

УДК 51-76

Онтологические модели алгоритмов роевого интеллекта для иммуносетевого моделирования лекарственных препаратов

Самигулина Г.А., д.т.н, доцент, заведующий лабораторией,
Институт информационных и вычислительных технологий КН МОН РК,
г. Алматы, Республика Казахстан, E-mail: galinasamigulina@mail.ru
Масимканова Ж.А., магистрант, Казахский национальный университет имени
аль-Фараби, г. Алматы, Республика Казахстан, +77078346177,
E-mail: masimkanovazh@gmail.com

Статья посвящена разработке информационной системы прогнозирования свойств новых лекарственных препаратов на основе методов роевого интеллекта и искусственных иммунных систем. Важным аспектом при прогнозировании зависимости между химическим строением вещества и их биологической активностью (QSAR, Quantitative Structure-Activity Relationship) является выбор дескрипторов. Перспективным направлением в области QSAR стало применение подходов искусственного интеллекта, которые обеспечивают высокую точность прогнозирования химических соединений с заданными свойствами. В статье рассматриваются методы пчелиной колонии и алгоритм роя частиц для решения задачи выделения информативных дескрипторов и дальнейшего иммуносетевого моделирования фармакологической активности химических соединений. Приведены существующие программные средства реализации данных алгоритмов для построения оптимального набора дескрипторов. Применение мультиалгоритмического подхода при иммуносетевом моделировании лекарств требует систематизации используемых методов и создание интегрированной онтологической модели. Разработка онтологических моделей позволяет структурировать входные и выходные данные, учитывать особенности функционирования и взаимосвязи, экономит временные и вычислительные ресурсы при разработке компонентно-ориентированного программного обеспечения для иммуносетевого моделирования новых химических веществ с заданной фармакологической активностью. Разработаны онтологические модели пчелиной колонии и алгоритма роя частиц для решения задачи выделения информативных дескрипторов в редактор онтологий Protege.

Ключевые слова: компьютерный молекулярный дизайн лекарств, QSAR, роевой интеллект, выделение информативных дескрипторов, онтологические модели.

Дәрілерді иммунды-желілік моделдеу үшін үйір интеллектісінің алгоритмдерінің онтологиялық моделдері

Самигулина Г.А., т.ғ.д., доцент, зертхана жетекшісі, Ақпараттық және есептеуіш технологиялар институты, Алматы, Қазақстан, galinasamigulina@mail.ru
Масимканова Ж.А., магистрант, аль-Фараби атындағы Қазақ ұлттық университеті, Алматы, Қазақстан, masimkanovazh@gmail.com

Мақала жасанды иммундық жүйелер мен үйір интеллектісінің әдістерінің негізінде жаңа дәрі-дәрмектердің қасиеттерін болжаудың ақпараттық жүйесін құрастыруға арналған. Дескрипторларды таңдау заттектердің химиялық құрылымы мен олардың биологиялық белсенділігі арасындағы байланысты (QSAR, Quantitative Structure-Activity Relationship) болжау барысында маңызды көрініс болып табылады. Белгіленген қасиеттері бар химиялық байланыстарды болжаудың жоғары дәлдігін қамтамасыз ететін жасанды интеллект әдістерін қолдану QSAR облысындағы келешекті бағыт болып табылады.

Мақалада ақпараттық дескрипторларды бөлу және ары қарай химиялық байланыстардың фармакологиялық белсенділіктерін иммунды-желілік моделдеу есептерін шешу үшін аралар шоғыры әдісі мен құстар алгоритмі қарастырылды. Дескрипторлардың оңтайлы жинағын құрастыру үшін берілген алгоритмдерді жүзеге асырудың қолда бар құралдары сипатталды. Дәрілерді иммунды-желілік моделдеу барысында мультиалгоритмдік әдісті қолдану пайдаланылатын әдістерді жүйелеуді және кіріктірілген онтологиялық моделін құрастыруды талап етеді. Онтологиялық моделдерді құрастыру белгіленген фармакологиялық белсенділігі бар жаңа химиялық заттектерді иммунды-желілік моделдеу үшін құрамдастық-бағытталған бағдарламалық қамтамасыз етуді құрастыру барысында кіріс және шығыс деректерін жүйелеуге, жұмыс істеу және байланысу ерекшеліктерін ескеруге, уақыт және есептеу ресурстарын азайтуға мүмкіндік береді. Ақпараттық дескрипторларды бөлу есебін шешу үшін аралар шоғыры әдісі мен құстар алгоритмінің онтологиялық моделдері Protege онтология редакторында құрастырылды.

Түйін сөздер: дәрілердің компьютерлік молекулярлық дизайны, QSAR, үйір интеллектісі, ақпараттық дескрипторларды бөлу, онтологиялық моделдер.

The ontological models of swarm intelligence algorithms for immune network modeling of drugs

Samigulina G. A., Dr.Sci. (Tech.), assistant professor, head of laboratory, Institute of Information and Computational Technologies, Almaty, Kazakhstan, galinasamigulina@mail.ru

Massimkanova Zh. A., student of master degree, al-Farabi Kazakh National University, Almaty, Kazakhstan, masimkanovazh@gmail.com

The article is devoted to development of information system of forecasting properties of new drugs based on methods of swarm intelligence and artificial immune systems. An In forecasting relationship between the chemical structure of substances and their biological activity (QSAR, Quantitative Structure-Activity Relationship) an important aspect is feature selection. A perspective direction in QSAR is the use of artificial intelligence, which provide high accuracy of forecasting of chemical compounds with predetermined properties. The article considers the methods of bee colonies and particle swarm algorithm for solving the problem of feature selection and further immune network modeling of pharmacological activity of chemical compounds. The existing software of implementation of these algorithms for constructing of optimum set of descriptors are given. Application of multialgorithmic approach in immune network modeling of drugs requires the systematization of using methods and development of an integrated ontological model. Development of ontological models allows to structure input and output data, to consider features of functioning and relationships, saves time and computational resources in the development of component-based software for immune network modeling of new substances with predetermined pharmacological activity. The ontological models of bee colony and particle swarm algorithm for solving the problem of feature selection are created in ontology editor Protege.

Key words: computer-aided molecular design, QSAR, swarm intelligence, feature selection, ontological model.

1 Введение

При компьютерном молекулярном дизайне лекарств исследование химических соединений связано с обработкой огромных массивов данных, представляющих собой определенную проблему. Решением одной из основных задач фармакологии является изучение зависимости между химическим строением веществ и их биологической активностью, что позволяет значительно сократить сроки разработки лекарственных препаратов и их стоимость. Количественные зависимости структура-свойства/активность (QSAR, Quantitative Structure-Activity Relationship) позволяют выявлять потенциально активные молекулы в базах химических соединений и осуществлять синтез веществ с заранее заданными свойствами. Правильное построение QSAR-модели зависит от качества

исходного набора соединений (Golla S., 2017: 478). При поиске зависимостей структура-активность химических веществ актуальной задачей является выбор оптимального набора дескрипторов на основе которого строится QSAR-модель. Задача предсказания новых соединений с заданными свойствами (Chin Y.L., 2012: 3) решается с использованием интеллектуальных методов, которые позволяют анализировать и предсказывать токсичность, метаболизм, реакционную способность соединений и др. К ним относятся нейронные сети, эволюционные алгоритмы, искусственные иммунные системы, алгоритмы роевого интеллекта и др.

В настоящее время методы роевого интеллекта являются перспективным направлением и широко применяются во многих областях медицины и фармакологии. К методам роевого интеллекта, основанным на моделировании поведения живых существ, относятся: методы муравьиной и пчелиной колонии, алгоритм роя частиц, альтруизма, капель воды, кукушки, серых волков, летучих мышей, гравитационного поиска, светлячков, оптимизация передвижения бактерий и многороевая оптимизация.

Актуально применение новых инновационных подходов при разработке системы прогнозирования зависимости структура-свойства/активность лекарственных препаратов. Одним из перспективных подходов являются искусственные иммунные системы. Они представляют собой сложную адаптивную структуру, эффективно использующую различные механизмы обучения, памяти и ассоциативного поиска для решения прикладных задач. Существует несколько направлений (Dasgupta D., 1998: 438) в искусственных иммунных системах, таких как клональный алгоритм отбора, негативный алгоритм отбора и иммунносетевой алгоритм.

Искусственные иммунные системы имитируют поведение иммунной системы человека в процессе защиты организма от внешних факторов (Samigulina G.A., 2015: 48) и обладают такими достоинствами как память, обучаемость, распределенность, самоорганизация и высокая степень параллелизма. В данной статье при решении задачи отбора информативных дескрипторов рассматриваются методы пчелиной колонии и алгоритмы роя частиц для построения оптимальной иммунносетевой модели.

2 Обзор литературы

При иммунносетевом моделировании особенно актуально применение онтологических моделей, позволяющих систематизировать и структурировать данные (Baranjuk V.V., 2015: 15). Ряд исследований посвящён построению онтологических моделей алгоритмов роевого интеллекта. Статья (Rangel C., 2015: 536) направлена на создание автоматизированной системы сравнения и выбора метода для лечения болезни на основе пчелиного алгоритма. В связи с разнообразием и сложностью вычислений авторами разработана онтологическая модель алгоритма. В исследовании (Yuan F., 2015: 52) изучается связь между симптомами и лечением в медицинских записях. Работа (Shao L., 2012: 222) посвящена разработке онтологической модели алгоритма роя частиц для решения проблемы подобию в наборах данных. В статье (Shi-xiong X., 2011: 6) предлагается система диагностики на основе метода онтологии и алгоритма роя иммунных частиц. Результаты эксперимента показывают, что новый алгоритм диагностики позволяет повысить точность принятия правильных решений.

Создание интеллектуальной технологии прогнозирования зависимости структура-

свойства/активность состоит из нескольких задач (Самигулина Г.А., 2015: 15). Сначала выбираются химические соединения для исследования, затем описываются структуры соединений на основе дескрипторов. После этого выполняется отбор информативных дескрипторов и осуществляется построение оптимального набора дескрипторов для дальнейшего иммунносетевого моделирования. Затем выполняется классификация и осуществляется обучение искусственной иммунной системы. Далее реализуется алгоритм распознавания образов на основе сингулярного разложения матриц (Тараканов А., 2007: 503) для определения класса и выполняется оценка энергетических ошибок, которая позволяет определить эффективность распознавания образов. В заключении производится отбор кандидатов лекарственных соединений с заданными свойствами для дальнейших исследований.

На рисунке 1 представлена схема интеллектуальной технологии прогнозирования фармакологической активности лекарственных препаратов на основе алгоритмов пчелиной колонии и роя частиц.

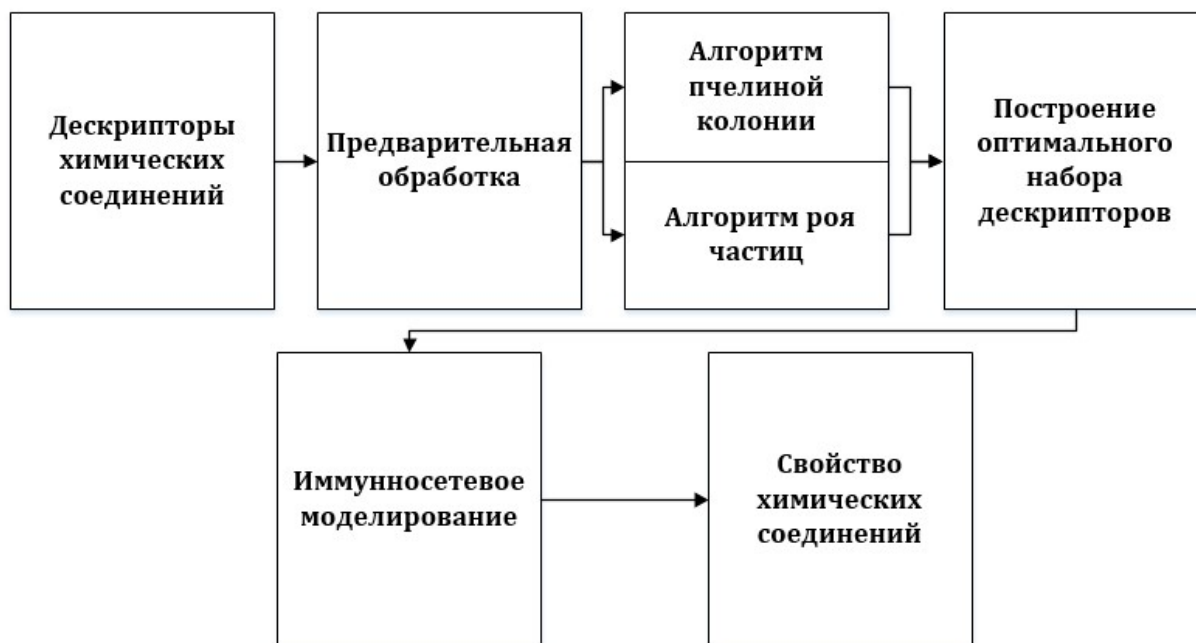


Рисунок 1 – Интеллектуальная технология прогнозирования зависимости структура-свойства/активность

На этапе предварительной обработки данных осуществляется нормализация, проверяется полнота дескрипторов, достоверность и выделение информативных дескрипторов на основе алгоритмов роевого интеллекта. Дескриптор - параметр, характеризующий структуру соединения, химические и биологические свойства веществ. Дескрипторы классифицируются по нескольким уровням: дескрипторы элементного уровня, дескрипторы структурной формулы, дескрипторы электронной структуры, дескрипторы молекулярной формулы, дескрипторы междумолекулярных взаимодействий (Раевский О.А., 2006: 99). Основными типами молекулярных дескрипторов являются: фрагментные дескрипторы, топологические индексы, физико-химические дескрипторы, квантово-

химические дескрипторы, дескрипторы молекулярных полей, константы заместителей, фармакофорные дескрипторы, дескрипторы молекулярного подобия и др. Всего в настоящее время из различных источников можно встретить более 7000 типов молекулярных дескрипторов (Ходашинский И.А., 2012: 142). Отбор информативных дескрипторов позволяет находить оптимальное количество дескрипторов, при котором после иммуносетевого моделирования и прогнозирования свойств новых химических соединений ошибка обобщения будет минимальной. Выделение информативных дескрипторов позволяет сократить затраты на вычисление ненужных дескрипторов, повышает скорость алгоритмов прогнозирования и качество прогноза.

В настоящее время увеличивается использование онтологических моделей для создания систематизированных структур информационных систем (Samigulina G.A., 2016: 3). Разработка интеллектуальной системы с использованием онтологических моделей позволяет учитывать особенности функционирования и взаимосвязи, снижает временные и вычислительные ресурсы. При разработке интеллектуальной системы компьютерного молекулярного дизайна лекарственных препаратов с заданными свойствами особенно важно построение интегрированной онтологической модели. Применение современных онтологических редакторов и построение OWL (Web Ontology Language) моделей облегчает решение задачи формирования оптимального набора дескрипторов.

В настоящее время для создания и поддержки онтологий существует целый ряд инструментов, которые помимо общих функций редактирования выполняют поддержку документирования онтологий, импорт и экспорт онтологий разных форматов и языков, поддержку графического редактирования, управление библиотеками онтологий и т.д. Редактор онтологии Protege является доступной программой для проектирования онтологических моделей. Protege включает редактор онтологий, позволяющий проектировать онтологии разворачивая иерархическую структуру абстрактных или конкретных классов и слотов. Структура онтологии сделана аналогично иерархической структуре каталога (Гладун А.Я., 2006: 1). Protege может генерировать формы получения знаний для введения классов и подклассов. Инструмент имеет графический интерфейс, удобный для использования.

3 Материал и методы

Постановка задачи формулируется следующим образом: необходимо разработать онтологические модели алгоритмов роевого интеллекта для решения задачи выделения информативных дескрипторов и дальнейшего иммуносетевого моделирования новых лекарственных препаратов с заданными свойствами.

Роевой интеллект основан на обратной связи при адаптации к изменяющейся среде и децентрализованности между частицами. Алгоритмы роевого интеллекта позволяют находить новые решения и улучшают существующие решения избегая ранней сходимости к локальному решению. Для поведения роя характерны автономность, распределение функций и самоорганизация. Размер популяции оказывает влияние на точность и время вычислений (Олейник А.А., 2012: 114). Большая популяция увеличивает время вычисления и ухудшает сходимость, а малая популяция сходимость улучшает. В алгоритме роя частица меняет свое положение основываясь на знании текущего положения, предыдущего состояния и предыдущих состояний соседей данной частицы.

Алгоритм пчелиной колонии (Олейник А.А., 2011: 85) является одним из эвристических алгоритмов для решения оптимизационных задач, который основан на имитации поведения колонии пчел при сборе нектара. Для описания поведения пчёл используются понятия, такие как рабочие пчелы, пчелы-исследователи, пчелы-разведчики. Все пчелы будут выбирать для исследования не только элитные участки, но и окрестности для разнообразия решений в последующих итерациях и для увеличения вероятности обнаружения оптимальных решений.

Для решения задачи выделения информативных дескрипторов используются модификации классического метода алгоритма пчелиной колонии. Алгоритмы GABC1 и GABC2 отличаются тем, что приводят к глобальному оптимуму. В модификации JA-ABC улучшена скорость сходимости и предусмотрена способность предотвращения преждевременной сходимости. Алгоритм ABC использует правила Деба, состоящие из трех эвристических правил и вероятностной схемы выбора выполнимых решений, основанных на их значениях фитнес-функции. Алгоритм MABC использует новый механизм, основанный на измерении различий между бинарными строками для генерации новых решений. Особенности пчелиного алгоритма являются динамическое разбиение поискового пространства на области и совместное исследование элитных областей и неэлитных окрестностей. При решении задачи возможно использование вышеуказанных модификаций пчелиного алгоритма и за счёт распараллеливания уменьшение времени работы алгоритмов (Schiezaro M., 2013: 30).

Алгоритм роя частиц может рассматриваться как многоуровневая система, в которой каждая частица-агент обменивается информацией по простым правилам (Agrawal S., 2015: 403). Поиск оптимального решения начинается с рассмотрения случайных частиц, называемых роем. Каждая частица имеет фитнес-значение, которое вычисляется с помощью функции затрат. Фитнес-функция создает пространство поиска, в котором каждое решение имеет фитнес-значение. Каждая частица настраивает позицию и скорость в пространстве поиска. На основе координаты частицы рассчитывается скорость частицы. Выбираются координаты лучшего решения для частицы и для роя частиц. При перемещении частица меняет свою скорость и позицию в соответствии со своим собственным опытом и соседей. Метод роя частиц позволяет проводить более ёмкий обмен информацией между частицами. В качестве связи используется так называемая общая память, суть которой в том, что каждая частица знает координаты наилучшей точки. То есть наилучшее решение, найденное роем в каждый момент времени известно всем его частицам (Hamed Z., 2012: 125). В итоге на движение частицы влияет стремление к наилучшему среди всех частиц положению.

Существует множество модификаций классического алгоритма с учетом уменьшения вероятности преждевременной сходимости путем изменения характеристик движения частиц и динамического изменения параметров алгоритма во время оптимизации. Алгоритм LBEST тщательно исследует пространство поиска, однако является более медленным, чем классический алгоритм (Particle Swarm Optimization). При этом, чем меньшее число соседей учитывается при формировании вектора скорости, тем ниже скорость сходимости алгоритма, но тем эффективней он избегает субоптимальных решений. В алгоритме InWPSO было изменено правило обновления векторов скоростей частиц. Алгоритм TVInWPSO показывает, что частицы сначала исследуют область поиска экстенсивно, находя множество субоптимальных решений, а со временем все более

концентрируются на исследовании их окрестностей. Возрастание инерции способствует сходимости алгоритма на поздних стадиях работы. Алгоритм CanPSO отличается способом вычисления векторов скоростей частиц, так как был введен дополнительный множитель для контроля скорости частицы. В модифицированном алгоритме FIParticleSwarm важную роль играет информированность всех частиц, что ведет к повышению эффективности алгоритма. Разработано следующее программное обеспечение на основе пчелиных алгоритмов для решения задачи отбора информативных дескрипторов: ABCoptim, ABCMATLABweb, Artificial-Bee-Colony, EABC, abc-clustering и др.

Существует ряд пакетов прикладных программ на основе роя частиц: ParadisEO, TribesPSO, EPSO, SwarmNLP, Co-PSO и др. Приложение ParadisEO - это программное обеспечение, посвященное проектированию, реализации и анализу метаэвристики для многокритериальной оптимизации. Программный продукт EPSO (Evolutionary Particle Swarm Optimization) является гибридной системой на основе эволюционного алгоритма и метода роя частиц. Программа SwarmNLP визуализирует оптимизацию роя частиц по многократным критериям. Эта программа является хорошим инструментом для решения задачи на основе оптимизации роя частиц и визуализации пространства решений. Программный пакет Co-PSO включает в себя расширяемые библиотеки тестовых функций и графов, определяющих топологии соседства, которые могут быть использованы алгоритмами роя частиц. В результате работы программа Co-PSO строит график изменения значений целевой функции с ростом числа итераций для частиц-победителя, а также графики изменения размеров частиц в функции номера итерации.

4 Результаты и обсуждение

Разработка онтологической модели начинается с общих понятий и переходит к детализации. Иерархическая структура онтологической модели состоит из основных классов, таких как входные (INPUTDATA) и выходные данные (OUTPUTDATA). В работе (Самигулина Г.А., 2015: 15) предложена интегрированная онтологическая модель иммуносетевой технологии, которая состоит из онтологической модели предварительной обработки данных, онтологической модели распознавания образов и онтологической модели оценки энергетической ошибки. Формальная интегрированная модель онтологии представляется в виде кортежа множеств:

$$OM_{ИТ} = \langle OM_{PR}, OM_{IR}, OM_{АЕЕ} \rangle$$

где OM_{PR} - онтологическая модель предварительной обработки данных;

OM_{IR} - онтологическая модель распознавания образов;

$OM_{АЕЕ}$ - онтологическая модель оценки энергетической ошибки.

При построении оптимальной иммуносетевой модели применяется мультиалгоритмический подход (Самигулина Г.А., 2016: 55), который позволяет использовать несколько алгоритмов и методов искусственного интеллекта для отбора дескрипторов. В работе предварительная обработка данных осуществляется на основе алгоритмов пчелиной колонии и роя частиц.

Онтологическая модель предварительной обработки данных состоит из онтологических моделей алгоритмов пчелиной колонии и роя частиц:

$$OM_{PR} = \langle OM_{ABC}, OM_{PSO} \rangle$$

где OM_{ABC} - онтологическая модель алгоритмов пчелиной колонии;

OM_{PSO} - онтологическая модель алгоритмов роя частиц.

Онтологическая модель алгоритмов пчелиной колонии состоит из нескольких модификаций классического алгоритма:

$$OM_{ABC} = \langle OM_{GABC1}, OM_{GABC2}, OM_{JA-ABC}, OM_{ABC}, OM_{MBABC}, OM_{BasicABC} \rangle$$

где OM_{GABC1} - онтологическая модель алгоритма GABC1;

OM_{GABC2} - онтологическая модель алгоритма GABC2;

OM_{JA-ABC} - онтологическая модель модификации JA-ABC;

OM_{ABC} - онтологическая модель алгоритма ABC;

OM_{MBABC} - онтологическая модель модификации MBAВ;

$OM_{BasicABC}$ - онтологическая модель классического алгоритма.

Так как не существуют универсальных методов для построения оптимального набора дескрипторов, то при разработке интеллектуальной системы прогнозирования фармакологической активности лекарственных соединений основной задачей является уменьшение ошибки обобщения и повышение качества прогноза.

Достоинством применения мультиалгоритмического подхода является возможность выбора алгоритма с наилучшими результатами после иммунносетевого моделирования, то есть выбирается тот алгоритм, использование которого дает минимальную ошибку обобщения. Алгоритмы, описанные выше, могут быть использованы для отбора дескрипторов и их результаты будут отличаться в зависимости от размера и качества данных, наличия независимых дескрипторов, критериев оптимальности и останова и т.д. При построении онтологической модели алгоритмов пчелиной колонии (Таблица 1) в качестве примера был рассмотрен классический алгоритм пчелиной колонии, так как отличительной чертой классического алгоритма является децентрализованность, взаимодействие агентов и простота поведения агентов. Данный алгоритм зависит от таких параметров как: общее число пчёл-разведчиков, общее число участков, число элитных участков, максимальное число итераций. Поэтому онтологическая модель классического алгоритма представляется следующим образом:

$$OM_{BasicABC} = \langle InitialPopulation, NumberOfLocation, Iteration, Gbest \rangle$$

где *InitialPopulation* - общее число пчёл-разведчиков;

NumberOfLocation - общее число участков;

Iteration - максимальное число итераций;

Gbest - лучшая координата пчелы.

С помощью инструмента OntoGraf в редакторе онтологий Protege было осуществлено графическое представление онтологической модели классического алгоритма пчелиной колонии (Рисунок 2). Алгоритм роя частиц также имеет несколько модификаций. Они различаются скоростью сходимости, наличием дополнительных множителей и др.

$$OM_{PSO} = \langle OM_{LBEST}, OM_{InWPSO}, OM_{TVInWPSO}, OM_{CanPSO}, OM_{FIPS}, OM_{BasicPSO} \rangle$$

где OM_{LBEST} - онтологическая модель алгоритма LBEST;
 OM_{InWPSO} - онтологическая модель модификации InWPSO;
 $OM_{TVInWPSO}$ - онтологическая модель алгоритма TVInWPSO;
 OM_{CanPSO} - онтологическая модель алгоритма CanPSO;
 OM_{FIPS} - онтологическая модель алгоритма FIParticleSwarm;
 $OM_{BasicPSO}$ - онтологическая модель классического алгоритма.

Таблица 1 – Онтологическая модель алгоритмов пчелиной колонии

Наименование	Содержание
Онтологическая модель алгоритмов пчелиной колонии	Алгоритмы пчелиной колонии:
	- Алгоритм GABC1
	- Алгоритм GABC2
	- Модификация JA-ABC
	- Алгоритм ABC
	- Модификация MBABC
	- Классический алгоритм BasicABC
	Создание размера популяции.
	Определение местоположения источников нектара (DetermineLocationOfFoodSource).
	Оценка фитнес-функций.
	Поиск рабочими пчелами новых источников и исследование лучшего источника (EmployedBeePhase).
	Выбор источника пчелой-исследователем (OnlookerBeePhase).
	Произведение локальной разведки в окрестностях выбранных точек (ScoutBeePhase).
	Определение соседства агентов (DetermineNeighbourhood).
	Выбор элитного (SelectEliteLocation) и неэлитного (SelectNon-eliteLocation) местоположения.
	Запоминание лучшего источника (Gbest). При достижении улучшения в рое сохраняется обновленное лучшее решение и соответствующий ему вектор параметров целевой функции.
	При выполнении условия останова (StopSearching) из множества решений выбирается наилучшее, которое является результатом вычисления алгоритма.
	Отобранный набор дескрипторов.

Критерием останова могут быть: достижение максимально допустимого количества итераций, нахождение приемлемого решения, отсутствие существенного улучшения значения целевой функции на протяжении некоторого количества итераций. За счет исключения дополнительных параметров классический алгоритм роя частиц снижает вычислительные сложности и время обработки данных. Онтологическая модель классического алгоритма роя частиц представляется следующим образом:

$$OM_{BasicPSO} = \langle PopulationSizeOfAgents, pbest, Iteration, Gbest \rangle$$

где $PopulationSizeOfAgents$ - общее число агентов;
 $pbest$ - наилучшая позиция частицы;

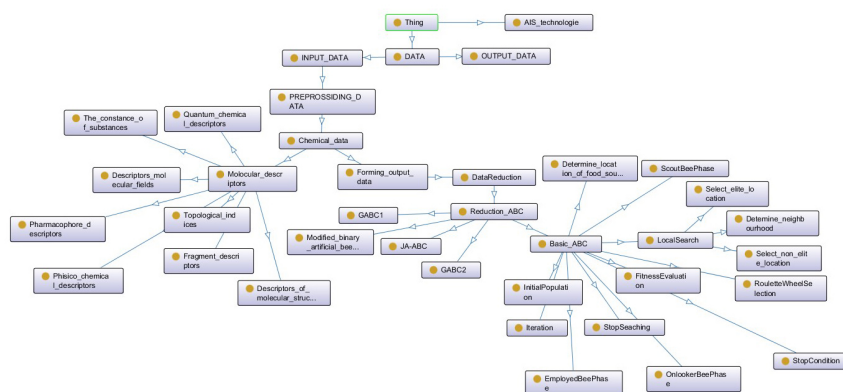


Рисунок 2 – Онтологическая модель алгоритма пчелиной колонии

Iteration - максимальное число итераций;

Gbest - наилучшая позиция роя.

В графическом виде онтологическая модель алгоритмов роя частиц представлена на рисунке 3.

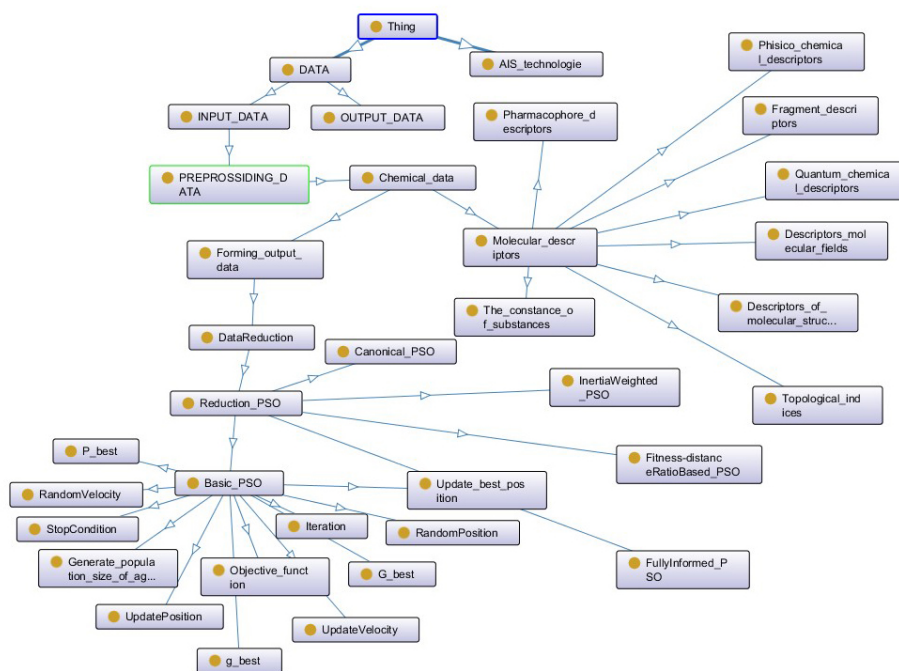


Рисунок 3 – Онтологическая модель роя частиц

В работе построена онтологическая модель классического алгоритма BasicPSO, так как алгоритм более глубоко исследует дескрипторное пространство и не имеет дополнительных параметров. Онтологическая модель алгоритмов роя частиц представлена в таблице 2.

Таблица 2 – Онтологическая модель алгоритмов роя частиц

Наименование	Содержание
Онтологическая модель алгоритмов роя частиц	Алгоритмы роя частиц:
	- Алгоритм LBEST
	- Модификация InWPSO
	- Алгоритм TVInWPSO
	- Алгоритм CanPSO
	- Алгоритм FIPS
	- Классический алгоритм BasicPSO
	Создание размера популяции частиц (PopulationSizeOfAgents).
	Генерация начальных положений (RandomPosition) и скоростей (RandomVelocity, при которой создается вектор равномерно распределенных случайных величин и произвольно выбирается наилучшая позиция частицы (pbest).
	Вычисление фитнес-функций.
	Определение наилучшего положения (gbest).
	Миграция на основе стохастического поведения частиц.
	Обновление позиций (UpdatePosition) и скоростей (UpdateVelocity).
	Обновление лучшей позиции частиц (UpdateBestPosition).
	Проверка условия завершения (StopCondition), которая происходит в результате достижения заданного числа итераций.
	Сохранение лучшего положения (Gbest). При достижении улучшения в рое сохраняется обновленное лучшее решение и соответствующий ему вектор параметров целевой функции.
	Отобранный набор дескрипторов.

5 Заключение

Интеллектуальная технология прогнозирования зависимости структура-свойства/активность с использованием мультиалгоритмического подхода позволяет интегрировать разные методы искусственного интеллекта для решения задачи компьютерного молекулярного дизайна новых лекарственных препаратов с заданными свойствами. Применение мультиалгоритмического подхода с использованием алгоритмов роевого интеллекта позволяет эффективно осуществлять предварительную обработку данных, выделять информативный набор дескрипторов, создавать оптимальную иммунносетевую модель. Выбирается тот алгоритм роевого интеллекта, который дает минимальную ошибку обобщения при иммунносетевом моделировании. Разработаны онтологические модели алгоритмов пчелиной колонии и алгоритмов роя частиц для выделения оптимального набора дескрипторов и дальнейшего иммунносетевого моделирования в редакторе онтологий Protege. Разработка интеллектуальной системы с использованием онтологических моделей рассматриваемых алгоритмов позволяет структурировать данные и анализировать скрытые взаимосвязи между дескрипторами. Работа выполнена по гранту КН МОН РК № ГР 0115РК00549 (2015-2017 г.г) по теме проекта: "Компьютерный молекулярный дизайн лекарственных препаратов на основе иммунносетевого моделирования".

Список литературы

- [1] *Agrawal S., Silakari S.* A review on application of Particle Swarm Optimization in Bioinformatics // Current bioinformatics. -2015. -P. 401-413.
- [2] *Baranjuk V.V., Smirnova O.S.* Expanding the bionics ontology by the description of swarm intelligence // International Journal of Open Information Technologies. -2015. -Vol. 3, No. 12. -P. 13-17.
- [3] *Chin Y.L., Chun W.Y.* Current Modeling Methods Used in QSAR/QSPR // Statistical Modelling of Molecular Descriptors in QSAR/QSPR. -2012. -P. 1-31.
- [4] *Dasgupta D.* Artificial Immune Systems and Their Applications. - NY: Springer-Verlag, 1998. -438 p.
- [5] *Golla S., Neely B., Whitebay E., Madihally S., Robinson R., Gasem K.* Virtual design of chemical penetration enhancers for transdermal drug delivery // Chem. Biol. Drug Design. -2012. -P. 478-487.
- [6] *Hamed Z., Aboozar K., Hamid M.* Application of modified particle swarm optimization as an efficient variable selection strategy in QSAR/QSPR studies // Journal of chemometrics. -2012. -P. 123-128.
- [7] *Rangel C., Aguilar J., Cerrada M., Altamiranda J.* An Approach for the Emerging Ontology Alignment based on the Bees Colonies // Proc. of Int'l Conf. Artificial Intelligence. -2015. -P. 536-541.
- [8] *Samigulina G.A., Samigulina Z.I.* Computational Molecular Design of Antiseptic Drugs based on Immune Network Modeling // Proceedings of the 12-th International Conference on Electronics Computer and Computation "ICECCO-2015". -Almaty: Suleyman Demirel University, 2015. -C. 47-52.
- [9] *Samigulina G.A., Samigulina Z.I.* Drug Design of sulfanilamide based on Immune Network Modeling and Ontological approach // Proc. of the 10th IEEE Int. Conf. on Application of Information and Communication Technologies AICT2016. -Baku, 2016.
- [10] *Schiezaro M., Pedrini H.* Data feature selection based on Artificial bee colony algorithm // EURASIP Journal on Image and Video processing. -2013. -P. 29-33.
- [11] *Shao L., Bai Y., Qiu Y., Du Zh.* Particle Swarm Optimization Algorithm Based on Semantic Relations and Its Engineering Applications // Systems Engineering Procedia. -2012. -Vol. 5. -P. 222-227.
- [12] *Shi-xiong X., Fei L., Qiang N.* Fault Diagnosis Method Based on Ontology and Particle Swarm-Immune Optimization Algorithm in the Motor // Multimedia and Signal Processing (CMSP). -2011. -P. 1-9.
- [13] *Tarakanov A., Nicosia G.* Foundations of immunocomputing // Proceedings of the 1-st IEEE Symposium of Computational Intelligence. -Honolulu, 2007. -P. 503-508.
- [14] *Yuan F., Chen S., Liu H., Xu L.* Artificial bee colony-based extraction of non-taxonomic relation between symptom and syndrome in TCM records [Электрон. ресурс]. -Vol. 6, Issue 6. -URL: <http://dx.doi.org/10.1504/IJCSM.2015.073600> (дата обращения: 07.01.2017)
- [15] *Гладун А.Я., Рогозина Ю.В.* Онтологии в корпоративных системах [Электрон. ресурс]. -2006. -URL: <http://www.management.com.ua/ims/ims116.html> (дата обращения: 19.12.2016)
- [16] Компьютерный молекулярный дизайн лекарственных препаратов на основе иммуносетевого моделирования: отчет о НИР (промежуточный) / Институт информационных и вычислительных технологий КН МОН РК: рук. Самигулина Г.А. -Алматы, 2015. -145 с. -№ГР 0115РК00549. - Инв. №0215РК01473.
- [17] *Олейник Ан.А., Олейник Ал.А., Субботин С.А.* Агентные технологии для отбора информативных признаков // Кибернетика и системный анализ. -2012. -№ 2. -С. 113-125.
- [18] *Олейник А.А.* Мультиагентный метод оптимизации с адаптивными параметрами // Искусственный интеллект. -2011. -№1. -С. 83-90.
- [19] *Раевский О.А.* Дескрипторы водородной связи в компьютерном молекулярном дизайне // Рос. Хим. Журнал. -2006. -№2. -С. 97-108.
- [20] *Самигулина Г.А., Самигулина З.И.* Построение оптимальной иммуносетевой модели для прогнозирования свойств неизвестных лекарственных соединений на основе мультиалгоритмического подхода // Проблемы информатики. -Новосибирск, 2013. -С. 21-29.
- [21] *Ходашинский И.А., Горбунов И.В., Дудин П.А., Синьков Д.С.* Построение нечетких систем прогнозирования эффективности немедикаментозного лечения // Интеллектуальные системы. -2012. -№3 (33). -С. 140-149.

References

- [1] Agrawal S., Silakari S., "A review on application of Particle Swarm Optimization in Bioinformatics," *Current bioinformatics* 10 (2015): 401-413.
- [2] Baranjuk V.V. and Smirnova O.S., "Expanding the bionics ontology by the description of swarm intelligence" *International Journal of Open Information Technologies* 3(12) (2015): 13-17.
- [3] Chin Y.L. and Chun W.Y., "Current Modeling Methods Used in QSAR/QSPR," *Statistical Modelling of Molecular Descriptors in QSAR/QSPR*(2012): 1-31.
- [4] Dasgupta D., *Artificial Immune Systems and Their Applications* (NY: Springer-Verlag, 1998), 438.
- [5] Gladun A.Ya., Rogushina Yu.V. (2006) Ontologii v korporativnyh sistemah [Ontologies in enterprise systems]. *Korporativnye sistemy*, vol. 1, accessed December 19, 2016, <http://www.management.com.ua/ims/ims116.html>.
- [6] Golla S. et al., "Virtual design of chemical penetration enhancers for transdermal drug delivery," *Chem. Biol. Drug Design* (2012): 478-487.
- [7] Hamed Z., Aboozar K., Hamid M., "Application of modified particle swarm optimization as an efficient variable selection strategy in QSAR/QSPR studies" *Journal of chemometrics* (2012): 123-128.
- [8] Oleinik An.A., Oleinik Al.A., Subbotin S.A. (2012) Agentnye texnologii dlya otbora informativnyh priznakov [Agent technologies for the selection of informative features]. *Kibernetika i sistemnyi analiz*, vol. 2, pp. 113-125.
- [9] Oleinik A.A. (2011) Multiagentnyi metod optimizacii s adaptivnymi parametrami [Multi algorithmic approach with adaptive parameters]. *Iskusstvennyi intellekt*, vol.1, pp. 83-90.
- [10] Raevskii O.A. (2006) Deskriptory bodorodnoi svyazi v komputernom molekulyarnom dizaine [Descriptors of hydrogen bond in computer molecular design]. *Ros. xim. jurnal*, vol. 2, pp. 97-108.
- [11] Rangel C. et al., "An Approach for the Emerging Ontology Alignment based on the Bees Colonies," *Int'l Conf. Artificial Intelligence* (2015): 536-541.
- [12] Samigulina G.A. and Samigulina Z.I., "Computational Molecular Design of Antiseptic Drags based on Immune Network Modeling," *Proceedings of the 12-th International Conference on Electronics Computer and Computation "ICECCO-2015"* (Almaty: Suleyman Demirel University, 2015): 47-52.
- [13] Samigulina G.A. and Samigulina Z.I. "Drug Design of sulfanilamide based on Immune Network Modeling and Ontological approach," *Proc. of the 10th IEEE Int. Conf. on Application of Information and Communication Technologies AICT2016*, accessed January 19, 2017, www.aict.info/2016.
- [14] Samigulina G.A. et al. (2015) Komputernyi molekulyarnyi dizain lekarstvennyh preparatov na osnove immunnosetenovogo modelirovaniya [Computer-aided molecular design of drugs based on immune network modeling]. *Otchet o NIR*, pp. 145.
- [15] Samigulina G.A., Samigulina Z.I. (2013) Postroenie optimalnoi immunnosetevoi modeli dlya prognozirovaniya svoystv neizvestnyh lekarstvennyh soedinenii na osnove multialgoritmicheskogo podhoda [Construction of an optimal immune-network model for predicting the properties of unknown drug compounds on the basis of a multi-algorithmic approach]. *Problemy informatiki*, pp. 21-29.
- [16] Schiezero M. and Pedrini H., "Data feature selection based on Artificial bee colony algorithm," *EURASIP Journal on Image and Video processing*(2013): 29-33.
- [17] Shao L. et al., "Particle Swarm Optimization Algorithm Based on Semantic Relations and Its Engineering Applications," *Systems Engineering Procedia* (2012): 222-227.
- [18] Shi-xiong X. et al., "Fault Diagnosis Method Based on Ontology and Particle Swarm-Immune Optimization Algorithm in the Motor," *Multimedia and Signal Processing (CMSP)*(2011): 1-9.
- [19] Tarakanov A. and Nicosia G., "Foundations of immunocomputing," *Proceedings of the 1-st IEEE Symposium of Computational Intelligence* (Honolulu, 2007): 503-508.
- [20] Xodashinskii I.A., Gorbunov I.V., Dudin P.A., Sinkov D.S. (2012) Postroenie nechetkih sistem prognozirovaniya effektivnosti nemedikamentoznogo lecheniya [Construction of fuzzy systems for predicting the effectiveness of non-drug treatment]. *Intellektualnye sistemy*, vol. 3, no 33, pp. 140-149.
- [21] Yuan F. et al., "Artificial bee colony-based extraction of non-taxonomic relation between symptom and syndrome in TCM records," *International Journal of Computing Science and Mathematics*(2015), accessed January 7, 2017, DOI: <http://dx.doi.org/10.1504/IJCSM.2015.073600>.