

«УТВЕРЖДАЮ»

Проректор по научной и
инновационной деятельности НИУ

«Белгородский государственный
университет»,

д.т.н., профессор

И.С. Константинов
Константинов И.С.

«09» июня 2014 г.



ОТЗЫВ

ведущей организации

на диссертационную работу Натальи Владимировны
Сапкиной «Восстановление закономерностей на основе
нечетких регрессионных моделей», представленную на
соискание ученой степени кандидата технических наук по
специальности

05.13.17 – «Теоретические основы информатики»

1. Актуальность темы диссертационной работы

Диссертационная работа Н.В. Сапкиной посвящена разработке информационной системы интеллектуального анализа данных, в которой реализуются нечеткие линейные регрессионные модели. Обнаружение закономерностей – одна из основных задач интеллектуального анализа данных. Особенный интерес к методам анализа данных возник в связи с развитием средств сбора и хранения информации, позволившим накапливать большие объемы данных. Конечно, существующие статистические методы представляют собой эффективный инструмент для построения функциональных зависимостей, но они основываются на обработке количественных данных и не применимы для тех случаев, когда информация является приближенной и представлена, например, нечеткими числами. Проблема разработки нечетких моделей для восстановления зависимостей решалась различными учеными, как отечественными (Р.А. Алиев, А.Э. Церковный, Г.А. Мамедова, Н.Г. Ярушкина), так и зарубежными (Н. Tanaka, P. Diamond, D. Dubois, M.S. Yang, M. Sakawa, M. Albrecht). Однако, в проведенных исследованиях отсутствует

комплексность подхода к реализации всех этапов моделирования. Необходимость совершенствования существующих методов нечеткого регрессионного моделирования за счет рассмотрения различных вариантов нечеткости исходных данных и/или параметров модели, и их реализации в рамках информационной системы интеллектуального анализа данных обуславливает актуальность диссертационного исследования.

Диссертационная работа выполнена в соответствии с одним из основных научных направлений Воронежского государственного университета «Математическое моделирование, программное и информационное обеспечение, методы вычислительной и прикладной математики и их применение к фундаментальным исследованиям в естественных науках».

2. Достоверность и новизна основных выводов и результатов диссертации

Основой проведенного Н.В. Сапкиной исследования является общая процедура линейного регрессионного моделирования, которая модифицирована для работы с приближенной информацией, представленной нечеткими числами L-R-типа. Исследование отличается логичностью и последовательностью рассмотрения проблем, соответствующих общей процедуре. К основным результатам диссертации, которые можно охарактеризовать как новые, относятся следующие:

- 1) Путем выявления свойств (коммутативность, ассоциативность, наличие нейтрального и единичного элементов) арифметических операций над нечеткими числами обосновано существование алгебр с одной операцией (сложения или умножения). Для группоидов со сложением и умножением определены нейтральные (единичные) элементы, поэтому каждый из этих группоидов является, в то же время, и моноидом. Также установлено, что группоид со сложением не имеет обратного элемента; для группоида с умножением найден обратный элемент и соответствующие ограничения на функции принадлежности, которые обеспечивают его существование. Показано, что группоид с умножением образует абелеву группу. Кроме того, выведены формулы для ряда конструкций (умножение на обратное число, на нечеткий нуль,

на нечеткую единицу, возведение нечеткого числа в квадрат и др.), которые были положены в основу преобразований при определении коэффициентов нечетких регрессионных моделей. Перечисленные теоретические результаты, основанные на определении арифметических операций над нечеткими числами, введенных А. Пегатом, являются новыми.

2) На основе метода наименьших квадратов определены оценки параметров нечеткой парной и множественной (в матричном виде формулы 3.35) линейных регрессионных моделей для случаев, когда исходная информация для построения модели является полностью или частично приближенной. Множественная регрессионная модель с нечеткими коэффициентами при условии, что независимые переменные принимают обычные числовые значения, а зависимая переменная – приближенные в виде нечетких чисел, рассмотрена впервые. Данная модель актуальна для случая, когда значения зависимой переменной прогнозируются. Отличительной особенностью подхода является возможность построения функций принадлежности нечетких коэффициентов на основе параметров функций принадлежности наблюдаемых данных.

3) Оценка качества (точность, адекватность) предложенных нечетких регрессионных моделей осуществляется на основе анализа остаточной последовательности классическими методами, при этом предложено несколько способов ее формирования, которые отличаются различными вариантами применения процедур дефазификации.

4) Для определения существенных переменных предложено несколько методов, которые основаны на коэффициенте корреляции, стандартизированном уравнении множественной линейной регрессии и нейросетевом подходе, отличительной особенностью которого является его адаптация для использования приближенной исходной информации в виде нечетких чисел в процедуре обучения.

5) Предложена структура информационной системы интеллектуального анализа данных и реализовано программное обеспечение ее компонент. Отличительной особенностью структуры, основанной на технологии информационного хранилища, является наличие приложений, отвечающих за администрирование и непосредственно анализ данных. Структура программного обеспечения включает модули, реализующие все этапы нечеткого регрессионного

моделирования, при этом в качестве инвариантной составляющей выступает блок нечеткой арифметики, который может использоваться как самостоятельное приложение.

Перечисленные научные результаты получены Н.В. Сапкиной самостоятельно, опубликованы в печати, докладывались и обсуждались на международных и российских конференциях.

При проведении исследований корректно были использованы методы математической статистики, в частности, корреляционно-регрессионного анализа, алгебры, нечеткой арифметики, теории информационных систем. Теоретические положения и выводы диссертации строго обоснованы и, в большей части, подтверждаются результатами вычислительных экспериментов. Диссертация содержит значительное количество иллюстративных примеров, поясняющих особенности нечеткого регрессионного анализа. Достоверность научных результатов диссертации также подтверждена созданием информационной системы интеллектуального анализа данных, работоспособность которой проверена в рамках вычислительного эксперимента, что позволило соискателю сравнить подходы к восстановлению закономерностей в данных на основе классического и нечеткого линейного регрессионного моделирования. Полученные выводы экспериментального исследования полностью согласуются с теоретическими положениями.

Таким образом, научные результаты, выводы и основные положения диссертационной работы Н.И. Сапкиной обоснованы и достоверны.

3. Ценность диссертационной работы для науки и практики

Работа вносит также вклад в развитие методов обработки приближенной информации.

Ценность результатов диссертационной работы Сапкиной Н.В. для науки заключается в следующем:

- получены результаты, касающиеся исследования свойств арифметических операций над нечеткими числами и существования алгебр определенного типа, которые позволят корректно выполнять преобразования при построении нечетких моделей;

- методы восстановления закономерностей для нечетких данных, предложенные в диссертации, по сути, представляют собой

обобщение классического линейного регрессионного моделирования на случай, когда исходная информация является приближенной, а коэффициенты моделей могут быть как числами, так и нечеткими числами.

Ценность результатов диссертационной работы для практики обусловлена созданием программного обеспечения информационной системы интеллектуального анализа данных, предназначенного для автоматизированного поиска закономерностей, которые позволяют выявить скрытые правила в базе данных. Информационная система может использоваться в качестве подсистемы в экспертных системах и системах поддержки принятия решений на этапе предварительного анализа больших объемов информации; выявления взаимосвязи между различными показателями, информация о которых содержится в базе данных (сокращение размерности, установление приоритетов); формирования знаний в виде правил о предметной области.

Научные результаты диссертационной работы Н.В. Сапкиной используются в учебном процесс Воронежского государственного университета и, кроме того, внедрены в практическую деятельность ЗАО ЛЦ «АВС Фарбен» (г. Воронеж), что подтверждено соответствующими актами.

В дальнейшем, результаты, полученные в диссертации, целесообразно использовать в вузах и организациях, занимающихся нечетким моделированием и вопросами разработки новых методов и алгоритмов восстановления закономерностей в данных. В частности, результаты Н.В. Сапкиной могут быть использованы в научных и прикладных исследованиях, проводимых в Воронежском государственном университете, Тверском государственном университете, Ульяновском государственном техническом университете, Санкт-Петербургском государственном техническом университете. Результаты исследований также целесообразно использовать в учебном процессе перечисленных и других вузов, осуществляющих подготовку бакалавров и магистров соответствующих направлений.

4. Замечания по диссертации

1) В диссертации отсутствует анализ возможностей использования для задач интеллектуального анализа данных модели ортогональной регрессии, позволяющей более адекватно отразить

согласованность поведения совокупности наблюдаемых переменных, особенно когда нет оснований выделения одной из них в качестве объясняемой.

2) В работе не уделяется должного внимания проблеме сравнения нечетких чисел, соискатель ограничивается простейшими вариантами.

3) Для оценки адекватности и точности разработанных нечетких регрессионных моделей предложено три варианта формирования остаточной последовательности, но отсутствует сравнительный анализ этих вариантов, нет рекомендаций по использованию.

4) Недостаточно глубоко рассмотрены вопросы эффективности практических применений предложенного аппарата интеллектуального анализа данных.

Перечисленные замечания не влияют на общую положительную оценку работы. Оформление диссертации соответствует требованиям ВАК, работа содержит 151 страницу текста, отличается логичность построения, изложена ясным и понятным языком. Автореферат полностью соответствует содержанию диссертации. Сапкиной Н.В. опубликовано 14 научных работ, из них 5 – из перечня изданий, рекомендованных ВАК РФ, 7 опубликованы в материалах конференций, имеется 2 свидетельства о государственной регистрации программ.

5. Заключение

Диссертационная работа Сапкиной Натальи Владимировны «Восстановление закономерностей на основе нечетких регрессионных моделей» является законченной научно-квалификационной работой, посвященной усовершенствованию методов обработки информации, содержит новые научные результаты в этой области, удовлетворяет требованиям «Положения о присуждении ученых степеней», предъявляемым к кандидатским диссертациям. Автор диссертации, Сапкина Наталья Владимировна, заслуживает присуждения ученой степени кандидата технических наук по специальности 05.13.17 – «Теоретические основы информатики».

Отзыв составлен д.т.н., профессором кафедры математического и программного обеспечения информационных систем (МиПОИС) Белгородского государственного национального исследовательского университета (НИУ «БелГУ») Корсуновым Н.И.

Отзыв ведущей организации обсужден и утвержден на заседании

кафедры математического и программного обеспечения информационных систем НИУ «БелГУ», протокол № 11 от 4 июня 2014 года.

Зав. каф. математического и программного обеспечения информационных систем ФГАОУ ВПО «Белгородский государственный национальный исследовательский университет», НИУ «БелГУ», кандидат технических наук, доцент

 Муромцев В.В.

Профессор каф. математического и программного обеспечения информационных систем ФГАОУ ВПО «Белгородский государственный национальный исследовательский университет», НИУ «БелГУ», доктор технических наук, профессор

 Корсунов Н.И.

Адрес: 308015, г. Белгород, ул. Победы, 85, корп.14, к.1-1
E-mail: muromtsev@bsu.edu.ru,
тел.(4722)301353,

