

В.М. Курейчик, И.Б. Сафроненкова**ОНТОЛОГИЧЕСКИЙ ПОДХОД К РЕАЛИЗАЦИИ ТЕХНОЛОГИЙ
РАСПРЕДЕЛЕННЫХ ВЫЧИСЛЕНИЙ В СЕТИ ИНТЕРНЕТ**

Развитие технологий распределенных вычислений позволило объединить географически-распределенные ресурсы, таким образом, предоставив возможности для эффективного решения ресурсоемких задач в различных областях науки и техники. Наряду с этим актуализировался ряд задач, который требует к своему решению новых подходов, учитывающих особенности реализации современных Интернет технологий. В настоящей работе рассмотрена проблема, связанная с переносом вычислительной нагрузки в распределенной системе автоматизированного проектирования (САПР), функционирующей в «туманной» среде. Целью данной работы является разработка онтологического подхода к решению задачи переноса вычислительной нагрузки в САПР с учетом особенностей «туманной» среды. Онтологический подход заключается в проведении процедуры онтологического анализа, которая позволяет «отсеивать» узлы-кандидаты, не отвечающие ресурсным требованиям для переноса части нагрузки. Научная новизна работы заключается в использовании моделей онтологии для решения задачи переноса вычислительной нагрузки в САПР. Это позволяет сократить число узлов-кандидатов в «тумане» для переноса нагрузки, тем самым сократить время моделирования процессов размещения и, соответственно, общее время решения задачи переноса нагрузки. Принципиальным отличием данного подхода является использование знаний о предметной области, отраженных в модели онтологии, для решения задачи переноса вычислительной нагрузки. Проведенные в работе вычислительные эксперименты доказали целесообразность использования онтологического анализа для решения задачи переноса вычислительной нагрузки.

Распределенная САПР; «туманные» вычисления; задача переноса вычислительной нагрузки; онтология.

V.M. Kureichik, I.B. Safronenkova**AN ONTOLOGICAL APPROACH TO DISTRIBUTED COMPUTING
TECHNOLOGIES IMPLEMENTATION ON THE INTERNET**

Distributed computing technologies development has allowed uniting geographically distributed resources and has provided an opportunity for effective resource intensive problem-solving in various fields of science and technology. At the same time a set of problems, which demands the development of new approaches, taking into account contemporary Internet technologies implementation, has risen. In this paper a problem of workload relocation in distributed computer-aided design system (DCAD) operating in the “fog” environment was considered. The goal of this paper is ontological approach development to workload relocation problem-solving in DCAD taking into account some “fog” environment features. The ontological approach involves an ontological procedure implementation, which allows “filtering” the candidate-nodes, which have insufficient resources for workload relocation. The scientific novelty of this paper is ontological models using for workload relocation problem-solving in DCAD. It allows reducing the number of candidate-nodes in the “fog” for workload relocation, thereby contributing to reduce the time of location process modeling and, consequently, the total time of workload relocation problem-solving is also reduced. The fundamental difference of presented approach is domain knowledge, represented in ontological model, applying for workload relocation problem-solving. The experimental study results have shown the expediency of ontological analysis for workload relocation problem-solving.

Distributed computer-aided design system; “fog”- computing; workload relocation problem; ontology.

Введение. Задача формирования расписания не нова, имеет различные постановки и решалась многократно множеством ученых [1, 2]. Однако данная проблема не теряет актуальности и в настоящее время, приобретая новые особенности в рамках развития технологий распределенных вычислений и Интернет-технологий [3–5].

Стоит отметить, что проектирование современных СБИС даже посредством распределенных систем автоматизированного проектирования (РСАПР) требует наличия значительных вычислительных и временных ресурсов. В этой связи быстроедействие РСАПР является определяющим фактором конкурентоспособности производства в целом [6].

Тот факт, что современные РСАПР успешно функционируют на базе Интернет-технологий, выполняющих задачи коммуникации и информационного обмена между приложениями [7], подтверждается появлением на рынке множества продуктов автоматизированных систем, реализованных на базе «облачных» и «туманных» концепций [8, 9].

Системы, в которых применяется «безсерверная архитектура» (serverless architecture) обеспечивают выполнение исходного кода тысяч и миллионов пользователей внутри вычислительной среды, не заботясь о масштабировании ресурсов. Это, в свою очередь, делает возможным обеспечить доступ к информационным ресурсам распределенных баз данных и организацию дистанционного взаимодействия распределенных коллективов пользователей САПР, что гарантирует высокий технический уровень проектов, уменьшение сроков их реализации и сведение к минимуму числа ошибок [10].

Туманные вычисления – это высоко-виртуализированная платформа, предоставляющая вычислительные, запоминающие и сетевые сервисы между конечными устройствами и центрами данных облачных вычислений, и которая обычно, но не всегда, располагается на границах сети [11]. «Туманную» среду можно рассматривать как расширение «облачных» вычислений до границ сети, которая дополняет централизованную «облачную» модель. В этом случае часть функций по обработке и хранению данных передается граничным узлам сети, что позволяет решать многие задачи, возникающие при экспоненциальном росте числа подключенных к сети устройств. В связи с этим, главной характеристикой «туманных» вычислений является возможность переноса части вычислительно нагрузки, что, в условиях «туманной» среды, предполагает решение оптимизационной задачи с определенными критериями и ограничениями, которые частично формируются требованиями самой вычислительной системы [12, 13].

В настоящей работе описан онтологический подход к решению задачи переноса вычислительной нагрузки в РСАПР, функционирующей на базе «туманных» вычислений.

Результаты проведенных вычислительных экспериментов подтвердили целесообразность использования онтологического подхода к реализации технологий распределенных вычислений.

Постановка задачи переноса вычислительной нагрузки в «туманной» среде. Пусть дан граф вычислительной задачи G , с некоторой вычислительной сложностью подзадачи x_i и объемом информации w_i , передаваемой между подзадачами. Граф G разделен на два подграфа G' и G'' . Необходимо разместить подграф G' на вычислительных устройства (ВУ) сегмента сети P' «туманного» слоя, в то время, как вычислительные задачи подграфа G'' продолжает исполняться на сегменте сети P'' (рис. 1).

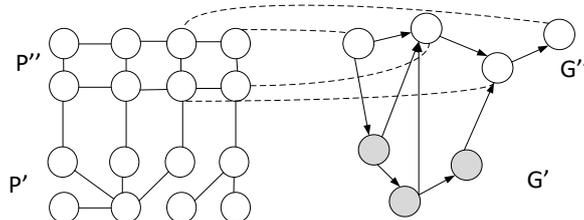


Рис. 1. Распределение вычислительных подзадач при переносе части вычислительной нагрузки

Будем считать, что основным критерием при решении оптимизационной задачи переноса вычислительной нагрузки является надежность системы.

Рассмотрим графовое описание набора задач

$$G = \{ \langle i, x_i, w_i \rangle \}, \quad (1)$$

где i – уникальный идентификатор вычислительной подзадачи;

x_i – вычислительная сложность i -ой подзадачи;

w_i – объем информации, передаваемый i -ой подзадачей в коммуникационную среду.

Вычислительные подзадачи графа G связаны с узлами вычислительных устройств из множества P . Причем P описывается графой структурой

$$P = \{ \langle j, p_j \rangle \text{ list} \}, \quad (2)$$

где j – идентификатор узла;

p_j – производительность узла;

$list$ – матрица пропускной способности каналов связи между инцидентными узлами сети.

Теперь рассмотрим подграф вычислительных подзадач G' , которые требуется перенести в то время, как вычислительные подзадачи подграфа G'' продолжают свое исполнение. Между данными подграфами существует несколько информационных потоков. Потоки, связанные с подграфом G' могут быть описаны набором кортежей:

$Flow_in = \{ \langle id_{out}, id_{in}, w_{out_in} \rangle \}$ – кортеж, описывающий количество информации, передаваемой между узлом $G'' id_{out}$ и узлом $G' id_{in}$ от G'' к G' .

$Flow_out = \{ \langle id_{in}, id_{out}, w_{in_out} \rangle \}$ – кортеж, описывающий количество информации, передаваемой между узлом $G' id_{out}$ и узлом $G'' id_{in}$ от G' к G'' .

Рассмотрим задачу размещения нагрузки. Пусть имеется подграф вычислительных задач G'' , связанный с P'' , и $Flow_in, Flow_out$. Необходимо разместить вычислительные подзадачи подграфа G' на множестве устройств сети P' таким образом, что общее время выполнения вычислительных подзадач G меньше заданного времени T с учетом выполнения критерия надежности системы.

Решением поставленной задачи является установление связи между вычислительными подзадачами G' и вычислительными узлами P' , которое может быть описано матрицей A :

$$A = \begin{pmatrix} \langle t_o^{ij}, u_{ij} \rangle & & & \\ & \dots & & \\ & & & \langle t_o^{NM}, u_{NM} \rangle \end{pmatrix}, \quad (3)$$

где t_o^{ij} – момент времени, когда начинается вычисление i -ой подзадачи j -м узлом;
 u^{ij} – доля общей производительности p_j заданного j -го узла для выполнения i -ой подзадачи.

Для дальнейшей разработки модели необходимо рассмотреть следующие параметры:

- ◆ $L_p(A)$ – нагрузка узла, порождаемая переносом вычислительной подзадачи на узел;
- ◆ $L_{dist}(A, Flow_in, Flow_out)$ – нагрузка узла, порождаемая обменом информации между подграфами G' и G'' ;
- ◆ $L_{tr}(A, Flow_in, Flow_out)$ – нагрузка узла, порождаемая передачей информации через узел;
- ◆ D_{lk} – список ребер графа P , который определяет маршрут между узлами l и k ;
- ◆ $ListD_{lk}$ – матрица, описывающая пропускную способность каналов связи между узлами l и k .

Рассмотрим целевую функцию как набор значений функций надежности определенных узлов F_j :

$$F_j = e^{-\lambda_j t}, \quad (4)$$

где λ_j – интенсивность отказов j -го узла, t – время функционирования устройства.

Поскольку

$$\lambda = \lambda_0 \cdot 2^{\Delta T / 10}, \quad (5)$$

$$\Delta T = kL, \quad (6)$$

где L – нагрузка устройства, k – коэффициент, который зависит от типа устройства.

Тогда зависимость между функцией надежности и нагрузкой описывается следующей формулой:

$$F_j = e^{-\lambda_{j0} \cdot 2^{kL/10}}. \quad (7)$$

Загрузка устройства зависит от распределения вычислительных задач по ВУ, которые описываются матрицей A . Рассмотренные выше параметры должны быть включены в модель задачи: $L_p(A)$, $L_{dist}(A, Flow_in, Flow_out)$, $L_{tr}(A, Flow_in, Flow_out)$.

Таким образом, полная нагрузка j -го ВУ описывается следующей формулой:

$$L_j = L_{pj}(A) + L_{distj}(A, Flow_in, Flow_out) + L_{trj}(A). \quad (8)$$

Главным ограничением для этой задачи является время исполнения T графа вычислительной задачи G , т.е. для:

$$G = G' \cup G'', \forall i \in G: \frac{x_i}{p_j u_{ij}} + t_{dist}(i) < T. \quad (9)$$

где $t_{dist}(i)$ – максимальное время доставки информации от i -ой подзадачи к подзадаче-получателю данных.

Поскольку модель рассматривает маршруты потока информации, $t_{dist}(i)$ описывается функцией:

$$t_{dist}(i) = \xi(A, G, P). \quad (10)$$

Более точно время доставки информации может быть вычислено на основе полной информации о подзадачах, назначаемых на вычислительные узлы с учетом параметров D_{ik} и $ListD_{ik}$ [14].

Онтологический подход к решению задачи переноса вычислительной нагрузки. Как известно, онтологии широко применяются для концептуального описания различных предметных областей [15–17].

В основе онтологического подхода к решению задачи переноса вычислительной нагрузки лежит онтологический анализ.

Введем понятие «узел-лидер». Условимся, что узел-лидер выбирается с использованием алгоритма избрания лидера из числа вычислительных узлов, на которых изначально выполнялась задача, и выполняет функцию сбора информации об узлах-кандидатах, расположенных в «тумане», потенциально готовых принять часть вычислительной нагрузки.

Общая схема процедуры онтологического анализа изображена на рис. 2 и включает в себя выполнение следующих этапов:

1. Входные данные, т.е. исходные данные параллельного популяционного алгоритма, реализуемого в РСАПР.
2. Классификация совокупности входных данных в соответствии с классами предметной онтологии;
3. Применение системы продукционных правил к классам онтологии с целью ограничения множества узлов, которое было получено в результате сбора сведений об имеющихся ресурсах узлом-лидером.
4. Принятие решения о выборе ограниченного множества узлов для переноса вычислительной нагрузки.



Рис. 2. Схема процедуры онтологического анализа

Ключевым элементом онтологического анализа является модель онтологии, которая позволяет структурировать знания об аспектах, являющихся ключевыми для решения задачи переноса вычислительной нагрузки, а именно:

Исходные данные переносимой задачи:

- ◆ классы параллельных алгоритмов, используемые в РСАПР (алгоритмы, вдохновленные живой и неживой природой, алгоритмы, инспирированные человеческим обществом и прочие) [18];

- ◆ модели распараллеливания алгоритмов (глобальная модель, островная модель) [19];
 - ◆ изменяемые параметры данных моделей (длительность сезона);
 - ◆ способы разбиения моделей распараллеливания (будем считать число способов разбиений конечно);
 - ◆ регламентированное время выполнения алгоритма;
- Информация о вычислительных ресурсах узлов-кандидатов:
- ◆ исходное размещение вычислительных узлов («туман», «облако»);
 - ◆ требования к вычислительным ресурсам (производительность узла-кандидата, удаленность узла-кандидата от узла-лидера, загруженность узла-кандидата).

Фрагмент разработанной модели онтологии изображен на рис. 3.

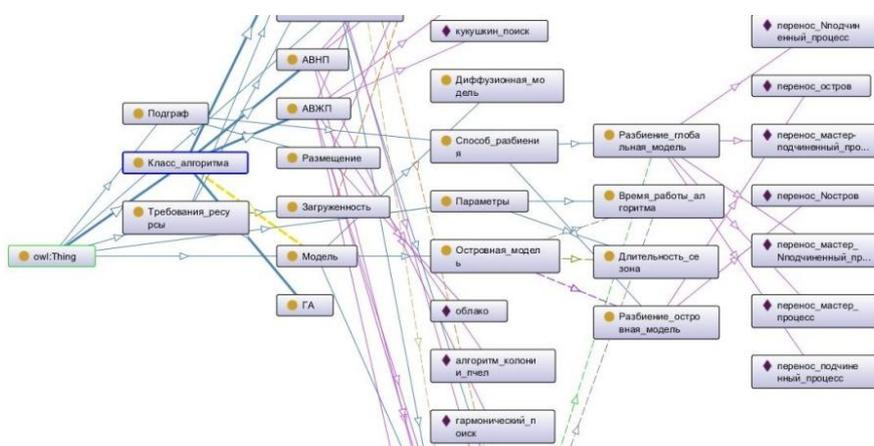


Рис. 3. Фрагмент разработанной онтологии

Таким образом, онтологический подход к решению задачи переноса вычислительной нагрузки, в основе которого лежит онтологический анализ позволяет «отсеять» узлы-кандидаты, заведомо не удовлетворяющие требованиям, необходимым для переноса части вычислительной нагрузки. Это позволяет сократить время, требуемое для моделирования процессов размещения, и, тем самым сократить время решения задачи переноса вычислительной нагрузки.

Реализация предложенной процедуры онтологического анализа производилась в среде программирования Python 3.8.3 для работы в операционной среде семейства MS Windows [20].

Вычислительный эксперимент. Для демонстрации эффективности работы процедуры онтологического анализа для распределенных систем автоматизированного проектирования, проведем две серии экспериментов в различных условиях. Для этого введем некоторые понятия.

1. В качестве критерия эффективности выберем время, необходимое для решения задачи переноса вычислительной нагрузки.

2. Эффективность работы модуля, реализующего процедуру онтологического анализа, будем определять по формуле:

$$\Delta = \frac{F_n - F_k}{F_n} \times 100\% , \quad (1)$$

где F_n – начальное значение целевой функции;

F_k – конечное значение целевой функции;

Δ – оценочная функция изменения времени, необходимого для решения задачи переноса вычислительной нагрузки в РСАПР.

3. Под «глубиной туманного» слоя будем понимать максимальное расстояние в хопах от узла-лидера до узла-кандидата, расположенного в «туманном» слое.

Первая серия вычислительных экспериментов заключалась в проведении экспериментов по реализации процедуры онтологического анализа на множестве доступных вычислительных узлов «туманного» слоя, глубина которого лежит в интервале от 2 до 5 хопов, доступное количество узлов в «туманном» слое $F=100$ и $F=1000$.

Вторая серия вычислительных экспериментов заключалась в проведении экспериментов по реализации процедуры онтологического анализа на множестве доступных вычислительных узлов «туманного» слоя, глубина которого лежит в интервале от 2 до 30 хопов, доступное количество узлов в «туманном» слое $F=100$ и $F=1000$.

Результаты двух серий вычислительных экспериментов отражены в табл. 1, 2 и на рис. 4, 5.

Таблица 1

Сравнение эффективности процедуры онтологического анализа при глубине «туманного» слоя 2 до 5 хопов ($\Delta\{2;5\}$) и от 2 до 30 хопов ($\Delta\{2;30\}$) в процентах, $F=100$

Номер эксперимента	Количество доступных вычислительных узлов, F	Количество узлов после процедуры онтологического анализа, F_s для глубины «туманного слоя» {2;5}	Количество узлов после процедуры онтологического анализа, F_s для глубины «туманного слоя» {2;30}	$\Delta\{2;5\}$, %	$\Delta\{2;30\}$, %
1	100	14	3	86	97
2	100	11	1	89	99
3	100	15	3	85	97
4	100	8	2	92	98
5	100	17	3	83	97
6	100	14	3	86	97
7	100	15	2	85	98
8	100	10	1	90	99
9	100	11	2	89	98
10	100	21	2	79	98

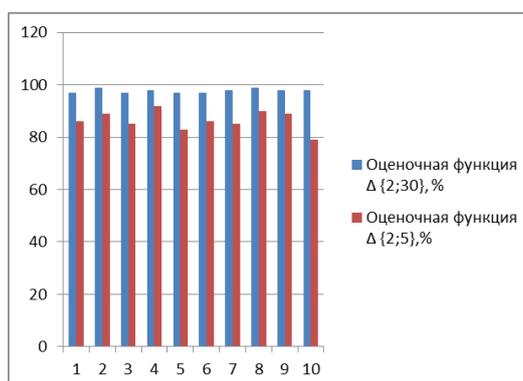


Рис. 4. Гистограмма сравнения эффективности процедуры онтологического анализа при глубине «туманного» слоя 2 до 5 хопов ($\Delta\{2;5\}$) и от 2 до 30 хопов ($\Delta\{2;30\}$) в процентах, $F=100$

Как видно из табл. 1 и рис. 4, наилучший результат улучшения целевой функции для глубины «туманного» слоя от 2 до 5 хопов составил 92 %, для глубины «туманного» слоя от 2 до 30 хопов – 99 % при количестве доступных узлов в «туманном» слое $F=100$.

Таблица 2

Сравнение эффективности процедуры онтологического анализа при глубине «туманного» слоя 2 до 5 хопов ($\Delta\{2;5\}$) и от 2 до 30 хопов ($\Delta\{2;30\}$) в процентах, $F=1000$

Номер эксперимента	Количество доступных вычислительных узлов, F	Количество узлов после процедуры онтологического анализа, F_s для глубины «туманного» слоя $\{2;5\}$	Количество узлов после процедуры онтологического анализа, F_s для глубины «туманного» слоя $\{2;30\}$	$\Delta\{2;5\}$, %	$\Delta\{2;30\}$, %
1	1000	138	15	86,2	98,5
2	1000	135	16	86,5	98,4
3	1000	119	13	88,1	98,7
4	1000	144	21	85,6	97,9
5	1000	149	10	85,1	99
6	1000	142	16	85,8	98,4
7	1000	139	15	86,1	98,5
8	1000	160	14	84	98,6
9	1000	152	28	84,8	97,2
10	1000	141	19	85,9	98,1

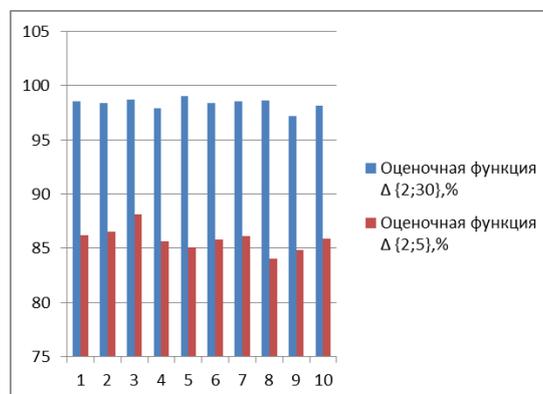


Рис. 5. Гистограмма сравнения эффективности процедуры онтологического анализа при глубине «туманного» слоя 2 до 5 хопов ($\Delta\{2;5\}$) и от 2 до 30 хопов ($\Delta\{2;30\}$) в процентах, $F=1000$

Как видно из табл. 2 и рис. 5, наилучший результат улучшения целевой функции для глубины «туманного» слоя от 2 до 5 хопов составил 88,1 %, для глубины «туманного» слоя от 2 до 30 хопов – 99 % при количестве доступных узлов в «туманном» слое $F=1000$.

Проведя анализ полученных результатов (табл. 1, 2 рис. 4, 5) можно сделать следующие выводы:

1. Наилучшие результаты получены при глубине «туманного» слоя в интервале от 2 до 30 хопов, при этом достигается улучшение целевой функции на 99 %.

2. Эффективность процедуры онтологического анализа зависит от глубины «туманного» слоя: чем глубже «туман», тем меньше времени требуется на решение задачи переноса вычислительной нагрузки в РСАПР.

Заключение. Актуальность разработки новых подходов и методов решения задачи переноса вычислительной нагрузки обусловлено развитием технологий распределенных вычислений, в том числе «туманной» концепции. В работе приведена постановка задачи переноса вычислительной нагрузки с учетом особенностей «туманной» среды. Предложен онтологический подход к решению данной задачи, основой которого является онтологический анализ. Разработана предметная онтология как ключевой компонент онтологического анализа, на основе которого разработан программный модуль в среде Python 3.8.3 и проведены вычислительные эксперименты.

Научная новизна предложенного подхода заключается в использовании онтологического анализа для решения задачи переноса вычислительной нагрузки в РСАПР, функционирующей в «туманной» среде. Это позволяет «отсеять» узлы-кандидаты, принадлежащие «туманному» слою, заведомо не отвечающие требованиям к ресурсам, необходимым для размещения вычислительной нагрузки. Принципиальным отличием предложенного подхода является использование знаний об используемых в РСАПР параллельных алгоритмах для решения задачи переноса вычислительной нагрузки.

Вычислительные эксперименты показали возможность улучшения по времени в случаях с различной глубиной «туманного» слоя: для интервала от 2 до 5 хопов улучшение целевой функции составило 92 % при $F=100$ и 88,1 % при $F=1000$; для интервала от 2 до 30 хопов улучшение целевой функции составило 99 % в обоих случаях.

Проведенный вычислительный эксперимент показал, что проведение процедуры онтологического анализа позволяет улучшить эффективность решения задачи переноса вычислительной нагрузки на 99%.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Кузюрин Н.Н., Грушин Д.А., Фомин С.А. Проблемы двумерной упаковки и задачи оптимизации в распределенных вычислительных системах // Тр. ИСП РАН. – 2014. – № 1. – С.483-501.
2. Жук. С.Н. О построении расписаний выполнения параллельных задач на группах кластеров с различной производительностью // Тр. Института системного программирования РАН. – 2012. – Т. 23. – С. 447-454.
3. Мельник Э.В., Клименко А.Б., Иванов Д.Я. Модель задачи распределения вычислительной нагрузки для информационно-управляющих систем на базе концепции туманных вычислений // Известия ТулГУ. Технические науки. – 2018. – № 2. – С. 174-187.
4. Melnik E.V., Klimenko A.B. A Novel Approach to the Reconfigurable Distributed Information and Control Systems Load-Balancing Improvement // 11th International Conference on Application of Information and Communication Technologies (AICT) – 2017. – P. 355-359.
5. Cisco Fog Computing Solutions: Unleash the Power of the Internet of Things. – Режим доступа: <https://docplayer.net/19735338-Unleashing-the-power-of-the-internet-of-things.html> (дата обращения: 16.07.2020).
6. Глушань В.М., Лаврик П.В. Распределенные САПР. Архитектура и возможности. – Старый Оскол: ТНТ, 2015. – 187 с.
7. Дмитриевич Г.Д., Анисимов Д.А. Построение систем автоматизированного проектирования на основе сервис-ориентированной архитектуры // Известия СПбГЭТУ «ЛЭТИ». – 2014. – № 2. – С. 14-18.
8. Облачные и интерактивные САПР. – Режим доступа: <https://www.autodesk.ru/solutions/cloud-based-online-cad-software> (дата обращения: 16.07.2020).
9. AWS IoT Greengrass. – Режим доступа: <https://aws.amazon.com/ru/greengrass/> (дата обращения: 16.07.2020).

10. Глушань В.М., Иванько Р.В. Анализ эффективности распределенных САПР // Известия ТРТУ. – 2006. – № 8 (63). – С. 115-120.
11. Chiang M., Balasubramanian B., Bonomi F. Fog for 5G and IoT. – Wiley, 2017. – 305 p.
12. Kalyaev I., Melnik E., Klimenko A. A Technique of Adaptation of the Workload Distribution Problem Model for the Fog-Computing Environment // Advances in Intelligent Systems and Computing. – 2019. – Vol. 986.
13. Гузик В.Ф., Каляев И.А., Левин И.И. Реконфигурируемые вычислительные системы: учебное пособие / под общ. ред. И.А. Каляева. – Таганрог: Изд-во ЮФУ, 2016. – 472 с.
14. Melnik E.V., Klimenko A.B., Ivanov D.Ya. Модель задачи формирования сообществ устройств информационно-управляющих систем в средах туманных вычислений // XIII Всероссийское совещание по проблемам управления ВСПУ-2019. – Тула: Тульский государственный университет, 2018. – Вып. 2.
15. Цуканова Н.И. Онтологическая модель представления и организации знаний: учеб. пособие для вузов. – М.: Горячая линия – Телеком, 2014. – 272 с.
16. Курейчик В.М., Сафроненкова И.Б. Создание онтологической модели систем автоматизированного проектирования в среде Protege 4.2 // Проблемы разработки перспективных микро- и наноэлектронных систем – 2016: Сб. трудов / под общ. ред. Академика РАН А.Л. Стемповского. – М.: ИПИМ РАН, 2016. – Ч. III. – С. 240-246.
17. Kureichik V., Safronenkova I.: Ontology-Based Decision Support System for the Choice of Problem-Solving Procedure of Commutation Circuit Partitioning // Creativity in Intelligent Technologies and Data Science. – 2017. – Vol. 754.
18. Гладков Л.А., Курейчик В.В., Курейчик В.М., Сороколетов П.В. Биоинспирированные методы в оптимизации. – М.: Физматлит, 2009. – 384 с.
19. Карпенко А.П. Современные алгоритмы поисковой оптимизации. Алгоритмы, вдохновленные природой: учеб. пособие. – М.: Изд-во МГТУ им. Н.Э. Баумана, 2014. – 446 с.
20. Свидетельство о государственной регистрации программы для ЭВМ 2020616523 Российская Федерация. Модуль онтологического анализа для распределенной системы автоматизированного проектирования / И.Б. Сафроненкова, Э.В. Мельник, В.М. Курейчик; заявители и правообладатели Федеральное государственное бюджетное учреждение науки «Федеральный исследовательский центр Южный научный центр Российской академии наук» (ЮНЦ РАН), Сафроненкова Ирина Борисовна. – No 2020614892; заявл. 18.05.2020; опубл 18.06.2020. – 1 с.

REFERENCES

1. Kuzyurin N.N., Grushin D.A., Fomin S.A. Problemy dvumernoy upakovki i zadachi optimizatsii v raspredelennykh vychislitel'nykh sistemakh [Two-dimensional packing problems and optimization in distributed computing systems], *Tr. ISP RAN* [Proceedings of the Institute for System Programming of the RAS], 2014, No. 1, pp. 483-501.
2. Zhuk. S.N. O postroenii raspisaniy vypolneniya parallel'nykh zadach na gruppakh klasterov s razlichnoy proizvoditel'nost'yu [On-line algorithm for scheduling parallel tasks on a group of related clusters], *Tr. Instituta sistemnogo programirovaniya RAN* [Proceedings of the Institute for System Programming of the RAS], 2012, Vol. 23, pp. 447-454.
3. Mel'nik E.V., Klimenko A.B., Ivanov D.Ya. Model' zadachi raspredeleniya vychislitel'noy nagruzki dlya informatsionno-upravlyayushchikh sistem na baze kontseptsii tumannykh vychisleniy [Model of the problem of distribution of computing load for information-control systems on the basis of the concept of miscalculation], *Izvestiya TulGU. Tekhnicheskie nauki* [Izvestiya Tula State University], 2018, No. 2, pp. 174-187.
4. Melnik E.V., Klimenko A.B. A Novel Approach to the Reconfigurable Distributed Information and Control Systems Load-Balancing Improvement, *11th International Conference on Application of Information and Communication Technologies (AICT) – 2017*, pp. 355-359.
5. Cisco Fog Computing Solutions: Unleash the Power of the Internet of Things. Available at: <https://docplayer.net/19735338-Unleashing-the-power-of-the-internet-of-things.html> (accessed 16 July 2020).

6. *Glushan' V.M., Lavrik P.V.* Raspredelelynye SAPR. Arkhitektura i vozmozhnosti [Distributed CAD systems. Architecture and possibilities]. Staryy Oskol: TNT, 2015, 187 p.
7. *Dmitrevich G.D., Anisimov D.A.* Postroenie sistem avtomatizirovannogo proektirovaniya na osnove servis-orientirovannoy arkhitektury [Computer-aided design systems engineering based on service-oriented architecture], *Izvestiya SPbGETU «LETI»* [Izvestia SPbETU "LETI"], 2014, No. 2, pp. 14-18.
8. Oblachnye i interaktivnye SAPR [Cloud and interactive CAD systems]. Available at: <https://www.autodesk.ru/solutions/cloud-based-online-cad-software> (accessed 16 July 2020).
9. AWS IoT Greengrass. Available at: <https://aws.amazon.com/ru/greengrass/> (accessed 16 July 2020).
10. *Glushan' V.M., Ivan'ko R.V.* Analiz effektivnosti raspredelelynykh SAPR [The analysis of CAD system effectiveness], *Izvestiya TRTU* [Izvestiya TSURE], 2006, No. 8 (63), pp. 115-120.
11. *Chiang M., Balasubramanian B., Bonomi F.* Fog for 5G and IoT. Wiley, 2017, 305 p.
12. *Kalyaev I., Melnik E., Klimenko A.* A Technique of Adaptation of the Workload Distribution Problem Model for the Fog-Computing Environment, *Advances in Intelligent Systems and Computing*, 2019, Vol. 986.
13. *Guzik V.F., Kalyaev I.A., Levin I.I.* Rekonfiguriruemye vychislitel'nye sistemy: uchebnoe posobie [Reconfigurable computer systems], ed. by I.A. Kalyaeva. Taganrog: Izd-vo YuFU, 2016, 472 p.
14. *Melnik E.V., Klimenko A.B., Ivanov D.Ya.* Model' zadachi formirovaniya soobshchestv ustroystv informatsionno-upravlyayushchikh sistem v sredakh tumannykh vychisleniy [A model of local group device of information and management system forming in the fog-environmental], *XIII Vserossiyskoe soveshchanie po problemam upravleniya VSPU-2019* [Proceedings of the XIII VSPU-2019.]. Tula: Tul'skiy gosudarstvennyy universitet, 2018, Issue 2.
15. *Tsukanova N.I.* Ontologicheskaya model' predstavleniya i organizatsii znaniy: ucheb. posobie dlya vuzov [An ontological model of knowledge representation and management: study guide for universities]. Moscow: Goryachaya liniya – Telekom, 2014, 272 p.
16. *Kureychik V.M., Safronenkova I.B.* Sozdanie ontologicheskoy modeli sistem avtomatizirovannogo proektirovaniya v srede Protege 4.2 [Creation of CAD-systems ontology using Protege 4.2] *Problemy razrabotki perspektivnykh mikro- i nanoelektronnykh sistem – 2016: Sb. trudov* [All-Russia Science & Technology Conference "Problems of Advanced Micro- and Nanoelectronic Systems Development"], ed. by Akademiya RAN A.L. Stempovskogo. Moscow: IPPM RAN, 2016, Part III, pp. 240-246.
17. *Kureichik V., Safronenkova I.*: Ontology-Based Decision Support System for the Choice of Problem-Solving Procedure of Commutation Circuit Partitioning, *Creativity in Intelligent Technologies and Data Science*, 2017, Vol. 754.
18. *Gladkov L.A., Kureychik V.V., Kureychik V.M., Sorokoletov P.V.* Bioinspirirovannyye metody v optimizatsii [Bio-inspired methods in optimization]. Moscow: Fizmatlit, 2009, 384 p.
19. *Karpenko A.P.* Sovremennyye algoritmy poiskovoy optimizatsii. Algoritmy, vdokhnovlennyye prirodoy: ucheb. posobie [Contemporary algorithms of search engine optimization. Bio-inspired algorithms: textbook]. Moscow: Izd-vo MGTU im. N.E. Baumana, 2014, 446 p.
20. *Safronenkova I.B., Mel'nik E.V., Kureychik V.M.* Svidetel'stvo o gosudarstvennoy registratsii programmy dlya EVM 2020616523 Rossiyskaya Federatsiya. Modul' ontologicheskogo analiza dlya raspredelelynoy sistemy avtomatizirovannogo proektirovaniya [State registration certificate 2020616523 Russian Federation. A module of ontological analysis for distributed computer-aided design system]; applicants and copyright holders Federal state budgetary institution of science "Federal research center southern scientific center of the Russian Academy of Sciences" (UNC RAS), Safronenkova Irina Borisovna, No 2020614892; declared 18.05.2020; published 18.06.2020, 1 p.

Статью рекомендовал к опубликованию д.т.н. С.Г. Капустян.

Курейчик Виктор Михайлович – Южный федеральный университет; e-mail: vmkureychik@sfedu.ru; 347928, г. Таганрог, пер. Некрасовский, 44; тел.: +78634311487; кафедра САПР; д.т.н.; профессор.

Сафроненкова Ирина Борисовна – Федеральный исследовательский центр Южный научный центр Российской академии наук; e-mail: safronenkova050788@yandex.ru; г. Ростов-на-Дону, ул. Чехова, 41; тел.: +79604678753; м.н.с.

Kureichik Viktor Mikhailovich – Southern Federal University, e-mail: vmkureychik@sfedu.ru; 44, Nekrasovskiy, Taganrog, 347928, Russia; phone: +78634311487; the department of computer aided design; dr. of eng. sc.; professor.

Safronenkova Irina Borisovna – Federal Research Centre The Southern Scientific Centre of the Russian Academy of Sciences; e-mail: safronenkova050788@yandex.ru; 41, Chekhov street, Rostov-on-Don, 344006, Russia; phone: +79604678753; junior researcher.

УДК 004.896

DOI 10.18522/2311-3103-2020-4-82-93

Б.К. Лебедев, В.Б. Лебедев, О.Б. Лебедев**ПОПУЛЯЦИОННЫЙ АЛГОРИТМ ПОСТРОЕНИЯ ДЕРЕВА РЕШЕНИЙ
МЕТОДОМ КРИСТАЛЛИЗАЦИИ РОССЫПИ АЛЬТЕРНАТИВ***

В ряде случаев возникает необходимость установления соответствия между заявленным и фактическим значением категориальной переменной на основе совокупности признаков объекта. В этом случае возникает потребность в классификаторе с оптимальной последовательностью рассматриваемых атрибутов с заданным значением целевой функции. Значением целевой переменной может быть: да, нет, номер сорта, номер класса и т.д. В работе решается задача построения классификационной модели в виде оптимальной последовательности рассматриваемых атрибутов и их значений, входящих в состав маршрута от корневой вершины к концевой вершине с заданным значением целевой переменной. Если требуется классификатор, включающий возможность альтернативных ответов, то вначале строятся независимо друг от друга оптимальные маршруты для каждого значения целевой переменной, а затем эти маршруты объединяются («склеиваются») в единое бинарное дерево решений. В алгоритме построения классификатора на основе метода кристаллизации россыпи альтернатив, каждое решение Q_k интерпретируется в виде ориентированного маршрута M_k на бинарном дереве решений. Назовем порядковый номер элемента в ориентированном маршруте M_k позицией $s_i \in S = \{s_i | i=1, 2, \dots, n_A\}$. Элементом маршрута M_k является пара (x_i, u_i) , где x_i соответствует A_i , u_i в маршруте M_k является ребром, выходящим из x_i и соответствует выбранному вместе с A_i значению A_i . Второй индекс элемента u_i определится после выбора A_i , помещенного в соседнюю с s_i позицию s_{j+1} . Работа алгоритма построения дерева решений базируется на использовании коллективной эволюционной памяти, под которой подразумевается информация, отражающая историю поиска решения. Алгоритм учитывает тенденции к использованию альтернатив из наилучших найденных решений. Особенности являются наличие непрямого обмена информацией – стигмержи. Совокупность данных об альтернативах и их оценках составляет россыпь альтернатив. Рассмотрены ключевые моменты анализа альтернатив в процессе эволюционной коллективной адаптации. Экспериментальные исследования показали, что разработанный алгоритм находит решения, не уступающие по качеству, а иногда и превосходящие своих аналогов в среднем на 3–4 %. Временная сложность алгоритма, полученная экспериментальным путем, лежит в пределах $O(n^2)$ – $O(n^3)$.

Классификация; дерево решений; оптимизация; популяционный алгоритм; адаптивное поведение; метод кристаллизации россыпи альтернатив.

В.К. Lebedev, V.B. Lebedev, O.B. Lebedev**POPULATION ALGORITHM FOR CONSTRUCTING A TREE
OF SOLUTIONS BY METHOD OF CRYSTALLIZATION OF ALTERNATIVES
FIELD**

In some cases, it becomes necessary to establish a correspondence between the declared and actual value of a categorical variable on the basis of a set of object characteristics. In this case, there is a need for a classifier with an optimal sequence of the considered attributes with a

* Работа выполнена при финансовой поддержке гранта РФФИ № 20–07–00260 А.