

ОТЗЫВ

официального оппонента

на диссертацию *Становова Владимира Вадимовича*

**«Самонастраивающиеся эволюционные алгоритмы
формирования систем на нечеткой логике»,**

представленную на соискание учёной степени

кандидата технических наук

по специальности 05.13.01 –

**«Системный анализ, управление и обработка информации
(космические и информационные технологии)»**

1. АКТУАЛЬНОСТЬ ПРОБЛЕМЫ ИССЛЕДОВАНИЯ

В настоящее время практически во всех сферах человеческой деятельности наблюдается рост числа нетривиальных и плохо структурированных задач, так называемых задач в условиях неопределенности, требующих принципиально новых подходов к своему решению. При этом приходится сталкиваться с вопросами адаптации применяемых алгоритмов и методов к изменяющимся параметрам решаемой задачи, а решать вопросы о возможности параллельной обработки больших массивов данных.

Актуальность темы диссертационной работы Становова В.В обусловлена необходимостью решения важной и актуальной задачи, связанной с развитием (в том числе, с модификаций и гибридизацией) эволюционных алгоритмов в контексте проблемы формирования классификационных систем, основанных на нечеткой логике.

Вопросам классификации данных посвящено огромное количество работ отечественных и зарубежных авторов. При этом в качестве основополагающих в контексте решения задач классификации можно выделить направления, основанные на принципах интеллектуального анализа данных и предполагающие применение искусственных нейронных сетей, искусственных иммунных систем, машины опорных векторов и систем нечеткого логического вывода.

Главным достоинством систем нечеткого логического вывода является возможность использования при их разработке лингвистических правил логического вывода с последующим их объединением в базы правил, которые представляют собой модель «белого ящика», то есть модель с явно прописанными правилами, определяющими причинно-следственные связи между входными и выходными параметрами системы. При этом количество классов может быть произвольным: нужное количество классов закладывается сразу в систему нечеткого логического вывода и нет необходимости в разработке бинарных классификаторов с последующим формированием мультиклассовых. Используемые

в системах нечеткого логического вывода правила оперируют лингвистическими понятиями, в связи с чем могут быть легко восприняты экспертом.

Главная проблема, с которой приходится сталкиваться разработчикам систем нечеткого логического вывода, связана с формированием структуры самих правил нечеткого вывода (так как требуется определить состав компонент правила и логическую связку между ними), с «тонкой» настройкой параметров правил (так как требуется определить тип и параметры применяемых функций принадлежности), с формированием самой базы правил нечеткого вывода (так как требуется определить весовые коэффициенты правил). Очевидно, что адекватное решение этой проблемы, от которого в итоге зависит качество классификации данных, невозможно без привлечения современного инструментария оптимизации, в роли которого в последнее время активно применяются различные эволюционные алгоритмы, в том числе, генетические алгоритмы. Последние достижения в сфере эволюционных алгоритмов оптимизации свидетельствуют о целесообразности проведения исследований в направлении их гибридизации, самоконфигурации и самонастройки с целью обеспечения более высокой сходимости оптимизационных алгоритмов и минимизации временных затрат на получение искомого решения.

Еще одна проблема, с которой приходится сталкиваться специалистам при разработке любых интеллектуальных классификаторов данных, связана с проблемой несбалансированности данных обучающей выборки, заключающейся в разном соотношении данных разных классов. От правильности решения данной проблемы напрямую зависит качество разрабатываемых классификаторов, поскольку зачастую именно миноритарные классы представляют наибольший целевой интерес при решении задач классификации данных. Для решения этой проблемы могут быть использованы подходы, основанные на балансировке данных, на разработке соответствующих версий самих алгоритмов классификации, на сочетании двух предыдущих подходов, на ансамблировании классификаторов для несбалансированных наборов данных.

Следует отметить, что ни для одной из вышеуказанных проблем в настоящее время не найдено надлежащих решений, несмотря на большой к ним интерес и огромное число публикаций по данным вопросам.

Таким образом, с учетом всего вышеизложенного можно сделать вывод об актуальности темы рассматриваемой диссертационной работы, направленной на разработку самонастраивающихся эволюционных алгоритмов формирования классификационных систем, основанных на нечеткой логике.

2. СТЕПЕНЬ ОБОСНОВАННОСТИ НАУЧНЫХ ПОЛОЖЕНИЙ, ВЫВОДОВ И РЕКОМЕНДАЦИЙ, СФОРМУЛИРОВАННЫХ В ДИССЕРТАЦИИ

Целью диссертационной работы является повышение качества и интерпретируемости нечетких классификаторов, а также снижение требуемых вычислительных ресурсов при формировании этих классификаторов посредством применения самонастраивающихся эволюционных алгоритмов.

Для достижения этой цели в диссертационной работе выполнен подробный анализ известных подходов к формированию правил нечеткого вывода и соответствующих им баз правил с целью выявления наиболее эффективных подходов и направлений. Кроме того, автором тщательно исследованы с применением репрезентативного множества тестовых задач из репозитория задач машинного обучения методы самонастройки эволюционных алгоритмов оптимизации. По результатам выполненных исследований в диссертационной работе разработан самонастраивающийся эволюционный алгоритм формирования баз нечетких правил для решения задач классификации несбалансированных данных; метод гибридизации Питтсбургского и Мичиганского подходов в эволюционном алгоритме формирования баз нечетких правил; метод селекции обучающих примеров для нечеткого классификатора.

Научные положения, теоретические выводы и практические рекомендации, предлагаемые автором, обоснованы строгими математическими описаниями и результатами сравнительного анализа разработанных автором алгоритмов и методов с известными аналогами, применяемыми в контексте решаемой задачи, с использованием основных метрик качества классификации. Программные системы, реализованные на основе предложенных алгоритмов и методов, прошли апробацию при решении тестовых и реальных прикладных задач.

Результаты исследования широко представлены на многочисленных конференциях и, а также опубликованы в ведущих реферируемых научных журналах, в том числе, 8 статей в журналах из Перечня ВАК РФ и 9 статей в изданиях, в изданиях, индексируемых в международных базах Web of Science и Scopus.

Достоверность научных положений, выводов и рекомендаций обеспечены корректным применением методов статистической обработки данных, теории вероятностей, эволюционных вычислений, оптимизации, нечеткой логики, системного анализа данных, моделирования динамических систем, выявления закономерностей в исходных данных, результатами апробации на Всероссийских и Международных конференциях, а также результатами исследований в рамках международных и отечественных проектов.

3. НАУЧНАЯ НОВИЗНА

В диссертации получены следующие результаты, характеризующиеся научной новизной:

- самонастраивающийся эволюционный алгоритм формирования нечетких систем для решения задач классификации с представлением баз правил в форме матриц переменной размерности, отличающийся от известных использованием оценки достоверности правил при назначении им весовых коэффициентов и за счет этого превосходящий по эффективности другие алгоритмы эволюционного построения нечетких систем;
- метод гибридизации Питтсбургского и Мичиганского подходов в эволюционном алгоритме формирования баз нечетких правил, отличающийся от

известных использованием при построении новых правил вероятностной процедуры выбора релевантных нечетких термов и позволяющий существенно повысить точность классификации на первых поколениях работы эволюционного алгоритма;

– метод селекции примеров для обучения классификаторов, отличающийся от известных адаптивной вероятностной процедурой организации подвыборок и назначения весовых коэффициентов и позволяющий одновременно повысить точность классификации и снизить объем требуемых для этого вычислительных ресурсов;

– метод самонастройки эволюционных алгоритмов, отличающийся от известных схемой оценки успешности операторов, применяемых многократно к каждому индивиду, и позволяющий настраивать вероятности применения эвристик в части алгоритма, основанной на Мичиганском подходе.

Вышеизложенным аргументирована новизна основных результатов диссертационной работы и их соответствие требованиям Положения ВАК РФ.

4. ТЕОРЕТИЧЕСКАЯ ЗНАЧИМОСТЬ

Теоретическая значимость результатов диссертационного исследования заключается в разработке самонастраивающегося эволюционного алгоритма формирования нечетких систем с целью получения компактных и точных баз правил посредством использования кодирования в форме матриц переменной размерности; метода гибридизации Питтсбургского и Мичиганского подходов на основе принципов самоконфигурации; метода активной селекции обучающих примеров, применяемых при разработке для нечетких классификаторов в контексте решения проблемы несбалансированности данных.

Предложенные в диссертационной работе решения вносят существенный вклад в теорию и практику технологий формирования эффективных классификационных систем нечеткого вывода с применением современных концептуальных идей теории эволюционных алгоритмов.

Полученные в ходе диссертационного исследования результаты реализованы в виде математического, алгоритмического и программного обеспечения.

Совокупность полученных результатов позволяет существенно расширить спектр практического применения классификационных систем, основанных на нечеткой логике.

5. ПРАКТИЧЕСКАЯ ЦЕННОСТЬ

Практическая ценность диссертационной работы Становова В.В. заключается в разработке с применением предложенных автором алгоритмов и методов программной системы, компоненты которой защищены 3 свидетельствами о государственной регистрации программы для ЭВМ. Данная программная система может использоваться в разных областях науки и техники при решении широкого спектра задач классификации.

Программная система позволяет ускорить процесс формирования базы в системе нечеткого вывода посредством использования алгоритмов самонастройки, а также благодаря применению алгоритмов, снижающих число пересчетов значений степеней принадлежности и весов правил.

Успешное тестирование программной системы на задачах классификации из различных прикладных областей (в частности, на задачах распознавания изображений, банковского скоринга и медицинской диагностики и т.п., взятых из известных репозиториях данных машинного обучения) позволяет сделать вывод о перспективности ее применения и при решении реальных практических задач.

Разработанные в диссертационной работе алгоритмы и методы были использованы при выполнении актуальных исследований в рамках международных проектов, в частности, двух российско-германских проектов «Распределенные интеллектуальные информационные системы обработки и анализа мультилингвистической информации в диалоговых информационнокоммуникационных системах» и «Математическое и алгоритмическое обеспечение автоматизированного проектирования аппаратно-программных комплексов интеллектуальной обработки мультилингвистической информации в распределенных высокопроизводительных системах космического назначения» (г. Ульм), российско-словенского проекта «Manpower control strategy determination with selfadapted evolutionary and biologically inspired algorithms» (г. Марибор).

Кроме того, предложенные автором решения были апробированы в проектах «Развитие теоретических основ автоматизации математического моделирования физических систем на основе экспериментальных данных» и «Разработка теоретических основ эволюционного проектирования интеллектуальных информационных технологий анализа данных» в рамках тематического плана ЕЗН СибГАУ.

Достоинством диссертационной работы, подтверждающим ее несомненную практическую направленность и значимость, является то, что она была поддержана Фондом содействия развитию малых форм предприятий в научно-технической сфере по программе «У.М.Н.И.К» в рамках НИОКР «Разработка программного обеспечения интеллектуального анализа данных "FuzzyMiner"» на 2014-2016 гг., а также Российским Фондом Фундаментальных Исследований в рамках проекта № 16-31-00349 «Разработка алгоритмов и подходов к повышению качества и скорости формирования технологий интеллектуального анализа данных посредством снижения размерности данных» на 2016-2017 гг.

Следует отметить востребованность выполненных автором разработок и в учебном процессе Института информатики и телекоммуникаций СибГАУ при выполнении лабораторных и курсовых работ.

6. КРИТИЧЕСКИЕ ЗАМЕЧАНИЯ

При изучении материалов диссертационной работы выявлены следующие недостатки.

1. Подробная информация о мерах качества классификации приведена в самом конце диссертационной работы – в главе 4, хотя по смыслу должна была бы находиться в главе 1. Автор не обосновывает, почему при расчетах для одних примеров он использует один набор мер качества, а для других – другой.

2. Следовало бы рассмотреть возможность использования при оценке качества классификаторов не только свертки критериев, но и современных эволюционных алгоритмов многокритериальной оптимизации.

3. Не исследован вопрос о перспективности/неперспективности использования иерархических систем нечеткого вывода, а также систем на основе нечетких множеств второго типа, реализующих дополнительные возможности при учете неопределенности.

4. Не следует использовать понятия «алгоритм» и «метод» как взаимозаменяемые.

5. Представление всех предложенных алгоритмов и методов носит описательный характер, необходимо было представить их в более структурированном виде, возможно, с применением соответствующих схем.

6. Отсутствие приложений к диссертационной работе, подтверждающих наличие актов внедрения и свидетельств о регистрации программных систем.

7. Наличие ряда орфографических ошибок и ошибок записи некоторых обозначений.

Однако указанные недостатки носят сугубо локальный характер и не снижают ценности полученных результатов.

7. ЗАКЛЮЧЕНИЕ

Представленная к защите диссертационная работа изложена последовательно, написана хорошим научно-техническим языком, является законченной научно-исследовательской работой, выполненной на актуальную тему, и содержит решение научной задачи, заключающейся в повышении качества и интерпретируемости нечетких классификаторов, а также в снижении требуемых вычислительных ресурсов при формировании данных классификаторов посредством применения самонастраивающихся эволюционных алгоритмов, и имеющей важное народно-хозяйственное значение при создании программных решений для различных отраслей народного хозяйства.

Представленная диссертационная работа может быть оценена как научно-квалификационная работа, отличающаяся научной новизной, а также теоретической и практической значимостью.

Диссертационная работа обладает внутренним единством, содержит новые научные результаты и положения, выдвигаемые для публичной защиты, и рекомендации по их практическому применению. Предложенные автором решения аргументированы и оценены по сравнению с другими известными решениями.

Оформление диссертационной работы соответствует требованиям, установленным Министерством образования и науки Российской Федерации.

Опубликованные автором научные работы достаточно полно отражают результаты исследований, представленные в диссертационной работе.

Основные результаты диссертационных исследований опубликованы в 38 научных работах, в том числе в 8 статьях в изданиях, входящих в Перечень ведущих периодических изданий ВАК Министерства образования и науки РФ, и 9 статьях, в изданиях, индексируемых в международных базах Web of Science и Scopus, материалах Всероссийских и Международных конференций. Автором получены 3 свидетельства о государственной регистрации программы для ЭВМ.

В автореферате полностью отражены основные результаты диссертационной работы и дано краткое изложение содержания проведенных исследований. Все требуемые разделы автореферата представлены в полном объеме, а их содержание соответствует содержанию диссертационной работы.

Считаю, что представленная работа соответствует всем требованиям, предъявляемым ВАК РФ к кандидатским диссертациям, а ее автор, Становов Владимир Вадимович, достоин присуждения ему ученой степени кандидата технических наук по специальности 05.13.01 – «Системный анализ, управление и обработка информации (космические и информационные технологии)».

Официальный оппонент:

доктор технических наук, профессор,

профессор кафедры

«Вычислительная и

прикладная математика»

ФГБОУ ВО «Рязанский государственный

радиотехнический университет»

Демидова Л.А.

« 22 » 11 2016 г.

Адрес места работы:

г. Рязань, 390005, ул. Гагарина, д. 59/1

<http://www.rsreu.ru>

Тел.: (4912) 46-03-03

E-mail: rgrtu@rsreu.ru

Подпись Демидовой Лилии Анатольевны заверяю.

Ученый секретарь
ученого совета РГРТУ



В.Н. Пржегорлинский