое, что } t_j = t \text{ и } s_j \in S_f\}| \leq 2, \quad f \in F. \quad (8)$$

Для формализации постановки задачи введём набор возможных углов направления антенн P . Обозначим через r_{citp} значение RSRP, принимаемое ячейкой $c \in C$, если антенна типа $t \in T$ с конфигурацией $p \in P$ помещена в слот $i \in I$. Будем использовать эти выходные параметры моделирования для вычисления общего значения R_{cs} сигнала RSRP, принимаемого ячейкой c от всех антенн, назначенных сектору s .

Введём следующие булевы переменные:

- $x_{itp} = 1$, если слот i занят антенной типа t с конфигурацией p , иначе $x_{itp} = 0$;
- $y_{ft} = 1$, если антенны типа t используются для обслуживания секторов на уровне f , иначе $y_{ft} = 0$;
- $z_{is} = 1$, если антенна в слоте i назначена сектору обслуживания s , иначе $z_{is} = 0$;
- $u_c = 1$, если ячейка c принимает сигнал достаточной мощности, иначе $u_c = 0$;
- $v_c = 1$, если ячейка c имеет хорошее качество сигнала, иначе $v_c = 0$;
- $w_c = 1$, если ячейка c получает наибольшее значение RSRP из сектора s_c , иначе $w_c = 0$.

Введя достаточно большое значение M , можно сформулировать задачу размещения антенн на стадионе [15]:

$$Q = \frac{1}{|C|} \sum_{c \in C} v_c \rightarrow \max, \quad (9)$$

$$S = \frac{1}{|C|} \sum_{c \in C} Q_c \rightarrow \max, \quad (10)$$

$$C = \frac{1}{|C|} \sum_{c \in C} w_c \rightarrow \max, \quad (11)$$

$$\sum_{t \in T} \sum_{p \in P} x_{itp} \leq 1, \quad i \in I, \quad (12)$$

$$\sum_{s \in S} z_{is} = 1, \quad i \in I, \quad (13)$$

$$\sum_{i \in I} \sum_{t \in T} \sum_{p \in P} x_{itp} \leq N_0, \quad (14)$$

$$R_{cs} = \sum_{i \in I} \sum_{t \in T} \sum_{p \in P} r_{citp} x_{itp} z_{is}, \quad c \in C, s \in S, \quad (15)$$

$$v_c Q_0 \leq Q_c, \quad c \in C, \quad (16)$$

$$Q_c = 10 \log \frac{\max_{s \in S} R_{cs}}{\sum_{s \in S} R_{cs} - \max_{s \in S} R_{cs}}, \quad c \in C, \quad (17)$$

$$R_{cs} + M(1 - w_c) \geq R_{cs'}, \quad s, s' \in S, s \neq s', c \in C_s, \quad (18)$$

$$u_c R_0 \leq \max_{s \in S} R_{cs}, \quad c \in C, \quad (19)$$

$$\sum_{c \in C} u_c \geq U_0, \quad (20)$$

$$x_{itp} \leq 1 - z_{is} + y_{ft}, \quad t \in T, i \in I, f \in F, s \in S_f, \quad (21)$$

$$\sum_{t \in T} y_{ft} \leq 2, \quad f \in F, \quad (22)$$

$$x_{itp}, y_{ft}, z_{is}, u_c, v_c, w_c \in \{0, 1\}. \quad (23)$$

Условия (9)–(11) определяют целевые функции, описанные выше. Согласно (12) в каждый слот можно поставить не более одной антенны в одной конфигурации, а (13) предписывает назначение антенны ровно на один сектор. Неравенство (14) ограничивает общее число антенн. Равенство (15) определяет значение RSRP в ячейке на основе расположения антенн. Условия (16)–(19) задают компоненты метрик качества, а (20) — минимальное число клиентов с хорошей силой сигнала. Наконец, неравенство (21) определяет типы антенн, используемых на каждом уровне, а (22) ограничивает число различных типов для одного уровня.

Одна из особенностей имитационной модели — достаточно большое время, которое занимают геометрические расчёты для определения потерь сигнала при прохождении через твёрдые предметы. Если отключить учёт этих потерь, то время моделирования значительно уменьшается, однако полученные значения RSRP отличаются от значений, посчитанных с учётом этих потерь. В разд. 3 будет приведён пример, как можно использовать эту особенность для ускорения процесса оптимизации.

2. Схема декомпозиции

2.1. Предыдущие результаты. В [15] для описанной задачи предложены конструктивная эвристика, процедура локального улучшения и алгоритм поиска с чередующимися окрестностями. Конструктивная

эвристика строит решение для каждого сектора независимо, перебирая заданный набор шаблонов расстановки антенн и используя в качестве метрики качества отношение

$$\frac{\sum_{c \in C_s} R_{cs}}{\sum_{c \in C \setminus C_s} R_{cs}}.$$

Процедура локального улучшения, представляющая собой локальный спуск по нескольким окрестностям, последовательно изменяет тип, расположение и углы каждой антенны, если это приводит к улучшению решения. Эта процедура и поиск с чередующимися окрестностями используют агрегированную целевую функцию

$$f(X, X_{\text{ref}}) = \left(\alpha \frac{\mathcal{Q}(X)}{\mathcal{Q}(X_{\text{ref}})} + \beta \frac{\mathcal{S}(X)}{\mathcal{S}(X_{\text{ref}})} + \gamma \frac{\mathcal{C}(X)}{\mathcal{C}(X_{\text{ref}})} \right) / (\alpha + \beta + \gamma), \quad (24)$$

которая на базе исходного решения X_{ref} и набора весов (α, β, γ) вычисляет относительное улучшение нового решения в сравнении со старым.

2.2. Основная идея. Идея декомпозиции схожа с техникой, применённой в конструктивной эвристике из [15]: каждый сектор оптимизируется независимо, после чего собирается итоговое решение. Однако в настоящей работе такая процедура повторяется итеративно с использованием решения, полученного на предыдущем шаге. Это позволяет брать в качестве целевых функций для задачи одного сектора целевые функции исходной задачи, вычисленные для всего стадиона. Расположение антенн в других секторах при этом считается таким же, как в решении с предыдущей итерации. Предлагаемая схема декомпозиции состоит из нескольких основных шагов.

1. Сформировать подзадачи для каждого сектора.
2. Независимо решить подзадачи.
3. Из частичных решений для секторов в отдельности собрать итоговое решение всей задачи.

2.3. Задачи одного сектора. Формируются как подзадачи исходной задачи с некоторым подмножеством переменных. Оптимизируются только антенны, назначенные на выбранный сектор. Все остальные антенны считаются зафиксированными. Рассматриваются те же целевые функции и ограничения (9)–(23), что и в исходной задаче. Однако оптимизируются только переменные, относящиеся к антеннам рассматриваемого сектора. Значения целевых функций вычисляются с использованием уровня сигнала на всём стадионе. Задачи одного сектора решаются при помощи известного популяционного алгоритма NSGA-II [16].

2.4. Сборка решения. Если есть найденные наборы недоминируемых решений для каждого сектора, то можно предложить два варианта, как собрать из них итоговое решение для исходной задачи.

- 1) Оптимизировать одну целевую функцию, не ухудшая остальные.
- 2) Оптимизировать агрегированную целевую функцию (24).

Обе эти задачи могут быть сформулированы как задачи целочисленного линейного программирования.

Оптимизация одной целевой функции. Пусть K_s — множество индексов решений для сектора $s \in S$. Обозначим значения выбранной целевой функции в решениях через f_{sk} , $s \in S$, $k \in K_s$, и пусть J — это множество остальных целевых функций. Тогда обозначим через $g_{j sk}$, $j \in J$, значения остальных целевых функций в найденных решениях, а через G_j^0 — соответствующие значения в предыдущем решении. Например, если выбрано качество покрытия, а X_s^k — k -е решение для сектора s , то

$$f_{sk} = \mathcal{Q}(X_s^k), \quad g_{1sk} = \mathcal{S}(X_s^k), \quad g_{2sk} = \mathcal{C}(X_s^k).$$

Используя эти обозначения, сформулируем первую задачу ЦЛП:

$$\sum_{s \in S} \sum_{k \in K_s} f_{sk} \delta_{sk} \rightarrow \max, \quad (25)$$

$$\sum_{k \in K_s} \delta_s^k = 1, \quad s \in S, \quad (26)$$

$$\sum_{s \in S} \sum_{k \in K_s} g_{j sk} \delta_{sk} \geq G_j^0, \quad j \in J, \quad (27)$$

$$\delta_{sk} \in \{0, 1\}, \quad s \in S, k \in K_s. \quad (28)$$

Целевая функция (25) оптимизирует выбранную метрику качества решения. Ограничение (26) предписывает взять ровно одно решение для каждого сектора, а в силу (27) значения остальных целевых функций не должны быть хуже заданных порогов.

Оптимизация агрегированной целевой функции. Аналогично если коэффициенты агрегированной целевой функции принять равными α, β, γ , то можем записать вторую модель следующим образом:

$$\sum_{s \in S} \sum_{k \in K_s} (\alpha \mathcal{Q}_{sk} + \beta \mathcal{S}_{sk} + \gamma \mathcal{C}_{sk}) \delta_{sk} \rightarrow \max, \quad (29)$$

$$\sum_{k \in K_s} \delta_s^k = 1, \quad s \in S, \quad (30)$$

$$\delta_{sk} \in \{0, 1\}, \quad s \in S, k \in K_s. \quad (31)$$

Целевая функция (29) представляет собой взвешенную сумму $\mathcal{Q}$, $\mathcal{S}$ и $\mathcal{C}$.

2.5. Итоговая схема. Пусть есть некоторое расписание *schedule* смены целевых функций или весов агрегированной функции. Обозначим значение, содержащее информацию о выбранной целевой функции или весах агрегированной функции, через *target*, а процедуру, которая выбирает и решает модель с набором решений для каждого сектора $P = \{P_1, \dots, P_{|S|}\}$, через *buildKnapsack*(*P*, *target*). С помощью этой процедуры представим итеративный алгоритм. Он использует генетический алгоритм для нахождения фронта Парето для каждого сектора, а затем собирает решение для исходной задачи, используя модель ЦЛП. Так как задачи каждого сектора решаются независимо, настоящие характеристики сигнала могут отличаться от вычисленных в модели (25)–(28) или (29)–(31). Однако на практике получаются значения, близкие к настоящим, а ошибки исправляются на следующих итерациях. Эта схема представлена в виде алгоритма 1.

Алгоритм 1. Итеративная эвристика

Вход: *schedule*, X^0 .

- 1: $X \leftarrow X^0$;
 - 2: **for** *target* $\in$ *schedule* **do**
 - 3: **for** $s \in S$ **do**
 - 4: $P_s \leftarrow \text{runSingleSector}(s)$;
 - 5: $X \leftarrow \text{buildKnapsack}(P, \text{target})$;
 - 6: **return** X .
-

Чтобы ускорить оптимизацию задач для одного сектора, популяции с предыдущей итерации используются в качестве начальных решений на следующей итерации.

Кроме того, можно добавить к ЦЛП-моделям ограничение на число типов антенн: не более двух типов антенн на одном уровне. Введём булевы переменные для указания распределения типов антенн по секторам. Пусть $v_{skt} = 1$, если сектор $s \in S$ в решении $k \in K_s$ содержит тип антенны $t \in T$, и $v_{skt} = 0$ иначе. Добавим соответствующие ограничения к обеим моделям (25)–(28) или (29)–(31) и выпишем ниже модифицированные модели (32)–(37) и (38)–(42) соответственно:

$$\sum_{s \in S} \sum_{k \in K_s} f_{sk} \delta_{sk} \rightarrow \max, \quad (32)$$

$$\sum_{k \in K_s} \delta_s^k = 1, \quad s \in S, \quad (33)$$

$$\sum_{s \in S} \sum_{k \in K_s} g_{jsk} \delta_{sk} \geq G_j^0, \quad j \in J, \quad (34)$$

$$y_{ft} \geq v_{skt} \delta_{sk}, \quad f \in F, s \in S_f, k \in K_s, t \in T, \quad (35)$$

$$\sum_{t \in T} y_{ft} \leq 2, \quad f \in F, \quad (36)$$

$$\delta_{sk}, y_{ft} \in \{0, 1\}, \quad s \in S, k \in K_s, f \in F, t \in T, \quad (37)$$

$$\sum_{s \in S} \sum_{k \in K_s} (\alpha Q_{sk} + \beta S_{sk} + \gamma C_{sk}) \delta_{sk} \rightarrow \max, \quad (38)$$

$$\sum_{k \in K_s} \delta_s^k = 1, \quad s \in S \quad (39)$$

$$y_{ft} \geq v_{skt} \delta_{sk}, \quad f \in F, s \in S_f, k \in K_s, t \in T, \quad (40)$$

$$\sum_{t \in T} y_{ft} \leq 2, \quad f \in F, \quad (41)$$

$$\delta_{sk}, y_{ft} \in \{0, 1\}, \quad s \in S, k \in K_s, f \in F, t \in T. \quad (42)$$

В этих моделях ограничения (35) и (40) указывают, что если тип используется на секторе, то он используется на уровне. Неравенства (36) и (41) гарантируют, что на каждом уровне используется не более двух типов антенн.

Вместе с тем, модель с ограничением на число типов используемых антенн может не иметь решений, если частичные решения для разных

Алгоритм 2. Итеративная эвристика с учётом типов антенн

Вход: schedule, sectorParameters, X^0 .

```

1:  $X \leftarrow X^0$ ;
2: for target  $\in$  schedule do
3:   for  $s \in S$  do
4:      $P_s \leftarrow \text{runSingleSector}(s)$ ;
5:    $X^{\text{tmp}} \leftarrow \text{buildKnapsack}(P, \text{target})$ ;
6:   if  $L \neq \emptyset$  then  $\triangleright L$  — множество уровней в решении  $X^{\text{tmp}}$ ,  
на которых больше двух типов антенн
7:     for  $l \in L$  do
8:       for  $(t_1, t_2) \in T_l(X^{\text{tmp}})$  do
9:         for  $s \in S$  do
10:           $P_s \leftarrow P_s \cup \text{runSingleSectorWithTypes}(s, t_1, t_2)$ ;
11:        $X \leftarrow \text{buildKnapsackWithTypes}(P, \text{target})$ ;
12:   else
13:      $X \leftarrow X^{\text{tmp}}$ ;
14: return  $X$ .
```

секторов содержат различные типы. Таким образом, чтобы гарантировать наличие решения, необходимо принудительно найти решения с определёнными типами антенн. Алгоритм перебирает все пары типов из числа тех, которые используются в недопустимом решении, и запрашивает вычисление дополнительных решений одного сектора с данной парой антенн. Итерационная схема с этой модификацией представлена в алгоритме 2 и далее будет упоминаться как IK (iterative knapsack).

3. Итоговая схема

Цель этой схемы — найти аппроксимацию Парето-множества с использованием алгоритмов, описанных ранее. Сначала строим исходное решение, используя конструктивную эвристику, описанную в [15]. Затем генерируем последовательность весовых коэффициентов для агрегированной целевой функции (24) и для каждого вектора выполняем процедуру локального улучшения `postoptimization`, описанную в [15], алгоритм 2 и повторное локальное улучшение. После этого сохраняем и возвращаем набор всех недоминируемых решений, найденных в ходе оптимизации. Представленная здесь схема выписана в виде алгоритма 3.

Алгоритм 3. Алгоритм взвешенного поиска

```

1:  $P \leftarrow \text{constructiveHeuristic}();$            ▷ Создать начальное решение
2: Выбрать набор весов  $W$ ;
3: for  $w \in W$  do
4:    $X \leftarrow \text{best}(P, w);$            ▷ Найти лучшее решение для данных весов
5:    $X \leftarrow \text{postoptimization}(X, w);$ 
6:    $X \leftarrow \text{IK}(X, w);$ 
7:    $X \leftarrow \text{postoptimization}(X, w);$ 
8:    $P \leftarrow P \cup \{X\};$ 
9: return  $P$ .
```

Отметим, что эта схема является модификацией схемы CPVNS из [15], которая заключается в замене поиска с чередующимися окрестностями эвристикой с использованием декомпозиции на шаге 6. Далее эту схему будем называть СРК (constructive heuristic, postoptimization, knapsack-based heuristic).

Как упоминалось в разд. 1, моделирование с учётом физических препятствий выполняется намного медленнее, поэтому в алгоритме используется линейная регрессия для прогнозирования значений целевых функций с потерями от препятствий по значениям, полученным при более

быстром моделировании без учёта таких потерь. Обозначим соответствующие значения с учётом потерь через $S_{\text{slow}}, Q_{\text{slow}}, C_{\text{slow}}, U_{\text{slow}}$, а значения при быстром моделировании через $S_{\text{fast}}, Q_{\text{fast}}, C_{\text{fast}}, U_{\text{fast}}$. Регрессионная модель использует значения быстрого симулятора $S_{\text{fast}}, Q_{\text{fast}}, C_{\text{fast}}, U_{\text{fast}}$ в качестве признаков и по ним прогнозирует настоящие значения целевых функций $S_{\text{slow}}, Q_{\text{slow}}, C_{\text{slow}}, U_{\text{slow}}$. Настоящие значения целевых функций вычисляются для лучших найденных решений, и модель обновляется после каждого запуска алгоритма на шагах 5, 6 и 7.

4. Численные эксперименты

Все эксперименты этого раздела проведены на компьютере с процессором Intel Core i7-8700 3,2 ГГц и 32 ГБ оперативной памяти под управлением Microsoft Windows 10 Pro. Алгоритмы использовали все доступные 12 потоков.

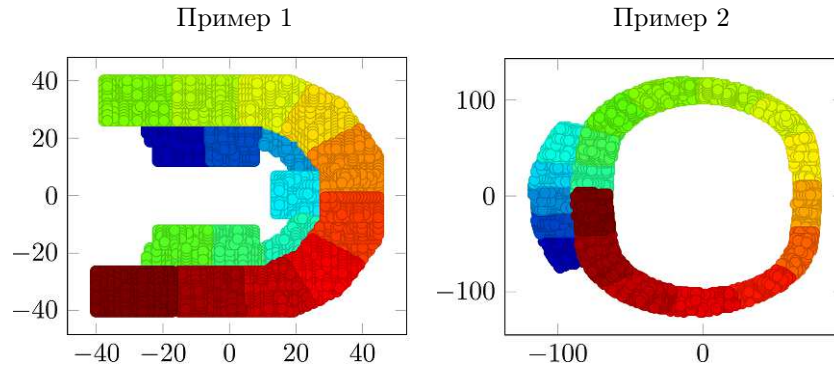


Рис. 1. Структура стадионов

Для оценки качества работы алгоритмов использовались два примера. Оба стадиона содержат 2 уровня. Первый стадион состоит из 17 секторов и 4426 ячейки, второй содержит 19 секторов и 3551 ячейку. На рис. 1 изображена структура указанных примеров. Разные секторы окрашены разными цветами.

Вместе с параметрами и данными для примеров заданы базовые решения, которые использовались для сравнения. В табл. 1 представлены

Таблица 1

Значения целевых функций для начальных решений

Пример	SINR	Качество покрытия	Согласованность
1	10,16	0,486	0,946
2	14,28	0,735	0,961

основные характеристики этих решений. В проведённых экспериментах установлен порог U_0 в 95 % ячеек с хорошим уровнем мощности сигнала, поэтому представленное решение для примера 0830 не удовлетворяет ограничениям (6) и (20). Однако это решение всё равно использовалось как некоторое хорошее для сравнения.

Проведены эксперименты для сравнения предложенной схемы с другими методами: предыдущим вариантом, использующим поиск с чередующимися окрестностями (CPVNS), и популяционным алгоритмом Borg MOEA [17]. Последний представляет собой генетический алгоритм общего назначения для многокритериальной оптимизации. Применена реализация из пакета BlackBoxOptim¹⁾. Генетический алгоритм работал 24 часа, а остальные — 2 часа. В схемах использован набор весов целевых функций $W = \{(1; 1; 1,5), (1; 1; 2,5), (2; 1; 1), (3; 1; 1)\}$. Результаты

¹⁾ <https://github.com/robertfeldt/BlackBoxOptim.jl>

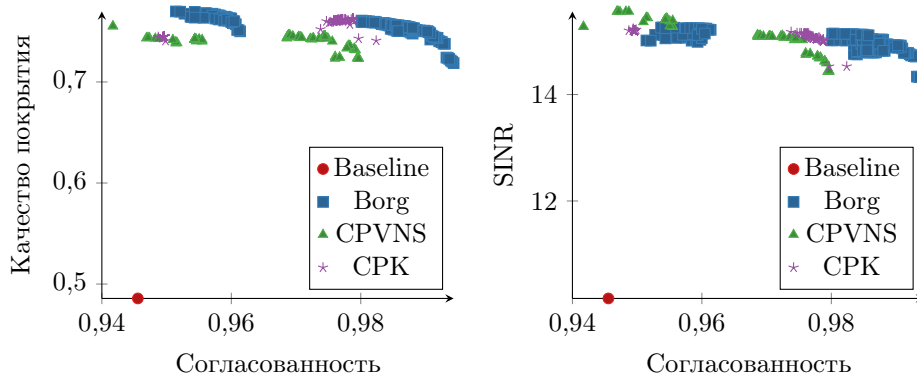


Рис. 2. Результаты алгоритмов для первого примера

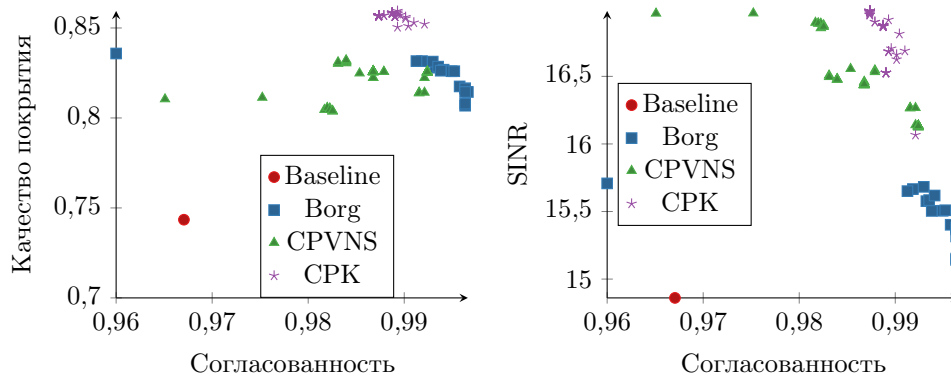


Рис. 3. Результаты алгоритмов для второго примера

Таблица 2

Относительные улучшения СРК относительно базовых решений

№ пп	Пример 1			Пример 2		
	SINR	Качество	Согл-ть	SINR	Качество	Согл-ть
1	42,9	25,7	3,4	8,1	10,9	2,5
2	42,9	25,5	3,7	12,3	11,6	2,2
3	47,6	27,7	3,3	12,3	10,9	2,4
4	49,6	25,4	0,4	12,4	11,4	2,3
5	49,8	25,8	0,4	12,1	11,3	2,3
6	49,5	25,9	0,3	13,1	10,8	2,3
7	49,2	26,6	2,8	13,6	11,5	2,1
8	47,9	27,1	3,3	13,7	11,3	2,1
9	49,5	25,9	0,4	13,8	10,7	2,2
10	48,6	27,5	3,1	14,3	11,3	2,0

представлены на рис. 2 и 3. Графики отражают попарные проекции целевых функций для двух примеров.

Видно, что обе схемы за разумное время демонстрируют хорошие результаты, сравнимые с долгим запуском алгоритма Borg MOEA. Использование предложенной в данной статье схемы декомпозиции улучшает результат по сравнению с использованием поиска с чередующимися окрестностями. Однако полученный Парето-фронт оказывается более узким из-за того, что в процессе оптимизации происходит просмотр меньшего числа полных решений для всей задачи. Также в табл. 2 представлены относительные улучшения целевых функций для некоторых 10 различных решений.

Для примера 1 удалось значительно улучшить базовое решение с точки зрения среднего уровня SINR и качества покрытия. Для примера 2 улучшение меньше, однако всё ещё удалось улучшить все три целевые функции одновременно.

Заключение

Рассмотрена задача размещения антенн на стадионе, и представлен эвристический алгоритм, основанный на декомпозиции задачи по секторам. Представленная схема состоит из трёх последовательно выполняемых процедур. Начальное решение строится при помощи простой конструктивной эвристики, после чего применяется процедура локального улучшения. Далее предложенная в этой работе эвристика, основанная на декомпозиции, решает подзадачи для каждого сектора в отдельности

и собирает полное решение при помощи модели ЦЛП. Эксперименты показали, что предложенный подход позволяет получать качественные решения за разумное время.

Финансирование работы

Исследование выполнено за счёт бюджета за счет бюджета Новосибирского гос. университета. Дополнительных грантов на проведение или руководство этим исследованием получено не было.

Конфликт интересов

Автор заявляет, что у него нет конфликта интересов.

Литература

1. **Kaya A. Ö., Calin D., Viswanathan H.** On the performance of stadium high density carrier Wi-Fi enabled LTE small cell deployments // Proc. 2015 IEEE Wireless Communications and Networking Conf. (New Orleans, USA, Mar. 9–12, 2015). Piscataway: IEEE, 2015. P. 855–860. DOI: 10.1109/WCNC.2015.7127581.
2. **Dapeng H., Ping C., Xiaoping T., Xingjian W.** Design and optimization of high density wireless network in gymnasium of Beijing Normal University // Proc. 2020 IEEE 3rd Int. Conf. Information Systems and Computer Aided Education (Dalian, China, Sept. 27–29, 2020). Piscataway: IEEE, 2020. P. 28–32. DOI: 10.1109/ICISCAE51034.2020.9236883.
3. **Hata M.** Empirical formula for propagation loss in land mobile radio services // IEEE Trans. Veh. Technol. 1980. V. 29, No. 3. P. 317–325. DOI: 10.1109/T-VT.1980.23859.
4. **Katev P. D.** Propagation models for WiMAX at 3.5 GHz // Proc. 9th Int. Conf. 2012 ELEKTRO (Žilina — Ražejské Teplice, Slovakia, May 21–22, 2012). Piscataway: IEEE, 2012. P. 61–65. DOI: 10.1109/ELEKTRO.2012.6225572.
5. **Naderi Soorki M., Saad W., Bennis M.** Optimized deployment of millimeter wave networks for in-venue regions with stochastic users' orientation // IEEE Trans. Wirel. Commun. 2019. V. 18, No. 11. P. 5037–5049. DOI: 10.1109/TWC.2019.2931535.
6. **Šlapak E., Gazda J., Bugár G., Maksymyuk T.** Review of cellular radio network cell placement design, from traditional to artificial intelligence based approaches // Proc. 61st Int. Symp. ELMAR-2019 (Zadar, Croatia, Sept. 23–25, 2019). Zadar: Croat. Soc. Electron. Marine, 2019. P. 93–96. DOI: 10.1109/ELMAR.2019.8918668.
7. **Vanhatupa T., Hannikainen M., Hamalainen T. D.** Genetic algorithm to optimize node placement and configuration for WLAN planning // Proc. 4th Int. Symp. Wireless Communication Systems 2007 (Trondheim, Norway, Oct. 17–19, 2007). Piscataway: IEEE, 2007. P. 612–616. DOI: 10.1109/ISWCS.2007.4392413.

8. Narayanan L., Subramanian B., Arokiaswami A., Iruthayaranjan M. W. Optimal placement of mobile antenna in an urban area using evolutionary multiobjective optimization // *Microw. Opt. Technol. Lett.* 2012. V. 54, No. 3. P. 737–743. DOI: 10.1002/mop.26627.
9. Raisanen L., Whitaker R. M. Comparison and evaluation of multiple objective genetic algorithms for the antenna placement problem // *Mobile Netw. Appl.* 2005. V. 10, No. 1. P. 79–88. DOI: 10.1023/B:MONE.0000048547.84327.95.
10. Goldsmith A. *Wireless communications*. Cambridge: Camb. Univ. Press, 2005. 644 p.
11. Audet C., Hare W. *Derivative-free and blackbox optimization*. Cham: Springer, 2017. 308 p. DOI: 10.1007/978-3-319-68913-5.
12. *Black box optimization, machine learning, and no-free lunch theorems*. Cham: Springer, 2021. 388 p. DOI: 10.1007/978-3-030-66515-9.
13. Yuskov A. D., Kulachenko I. N., Melnikov A. A., Kochetov Yu. A. Decomposition approach to simulation-based optimization of inventory management // *Mathematical optimization theory and operations research: Recent trends*. Rev. Sel. Pap. 22nd Int. Conf. MOTOR 2023 (Yekaterinburg, Russia, July 2–8, 2023). Cham: Springer, 2023. P. 259–273. (*Commun. Comput. Inf. Sci.*; V. 1881). DOI: 10.1007/978-3-031-43257-6_20.
14. Yuskov A. D., Kulachenko I. N., Melnikov A. A., Kochetov Yu. A. Two-stage algorithm for bi-objective black-box traffic engineering // *Optimization and applications*. Rev. Sel. Pap. 14th Int. Conf. OPTIMA 2023 (Petrovac, Montenegro, Sept. 18–22, 2023). Cham: Springer, 2023. P. 110–125. (*Lect. Notes Comput. Sci.*; V. 14395). DOI: 10.1007/978-3-031-47859-8_9.
15. Yuskov A. D., Kulachenko I. N., Melnikov A. A., Kochetov Yu. A. Stadium antennas deployment optimization // *Mathematical optimization theory and operations research*. Proc. 23rd Int. Conf. MOTOR 2024 (Omsk, Russia, June 30–July 6, 2024). Cham: Springer, 2024. P. 449–461. (*Lect. Notes Comput. Sci.*; V. 14766). DOI: 10.1007/978-3-031-62792-7_30.
16. Deb K., Agrawal S., Pratap A., Meyarivan T. A fast elitist non-dominated sorting genetic algorithm for multi-objective optimization: NSGA-II // *Parallel problem solving from nature — PPSN VI*. Proc. 6th Int. Conf. (Paris, France, Sept. 18–20, 2000). Heidelberg: Springer, 2000. P. 849–858. (*Lect. Notes Comput. Sci.*; V. 1917). DOI: 10.1007/3-540-45356-3_83.
17. Hadka D., Reed P. Borg: An auto-adaptive many-objective evolutionary computing framework // *Evol. Comput.* 2013. V. 21, No. 2. P. 231–259. DOI: 10.1162/EVCO_a_00075.

Юськов Александр Дмитриевич

Статья поступила

19 июля 2024 г.

После доработки —

17 августа 2024 г.

Принята к публикации

22 сентября 2024 г.

A DECOMPOSITION APPROACH
TO A STADIUM ANTENNA DEPLOYMENT PROBLEM*A. D. Yuskov*Novosibirsk State University,
2 Pirogov Street, 630090 Novosibirsk, Russia
E-mail: a.yuskov@g.nsu.ru

Abstract. We consider a stadium antenna deployment problem. The stadium is divided into sectors. Several antennas are assigned to each sector. Users should receive a signal of a certain quality from antennas assigned to their sector. The problem is to choose locations of antennas, their types, angles, and assignments to sectors to maximize three quality criteria: the mean signal to interference ratio (SIR), the number of clients with good signal quality, and the assignment consistency. We use a simulation to compute the signal quality. We present a three-stage heuristic approach to the problem. It uses a constructive heuristic, a local improvement procedure, and a decomposition-based MIP heuristic. We carry out numerical experiments on test instances with 94 antennas of 7 types, 19 sectors, and 4426 clients. It is possible to improve the provided baseline solutions in 2 h and obtain solutions comparable to running a metaheuristic package for 24 h. Tab. 2, illustr. 3, bibliogr. 17.

Keywords: black box optimization, metaheuristics, wireless network, signal quality, SINR.

References

1. **A. Ö. Kaya, D. Calin, and H. Viswanathan,** On the performance of stadium high density carrier Wi-Fi enabled LTE small cell deployments, in *Proc. 2015 IEEE Wireless Communications and Networking Conf., New Orleans, USA, Mar. 9–12, 2015* (IEEE, Piscataway, 2015), pp. 855–860, DOI: 10.1109/WCNC.2015.7127581.

2. **H. Dapeng, C. Ping, T. Xiaoping, and W. Xingjian**, Design and optimization of high density wireless network in gymnasium of Beijing Normal University, in *Proc. 2020 IEEE 3rd Int. Conf. Information Systems and Computer Aided Education, Dalian, China, Sept. 27–29, 2020* (IEEE, Piscataway, 2020), pp. 28–32, DOI: 10.1109/ICISCAE51034.2020.9236883.
3. **M. Hata**, Empirical formula for propagation loss in land mobile radio services, *IEEE Trans. Veh. Technol.* **29** (3), 317–325 (1980), DOI: 10.1109/T-VT.1980.23859.
4. **P. D. Katev**, Propagation models for WiMAX at 3.5 GHz, in *Proc. 9th Int. Conf. 2012 ELEKTRO, Žilina — Rajecské Teplice, Slovakia, May 21–22, 2012* (IEEE, Piscataway, 2012), pp. 61–65, DOI: 10.1109/ELEKTRO.2012.6225572.
5. **Naderi M. Soorki, W. Saad, and M. Bennis**, Optimized deployment of millimeter wave networks for in-venue regions with stochastic users' orientation, *IEEE Trans. Wirel. Commun.* **18** (11), 5037–5049 (2019), DOI: 10.1109/TWC.2019.2931535.
6. **E. Šlapak, J. Gazda, G. Bugár and T. Maksymyuk**, Review of cellular radio network cell placement design, from traditional to artificial intelligence based approaches, in *Proc. 61st Int. Symp. ELMAR-2019, Zadar, Croatia, Sept. 23–25, 2019* (Croat. Soc. Electron. Marine, Zadar, 2019), pp. 93–96, DOI: 10.1109/ELMAR.2019.8918668.
7. **T. Vanhatupa, M. Hannikainen, and T. D. Hamalainen**, Genetic algorithm to optimize node placement and configuration for WLAN planning, in *Proc. 4th Int. Symp. Wireless Communication Systems 2007, Trondheim, Norway, Oct. 17–19, 2007* (IEEE, Piscataway, 2007), pp. 612–616, DOI: 10.1109/ISWCS.2007.4392413.
8. **L. Narayanan, B. Subramanian, A. Arokiaswami, and M. W. Iruthayarajan**, Optimal placement of mobile antenna in an urban area using evolutionary multiobjective optimization, *Microw. Opt. Technol. Lett.* **54** (3), 737–743 (2012), DOI: 10.1002/mop.26627.
9. **L. Raisanen and R. M. Whitaker**, Comparison and evaluation of multiple objective genetic algorithms for the antenna placement problem, *Mobile Netw. Appl.* **10** (1), 79–88 (2005), DOI: 10.1023/B:MONE.0000048547.84327.95.
10. **A. Goldsmith**, *Wireless Communications* (Camb. Univ. Press, Cambridge, 2005).
11. **C. Audet and W. Hare**, *Derivative-Free and Blackbox Optimization* (Springer, Cham, 2017), DOI: 10.1007/978-3-319-68913-5.
12. *Black Box Optimization, Machine Learning, and No-Free Lunch Theorems* (Springer, Cham, 2021), DOI: 10.1007/978-3-030-66515-9.
13. **A. D. Yuskov, I. N. Kulachenko, A. A. Melnikov, and Yu. A. Kochetov**, Decomposition approach to simulation-based optimization of inventory management, in *Mathematical Optimization Theory and Operations Research: Recent Trends* (Rev. Sel. Pap. 22nd Int. Conf. MOTOR 2023, Yekaterinburg, Russia, July 2–8, 2023) (Springer, Cham, 2023), pp. 259–273 (Commun. Comput. Inf. Sci., Vol. 1881), DOI: 10.1007/978-3-031-43257-6_20.

14. **A. D. Yuskov, I. N. Kulachenko, A. A. Melnikov, and Yu. A. Kochetov**, Two-stage algorithm for bi-objective black-box traffic engineering, in *Optimization and Applications* (Rev. Sel. Pap. 14th Int. Conf. OPTIMA 2023, Petrovac, Montenegro, Sept. 18–22, 2023) (Springer, Cham, 2023), pp. 110–125 (Lect. Notes Comput. Sci., Vol. 14395), DOI: 10.1007/978-3-031-47859-8_9.
15. **A. D. Yuskov, I. N. Kulachenko, A. A. Melnikov, and Yu. A. Kochetov**, Stadium antennas deployment optimization, in *Mathematical optimization theory and operations research* (Proc. 23rd Int. Conf. MOTOR 2024, Omsk, Russia, June 30 – July 6, 2024) (Springer, Cham, 2024), pp. 449–461 (Lect. Notes Comput. Sci., Vol. 14766), DOI: 10.1007/978-3-031-62792-7_30.
16. **K. Deb, S. Agrawal, A. Pratap, and T. Meyarivan**, A fast elitist non-dominated sorting genetic algorithm for multi-objective optimization: NSGA-II, in *Parallel Problem Solving from Nature — PPSN VI* (Proc. 6th Int. Conf., Paris, France, Sept. 18–20, 2000) (Springer, Heidelberg, 2000), pp. 849–858 (Lect. Notes Comput. Sci., Vol. 1917), DOI: 10.1007/3-540-45356-3_83.
17. **D. Hadka and P. Reed**, Borg: An auto-adaptive many-objective evolutionary computing framework, *Evol. Comput.* **21** (2), 231–259 (2013), DOI: 10.1162/EVC0_a_00075.

Aleksandr D. Yuskov

Received July 19, 2024

Revised August 17, 2024

Accepted September 22, 2024

ДИСКРЕТНЫЙ АНАЛИЗ
И ИССЛЕДОВАНИЕ ОПЕРАЦИЙ

2025. Том 32, № 1

Зав. редакцией Ю. В. Шамардин

Журнал подготовлен с использованием макропакета $\text{\LaTeX 2}_{\epsilon}$.

The present publication has been typeset using $\text{\LaTeX 2}_{\epsilon}$.

Журнал зарегистрирован в Федеральной службе по надзору
в сфере связи, информационных технологий и массовых коммуникаций.

Свидетельство о регистрации ЭЛ № ФС77-85978 от 26.09.2023 г.

Размещение в сети Интернет: math-sobolev.ru.

Дата размещения в сети Интернет 22.09.2025 г.

Формат 70×100 1/16. Усл. печ. л. 13,3. Объём 2,62 МБ.

Издательство Института математики,
пр. Академика Коптюга, 4, 630090 Новосибирск, Россия