

На правах рукописи



Кар

Каратач Сергей Александрович

РАЗРАБОТКА МЕТОДОВ И АЛГОРИТМОВ
ОБРАБОТКИ ВРЕМЕННЫХ РЯДОВ В ЗАДАЧАХ
ИХ ПРОГНОЗИРОВАНИЯ НА ОСНОВЕ
НЕЧЕТКИХ СИСТЕМ С УЧЕТОМ НЕЧЕТКОСТИ
ВХОДОВ

Специальность 2.3.1 —
«Системный анализ, управление и обработка информации,
статистика»

Автореферат
диссертации на соискание учёной степени
кандидата технических наук

Работа выполнена в федеральном государственном бюджетном образовательном учреждении высшего образования «Белгородский государственный технологический университет им. В.Г.Шухова» (БГТУ им.В.Г.Шухова).

Научный руководитель: **кандидат технических наук, профессор
Синюк Василий Григорьевич**

Официальные оппоненты: **Фамилия Имя Отчество,
доктор физико-математических наук, профессор,
Не очень длинное название для места работы,
старший научный сотрудник**
**Фамилия Имя Отчество,
кандидат физико-математических наук,
Основное место работы с длинным длинным
длинным длинным названием,
старший научный сотрудник**

Захита состоится **DD mmmmmmmm YYYY г. в XX часов** на заседании диссертационного совета БелГУ.22.08 при Белгородском государственном национальном исследовательском университете (НИУ «БелГУ») по адресу: 308015, г. Белгород, ул. Победы, 85, корпус 14, каб. 1-1.

С диссертацией можно ознакомиться в библиотеке **ФГАОУ ВО «Белгородский государственный национальный исследовательский университет» (НИУ «БелГУ»)**.

Отзывы на автореферат в двух экземплярах, заверенные печатью учреждения, просьба направлять по адресу: 308015, г. Белгород, ул. Победы, 85, корпус 14, каб. 1-1, ученому секретарю диссертационного совета БелГУ.22.08.

Автореферат разослан **DD mmmmmmmm2025 года.**
Электронная почта совета: BSU.22.08@bsuedu.ru; zhikharev@bsuedu.ru.

Ученый секретарь
диссертационного совета
БелГУ.22.08,
кандидат технических наук



Жихарев Александр Геннадиевич

Общая характеристика работы

Актуальность темы. Процесс управления сводится к принятию решений и контролю за их реализации. Принимаемые решения должны быть обоснованными. Методы обоснования решений входят в системы информационного обеспечения управления. Одним из таких методов является использование результатов анализа возможных исходов при различном воздействии на объект управления. Для анализа выполняется прогнозирование этих исходов. Прогнозирование часто осуществляется по временным рядам, которые генерируются при регистрации значений параметров некоторого контролируемого процесса. Обоснованность принимаемых на основе прогноза решений зависит от адекватности извлечения и учета неопределенности временных данных используемой моделью прогнозирования.

Решением задача прогнозирования временных рядов занималось большое количество исследователей. В результате среди прочих выделился набор широко используемых моделей прогнозирования, таких как линейный предсказатель и модель Бокса-Дженкинса. Такие подходы используют для получения вероятностных оценок характеристик временных рядов свойство стационарности и статистическое усреднение по ансамблю характеристик объектов. Неопределенность при этом часто рассматривается как ошибки измерений, подчиненные конкретному закону распределения, то есть происходит «навязывание законов природе». Таким образом, из-за наложения этих ограничений вероятностные модели неадекватны.

Альтернативным подходом может служить использование нечетких систем с фазификацией типа non-singleton, на основе разработанной Л. Заде теории нечеткой логики. Тогда как, вероятностные модели основаны на усреднении характеристик временных рядов, в нечетких моделях усредняется реакция функции влияние этих характеристик. Для задания функций влияния при этом не требуется иметь ансамбль временных рядов или свойства стационарности отдельного временного ряда.

Распространенные подходы нечеткого вывода, выработанные Э. Мамдани, П. Ларсеном, Т. Такаги, М. Сугено и Ю. Цукамато, как правило, используют четкие значения входов, полученные в результате фазификации singleton, для простоты реализации. Из-за такого упрощения теряется информация о неопределенности входных значений временных рядов. Использование фазификации типа non-singleton позволяет сохранить информацию о нечеткости входных данных, но приводит к экспоненциальной вычислительной сложности нечеткого логического вывода при многих входах модели. Тогда чтобы использовать нечеткий логический вывод при фазификации non-singleton нужно создать нечеткую модель, разработать эффективные алгоритмы и способ быстрой их реализации.

Таким образом, задача разработки информационной технологии прогнозирования на основе нечеткого вывода при фазификации типа non-singleton является актуальной.

Проблемы связанные с нечетким выводом и нечетким моделированием в России изучались и прорабатывались А.Н. Аверкин, В.В Борисов, В.В. Курейчик, И.А. Ходашинский, А.В. Язенин, Н.Г. Ярушкина. За рубежом развитием нечеткой теории занимались Р. Angelov, D. Dubois, H. Ishibuchi, E.H. Mamdani, J.M. Mendel, L. Rutkowski, M. Sugeno, T. Takagi, H. Tanaka, R.R. Yager, T. Yasukawa, L. Wang, L. Zadeh и др..

Объектом исследования являются информационная технология управления на основе анализа временных рядов.

Предметом исследования являются методы и алгоритмы прогнозирования временных рядов на основе нечеткого моделирования.

Целью данной работы является совершенствование информационного обеспечения управления на основе разработки информационной технологии и алгоритмов прогнозирования временных рядов с использованием нечетких систем, адекватно учитывающих уникальность объектов.

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

1. Анализ методов прогнозирования временных рядов в задачах управления с позиции адекватности учета уникальности свойств генерирующих их объектов.
2. Разработка метода прогнозирования временных рядов на основе использования нечетких систем, позволяющих описать нечеткий вход.
3. Разработка алгоритмической реализации метода прогнозирования временных рядов на основе нечетких моделей.
4. Разработка информационной технологии прогнозирования временных рядов и прототипа ее программной реализации на основе нечеткой модели.
5. Оценка работоспособности программно-алгоритмической реализации на основе вычислительных экспериментов.

Научной новизной обладают следующие результаты диссертационного исследования:

1. Установлена неадекватность отражения свойств уникальных объектов на основе существующих методов прогнозирования временных рядов.
2. Модель нечеткого вывода на основе нечеткого значения истинности, позволяющая адекватно учесть изменчивость характеристик входных данных.
3. Метод прогнозирования временных рядов на основе разработанной системы нечеткого вывода.

4. Алгоритмы реализации метода прогнозирования на основе нечеткого вывода с учетом параллельности обеспечивающих полиномиальную вычислительную сложность.

Теоретическая значимость определяет принцип учета нечеткости входных данных, модель нечеткого вывода на этой основе и способы получения результата.

Практическая значимость: определяется возможностью построения систем прогнозирования временных рядов в задачах принятия решения с учетом неопределенности временных рядов на основе разработанных результатов. Применение разработанной информационной технологии позволяет осуществить прогнозирование временных рядов с изменчивыми характеристиками в задачах принятия решения. Теоретические результаты и разработанный прототип программного обеспечения используется в учебном процессе БГТУ им. В. Г. Шухова.

Методология и методы исследования. В работе использовалась методология принятия решений, методы теории временных рядов, теории нечетких множеств и машинного обучения.

Основные положения, выносимые на защиту:

1. Применение разработанной информационной технологии позволяет расширить системы управления для принятия обоснованных решений на основе временных рядов, описывающих уникальные объекты.
2. Разработанная программно-алгоритмическая реализация не предъявляет высоких требований к архитектуре вычислительных сред.
3. Результаты проведенных вычислительные эксперименты иллюстрируют работоспособность информационной технологии.

Область исследования. Содержание диссертации соответствует следующим пунктам паспорта специальности 2.3.1. «Системный анализ, управление и обработка информации, статистика» по следующим направлениям исследований:

- п. 1 паспорта специальности: Теоретические основы и методы системного анализа, оптимизации, управления, принятия решений, обработки информации и искусственного интеллекта.
- п. 4 паспорта специальности: Разработка методов и алгоритмов решения задач системного анализа, оптимизации, управления, принятия решений, обработки информации и искусственного интеллекта.

Достоверность полученных результатов обеспечивается корректностью математических преобразований, отсутствием противоречий с известными положениями теории и практики прогнозирования временных рядов, адекватным учетом нечеткость входных данных и иллюстрируется

результатами вычислительных экспериментов о работоспособности информационной технологии.

Апробация работы. Основные результаты работы докладывались на следующих конференциях:

1. Международная конференция «Перспективные компьютерные и цифровые технологии» (ACDT 2021), г. Белгород, 2021.
2. XV Международная научная конференция «Параллельные вычислительные технологии (ПаВТ) 2021», г. Волгоград, 2021.
3. XI Международной научно-практической конференции «Интегрированные модели и мягкие вычисления в искусственном интеллекте (ИММВ-2022)», г. Коломна, 2022 г.
4. XX Национальная конференция по искусственному интеллекту с международным участием (КИИ-2022), г. Москва, 2022.
5. XVII Международная научная конференция «Параллельные вычислительные технологии (ПаВТ) 2023», г. Санкт-Петербург, 2023.
6. XXI Национальная конференция по искусственному интеллекту с международным участием (КИИ-2023), г. Смоленск, 2023.
7. X Всероссийская научно-техническая конференция «Информационные технологии в науке, образовании и производстве» (ИТ-НОП-2025), г. Орел, 2025.

Личный вклад. все результаты диссертационного исследования получены либо автором лично, либо при его непосредственном участии.

Публикации. Основные результаты по теме диссертации изложены в 11 печатных изданиях, 3 из которых изданы в журналах, рекомендованных ВАК, 3 — в периодических научных журналах, индексируемых Web of Science и Scopus, 5 — в тезисах докладов. Зарегистрированы 2 программы для ЭВМ.

Содержание работы

Во введении описана актуальность работы, сформулированы цель и задачи исследования, изложены основные результаты, их теоретическая и практическая значимость, приведена новизна исследования и защищаемые положения.

Глава 1. «Анализ методов прогнозирования временных рядов в задачах управления с позиции адекватности учета уникальности свойств генерирующих их объектов»

В главе проведен анализ адекватности учета оценок различного характера неопределенности во временных рядах моделями их прогнозирования. Особое внимание удалено распространенным моделям прогнозирования, таким как, линейный предсказатель и модель Бокса-Дженкинса. В результате их анализа показано, что эти модели предполагают стационарность временного ряда, а статистические характеристики оцениваются

для ансамбля объектов. Это делает применение данных моделей неадекватным в случае описания временным рядом некоторого уникального объекта или нестационарности самого временного ряда.

На основе проведенного анализа сформулированы задачи исследования, решение которых направлено на достижение цели диссертационной работы.

Глава 2. «Разработка метода прогнозирования временных рядов на основе использования нечетких систем с учетом нечеткости входа»

посвящена разработке метода прогнозирования временных рядов на основе нечетких систем, адекватно учитывающей нечеткость временных рядов.

Задача нечеткого логического вывода Нечеткая система представляется собой базу правил вида:

$$R_k : \text{Если } x_1 \text{ есть } A_{k1} \text{ и } x_2 \text{ есть } A_{k2} \text{ и } \dots \text{ и } x_n \text{ есть } A_{kn}, \text{ то } y \text{ есть } B_k, \quad (1)$$

где N — количество нечетких правил, $A_{ki} \subseteq X_i, i = \overline{1, n}$, $B_k \subseteq Y$ — нечеткие множества, которые характеризуются функциями принадлежности $\mu_{A_{ki}}(x_i)$ и $\mu_{B_k}(y)$ соответственно; x_1, x_2, \dots, x_n — входные переменные, причем

$$[x_1, x_2, \dots, x_n]^T = \mathbf{x} \in X_1 \times X_2 \times \dots \times X_n.$$

Символами $X_i, i = \overline{1, n}$ и Y обозначаются соответственно пространства входных и выходной переменных. Если ввести обозначения $\mathbf{X} = X_1 \times X_2 \times \dots \times X_n$ и $\mathbf{A}_k = A_{k1} \times A_{k2} \times \dots \times A_{kn}$, причем

$$\mu_{\mathbf{A}_k}(\mathbf{x}) = \bigwedge_{i=1}^n \mu_{A_{ki}}(x_i),$$

где T_1 — произвольная t -норма, то правило 1 представляется в виде нечеткой импликации

$$R_k : \mathbf{A}_k \rightarrow B_k, k = \overline{1, N}. \quad (2)$$

Правило R_k можно формализовать как нечеткое отношение, определенное на множестве $\mathbf{X} \times Y$, т.е. $R_k \subseteq \mathbf{X} \times Y$ — нечеткое множество с функцией принадлежности

$$\mu_{R_k}(\mathbf{x}, y) = \mu_{\mathbf{A}_k \rightarrow B_k}(\mathbf{x}, y).$$

Модель логического типа определяет задание функции $\mu_{\mathbf{A}_k \rightarrow B_k}(\mathbf{x}, y)$ на основе известных функций принадлежности $\mu_{\mathbf{A}_k}(\mathbf{x})$ и $\mu_{B_k}(y)$ с помощью одной из функций нечеткой импликации:

$$\mu_{\mathbf{A}_k \rightarrow B_k}(\mathbf{x}, y) = I(\mu_{\mathbf{A}_k}(\mathbf{x}), \mu_{B_k}(y)),$$

где I — некоторая импликация.

Ставится задача определить нечеткий вывод $B'_k \subseteq Y$ для системы, представленной в виде (1), если на входах — нечеткие множества. $\mathbf{A}' = A'_1 \times A'_2 \times \cdots \times A'_n \subseteq \mathbf{X}$ или x_1 есть A'_1 и x_2 есть A'_2 и … и x_n есть A'_n с соответствующей функцией принадлежности $\mu_{\mathbf{A}'}(\mathbf{x})$, которая определяется как

$$\mu_{\mathbf{A}'}(\mathbf{x}) = \bigwedge_{i=1,n} T_3 \mu_{A'_i}(x_i). \quad (3)$$

Несинглтонный фаззификатор отображает измеренное $x_i = x'_i, i = \overline{1,n}$ в нечеткое число, для которого $\mu_{A'_i}(x'_i) = 1$ и $\mu_{A'_i}(x_i)$ уменьшается от единицы по мере удаления от x'_i . В соответствии с обобщенным нечетким правилом modus ponens, нечеткое множество B'_k определяется композицией нечеткого множества \mathbf{A}' и отношения \mathbf{R}_k , т.е.

$$B'_k = \mathbf{A}' \circ (\mathbf{R}_k \rightarrow B_k),$$

или, на уровне функций принадлежности

$$\mu_{B'_k}(y|\mathbf{x}') = \sup_{\mathbf{x} \in \mathbf{X}} \left\{ \mu_{\mathbf{A}'}(\mathbf{x}') \star^{T_2} I(\mu_{\mathbf{A}_k}(\mathbf{x}), \mu_{B_k}(y)) \right\}. \quad (4)$$

В (4) применена условная нотация, так как ввод в нечеткую систему происходит при определенном значении \mathbf{x} , а именно \mathbf{x}' . Обозначение $\mu_{B'_k}(y|\mathbf{x}')$ показывает, что $\mu_{B'_k}$ изменяется с каждым значением \mathbf{x}' .

Применение предложенного метода вывода в нечеткой модели для задачи регрессии временных рядов

Пусть задан временной ряд $\{y_t\}_{t=1}^T = \{y_1, \dots, y_T\}$, где $y_t \in \mathbb{R}$ — измеренное значение наблюдаемой переменной в момент времени t , а T — длина доступной выборки. При моделировании временных последовательностей с использованием нейро-нечетких систем каждое значение $y_t \in \mathbb{Y} \subseteq \mathbb{R}$ фаззифицируется в нечеткое множество A'_t . Тогда для прогнозирования значения \hat{y}_{t+h} с горизонтом h на основании среза наблюдений y_{t-p+1}, \dots, y_t можно использовать нечеткую систему с базой из N правил вида:

$$R_k : \text{Если } \bigwedge_{i=1}^p (y_{t-i+1} \text{ есть } A_{ki}), \text{ то } y_{t+1} \text{ есть } A_{kp+1}, k = \overline{1,N},$$

где p — размер окна запаздывания (порядок модели, количество входов нечеткой системы).

Процедура фаззификации. Если значения входных нечетких множеств нечеткой системы задаются с использованием синглтонного типа фаззификации теряется информация о неопределенности измеренных значений x' и они рассматриваются как достоверные. Поэтому в работе используется *несинглтонный способ фаззификации*.

Для фазификации с учетом неопределенности значений временного ряда в работе описана процедура адаптивной несинглтонной фазификации временной последовательности, приведенная на рисунке 1, где получаемая в результате фазификации функция принадлежности меняет форму на участках с низкой и высокой нестабильностью временных значений, как показано на рисунке 2.



Рисунок 1 — Схема обобщенной процедуры адаптивной несинглтонной фазификации временной последовательности.

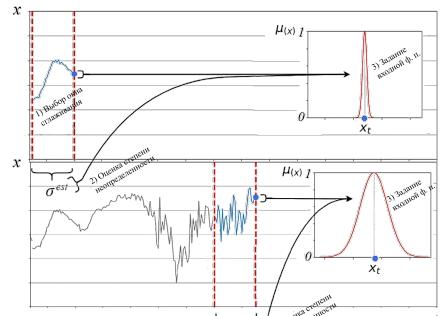


Рисунок 2 — Иллюстрация процедуры несинглтонной фазификации временного ряда с низким (вверху) и высоким (внизу) уровнем шума.

Выбор схемы дефазификации.

В задаче регрессии дискретная формула дефазификации по центру тяжести или другие более простые схемы не показывает достаточной точности, а непрерывная ее формулировка первой имеет большую вычислительную сложность. Поэтому в работе для регрессии используется дефазификация по среднему максимуму.

Глава 3. «Разработка алгоритмической реализации метода прогнозирования временных рядов на основе нечетких моделей»

посвящена разработке эффективного алгоритма прогнозирования временных рядов нечеткой системой с использованием несинглтонной фазификации посредством разработки алгоритмов нечеткого логического вывода на основе нечеткого значения истинности (НЗИ), а также параллельного алгоритма свертки НЗИ.

Вычислительная сложность выражения композиционного правила (4) для вывода логического типа при использовании non-singleton фазификации составляет $O(|X_1| \cdot |X_2| \cdots \cdot |X_n| \cdot |Y|)$ т.е. экспоненциальная. Это является актуальным препятствием использования нечеткого вывода логического типа совместно с фазификацией non-singleton.

Альтернативный метод нечеткого вывода с полиномиальной вычислительной сложностью

Нечеткое значение истинности нечеткого множества A относительно н. м. A' представляет собой нечеткое множество с функцией принадлежности совместимости $CP(A, A')$ A по отношению к A' , причем A' рассматривается как достоверное:

$$\tau_{A_k|A'}(v) = \mu_{CP(A_k, A')}(v) = \sup_{\substack{\mu_{A_k}(x)=v \\ x \in X}} \{\mu_{A'}(x)\}. \quad (5)$$

Для нечеткой системы с одним входом истинностное преобразование позволяет выполнить переход к новому виду формулы композиционного правила вывода:

$$\mu_{B'_k}(y|\mathbf{x}') = \sup_{v \in [0,1]} \left\{ \tau_{A_k|A'}(v) \overset{T_2}{\star} I(v, \mu_{B_k}(y)) \right\}. \quad (6)$$

Это соответствует новой структуре правил в базе правил:

$$\text{Если } \text{нзи есть ИСТИННО, то } y \text{ есть } B'_k \quad (7)$$

Для нечеткой системы с несколькими входами НЗИ вычисляются по каждому входу отдельно, а затем производится их свертка по расширенной по принципу обобщения \tilde{T} -норме.

$$\tau_{\mathbf{A}_k|\mathbf{A}'}(v) = \tilde{\tau}_{\substack{i=1, n \\ (v_1, \dots, v_n) \in [0,1]^n}} \tau_{A_{ki}|A'_i} = \sup_{\substack{\mathbf{T}_1 \\ i=1, n \\ v_i=v}} \left\{ \mathbf{T}_3 \tau_{A_{ki}|A'_i}(v_i) \right\} \quad (8)$$

$$= \sup_{\substack{\mathbf{T}_1 \\ i=1, n \\ (x_1, \dots, x_n) \in \mathbf{x}}} \left\{ \mathbf{T}_3 \mu_{A'_i}(x_i) \right\}, v \in [0, 1], \quad (9)$$

Рекурсивная схема вычисления свертки НЗИ по формуле (8) иллюстрируется выражением:

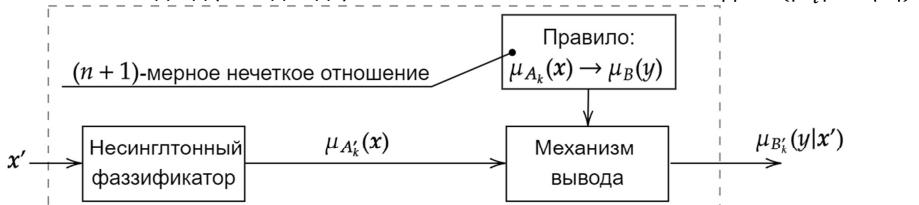
$$\tau_{A_k, A'}(v) = \tilde{\tau}_{\substack{i=1, n}} \tau_{A_{ki}|A'_i}(v_i) \quad (10)$$

$$= \left(\dots \left(\left(\mu_{CP(A_{k1}, A'_1)}(v_1) \tilde{T}_1 \mu_{CP(A_{k2}, A'_2)}(v_2) \right) \tilde{T}_1 \dots \right) \tilde{T}_1 \mu_{CP(A_{kn}, A'_n)} \right). \quad (11)$$

Тогда для системы с n входами выражения нечеткого вывода на основе НЗИ (6) примет вид:

$$\mu_{B'_k}(y|\mathbf{x}') = \sup_{v \in [0,1]} \left\{ \tau_{\mathbf{A}_k|\mathbf{A}'}(v) \overset{T_2}{\star} I(v, \mu_{B_k}(y)) \right\} \quad (12)$$

Классический подход (метод Заде):



Альтернативный подход (предлагаемый метод):

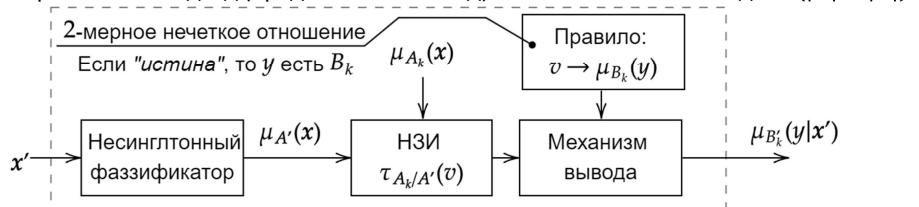


Рисунок 3 — Сравнение классической схемы нечеткого вывода и схемы нечеткого вывода на основе НЗИ

Порядок функции временной сложности вычисления B'_k на основе выражения (12) составляет $O(n|V|^2 + |V| \cdot |Y|)$, где $V = CP(A_k, A')$. Сравнение схем нечетких выводов с соответствием соотношениями (4) и (12) представлены на рис. 3.

Выход типа Мамдани

В [4] ослаблено ограничение на использование одной и той же T -нормы в формуле композиционного правила вывода в работах Менделля для случая вывода типа Мамдани, а также показано, что вывод по отдельному правилу в случае $T_2 = T_4 = T$ может быть записан через *меру возможности*:

$$\mu_{B'_k}(y) = \sup_{v \in [0;1]} \left\{ \tau_{A_k|A'} \star^T (v \star^T \mu_{B_k}(y)) \right\} = \prod_{A_k|A'} \star^T \mu_{B_k}(y),$$

где $\prod_{A_k|A'} = \sup_{v \in [0;1]} \left\{ \tau_{A_k|A'} \star^T v \right\}$.

Если при этом используется дефаззификация *по центру сумм* (*CoS*) и T -норма Ларсена, то, как доказано в [4], результат дефаззификации зависит от ширины гауссовой или треугольной функции принадлежности консеквента, тогда как дефаззификация *по среднему центру* учитывает только параметр центра. Например, при использовании в качестве выходной ф. п. гауссовой функции $\mu_{B_k}(y) = \exp(-((y - \bar{y}_k)/\sigma_k)^2)$ формула дефаззификации имеет вид:

$$\hat{y}_{CoS} = \frac{\int_{\mathbb{Y}} y \sum_{k=1}^N \prod_{\mathbf{A}_k|\mathbf{A}'} \overset{\text{T}_2}{\star} \mu_{B_k}(y)}{\int_{\mathbb{Y}} \sum_{k=1}^N \prod_{\mathbf{A}_k|\mathbf{A}'} \overset{\text{T}_2}{\star} \mu_{B_k}(y)} = \frac{\sum_{k=1}^N \prod_{\mathbf{A}_k|\mathbf{A}'} \bar{y}_k \sigma_k}{\sum_{k=1}^N \prod_{\mathbf{A}_k|\mathbf{A}'} \sigma_k}, \quad (13)$$

поскольку $\int_{-\infty}^{\inf} \mu_{B_k}(y) dy = \sigma_k \sqrt{\pi}$ и $\int_{-\infty}^{\inf} y \mu_{B_k}(y) dy = \bar{y}_k \sigma_k \sqrt{\pi}$.

Для предложенного метода вывода на основе нечеткого значения истинности в [5] показано, что для достаточно удаленных и непересекающихся выходных ф. п. нечетких множеств, т. е. когда $\mu_{B_k}(\bar{y}_r) = 0$ для $k \neq r$, сложность вычисления дефазификации по центру тяжести сокращается за счет упрощения выражений импликаций:

– для S -импликации

$$\hat{y}_{CoG} = \frac{\sum_{k=1}^N \bar{y}_k \text{T}_{r=1}^N \left\{ \sup_{v \in [0,1]} \left\{ \tau_{A_r|A'} \overset{\text{T}_2}{\star} (1-v) \right\} \right\}}{\sum_{k=1}^N \text{T}_{r=1}^N \left\{ \sup_{v \in [0,1]} \left\{ \tau_{A_r|A'} \overset{\text{T}_2}{\star} (1-v) \right\} \right\}},$$

– для R -импликации

$$\hat{y}_{CoG} = \frac{\sum_{k=1}^N \bar{y}_k \text{T}_{r=1}^N \left\{ \tau_{A_r|A'}(0) \right\}}{\sum_{k=1}^N \text{T}_{r=1}^N \left\{ \tau_{A_r|A'}(0) \right\}},$$

– для Q -импликации

$$\hat{y}_{CoG} = \frac{\sum_{k=1}^N \bar{y}_k \text{T}_2 \left\{ \sup_{v \in [0,1]} \left\{ \tau_{A_k|A'} \overset{\text{T}_2}{\star} \max(1-v, v) \right\} \text{T}_{r=1}^N \left\{ \sup_{v \in [0,1]} \left\{ \tau_{A_r|A'} \overset{\text{T}_2}{\star} \max(1-v, v) \right\} \right\} \right\}}{\sum_{k=1}^N \text{T}_2 \left\{ \sup_{v \in [0,1]} \left\{ \tau_{A_k|A'} \overset{\text{T}_2}{\star} \max(1-v, v) \right\} \text{T}_{r=1}^N \left\{ \sup_{v \in [0,1]} \left\{ \tau_{A_r|A'} \overset{\text{T}_2}{\star} \max(1-v, v) \right\} \right\} \right\}},$$

Использованная в работе дефазификация по среднему максимуму для вывода на основе нечеткого значения истинности:

$$\tau_{A_k|A'}(v) = \tilde{\text{T}}_{i=1,n} \left\{ \mu_{CP(A_{ki}, A'_{t-p+i})}(v_i) \right\}, k = \overline{1, N}, \quad (14)$$

$$\hat{y}_{t+h} = \arg \max_{y \in \mathbb{Y}} \left\{ \text{T}_{k=1}^N \left\{ \sup_{v \in [0,1]} \left\{ \tau_{\mathbf{A}_k|\mathbf{A}'} \overset{\text{T}_2}{\star} I(v, \mu_{A_{k,p+1}}(y)) \right\} \right\} \right\}. \quad (15)$$

Параллельный алгоритм свертки НЗИ

При программной реализации вычисления и свертки НЗИ $\tau_{A_{ki}|A'_i}$ вычисление производится в точках расчетной сетки. Значение НЗИ по i -му входу в точке расчетной сетки v_j в данной работе обозначается $ftv_i[v_j]$ (ftv

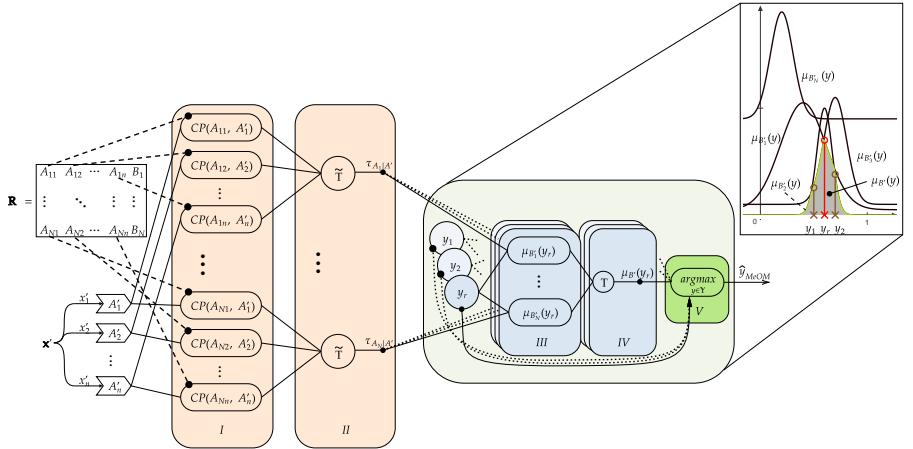


Рисунок 4 — Схема нейро-нечеткой системы с вычислением НЗИ и дефазификацией по среднему максимуму, а также пример работы процедуры дефазификации.

— *fuzzy truth value*). Расчетная сетка размера D_{ftv} задается на пространстве $\mathbb{V} = [0; 1]$ мощности $|\mathbb{V}|$.

Для нахождения свертки НЗИ по одному правилу можно составить алгоритм на основе формулы (10). Вычислительная сложность при параллельной реализации такого алгоритма составит $O(D_{ftv}^2 \cdot \log n)$. Значения $ftv_i[v_j]$ необходимо вычислить до запуска алгоритма свертки, что потребует сложности по памяти $O(D_{ftv} \cdot n)$.

Algorithm 1 Алгоритм свертки НЗИ при $T_1 = \min$

```

Require:  $ftv_i, i = \overline{1, n}$  — это  $\tau_{A_{ki}|A'_i}$  дискретизированная в точках  $v_j$ 
 $max\_ftv[i] = 0;$ 
for  $v_j = 1 \dots 0$  do
     $s \leftarrow \{ftv_i[v_j] \mid ftv_i[v_j] \geq max\_ftv[i]\};$ 
     $max\_ftv[i] \leftarrow max(max\_ftv[i], ftv_i[v_j]);$ 
     $v\_max\_index \leftarrow arg\max_i \{ftv_i[v_j]\};$ 
    if  $s = \emptyset \& i = v\_max\_index$  then
         $r[i] \leftarrow ftv_i[v_j];$ 
    else
         $r[i] \leftarrow max\_ftv[i];$ 
    end if
     $ftv\_reduced[v_j] \leftarrow T_3 \{r[i]\};$ 
end for
return  $ftv\_reduced$ 

```

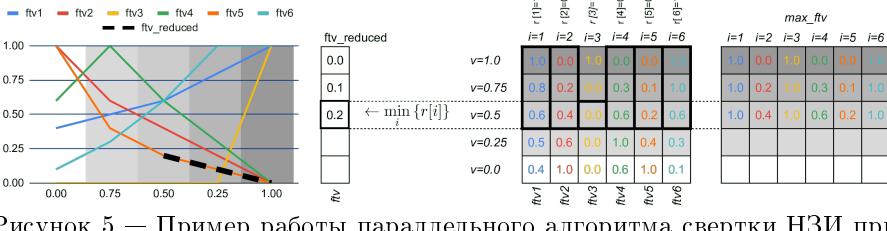


Рисунок 5 — Пример работы параллельного алгоритма свертки НЗИ при расчетной сетке состоящей из 5 точек.

При выполнении работы был разработан параллельный алгоритм свертки НЗИ [6] с вычислительной сложностью $O(D_{ftv} \cdot \log n)$. Алгоритм 1 разработан при допущении, что $T_1 = \min$, тогда $\frac{T_1}{i} v_i = \min_{i=1,n} v_i$ в формуле (8). При работе алгоритм итеративно продвигается от 1 к 0 в области \mathbb{V} , как показано на рисунке 5. На каждой j -й итерации вычисляются значения $f_{tv_i}[v_j]$, которые агрегируются в одном вспомогательном массиве \max_ftv , который требует сложности по памяти $O(n)$.

Глава 4. «Разработка информационной технологии прогнозирования временных рядов, прототипа ее программной реализации на основе нечеткой модели и оценка работоспособности»

содержит описание нескольких проведенных экспериментов для оценки характеристик разработанной нечеткой модели прогнозирования временных рядов. Первый эксперимент направлен на подтверждение полиномиальной зависимости времени нечеткого вывода от количества входов нечеткой системы, а также на прирост качества прогнозирования при использовании нечеткого вывода логического типа с несинглтонной фазификацией. Второй эксперимент был проведен для оценки качества в прикладной задаче прогнозирования временных рядов.

Библиотека с параллельной реализацией нечеткого вывода на основе технологии CUDA

Для более оперативного проведения экспериментов и практического применения в нагруженных промышленных приложениях была выполнена параллельная реализация нечеткого вывода на основе НЗИ с использованием языка программирования C++ и технологии CUDA. Для удобства использования к разработанному модулю вывода был реализован интерфейс из языка Python с помощью расширения Cython. Схема использования изображена на рисунке 6.

Важным достоинством реализации является полная организация вычислений внутри *потоковых мультипроцессоров* графических ускорителей за счет размещения всей базы правил и пакета экземпляров входных данных в разделяемой памяти на чипе потокового мультипроцессора,

что предотвращает возникновение простых арифметико-логических модулей при ожидании загрузки порции данных (например, базы правил) из глобальной памяти графического ускорителя. Также, вычисления организованы внутри группы из 32 CUDA-нитей, что избавляет от необходимости синхронизации внутри CUDA-блока и позволяет использовать инструкции аппаратной свертки массивов чисел внутри таких групп.

В библиотеке реализована дискретизированная дефазификация по центру тяжести и дефазификация по среднему максимуму. Для точного вычисления значения дефазификации выхода нечеткой системы для задачи регрессии используется метод оптимизации *Gradient-aware Particle Swarm Optimization*.

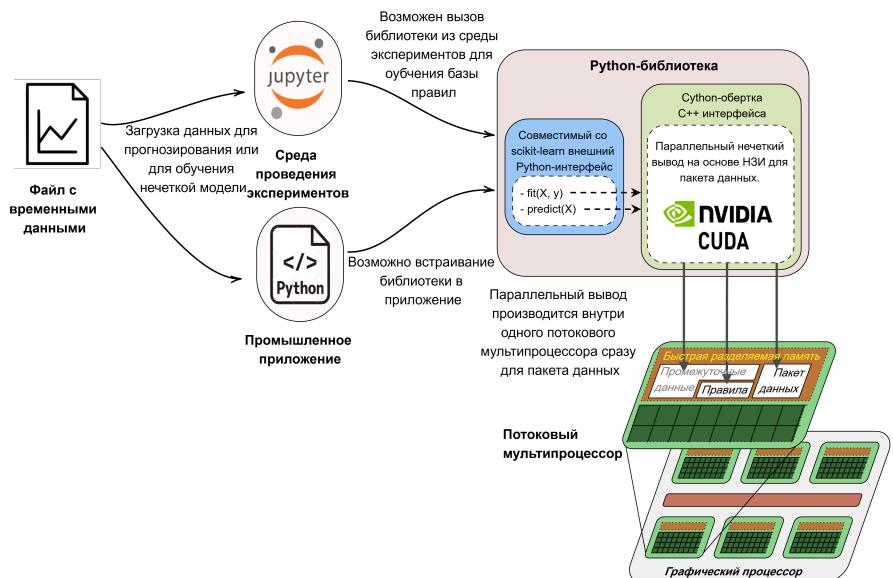


Рисунок 6 — Схема использования библиотеки для нечеткого моделирования и прогнозирования временных рядов.

Первый эксперимент проводился с использованием синтетического набора данных Mackey-Glass (M-G). В данной работе этот набор данных был сгенерирован в результате решения дифференциального уравнения:

$$\frac{dx(t)}{dt} = \beta \frac{x(t - \tau)}{1 + x(t - \tau)^n} - \gamma x(t), \quad (16)$$

со значениями параметров $\tau = 30$, $\beta = 0.2$, $\gamma = 0.1$.

В эксперименте использовался участок временного ряда $t = \overline{1,1000}$, а также применялся адаптивный метод оценки зашумленности временной

последовательности в каждой точке t на основе экспоненциально взвешенного скользящего среднего.

Нечеткие множества для значений временного ряда были получены с использованием этой процедуры фазификации временных рядов (1) для обеспечения адаптивности оценки неопределенности в конкретной точке. Формулы для вычисления степени неопределенности показаны ниже

$$d_1 = d_2, \\ d_t = \frac{1}{\sqrt{2}}(x_t - x_{t-1}), \quad (17) \quad \hat{\sigma}_t^2 = (1 - \alpha)\hat{\sigma}_{t-1}^2 + \alpha(d_t - \hat{d}_t)^2$$

$$\hat{d}_t = (1 - \alpha)\hat{d}_{t-1} + \alpha d_t \quad (19) \\ = \sum_{p=1}^t \alpha(1 - \alpha)^{t-p} d_p. \quad (18) \quad \hat{\sigma}_t = \sqrt{\hat{\sigma}_t^2}, \quad (20)$$

После проведенного вычислительного эксперимента были получены показатели удельного (на одну точку из набора точек одной итерации алгоритма PSO) времени работы параллельного алгоритма нечеткого вывода на основе НЗИ на обучающем наборе данных и времени работы на тренировочном наборе данных для различного размера окна запаздывания, изображенные на рисунке 8. На этом рисунке наблюдается линейный рост времени выполнения алгоритма с увеличением количества входов нечеткой системы, что подтверждает утверждение о полиномиальной зависимости временной сложности метода нечеткого вывода на основе НЗИ от количества входов.

Для сравнения с альтернативными нечеткими моделями условия этого эксперимента были выбраны такими же как в публикации [`<empty citation>`], приводящей результаты прогнозирования временного ряда M-G для нечетких систем типа Мамдани при синглтонной и несинглтонной фазификации. Качество прогнозирования оценивалось с использованием метрики sMAPE:

$$sMAPE = \frac{100\%}{n} \sum_{t=1}^n \frac{|y_t - \hat{y}_t|}{\frac{|y_t| + |\hat{y}_t|}{2}},$$

где y_t и \hat{y}_t соответствуют истинному и предсказанному значению в момент времени t .

По итогу проведенного эксперимента лучший показатель качества прогнозирования по этой метрике на рисунке 9 достигается при размере окна запаздывания — 3 точки.

Достигнутое в этом случае значение $sMAPE = 8\%$ при количестве правил 30, заметно превосходит точность моделирования незашумленной последовательности Mackey-Glass (M-G) с использованием синглтонной

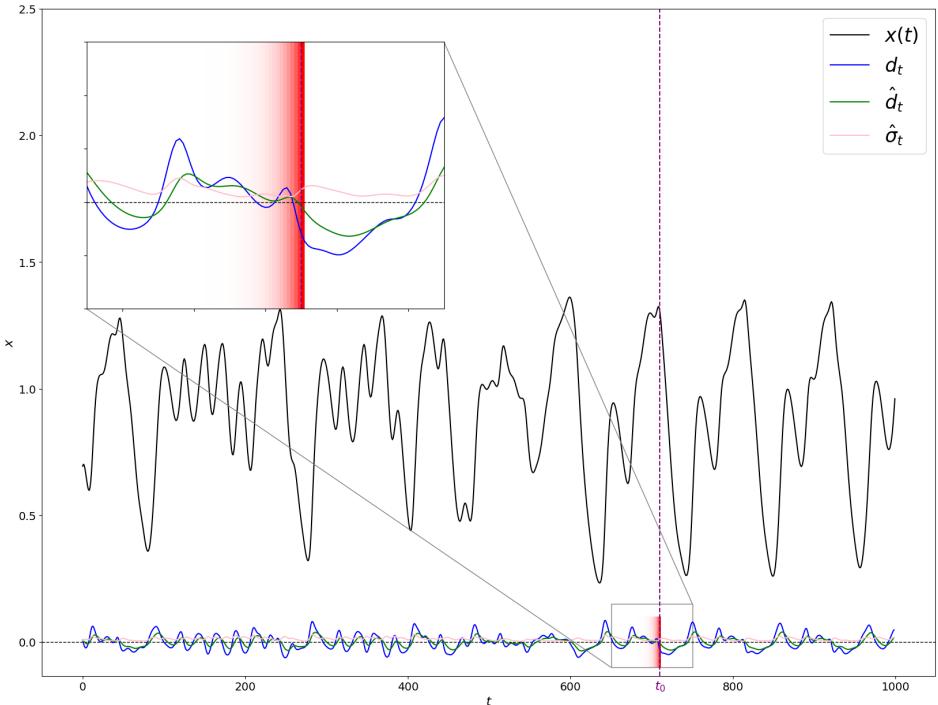


Рисунок 7 — График сгенерированной последовательности Mackey-Glass $x(t)$, $t \in [0; 999]$, график разностей соседних точек d_t последовательности $x(t)$, график \hat{d}_t с наложением экспоненциально взвешенного слаживания на последовательность разностей и график экспоненциально взвешенного скользящего среднеквадратичного отклонения разностей $\hat{\sigma}_t$. На вложенном изображении участка $t \in [650, 750]$ яркостью красного цвета показано значения весового коэффициента $\alpha(1 - \alpha)^{t-p}$, $p = \overline{1, t_0}$ при $\alpha = 0.2$.

фазификации со значением $sMAPE \approx 40\%$ в [14]. Также полученное качество прогнозирования сопоставимо со значением этой метрики при аналогичной конфигурации эксперимента прогнозирования временного ряда M-G в той же публикации, где для достижения точности моделирования временной последовательности $sMAPE \approx 10\%$ используется нечеткая система типа Мамдани, содержащая 184 правила¹ в базе правил для не зашумленного временного ряда M-G и 597, 402, 312 правил для различных конфигураций и амплитуды добавленного шума. Данные наблюдения

¹Непосредственно в самой статье не указано количество используемых правил. Однако авторами этой статьи опубликован программный код описанного в статье метода, запустив который удалось воспроизвести проводимые авторами эксперименты и восстановить количество правил в их нечетких системах.

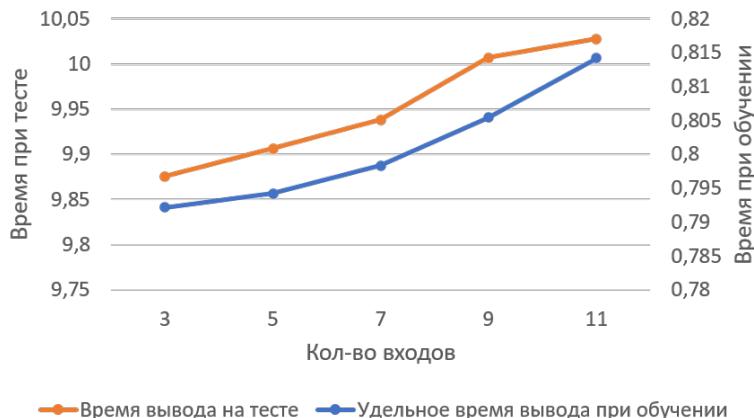


Рисунок 8 — График длительности выполнения параллельной реализации нечеткого вывода для обучающего и тестового набора данных при количестве правил $N = 30$.

показывают, что метод регрессии временных рядов на основе логического нечеткого вывода при несинглтонной фазификации значительно превосходит качество регрессии с использованием синглтонной фазификации, а также имеет сопоставимую с методом регрессии Мамдани при несинглтонной фазификации

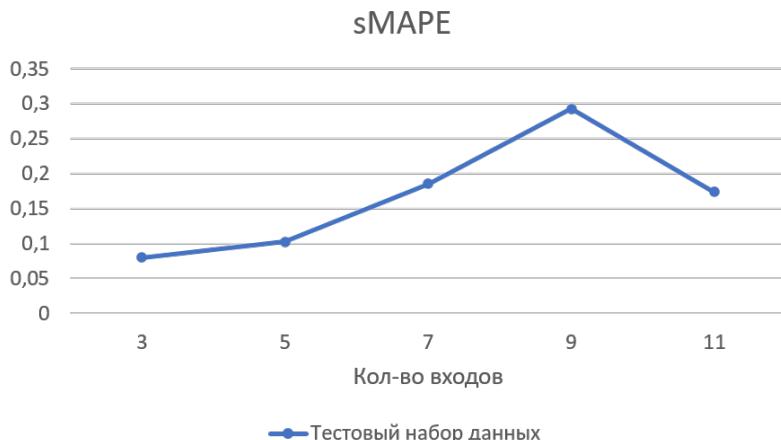


Рисунок 9 — График значений метрики sMAPE на обучающем и тестовом наборах данных при различных размерах окна запаздывания, количестве правил — 30.

Таблица 1 — Значения метрики MAE относительно $\mathbb{E}[\mathbb{E}[s_t]]$ для разработанной NSFLS+FTV и альтернативных моделей.

Модель	Relative Mean Absolute Error (RMAE)	
	Прогноз на 1 мес.	Средняя за 3 мес.
Наивная (среднее по ряду)	0.166	0.204
ARIMA	0.187	0.230
Градиентный бустинг	0.152	0.187
Многослойный перцептрон	0.128	0.157
Рекуррентная нейросеть (DeepAR)	0.188	0.231
NSFLS+FTV	0.112	0.142

точность, но при меньшем количестве правил за счет возможности построения более сложной функции аппроксимации.

Второй эксперимент посвящен применению разработанного метода для прогнозирования помесячного объема транзакций безналичных платежей корпоративным клиентам банка. Источником неопределенности в этом наборе данных является стохастический характер динамики временной последовательности. Поэтому в этом эксперименте также использовалась процедура аддитивной несинглтонной фазификации. В качестве метрики использовалась RMAE (relative MAE) — отношение средней абсолютной ошибки к математическому ожиданию средних по рядам клиентов:

$$RMAE = \frac{MAE}{\mathbb{E}[\mathbb{E}[s_t]]}.$$

В таблице 1 приведены значения метрики качества при прогнозирования на 1 и 3 месяца для моделей: наивная, ARIMA, градиентный бустинг, многослойный перцептрон, DeepAR.

Стоит описать некоторые детали применения и используемую конфигурацию других моделей:

- В наивной модели прогноз формируется как среднее значение в окне запаздывания.
- При использовании ARIMA для каждого клиента обучалась отдельная авторегрессионная модель.
- В качестве градиентного бустинга использовался CatBoostRegressor. На вход модели подавались признаки для различных размеров окон запаздывания: 3, 6, 9 и 12. Для каждого прогнозного месяца была построена отдельная модель бустинга.
- Многослойный перцептрон имел размерность скрытых слоев — 1024 и количество полносвязных слоев — 4.
- В рекуррентной модели DeepAR использовалась размерность скрытого состояния LSTM-слоя — 128 и количество слоев LSTM — 4.

По итогу эксперимента было получено значение метрики $RMAE = 0.109/0.142$ при прогнозировании на один/три месяца соответственно. Это

превосходит на 17%/11% качество прогнозирования с использованием многослойного перцентрона со значениями $RMAE = 0.128/0.157$.

В **заключении** сделаны выводы о полученных в процессе работы результатах.

Основные результаты и выводы

1. Проведенный анализ показал, что существующие методы не позволяют адекватно отразить уникальность исследуемого объекта и учесть изменчивость характеристик данных входных временных рядов.
2. Разработан метод прогнозирования временных рядов в задачах управления на основе нечеткого вывода позволяет адекватно отразить уникальность и изменчивость характеристик входных данных.
3. Разработаны вычислительные процедуры предложенного метода прогнозирования временных рядов, обеспечивающая вычислительную эффективность за счет использования предложенных алгоритмов нечеткого вывода и свертки НЗИ.
4. Разработанный прототип программной поддержки информационной технологии прогнозирования на основе параллельной архитектуры вычислений продемонстрировал высокую работоспособность при сравнении с другими методами прогнозирования.

Выводы

- Разработанная информационная технология позволяет усовершенствовать инструменты информационного обеспечения прогнозирования на основе временных рядов.
- Реализация информационной технологии может быть выполнена стандартных компьютерах, оснащенных графическими ускорителями с поддержкой технологии CUDA.

Перспективы дальнейших исследований

Повышение эффективности алгоритма прогнозирования временных рядов с использованием нечеткой системы за счет более эффективного способа агрегации базы правил.

Рекомендации по использованию

Рекомендуется организациям разрабатывающим программы для информационного обеспечения.

ОСНОВНЫЕ ПУБЛИКАЦИИ ПО ТЕМЕ ДИССЕРТАЦИИ

В изданиях из списка ВАК РФ

1. *Sinuk V. G., Karatach S. A.* Inference Method and Parallel Implementation for MISO Structure Systems for Inputs with Linguistic Values [Текст] // Journal of Information Technologies and Computing Systems. — 2020. — № 3. — С. 85—93. — DOI: 10.14357/20718632200308.
2. *Karatach S. A., Sinuk V. G.* Machine Learning of a Fuzzy System with Linguistic Inputs Using Parallel Technologies [Текст] // Journal of Information Technologies and Computing Systems. — 2021. — № 3. — С. 60—69. — DOI: 10.14357/20718632210306.
3. *Karatach S. A., Sinuk V. G.* Parallel Implementation of Evolutionary Learning of a Fuzzy System with Non-Singleton Fuzzification [Текст] // Journal of Information Technologies and Computing Systems. — 2023. — № 2. — С. 113—122. — DOI: 10.14357/20718632230212.

В изданиях, входящих в международную базу цитирования Web of Science

4. *Sinuk V. G., Karatach S. A.* The Inference Method for a Mamdani Type System with Nonsingleton Fuzzification [Текст] // Pattern Recognition and Image Analysis. Advances in Mathematical Theory and Applications. — 2023. — Т. 33, № 3. — С. 506—510. — DOI: 10.1134/S1054661823030422.
5. *Karatach S. A., Sinuk V. G.* Method of Inference of Fuzzy Logical Type Systems with Nonsingleton Fuzzification [Текст] // Pattern Recognition and Image Analysis. — 2024. — Сент. — Т. 34, № 3. — С. 645—651. — DOI: 10.1134/S1054661824700470. — URL: <https://doi.org/10.1134/S1054661824700470>.

В изданиях, входящих в международную базу цитирования Scopus

6. *Karatach S., Sinuk V.* Implementation of a Fuzzy Inference Method with Nonsingleton Fuzzification Based on CUDA and GPGPU Technologies [Текст] // Parallel Computational Technologies: Proceedings of the International Scientific Conference (PCT 2023). Т. 1868 / под ред. L. Sokolinsky, M. Zymbler. — Cham : Springer, 2023. — С. 43—58. — (Communications in Computer and Information Science (CCIS)). — DOI: https://doi.org/10.1007/978-3-031-38864-4_4.

В сборниках трудов конференций

7. *Karatach S. A., Sinuk V. G.* Parallel implementation of methods for inference systems of computational intelligence MISO-structure with fuzzy inputs [Текст] // Параллельные вычислительные технологии (ПаВТ'2020): Короткие статьи и описания плакатов (31 марта—2, апр. 2020). — Пермь : Издательский центр ЮУрГУ, 2020. — С. 29—39.

8. *Karatach S. A., Sinuk V. G.* Parallel implementation of evolutionary learning of a fuzzy system with quality inputs using CUDA technology [Текст] // Параллельные вычислительные технологии (ПаВТ'2021): Короткие статьи и описания плакатов. XV международная конференция (30 марта—1, апр. 2021). — Волгоград : Издательский центр ЮУрГУ, 2021. — С. 67—80.
9. *Синюк В. Г., Карапач С. А.* Метод вывода для систем типа Мамдани с несинглтонной фаззификацией [Текст] // Двадцатая Национальная конференция по искусственному интеллекту с международным участием, КИИ-2022: Труды конференции. В 2-х томах. Т. 2 (21—23 дек. 2022). — Москва : Национальный исследовательский университет "МЭИ", 2022. — С. 55—64.
10. *Карапач С. А., Синюк В. Г.* Метод и параллельная реализация нечеткой классификации объектов с несинглтонной фаззификацией [Текст] // Интегрированные модели и мягкие вычисления в искусственном интеллекте ИММВ-2022: Сборник научных трудов XI Международной научно-практической конференции. В 2-х томах. Т. 1 (16—19 мая 2022). — Коломна : Общероссийская общественная организация «Российская ассоциация искусственного интеллекта», 2022. — С. 67—73.
11. *Синюк В. Г., Карапач С. А.* Метод вывода нечетких систем логического типа при несинглтонной фаззификации [Текст] // Двадцать первая Национальная конференция по искусственному интеллекту с международным участием (КИИ-2023): Труды конференции. В 2-х томах. — Смоленск : Принт-Экспресс, 2023. — С. 118—130.

Зарегистрированные программы для ЭВМ

12. *Свидетельство о гос. регистрации программы для ЭВМ.* Среда для параллельного обобщенного нечеткого вывода на основе нечеткого значения истинности [Текст] / В. Г. Синюк, С. А. Карапач ; федеральное государственное бюджетное образовательное учреждение высшего образования «Белгородский государственный технологический университет им. В.Г. Шухова». — № 2025661081 ; заявл. 05.05.2025 ; опубл. 28.05.2025, 2025663372 (Рос. Федерация).
13. *Свидетельство о гос. регистрации программы для ЭВМ.* Библиотека моделирования временных рядов на основе параллельного нечеткого вывода с использованием нечеткого значения истинности [Текст] / В. Г. Синюк, С. А. Карапач ; федеральное государственное бюджетное образовательное учреждение высшего образования «Белгородский государственный технологический университет им. В.Г. Шухова». — № 2025660990 ; заявл. 05.05.2025 ; опубл. 28.05.2025, 2025663464 (Рос. Федерация).

Каратач Сергей Александрович

Разработка методов и алгоритмов обработки временных рядов в задачах их прогнозирования на основе нечетких систем с учетом нечеткости входов

Автореф. дис. на соискание ученой степени канд. техн. наук

Подписано в печать _____. Заказ № _____

Формат 60×90/16. Усл. печ. л. 1. Тираж 100 экз.

Типография _____

