

На правах рукописи



Сорокина Ирина Владимировна

Методы оценки параметров возможностных
распределений и их применение для
прогнозирования неисправностей
электрооборудования

Специальность 05.13.01 — «Системный анализ, управление
и обработка информации (в промышленности)»

Автореферат
диссертации на соискание учёной степени
кандидата технических наук

Тверь — 2018

Работа выполнена на кафедре информационных технологий ФГБОУ ВО «Тверской государственной университет» (ФГБОУ ВО «ТвГУ»).

Научный руководитель: **Язенин Александр Васильевич**,
доктор физико-математических наук, профессор

Официальные оппоненты: **Рыжов Александр Павлович**,
доктор технических наук, профессор,
Федеральное государственное бюджетное образовательное учреждение высшего образования «Московский государственный университет имени М.В.Ломоносова», г. Москва, доцент кафедры математической теории интеллектуальных систем

Тулупьев Александр Львович,
доктор физико-математических наук, доцент
Федеральное государственное бюджетное учреждение науки Санкт-Петербургский институт информатики и автоматизации Российской академии наук, г. Санкт-Петербург, главный научный сотрудник с возложенными обязанностями заведующего лабораторией теоретических и междисциплинарных проблем информатики

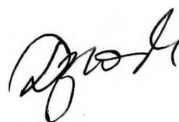
Ведущая организация: Государственное бюджетное образовательное учреждение высшего образования Московской области «Университет «Дубна»

Защита состоится « 22 » февраля 2019г. в 14 часов 00 минут на заседании диссертационного совета Д 212.262.06 в ФГБОУ ВО «ТвГТУ» по адресу: 170026, г. Тверь, наб. Аф. Никитина, 22.

С диссертацией можно ознакомиться в библиотеке Тверского государственного технического университета и на сайте www.tstu.tver.ru

Автореферат разослан « ___ » _____ 2019 г.

Ученый секретарь
диссертационного совета
доктор физико-математических наук,
профессор



С.М.Дзюба

Общая характеристика работы

Актуальность темы исследования

В задачи, поставленные ОАО «РЖД» в «Стратегии развития железнодорожного машиностроения до 2030 года», входит создание комплексных систем диагностики и прогнозирования технического состояния инфраструктуры подвижного состава.

Внедрение автоматизированных систем выявления и прогнозирования неисправностей позволяет одновременно снизить расходы на обслуживание и ремонт электрооборудования и увеличить его безопасность. Раннее выявление повреждённых частей позволяет сократить их воздействие на другие механизмы.

Проанализировав литературу, посвящённую решению данной проблемы, можно сделать вывод о необходимости разработки такой системы с применением современных методов обработки информации.

Для создания систем подобного рода могут быть использованы машины нечёткого вывода, обучение которых может производиться на основе аппарата теории неопределённостей. Данный инструмент отлично зарекомендовал себя для решения задач интеллектуального анализа данных. Теоретической основой для решения этой задачи должны быть математически обоснованные методы оценки параметров нечётких распределений.

Однако, к настоящему времени такие методы были разработаны только для одномерных распределений, что существенно ограничивает возможность использования этих результатов при решении практических задач, в том числе при обучении машин нечёткого вывода.

Таким образом, можно ожидать, что разработка методов оценки параметров многомерных возможных распределений и их внедрение в алгоритмы обучения позволит повысить качество работы машин нечёткого вывода и расширить круг практических задач, решаемых с их помощью, в том числе задачи прогнозирования неисправностей электрооборудования.

Этим определяется актуальность темы диссертации.

Степень разработанности темы

Развитие теории нечетких множеств и ее применение для моделирования неопределенности привело к появлению и применению систем нечёткого вывода, которые анализируют входные значения в терминах лингвистических переменных, осуществляют их отображение во множество выходных

значений на основе базы нечётких правил. Эта база может быть построена с применением методов оценки параметров вероятностных распределений.

Данная тема получила развитие в статье Näther W. “On possibilistic inference”, в которой рассматриваются некоторые идеи получения оценок параметров. Одновременно Zhen-Yuan W. и Shou-Mei L. в “Fuzzy linear regression analysis of fuzzy-valued variables” предложили μ/E метод оценки параметров вероятностных распределений, который далее получил развитие в работах Xizhao W. и Minghu H.

Badard R. рассматривал “минимальную оценку нечеткости” для нечётких моделей и изучал сходимость этих оценок. Dishkant H. в статье “About membership functions estimation” изучал оценку параметров нечетких переменных в значении Заде Л. Cai К.-У. занимался изучением оценки параметров для нормальных нечётких переменных, основываясь на понятиях Nahmias S. На основе его работ Hong D.H. в работе “Parameter estimations of mutually T-related fuzzy variables” оценивал параметры обобщённых T-связанных нечётких величин.

Таким образом, можно сделать вывод о том, что тема активно развивается, получены оценки в одномерном случае, но ряд вопросов остаётся открытым. Одним из таких вопросов является исследование методов оценки параметров в многомерном случае.

Что касается задачи прогнозирования неисправностей электрооборудования применительно к железнодорожным вагонам, можно отметить, что большинство исследований и существующих систем направлено на анализ состояния рельсов, колёсных пар и подвески железнодорожных вагонов. В таких системах ведётся обработка сравнительно небольшого числа аналоговых сигналов высокой частоты, собранных с помощью дополнительных датчиков. Анализ данных осуществляется с использованием Фурье или вейвлет-преобразований для начальной обработки данных.

Характер данных, доступных в современных системах управления электрооборудованием пассажирских вагонов, существенно отличается. Число сигналов достигает нескольких сотен, большинство из них является бинарными и редко изменяется. В таких условиях применение методов спектрального анализа не является оправданным и прямой перенос имеющегося опыта на другие системы вагонов невозможен. Необходима разработка новых методов поиска и выявления признаков нештатных режимов работы оборудования.

Этим и определяются цель и задачи диссертационного исследования.

Целью данной работы является повышение надёжности работы электрооборудования на основе применения методов оценки параметров возможных распределений.

Данная цель предполагает разработку методов определения параметров многомерных распределений, ориентированных на использование для обучения машин нечёткого вывода, которые могут быть применены на практике при решении задач прогнозирования неисправностей электрооборудования.

Для достижения поставленной цели необходимо было решить следующие **задачи**:

- рассмотреть существующие методы построения оценки параметров возможных распределений в одномерном случае;
- разработать методы оценивания параметров возможных распределений в многомерном случае для сильнейшей и Архимедовых t -норм;
- определить свойства полученных оценок;
- разработать алгоритм обучения машины нечёткого вывода на основе метода оценки параметров многомерных возможных распределений;
- реализовать библиотеку программ для восстановления параметров многомерных возможных распределений;
- провести сравнение разработанного алгоритма с существующими аналогами с помощью численных экспериментов;
- разработать алгоритмы идентификации сложных систем на основе объединения нечёткого вывода с эффективными методами машинного обучения;
- на основе разработанных алгоритмов реализовать прототипы систем прогнозирования, ориентированные на применение в промышленности.

Научная новизна

Научная новизна заключается в разработке комплекса методов оценки параметров многомерных возможных распределений:

1. Оценка параметров многомерных распределений минисвязанных возможных величин:
 - 1.1. Метод нахождения оценки параметров многомерных возможных распределений при заданном уровне возможности;

1.2. Максиминная оценка параметров многомерных возможностных распределений;

2. Оценка параметров многомерных возможностных распределений в случае Архимедовых t-норм.

Также научную новизну составляет разработка алгоритма обучения машин нечёткого вывода, алгоритма идентификации сложных систем, основанного на нейронных сетях и нечётком выводе, и алгоритма идентификации сложных систем, использующего бустинг нечётких контроллеров.

Практическая значимость

Предложенный алгоритм обучения машин нечёткого вывода обеспечивает более высокую точность по сравнению с существующими аналогами, что создаёт предпосылки для повышения эффективности функционирования различных систем управления и анализа данных, в том числе для повышения надёжности работы электрооборудования на предприятиях вагоностроения и электронной промышленности.

На основе предложенных алгоритмов идентификации в диссертации разработаны:

- система анализа данных для прогнозирования неисправностей вагонного электрооборудования на железнодорожном транспорте;
- система прогнозирования параметров прозрачных проводниковых материалов.

Соответствие паспорту специальности

Проблематика, обозначенная в диссертации, соответствует специальности 05.13.01 «Системный анализ, управление и обработка информации».

Пункту 1 формулы специальности: "теоретические основы и методы системного анализа, оптимизации, управления, принятия решений и обработки информации" соответствует разработка методов оценки параметров многомерных возможностных распределений.

Пункту 4 формулы специальности "разработка методов и алгоритмов решения задач системного анализа, оптимизации, управления, принятия решений и обработки информации" и пункту 6 "методы идентификации систем управления на основе ретроспективной, текущей и экспертной информации" соответствует разработанный алгоритм обучения машин нечёткого вывода.

Пункту 5 формулы специальности "разработка специального математического и алгоритмического обеспечения систем анализа, оптимизации, управления, принятия решений и обработки информации" соответствует библиотека программ для восстановления параметров многомерных нечётких распределений

Пункту 7 формулы специальности "методы и алгоритмы структурно-параметрического синтеза и идентификации сложных систем" соответствует:

1. Алгоритм идентификации сложных систем, основанный на нейронных сетях и нечётком выводе.
2. Алгоритм идентификации сложных систем, использующий бустинг нечётких контроллеров.

Пункту 9 формулы специальности "разработка проблемно-ориентированных систем управления, принятия решений и оптимизации технических объектов" соответствует:

1. Система прогнозирования неисправностей вагонного электрооборудования на железнодорожном транспорте.
2. Система прогнозирования свойств прозрачных проводниковых материалов, использующая алгоритм Adaboost совместно с разработанным алгоритмом обучения машин нечёткого вывода.

Пункту 11 формулы специальности "методы и алгоритмы прогнозирования и оценки эффективности, качества и надежности сложных систем" соответствует двухуровневая система анализа данных, основанная на нейронных сетях и нечётком выводе для прогнозирования неисправностей вагонного электрооборудования на железнодорожном транспорте.

Методология и методы исследования

Методология опирается на формулирование математических утверждений и их доказательство, описание и изучение свойств алгоритмов, апробацию теоретических результатов через реализацию в коде программ и проведение вычислительных экспериментов.

В работе используются методы теории возможностей, системного анализа, общей теории систем, моделирования систем, теории управления, нейронные сети, методы оптимизации, а также технологии сбора информации, проведение численных экспериментов, разработки и испытания опытного образца.

Для проектирования библиотеки программ для восстановления параметров многомерных нечётких распределений и систем прогнозирования использовались принципы объектно-ориентированного программирования. Для их разработки были использованы технологии, связанные с языками реализации (Matlab и Python) и средами разработки (Matlab и Anaconda).

Полученные результаты диссертационной работы проверялись с помощью проведения вычислительных экспериментов и применения результатов исследований в условиях реальной проектной деятельности.

Основные положения, выносимые на защиту:

1. Методы оценки параметров многомерных возможностных распределений.
2. Алгоритм обучения машин нечёткого вывода.
3. Библиотека программ для восстановления параметров многомерных нечётких распределений.
4. Алгоритм идентификации сложных систем, основанный на нейронных сетях и нечётком выводе.
5. Алгоритм идентификации сложных систем, использующий бустинг нечётких контроллеров.
6. Система прогнозирования свойств прозрачных проводниковых материалов, использующая алгоритм Adaboost совместно с разработанным алгоритмом обучения машин нечёткого вывода.
7. Двухуровневая система анализа данных, основанная на нейронных сетях и нечётком выводе для прогнозирования неисправностей вагонного электрооборудования на железнодорожном транспорте.

Достоверность результатов

Достоверность полученных в диссертационной работе результатов подтверждается строгостью проводимых математических обоснований при формулировании и доказательстве теорем, результатами численных расчётов, сравнительным анализом полученных в ходе модельных экспериментов результатов с известными.

По результатам диссертационной работы разработан и внедрен программный комплекс прогнозирования неисправностей вагонного электрооборудования на железнодорожном транспорте в рамках совместных проектов РФФИ и РЖД № 12-07-13117-офи_м_РЖД и №13-07-13160-офи_м_РЖД.

Методы оценки параметров многомерных возможностных распределений использованы в ЗАО «РТИС ВКО» для разработки программного обеспечения по заказу Министерства обороны РФ.

Результаты диссертационной работы внедрены в учебный процесс на факультете прикладной математики и кибернетики Тверского государственного университета в курсе "Анализ нечётких информационных систем".

Апробация результатов

Основные результаты работы докладывались на 6 конференциях:

1. Студенческая научно-практическая конференция факультета прикладной математики и кибернетики;
2. Интегрированные модели и мягкие вычисления в искусственном интеллекте. VII-я Международная научно-техническая конференция;
3. 20th East West Fuzzy Colloquium 2013. 20th Zittau Fuzzy Colloquium;
4. Комплексное использование и охрана подземных пространств: Международ. науч.-практ. конф., посвящ. 100-летию юбилею науч. и туристско-экскурсионной деятельности в Кунгурской Ледяной пещере и 100-летию со дня рожд. В.С. Лукина;
5. Гибридные и синергетические интеллектуальные системы: II международный Поспеловский симпозиум;
6. ICFLS 2015: XIII International Conference on Fuzzy Logic Systems, Barcelona, Spain.

Личный вклад автора

Все основные научные результаты, разработка и написание программного обеспечения принадлежат лично автору.

Публикации

Основные результаты диссертации опубликованы в 11 печатных изданиях, 5 из которых в журналах, из перечня ВАК РФ, 6 — в тезисах докладов конференций. Получено 2 свидетельства о регистрации программных продуктов для ЭВМ.

Объем и структура работы

Диссертация состоит из введения, пяти глав, заключения, списка литературы и 6 приложений. Общий объем диссертации, не включая приложения, 115 страниц текста с 30 рисунками и 7 таблицами. Список литературы содержит 104 наименования.

Содержание работы

Во **введении** обоснована актуальность работы, приведён обзор научной литературы, определяющий степень разработанности темы исследования, сформулированы цель, задачи диссертационного исследования, сформулирована научная новизна и практическая значимость представляемой работы.

В **первой главе** рассмотрены понятия и термины теории возможностей, которые будут использоваться далее по ходу диссертации.

В параграфах 1.1 и 1.2 введено понятие возможностного пространства, определены его свойства, введены понятия мер неопределённости, возможности и необходимости. Определена возможностная величина (переменная), её носитель, множество α -уровня. Рассмотрены операции над минисвязанными возможностными величинами.

Параграф 1.3 посвящён агрегированию нечёткой информации на основе t -норм. Вначале дано понятие треугольной нормы (t -нормы), описаны её основные свойства, особое внимание уделено сильнейшей и слабейшей t -нормам.

Рассмотрен класс Архимедовых t -норм и его взаимосвязь с остальными треугольными нормами.

Далее введено понятие аддитивного генератора, обозначена важность данного математического инструмента. Показана его связь с Архимедовыми t -нормами.

В параграфе 1.4 рассмотрены существующие подходы к интерпретации понятия функции распределения (принадлежности) и способам измерения её значений. Вначале с помощью каждого из подходов дана интерпретация предложенного нечёткого утверждения, а далее подробно рассмотрен каждый из них.

После рассмотрения существующих способов измерения значений функции распределения в качестве основополагающего для данной работы был выбран подход, основанный на интерпретации принадлежности как меры сходства объектов, который предполагает, что существует идеальный элемент множества, который принадлежит этому множеству в полной мере, а принадлежность других элементов определяется расстоянием от них до идеального элемента.

Во **второй главе** были рассмотрены существующие подходы к оценке параметров возможностных распределений в одномерном случае.

В параграфе 2.1 произведена постановка задачи, определены свойства, которым должна отвечать полученная оценка.

В параграфе 2.2 сделан обзор двух существующих методов получения оценки в одномерном случае: в параграфе 2.2.1 метод Cai Kai-Yuan и его расширение, сделанное Dug Hun Hong, и в параграфе 2.2.2. подробно описан метод Wang Xizhao и Ha Minghu для нахождения максиминной μ/E оценки параметров.

В отличие от Hong, Wang Xizhao и Ha Minghu доказали, что их оценка является состоятельной, достаточной и максимально правдоподобной.

В параграфе 2.3 произведён сравнительный анализ рассмотренных методов, который позволяет утверждать, что метод μ/E оценки является частным случаем метода Dug Hun Hong.

Третья глава посвящена методам нахождения оценки параметров многомерных возможностных распределений, основанным на геометрическом подходе, заключающемся в построении эллипсоида минимального объёма, охватывающего все точки распределения.

В параграфе 3.1 рассматривается случай минисвязанных возможностных величин.

Параграф 3.1.1 посвящён оценке параметров в случае заданного уровня возможности и описывает метод оценивания.

Была доказана теорема, позволяющая найти оценку параметров многомерных возможностных распределений при заданном уровне возможности:

Теорема 1. Пусть:

1. $f : \mathbb{R}^1 \rightarrow \mathbb{R}^1$ — строго убывающая на луче $[0, +\infty)$ функция, $f(0) = 1$.

2. $Q = \{\mu_{\xi}(\mathbf{x}, \mathbf{c}, A)\}$ — семейство распределений, где $\mathbf{c} \in \mathbb{R}^m$, A — положительно определённая симметричная матрица $m \times m$, $\mu_{\xi}(\mathbf{x}, \mathbf{c}, A) = \max \{0, f(\sqrt{(\mathbf{x} - \mathbf{c})^T A (\mathbf{x} - \mathbf{c})})\}$.

3. $X = \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ — наблюдаемые значения выборки (ξ_1, \dots, ξ_n) минисвязанных возможностных величин, принадлежащих семейству распределений Q , $\mathbf{x}_i \in \mathbb{R}^m, i = 1, \dots, n$.

4. W — многомерный эллипсоид минимального объёма, описанный вокруг множества X , заданный параметрами $(\check{\mathbf{c}}, \check{A})$, где $\check{\mathbf{c}}$ — центр эллипсоида W , \check{A} — матрица, задающая ориентацию осей эллипсоида и длину его радиусов.

Тогда оценкой параметра $\theta_{\alpha} = (\mathbf{c}, A)$ является $\hat{\theta}_{\alpha} = (\hat{\mathbf{c}}, \hat{A}) = (\check{\mathbf{c}}, k\check{A})$, где $k = (f^{-1}(\alpha))^2$.

Также в параграфе рассмотрен пример нахождения оценки параметров распределения в двумерном случае для функции $f(t) = e^{-t^2}$.

На основе проведённого исследования и рассмотренного примера сделан вывод, что параметр α влияет на форму распределения: чем больше значение параметра, тем больше увеличивается значение коэффициента нечёткости и распределение становится более размытым.

Параграф 3.1.2 описывает метод нахождения максиминной μ/E оценки параметров многомерных возможностных распределений и является расширением подхода Wang Xizhao и Ha Minghu на многомерный случай.

В ходе исследования была доказана теорема, описывающая способ определения оценки параметров \mathbf{c}, A .

Теорема 2. Пусть:

1. $f : \mathbb{R}^1 \rightarrow \mathbb{R}^1$ — строго убывающая на луче $[0, +\infty)$ функция, $f(0) = 1$.
2. $Q = \{\mu_{\xi}(\mathbf{x}, \mathbf{c}, A)\}$ — семейство распределений, где $\mathbf{c} \in \mathbb{R}^m$, A — положительно определённая симметричная матрица, $\mu_{\xi}(\mathbf{x}, \mathbf{c}, A) = \max\{0, f(\sqrt{(\mathbf{x} - \mathbf{c})^T A (\mathbf{x} - \mathbf{c})})\}$.
3. $X = \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ — наблюдаемые значения выборки (ξ_1, \dots, ξ_n) минимально связанных возможностных величин, принадлежащих семейству распределений Q , $\mathbf{x}_i \in \mathbb{R}^m, i \in [1, \dots, n]$.
4. W — многомерный эллипсоид минимального объёма, описанный вокруг множества X .

Тогда максиминной оценкой параметра $\theta = (\mathbf{c}, A)$ является $(\hat{\mathbf{c}}, \hat{A})$, такие, что эллипсоид W задаётся уравнением $(\mathbf{x} - \hat{\mathbf{c}})^T \hat{A} (\mathbf{x} - \hat{\mathbf{c}}) = 1/q^2$, $q = \arg \max_{t \geq 0} t^m \bar{f}(t)$.

Рассмотрены примеры нахождения оценок параметров двумерных распределений.

В параграфе 3.1.3 доказываются свойства полученной оценки параметров.

В параграфе 3.2 производится построение оценки параметров в случае Архимедовой t -нормы. Найдена оценка для параметра c , отвечающего за расположение центра распределения, и представлен алгоритм численного нахождения параметра A , отвечающего за поворот и масштабирование.

В **четвертой главе** в параграфе 4.1.1 дано понятие нечёткого правила, в параграфе 4.1.2 произведено описание машины нечёткого вывода. Перечислены составные элементы таких систем.

Параграф 4.1.3 посвящён разработке алгоритма обучения машин нечёткого вывода, основанного на идее алгоритма Cluster Estimation, который был предложен Stephen L. Chiu.

Для проверки работоспособности разработанного алгоритма в параграфе 4.1.4 рассмотрена задача аппроксимации нелинейной функции одной переменной $z = \sin(y)/y$. После чего произведено сравнение результатов с результатами алгоритма Cluster Estimation. Показано, что действуя согласно разработанному алгоритму, наиболее качественные результаты получаются при использовании модели нечёткого вывода Сугено 1 и экспоненциальной функции распределения, и в целом авторский алгоритм даёт меньшую ошибку в сравнении с алгоритмом Cluster Estimation.

В параграфе 4.1.5 рассматривается задача прогнозирования значений нестационарного временного ряда, порождаемого дифференциальным уравнением Mackey-Glass:

$$\dot{x}(t) = \frac{0.2x(t - \tau)}{1 + x^{10}(t - \tau)} - 0.1x(t).$$

Результаты работы над этой задачей с использованием разработанного алгоритма показали, что он является достаточно эффективным, но, тем не менее, иногда уступает алгоритму Cluster Estimation.

Для улучшения результатов в параграфе 4.1.6 была произведена модификация алгоритма путём замены метода кластеризации на Mountain method.

По результатам сравнения работы нового алгоритма с предыдущим разработанным алгоритмом на первой задаче – аппроксимации нелинейной функции $z = \sin y/y$ выяснилось, что модифицированный алгоритм работает значительно лучше.

После применения нового алгоритма к задаче прогнозирования значений нестационарного временного ряда был получен результат, уступающий лишь методу ANFIS.

В параграфе 4.2 описывается разработанная двухуровневая система анализа данных, основанная на нейронных сетях и нечётком выводе. Применение ограниченных машин Больцмана для построения нейронной сети позволяет выявить существенные закономерности в данных и перейти от пространства входов размерностью в несколько сотен или тысяч измерений к пространству скрытых переменных небольшой размерности. Далее для выделения областей, соответствующих различным классам (при решении задач классификации), в пространстве выходов нейронной сети используется маши-

на нечеткого вывода. Для её обучения применяется разработанный алгоритм обучения машин нечёткого вывода.

В параграфе 4.3 рассматривается процедура бустинга и алгоритм Adaboost. Рассматривается возможность использования алгоритма бустинга Adaboost для построения композиций машин нечёткого вывода, обучающихся разработанным алгоритмом.

В **пятой главе** описаны библиотеки программ и прототипы систем прогнозирования, разработанные в ходе работы над диссертацией.

В параграфе 5.1 описывается реализация алгоритма обучения машины нечёткого вывода в среде Matlab. Разработанная программа поддерживает модели нечёткого вывода Сугено 0-го и 1-го порядка.

В параграфе 5.2 производится реализация алгоритма обучения машины нечёткого вывода на языке программирования Python. Алгоритм реализован в соответствии с соглашениями библиотеки scikit-learn, что позволяет интегрировать его с современными системами анализа данных.

Параграф 5.2.1 посвящён решению проблемы прогнозирования неисправностей вагонного электрооборудования на железнодорожном транспорте с применением разработанных технологий. Производится поставка задачи и первичный анализ данных. Далее происходит апробация двухуровневой системы анализа данных, приведённой в параграфе 4.2 и оценка её работы по сравнению с алгоритмам ANFIS.

Анализ работы алгоритмов показал, что использование разработанно метода оценки параметров функций распределения нечётких термов позволяет более точно описать структуру данных и получить более адекватные результаты, чем с помощью ANFIS (таблица 1).

После проведения экспериментов были получены следующие результаты:

Таблица 1

Результаты тестирования алгоритмов обучения машин нечёткого вывода

Алгоритм	Среднее число ошибочных классификаций	
	Без нормализации данных	С нормализацией данных
ANFIS	5.96 (37%)	5.40 (34%)
μ/E алгоритм	4.03 (25%)	4.16 (26%)

В целом можно сказать, что разработанная двухуровневая система анализа данных справляется с задачей прогнозирования лучше, чем ANFIS.

Пример работы реализованной системы представлены на рисунке 1, где * – прогноз, о – значение из контрольной выборки.

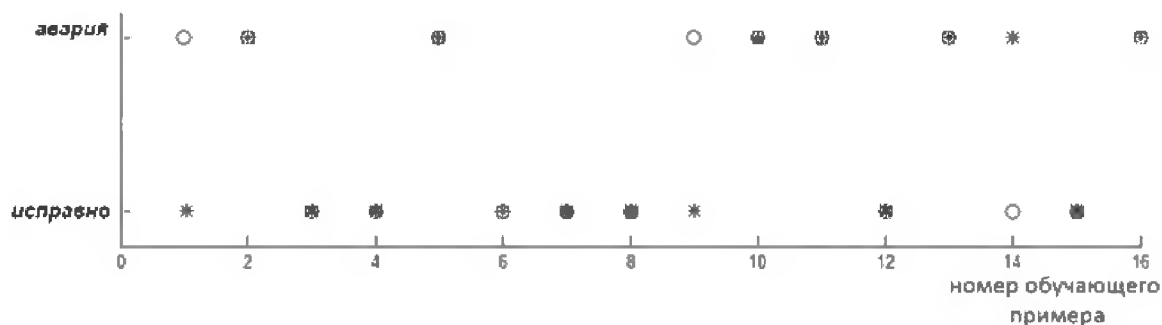


Рис. 1. Пример работы обученной машины нечёткого вывода

Полученная архитектура может являться основой системы прогнозирования возникновения неисправностей в электрооборудовании вагона.

В параграфе 5.2.2 затрагивается вопрос применение разработанного математического аппарата к проблеме идентификации состояния воздушного пространства.

В параграфе 5.2.3 описывается применение алгоритма Adaboost совместно с разработанным алгоритмом обучения машин нечёткого вывода к типовой задаче классификации (параграф 5.2.3.1) и к задаче идентификации параметров прозрачных проводниковых материалов (параграф 5.2.3.2).

Большое значение для радиоэлектронной промышленности имеет поиск материалов, которые могут быть использованы для изготовления прозрачных проводников. Чтобы отобрать наиболее многообещающие материалы для проведения реальных экспериментов, можно применить прогнозирование для оценки интересующих свойств материала на основе информации о его составе.

В ходе работы над диссертацией была создана система для решения задачи прогнозирования значения такого ключевого свойства проводников как энергии образования – важный показатель устойчивости материала, обеспечивающий надёжность произведённых проводников. В качестве входа данная система получает набор данных, где каждый проводник описывается 11 свойствами, и на основе их значений прогнозируется значение энергии образования.

Была проведена серия экспериментов по определению качества работы созданной системы: при количестве алгоритмов в композиции равном 10 и радиусе равном 0.9, коэффициент детерминации составил 0.7115, при 15

алгоритмах в композиции – 0.7336 при радиусе 0.7(рис.2). Данные значения достигаются на тестовой выборке.

Для сравнения данная задача решалась также с помощью одного из самых популярных в настоящее время методов анализа данных – деревьев решений с применением бустинга. Бустинг деревьев решений показал следующие результаты: при количестве алгоритмов в композиции равном 10, коэффициент детерминации составил 0.6310, при количестве алгоритмов в композиции равном 15 – 0.6296 (рис. 2).

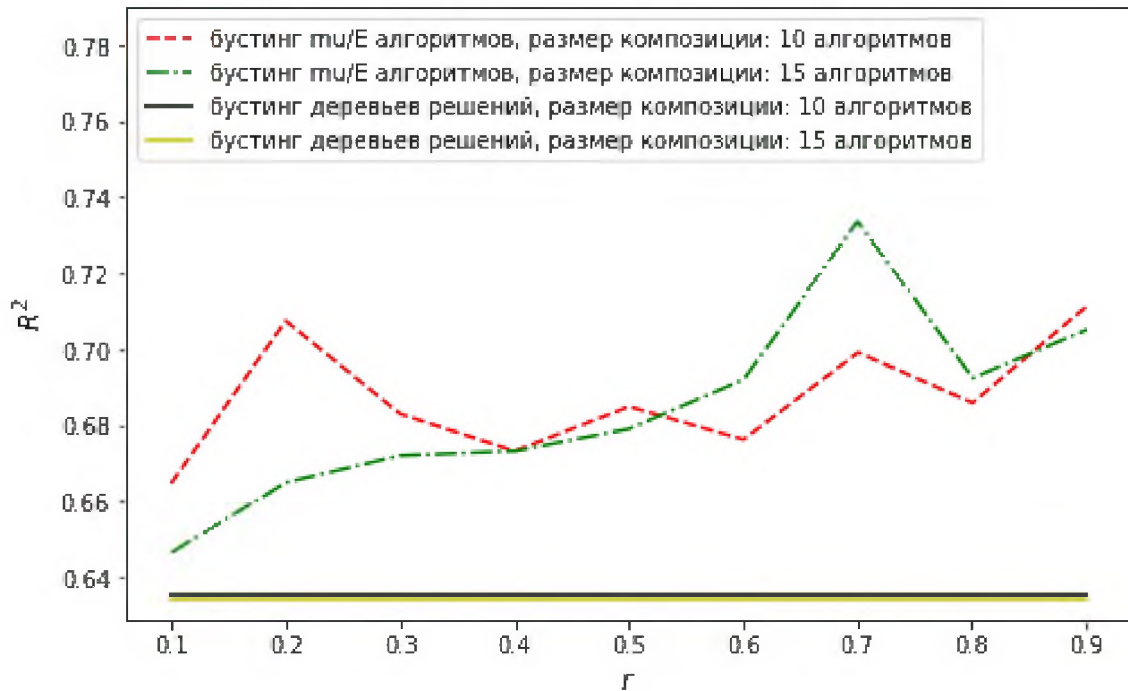


Рис. 2. График ошибок детерминации при бустинге μ/E алгоритмов и деревьев решений

Таким образом, можно сделать вывод о том, что разработанный алгоритм по эффективности превосходит бустинг деревьев решений для задач такого рода.

В заключении диссертационной работы подводятся итоги диссертационного исследования, излагаются основные научные результаты, полученные в ходе работы над диссертацией.

Проведённые исследования показали, что в случае использования модели, состоящей из одной машины нечёткого вывода, в зависимости от характера исследуемой системы, возможны два варианта:

- в части случаев наиболее эффективными оказываются системы, построенные на одномерных термах с помощью алгоритма ANFIS;
- в других случаях более эффективным является использование многомерных возможностных распределений, и в этом случае предложенный

в работе алгоритм позволяет достичь более высоких результатов по сравнению с существующими аналогами.

Заключение

В результате диссертационного исследования были выполнены все поставленные задачи и, таким образом, достигнута цель работы.

Основные результаты диссертации:

1. Предложены методы оценивания параметров возможных распределений в многомерном случае для сильнейшей и Архимедовых t -норм.
2. Показано, что оценка, полученная для случая сильнейшей t -нормы, является состоятельной, достаточной и максимально правдоподобной.
3. На основе полученных методов разработан алгоритм обучения машины нечёткого вывода.
4. Реализована библиотека программ для восстановления параметров многомерных возможных распределений.
5. Разработан алгоритм идентификации сложных систем, основанный на нейронных сетях и нечётком выводе.
6. Разработан алгоритм идентификации сложных систем, использующий бустинг нечётких контроллеров.
7. Реализована система прогнозирования свойств прозрачных проводниковых материалов.
8. Реализована система анализа данных для прогнозирования неисправностей вагонного электрооборудования на железнодорожном транспорте.

Подводя итоги исследования, можно утверждать, что при решении задачи прогнозирования неисправностей электрооборудования применение разработанных в ходе диссертации гибридных систем идентификации позволяет достичь более высокой точности, чем при применении систем, использующих одну машину нечёткого вывода. Дальнейшее совершенствование алгоритмов обучения машин нечёткого вывода поможет увеличить точность прогнозирования неисправностей, тем самым повысив качество работы созданных систем.

Публикации автора по теме диссертации

Статьи в изданиях, включенных в список ВАК РФ

1. Иншина, И. В. Оценка параметров двумерных возможных распределений /И. В. Иншина, С. В. Сорокин // Нечеткие системы и мягкие вычисления. — 2012. — Т. 7, № 2. — С. 99–108.
2. Сорокина, И. В. К задаче оценки параметров многомерных возможных распределений /И. В. Сорокина, С. В. Сорокин // Нечеткие системы и мягкие вычисления. — 2013. — Т. 8, № 2. — С. 101–113.
3. Сорокин, С. В. Прогнозирование неисправностей оборудования с использованием нейронных сетей и нечёткого вывода / С. В. Сорокин, И. В. Сорокина, А. В. Язенин // Техника железных дорог. — 2014. — №2(26). — С. 64–74.
4. Сорокина, И. В. Оценка параметров многомерных возможных распределений при заданном уровне риска / И. В. Сорокина, С. В. Сорокин // Нечёткие системы и мягкие вычисления. — 2015. — Т. 10, № 2. — С. 181–193.
5. Сорокина, И. В. Оценка параметров многомерных возможных распределений в случае Архимедовой t-нормы / И. В. Сорокина // Нечёткие системы и мягкие вычисления. — 2018. — Т. 13, № 1. — С. 45–58.

Публикации в прочих изданиях

6. Иншина, И. В. Метод нахождения оценки параметров многомерных возможных распределений / И. В. Иншина // Студенческие научно-практические конференции факультета прикладной математики и кибернетики: сб. трудов конференций от 26 апреля 2012 года и 25 апреля 2013 года. — Тверь, 2013. — С. 66–76.
7. Иншина, И. В. Оценка параметров многомерных возможных распределений /И. В. Иншина, С. В. Сорокин // Интегрированные модели и мягкие вычисления в искусственном интеллекте: сб науч. трудов VII-й Международной научно-технической конференции. — Коломна, 2013 г. — Т.1. — С. 420–429.
8. Inshina, I. V. Parameter Estimation of n-dimensional Fuzzy Distributions in Fuzzy Model Identification / I. V. Inshina, S. V. Sorokin // Proceeding of East West Fuzzy Colloquium 2013. 20th Zittau Fuzzy Colloquium. — Zittau, 2013. — С. 55–62.

9. Сорокин, С. В. Модель динамики уровней подземных вод в пещере Кулогорская-Троя / С. В. Сорокин, Н. А. Франц, И. В. Сорокина // Комплексное использование и охрана подземных пространств: междунар.науч.-практ. конф., посвящ. 100-летию юбилею науч. и туристско-экскурсионной деятельности в Кунгурской Ледяной пещере и 100-летию со дня рожд. В.С. Лукина / ГИ УрО РАН; под общ. ред. О. Кадебской, В. Андрейчука. – Пермь, 2014. – С. 92–99.
10. Сорокина, И. В. Применение метода оценки параметров многомерных возможностных распределений для создания машин нечёткого вывода / И. В. Сорокина, С. В. Сорокин // Гибридные и синергетические интеллектуальные системы: материалы II международного Поспеловского симпозиума / под ред. д.т.н., проф. А.В. Колесникова. – Калининград, 2014. – С. 327–335.
11. Sorokina, I. V. Parameter Estimation of Multidimensional Possibility Distributions / I. V. Sorokina, S. V. Sorokin, A. V. Yazenin // ICFLS 2015: XIII International Conference on Fuzzy Logic Systems. – Barcelona, 2015. – №9(2). – С. 63–66.

Свидетельства о регистрации программ для ЭВМ

12. U/E анализатор данных эксперимента ATLAS: свидетельство о государственной регистрации программ для ЭВМ / И. В. Сорокина, С. В. Сорокин; правообладатель ФГБОУ ВПО "Тверской государственный университет". — №2014662971; дата регистрации 12.12.2014г.
13. Библиотека для восстановления параметров многомерных нечётких распределений: свидетельство о государственной регистрации программ для ЭВМ / И. В. Сорокина, С. В. Сорокин; правообладатели: И. В. Сорокина, С. В. Сорокин. — №2018617469; дата регистрации 25.06.2018г.

Подписано в печать 02.11.2018. Формат 60x84 1/16.

Усл. печ. л. 1,18. Тираж 100. Заказ № 565.

Редакционно-издательское управление

Тверского государственного университета.

Адрес: 170100, г. Тверь, Студенческий пер. 12, корпус Б.

Тел. РИУ (4822) 35-60-63.