

Комплексный метод определения эффективных решений по составам партий данных и расписаниям их обработки в конвейерных системах

К. В. КРОТОВ

Севастопольский государственный университет, Россия

Контактный e-mail: krotov_k1@mail.ru

Построение комплексных расписаний обработки партий данных предполагает формирование взаимосвязанных решений по составам партий и расписаниям их обработки. В результате выполненной декомпозиции обобщенной задачи построения расписаний обработки партий ее решение представляется в виде решений двух подзадач — определения составов партий и непосредственно расписаний. В рассмотрение введена модель двухуровневого программирования для определения эффективных решений по составам партий и расписаниям их обработки. Для нахождения эффективных решений выполнено обоснование методов определения составов партий и расписаний. Проведены исследования особенностей обработки партий в конвейерных системах с точки зрения входных параметров, характеризующих обрабатываемые данные и конвейерную систему.

Ключевые слова: расписание обработки партий, конвейерная система, модель двухуровневой игры, метод определения составов партий.

Введение

При реализации современных информационных технологий возникает необходимость обработки больших объемов разнотипных данных. Такой информационной технологией является мониторинг окружающей среды с использованием данных дистанционного зондирования (ДДЗ) Земли. Задачей мониторинга окружающей среды является идентификация различных видов природных явлений и загрязнений техногенного характера, определение условий их развития. Решение этой задачи реализуется путем создания Интернет-ресурсов, предоставляющих пользователям услуги по тематической обработке спутниковых данных. Поскольку в буфере приема центра обработки, реализующего соответствующие услуги, накапливаются данные от различных спутников, системой обрабатываются данные разных типов, на обработку поступает не единственный спутниковый снимок определенного типа, а их набор, т. е. в буфере хранения данных Интернет-сервиса одновременно находятся несколько наборов данных разных типов. В силу больших объемов обрабатываемых данных, наличия ограничений на время получения результатов требуется выполнять тематическую обработку в составе высокопроизводительных вычислительных систем. В связи с этим является актуальной задача управления обработкой данных в этих системах и разработки способов его реализации.

Работы по созданию систем мониторинга окружающей среды с использованием ДДЗ ведутся в различных научных организациях (ИКИ РАН, ИВТ СО РАН, ИАПУ ДВО РАН). Результаты работ по реализации таких систем изложены в [1–9]. Управление процессом обработки данных в системах, рассмотренных в [1–3], предусматривает, что диспетчер заданий идентифицирует наличие свободных ресурсов (вычислительного ресурса, ресурса оперативной памяти (ОП) и т. д.) для каждого типа данных из очереди, формирует сценарии обработки, представляющие собой последовательности запуска требуемых процедур, передает сценарии в виде файлов конфигурации на определенные вычислительные устройства (ВУ). На каждом из устройств данные определенных типов обрабатываются независимо.

В работах [4, 5] обобщены введенные в [1–3] принципы построения рассматриваемых систем и способы управления обработкой данных в них. Для каждого типа данных выделяются: сервер-диспетчер, формирующий запросы на получение данных от центров их хранения и предоставления, сервер-хранилище данных, реализующий хранение и тематическую обработку снимков определенного типа в соответствии с заданным для этого устройства сценарием. При освобождении сервера-хранилища данных он будет выделяться для обработки новых данных соответствующего типа.

В работе [6] рассмотрены модификации базовой технологии обработки, представленной в [1–5]. Модификация предусматривает, что при поступлении в систему новых данных диспетчер заданий не ожидает освобождения ВУ, а определяет на них наличие требуемого для обработки количества ресурсов. В случае идентификации такого устройства ему передаются данные вместе со сценарием обработки. Управление процессом обработки предусматривает не только идентификацию ВУ с требуемым количеством ресурсов, но и реализацию равномерного распределения данных по устройствам (управляющая процедура предотвращает загрузку всех ВУ обработкой данных одного типа).

В соответствии с [7] планирование выполнения заданий, находящихся в очереди, предусматривает использование освободившихся устройств, подходящих (с точки зрения наличия в ОП требуемых процедур) для интерпретации соответствующего сценария. Планирование выполнения заданий в системе, рассмотренной в [8], предусматривает использование системы управления обработкой Taverna. В ней процесс обработки представляется в виде графа потока управления. Выполнение различных рабочих процессов, связанных с обработкой данных, распределяется по вычислительным устройствам, в ОП которых имеются требуемые процедуры. При этом рабочему процессу для реализации обработки выделяется тот ресурс, который менее всего загружен выполнением других рабочих процессов. Подход к управлению процессом обработки данных, описанный в [9], аналогичен способу управления, рассмотренному в [1–3, 6, 7].

Развитие систем обработки спутниковых данных за рубежом описано в [10, 11]. Работа [10] посвящена организации распределенной обработки данных средствами Grid-технологии. Брокер ресурсов, получив информацию о наличии свободных ресурсов, направляет на них как обрабатываемые данные, так и сами процедуры и сценарии обработки. Особенность — использование распределенных вычислений для обработки данных только одного типа (т. е. управление обработкой связано с распределением определенного количества снимков между ресурсами, выполняющими однотипную обработку) [11].

Анализ способов управления обработкой показывает, что при наличии требуемого количества ресурсов на ВУ для данных определенного типа диспетчирующее устрой-

ство повторно формирует сценарий обработки, который вновь передается на него для интерпретации. Это приводит к повторной активизации обрабатывающих процедур (загрузка в ОП, запуск на выполнение и т. д.). Исключить указанные недостатки позволяет подход, предусматривающий объединение однотипных данных в группы, называемые партиями. Партия данных — это совокупность данных одного типа, обрабатываемых в вычислительной системе без ее переналадки на обработку данных другого типа. Таким образом, формирование партий позволяет исключить повторяющееся построение сценариев обработки для однотипных данных и передачу их на вычислительные устройства.

Поскольку методы стандартной тематической обработки спутниковых данных предполагают ее реализацию в виде последовательности этапов (стадий), для снижения временных затрат на выполнение эта обработка должна быть конвейеризирована. В том случае, когда в конвейерной системе выполняются операции с наборами спутниковых данных разных типов, возникает задача управления процессом обработки. Задача управления процессом обработки является комплексной и предполагает решение двух подзадач: определения составов партий данных и построения расписаний их обработки.

Анализ современных методов решения задач определения составов партий и расписаний их обработки представлен в работе [12]. Для решения задач определения эффективных составов партий используются подходы на основе частично целочисленного линейного программирования (ЧЦЛП), эвристических процедур, а также алгоритмы локальной оптимизации для задач в упрощенной постановке (один обрабатывающий прибор, директивные сроки обработки партий). Однако упомянутые методы обладают рядом недостатков:

- 1) для аппарата ЧЦЛП — при большой размерности задачи получение решения затруднительно, решения по составам партий формируются без учета расписаний обработки;
- 2) для методов формирования партий с учетом директивных сроков обработки на одном приборе — решения по составам партий не являются эффективными, а только удовлетворяют директивным срокам;
- 3) для эвристических процедур — методы не позволяют получить эффективные решения, приближающиеся к оптимальным.

Исходя из сказанного разработка методов определения эффективных составов партий и расписаний их обработки является актуальной задачей.

1. Постановка цели и задач научного исследования

Цель работы состоит в совершенствовании методов решения комплексной задачи определения составов партий и расписаний их обработки в конвейерных системах. Достижение цели обеспечивается декомпозицией обобщенной функции системы на совокупность иерархически упорядоченных подфункций. На основании предложенного иерархического подхода достижение цели обеспечивается решением задач обоснования: а) модели процесса обработки партий данных в конвейерной системе, вида критериев оптимизации решений; б) метода поиска локально оптимальных решений по составам партий данных; в) метода поиска локально оптимальных решений по расписаниям обработки партий.

2. Обоснование модели двухуровневого программирования для формирования составов партий данных и расписаний их обработки в конвейерных системах

Под конвейеризацией выполнения программы подразумевается разделение ее на фрагменты, каждый из которых закреплен за соответствующим сегментом конвейера. Маршруты обработки данных всех типов одинаковы, строго фиксированы и предполагают прохождение ими всех сегментов конвейера. Введем обозначения: l — индекс сегмента конвейера ($l = \overline{1, L}$); n — количество типов данных, обрабатываемых в системе; i — идентификатор типа данных ($i = \overline{1, n}$). Количество элементов в наборе данных, характеризуемых индексом i , обозначим как n^i . Данные i -го типа ($i = \overline{1, n}$) обрабатываются соответствующей программой. В системе используются n типов программ, обрабатывающих данные i -х типов.

Для формирования решений по составам партий данных введем обозначения: m_i — количество партий данных i -го типа ($i = \overline{1, n}$), формируемых на первом уровне принятия решений; элементы m_i образуют вектор \mathbf{M} , соответствующий количеству партий данных n типов; A — матрица, элемент a_{ih} которой — это количество данных i -го типа в h -й партии ($h = \overline{1, m_i}$). Решение, формируемое на верхнем уровне иерархии системы, имеет вид $[\mathbf{M}, A]$, где \mathbf{M} — вектор количества партий данных i -х типов ($i = \overline{1, n}$), A — матрица составов партий.

В соответствии с решением по составам партий требуется определить порядки (последовательности) их обработки на сегментах конвейера, т.е. расписание обработки партий. Расписание обработки партий обозначим как π , оно представляет собой набор (множество) последовательностей π^l запуска партий на обработку на l -х сегментах конвейера ($l = \overline{1, L}$). Расписание π формируется в предположении, что порядок обработки партий одинаков на всех сегментах конвейера.

Для формализации вида последовательностей π^l расписания π в рассмотрение введена матрица порядка обработки партий P . Элемент $p_{ij} = 1$, если партия данных i -го типа в последовательности π^l занимает j -ю позицию, $p_{ij} = 0$ в случае, если партия данных i -го типа не занимает в последовательности π^l j -ю позицию, размерность матрицы $n \times n_p$, где n_p — количество партий в последовательностях π^l ($n_p = \sum_{i=1}^n m_i$). Порядок обработки партий на всех сегментах одинаков, поэтому достаточно определения одной матрицы порядка P . В рассмотрение введена матрица R — матрица количества данных i -х типов в партиях, занимающих в последовательностях π^l j -е позиции (r_{ij} — количество данных i -го типа в партии, занимающей j -ю позицию в π^l). Тогда решение, формируемое на нижнем уровне иерархии в системе, имеет вид $[P, R]$.

Для формализации вида двухуровневой модели принятия решений по составам партий и расписаниям их обработки в конвейерной системе введены следующие обозначения [13]: (t_i^l) — длительность обработки данных i -го типа на l -м сегменте конвейера ($l = \overline{1, L}$); (t_{jq}^{0l}) — матрица моментов времени начала обработки q -х данных в партиях, занимающих в π^l j -е позиции (q — порядковый номер данных в партии в j -й позиции в π^l , $q = \overline{1, n_j}$, $n_j = \sum_{h=1}^n r_{hj}$, где n_j — количество данных в партии, занимающей j -ю позицию в π^l). Значения t_{jq}^{0l} определяются в соответствии с выражениями модели вычислительного процесса обработки партий данных, рассмотренной в [13]. Аналогичным образом в [13] представлены рассуждения, позволившие сформировать вид критерия

на нижнем уровне принятия решений, используемого при определении эффективных расписаний обработки партий.

Если n_p — количество партий, сформированных на верхнем уровне иерархии (индекс последней сформированной партии), то n_{n_p} — количество данных, входящих в n_p -ю партию (индекс последних данных в n_p -й партии), $t_{n_p, n_{n_p}}^{0L}$ — момент времени начала обработки последних данных в партии с индексом n_p на L -м сегменте конвейера. Тогда момент времени окончания обработки этой партии на L -м сегменте и момент времени окончания обработки всех данных в системе определяются выражением вида

$$t_{n_p, n_{n_p}}^{0L} + \sum_{h=1}^n t_h^L p_{h, n_p}.$$

Двухуровневая модель определения эффективных составов партий и расписаний их обработки (с учетом [13]) имеет первый (верхний) уровень $\min f_1$ и второй (нижний) уровень $\min f_2$, где

$$\begin{aligned} f_1 &= t_{n_p, n_{n_p}}^{0L} + \sum_{h=1}^n t_h^L p_{h, n_p}, \\ f_2 &= \sum_{l=2}^L t_{11}^{0l} + \sum_{l=1}^L \sum_{j=2}^{n_p} \left[t_{j1}^{0l} - \left(t_{j-1, n_{j-1}}^{0l} + \sum_{h=1}^n t_h^l p_{h, j-1} \right) \right] + \\ &+ \sum_{l=2}^L \sum_{j=1}^{n_p} \sum_{q=2}^{n_j} \left[t_{jq}^{0l} - \left(t_{j, q-1}^{0l} + \sum_{h=1}^n t_h^l p_{h, j} \right) \right]. \end{aligned} \quad (1)$$

3. Обоснование метода определения эффективных составов партий данных

Для обоснования метода формирования локально оптимальных решений по составам партий данных i -х типов ($i = \overline{1, n}$) в рассмотрение введены следующие обозначения: i' — идентификатор типа данных, составы партий которого изменяются на текущем шаге алгоритма; $m_{i'}$ — количество партий данных i' -го типа, составы которых определяются; s — индекс текущего улучшаемого решения путем поиска более эффективного решения в рамках окрестности O_k ; g — индекс шага алгоритма, выполняемого по отношению к шагу s , который соответствует новому формируемому решению, находящемуся в окрестности O_k ($(s+g)$ — номер шага алгоритма по формированию решения в окрестности O_k текущего локально оптимального решения); h' — индекс партии, состав которой изменяется при реализации алгоритма на текущей итерации; I — множество типов данных, для которых выполняется формирование составов партий (первоначально $I = \{1, 2, \dots, n\}$); I' — множество типов данных, используемое при формировании локально оптимального решения по составам партий; A' — матрица (аналог матрицы составов партий A), используемая при определении наилучшего решения в окрестности O_k текущего локально оптимального решения (матрица A' используется при формировании промежуточных решений); $A_{1i}^{6y\Phi}$ и $A_{2i}^{6y\Phi}$ ($i = \overline{1, n}$) — матрицы, предназначенные для хранения (буферизации) составов партий данных i -х типов, сформированных на различных итерациях алгоритма (в матрицах $A_{1i}^{6y\Phi}$ ($i = \overline{1, n}$) хранятся решения по составам партий данных i -х типов, входящие в окрестность O_k с меньшей метрикой, в матрицах $A_{2i}^{6y\Phi}$ ($i = \overline{1, n}$) хранятся решения по составам партий данных i -х типов, входящие

в окрестность O_{k+1} , сформированные на основе решений из окрестности O_k , т. е. на основе решений из матриц $A_{1i}^{\text{бy}\Phi}$ ($i = \overline{1, n}$); n_{p1}^i — количество решений по составам партий данных i -го типа в матрице $A_{1i}^{\text{бy}\Phi}$ размерностью $n_{p1}^i \times m_i$; n_{p2}^i — количество решений по составам партий данных i -го типа в матрице $A_{2i}^{\text{бy}\Phi}$ размерностью $n_{p2}^i \times m_i$; $q1_i$ — индекс текущей строки в матрице $A_{1i}^{\text{бy}\Phi}$ ($q1_i = \overline{1, n_{p1}^i}$) (рассматриваемого решения по составам партий i -го типа в окрестности O_k); $q2_i$ — индекс текущей строки в матрице $A_{2i}^{\text{бy}\Phi}$ ($q2_i = \overline{1, n_{p2}^i}$) (рассматриваемого решения по составам партий i -го типа в окрестности O_{k+1}); $q2'_i$ — параметр, предназначенный для хранения номера строки в $A_{2i}^{\text{бy}\Phi}$ — номера решения, гарантирующего максимальное по модулю значение левого дискретного градиента критерия f_1 [4]; G — максимальное по модулю значение левого дискретного градиента $-\nabla_g f_1$, достигаемое в окрестностях различного вида текущего локально оптимального решения (максимального по модулю значения $-\nabla_g f_1$).

Для текущего локально оптимального решения по составам партий данных i -х типов рассматриваются два вида окрестностей. Построение решений, включаемых в окрестность первого вида, связано с реализацией: а) изменения составов партий данных каждого i -го типа в заданном количестве m_i ($i = \overline{1, n}$) при неизменных составах партий данных других типов; б) увеличения количества m_i партий данных i -го типа в случае, если построение решений путем изменения составов партий данных этого типа в количестве m_i невозможно (выполнены условия окончания формирования составов партий в количестве m_i). Тогда реализуются увеличение количества m_i партий данных этого типа и задание начального решения для этого количества партий. Построение решений, включаемых в окрестность второго вида, предполагает совместное использование сформированных решений по составам партий данных всех типов при заданном количестве партий m_i для них.

Для обоснования алгоритма определения решений по составам партий сформулирован способ изменения составов партий данных одного i' -го типа, в котором используются следующие условия и рассуждения (для i -го типа данных):

- количество данных i -го типа в партиях не может быть менее 2 ($a_{ih} \geq 2$, $h = \overline{1, m_i}$); если при формировании начального решения по составам m_i партий данных i -го типа для $h = 1$ (первая партия) получено $a_{i1} < 2$, то исследование решений по составам партий данных этого i -го типа не выполняется;
- способ формирования начального решения для количества m_i партий данных i -го типа предполагает, что $a_{ih} = 2$ ($h = \overline{2, m_i}$), а элемент a_{i1} определяется как $a_{i1} = n^i - \sum_{h=2}^{m_i} a_{ih}$; в дальнейшем при формировании составов партий данные извлекаются из партии с $h = 1$ и распределяются по остальным партиям ($h = \overline{2, m_i}$);
- значения параметров m_i , задаваемые первоначально для данных всех i -х типов ($i = \overline{1, n}$), равны 2 ($m_i = 2$);
- модификация количества партий i' -го типа данных предполагает, что параметр $m_{i'}$ увеличивается (при определении составов партий) до тех пор, пока в начальном решении для $m_{i'}$ выполняется условие $a_{i'h} \geq 2$ ($h = \overline{1, m_{i'}}$); при $a_{i'1} < 2$ (для начального решения при значении $m_{i'}$) формирование составов партий i' -го типа прекращается;
- формирование решений по составам партий предполагает увеличение количества данных в партии с индексом $h' > 1$ (при неизменном составе h -х партий ($h = \overline{2, h' - 1}$) & ($h = \overline{h' + 1, m_i}$)) и уменьшение количества данных в партии с $h = 1$

$\left(a_{i'1} = n^{i'} - \sum_{h=2}^{m_{i'}} a_{i'h} \right)$; при выполнении условия $a_{i'1} \geq a_{i'h}$ ($h = \overline{2, m_{i'}}$) модификация составов партий продолжается; при выполнении $a_{i'1} < a_{i'h}$ полученное решение по составам партий в количестве $m_{i'}$ не рассматривается, так как является аналогичным полученному на предыдущих шагах алгоритма.

Способ модификации составов партий i' -го типа предусматривает, что на основе решений, находящихся в окрестности первого вида текущего локально оптимального решения с метрикой k (окрестность O_k), формируются решения, находящиеся в окрестности с большей метрикой ($k + 1$) (окрестность O_{k+1}). Метрика k окрестности O_k , содержащей решения по составам партий i' -го типа, определяется выражением

$$k = \sum_{i'=1}^n \sum_{h=2}^{m_{i'}} [a'_{i',h}(s+g) - a_{i',h}(s)],$$

где $a_{i',h}(s)$ — элемент i' -й строки матрицы $A(s)$, соответствующей текущему локально оптимальному решению по составам партий данных; $a'_{i',h}(s+g)$ — элемент i' -й строки матрицы $A'(s+g)$, соответствующей решению по составам партий в окрестности O_k .

В матрице $A_{1i}^{\text{бyf}}$ выполняется хранение решений, находящихся в окрестности первого вида O_k , на основе которых реализуется формирование решений, входящих в окрестность O_{k+1} (в матрицах $A_{1i}^{\text{бyf}}$ хранятся решения, используемые при формировании решений в окрестности O_{k+1}). В матрицах $A_{2i}^{\text{бyf}}$ реализуется хранение сформированных решений, находящихся в окрестности O_{k+1} . Тогда на основе одного решения по составам партий данных ($q1_i$ -й строки матрицы $A_{1i}^{\text{бyf}}$) формируется совокупность решений, хранимых в матрице $A_{2i}^{\text{бyf}}$, находящихся в окрестности O_{k+1} . Если для текущего рассматриваемого решения ($q1_i$ -й строки матрицы $A_{1i}^{\text{бyf}}$) выполняется условие $(a_{q1_i,h})_{1i}^{\text{бyf}} = (a_{q1_i,h+1})_{1i}^{\text{бyf}}$, то увеличение на 1 значения $(a_{q1_i,h})_{1i}^{\text{бyf}}$ (при неизменном значении $(a_{q1_i,h+1})_{1i}^{\text{бyf}}$), а затем увеличение на 1 значения $(a_{q1_i,h+1})_{1i}^{\text{бyf}}$ (при неизменном значении $(a_{q1_i,h})_{1i}^{\text{бyf}}$) обуславливают получение одинаковых решений по составам партий данных i -го типа.

Таким образом, формулировка условия, ограничивающего количество решений в окрестности O_{k+1} , имеет следующий вид: если при формировании нового решения по составам $m_{i'}$ партий данных i' -го типа на основе решения, представленного строкой $(A_{q1_{i'}})_{1i'}^{\text{бyf}}$ матрицы $A_{1i'}^{\text{бyf}}$, для $q1_{i'}$ -й строки выполняется условие $(a_{q1_{i'},h})_{1i'}^{\text{бyf}} = (a_{q1_{i'},h+1})_{1i'}^{\text{бyf}}$, то не требуется увеличивать значение элемента $(a_{q1_{i'},h+1})_{1i'}^{\text{бyf}}$. Должен быть осуществлен переход к элементу $(a_{q1_{i'},h+2})_{1i'}^{\text{бyf}}$, который увеличивается в случае выполнения условия $(a_{q1_{i'},h})_{1i'}^{\text{бyf}} \neq (a_{q1_{i'},h+2})_{1i'}^{\text{бyf}}$.

Для определения локально оптимального решения по составам партий данных i -х типов ($i = \overline{1, n}$) сформулированы:

- способ формирования решений по составам партий данных одного i' -го типа, включаемых в окрестности O_{k+1} первого вида, на основе решений, входящих в окрестность O_k (изменение составов партий в количестве m_i);
- способ формирования решений, представляющих собой комбинацию решений по составам партий данных всех i -х типов ($i = \overline{1, n}$), включаемых в окрестности O_{k+1} второго вида;
- обобщенный алгоритм определения локально оптимального решения по составам партий данных.

Входными параметрами алгоритма формирования решений в окрестности O_{k+1} первого вида, получаемых на основе решений из окрестности O_k , являются: идентификатор i' -го типа данных, для которого определяются составы партий; количество $n_{p1}^{i'}$ решений в окрестности O_k ; количество партий $m_{i'}$, составы которых изменяются; матрица $A_{1i'}^{\text{буф}}$, содержащая решения в окрестности O_k .

Алгоритм формирования решений по составам партий данных одного i' -го типа в окрестности O_{k+1} первого вида на основе решений из окрестности O_k имеет следующий порядок шагов.

1. Инициализация значений параметров: $q1_{i'} = 1, n_{p2}^{i'} = 0, q2_{i'} = 0, q2'_{i'} = 0, g = 1$.

2. Инициализация номера партии h' , состав которой изменяется, значением 2 ($h' = 2$).

3. Изменение $q2_{i'} = q2_{i'} + 1$; в партии с индексом h' — увеличение количества данных на 1; построение решения по составам партий i' -го типа ($q2_{i'}$ -й строки матрицы $A_{2i'}^{\text{буф}}$): $(a_{q2_{i'},h})_{2i'}^{\text{буф}} = (a_{q1_{i'},h})_{1i'}^{\text{буф}}$ ($h = 2, h' - 1 \& h = h' + 1, m_{i'}$), $(a_{q2_{i'},h'})_2^{\text{буф}} = (a_{q1_{i'},h'})_1^{\text{буф}} + 1$, $(a_{q2_{i'},1})_{2i'}^{\text{буф}} = n_{i'} - \sum_{h=1}^{m_{i'}} (a_{q1_{i'},h'})_{1i'}^{\text{буф}}$; $n_{p2}^{i'} = n_{p2}^{i'} + 1$.

4. Проверка условия $(a_{q2_{i'},1})_{2i'}^{\text{буф}} < (a_{q2_{i'},h'})_{2i'}^{\text{буф}}$. В случае его выполнения сформированное в виде $q2_{i'}$ -й строки $(A_{q2_{i'}})_{2i'}^{\text{буф}}$ матрицы $A_{2i'}^{\text{буф}}$ решение не интерпретируется (не включается в окрестность O_{k+1}), тогда $n_{p2}^{i'} = n_{p2}^{i'} - 1$ и $q2_{i'} = q2_{i'} - 1$ (составы партий i' -го типа в количестве $m_{i'}$ не модифицируются). При условии $(a_{q2_{i'},1})_{2i'}^{\text{буф}} \geq (a_{q2_{i'},h'})_{2i'}^{\text{буф}}$ полученное решение интерпретируется (включается в окрестность O_{k+1}), выполняется задание значения параметра $j = 1$ (j — значение шага изменения номера партии h'), реализуется переход на шаг 5.

5. Проверка условия $(h' + j) \leq m_{i'}$. При $(h' + j) \leq m_{i'}$ реализуется проверка условия формирования решения по составам партий i' -го типа в виде $(a_{q1_{i'},h'})_{1i'}^{\text{буф}} > (a_{q1_{i'},h'+j})_{1i'}^{\text{буф}}$. Если условие $(a_{q1_{i'},h'})_{1i'}^{\text{буф}} > (a_{q1_{i'},h'+j})_{1i'}^{\text{буф}}$ не выполняется, то реализуется переход на шаг 6. При выполнении условия существует возможность формирования нового решения по составам партий данных i' -го типа, тогда $h' = h' + j$, реализуется переход на шаг 3 (переход к формированию следующего решения в окрестности O_{k+1}). При $(h' + j) > m_{i'}$ — переход на шаг 7.

6. В случае выполнения условия $(a_{q1_{i'},h'})_{1i'}^{\text{буф}} = (a_{q1_{i'},h'+j})_{1i'}^{\text{буф}}$ при модификации $(h' + j)$ -й партии будет получено решение, дублирующее сформированное ранее. Осуществляется модификация параметра $j = j + 1$. Выполняется переход на шаг 5.

7. При $(h' + j) > m_{i'}$ сформированы все решения (хранящиеся в матрице $A_{2i'}^{\text{буф}}$) с использованием одного решения, хранящегося в $q1_{i'}$ -й строке матрицы $A_{1i'}^{\text{буф}}$. Переход на шаг 8.

8. Модификация значения $q1_{i'}$ — индекса решения, хранимого в матрице $A_{1i'}^{\text{буф}}$ (переход к следующему решению из окрестности O_k , на основании которого формируются новые решения): $q1_{i'} = q1_{i'} + 1$. При $q1_{i'} \leq n_{p1}^{i'}$ (рассмотрены не все решения из окрестности O_k , хранимые в матрице $A_{1i'}^{\text{буф}}$) реализуется переход на шаг 2. При условии $q1_{i'} > n_{p1}^{i'}$ все решения, хранящиеся в матрице $A_{1i'}^{\text{буф}}$, использованы при формировании решений в окрестности O_{k+1} . Выполняется переход на шаг 9.

9. При $n_{p2}^{i'} > 1$ выполняется сравнение решений, хранимых в матрице $A_{2i'}^{\text{буф}}$, с точки зрения дублирования ими друг друга. Процедура сравнения предполагает формирование копии матрицы $A_{2i'}^{\text{буф}}$. В копии матрицы $A_{2i'}^{\text{буф}}$ производятся упорядочивание эле-

ментов в каждой $q2_{i'}$ -й строке ($q2_{i'} = \overline{1, n_{p2}^{i'}}$) по убыванию их значений, поэлементное сравнение каждой $q2_{i'}$ -й строки ($q2_{i'} = \overline{1, n_{p2}^{i'} - 1}$) с другими строками, удаление из матрицы $A_{2i'}^{\text{буф}}$ строк, дублирующих рассматриваемую $q2_{i'}$ -ю строку. При удалении строки из матрицы $A_{2i'}^{\text{буф}}$ реализуются изменение $n_{p2}^{i'} = n_{p2}^{i'} - 1$ и переупорядочивание решений в матрице $A_{2i'}^{\text{буф}}$. В результате повторяющиеся решения из матрицы $A_{2i'}^{\text{буф}}$ удалены; задается значение $q2_{i'} = 1$; при $n_{p2}^{i'} > 0$ осуществляется переход на шаг 11.

10. При $n_{p2}^{i'} = 0$ не были сформированы новые решения в окрестности O_{k+1} . Выполняется переход на шаг 16.

11. Реализуется исследование решения по составам партий данных i' -го типа в окрестности O_{k+1} с точки зрения его эффективности. Для этого выполняется инициализация i' -й строки матрицы A' решением по составам партий данных, хранимым в $q2_{i'}$ -й строке $(A_{q2_{i'}})_{2i'}^{\text{буф}}$ матрицы $A_{2i'}^{\text{буф}}$: $(a_{i',h})' = (a_{q2_{i'},h})_{2i'}^{\text{буф}}$ ($h = \overline{1, m_{i'}}$). Выполняется передача сформированного решения $[\mathbf{M}(s+g), A'(s+g)]$ на второй уровень для определения эффективного расписания обработки партий (при этом исходное решение $A(s)$ не изменяется).

12. Использование полученного со второго уровня локально оптимального решения вида $[P, (t_{jq}^0) | l = \overline{1, L}]^*$ для вычисления значения $f_1(s+g)$.

13. Для решения $[\mathbf{M}(s+g), A'(s+g)]$ вычисляются левый (правый) дискретный градиент $-\nabla_g f_1$ ($+\nabla_g f_1$). Вычисление $-\nabla_g f_1(s)$ и $+\nabla_g f_1(s)$ выполняется по формулам [14]:

$$\text{а) } -\nabla_g f_1(s) = [f_1(s+g) - f_1(s)] \leq 0; \quad \text{б) } +\nabla_g f_1(s) = [f_1(s+g) - f_1(s)] > 0.$$

14. При $+\nabla_g f_1(s) > 0$ решение $[\mathbf{M}(s+g), A'(s+g)]$ не является более эффективным, чем решение $[\mathbf{M}(s), A(s)]^*$. Осуществляется модификация значения $q2_{i'} = q2_{i'} + 1$. Если $q2_{i'} \leq n_{p2}^{i'}$ (рассмотрены не все решения из окрестности O_{k+1}), то выполняется модификация индекса шага алгоритма $g = g + 1$, реализуется переход на шаг 11 (переход к анализу следующего решения из O_{k+1}). При $q2_{i'} > n_{p2}^{i'}$ выполняется переход на шаг 16.

15. Если для решения по составам партий ($q2_{i'}$ -й строки $(A_{q2_{i'}})_{2i'}^{\text{буф}}$ матрицы $A_{2i'}^{\text{буф}}$) реализуется $-\nabla_g f_1(s) \leq 0$, то выполняется сравнение значений $-\nabla_g f_1(s)$ и G — градиента целевой функции (являющегося “глобальным” для всех типов данных). Если $-\nabla_g f_1(s) < G$, то текущее решение по составу партий i' -го типа данных является наилучшим среди всех рассматриваемых решений по составам партий данных n типов в окрестностях O_{k+1} первого вида. При $-\nabla_g f_1(s) < G$ значения всех параметров $q2'_i$ ($i \neq i'$) обнуляются, индекс $q2_{i'}$ решения ($q2_{i'}$ -й строки $(A_{q2_{i'}})_{2i'}^{\text{буф}}$ матрицы $A_{2i'}^{\text{буф}}$) сохраняется: $q2'_{i'} = q2_{i'}$, модификация значения G : $G = -\nabla_g f_1(s)$. Реализуется переход к следующему решению в окрестности O_{k+1} , для этого $q2_{i'} = q2_{i'} + 1$. Если $q2_{i'} \leq n_{p2}^{i'}$, то $g = g + 1$, реализуется переход на шаг 11. Если $q2_{i'} > n_{p2}^{i'}$, то выполняется переход на шаг 16.

16. Если $q2'_{i'} \neq 0$, то в окрестности O_{k+1} определено решение по составам партий i' -го типа, лучшее, чем текущее оптимальное решение $[\mathbf{M}(s), A(s)]^*$. Осуществляется останов алгоритма определения решений по составам партий данных i' -го типа в окрестности O_{k+1} .

Выходные параметры алгоритма формирования составов партий одного i' -го типа в окрестности O_{k+1} : $n_{p2}^{i'}$ — количество сформированных решений по составам партий, входящих в O_{k+1} ; значение $q2'_{i'}$ идентификатора строки в матрице $A_{2i'}^{\text{буф}}$, соответствующее наилучшему решению по составам партий данных i' -го типа в окрестности O_{k+1} ; матрица $A_{2i'}^{\text{буф}}$, содержащая (при $n_{p2}^{i'} > 0$) решения, входящие в окрестность O_{k+1} . При $n_{p2}^{i'} > 0$ и $q2'_{i'} = 0$ в окрестности O_{k+1} отсутствует решение, являющееся лучшим, чем текущее локально оптимальное, необходимо формирование решений в окрест-

ности O_{k+2} первого вида с большей метрикой. Рассмотренный способ определения решений, входящих в окрестность O_{k+1} , последовательно применяется ко всем типам данных ($i = \overline{1, n}$). Он предполагает (для i -го типа, $i = \overline{1, n}$) формирование на основе решений из матриц $A_{1i}^{\text{буф}}$ новых решений, сохраняемых в матрицах $A_{2i}^{\text{буф}}$. Тогда в окрестность O_{k+1} первого вида входят решения по составам партий i -х типов ($i = \overline{1, n}$), сформированные путем изменения составов партий, сохраненных в матрицах $A_{1i}^{\text{буф}}$.

В том случае, если в окрестности O_{k+1} первого вида не найдено решение, лучшее, чем локально оптимальное решение $[\mathbf{M}(s), A(s)]^*$, то поиск выполняется в окрестности O_{k+1} второго вида.

Решения, входящие в окрестность второго вида, формируются путем совместного рассмотрения сформированных решений по составам партий i -х типов, сохраненных в матрицах $A_{2i}^{\text{буф}}$ ($i = \overline{1, n}$). Формулировка способа формирования решений по составам партий данных, входящих в окрестность O_{k+1} второго вида, предваряется введением обозначений: G_p, G_{p1} — множества идентификаторов типов данных, для которых при формировании решений в окрестности O_{k+1} первого вида получено $n_{p2}^i > 0$, элементы этих множеств упорядочены по возрастанию, G'_p — множество типов данных, для которых при формировании решений в окрестности O_{k+1} первого вида получено $n_{2p}^i > 0$, элементы этого множества упорядочиваются по убыванию.

Алгоритм формирования решений по составам партий, входящих в окрестность O_{k+1} второго вида для решения $[\mathbf{M}(s), A(s)]^*$, реализует следующие шаги.

1. Формирование множеств G_p и G_{p1} в виде $G_p = \{i_j, \dots, i_k, \dots\}$, $G_{p1} = \{i_j, \dots, i_k, \dots\}$, при этом типы данных $i \in G_p$ и $i \in G_{p1}$ такие, что $n_{p2}^i > 0$, $i_j < i_k$, где j, k — порядковые номера идентификаторов типов данных в множествах G_p и G_{p1} ; формирование множества G'_p : $G'_p = \{\dots, i_k, \dots, i_j\}$ при $n_{p2}^i > 0$, $i_j \leq i_k$.

2. Если $G_p \neq \emptyset$, то извлечение из G_p идентификатора i' -го типа данных осуществляется следующим образом: $i' = \min_j (i_j | i_j \in G_p)$; модификация множества G_p : $G_p = G_p \setminus \{i'\}$; задание значения параметра $q_{2i'} = 1$ для рассматриваемого i' -го типа данных; для i' -го типа данных определение текущего решения по составам партий (формирование i' -й строки матрицы A'): $(a_{i',h})' = (a_{q_{2i'},h})_{2i'}^{\text{буф}}$ ($h = \overline{1, m_i}$).

3. Проверка условия $G_p \neq \emptyset$, при его выполнении осуществляется переход на шаг 2. При выполнении условия $G_p = \emptyset$ сформирована матрица A' , входящая в решение $[\mathbf{M}, A']$.

4. Передача решения $[\mathbf{M}, A']$ на второй уровень иерархии для формирования на нем решения вида $\left[P, (t_{jq}^{0l}) | l = \overline{1, L} \right]$ по порядку обработки партий; получение со второго уровня иерархии решения $\left[P, (t_{jq}^{0l}) | l = \overline{1, L} \right]^*$, соответствующего решению $[\mathbf{M}, A']$; определение на основе $\left[P, (t_{jq}^{0l}) | l = \overline{1, L} \right]^*$ значения критерия f_1 , соответствующего решению $[\mathbf{M}, A']$; для решения $[\mathbf{M}, A']$ вычисление левого (правого) дискретного градиента $-\nabla_g f_1$ ($+\nabla_g f_1$); при $+\nabla_g f_1(s) > 0$ решение $[\mathbf{M}, A']$ не является более эффективным, чем решение $[\mathbf{M}(s), A(s)]^*$, поэтому оно не интерпретируется, выполняется переход на шаг 5. При реализации $-\nabla_g f_1(s) \leq 0$ сравниваются значения $-\nabla_g f_1(s)$ и G — градиента целевой функции f_1 . Если $-\nabla_g f_1(s) < G$, то текущее решение по составам партий данных i -х типов ($i \in G_p$) — наилучшее среди рассмотренных решений по составам партий данных в окрестностях O_{k+1} второго вида, для всех i -х типов данных ($i \in G_{p1}$) выполняется инициализация параметров q_{2i}' алгоритма: $q_{2i}' = q_{2i}$ (при $i \in G_{p1}$), модификация значения G : $G = -\nabla_g f_1(s)$; реализуется переход на шаг 5.

5. Извлечение из G'_p идентификатора i' -го типа данных, для которого изменяется решение по составам партий: $i' = \max_j(i_j | i_j \in G'_p)$; определение позиции j' (номера) этого типа данных в множестве G_{p1} ; модификация значения параметра $q_{2i'}$ для рассматриваемого i' -го типа данных: $q_{2i'} = q_{2i'} + 1$; если $q_{2i'} \leq n_{p2}^{i'}$, то формируется i' -я строка матрицы A' : $(a_{i',h})' = (a_{q_{2i'},h})_{2i'}^{\text{буф}}$ ($h = \overline{1, m_i}$), выполняется переход на шаг 7.

6. Если $q_{2i'} > n_{p2}^{i'}$, то $G'_p = G'_p \setminus \{i'\}$. Если $G'_p \neq \emptyset$, то реализуется переход на шаг 5. При $G'_p = \emptyset$ выполняется переход на шаг 8.

7. Формирование множеств G_p, G'_p в виде $G_p = \{i_{j'+1}, \dots\}$, $G'_p = \{\dots, i_k, \dots, i_j\}$ (множество G_p включает типы данных, следующие в множестве G_{p1} за текущим рассматриваемым i' -м типом (начиная с $(j'+1)$ -й позиции), если идентификатор рассматриваемого типа данных i' является “последним” в множестве G_{p1} , то $G_p = \emptyset$, вид множества G'_p соответствует первоначальному); выполняется переход на шаг 2.

8. Останов алгоритма.

На основе способов формирования решений по составам партий данных i -х типов ($i = \overline{1, n}$), входящих в окрестности O_{k+1} первого и второго вида, сформулирован алгоритм определения локально оптимальных решений по составам партий. Алгоритм предполагает формирование решений путем: 1) изменения составов партий данных каждого типа при неизменных составах партий других типов, увеличения количества m_i партий данных i -х типов и задания начальных составов этих партий; 2) совместного использования измененных составов партий данных всех n типов и определения окрестности решений второго вида при неизменных значениях параметров m_i .

Перед реализацией алгоритма определения локально оптимальных решений по составам партий данных выполняются инициализация множества I типов данных, для которых формируются составы партий: $I = \{1, 2, 3, \dots, n\}$, инициализация индекса шага алгоритма s и значения метрики окрестности k значением 0.

Алгоритм определения локально оптимального решения по составам партий данных всех i -х типов ($i = \overline{1, n}$) содержит следующий порядок шагов.

1. Задание для каждого i -го типа данных значений параметров n_{p1}^i ($i = \overline{1, n}$) равными 1 ($n_{p1}^i = 1$ — количество решений по составам партий данных i -х типов, сформированных на текущей итерации и хранимых в матрице $A(s)$ (буферизируемых в матрицах $A_{1i}^{\text{буф}}$).

2. Формирование начального решения по составам партий данных i -х типов ($i = \overline{1, n}$):

а) $m_i = 2$ ($i = \overline{1, n}$);

б) инициализация h -х элементов i -х строк матрицы A : $a_{ih} = 2$ ($h = \overline{2, m_i}$),

$$a_{i1} = n^i - \sum_{h=2}^{m_i} a_{ih};$$

в) проверка условия корректности решений по составам партий данных i -х типов ($i = \overline{1, n}$): $a_{i1} \geq 2$ и $a_{i1} \geq a_{ih}$ ($h = \overline{2, m_i}$);

г) если для какого-либо i -го типа данных ($i = \overline{1, n}$) в начальном решении это условие не выполняется, то для данных этого типа формируется фиксированная партия: $m_i = 1$, $a_{i1} = n^i$, $n_{p1}^i = 0$; модификация множества I типов данных: $I = I \setminus \{i\}$; параметр G инициализируется значением 0; инициализация множества I' : $I' = I$; результатом является сформированное начальное решение по составам партий данных $[M(s), A(s)]$.

3. Передача решения $[\mathbf{M}(s), A(s)]$ на второй уровень для формирования расписания обработки партий; получение со второго уровня расписания для текущего решения по составам партий $[\mathbf{M}(s), A(s)]$ в виде $\left[P, (t_{jq}^{0l}) | l = \overline{1, L} \right]^*$; определение значения $f_1(s)$ для решения $[\mathbf{M}(s), A(s)]$.

4. Определение i' -го типа данных, составы партий которого изменяются: $i' = \min\{i | i \in I'\}$, $I' = I \setminus \{i'\}$.

5. Для определения модифицированных составов партий данных i' -го типа (формирования решений в окрестности O_{k+1} текущего локально оптимального решения $[\mathbf{M}, A]^*$) реализуется инициализация параметра $q_{1i'} = 1$ и элементов матрицы $A_{1i'}^{\text{буф}}$ для i' -го типа данных $(a_{q_{1i'}, h})_{1i'}^{\text{буф}} = a_{i'h}$ ($h = \overline{1, m_i}$).

6. Определение решений по составам партий данных i' -го типа, входящих в окрестность O_{k+1} первого вида, формируемых на основе решений, находящихся в матрице $A_{1i'}^{\text{буф}}$. (Для этого вызывается алгоритм формирования составов партий данных i' -го типа, рассмотренный выше, использующий решения, хранящиеся в матрице $A_{1i'}^{\text{буф}}$, для формирования новых решений в окрестности O_{k+1} первого вида.) При выполнении для рассматриваемого i' -го типа данных условия $n_{p2}^{i'} > 0$ реализуется переход на шаг 14.

7. Если для данных i' -го типа после формирования решений в окрестности O_{k+1} первого вида выполняется условие $n_{p2}^{i'} = 0$ (в окрестности O_{k+1} текущего локально оптимального решения $[\mathbf{M}, A]^*$ не получены решения, для которых выполняется условие корректных составов партий), то реализуется модификация количества партий: $m_{i'} = m_i + 1$. Выполняется формирование начального решения по составам партий в количестве $m_{i'}$: осуществляется инициализация значений $n_{p2}^{i'} = 1$; $q_{2i'} = 1$, $q_{2i'}' = 0$, $g = g + 1$, элементов матрицы $A_{2i'}^{\text{буф}}$: $(a_{q_{2i'}, h})_{2i'}^{\text{буф}} = 2$ ($h = \overline{2, m_{i'}}$), $(a_{q_{2i'}, 1})_{2i'}^{\text{буф}} = n_{i'} - \sum_{h=1}^{m_{i'}} (a_{q_{2i'}, h})_{2i'}^{\text{буф}}$.

8. Проверка корректности сформированного начального решения по составам партий данных в количестве $m_{i'}$: $(a_{q_{2i'}, 1})_{2i'}^{\text{буф}} \geq 2$; $(a_{q_{2i'}, h})_{2i'}^{\text{буф}} \geq (a_{q_{2i'}, h})_{2i'}^{\text{буф}}$ ($h = \overline{2, m_{i'}}$); при его выполнении реализуется переход на шаг 9. Если условие корректных составов партий для начального решения не выполняется, то формирование новых решений по составам партий данных i' -го типа невозможно, текущая i' -я строка матрицы A соответствует эффективному решению по составам партий данных этого типа; удаление i' -го типа данных из множества I : $I = I \setminus \{i'\}$ (формирование решений по составам партий данных для i' -го типа данных прекращается), переход на шаг 4.

9. Анализ эффективности решения по составам партий данных в количестве $m_{i'}$ (решение в окрестности O_{k+1} первого вида, $n_{p2}^{i'} = q_{2i'} = 1$). Для этого выполняется инициализация i' -й строки матрицы A' решением по составам партий данных, хранимым в строке $(A_{q_{2i'}})_{2i'}^{\text{буф}}$ матрицы $A_{2i'}^{\text{буф}}$: $(a_{i', h})' = (a_{q_{2i'}, h})_{2i'}^{\text{буф}}$ ($h = \overline{1, m_i}$ при $q_{2i'} = 1$); передача сформированного решения $[\mathbf{M}(s+g), A'(s+g)]$ для определения расписания обработки.

10. Получение для рассматриваемого решения $[\mathbf{M}(s+g), A'(s+g)]$ со второго уровня расписания обработки партий в виде $\left[P, (t_{jq}^{0l}) | l = \overline{1, L} \right]^*$, его использование для вычисления значения критерия $f_1(s+g)$.

11. Для решения $[\mathbf{M}(s+g), A'(s+g)]$ вычисление левого $-\nabla_g f_1(s)$ (либо правого $+\nabla_g f_1(s)$) дискретного градиента целевой функции f_1 [14].

12. При условии $+\nabla_g f_1(s) > 0$ сформированное решение $[\mathbf{M}(s+g), A'(s+g)]$ не является более эффективным, чем решение $[\mathbf{M}, A]^*$ (для количества партий $m_{i'}$), тогда реализуется переход на шаг 14 (параметр $q_{2i'}'$ не инициализируется).

13. Если для сформированного начального решения по составам $m_{i'}$ партий (строки $(A_{q_{2i'}})_{2i'}^{\text{буф}}$ матрицы $A_{2i'}^{\text{буф}}$) выполняется условие $-\nabla_g f_1(s) \leq 0$, то реализуется сравнение значений $-\nabla_g f_1(s)$ и G — градиента целевой функции f_1 . Если $-\nabla_g f_1(s) < G$, то текущее решение по составам партий данных i' -го типа является наилучшим среди рассмотренных решений по составам партий данных в окрестности O_{k+1} первого вида. Если $-\nabla_g f_1(s) < G$, то значения всех параметров q_{2i}' ($i \in I$) обнуляются, индекс q_{2i}' решения по составам партий данных i' -го типа (q_{2i}' -й строки $(A_{q_{2i'}})_{2i'}^{\text{буф}}$ матрицы $A_{2i'}^{\text{буф}}$) сохраняется: $q_{2i}' = q_{2i}$. Изменение значения G : $G = -\nabla_g f_1(s)$, переход на шаг 14.

14. Осуществляется проверка условия $I' \neq \emptyset$; в случае его выполнения реализуется переход на шаг 4; при $I' = \emptyset$ (для каждого i -го типа данных ($i \in I$) сформированы новые решения по составам партий) осуществляется переход на шаг 15.

15. Проверка условия $I = \emptyset$. При его выполнении отсутствуют i -е типы данных ($i \in I$), для которых сформированные решения удовлетворяют условию корректного состава партий (отсутствует возможность анализа новых решений). Осуществляется переход на шаг 24. При $I \neq \emptyset$ для i -х типов данных ($i \in I$) реализуется проверка условия $\exists i \in I: q_{2i}' \neq 0$. При выполнении для i' -го типа данных ($i' \in I$) условия $q_{2i}' \neq 0$ решение по составам его партий (при неизменных составах партий других i -х типов ($i \neq i', i \in I$)) обеспечивает максимальное уменьшение значения целевой функции $f_1(s)$ в рамках окрестности O_{k+1} первого вида; реализуется переход на шаг 16. В случае, когда для всех i -х типов данных ($i \in I$) выполняется условие $q_{2i}' = 0$, в рамках окрестности O_{k+1} первого вида не определено решение, улучшающее решение $[\mathbf{M}, A]^*$; переход на шаг 17.

16. Формирование нового локально оптимального решения $[\mathbf{M}, A]^*$ с учетом составов партий данных i' -го типа, для которого $q_{2i}' \neq 0$: инициализация i' -й строки матрицы A текущего решения $[\mathbf{M}, A]$ значениями элементов q_{2i}' -й строки $(A_{q_{2i'}})_{2i'}^{\text{буф}}$ матрицы $A_{2i'}^{\text{буф}}$: $a_{i'h} = (a_{q_{2i}', h})_{2i'}^{\text{буф}}$ ($h = \overline{1, m_i}$); в результате получено новое локально оптимальное решение $[\mathbf{M}, A]^*$; инициализация $n_{p1}^i = 1$ ($i = \overline{1, n}$); $I' = I$; переход на шаг 4.

17. Формирование решений по составам партий данных i -х типов, включаемых в окрестность O_{k+1} второго вида, путем комбинирования решений по составам партий данных i -х типов, находящихся в матрицах $A_{2i}^{\text{буф}}$ ($i = \overline{1, n}$). Для этого вызывается алгоритм формирования составов партий данных i -х типов, рассмотренный выше, использующий решения, хранящиеся в матрицах $A_{2i}^{\text{буф}}$, для формирования новых решений в окрестности O_{k+1} второго вида и определяющий лучшее из них в этой окрестности.

18. Для i -х типов данных ($i \in I$) проверка условий $q_{2i}' = 0$. При выполнении этих условий в окрестности O_{k+1} второго вида не сформировано решение, лучшее, чем текущее локально оптимальное решение $[\mathbf{M}, A]^*$. Реализуется переход на шаг 20. Если для i -х типов ($i \in I$) условия $q_{2i}' = 0$ не выполняются, то в окрестности O_{k+1} второго вида сформировано решение, являющееся лучшим, чем текущее локально оптимальное решение $[\mathbf{M}, A]^*$; реализуется переход на шаг 19.

19. Выполняется формирование нового локально оптимального решения $[\mathbf{M}, A]^*$ с учетом составов партий данных i -х типов, для которых $q_{2i}' \neq 0$. Для каждого i -го идентификатора типа данных, для которого выполняется условие $q_{2i}' \neq 0$, реализуется инициализация i -й строки матрицы A решения $[\mathbf{M}, A]^*$ значениями элементов q_{2i}' -й строки $(A_{q_{2i'}})_{2i'}^{\text{буф}}$ матрицы $A_{2i'}^{\text{буф}}$: $a_{ih} = (a_{q_{2i}', h})_{2i'}^{\text{буф}}$ при $h = \overline{1, m_i}$. Результат — новое локально оптимальное решение $[\mathbf{M}, A]^*$; инициализация $n_{p1}^i = 1$ ($i \in I$), $I' = I$; переход на шаг 4.

20. Инициализация множества типов данных I' : $I' = I$.

21. Идентификация i' -го типа данных, решения по составам партий данных которого буферизируются в матрице $(A)_{1i'}^{\text{буф}}$: $i' = \min\{i | i \in I\}$, $I' = I' \setminus \{i'\}$.

22. Инициализация строк матрицы $(A)_{1i'}^{\text{буф}}$ решениями по составам партий данных i' -го типа, находящимися в матрице $(A)_{2i'}^{\text{буф}}$, в количестве $n_{2p}^{i'}$: $(a_{q1_{i'},h})_{1i'}^{\text{буф}} = (a_{q2_{i'},h})_{2i'}^{\text{буф}}$ ($h = \overline{1, m_i}$, $q1_{i'} = \overline{1, n_{p2}^{i'}}$, $q2_{i'} = \overline{1, n_{p2}^{i'}}$), присваивание $n_{p1}^{i'} = n_{p2}^{i'}$. В результате сформирована матрица $(A)_{1i'}^{\text{буф}}$, содержащая решения, на основе которых будут формироваться решения в окрестности O_k первого вида с большей метрикой.

23. Проверка условия $I' \neq \emptyset$. Если условие выполняется, то реализуется переход на шаг 21. В случае выполнения условия $I' = \emptyset$ промежуточные решения по составам партий данных всех i -х типов ($i \in I$) буферизированы в соответствующих матрицах $(A)_{1i}^{\text{буф}}$; осуществляется присваивание $I' = I$, $k = k + 1$, реализуется переход на шаг 4.

24. Окончание алгоритма.

Вычислительная сложность сформулированного алгоритма определения локально оптимальных решений по составам партий составляет $O(n(n^i)^2)$. Точность приближения получаемых решений к глобально оптимальному $\varepsilon = 0.001$.

Метод определения составов партий данных i -х типов ($i = \overline{1, n}$) взаимодействует с методом формирования расписаний обработки партий, предложенным в [13].

Основным компонентом рассматриваемой иерархической системы построения комплексных расписаний обработки партий является подсистема формирования составов партий, функционирующая на верхнем уровне. Для анализа эффективности метода определения решений по составам партий введены следующие обозначения параметров задачи:

— $\max(t_i^l) / \min(t_i^l)$ — отношение максимальной длительности обработки данных i -х типов ($i = \overline{1, n}$) на всех l -х сегментах конвейера ($l = \overline{1, L}$) к минимальной длительности обработки данных этих типов;

— $\max(t_{ij}^l) / \min(t_{ij}^l)$ — отношение максимальной длительности переналадки сегментов конвейера ($l = \overline{1, L}$) к минимальной длительности переналадки этих сегментов.

В исследованиях параметр n принимал значения 5, 10 (при увеличении значений этого параметра тенденция изменения эффективности метода определения составов партий сохраняется), параметр n^i ($i = \overline{1, n}$) принимал значения 12, 16, 24. Значения параметра L задавались равными 5, 10. Значения отношения $\max(t_{ij}^l) / \min(t_{ij}^l)$ задавались равными 1, 2, 4, 8, 16. Аналогичным образом значения отношения $\max(t_i^l) / \min(t_i^l)$ задавались равными 1, 2, 4, 8, 16.

При исследовании эффективности метода построения решений по составам партий данных для определенного набора исходных входных характеристик формировались решения по обработке:

— фиксированных партий (набор данных определенного i -го типа в количестве n^i интерпретировался как одна фиксированная партия);

— нефиксированных партий (формировались эффективные решения по составам партий данных и расписания их обработки).

В качестве параметра эффективности полученных решений определялось отношение разности значений моментов времени окончания обработки фиксированных партий и партий, сформированных в соответствии с предложенным методом, к моменту времени окончания обработки фиксированных партий (соответствующие зависимости

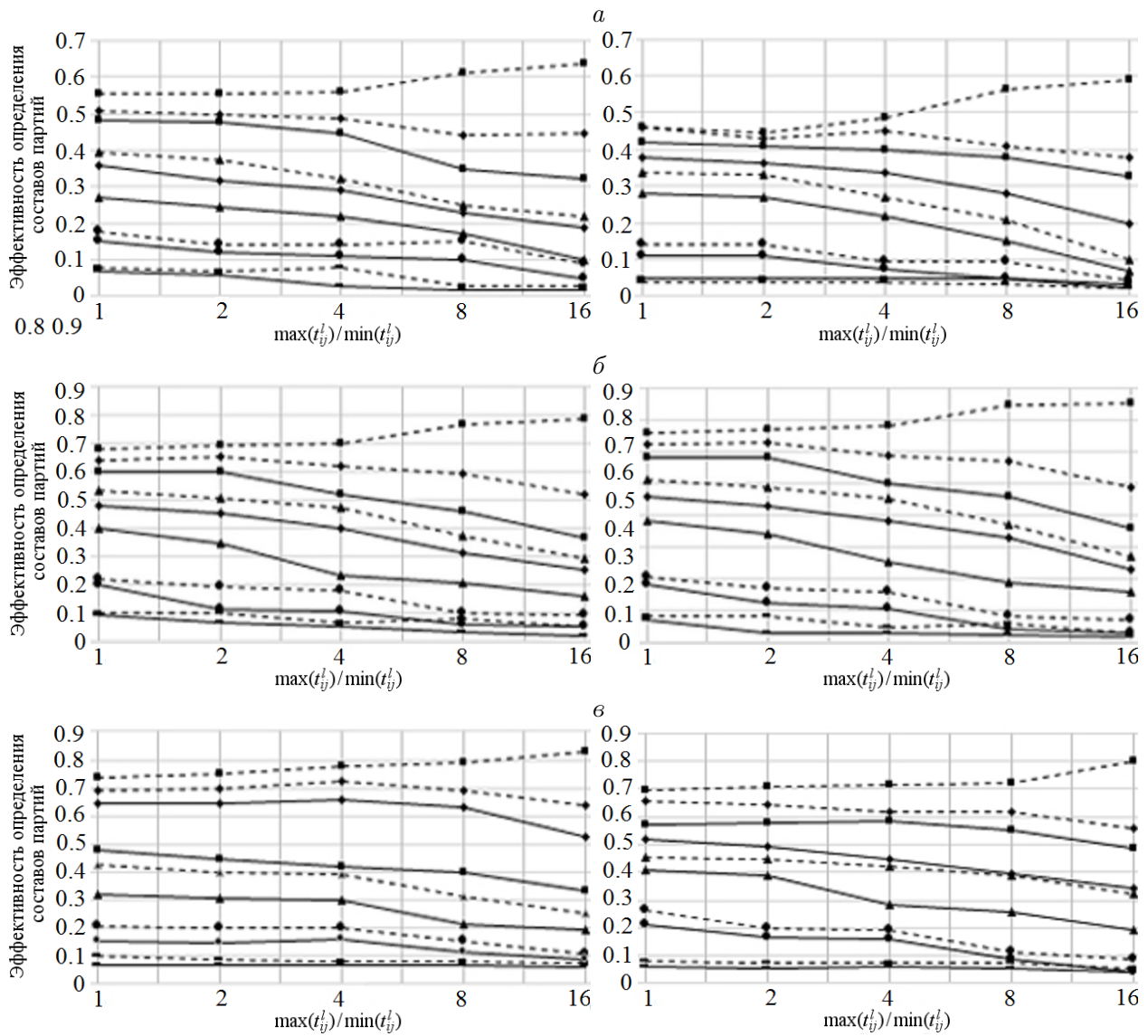


Рис. 1. Эффективность метода формирования составов партий при $n^i = 12$ (а), $n^i = 16$ (б), $n^i = 24$ (в). Слева — $n = 10, L = 5$; справа — $n = 10, L = 10$. Сплошные линии — ГА, штриховые — МОСП при $\max(t_{ij}^l) / \min(t_{ij}^l) = 1, 2$ (●), 4 (◆), 8 (▲), 16 (■) соответственно

для метода оптимизации составов партий (МОСП) и генетических алгоритмов (ГА) представлены на рисунке).

На основе полученных оценок эффективности определения решений по составам партий данных установлены особенности процесса оптимизации:

1) при одинаковых значениях n и n^i увеличение количества приборов L обуславливает снижение эффективности метода определения составов партий данных;

2) увеличение значения n^i при прочих неизменных значениях входных параметров обуславливает увеличение эффективности метода формирования решений по составам партий;

3) увеличение значений отношения $\max(t_{ij}^l) / \min(t_{ij}^l)$ обуславливает увеличение эффективности формирования решений по составам партий данных, т. е. увеличение неоднородности в длительностях переналадок приборов вызывает рост эффективности метода формирования решений по составам партий данных;

4) увеличение значений отношения $\max(t_i^l) / \min(t_i^l)$ при зафиксированных значениях отношения $\max(t_{ij}^l) / \min(t_{ij}^l)$ обуславливает увеличение эффективности формирования решений по составам партий данных.

Также реализовано сравнение результатов выполненных исследований по определению эффективности разработанного метода формирования составов партий с результатами других методов определения решения задачи (в частности, с генетическими алгоритмами). Для этого применен алгоритм, предложенный в [13]. Аппарат генетических алгоритмов выбран в качестве способа получения решений рассматриваемой задачи, так как он достаточно проработан к настоящему времени и позволяет получать эффективные решения.

Выводы

В работе сформулированы способ определения составов партий и метод формирования локально оптимальных решений по составам партий данных. В целом (при различных значениях входных параметров) реализация метода формирования решений по составам партий позволяет более чем на 80 % сократить время обработки данных в конвейерной системе по сравнению с обработкой фиксированных партий (т.е. без применения предложенного метода) и на 30 % по сравнению с решениями, полученными с использованием аппарата генетических алгоритмов. Дальнейшие исследования направлены на разработку методов построения комплексных расписаний обработки партий данных при наличии ограничений на длительность интервалов функционирования системы при условии формирования комплектов из обработанных данных и установленных для них (комплектов) директивных сроков.

Список литературы / References

- [1] Миклашевич С.Э., Балашов И.В., Бурцев М.А. и др. Программно-аппаратный комплекс для сбора, обработки, архивации и распространения спутниковых данных // *Соврем. пробл. дистанционного зондирования Земли из космоса*. 2012. Т. 9, № 4. С. 47–56. Miklashevich, S.E., Balashov, I.V., Burtsev, M.A. et al. Complex system for the receiving, processing, archiving and distribution of satellite data and products of thematic processing // *Current Problems in Remote Sensing of the Earth from Space*. 2012. Vol. 9, No. 4. P. 47–56. (In Russ.)
- [2] Лупян Е.А., Балашов И.В., Бурцев М.А. и др. Создание технологий построения информационных систем дистанционного мониторинга // *Соврем. пробл. дистанционного зондирования Земли из космоса*. 2015. Т. 12, № 5. С. 53–75. Lupyan, E.A., Balashov, I.V., Burtsev, M.A. et al. Development of information systems design technologies // *Current Problems in Remote Sensing of the Earth from Space*. 2015. Vol. 12, No. 5. P. 53–75. (In Russ.)
- [3] Лупян Е.А., Прошин А.А., Бурцев М.А. и др. Центр коллективного пользования системами архивации, обработки и анализа спутниковых данных ИКИ РАН для решения задач изучения и мониторинга окружающей среды // *Соврем. пробл. дистанционного зондирования Земли из космоса*. 2016. Т. 13, № 5. С. 11–28. Lupyan, E.A., Proshin, A.A., Burtsev, M.A. et al. IKI center for collective use of satellite data archiving, processing and analysis systems aimed at solving the problems of environmental study and monitoring // *Current Problems in Remote Sensing of the Earth from Space*. 2016. Vol. 13, No. 5. P. 11–28. (In Russ.)

- [4] **Барталев С.А., Егоров В.А., Жарко В.О. и др.** Спутниковое картографирование растительного покрова России. М.: ИКИ РАН, 2016. 208 с.
Bartalev, S.A., Egorov, V.A., Zharko, V.O. et al. Satellite mapping of vegetation cover of Russia. M.: IKI RAN, 2016. 208 p. (In Russ.)
- [5] **Кашницкий А.В., Балашов И.В., Лупян Е.А. и др.** Создание инструментов для удаленной обработки спутниковых данных в современных информационных системах // Современ. пробл. дистанц. зондирования Земли из космоса. 2015. Т. 12, № 1. С. 156–170.
Kashnitsky, A.V., Balashov, I.V., Lupyay, E.A. et al. Development of software tools for satellite data remote processing in contemporary information systems // Current Problems in Remote Sensing of the Earth from Space. 2015. Vol. 12, No. 1. P. 156–170. (In Russ.)
- [6] **Кобец Д.А., Матвеев А.М., Мазуров А.А., Прошин А.А.** Организация автоматизированной многопоточковой обработки спутниковой информации в системах дистанционного мониторинга // Современ. пробл. дистанц. зондирования Земли из космоса. 2015. Т. 12, № 1. С. 145–155.
Kobets, D.A., Matveev, A.M., Mazurov, A.A., Proshin, A.A. Organization of automated multithreaded processing of satellite information in remote monitoring systems // Current Problems in Remote Sensing of the Earth from Space. 2015. Vol. 12, No. 1. P. 145–155. (In Russ.)
- [7] **Шокин Ю.И., Пестунов И.А., Смирнов В.В. и др.** Распределенная информационная система сбора, хранения и обработки спутниковых данных для мониторинга территорий Сибири и Дальнего Востока // Журн. Сибирского федерального ун-та. Серия: Техника и технологии. 2008. № 4. С. 291–314.
Shokin, Yu.I., Pestunov, I.A., Smirnov, V.V. et al. Distributed information system of collection, storage and processing of satellite data for monitoring of the territories of Siberia and the Far East // J. of Siberian Federal Univ. Engineering & Technologies. 2008. No. 4. P. 291–314. (In Russ.)
- [8] **Шокин Ю.И., Добрецов Н.Н., Кихтенко В.А. и др.** О распределенной инфраструктуре системы оперативного спутникового мониторинга ЦКП ДДЗ СО РАН // Вычисл. технологии. 2013. Т. 18. Спец. выпуск. С. 86–94.
Shokin, Yu.I., Dobretsov, N.N., Kikhtenko, V.A. et al. On a distributed infrastructure for the monitoring of satellite remote sensing data using the center for shared access // Comput. Technologies. 2013. Vol. 18, Special issue. P. 86–94. (In Russ.)
- [9] **Бабяк П.В., Недолужко И.В., Фомин Е.В.** Подход к предоставлению услуг по обработке спутниковых данных в Центре коллективного пользования регионального спутникового мониторинга окружающей среды ДВО РАН // Тр. XIV Всерос. объедин. конф. “Интернет и современное общество” (IMS-2011), 10–12 октября 2011 г., Санкт-Петербург, 2011. С. 27–32.
Babiak, P.V., Nedoluzhko, I.V., Fomin, E.V. Approach to the provision of services for the processing of satellite data in the Center of collective use for regional satellite environmental monitoring of the FEB of RAS // Proc. of the XIVth all-Russian Joint Conf. “Internet and Modern Society” (IMS-2011), 10–12 Oct. 2011, St. Petersburg, 2011. P. 27–32. (In Russ.)
- [10] **Тео, У.М., Тай, С.С., Гозали, Ж.П.** Distributed geo-rectification of satellite images using grid computing // Proc. of Intern. Parallel and Distributed Proc. Symp. IEEE Comput. Soc. Press, 2003. P. 152–157.
- [11] **Suarez, R.A., Nasmachnow, S.** Parallel computing applied to satellite images processing for solar resource estimates // CLEI Electronic J. 2012. Vol. 15, No. 3. Paper 4.
- [12] **Кротов К.В.** Многоуровневая модель построения расписаний обработки партий данных в конвейерных системах при формировании комплектов и наличии ограничений // Тр. СПИИРАН. 2016. Вып. 4(47). С. 65–91.

Krotov, K.V. Multilevel model of construction of the processing schedules batch data in conveyor systems in the formation of sets and availability restrictions // Trudy SPIIRAN. 2016. Iss. 4(47). P. 65–91. (In Russ.)

- [13] **Кротов К.В.** Использование аппарата генетических алгоритмов при формировании решений по составам партий данных в двухуровневой задаче построения комплексных расписаний их обработки // Автоматизированные технологии и производства. 2017. № 2. С. 23–34.

Krotov, K.V. Using the method of genetic algorithms for designing the decisions for the composition of data batches in a two-level problem of building complex schedules within their processing // Avtomatizirovannye Tekhnologii i Proizvodstva. 2017. No. 2. P. 23–34. (In Russ.)

- [14] **Ковалев М.М.** Матроиды в дискретной оптимизации. М.: Изд-во “Едиториал УРСС”, 2003. 224 с.

Kovalev, M.M. Matroids in discrete optimization. Moscow: Izd-vo “Editorial URSS”, 2003. 224 p. (In Russ.)

*Поступила в редакцию 29 июня 2017 г.,
с доработки — 27 ноября 2017 г.*

The complex method for definition of effective decisions on data set composition and shedule for their processing in conveyor systems

KROTOV, KIRILL V.

Sevastopol State University, Sevastopol, 290053, Russia

Corresponding author: Krotov, Kirill V., e-mail: krotov_k1@mail.ru

The task of controlling the pipelined processing for various types of remote sensing data of the Earth is important. The task of managing the data processing process is complex and represents a set of two interrelated subtasks — the definition of the composition of batches and the schedules of their processing. The complex task of managing the processing of data lots is represented in the form of hierarchy of the required subtasks. The realization of effective decisions on the composition of parties is realized at the top level, schedules and their processing — on the lower level. A hierarchical game model (two-level programming) for determining solutions at the appropriate levels is accordingly designed. The final time for processing all data entering the system is considered as a criterion at the upper level (to determine the composition of data lots). The total downtiming of the pipeline segments when processing batches of data is the criterion at the lower level. Local optimization methods are proposed at each level of the hierarchy to determine effective solutions.

The method for determining solutions by party composition provides for the search for a better solution in the vicinity of different types of the current locally optimal solution. The approaches of making decisions on the composition of parties in the vicinity of the respective species are formulated. The proof of both theorems determining the conditions for the end of the change in the composition of a given number of batches and that eliminates duplication of solutions in a certain neighborhood is accomplished. A generalized algorithm for determining locally optimal solutions by party composition, acting on the upper level of the hierarchy, is formulated based on these methods.

The algorithm for determining effective schedules is based on “greedy” strategies and involves adding each new batch to the processing sequences optimized in the previous steps and determining an effective position for them in this order.

Software implementation of the developed methods of local optimization of solutions at each level of the hierarchy is completed. Studies for the dependence of the effectiveness of the proposed methods (in particular, the method of forming batches of batches) on the parameters of the system and the data processed are carried out. The obtained results have shown that use of the offered method of the definition of batches allows considerable reducing for the time of processing of batches of the data in the conveyor system. The efficiency of the treatment is then increased by 10 % to 80 %.

Keywords: batch processing schedules, conveyor system, two-level game model, batch composition determination method.

Received 29 Juny 2017

Received in revised form 27 November 2017