На правах рукописи

Ахмедова Шахназ Агасувар кызы

КОЛЛЕКТИВНЫЙ САМОНАСТРАИВАЮЩИЙСЯ МЕТОД ОПТИМИЗАЦИИ НА ОСНОВЕ БИОНИЧЕСКИХ АЛГОРИТМОВ

05.13.01 – Системный анализ, управление и обработка информации
(информатика, вычислительная техника, управление)

АВТОРЕФЕРАТ

диссертации на соискание ученой степени
кандидата технических наук

Красноярск – 2016

Работа выполнена в ФГБОУ ВО «Сибирский государственный аэрокосмический университет имени академика М.Ф. Решетнева» (СибГАУ), г. Красноярск

Научный руководитель: доктор технических наук, профессор **Семенкин Евгений Станиславович**

Официальные оппоненты: **Дивеев Асхат Ибрагимович**, доктор технических наук, профессор, ФГБУН «Вычислительный центр им. А. А. Дородницына Российской академии наук», г. Москва, сектор проблем кибернетики отдела нелинейного анализа и проблем безопасности, заведующий сектором

Хамисов Олег Валерьевич, доктор физико-математических наук, с.н.с., ФГБУН «Институт систем энергетики имени Л. А. Мелентьева Сибирского отделения Российской академии наук», г. Иркутск, отдел прикладной математики, заведующий отделом

Ведущая организация: Национальный исследовательский университет МГТУ им. Н. Э. Баумана

Защита состоится “_07_” _апреля_ 2016 г. в 14:00 на заседании диссертационного совета Д 999.007.02, созданного на базе Сибирского федерального университета и Института вычислительного моделирования СО РАН, по адресу 660074, Красноярск, ул. Киренского, 26Б, СФУ, ИКИТ, ауд. УЛК115.

С диссертацией можно ознакомиться в библиотеке и на сайте Сибирского федерального университета http://research.sfu-kras.ru/attestation/dissertations.

Автореферат разослан "___"_____20__г.

Ученый секретарь диссертационного совета — Бронов Сергей Александрович

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность. В настоящее время бионические стохастические алгоритмы оптимизации многоэкстремальных функций вещественных переменных, основанные на имитации коллективного поведения различных животных, продемонстрировали свою эффективность и активно используются при решении различных практических задач. Процедура поиска наилучшего решения такими алгоритмами имитирует некоторый природный процесс либо поведение определенных видов животных, учитывающее их видовые особенности.

Одним из наиболее изученных среди таких алгоритмов является так называемый стайный (роевый) алгоритм (Particle Swarm Optimization, PSO), разработанный в 1995 году. Исследования показали эффективность алгоритма и целесообразность его применения при решении задач как безусловной, так и условной оптимизации функций вещественных переменных. Помимо стайного алгоритма наибольший интерес из последних разработок представляют аналогичные ему следующие бионические алгоритмы: алгоритм стай волков (Wolf Pack Search, WPS), алгоритм светлячков (Firefly Algorithm, FFA), алгоритм поиска кукушек (Cuckoo Search Algorithm, CSA) и алгоритм летучих мышей (Bat Algorithm, BA). Перечисленные бионические методы оптимизации, как и стайный алгоритм, изначально были разработаны для решения оптимизационных задач с вещественными переменными и идейно наиболее близки к нему, что отличает их от других аналогичных подходов (пчелиные алгоритмы, алгоритм умных капель и т.п.). Каждый из упомянутых алгоритмов имитирует некоторый аспект поведения соответствующего вида животных.

Эффективность таких алгоритмов при решении реальных практических задач оценивается по найденному минимальному (максимальному) значению целевой функции, а также по количеству вычислений целевой функции, необходимых для достижения данных значений. В то же время, при тестировании требуется проводить значительное число запусков, поэтому эффективность таких алгоритмов оценивается как их надежность, т.е. доля успешных запусков, когда было найдено оптимальное решение с заданной точностью.

Основной проблемой применения бионических алгоритмов является необходимость довольно точной настройки их многочисленных параметров, от которых существенно зависит эффективность оптимизации. Кроме того, заранее не ясно, какой именно алгоритм более всего подходит для решения конкретной задачи, т.к. они обладают схожей эффективностью. Поэтому исследователи зачастую выбирают алгоритмы наугад. В этой связи разработка универсального самонастраивающегося бионического метода оптимизации является **актуальной научно-технической задачей**.

Цель диссертационной работы состоит в повышении эффективности

решения задач оптимизации бионическими алгоритмами за счет автоматизации их выбора и настройки их параметров.

Сформулированная цель предопределила следующую совокупность решаемых **задач***:*

1. Исследовать эффективность различных бионических методов для решения задач условной и безусловной оптимизации с вещественными и бинарными переменными, определить наиболее успешный из них.
2. Разработать коллективный самонастраивающийся бионический алгоритм для решения задач условной и безусловной оптимизации с вещественными переменными.
3. Разработать модификации нового алгоритма для решения задач условной и безусловной оптимизации с бинарными переменными;
4. Реализовать разработанные алгоритмы в виде программных систем и оценить их эффективность на репрезентативном множестве тестовых задач.
5. Провести апробацию разработанных алгоритмов при решении реальных практических задач.

Методы исследования. В данной работе использовались методы эволюционных вычислений, оптимизации, теории вероятностей и математической статистики, системного анализа, методика разработки интеллектуальных информационных систем и другие.

Научная новизна работы заключается в следующем:

1. На основе пяти известных бионических алгоритмов разработан, реализован и исследован новый коллективный метод решения задач безусловной оптимизации с вещественными переменными, отличающийся от известных способом организации взаимодействия популяций и настройки параметров.
2. Разработана, реализована и исследована модификация нового метода для решения задач безусловной оптимизации с бинарными переменными.
3. Разработана, реализована и исследована модификация нового метода для решения задач условной оптимизации с вещественными и бинарными переменными.
4. На основе разработанного метода оптимизации предложены новые алгоритмы автоматического построения нейронных сетей и машин опорных векторов.

Теоретическая значимость результатов диссертационного исследования состоит в разработке нового самонастраивающегося метода оптимизации, основанного на коллективной работе бионических алгоритмов, обладающего большей эффективностью по точности, скорости и надежности по сравнению с каждым из используемых в нем алгоритмов и позволяющего исключить значительные временные затраты

на выбор наиболее эффективного из них за счет автоматической настройки на задачу в ходе ее решения.

Практическая ценность. Разработанные на базе предложенного метода самонастраивающиеся алгоритмы моделирования и оптимизации исключают многократные прогоны для выбора их настроек под конкретную задачу, позволяя тем самым экономить время и вычислительные ресурсы, не снижая при этом эффективности их применения конечным пользователем. Разработанная процедура позволяет использовать полезные свойства нескольких методов при решении практических задач и не требует от конечного пользователя экспертных знаний в области эволюционного моделирования и оптимизации. В ходе выполнения работы были успешно решены задачи оптимизации и анализа данных из области технической и медицинской диагностики, банковского скоринга, категоризации текстовых документов, и другие.

Реализация результатов работы. Разработанные алгоритмы использованы при выполнении исследований в рамках российско-германских проектов (совместно с университетом г. Ульм) «Распределенные интеллектуальные информационные системы обработки и анализа мультилингвистической информации в диалоговых информационно-коммуникационных системах» (ФЦП ИР, ГК №11.519.11.4002) и «Математическое и алгоритмическое обеспечение автоматизированного проектирования аппаратно-программных комплексов интеллектуальной обработки мультилингвистической информации в распределенных высокопроизводительных системах космического назначения» (ФЦП НПК, ГК № 16.740.11.0742), российско-словенского проекта (совместно с университетом г. Марибор) «Manpower control strategy determination with self-adapted evolutionary and biologically inspired algorithms» (ARRS Project BI-RU/14-15-047), а также в рамках проекта №8.5541.2011 «Развитие теоретических основ автоматизации математического моделирования физических систем на основе экспериментальных данных» и проекта № 140/14 «Разработка теоретических основ эволюционного проектирования интеллектуальных информационных технологий анализа данных» тематического плана ЕЗН СибГАУ. Данная работа поддержана грантом Фонда содействия развитию малых форм предприятий в научно-технической сфере в рамках программы У.М.Н.И.К., стипендией имени Аниты Борг от компании Google (Google Anita Borg Scholarship) и стипендией Президента РФ молодым ученым и аспирантам, осуществляющим перспективные научные исследования и разработки по приоритетным направлениям модернизации российской экономики.

Шесть программных систем зарегистрированы в Роспатенте. Разработанные в диссертации программные системы используются в учебном процессе Института информатики и телекоммуникаций СибГАУ

при выполнении лабораторных и курсовых работ и, кроме того, переданы для использования в две инновационные IT- компании.

Основные положения, выносимые на защиту:

1. Разработанные самонастраивающиеся коллективные бионические методы позволяют эффективно по скорости, точности и надежности решать задачи условной и безусловной оптимизации с вещественными и бинарными переменными и превосходят алгоритмы-компоненты по надежности и скорости.

2. Предложенный способ самонастройки бионических алгоритмов позволяет автоматически определять наиболее эффективный из них и избегать излишних затрат на подбор их параметров, не снижая при этом точности и надежности моделирования и оптимизации.

3. Кооперативный бионический алгоритм автоматического генерирования искусственных нейронных сетей позволяет автоматически генерировать структуру и настраивать весовые коэффициенты нейросетевых моделей, применяемых в различных задачах анализа данных.

4. Кооперативный бионический алгоритм построения машин опорных векторов позволяет автоматически генерировать эффективные интеллектуальные информационные технологии на их основе.

Апробация работы. Процесс разработки алгоритмов и результаты проведенных исследований докладывались в период 2010-2015 гг. на 35 конференциях различного уровня, среди которых: International Conference on Swarm Intelligence (ICSI, Hefei, China, 2014; Beijing, China, 2015), International Conference on Informatics in Control, Automation and Robotics (ICINCO, Vienna, Austria, 2014; Colmar, France, 2015), International Congress on Advanced Applied Informatics (AAI, Okayama, Japan, 2015), Congress on Evolutionary Computations of the IEEE World Congress on Computational Intelligence (CEC WCCI, Beijing, China, 2014), International Conference on Computer Science and Artificial Intelligence (ICCSAI, Wuhan, China, 2014), Conference on Engineering and Applied Sciences Optimization (OPT-I, Kos Island, Greece, 2014), Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis, (Baltimore, USA, 2014), Congress on Evolutionary Computations (CEC, Cancun, Mexico, 2013), Genetic and Evolutionary Computation Conference (GECCO, Amsterdam, Holland, 2013), International Workshop on Mathematical Models and its Applications (IWMMA, Baykal, Russia, 2012; Krasnoyarsk, Russia, 2013, 2014, 2015), XIV Национальная конференция по искусственному интеллекту с международным участием (КИИ, г. Казань, 2014), V Международная конференция «Системный анализ и информационные технологии» (САИТ, Красноярск, 2013), XVI, XVIII и XIX Международные научные конференции «Решетневские чтения» (г. Красноярск, 2012, 2014, 2015 гг.), и др. Отдельные результаты работы обсуждались на научном семинаре института информационных технологий университета г. Ульм (Германия,

2014). Диссертация в целом обсуждалась на научно-технических семинарах кафедры системного анализа и исследования операций СибГАУ и кафедры систем автоматизированного проектирования (РК6) НИУ МГТУ им. Н.Э. Баумана.

Публикации. По материалам данной работы опубликовано 22 печатные работы, в том числе 6 статей в научных изданиях Перечня ВАК ([1-6]), 10 в изданиях, индексируемых в международной базе Scopus [7-16], из них 4 индексированы также в Web of Science [9, 10, 11, 16].

Структура работы. Работа состоит из введения, четырех глав, заключения и списка литературы.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Во **введении** обоснована актуальность работы, сформулирована цель и поставлены задачи исследования, рассмотрены вопросы научной новизны и практической ценности проведенных исследований, изложены основные положения, выносимые на защиту.

Первая глава работы посвящена исследованию различных бионических алгоритмов оптимизации, получивших свое развитие в последние десятилетия.

Установлено, что одним из наиболее эффективных методов для решения сложных задач оптимизации с вещественными переменными является так называемый стайный (роевый) алгоритм PSO, идея которого почерпнута из социального поведения групп животных (стаи птиц, косяка рыб или стада копытных). Работа алгоритма начинается с создания множества потенциальных решений (популяции индивидов) случайным образом. Каждый индивид имеет свои координаты и передвигается с некоторой скоростью, которые изменяются в процессе работы алгоритма в зависимости от результатов, достигнутых как самим индивидом, так и популяцией в целом. Таким образом, моделируется многоагентная система, где агенты-индивиды двигаются к оптимальному решению, обмениваясь при этом информацией с соседями.

Из последних разработанных бионических методов наиболее похожи на стайный алгоритм следующие: алгоритм поиска кукушек, алгоритм летучих мышей, алгоритм светлячков и алгоритм поиска стаей волков. Алгоритм поиска стаей волков копирует процесс охоты. Алгоритмом светлячков имитируется привлечение светлячков друг к другу, которое зависит от яркости свечения особей. Алгоритм поиска кукушек копирует их поведение во время откладывания яиц в процессе вынужденного гнездового паразитизма. За основу алгоритма летучих мышей взято их свойство эхолокации, используемое для обнаружения добычи и препятствий, а также для обеспечения возможности разместиться в темноте на насесте.

Все эти алгоритмы продемонстрировали свою эффективность при решении различных сложных прикладных задач и являются

взаимозаменяемыми. В ходе выполнения диссертации было проведено исследование эффективности перечисленных бионических алгоритмов с целью выявления наиболее успешного из них при решении различных задач условной и безусловной оптимизации. Исследование проводилось на репрезентативном множестве тестовых задач оптимизации, принятых в мировом сообществе по эволюционным вычислениям. Размерность решаемых задач варьировалась от двух до пятидесяти. Результаты оценивались по двум критериям: минимальное количество вычислений целевой функции, требуемое для получения решения с заданной точностью, и минимальная ошибка, получаемая алгоритмом после выполнения заданного количества вычислений целевой функции.

В итоге было установлено, что нельзя с уверенностью утверждать, что тот или иной алгоритм предпочтительнее для конкретной задачи, так как лучшие результаты достигались для разных задач различными алгоритмами. Установлено также, что лучший результат даже для одной и той же задачи при изменении ее размерности получается различными алгоритмами. Кроме того, заранее определить, сколько индивидов достаточно для осуществления поиска решения, невозможно, так как нужно определить такое их число, чтобы за определенное количество вычислений целевой функции было достигнуто оптимальное решение, с одной стороны, и чтобы этих вычислений было как можно меньше, с другой. Таким образом, помимо сложности выбора алгоритма для решения той или иной задачи, возникает проблема с настройкой главного параметра любого бионического метода – размера популяции. Все перечисленное серьезно затрудняет применение данных алгоритмов на практике даже для специалистов, не говоря уже о конечных пользователях, не являющихся экспертами в области стохастической оптимизации.

Полученные выводы послужили предпосылкой для разработки нового самонастраивающегося оптимизационного метода и его модификаций для решения задач условной и безусловной оптимизации с бинарными или вещественными переменными, основанных на коллективном поведении перечисленных выше алгоритмов. Данный подход получил наименование «Co-Operation of Biology Related Algorithms» (COBRA) и описан во **второй главе** диссертационной работы.

Главная идея метода заключается в генерировании пяти популяций (по одной для каждого бионического алгоритма из первой главы), которые в дальнейшем коллективно решают задачу оптимизации на основе конкуренции и кооперации. При этом алгоритм является самонастраивающимся, то есть процесс настройки размера популяции для каждого алгоритма-компонента автоматизирован: размер каждой популяции может как увеличиваться, так и уменьшаться в зависимости от того, как менялось значение функции пригодности на итерации. Другими словами, если на t-ой итерации средняя пригодность индивидов k-й (k =

1,2,3,4,5) популяции лучше средней пригодности других популяций, то *k*-я популяция считается «победителем», а все остальные «проигравшими». Из «проигравших» популяций удаляются индивиды – они добавляются к «победившей» популяции. Таким образом, определяется лучший алгоритм для задачи на каждой итерации. С другой стороны, общее число всех индивидов может тоже либо увеличиваться, либо уменьшаться. Если на протяжении некоторого заданного числа поколений значение функции пригодности не улучшается, то увеличивается размер всех популяций, и наоборот, если на протяжении некоторого заданного числа поколений значение функции пригодности только улучшается, то размер всех популяций уменьшается.

Кроме того, все популяции сотрудничают друг с другом. Они обмениваются индивидами: худшие индивиды одной популяции заменяются лучшими индивидами другой популяции, тем самым передается информация о наилучших решениях полученных всем коллективом алгоритмов в целом. Схема алгоритма COBRA представлена ниже.

Схема алгоритма Co-Operation of Biology Related Algorithms

установить минимальный размер будущих популяций
установить максимальный размер каждой из будущих популяций и общий максимальный размер
инициализация 5 популяций ***P, W, F, C, B*** *минимального размера случайным образом*
инициализация соответствующих параметров для каждого алгоритма
определение ***GBest***
пока *не выполнится критерий остановки*
    выполнить для популяции ***P*** *стайный алгоритм*
    выполнить для популяции ***W*** *алгоритм стай волков*
    выполнить для популяции ***F*** *алгоритм светлячков*
    выполнить для популяции ***C*** *алгоритм поиска кукушек*
    выполнить для популяции ***B*** *алгоритм летучих мышей*
    вычисление средней пригодности популяции: популяция с максимальной пригодностью – «победитель», прочие популяции – «проигравшие»
    для *каждой «проигравшей» популяции*
        сократить размер «проигравшей» популяции на 10% (не меньше минимального)
    конец цикла
    пока *размер победившей популяции и всех популяций в совокупности не превысит максимальное допустимое значение*
        увеличить размер победившей популяции на суммарное число удаленных индивидов из проигравших популяций
    конец цикла
    если *совершенно* $100 \cdot k, k = 1,2,...$ *вычислений целевой функции*
        осуществить миграцию лучших индивидов
    если *сменилось* $10 \cdot k, k = 1,2,...$ *поколений*
        если *GBest не улучшалось*

пока *размер всех популяций в целом не превысит максимального*
увеличить размер каждой популяции на K индивидов
конец цикла
если *GBest все время улучшалось*
пока *размер всех популяций больше минимального без K индивидов уменьшить размер каждой популяции на K индивидов*
конец цикла
обновить значение ***GBest***
конец цикла

Все пять алгоритмов-компонент и COBRA были исследованы на 28 специально спроектированных тестовых задачах из CEC 2013 Special Session on Real-Parameter Optimization, среди которых пять унимодальных функций, 15 базовых многоэкстремальных функций и 8 составных функций, являющихся специальными комбинациями базовых. Размерность задач – 2, 3, 5, 10, 30, 50 переменных. Для каждой задачи программа запускалась 51 раз. В итоге сравнение проводилось по двум критериям: самый лучший результат, полученный алгоритмом для данной задачи, и лучший средний результат. Количество разрешенных вычислений целевой функции для всех алгоритмов задавалось одинаковым.

Численные эксперименты показали, что коллективный метод COBRA показывает все более и более лучшие результаты по обоим критериям с увеличением размерности задач. По первому критерию коллективный алгоритм «выигрывал» у своих компонент на 11 задачах при размерности 5, на 17 при размерности 10 и на 26 при размерности 30. По второму критерию были следующие результаты: 19 «побед», 24 «победы», 28 «побед» при размерностях 5, 10 и 30, соответственно. На задачах с пятьюдесятью переменными COBRA выигрывает на всех тестовых задачах по обоим критериям.

Таким образом, описанный новый самонастраивающийся алгоритм оптимизации на основе коллективного поведения алгоритмов стайного интеллекта показал свою эффективность на репрезентативном множестве тестовых задач безусловной оптимизации с вещественными переменными. С ростом размерности задачи превосходство коллективного алгоритма становится неоспоримым, а следовательно, этот алгоритм можно применять вместо составляющих его методов, снимая, тем самым, вопрос о выборе наиболее эффективного из них.

Сравнение алгоритма COBRA с участниками соревнования специальной сессии Конгресса по эволюционным вычислениям (CEC 2013) на задачах размерности 50 показало высокую конкурентоспособность предложенного подхода.

Разработчики бионических алгоритмов часто модифицируют их с целью обеспечения возможности решения комбинаторных задач. Поэтому в дальнейшем алгоритм COBRA был адаптирован для решения задач

безусловной оптимизации с бинарными переменными. Так как первоначально все алгоритмы-компоненты были созданы для решения оптимизационных задач с вещественными переменными, для бинаризации эвристики COBRA была применена методика, использующая логистическую функцию (Kennedy J., Eberhart R., 1997):

$$s(v) = \frac{1}{1 + \exp(-v)}.$$

Логистическая функция вычисляется либо от значения скорости, либо от значения координат индивида (зависит от алгоритма). Далее генерируется случайное число из равномерного распределения на отрезке [0;1] и, если оно меньше значения логистической функции для некоторого индивида из популяции, то соответствующая координата индивида принимает значение 1, в противном случае – 0. Данная модификация была названа COBRA-b. Работоспособность алгоритма COBRA-b и целесообразность его применения были также установлены проверкой на том же множестве тестовых задач. Исследования показали, что алгоритм COBRA-b показывает лучшие результаты, чем соответствующие модификации его компонент, а значит и этот метод можно применять вместо составляющих его алгоритмов.

В дальнейшем эвристика COBRA была модифицирована для решения задач условной оптимизации, разработанная модификация была названа COBRA-c. Модификация коллективного алгоритма заключается в учете ограничений следующими способами: методом динамических штрафов, с помощью правил Дэба, а также методом, использующим принципы селекции.

Метод динамических штрафов – алгоритм учета ограничений, для которого пригодность каждого индивида x на t-ой итерации вычисляется по следующей формуле:

$$fitness(x) = f(x) + \delta \cdot (C \cdot t)^{\alpha} \cdot \sum_{j=1}^{m} f_j^{\beta}(x),$$

где $f(x)$ – целевая функция, m – количество функций ограничений, $f_j^{\beta}(x)$ – штраф за нарушение j-го ограничения, C – константа, $\delta = 1$ для задач минимизации и $\delta = -1$ для задач максимизации.

По правилам Дэба лучшее из двух решений определяется следующим образом:

1) если оба решения удовлетворяют ограничениям задачи, то лучшее определяется по значению целевой функции;

2) если одно из решений удовлетворяет ограничениям задачи, а другое нет, то лучшим решением будет то, которое удовлетворяет ограничениям, независимо от значения целевой функции;

3) если оба решения не удовлетворяют ограничениям, то лучшим считается то, для которого сумма нарушений ограничений меньше.

Метод учета ограничений, использующий селекцию, заключается в том, что и оптимизируемая функция, и функции степени нарушенности ограничений изначально считаются целевыми функциями, однако, какую из них нужно минимизировать (максимизировать) на каждой итерации для каждого индивида, определяется с помощью турнирной селекции.

Модификация COBRA-c была протестирована с помощью 18 задач, представленных на конкурсе алгоритмов условной оптимизации конгресса по эволюционным вычислениям CEC 2010. В соответствии с правилами конкурса размерность задач была равна 10 и 30, программа запускалась для каждой задачи 25 раз. Максимальное число вычислений целевой функции было установлено равным 200000 для задач размерности 10 и 600000 для задач размерности 30. По результатам тестирования была доказана эффективность разработанного метода и его преимущество над компонентами. Кроме того, алгоритм COBRA-c сравнивался с алгоритмами, представленными в упомянутом конкурсе, и была установлена его высокая конкурентоспособность в сравнении с альтернативными подходами.

Третья глава диссертации посвящена разработке алгоритмического обеспечения для автоматизации проектирования искусственных нейронных сетей и машин опорных векторов, при помощи методов, описанных во второй главе.

Для автоматического генерирования нейронной сети необходимо решить две задачи безусловной оптимизации: одна задача связана с выбором структуры сети, а другая – с настройкой весовых коэффициентов. Для их решения были использованы алгоритмы COBRA-b и COBRA, соответственно. Каждая структура нейронной сети представляется в виде бинарной строки следующим образом:

1) Устанавливается максимально допустимое число скрытых слоев и число нейронов на каждом слое.

2) Для настройки структуры нейронной сети используются 15 наиболее известных функций активации (сигмоиды, гауссиан, пороговая, линейная и т.д).

3) Каждый нейрон кодируется бинарной строкой из четырех бит (нейрон, закодированный бинарной строкой вида «0000», в структуре отсутствует).

4) Для определения активационной функции, закрепленной за нейроном, двоичное число, которым он был закодирован, переводится в десятичное число, которое определяет номер активационной функции (1, 2, …, 15).

Таким образом, каждая структура нейронной сети представляется бинарной строкой определенной длины. То есть проблему выбора структуры нейронной сети можно рассматривать как задачу безусловной оптимизации с бинарными переменными. В итоге при генерировании

нейронной сети алгоритмом COBRA-b сначала инициализировалась популяция структур в виде заданного множества бинарных строк. Для каждой структуры приближенно подбирались весовые коэффициенты, то есть решалась задача безусловной оптимизации с вещественными переменными алгоритмом COBRA, а затем выполнялась «основная» работа эвристики и выбиралась та структура, которая решает исходную задачу лучше всего. Конечным этапом автоматического генерирования нейросети под решаемую задачу является полное обучение алгоритмом COBRA найденной алгоритмом COBRA-b лучшей структуры.

Далее разработанные алгоритмы были применены для автоматической настройки машин опорных векторов. Основная идея метода опорных векторов – перевод исходных обрабатываемых векторов в пространство заданной размерности и поиск разделяющей гиперплоскости с максимальным зазором в этом пространстве. Предлагается рассматривать общий случай, когда неизвестно можно ли разделить объекты из разных классов линейно. В этом случае наиболее эффективным способом решения задачи классификации или прогнозирования методом опорных векторов является перенос данных-векторов в пространство большей размерности и поиск оптимальной разделяющей гиперплоскости уже для данных в новом пространстве. То есть основная идея этого подхода заключается в том, что исходные данные в пространстве большей размерности будут линейно разделимы. Проще всего реализовать данную идею, применив функцию-ядро. В диссертации описана работа машин опорных векторов с полиномиальной функцией-ядром. Таким образом, автоматическое генерирование машин опорных векторов сводилось к решению двух задач оптимизации с вещественными переменными: условной и безусловной. Сначала для решения задачи анализа данных методом опорных векторов определялись параметры функции-ядра алгоритмом COBRA (задача безусловной оптимизации), затем для полученного ядра алгоритмом COBRA-c строилась оптимальная разделяющая гиперплоскость (задача условной оптимизации).

Далее был разработан алгоритм генерирования коллективов классификаторов на основе метода опорных векторов.

В **четвертой** главе проведена практическая апробация методов, описанных в третьей главе, на различных задачах классификации и прогнозирования. Разработанные алгоритмы генерирования технологий интеллектуального анализа данных были использованы для решения двух задач банковского скоринга и двух задач медицинской диагностики из репозитория UCI, двух задач распознавания образов из репозитория KEEL. Затем ими же решались четыре задачи категоризации текста, взятые с конкурса DEFT’07 и DEFT’08. Все перечисленные задачи анализа данных являются задачами классификации. Далее были решены задача прогнозирования успешности обучения абитуриентов и задача

прогнозирования процесса деградации солнечных батарей космических аппаратов. В итоге, результаты, полученные для всех задач, оказались в числе лучших или просто лучшими по сравнению с результатами, полученными другими известными подходами.

Для примера в таблице 1 представлены полученные различными алгоритмами, включая разработанные в диссертации методы (машины опорных векторов «COBRA+SVM» и нейронные сети «COBRA+ANN»), результаты (доля правильно классифицированных примеров тестовой выборки) для трех задач категоризации текста, взятых с конкурса DEFT'07. В скобках указано место, занимаемое методом среди альтернативных подходов. Последний столбец содержит общий ранг алгоритма по трем задачам.

Таблица 1. Результаты, полученные для задач конкурса DEFT07

Исследовательская группа	Книги	Игры	Дебаты	Ранг
J.-M. Torres-Moreno	0.603 (2)	0.784 (1)	0.720 (1)	1
G. Denhiere	0.599 (3)	0.699 (6)	0.681 (6)	4
S. Maurel	0.519 (7)	0.706 (5)	0.697 (5)	5
M. Vernier	0.577 (4)	0.761 (3)	0.673 (7)	6
E. Crestan	0.529 (6)	0.673 (8)	0.703 (4)	7
M. Plantie	0.472 (9)	0.783 (2)	0.671 (9)	8
A.-P. Trinh	0.542 (5)	0.659 (9)	0.676 (8)	9
M. Genereux	0.464 (10)	0.626 (10)	0.569 (12)	11
E. Charton	0.504 (8)	0.619 (11)	0.616 (10)	10
A. Acosta	0.392(11)	0.536(12)	0.582 (11)	12
COBRA+SVM	0.619 (1)	0.696 (7)	0.714 (2)	3
COBRA+ANN	0.613 (2)	0.727 (4)	0.709 (3)	2

Нетрудно видеть, что предложенные в диссертационной работе подходы, хотя и являются универсальными, превосходят по эффективности большинство специализированных методов, попавших в число победителей соревнования.

В **заключении** диссертации приведены основные результаты и выводы.

ОСНОВНЫЕ РЕЗУЛЬТАТЫ И ВЫВОДЫ

В ходе выполнения диссертационной работы получены следующие результаты:

1) Разработан, реализован и апробирован новый коллективный самонастраивающийся метод решения задач безусловной оптимизации с вещественными переменными.

2) Разработан, реализован и апробирован новый коллективный

самонастраивающийся метод решения задач безусловной оптимизации с бинарными переменными.

3) Выполнены модификации разработанных алгоритмов, позволяющие решать задачи условной оптимизации.

4) На основе разработанных методов и их модификаций построены алгоритмы автоматизированного генерирования информационных технологий интеллектуального анализа данных на основе искусственных нейронных сетей и машин опорных векторов.

5) Эффективность разработанных подходов продемонстрирована в ходе решения 12 прикладных задач оптимизации и анализа данных из различных областей науки и техники.

Таким образом, в данной работе разработаны, реализованы и проверены новые оптимизационные методы, а также эффективные алгоритмические схемы их применения при автоматизированном генерировании информационных технологий интеллектуального анализа данных, позволяющие успешно решать широкий круг практических задач и имеющие существенное значение для развития теории и практики системного анализа и обработки информации.

ПУБЛИКАЦИИ ПО ТЕМЕ ДИССЕРТАЦИИ:

Статьи в ведущих рецензируемых научных изданиях Перечня ВАК

1. **Akhmedova, Sh.A.** SVM-based classifier ensembles design with co-operative biology inspired algorithm / Sh.A. Akhmedova, E.S. Semenkin // Вестник СибГАУ. – 2015. - № 1 (16). – С.22-27.
2. **Ахмедова, Ш.А.** Проектирование систем автоматического определения субъективного мнения кооперативными бионическими алгоритмами / Ш.А. Ахмедова, Е.С. Семенкин // Системы управления и информационные технологии. – 2014. – №3.1 (57). – С. 104-108.
3. **Ахмедова, Ш.А.** Интеллектуальные информационные технологии для прогнозирования успешности учебной деятельности абитуриентов / Ш.А. Ахмедова, С.Р. Вишневская, А.А. Коромыслова // Вестник СибГАУ. – 2014. – № 3 (54). – С. 16-20.
4. **Ахмедова, Ш.А.** Кооперативный бионический алгоритм безусловной оптимизации / Ш.А. Ахмедова, Е.С. Семенкин // Программные продукты и системы. – 2013. – № 4 (104). – С. 133-136.
5. **Akhmedova, Sh.A.** New optimization metaheuristic based on co-operation of biology related algorithms / Sh.A. Akhmedova, E.S. Semenkin // Вестник СибГАУ. – 2013. – № 4 (50). – С. 92-99.
6. **Akhmedova, Sh.A.** Development and investigation of the effectiveness of the particle swarm optimization algorithm / Sh.A. Akhmedova // Вестник СибГАУ. – 2012. – № 4 (44). – С. 78-79.

Публикации в изданиях, индексируемых в международных базах

7. **Akhmedova, Sh.** Genetic Algorithm Based X-Ray Diffraction Analysis for Chemical Control of Aluminium Smelters Baths / Sh. Akhmedova, I. Yakimov, A. Zaloga, S. Burakov, E. Semenkin, P. Dubinin, O. Piksina, E. Andryushchenko // 12th International Conference on Informatics in Control, Automation and Robotics (ICINCO'2015) – Colmar, France. – 2015. – P. 32-39.

8. Skraba, A. Modelling and optimization of strictly hierarchical manpower system / A. Skraba, E. Semenkin, D. Kofjac, M. Semenkina, A. Znidarsic, M. Maletic, **Sh. Akhmedova**, C. Rozman, V. Stanovov // 12th International Conference on Informatics in Control, Automation and Robotics (ICINCO'2015)) – Colmar, France. – 2015. – P. 215-222.

9. Zaloga, A. On the Application of Co-Operative Swarm Optimization in the Solution of Crystal Structures from X-Ray Diffraction Data / A. Zaloga, S. Burakov, E. Semenkin, I. Yakimov, **Sh. Akhmedova**, M. Semenkina, E. Sopov // Advances in Swarm and Computational Intelligence, ICSI-CCI 2015, LNCS 9140. – 2015. – Part I. – P. 89-96.

10. **Akhmedova, Sh.** Data Mining Tools Design with Co-Operation of Biology Related Algorithms / Sh. Akhmedova, E. Semenkin // Advances in Swarm Intelligence, ICSI 2014, LNCS 8794. – 2015. – Part I. – P. 499-506.

11. **Akhmedova, Sh.** Co-Operation of Biology Related Algorithms Meta-Heuristic in ANN-Based Classifiers Design / Sh. Akhmedova, E. Semenkin // Congress on Evolutionary Computations of the IEEE World Congress on Computational Intelligence (CEC WCCI 2014). – Beijing, China. – 2014. – P. 867-873.

12. **Akhmedova, Sh.** Co-Operation of Biology Related Algorithms for Support Vector Machine Automated Design / Sh. Akhmedova, E. Semenkin, T. Gasanova, W. Minker // International Conference on Engineering and Applied Sciences Optimization (OPT-i'2014). – Greece. – 2014. – P. 1831-1837.

13. **Akhmedova, Sh.** Automatically Generated Classifiers for Opinion Mining with Different Term Weighting Schemes / Sh. Akhmedova, E. Semenkin, R. Sergienko // 11th International Conference on Informatics in Control, Automation and Robotics (ICINCO'2014). – Vienna, Austria. – 2014. – Vol. 2. – P. 845-850.

14. Semenkina, M. Spacecraft Solar Arrays Degradation Forecasting with Evolutionary Designed ANN-based Predictors / M. Semenkina, **Sh. Akhmedova**, E. Semenkin, I. Ryzhikov // Proceedings of the 11th International Conference on Informatics in Control, Automation and Robotics (ICINCO'2014). – Vienna, Austria. – 2014. – Vol. 1. – P. 421-428.

15. **Akhmedova, Sh.** Development and Investigation of Bio-logically Inspired Algorithms Cooperation Metaheuristic / Sh. Akhmedova, A. Shabalov // Genetic and Evolutionary Computation Conference Companion (GECCO'2013). – Amsterdam, Holland. – 2013. – P. 1417-1418.

16. **Akhmedova, Sh.** Co-Operation of Biology Related Algorithms /

Sh. Akhmedova, E. Semenkin // 2013 IEEE Congress on Evolutionary Computation (CEC'2013). – Mexico. – 2013. – P. 2207-2214.

Публикации в сборниках трудов конференций

17. Sergienko, R. Opinion Mining and Topic Categorization with Novel Term Weighting / R. Sergienko, T. Gasanova, **Sh. Akhmedova**, E. Semenkin, W. Minker // 5th Workshop on Computational Approaches to Subjectivity, Sentiment & Social Media Analysis (WASSA'2014). – Baltimore, USA. – 2014. – P. 84-89.

18. **Ахмедова, Ш.А.** Обучение машин опорных векторов для решения задач категоризации текста / Ш.А. Ахмедова, Е.С. Семенкин, Т.О. Гасанова, В. Минкер // XIV национальная конференция по искусственному интеллекту с международным участием КИИ-2014. – Казань. – 2014. – Том 2. – С. 222-230.

19. **Ахмедова, Ш.А.** Коллективный бионический алгоритм для решения задач безусловной оптимизации с бинарными переменными / Ахмедова Ш.А. // Труды XVIII Международной научной конференции «Решетневские чтения». – Красноярск. – 2014. – Том 2. – С. 17-18.

20. **Ахмедова, Ш.А.** Генерирование нейросетевых классификаторов кооперативными бионическими алгоритмами оптимизации / Ахмедова Ш.А. // Труды XVIII Международной научной конференции «Решетневские чтения». – Красноярск. – 2014. – Том 2. – С. 224-226.

21. **Ахмедова, Ш.А.** Коллективный метод оптимизации на основе стайных бионических алгоритмов / Ш.А. Ахмедова, Е.С. Семенкин // Системный анализ и интеллектуальные технологии (САИТ-2013). – Красноярск. – 2013 – Том 2. – С. 322-331.

22. **Ахмедова, Ш.А.** Формирование оптимального инвестиционного портфеля предприятия многокритериальным «стайным» алгоритмом / Ш.А. Ахмедова // Труды XVI Международной научной конференции «Решетневские чтения». – Красноярск. – 2012. – Том 2. – С. 471-472.

Зарегистрированные программные системы

23. **Ахмедова Ш.А.**, Семенкин Е.С. Коллективный метод условной оптимизации на основе стайных бионических алгоритмов. Свидетельство о государственной регистрации программы для ЭВМ №2013661150; зарег. в Реестре программ для ЭВМ 29.11.2013.

24. **Ахмедова Ш.А.**, Семенкин Е.С. Коллективный метод безусловной оптимизации на основе стайных бионических алгоритмов. Свидетельство о государственной регистрации программы для ЭВМ №2014610132; зарег. в Реестре программ для ЭВМ 09.01.2014.

25. **Ахмедова Ш.А.**, Семенкин Е.С. Коллективный метод безусловной оптимизации на основе стайных бионических алгоритмов для решения задач с бинарными переменными. Свидетельство о

государственной регистрации программы для ЭВМ №2014613715; зарег. в Реестре программ для ЭВМ 03.04.2014.

26. **Ахмедова Ш.А.**, Семенкин Е.С. Машины опорных векторов, автоматически сгенерированные коллективными методами условной и безусловной оптимизации на основе стайных бионических алгоритмов. Свидетельство о государственной регистрации программы для ЭВМ №2014613716; зарег. в Реестре программ для ЭВМ 03.04.2014.

27. **Ахмедова Ш.А.**, Семенкин Е.С. Автоматическое генерирование нейронных сетей алгоритмами стайного типа. Свидетельство о государственной регистрации программы для ЭВМ №2014616237; зарег. в Реестре программ для ЭВМ 26.08.2014.

28. **Ахмедова Ш.А.**, Семенкин Е.С. Проектирование SVM-классификаторов коллективными бионическими алгоритмами с выбором информативных входов. Свидетельство о государственной регистрации программы для ЭВМ №2015611847; зарег. в Реестре программ для ЭВМ 06.02.2015.

Ахмедова Шахназ Агасувар кызы

Коллективный самонастраивающийся метод оптимизации
на основе бионических алгоритмов

Автореферат

Подписано к печати Формат 60x84/16

Уч. изд. л. 1.0 Тираж 100 экз. Заказ № _____

Отпечатано в отделе копировальной и множительной техники СибГАУ
660037, г. Красноярск, пр. им. газ. «Красноярский рабочий», 31