

Самигулина Г.А.¹, Масимканова Ж.А.¹

¹Институт информационных и вычислительных технологий,
г. Алматы, Казахстан

ПОСТРОЕНИЕ ОПТИМАЛЬНОГО НАБОРА ДАННЫХ НА ОСНОВЕ МОДИФИЦИРОВАННЫХ АЛГОРИТМОВ МУРАВЬИНЫХ КОЛОНИЙ И ОНТОЛОГИЧЕСКОГО ПОДХОДА ДЛЯ ИММУНОСЕТЕВОГО МОДЕЛИРОВАНИЯ

Аннотация. Статья посвящена актуальной проблеме применения биоинспирированных интеллектуальных подходов для предварительной обработки данных и решения задачи прогнозирования в нефтегазовой отрасли. Цель исследования - построение оптимальной иммуносетевой модели на основе модифицированных алгоритмов муравьиных колоний и искусственных иммунных систем. Выделение информативных дескрипторов осуществляется на основе мультиалгоритмического подхода, который позволяет использовать несколько модифицированных алгоритмов. Для систематизации применяемых методов и облегчения написания программного обеспечения предложены онтологические модели рассматриваемых алгоритмов. Модифицированные алгоритмы муравьиных колоний отличаются между собой наличием дополнительных переменных, различных ограничений, топологией и т.д. Большое значение имеют такие параметры, как размер популяций агентов и количество итераций.

Ключевые слова: модифицированные алгоритмы муравьиных колоний, иммуносетевое моделирование, выделение информативных дескрипторов.

• • •

Түйіндеме. Зерттеу мұнай-газ саласында болжамдау тапсырмасын шешу үшін және деректерді алдын-ала өңдеу үшін биоинспирленген интеллектуалдық әдістерді қолданудың маңызды мәселесіне арналған. Зерттеудің мақсаты – жасанды иммундық жүйелер мен модификацияланған құмырсқа алгоритмдері негізінде оңтайлы иммундық желілік модельді құрастыру болып табылады. Ақпараттық дескрипторларды белгілеу бірнеше модификацияланған алгоритмдерді қолдануға мүмкіндік беретін мультиалгоритмдік әдіс негізінде жүзеге асырылады. Қолданылатын әдістерді жүйелеу және бағдарламалық камтамасыз етуді құрастыруды жеңілдету мақсатында қарастырылған алгоритмдердің он-

тологиялық модельдері ұсынылды. Құмырсқа илеуінің модификацияланған алгоритмдері бір-бірінен қосымша айнаымалылардың, әртүрлі шектеулердің, топологияның және т.б. болуымен ерекшеленеді. Агенттер популяциясының көлемі және итерациялардың саны сияқты параметрлер маңызды болып табылады.

Түйінді сөздер: құмырсқа илеуінің модификацияланған алгоритмдері, иммундық желілік модельдеу, ақпараттық дескрипторларды белгілеу.

• • •

Abstract. The article is devoted to the actual problem of application of bioinspired intellectual approaches for preliminary data processing and solving the problem of prediction in the oil and gas industry. The aim of the research is to construct an optimal immune network model based on modified algorithms of ant colonies and artificial immune systems. The selection of informative descriptors is carried out based on a multi-algorithm approach, which allows to use of several modified algorithms. The ontological models of considered algorithms have been proposed for the systematization of used methods and the facilitation of writing of software. Modified algorithms of ant colony differ by the presence of additional variables, various constraints, topology, etc. Such parameters as the size of populations of agents and the amount of iterations have great importance.

Keywords: modified algorithms of ant colonies, immune network modeling, the selection of informative descriptors.

Введение. В настоящее время актуальной задачей является применение последних достижений искусственного интеллекта при создании современных информационных технологий прогнозирования. Активно разрабатываются инновационные информационные системы на основе биоинспирированных интеллектуальных подходов для обработки и анализа больших данных и решения задач прогнозирования.

Искусственный интеллект широко применяется в нефтегазовой отрасли, робототехнике, банковских системах, образовании и медицине. Большие перспективы имеет модернизация нефтегазовой отрасли на основе эффективного внедрения инновационных и интеллектуальных технологий. Особое внимание уделяется этой проблеме в крупных международных нефтегазовых компаниях таких как: Shell, Chevron, Saudi Aramco, Petrobras, Kuwait Oil и др. Применение интеллектуальных методов в нефтегазовой промышленности позволяет улучшить точность прогнозирования разведки и добычи нефти. Наибольшее

Работа выполнена по гранту КН МОН РК по теме: «Разработка когнитивной Smart-технологии для интеллектуальных систем управления сложными объектами на основе подходов искусственного интеллекта» (2018-2020 гг.).

распространение получили интеллектуальные технологии с использованием нейронных сетей, эволюционных алгоритмов, искусственных иммунных систем, алгоритмов роевого интеллекта и др. Нейронные сети (НС) широко применяются при решении геотехнических задач, которые позволяют сократить количество скважин, а также проводимых исследований для определения свойств грунта [1]. Нейронные сети обладают способностью обработки больших данных, позволяют адаптироваться к изменяющимся условиям, обобщать и обучаться [2].

Активно развиваются алгоритмы роевого интеллекта, такие как муравьиные и пчелиные колонии, метод роя частиц, алгоритм перемещения бактерий, серых волков и многие другие. Данные подходы успешно применяются при решении оптимизационных задач. Они отличаются способностью быстро и достаточно точно решать проблему сложного поиска. Алгоритмы роевого интеллекта широко используются для построения оптимального набора дескрипторов и исключения малоинформативных данных. Существует множество работ по данной тематике. В работе [3] используется алгоритм муравьиной колонии (ACO) для определения оптимального количества газа в скважинах для трех месторождений. Предложенный алгоритм по полученным результатам показывает лучшее распределение газа в скважинах по сравнению с другими методами оптимизации. В статье [4] исследуется совместное применение методов роя частиц (PSO) и опорных векторов (SVR) для формирования оптимального набора гиперпараметров. В работе PSO-SVR модель сравнивается с SVR-моделями, полученными методом случайного поиска (RAND-SVR) и методом проб и ошибок (TE-SVR). Сравнение осуществляется с использованием реальных промышленных наборов данных, полученных при добычи нефти с четырех различных нефтяных скважин. Моделирование показывает, что PSO-SVR модель демонстрирует наилучшие результаты. В исследовании [5] представляется метод роя частиц (PSO) для прогнозирования, который позволяет минимизировать общую стоимость работы скважин. Метод роя частиц находит оптимальную комбинацию параметров бурения, а также оптимальную глубину выталкивания. Данный алгоритм может быть применен для планирования новых скважин или в качестве симулятора бурения. В работе [6] предлагается пчелиный алгоритм для решения задач размещения объектов в нефтегазовой промышленности, позволяющий за короткое время находить оптимальное решение с заданной точностью. Алгоритм основан на децентрализованном поведении интеллектуальных агентов, которые представляют собой самоорганизующуюся систему.

Интересны разработки на основе применения Искусственных Иммунных Систем (ИИС). В статье [7] представлены алгоритмы искусственных иммунных систем (CLONALG и Parallel AIRS2) для распознавания образов при классификации этапов и параметров бурения. Полученные результаты сравнивались с другими методами такими как: нейронные сети и метод опорных векторов.

Актуальна разработка онтологических моделей для анализа применяемых алгоритмов, структуризации данных и создания программного обеспечения. Применение современных онтологических редакторов и разработка онтологических моделей облегчает решение задачи выделения информативных дескрипторов и построение оптимальной модели [8]. Например, работа [9] посвящена построению OWL (Web Ontology Language) модели муравьиного алгоритма для решения задачи выбора входных параметров.

Постановка задачи. Необходимо решить задачу выделения информативных дескрипторов на основе модифицированных алгоритмов муравьиных колоний и онтологического подхода для построения оптимального набора данных и дальнейшего прогнозирования с использованием иммунносетевого моделирования.

Методы исследования. Не существует универсальных алгоритмов для предварительной обработки данных. Отбор информативных дескрипторов осуществляется на основе модифицированных алгоритмов муравьиных колоний согласно концепции мультиалгоритмического подхода [10,11], при котором используется несколько модифицированных алгоритмов.

В данном случае рассматривается классический алгоритм муравьиной колонии, основанный на моделировании поведения муравьев в процессе решения оптимизационных задач. При решении задачи выделения информативных дескрипторов набор данных представляется в виде графа, узлы которого являются дескрипторами [12]. Сначала задается размер популяции и количество итераций. Затем каждый муравей-агент устанавливается в какую-то начальную точку. Задается количество феромона, которое пропорционально проценту ошибок, полученному при классификации дескрипторов. Следующий дескриптор выбирается в зависимости от количества феромона на гранях. На каждом шаге происходит испарение феромона и при каждой итерации выбирается набор дескрипторов с минимальной ошибкой. После прохождения требуемого количества итераций завершается алгоритм. В таком подходе предполагается, что существует какой-то

оптимальный набор дескрипторов, на котором классификаторы дают минимальный процент ошибки и подграф, на гранях которого будет максимальное количество феромона. На гранях, соединяющих неинформативные дескрипторы феромона должно быть минимальное количество. Затем муравей-агент составляет набор дескрипторов, который предлагается использовать для построения иммунносетевой модели.

Алгоритм муравьиных колоний имеет множество модификаций, например: Elitist ant system (EAS), Ant Colony System (ACS), Max-min ant system (MMAS), Ant-Q, AntSrank, Continuous ant colony optimization (CACO), Continuous Interacting Ant Colony (CIAC), Direct ant colony optimization (DACO), ACO extended to continuous domains и др. Модификация EAS [13] отличается введением в алгоритм «элитных муравьев-агентов» для определения кратчайшего пути из выбранных путей. В алгоритме ACS [14] изменяется время обновления феромона. Модификация MMAS [15] отличается вводом ограничений на максимальное и минимальное количество феромона на грани. Алгоритм CACO [16] основан на локальном поиске вокруг гнезда. Основным отличием является пространство поиска, в котором существует основная точка, из которой исходят направления поиска в виде векторов. На каждой следующей итерации муравей-агент выбирает одно из направлений по количеству феромона. В модифицированном алгоритме CIAC используются двойные каналы связи для обмена информацией. В алгоритме DACO определяется количество феромона и формируются правила испарения феромона.

Результаты исследования. Разработана общая онтологическая (OWL) модель, состоящая из онтологических моделей модифицированных алгоритмов муравьиных колоний (Таблица), которая представляется в виде:

$$OM_{ACO} = \langle OM_{BasicACO}, OM_{EAS}, OM_{ACS}, OM_{MMAS}, OM_{CACO} \rangle,$$

где: OM_{ACO} - OWL модель алгоритмов муравьиных колоний;

$OM_{BasicACO}$ - OWL модель классического алгоритма муравьиной колонии;

OM_{EAS} - OWL модель алгоритма Elitist ant system;

OM_{ACS} - OWL модель алгоритма Ant Colony System;

OM_{MMAS} - OWL модель алгоритма Max-min ant system;

OM_{CACO} - OWL модель алгоритма Continuous ant colony optimization.

Таблица - Онтологические модели модифицированных алгоритмов муравьиных колоний

Онтологические модели алгоритмов муравьиных колоний	Описание онтологической модели
1	2
Онтологическая модель классического алгоритма муравьиной колонии	<p>OWL модель классического алгоритма муравьиной колонии:</p> <ul style="list-style-type: none"> - создание популяции агентов; - произвольное разделение агентов между узлами; - вычисление фитнес-функций и определение количества феромона; - перемещение агентов; - испарение феромона; - обновление локального и глобального количества феромона; - проверка условия завершения; - определение веса грани и сохранение лучшего положения агента; - построение оптимальной иммуносетевой модели на основе выделенного набора дескрипторов.
Онтологическая модель алгоритма Elitist ant system	<p>OWL модель алгоритма EAS:</p> <ul style="list-style-type: none"> - инициализация популяции агентов; - построение соседства агентов; - распределение агентов; - вычисление фитнес-функций; - сохранение лучшего решения агентов; - обновление пространства поиска; - случайное исследование соседства агентов для нахождения лучшего глобального решения; - обновление феромона; - испарение феромона; - генерация новой популяции элитных агентов; - обновление локального и глобального лучшего решения; - построение набора лучших решений.
Онтологическая модель алгоритма Ant Colony System	<p>OWL модель алгоритма ACS:</p> <ul style="list-style-type: none"> - создание популяции агентов; - произвольное разделение агентов между узлами; - равномерное распределение феромона; - вычисление фитнес-функций; - анализ конечных решений и определение частоты обновления феромона; - перемещение агентов; - испарение феромона; - обновление локального и глобального лучшего решения;

1	2
	<ul style="list-style-type: none"> - проверка условия завершения; - сохранение лучшего положения агента; - построение оптимального набора дескрипторов.
<p>Онтологическая модель алгоритма Max-min ant system.</p>	<p style="text-align: center;">OWL модель алгоритма MMAS:</p> <ul style="list-style-type: none"> - инициализация минимального и максимального количества феромона, размера популяции, количества итераций; - сохранение минимального количества феромона на каждой грани; - распределение агентов между узлами; - вычисление фитнес-функций; - сохранение лучшего локального решения; - обновление количества феромона; - перемещение агентов; - испарение феромона; - обновление локального и глобального количества феромона; - обновление лучшего глобального решения; - проверка условия завершения; - сохранение глобального решения; - построение оптимального набора дескрипторов из выбранных решений.
<p>Онтологическая модель алгоритма Continuous ant colony optimization.</p>	<p style="text-align: center;">OWL модель алгоритмаCASO:</p> <ul style="list-style-type: none"> - создание агентов; - установка начального пространства поиска; - установка количества феромона и скорости испарения; - создание дополнительных переменных; - вычисление фитнес-функций; - обновление локального решения; - обновление местоположения до начальных координат; - обновление количества феромона; - обновление пространства поиска.
	<ul style="list-style-type: none"> - обновление глобального решения; - испарение феромона; - обновление местоположения до текущих координат; - выбор локального направления поиска; - обновление локального решения; - сравнение локального и глобального решения; - обновление глобального решения; - проверка условия завершения; - сохранение лучших координат; - построение оптимального набора дескрипторов.

Модифицированные алгоритмы муравьиных колоний отличаются между собой наличием дополнительных переменных, ограничени-

ем на максимальное и минимальное количество феромона, топологией построения пространства поиска и др.

Заключение. Таким образом, разработанная технология на основе модифицированных алгоритмов муравьиной колонии позволяет осуществлять предварительную обработку данных и формировать оптимальный набор дескрипторов для дальнейшего иммуносетевого моделирования объектов нефтегазовой отрасли. Применение мультиалгоритмического подхода требует систематизации используемых алгоритмов. Создание онтологических моделей позволяет анализировать и структурировать данные, а также экономить время при разработке программного обеспечения.

Список литературы

1 *Abouzar M., Soroor S.* The Application of Artificial Neural Networks for the Prediction of Oil Production Flow Rate // *Energy Sources, Part A: Recovery Utilization and Environmental Effects.* - 2012. - №34 (19). - С.1834-1843.

2 *Казначеев П.Ф., Самойлова Р.В., Курчиски Н.В.* Применение методов искусственного интеллекта для повышения эффективности в нефтегазовой и других сырьевых отраслях // *Экономика и экономические науки.* - 2016.-Т. 11. № 5. - С. 188-197

3 *Ghaedi M., Ghotbi C., Aminshahidy B.* The optimization of gas allocation to a group of wells in a gas lift using an efficient Ant Colony Algorithm (ACO) // *Energy Sources, Part A: Recovery, Utilization and Environmental Effects.* - 2014. - Vol. 36, Issue 11. - С. 1234-1248.

4 *Akande K.O., Owolabi T.O., Olatunji S.O., Abdul Raheem, A.* A hybrid particle swarm optimization and support vector regression model for modelling permeability prediction of hydrocarbon reservoir // *Journal of Petroleum Science and Engineering.* - 2017. - Vol. 150. - С.43-53.

5 *Self R., Atashnezhad A., Hareland G.* *Reducing*. Drilling Cost by finding Optimal Operational Parameters using Particle Swarm Algorithm // *Proc. of Society of Petroleum Engineers Source SPE Deepwater Drilling and Completions Conference.* URL: <https://www.researchgate.net/publication/304998993> (дата обращения: 04/03/2018)

6 *Абрамов А.С.* Методы искусственного интеллекта для решения задачи оптимального размещения кустовых площадок и распределения скважин между ними. - 2016. - №4 (736). - С. 25-29.

7 *Serapiao A.B.S., Mendes J.R.P., Miura K.* Artificial Immune Systems for Classification of Petroleum Well Drilling Operations // *Artificial Immune*

Systems. - Springer- Verlag, Heidelberg, 2007. - Vol. 4628. - С. 47-58.

8 *Самигулина Г.А., Масимканова Ж.А.* Онтологические модели алгоритмов роевого интеллекта для иммунносетевого моделирования лекарственных препаратов // Вестник КазНУ им. Аль-Фараби (Journal of Mathematics, Mechanics and Computer Science). - 2017. - №1(93). - С.92-104.

9 *Zong N., Zheng Y., Wang Z.* A SWRL Rules Parser Algorithm Based on the Topic Semantic Comparability and Ant Colony Algorithm // Fuzzy Systems and Knowledge Discovery. - IEEE, 2009. DOI: 10.1109/FSKD.2009.283

10 *Samigulina G.A.* Immune network modeling technology for complex objects intellectual control and forecasting system: monograph. Science Book Publishing House. - USA, 2015. - 172 с.

11 *Самигулина Г.А., Самигулина З.И.* Построение оптимальной иммунносетевой модели для прогнозирования свойств неизвестных лекарственных соединений на основе мультиалгоритмического подхода // Проблемы информатики. - Новосибирск, 2013. - №2. - С. 21 -29.

12 *Chen B., Chen L., Chen Y.* Efficient ant colony optimization for image feature selection // Signal Processing. - Elsevier, 2013. - Vol. 93, Issue 6. - С. 1566-1576.

13 *Negulescu S.C., Oprean C., Kifor C.V., Carabulea I.* Elitist Ant System for Route Allocation Problem // 8th WSEAS International Conference on APPLIED INFORMATICS AND COMMUNICATIONS (AIC'08). - Rhodes, Greece, 2008. - С. 62-67.

14 *Yoo K., Han S.* A modified ant colony optimization algorithm for dynamic topology optimization // Computers & Structures. - Elsevier, 2013. - Vol. 123. - С. 68-78.

15 *Crawford B., Soto R., Johnson F., Monfroy E., Paredes F.* A Max-Min Ant System algorithm to solve the Software Project Scheduling Problem // Expert Systems with Applications. - №41. - 2014. - С. 6634-6645.

16 *Aidov A., Dulikravich G.S.* Modified continuous ant colony algorithm // Proceedings of 2nd Intern. Congress of Serbian Society of Mechanics. - Serbia, 2009. URL: <https://www.researchgate.net/publication/228460180> (дата обращения: 11/03/2018).

Самигулина Г.А., доктор технических наук, доцент,
e-mail: galinasamigulina@mail.ru

Масимканова Ж.А., магистр, e-mail: masimkanovazh@gmail.com