DOI: 10.15587/2312-8372.2017.108909

РЕШЕНИЕ НЕКОТОРЫХ ЗАДАЧ ТЕОРИИ РАСПИСАНИЙ МЕТОДАМИ ВЫПУКЛОЙ ОПТИМИЗАЦИИ

Косолап А. И., Нестеренко А. Н.

1. Введение

Вопрос упорядочения процессов и алгоритмов необходим в любой сфере деятельности, где количество таких процессов достаточно велико [1]. Отдельный класс задач такого рода связан с упорядочиванием выполнения некоторого количества процессов (заданий) на заданном количестве устройств так называемые «задачи теории расписания» [2]. Проблемы обработки информации, технологические процессы промышленных предприятий, транспортная задача – все они (и не только) сводятся к задачам теории расписаний. По своему содержанию многие задачи теории расписаний являются оптимизационными. Они состоят в выборе (нахождении) среди множества допустимых расписаний (расписаний, допускаемых условиями задачи) тех решений, на которых достигается «оптимальное» значение целевой функции. Предполагается, что в начале периода планирования известен перечень заданий, подлежащих выполнению. Каждое задание представляет собой совокупность взаимозависимых работ, выполнение которых осуществляется на отдельных устройствах. Для каждого задания указывается время обработки на каждой машине, порядок обслуживания и сроки выполнения. Степень трудоемкости задач теории расписаний определяется числом заданий (n) и количеством устройств обработки (m). Отдельный класс задач «потоковая линия» (flow shop) фиксирует порядок использования машин и предполагает последовательное многостадийное выполнение каждого задания на каждой машине в установленном порядке. Решением задачи является последовательность выполнения заданий, при которой минимизируется общее время обработки. Несмотря на простоту постановки задач теории расписаний и все увеличивающееся количество работ, посвященных их решению, алгоритм, который позволит находить оптимальное или приближенное решение за «разумное» время, не найден. Предлагаемый подход к решению задач теории расписаний методами выпуклой оптимизации позволяет за полиноминальное время получить оптимальное или приближенное решение.

2. Объект исследования и его технологический аудит

Объектом данного исследования являются задачи теории расписаний, включающие в себе определенные ресурсы (устройства обработки), набор заданий, ограничения предшествования и методы оценки расписания. Рассмотрена так называемая «конвейерная задача» теории расписаний, которая включает в себя все перечисленные выше составляющие, при этом очередность обработки каждого задания на устройствах одинакова и определена. Искомое расписание должно обеспечивать минимальное время Т выполнения всех заданий. Граф конвейерной задачи показан на рис. 1.

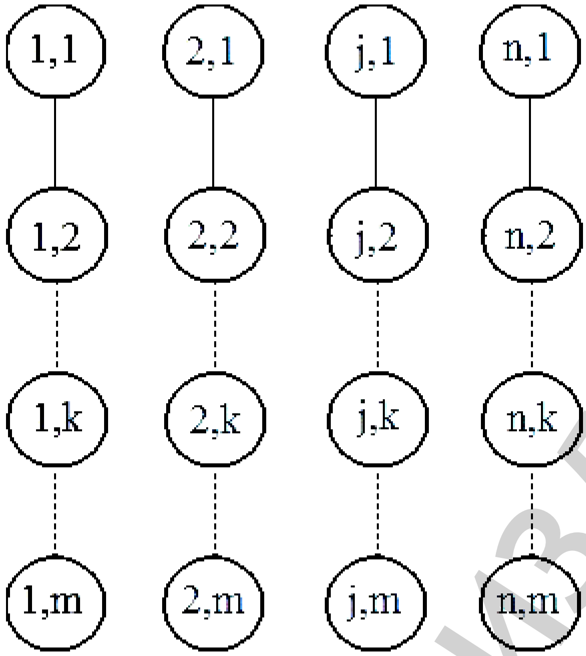

Рис. 1. Конвейерная задача. Каждое задание в одинаковой последовательности проходит все устройства

Оптимальное решение сформулированной проблемы можно получить, рассматривая всевозможные очередности выполнения исходных заданий. Количество возможных вариантов очередности выполнения заданий равно числу перестановок из *n* элементов, т. е. *n!* вариантов. Поэтому при *n~10-20* получить оптимальное решение уже невозможно, так как объем вычислений и время их выполнения будут превышать все разумные пределы.

3. Цель и задачи исследования

Целью данной работы является разработка способа решения конвейерных задач теории расписаний на основе методов выпуклой оптимизации. Этот метод должен обеспечивать расчет оптимального расписания с заданной точностью за приемлемое время.

Для достижения поставленной цели необходимо:

1. Переформулировать ограничения конвейерной задачи в виде набора аналитических функций переменных этой задачи.

2. Переписать ограничения в виде выпуклых функций и сформулировать задачу выпуклой оптимизации.

3. Выбрать оптимальный метод решения многоэкстремальных задач выпуклой оптимизации.

4. Исследование существующих решений проблемы

Классификация различных задач теории расписаний рассмотрены в [3]. Как известно, задачи теории расписаний в общей формулировке при количестве устройств обработки большем, чем $m>2$ не имеют алгоритма решения и являются NP-полными [1, 2]. В работе [4] предпринята попытка использовать алгоритм Джонсона (*m=2)* для построения оптимального расписания с большим количеством устройств обработки. Но, как отмечают авторы, подобный подход не приводит к оптимальному решению, а позволяет найти некоторое приближение к нему. Оценка отличия найденного решения от оптимального практически неосуществима. Традиционный подход к решению таких задач использует метод последовательного раскрывания модулей или метод перестановок заданий [5], что приводит к большим затратам времени расчета вариантов. Метод ветвей и границ [1], а также его различные модификации [6], усовершенствуют методы перестановок заданий, применяемые при решении конвейерной задачи теории расписаний. Но они включают в себя дополнительные процедуры – процедуру оценок расписаний, процедуру ветвления, процедура отсева и процедура останова. Основным недостатком метода ветвей и границ является необходимость определения оценок в каждой вершине дерева ветвления, а при большом числе заданий число вершин становится значительным [7]. Это не позволяет провести рассмотрение всех вершин дерева ветвлений. В [8] отмечено, что современные модели и методы теории расписаний можно разделить на две группы: приближенные (эвристические) и перестановочные с различными вариантами оптимизации перестановок. В результате, при использовании методов первой группы теряется точность решений, а время решения задачи методами второй группы может выходить за разумные границы. Поэтому использование общих методов глобальной оптимизации [9, 10], позволяющих за разумное время получить решение задачи теории расписаний, является перспективным. Данная работа посвящена математической формулировке задач теории расписаний в виде, удобном для использования методов глобальной оптимизации.

5. Методы исследования

Пусть m – количество устройств обработки; n – количество заданий, выполняемых на этих устройствах, причем каждое задание включает обработку на каждом из m устройств. Обозначим:

$$t_{\min} = \min(t_{ij}),$$

где t_{ij}– время выполнения i-ого задания на j-ом устройстве, время начала выполнения этого задания на выбранном устройстве обозначим x_{ij} (переменные задачи расписания). Введем вспомогательную функцию занятости устройств [8] в виде:

$$g_{ij}\left(t-x_{ij}\right)=\frac{1}{2}\left(\frac{\left|t-x_{ij}\right|}{t_{ij}}+\frac{\left|t-x_{ij}-t_{ij}-t_{\min}\right|}{t_{\min}}-\left|t-x_{ij}-t_{ij}\right|*\left(\frac{1}{t_{ij}}+\frac{1}{t_{\min}}\right)\right), \qquad (1)$$

график которой представлен на рис. 2.

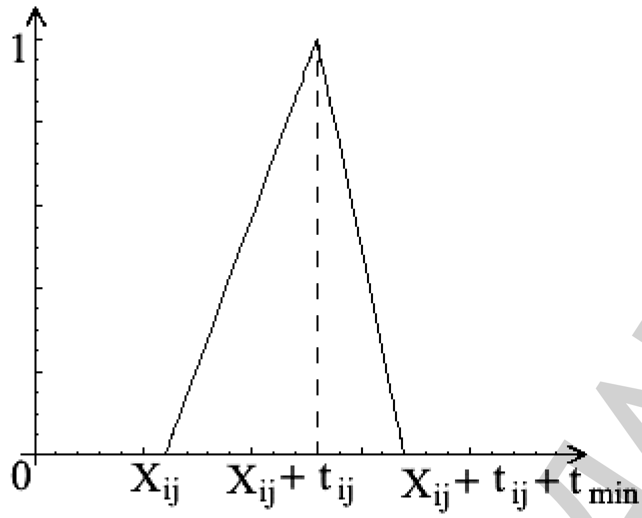

Рис. 2. График функции $g_{ij}\left(t-x_{ij}\right)$

Условие: на одном устройстве в данный момент времени обрабатывается только одно задание равносильно неравенству:

$$\sum_{i=1}^{n} g_{ij}\left(t-x_{ij}\right) \le 1,\ \forall t \ge 0,\ j=1,...,m. \tag{2}$$

Аналогично, второе условие: каждое задание не может одновременно выполняться на двух устройствах, принимает вид:

$$\sum_{j=1}^{m} g_{ij}\left(t-x_{ij}\right) \le 1,\ \forall t \ge 0,\ i=1,...,n. \tag{3}$$

Общее время выполнения всех заданий на устройствах удовлетворяет дополнительному набору линейных ограничений:

$$T \ge x_{im} + t_{im},\ i=1,...n, \tag{4}$$

и представляет собой целевую функцию конвейерной задачи. Минимизации величины $\boldsymbol{T}$ добиваются, изменяя переменные (x_{kj}). Приходим к следующей задаче минимизации:

$$\min(T|\sum_{j=1}^{m} g_{ij}\left(t-x_{ij}\right) \le 1, \sum_{i=1}^{n} g_{ij}\left(t-x_{ij}\right) \le 1, T \ge x_{im} + t_{im}). \tag{5}$$

Оценим ограничения на переменные (x_{kj}), которые следуют из соотношений (2), (3) задачи (5).

Рассмотрим систему неравенств (2), которая должна выполняться в любой момент времени t, в том числе и для моментов времени, соответствующих максимуму каждой из входящих в сумму функций g_{ij}, т. е. моменты времени $t = x_{1j} + t_{1j}, t = x_{2j} + t_{2j}, t = x_{3j} + t_{3j}$ и так далее. В результате получим следующую систему неравенств:

$$\sum_{i=1}^{n} g_{ij}\left(x_{kj} + t_{kj} - x_{ij}\right) \le 1,\ j = 1,...m,\ k = 1,....n. \tag{6}$$

В случае $i=k$ в каждом из неравенств (6) окажется слагаемое вида $g_{ij}\left(t_{ij}\right)$, которое в силу определения (1) тождественно равно единице, и систему неравенств (6) можно переписать следующим образом:

$$\sum_{i=1}^{n} g_{ij}\left(x_{kj} + t_{kj} - x_{ij}\right) \le 0,\ j = 1,...m,\ k = 1,...n,\ k \ne i. \tag{7}$$

Так как функции g_{ij} по определению неотрицательны, то каждое из входящих в сумму слагаемых должно быть равно нулю. Приходим к набору ограничений на разности $x_{kj} - x_{ij}$, причем $k \ne i$. Указанные разности встречаются дважды в уравнениях с переставленными индексами i и k, причем знаки разностей противоположны, поэтому одновременно должны выполняться два условия:

$$\begin{gathered} g_{ij}\left(x_{kj} - x_{ij} + t_{kj}\right) = 0 \text{ и } g_{kj}\left(x_{ij} - x_{kj} + t_{ij}\right) = 0, \\ j = 1,...m,\ k = 1,...n,\ k \ne i. \end{gathered} \tag{8}$$

Нуль функции $g_{ij}(z)$ достигается при значениях аргумента $z<0$ или $z > t_{ij} + t_{\min}$, что приводит к неравенствам:

$$\begin{gathered} x_{kj} - x_{ij} + t_{kj} \le 0,\ x_{kj} - x_{ij} + t_{kj} \ge t_{ij} + t_{\min}, \\ x_{ij} - x_{kj} + t_{ij} \le 0,\ x_{ij} - x_{kj} + t_{ij} \ge t_{kj} + t_{\min}, \end{gathered}$$

или

$$\begin{gathered} x_{kj} - x_{ij} \le -t_{kj},\ x_{kj} - x_{ij} \ge t_{ij} - t_{kj} + t_{\min}, \\ x_{kj} - x_{ij} \ge t_{ij},\ x_{kj} - x_{ij} \le -t_{kj} + t_{ij} - t_{\min}. \end{gathered} \tag{9}$$

В выражении (9) выполнение первого неравенства автоматически приводит к выполнению четвертого неравенства, а выполнение третьего неравенства автоматически приводит к выполнению второго. Таким образом, из набора неравенств (9) действующими оказываются два неравенства:

$$x_{kj} - x_{ij} \le -t_{kj},\ x_{kj} - x_{ij} \ge t_{ij}, \tag{10}$$

и выполнение любого из них обнуляет каждую из функций в выражении (8).

Первое из неравенств (10) указывает, что выполнение *i*-го задания на *j*-ом устройстве не может начинаться до тех пор, пока на этом устройстве не закончится его обработка на *k*-го задания. Второе выражение указывает, что выполнение *k*-го задания на *j*-ом устройстве не может начинаться до тех пор, пока не закончится обработка на этом устройстве *i*-го задания. В зависимости от очередности выполнения заданий будет реализовано одно из этих неравенств, но выполнение обоих одновременно невозможно. Объединение этих условий возможно в выражении:

$$(x_{kj} - x_{ij} + t_{kj})(x_{kj} - x_{ij} - t_{ij}) \ge 0. \tag{11}$$

Выражение (11) определяет две параллельные гиперплоскости в пространстве переменных (x_{kj}). Область пространства между гиперплоскостями не удовлетворяет условию (11) и допустимых решений задачи (5) здесь быть не может. Всевозможные решения задачи (5) (в том числе и оптимальное) будут находится снаружи от гиперплоскостей. Таким образом, полная система неравенств (9), накладывающая ограничения на переменные x_{kj}, x_{ij} (на одном устройстве в данный момент времени обрабатывается только одно задание), может быть представлена в виде:

$$\begin{gathered}(x_{kj} - x_{ij} + t_{kj})(x_{kj} - x_{ij} - t_{ij}) \ge 0,\ j = 1,\dots m,\\ i = 1,\dots n-1,\ k = i+1,\dots n.\end{gathered} \tag{12}$$

Для случая трех ограничений вида (12) на рис. 3 показано взаимное расположение пар гиперплоскостей и области возможных решений задачи (5).

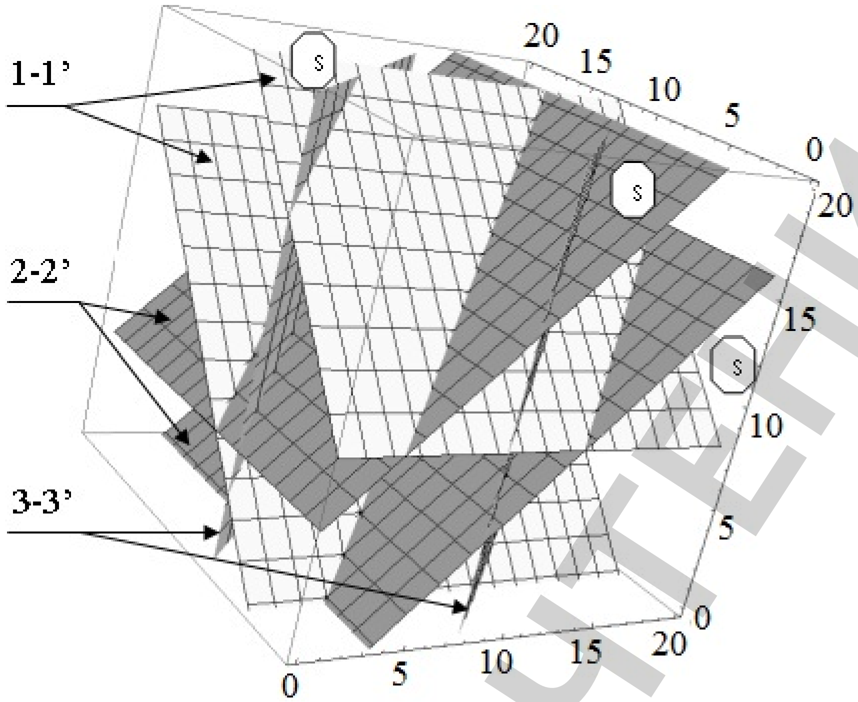

Рис. 3. Взаимное расположение пар гиперплоскостей и области возможных решений задачи (5): $1-1^{'},2-2^{'},3-3^{'}$ – пары гиперплоскостей, соответствующие ограничениям (11)

Допустимые решения на рис. 1 возможны в областях, находящихся снаружи всех пар гиперплоскостей. Эти области обозначены как S. Система неравенств (3) рассматривается аналогично, с соответствующей заменой индексов, что приводит к следующим ограничениям:

$$(x_{ij}-x_{iq}+t_{ij})(x_{ij}-x_{iq}-t_{iq})>0,$$
$$i=1,...n,\ q=1,...m,\ j=q+1,...m. \qquad (13)$$

В случае конвейерной задачи последовательность выполнения заданий на устройствах определена и неравенства (13) необходимо заменить условиями:

$$x_{ik+1}\geq x_{ik}+t_{ik},\ i=1,....n,\ k=1,....m-1. \qquad (14)$$

Сопоставляя полученные результаты с графом конвейерной задачи на рис. 1, можно утверждать, что для каждой линии графа, определяющей последовательность выполнения операций в задании, записывается линейное ограничение вида (14). Если для выбранного задания последовательность операций на устройствах безразлична, то в этом случае необходимо использовать ограничение вида (13). Для заданий, выполняемых на одном устройстве, но не связанных линиями графа, должны быть записаны ограничения вида (12).

6. Результаты исследований

Предлагаемая в данной работе процедура позволяет записать ограничения задачи теории расписаний в виде набора гладких, выпуклых, дважды дифференцируемых функций, количество которых не превосходит:

$$N_{общ} = \frac{1}{2} n \cdot m(m + n - 2). \tag{15}$$

Ограничений вида (14) всегда меньше, чем ограничений вида (13), поэтому для конвейерной задачи их общее число будет меньше, чем в (15):

$$N_{общ} = n \cdot (m-1) + \frac{1}{2} m \cdot n(n-1) = \frac{1}{2} n \cdot m(n+1) - n. \tag{16}$$

Таким образом, любая конвейерная задача теории расписаний сводится к задаче минимума линейной функции (min(T)) (4) с набором линейных и квадратичных ограничений (12), (14), т. е. к так называемой проблеме выпуклой оптимизации. Развитие теории выпуклой оптимизации в последние годы проходило очень интенсивно и были разработаны достаточно надежные и эффективные методы нахождения оптимального решения задачи с заданной точностью [10]. Эти методы представлены несколькими «соперничающими» направлениями, имеющими свои сильные и слабые стороны, описание которых подробно изложены в [11]. Наибольший прогресс в глобальной оптимизации получен с помощью метода точной квадратичной регуляризации EQR [12]. Этот метод применим для широкого класса многоэкстремальных задач и позволяет разбить их на два класса сложности. Первый класс сводится к минимизации нормы вектора на выпуклом множестве, а второй – к максимизации нормы вектора на выпуклом множестве. Задача нахождения минимума линейной функции (min(T)) (4) с набором линейных и квадратичных ограничений (12), (14) относится ко второму классу сложности [9].

Использование метода точной квадратичной регуляризации преобразует задачу (min(T)) с ограничениями (12)–(14), к следующему виду:

$$\begin{aligned}
&\max\{\| x\|^2 \mid T + s + (r-1)\| x\|^2 \le d\},\\
&(x_{kj} - x_{ij} + t_{kj})(x_{ij} - x_{kj} + t_{ij}) + r\| x\|^2 \le d,\\
&x_{ik+1} \ge x_{ik} + t_{ik},
\end{aligned} \tag{17}$$

где s – фиксированный параметр, удовлетворяющий условию:

$$T^* + s \ge \| x^*\|^2, \tag{18}$$

а T^* и x^* – решение задачи (17), значение параметра $r>0$ выбирается таким, чтобы допустимая область решений задачи (17) была выпуклой. Действительно, параметр r входит во все ограничения задачи (17), гессианы этих функций при соответствующем выборе параметра r будут положительно определенными матрицами и функции ограничений будут выпуклыми. Компонентами вектора x являются искомые переменные задачи расписания. В задаче (17) необходимо найти минимальное значение параметра d>0, для которого одновременно с условием (18) выполняется условие:

$$r\|x\|^2 = d. \tag{19}$$

Параметр s наряду с параметром d определяет область поиска решения и его выбор в соответствии с условием (18) приводит к нахождению глобального минимума времени прохождения расписания T^* и соответствующих переменных этого расписания x^*, т. е. решению задачи.

Таким образом, конвейерная задача теории расписаний сводится к одной из традиционных задач глобальной оптимизации, а именно минимизации квадрата нормы вектора x с квадратичными ограничениями (17).

Для демонстрации работоспособности изложенной процедуры была выбрана модель конвейерной обработки, состоящая из пяти заданий (n=5) и трех устройств последовательной обработки (m=3). Время выполнения каждого из заданий на устройствах t_{ij} выбирались случайным образом и заданы табл. 1.

Таблица 1

Время выполнения каждого из заданий на отдельных устройствах

Устройство обработки	Задания				
	1	2	3	4	5
1	5	4	7	2	3
2	8	4	1	8	5
3	1	9	4	4	1

В результате проведенной оптимизации методом точной квадратичной регуляризации (EQR) были получены значения x_{ij} времени начала обработки заданий на устройствах, значения которых приведены в табл. 2.

Таблица 2

Время начала выполнения каждого из заданий на отдельных устройствах

Устройство обработки	Задания				
	1	2	3	4	5
1	6	2	11	0	18
2	14	10	22	2	23
3	23	14	24	10	28

Время окончания выполнения всех заданий равно 29 и является минимальным для данной задачи, а полученное расписание оптимально. При этом выполняются все ограничения последовательности операций (14), представленные в табл. 3 и ограничения (12) обработки только одного задания на любом устройстве в данный момент времени (табл. 4). Для наглядности ограничения (14) переписаны в виде:

$$x_{ik+1} - x_{ik} - t_{ik} \geq 0, \ i = 1,....n, \ k = 1,....m-1. \tag{20}$$

Таблица 3
Разница времен начала выполнения последующей операции и окончания предыдущей для каждого из заданий

Пара операций	Задания				
	1	2	3	4	5
1–2	3	4	4	0	2
2–3	1	0	1	0	0

Таблица 4
Численные значения ограничений (12) для каждой из машин и возможных пар заданий

Номер устройства	Пара заданий									
	1–2	1–3	1–4	1–5	2–3	2–4	2–5	3–4	3–5	4–5
1	00	00	44	105	80	00	228	162	00	336
2	00	00	80	14	104	00	162	252	00	338
3	00	00	126	24	14	00	75	180	00	266

Приведенные решения были получены численными расчетами с использованием программы Excel Solver (США).

7. SWOT-анализ результатов исследований

Strengths. При использовании традиционных методов теории расписаний увеличение числа заданий или устройств приводит к экспоненциальному возрастанию времени решения задачи. Предлагаемый метод преобразования конвейерной задачи теории расписаний сводит ее к задаче выпуклой оптимизации. Количество ограничений в этой задаче оказывается полиноминальным по отношению к числу заданий и числу устройств обработки. Поэтому увеличение числа заданий или числа устройств обработки приводит только к полиноминальному возрастанию времени решения задачи. Поэтому время, необходимое для расчета оптимального расписания, будет на порядки меньше, чем в случае использования традиционных методов.

Weaknesses. Задача выпуклой оптимизации, полученная преобразованием конвейерной задачи, оказывается многоэкстремальной. Поэтому не все методы выпуклой оптимизации позволяют найти глобальный минимум, т. е. оптимальное решение.

Найденное решение может быть близким к оптимальному, но не совпадать с ним. Поэтому возможно потребуется повторить решение задачи с другими начальными условиями.

Opportunities. Метод точной квадратичной регуляризации (EQR) позволяет найти глобальный минимум многоэкстремальной задачи. Алгоритмизация определения дополнительных параметров метода (EQR) и их изменения позволит автоматизировать процесс нахождения оптимального расписания. В этом случае затраченное время на расчет оптимального расписания сократится. Это даст возможность оптимизировать использование ресурсов в зависимости от изменяющегося со временем потока заданий.

Threats. Дополнительные затраты будут связаны с разработкой и внедрением нового программного обеспечения для определения оптимальных расписаний.

8. Выводы

1. Изменена формулировка ограничений конвейерной задачи и приведена к набору аналитических функций переменных этой задачи. Тем самым переходим от задач дискретной математики к задачам классического математического анализа.

2. Аналитические функции ограничений являются выпуклыми. Таким образом, исходная задача переформулирована в задачу выпуклой оптимизации. Это позволяет решать задачу методами выпуклой оптимизации.

3. Выбран метод точной квадратичной регуляризации (EQR) для решения многосвязной задачи выпуклой оптимизации. Его использование позволяет за полиноминальное время найти оптимальное расписание конвейерной задачи.

Литература

1. Coffman, E. G. Computer and Job-shop Scheduling Theory [Text] / E. G. Coffman. – John Wiley&Sons, 1976. – 312 p.
2. Conway, R. W. Theory of Scheduling [Text] / R. W. Conway, W. L. Maxwell, L. W. Miller. – Addison-Wesley Educational Publishers Inc., 1967. – 294 p.
3. Lazarev, A. A. Teoriia raspisanii zadachi i algoritmy [Text] / A. A. Lazarev, E. R. Gafarov. – Moscow: Nauka, 2011. – 222 p.
4. Wang, J.-B. Flow shop scheduling with deteriorating jobs under dominating machines to minimize makespan [Text] / J.-B. Wang // The International Journal of Advanced Manufacturing Technology. – Vol. 48, No. 5–8. – P. 719–723. doi:10.1007/s00170-009-2314-2
5. Prilutsky, M. Kh. Multi-stage tasks of the scheduling theory with alternative choice of the works fulfilment [Text] / M. Kh. Prilutsky, S. Ye. Vlasov // Sistemy upravleniia i informatsionnye tehnologii. – 2005. – No. 2. – P. 44–48.
6. Prilutsky, M. Kh. The method of paths and boundaries with heuristic evaluation for a conveyor problem of scheduling theory [Text] / M. Kh. Prilutsky, V. S. Vlasov // Vestnik of Lobachevsky State University of Nizhni Novgorod. – 2008. – No. 3. – P. 147–153.
7. Emmons, H. Flow Shop Scheduling [Text] / H. Emmons, G. Vairaktarakis // International Series in Operations Research&Management Science. – Springer US, 2013. – 334 p. doi:10.1007/978-1-4614-5152-5

8. Anichkin, A. S. A survey of emerging models and methods of scheduling [Text] / A. S. Anichkin, V. A. Semenov // Proceedings of the Institute for System Programming of the RAS. – 2014. – Vol. 26, No. 3. – P. 5–50.

9. Kosolap, A. I. Metody global'noi optimizatsii [Text] / A. I. Kosolap. – Dnipro: Nauka i obrazovanie, 2013. – 316 p.

10. Boyd, S. Convex Optimization [Text] / S. Boyd, L. Vandenberghe. – Cambridge University Press, 2004. – 716 p. doi:10.1017/cbo9780511804441

11. Nesterov, Yu. E. Metody vypukloi optimizatsii [Text] / Yu. E. Nesterov. – Moscow: MTsNMO, 2010. – 281 p.

12. Kosolap, A. Method of quadratic relaxation for solution of nonconvex quadratic program [Text] / A. Kosolap. – Remagen, Germany, 2009. – 20 p.