

Нечеткая модель второго типа прогнозирования временных рядов с хедж- алгеброй и алгоритмом генетической ОПТИМИЗАЦИИ

Нгуен Тхи Тху Зунг, Л. В. Черненькая

Аннотация—Чтобы соответствовать современным требованиям развития социально-экономических проблем необходимо разрабатывать и совершенствовать модели прогнозирования. Существующие модели прогнозирования нечетких временных рядов (FTS) построены на основе теории нечеткой логики первого типа, однако теория нечеткой логики второго типа показывает больший охват предметных областей и более точное моделирование состояния объектов и систем. Это важно, поскольку в реальности степень принадлежности элемента к конкретному множеству не может быть определена точно, а только в пределах диапазона. В данной работе предложена нечеткая модель прогнозирования временных рядов, основанная на теории нечеткой логики второго типа и структуры алгебры Хеджа. Параметры предложенной модели оптимизированы с помощью генетических алгоритмов. Предложенная модель тестируется на прогнозе ежедневных значений данных тайваньского фондового индекса (TAIEX), а результативность прогнозирования оценивается по метрикам RMSE, MAPE и MSE.

Ключевые слова—Нечеткая логика второго типа, нечеткая модель прогнозирования временных рядов второго типа, хедж-алгебра, генетический алгоритм, оптимизация модели прогнозирования, фазификация, дефазификация второго типа.

I. ВВЕДЕНИЕ

Одним из значительных достижений в информационных технологиях является появление нечеткой теории, предложенной Заде в 1965 году [1]. В дальнейшем, в 1993 году Сонгом и Чисомом разработана нечеткая модель прогнозирования временных рядов [2]. С тех пор нечеткие модели прогнозирования временных рядов применяются и совершенствуются для прогнозирования различных ситуаций [3]. Прогнозирование нечетких временных рядов - это краткое название прогнозирования нечетких временных рядов первого порядка. Для совершенства модели нечетких временных рядов первого порядка по качеству точности прогнозирования целесообразно использовать следующие подходы: техника дискретизации универсума дискурса с помощью методов на основе

отношений, методов оптимизации (метод роя частиц (PSO), генетический алгоритм), методы кластеризации при фазификации (метода K -средних (KM) и нечеткой кластеризации C -средних (FCM)), построение моделей высокого порядка, эвристических моделей, гибридных моделей и многофакторных моделей [4], [5].

На практике во многих реальных случаях невозможно однозначно определить степень принадлежности элемента к нечеткому множеству. Это затрудняет применение модели прогнозирования нечетких временных рядов первого типа. Для преодоления указанной проблемы Хуанг в своей работе предложил модель прогнозирования нечетких временных рядов второго типа [6]. В данной модели нечеткие логические группы отношений (FLRG) получены из модели нечеткого временного ряда первого типа, а затем расширены до модели нечеткого временного ряда второго типа путем определения дополнительных переменных, таких как "высокий", "низкий", с применением операторов объединения и пересечения на нечетком множестве второго типа в процессе дефазификации. Преимуществом нечеткого множества второго типа является способность выражать больше информации. В работе [7] гибридные интеллектуальные системы, нейронные сети и интервальная нечеткая логика второго типа представлены и применены для прогнозирования хаотических временных рядов. Результаты прогнозирования показали, что интервальная нечеткая логика второго типа превосходит некоторые гибридные интеллектуальные подходы, а нейронные сети дают результат, сопоставимый с нечеткими системами второго типа.

Модели нечеткого прогнозирования второго типа для временных рядов разрабатываются и совершенствуются многими исследователями. В работе [8] предложен новый метод прогнозирования тайваньского фондового индекса на основе оптимизированного нечеткого временного ряда высокого порядка второго типа. В работе [9] предлагается интервальная нечеткая весовая подстройка второго типа для нейронных сетей с обратным распространением, которая применяется в задаче прогнозирования временных рядов. Кроме того, в

Статья получена 19 ноября 2023.

Нгуен Тхи Тху Зунг, аспирантка, Россия, Санкт-Петербург, Санкт-Петербургский Политехнический университет Петра Великого (e-mail: thudung.mta.tb@gmail.com; ORCID: <https://orcid.org/0000-0001-9206-5968>).

Черненькая Людмила Васильевна, д-р техн. наук, проф., Россия, Санкт-Петербург, Санкт-Петербургский Политехнический университет Петра Великого (e-mail: ludmila@qmd.spbstu.ru).

работе [10] нечеткие нейронные сети интервального второго типа (IT2) применяются для нелинейной идентификации и прогнозирования временных рядов. Разработана новая модель прогнозирования временных рядов второго типа на основе трехфакторных нечетких логических отношений, при этом используется новая техника разделения пространства рассуждений на интервалы различной длины для разных факторов [11]. Предложена интервальная система нечеткой логики второго типа (IT2FLS) для прогнозирования фондового индекса на основе нечеткого временного ряда и нечеткой карты логических связей (FLRM) [12]. Кроме того, в работе [13] предложен гибридный метод прогнозирования на основе весовой корректировки нейронной сети с использованием обобщенного нечеткого множества второго типа. Также в работе [14] предложена модель временных рядов на основе нечеткого множества второго типа. Прогнозирование временных рядов на основе интервального нечеткого множества второго типа с использованием подхода, основанного на разбиении данных, описано в работе [15]. Новая модель IT2-FCM-FTS, использующая алгоритм интервального второго типа FCM вместо традиционного FCM для разделения области выборки и улучшения производительности модели прогнозирования временных рядов, представлена в работе [16].

В моделях прогнозирования нечетких временных рядов первого типа широко применяется метод Хеджа-алгебры, однако его применение в моделях прогнозирования нечетких временных рядов второго типа пока ограничено. Наряду с этим комбинированное применение методов оптимизации для моделей прогнозирования нечетких временных рядов второго типа по-прежнему требует большого внимания и развития [17], [18], [19].

В данной статье рассмотрена новая модель прогнозирования нечетких временных рядов второго типа с использованием подхода Хеджа-алгебры. В частности, процессы семантизации и десемантизации используются для эффективной поддержки фазификации и дефазификации на основе выявления существенных связей между лингвистическими метками (а также семантическим значением) и граничными интервалами соответствующих нечетких множеств. Для оптимизации результатов прогнозирования путем минимизации значения ошибки MARE в модели применен генетический алгоритм для получения величины основных входных параметров.

Данная статья разделена на 6 частей. Первая часть представляет собой введение. В следующей части представлены теоретические основы модели. В третьей части приведена структура предлагаемой модели и этапы ее реализации. В четвертой части представлено применение предложенной модели для прогнозирования реальных данных. В пятой части приводится оценка эффективности предложенной модели. Обобщенные выводы представлены в Заключение.

II. ТЕОРЕТИЧЕСКИЕ ОСНОВЫ

A. Основы теории нечеткого множества второго типа

Определение 1. Нечеткое множество второго типа

можно определить как расширение нечеткого множества первого типа.

Пусть U — универсума дискурса; \tilde{A} — нечеткое множество в U . Тогда нечеткое множество \tilde{A} второго типа в X — это нечеткое множество, степени принадлежности которого сами по себе являются нечеткими и тогда

$$\tilde{A} = \{(x, \mu_{\tilde{A}}(x)) \mid \mu_{\tilde{A}}(x) \in \tilde{A}(U) : x \in X\},$$

где $\mu_{\tilde{A}} : X \rightarrow \tilde{A}(U)$. Заметили, что в нечетком множестве первого типа степень принадлежности характеризуется четким значением, тогда в нечетком множестве второго типа степень принадлежности рассматривается как нечеткая.

B. Прогнозирование нечетких временных рядов второго типа

Определение 2. Нечеткая модель временных рядов второго типа использует нечеткие отношения, установленные моделью первого типа на основе наблюдений первого типа. Операторы используются для включения или исключения нечетких связей, полученных из наблюдений первого типа и второго типа. Прогнозы второго типа затем рассчитываются на основе этих нечетких отношений.

Определение 3. Операторы объединения и пересечения определены для вычисления отношений между двумя нечеткими логическими группами отношений:

$$\vee(LHS_d, LHS_e) = RHS_d \cup RHS_e$$

$$\wedge(LHS_d, LHS_e) = RHS_d \cap RHS_e,$$

где \vee — оператор объединения и \wedge — оператор пересечения в теории множеств; LHS и RHS — это левая и правая стороны группы нечетких логических отношений d и e соответственно.

C. Хедж-алгебраический подход

Определение 5. Хедж-алгебраическая структура имеет вид $AX = (X, G, C, H, \leq)$, где: $G = \{c^+, c^-\}$ — множество первичных образующих, c^+ и c^- отрицательный и положительный первичный термин лингвистической переменной X ; $C = \{0, W, 1\}$ — набор констант, которые различаются элементами в X ; H — набор хеджа, где $H^- = \{h_{-1}, h_{-2}, \dots, h_{-q}\}, h_{-1} < h_{-2} < \dots < h_{-q}$ и $H^+ = \{h_{+1}, h_{+2}, \dots, h_{+p}\}, h_{+1} < h_{+2} < \dots < h_{+p}$.

Определение 6. Семантическая квантификация

Пусть $v : X \rightarrow [0, 1]$ — семантическое количественное отображение, порожденное fm на X . Тогда семантическая квантификация метки определяется следующим образом:

$$v(h_j, x) = v(x) + \text{sign}(h_j, x) \left(\sum_{i=\text{sign}(j)}^j fm(h_i, x) - \omega(h_j, x) fm(h_j, x) \right)$$

где

$$\omega(h_j, x) = \frac{1}{2} [1 + \text{sign}(h_j, x) \text{sign}(h_p, h_j, x) (\beta - \alpha)], j \in [-q^+ p], j \neq 0$$

где функция $Sign(x)$ - представляет собой знаковую функцию; fm называется мерой нечеткости.

III. ПРЕДЛОЖЕННАЯ МОДЕЛЬ

В данной работе предложена новая модель для повышения эффективности прогнозирования модели нечетких временных рядов второго типа с использованием подхода хедж-алгебры. Для оптимизации результатов модели прогнозирования путем минимизации значения MAPE при изменении входных параметров в предложенной модели используются генетические алгоритмы. Модель выполнена для различного количества входных наблюдений: 50, 100, 150, 200, 250 соответственно. Структура предложенной модели прогнозирования представлена на рис. 1. Этапы реализации предложенной модели описаны следующим образом:

Шаг 1. Ввод данных

Входной временной ряд имеет вид $F_i(t)$ со значениями в наборе X , где $X = \{x_i\}$, $i = \overline{1, n}$, n — количество наблюдений, соответствующее количеству моментов во входном временном ряду. Предположим, что в каждый момент времени t наблюдаемое значение имеет верхнее и нижнее предельное значение \bar{x}_i и \underline{x}_i соответственно, тогда временные ряды второго типа имеют вид $\bar{F}_i(t)$ и $\underline{F}_i(t)$, расположенный в соответствующих множествах \bar{X} и \underline{X} , где $\bar{X} = \{\bar{x}_i\}$, $\underline{X} = \{\underline{x}_i\}$, $i = \overline{1, n}$.

Шаг 2. Дискретизация

Универсум дискурса имеет вид $U = [X_{\min} - U_1, X_{\max} + U_2]$. На данном шаге выполняется деление универсума дискурса U на N интервалов, т.е. $U = \{u_j\}$; $j = \overline{1, N}$.

Шаг 3. Фаззификатор

Целью данного шага является построение нечетких множеств A_j на основе интервалов u_j , затем определение принадлежности наблюдений к нечетким множествам, одновременно определение семантики и семантической квантификации для лингвистических переменных соответствующих нечетких множеств.

Шаг 3.1. Определение нечетких множеств.

Нечеткие множества A_j определяются следующим образом:

$$\begin{aligned} A_1 &= 1/u_1 + 0.5/u_2 + 0/u_3 + \dots + 0/u_N, \\ A_2 &= 0.5/u_1 + 1/u_2 + 0.5/u_3 + \dots + 0/u_N, \\ &\dots \\ A_{N-1} &= 0/u_1 + 0/u_2 + \dots + 0.5/u_{N-2} + 1/u_{N-1} + 0.5/u_N, \\ A_N &= 0/u_1 + 0/u_2 + \dots + 0/u_{N-2} + 0.5/u_{N-1} + 1/u_N. \end{aligned}$$

Шаг 3.2. Семантизация и семантическая квантификация

Семантизация объясняется тем, что лингвистические метки $\{hA_jc\}$ соответствуют нечетким множествам A_j , структурированным в соответствии с подходом хедж-алгебры, т.е. $A_1 \rightarrow \{hA_1c\}$, $A_2 \rightarrow \{hA_2c\}$, ..., $A_N \rightarrow \{hA_Nc\}$,

где hA_j представляет собой хедж-ряд, действующий с $c = \{c^-, c^+\}$.

Семантическая квантификация выполняется путем расчета семантических значений по параметрам α, β, θ следующим образом: $\{hA_jc\} \rightarrow f_j(\theta, \alpha, \beta)$ посредством отображений $v, fm, sign$.

Шаг 3.3. Фаззификация

Наблюдаемые значения $x_i, \bar{x}_i, \underline{x}_i$ временного ряда $F_i(t), \bar{F}_i(t), \underline{F}_i(t)$ фаззифицируются путем определения их принадлежности к нечетким множествам A_j .

Шаг 4. Построение групп нечетких логических отношений

После фаззификации наблюдаемых значений временного ряда $F_i(t)$, на данном шаге создаются группы нечетких логических отношений. Тогда m -я группа имеет вид $SA_m \rightarrow SA_k$, где $m, k \in [1, N]$.

Шаг 5. Определение нечетких отношений

Из построенных групп нечетких логических отношений прогнозирование наблюдения в момент i определяется через наблюдаемое значение в момент времени $i-1$. То есть для целевого временного ряда $F_i(t)$ получается $SA_m^{(i-1)} \rightarrow SA_k^{(i-1)}$, для ряда $\bar{F}_i(t)$ и $\underline{F}_i(t)$, получаются $\overline{SA}_m^{(i-1)} \rightarrow \overline{SA}_k^{(i-1)}$ и $\underline{SA}_m^{(i-1)} \rightarrow \underline{SA}_k^{(i-1)}$ соответственно.

Шаг 6. Дефаззификация

Шаг 6.1. Дефаззификация второго типа

Для дефаззификации второго типа необходимо определить объединение и пересечение, где объединение имеет вид $SA_V^{(i)} = V(SA_k^{(i-1)}, \overline{SA}_k^{(i-1)}, \underline{SA}_k^{(i-1)})$ и пересечение имеет вид $SA_\Lambda^{(i)} = \Lambda(SA_k^{(i-1)}, \overline{SA}_k^{(i-1)}, \underline{SA}_k^{(i-1)})$.

Шаг 6.2. Десемантизация

Десемантизация определяется на основе последствий нелинейной десемантизации со значениями параметра семантизации sp и параметра десемантизации dp следующим образом:

Для объединения: $x_V^{(i)} = g_{norm}(sp, dp, SA_V^{(i)})$, для пересечения: $x_\Lambda^{(i)} = g_{norm}(sp, dp, SA_\Lambda^{(i)})$.

Шаг 6.3. Дефаззификация первого типа

Дефаззификация первого типа — это определение прогнозируемых значений для целевого временного ряда, то есть прогнозируемое значение, наблюдаемое в момент времени i , равно:

$$x^{(i)} = \frac{x_V^{(i)} + x_\Lambda^{(i)}}{2}$$

Шаг 7. Оптимизация

Оптимизация осуществляется с помощью генетического алгоритма. Чем точнее модель прогнозирования, тем меньше значение MAPE. Таким образом, генетический алгоритм (GA) оптимизирует значения входных параметров $j, \theta, \alpha, \beta, sp, dp$, так, чтобы минимизировать значение MAPE, т.е. $ga(RMSE, j, \theta, \alpha, \beta, sp, dp)$.

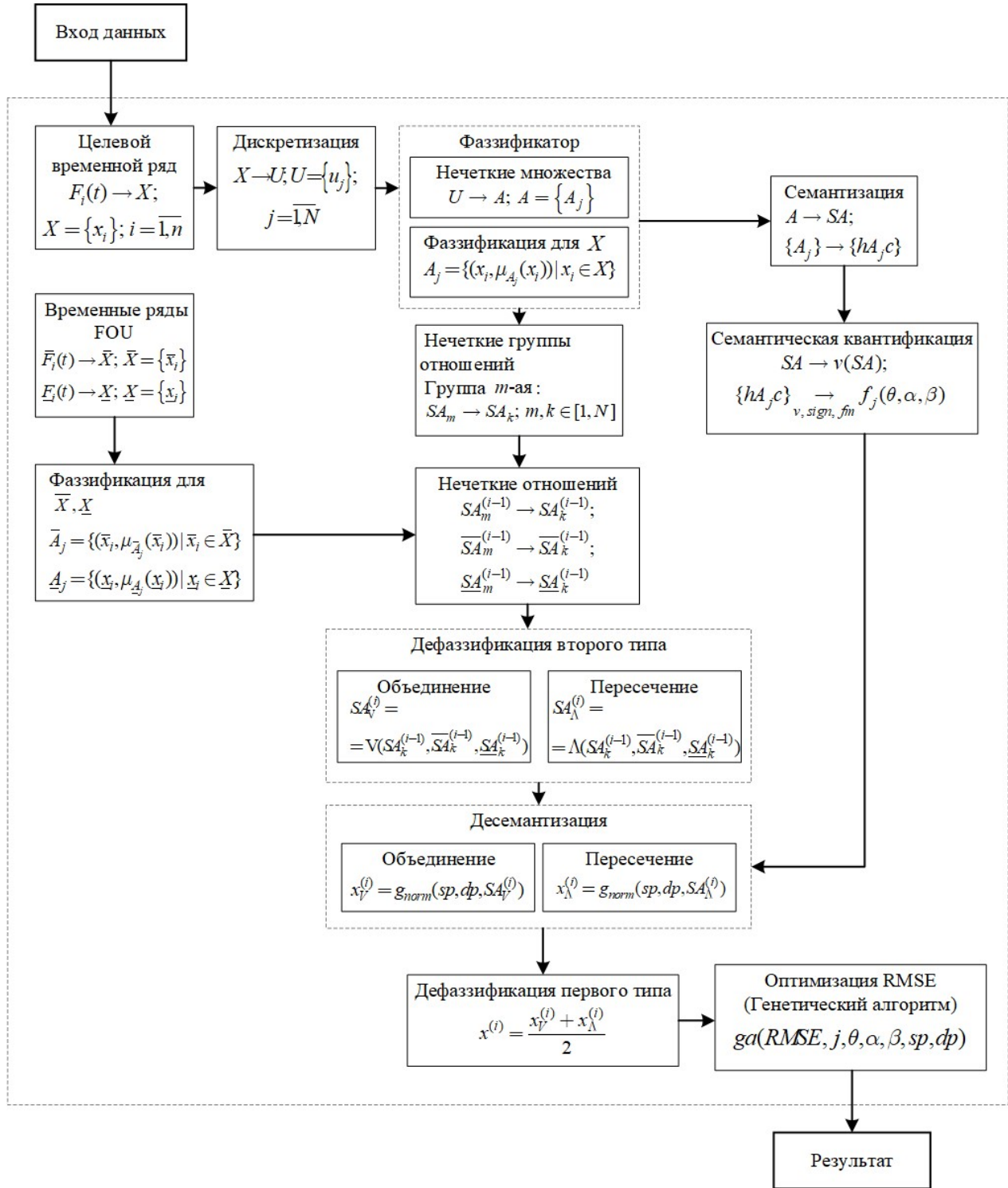


Рис. 1. Структура предложенной модели

IV. ЭКСПЕРИМЕНТ И АНАЛИЗ РЕЗУЛЬТАТОВ

A. Эксперимент

В работе предложенная модель была протестирована на прогнозировании ежедневных данных ТАИEX в период с 04.01.2000 до 21.3.2000. Значения ежедневных данных показаны в таблице 1.

Шаг 1. Ввод данных

Целевой временной ряд — это ряд ежедневных значений данных ТАИEX на момент закрытия. Временные ряды второго типа представляют собой

ежедневный ряд самых высоких значений и ряд самых низких значений данных ТАИEX. Тогда наблюдаемое множество значения $X = [8250, 10394]$.

Таблица 1. Ежедневные данные ТАИEX

Момент	$F(t)$	$\bar{F}(t)$	$\underline{F}(t)$
...
01.04.00	8757	8804	8643
01.05.00	8849	8868	8668
01.06.00	8922	9024	8864
01.07.00	8846	8941	8739

01.10.00	9103	9126	8892
01.11.00	8927	9333	8891
...

Шаг 2. Дискретизация

Универсум дискурса определяется $U = [8200, 10400]$. В данной работе универсум дискурса U разбит на 47 равных интервалов, т.е. $U = \{[u_j]\}$, $u_1 = [8200, 8246.8]$, $u_2 = [8246.8, 8293.6], \dots, u_{47} = [10353, 10400]$

Шаг 3. Фаззификатор

Шаг 3.1. Определите нечеткие множества

Треугольные нечеткие множества определяются разделенными интервалами следующим образом:

$$A_1 = 1/u_1 + 0.5/u_2 + 0/u_3 + \dots + 0/u_{48},$$

$$A_2 = 0.5/u_1 + 1/u_2 + 0.5/u_3 + \dots + 0/u_{48},$$

...

$$A_{46} = 0/u_1 + 0/u_2 + \dots + 0.5/u_{45} + 1/u_{46} + 0.5/u_{47},$$

$$A_{47} = 0/u_1 + 0/u_2 + \dots + 0/u_{45} + 0.5/u_{46} + 1/u_{47}.$$

Функции принадлежности, соответствующие нечетким множествам, показаны на рис. 2.

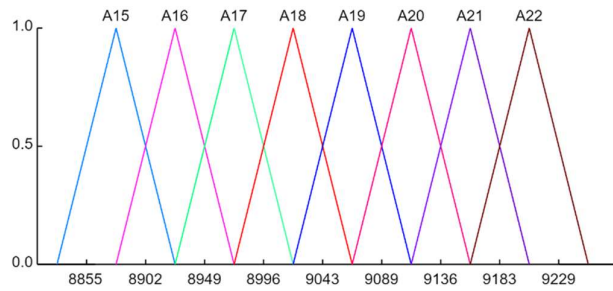


Рис 2. Треугольная нечеткая функция принадлежности типа I

Шаг 3.2. Семантизация и семантическая квантификация

Семантизация лингвистических переменных выполняется в соответствии со структурой хедж-алгебры с множеством первичных образующих $G = \{c^-, c^+\} = \{low, high\}$, при наборе констант $C = \{0, W, 1\}$, $W = \{middle\}$, и наборе хеджа $H^- = \{h_{-1}\} = \{little\}$, $H^+ = \{h_{+1}\} = \{very\}$.

Тогда лингвистические переменные структурированы следующим образом:

$$A_1 = [\underbrace{h_{+1} \dots h_{+1}}_{11} c^-] = \underbrace{V..V}_{11} Lo; \dots; \quad A_{11} = [h_{+1} c^-] = VLo;$$

$$A_{12} = [c^-] = Lo; \quad A_{13} = [h_{-1} c^-] = LLo; \dots;$$

$$A_{23} = [\underbrace{h_{-1} \dots h_{-1}}_{11} c^-] = \underbrace{L..L}_{11} Lo; \quad A_{24} = [W] = middle;$$

$$A_{25} = [\underbrace{h_{-1} \dots h_{-1}}_{11} c^+] = \underbrace{L..L}_{11} Hi; \dots; \quad A_{35} = [h_{-1} c^+] = LHi;$$

$$A_{36} = [c^+] = Hi; \quad A_{37} = [h_{+1} c^+] = VHi; \dots;$$

$$A_{47} = [\underbrace{h_{+1} \dots h_{+1}}_{11} c^+] = \underbrace{V..V}_{11} Hi;$$

Положительные хеджи усиливают семантическое значение, то есть оттягивают лингвистическую

переменную от положения равновесия W , а отрицательные хеджи действуют наоборот (см. рис. 3).



Рис. 3. Семантический уровень лингвистических переменных

При семантической квантификации лингвистических переменных параметры определены следующим образом: $fm(0) = fm(W) = fm(1) = 0$, $fm(c^+) = 1 - \theta$, $fm(c^-) = \theta$, $v(c^+) = \alpha + \beta\theta$, $v(c^-) = \beta\theta$. В данной работе, значения $\alpha = 0.0085$ и $\theta = 0.4745$ выбраны.

Тогда семантическая квантификация выполняется следующим образом:

$$\begin{aligned} v(A_{11}) &= v(h_{+1} c^-) \\ &= v(c^-) + sign(h_{+1} c^-) [fm(h_{+1} c^-) - \omega(h_{+1} c^-) \cdot fm(h_{+1} c^-)] \\ &= \theta \beta^2 \approx 0.4665 \end{aligned}$$

где

$$\omega(h_{+1} c^-) = \frac{1}{2} [1 + sign(h_{+1} c^-) sign(h_{+1} h_{+1} c^-) (\beta - \alpha)]$$

Аналогично определяются семантические квантификации для всех лингвистических переменных.

Шаг 3.3. Фаззификация

Для фаззификации временного ряда необходимо определить принадлежность наблюдений к нечетким множествам. Наблюдения относят к тому нечеткому множеству, в котором они имеют наибольшую принадлежность (см. таблица 2).

Таблица 2. Фаззификация

Момент	$F(t)$	НМ ¹	$\bar{F}(t)$	НМ	$\underline{F}(t)$	НМ
...
01.10.00	9103	A_{20}	9126	\bar{A}_{20}	8892	\underline{A}_{15}
01.11.00	8927	A_{16}	9333	\bar{A}_{25}	8891	\underline{A}_{15}
01.12.00	9145	A_{21}	9145	\bar{A}_{21}	8937	\underline{A}_{16}
01.13.00	9107	A_{20}	9238	\bar{A}_{23}	9095	\underline{A}_{20}
01.14.00	9023	A_{18}	9237	\bar{A}_{23}	8968	\underline{A}_{17}
01.17.00	9315	A_{24}	9385	\bar{A}_{26}	9230	\underline{A}_{23}
01.18.00	9250	A_{23}	9354	\bar{A}_{25}	9210	\underline{A}_{22}
01.19.00	9151	A_{21}	9336	\bar{A}_{25}	9151	\underline{A}_{21}
01.20.00	9137	A_{21}	9211	\bar{A}_{22}	9078	\underline{A}_{19}
...

Шаг 4. Построить группы нечетких логических отношений

На основе фаззификации группы нечетких логических отношений представлены в таблице 3.

Таблица 3. Группы нечетких логических отношений

Закрытый	Высокий	Низкий
...
$A_{14} \rightarrow A_{10}, A_{14}, A_{16}, A_{20}$	$\bar{A}_{18} \rightarrow A_{16}$ $\bar{A}_{20} \rightarrow A_{25}$	$\underline{A}_{15} \rightarrow A_{12}, A_{15}, A_{16}$ $\underline{A}_{16} \rightarrow A_{20}$

¹ НМ – Нечеткое множество

$A_{16} \rightarrow A_{14}, A_{21}$	$\bar{A}_{21} \rightarrow A_{23}$	$\underline{A}_{17} \rightarrow A_{23}$
$A_{18} \rightarrow A_{24}$	$\bar{A}_{22} \rightarrow A_{24}$	$\underline{A}_{19} \rightarrow A_{21}$
$A_{20} \rightarrow A_{16}, A_{18}$	$\bar{A}_{23} \rightarrow A_{23}$	$\underline{A}_{20} \rightarrow A_{17}$
$A_{21} \rightarrow A_{20}, A_{21}, A_{23}$	$\bar{A}_{24} \rightarrow A_{15}, A_{27}$	$\underline{A}_{21} \rightarrow A_{19}, A_{25}$
$A_{23} \rightarrow A_{21}, A_{26}$	$\bar{A}_{25} \rightarrow A_{21}, A_{22}, A_{25}$	$\underline{A}_{22} \rightarrow A_{21}$
$A_{24} \rightarrow A_{23}$	$\bar{A}_{26} \rightarrow A_{25}, A_{31}$	$\underline{A}_{23} \rightarrow A_{22}$
...

Шаг 5. Определение нечетких отношений

Чтобы определить прогнозируемое значение для времени t , используется группа отношений нечеткой логики, соответствующая времени $t - 1$.

Например, для прогнозирования значения наблюдения

Таблица 4. Нечеткие отношения

Момент	$F(t)$	$\bar{F}(t)$	$\underline{F}(t)$
...
01.11.00	$A_{20} \rightarrow A_{16}, A_{18}$	$\bar{A}_{20} \rightarrow A_{25}$	$\underline{A}_{15} \rightarrow A_{12}, A_{15}, A_{16}$
01.12.00	$A_{16} \rightarrow A_{14}, A_{21}$	$\bar{A}_{25} \rightarrow A_{21}, A_{22}, A_{25}$	$\underline{A}_{15} \rightarrow A_{12}, A_{15}, A_{16}$
01.13.00	$A_{21} \rightarrow A_{20}, A_{21}, A_{23}$	$\bar{A}_{21} \rightarrow A_{23}$	$\underline{A}_{16} \rightarrow A_{20}$
01.14.00	$A_{20} \rightarrow A_{16}, A_{18}$	$\bar{A}_{23} \rightarrow A_{23}$	$\underline{A}_{20} \rightarrow A_{17}$
01.17.00	$A_{18} \rightarrow A_{24}$	$\bar{A}_{23} \rightarrow A_{23}$	$\underline{A}_{17} \rightarrow A_{23}$
01.18.00	$A_{24} \rightarrow A_{23}$	$\bar{A}_{26} \rightarrow A_{25}, A_{31}$	$\underline{A}_{23} \rightarrow A_{22}$
01.19.00	$A_{23} \rightarrow A_{21}, A_{26}$	$\bar{A}_{25} \rightarrow A_{21}, A_{22}, A_{25}$	$\underline{A}_{22} \rightarrow A_{21}$
01.20.00	$A_{21} \rightarrow A_{20}, A_{21}, A_{23}$	$\bar{A}_{25} \rightarrow A_{21}, A_{22}, A_{25}$	$\underline{A}_{21} \rightarrow A_{19}, A_{25}$
...

Шаг 6. Дефаззификация

Шаг 6.1. Дефаззификация второго типа

Для дефаззификации второго типа необходимо выполнить два действия: определение объединения и определение пересечения.

Например, для прогноза значения наблюдения 13 января 2000 г. используются следующим нечеткие отношения: Закрытый: $A_{21} \rightarrow A_{20}, A_{21}, A_{23}$; Низкий: $\underline{A}_{16} \rightarrow A_{20}$; Высокий: $\bar{A}_{21} \rightarrow A_{23}$.

По оператору объединения:

$$A_{\vee}[13.1.2000] = \vee((A_{20}, A_{21}, A_{23}), A_{20}, A_{23}) = A_{20}, A_{21}, A_{23}$$

По оператору пересечения:

$$A_{\wedge}[13.1.2000] = \wedge((A_{20}, A_{21}, A_{23}), A_{20}, A_{23}) = \emptyset$$

Шаг 6.2. Десемантизация

Десемантизация выполняется на основе последствий нелинейной десемантизации по оператору объединения и пересечения, значения параметров семантизации и десемантизации выбираются $sp = 0.6774$ и $dp = -0.626$. Результаты расчета представлены в таблице 5.

Таблица 5. Эвристические нечеткие отношения высокого порядка

Момент	объединение	пересечение
01.10.00	8874	8855
01.11.00	9135	9136
01.12.00	8922	8949
01.13.00	9205	9183

13 января 2000 г. необходимо учитывать, что значение наблюдения 12 января 2000 г. принадлежит нечеткой группе A_{21} для временных рядов “закрытый” и \bar{A}_{21} , \underline{A}_{16} для временных рядов «высокий» и «низкий» соответственно (согласно табл. 2), поэтому используется группа нечетких логических отношений $A_{21} \rightarrow A_{20}, A_{21}, A_{23}$ для временных рядов “закрытый” и $\bar{A}_{21} \rightarrow A_{23}$; $\underline{A}_{16} \rightarrow A_{20}$ для временных рядов «высокий» и «низкий» соответственно (согласно табл. 3) для прогноза на 13 января 2000 года.

Нечеткие отношения, определенные для прогнозирования каждого момента времени, показаны в таблице 4.

01.14.00	9135	9136
01.17.00	9112	9043
01.18.00	9184	9323
01.19.00	9276	9277
01.20.00	9205	9183
01.21.00	9205	9183
01.24.00	9276	9277
01.25.00	9489	9417
01.21.00	9205	9183

Например, вычисляется значение прогноза для [13.1.2000] следующим образом:

Из таблицы 7 нечеткие соотношения используется для прогноза на момент [13.1.2000]:

По оператору объединения: $A_{\vee}[13.1.2000] = A_{20}, A_{21}, A_{23}$

Семантическое значение определяется следующим образом:

$$\begin{aligned} x_s[13.1.2000] &= (2/5) * A_{20} + (1/5) * A_{21} + (2/5) * A_{23} \\ &\approx (2/5) * 0.4705 + (1/5) * 0.4705 + (2/5) * 0.4705 \\ &\approx 0.4705 \end{aligned}$$

Как показано на рис. 2, A_{20}, A_{21}, A_{23} находятся на отрезке от $a = 9066$ до $b = 9300$, тогда:

$$g_{norm}(x_s, sp) = g_{norm}(0.4705, 0.6774) \\ = (0.6774 * 0.4705 * (1 - 0.4705) + \\ + 0.4705)(9300 - 9066) + 9066 \\ \approx 9216$$

Спрогнозированное значение определяется из нелинейной денормализации:

$$x_v[13.1.2000] = g_{norm}(x_s, dp) = g_{norm}(0.4705, -0.626) \\ \approx -0.626 * (9216 - 9066) * (9300 - 9216) / (9300 - 9066) + 9216 \\ \approx 9205^{**}$$

*, ** Расчетные значения могут отличаться из-за аппроксимации

По оператору пересечения: $A_{\wedge}[13.1.2000] = \emptyset$

Поскольку оно равно пустому множеству, для прогнозирования берем нечеткое множество A_{21} на 12 января 2000 г. Аналогично, после выполнения десемантизации и денормализации получаем $x_{\wedge}[13.1.2000] = 9183$.

Шаг 6.3. Дефаззификация первого типа

Дефаззификация первого типа дает результаты прогнозирования временного ряда, показанные в таблице 6.

Таблица 6. Результаты дефаззификации первого типа

Момент	Прогноз
01.10.00	8865
01.11.00	9136
01.12.00	8936
01.13.00	9194
01.14.00	9136

01.17.00	9078
01.18.00	9254
01.19.00	9277
01.20.00	9194
01.21.00	9194
01.24.00	9277
01.25.00	9453
01.26.00	9453
01.27.00	9570
01.28.00	9629
01.31.00	9676

Например, прогнозное значение на 13 января 2000 года рассчитывается следующим образом:

$$x[13.1.00] = \frac{x_v[13.1.00] + x_{\wedge}[13.1.00]}{2} = \frac{9205 + 9183}{2} = 9194$$

Шаг 7. Оптимизация

После проведения оптимизации путем применения генетического алгоритма с целевым значением MAPE результаты оптимизации при разных количествах наблюдений показаны в таблице 7.

Параметры для получения оптимальных результатов прогнозирования отображаются в столбцах "параметров", которым будут соответствовать наименьшие найденные значения ошибок для MSE, MAPE, RMSE, как и в столбце "метрика".

Значение MAPE наименьшее, когда входные данные равны $n = 50$, количество нечетких множеств равно $N = 35$, а соответствующие параметры $\alpha = 0.058$, $\theta = 0.4615$, $sp = 0.3827$, $dp = 0.0389$.

Таблица 7. Оптимизация по MAPE

n	Параметры					Метрика		
	N	α	θ	sp	dp	MSE	MAPE	RMSE
50	35	0.058	0.4615	0.3827	0.0389	27342	0.0130	165.35
100	75	0.0589	0.6197	0.1074	-0.3123	27733	0.0140	166.53
150	139	0.0508	0.5241	-0.1574	-0.1552	23487	0.0134	153.25
200	187	0.0502	0.5174	-0.6163	-0.0736	25782	0.01527	160.56
250	235	0.05	0.4248	-0.1965	0.1537	25481	0.0165	159.63

В. Анализ результатов прогнозирования

Сходство фактических значений и прогнозных значений по предложенной модели изображено на рис. 4, 5 для случая с $n = 50$ и $n = 250$ соответственно. Видно, что прогнозные и фактические значения для каждого случая достаточно близки друг к другу при средних разницах, составляющих примерно 1,3% и 1,65%, соответственно.

Для оценки эффективности прогнозирования предложенной модели в данной статье применены следующие меры погрешности между фактическими значениями и прогнозируемыми значениями:

Средняя абсолютная ошибка в процентах:

$$MAPE = \frac{1}{n} \sum_{i=1}^n \frac{|x(t) - \tilde{x}(t)|}{x(t)}$$

Среднеквадратичная ошибка:

$$MSE = \frac{1}{n} \sum_{i=1}^n (x(t) - \tilde{x}(t))^2$$

Корень из среднеквадратичной ошибки:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (x(t) - \tilde{x}(t))^2}$$

где $x(t)$ – фактическое значение в момент t ;

- $\tilde{x}(t)$ – прогнозируемое значение в момент t ;
- n – количество наблюдений.

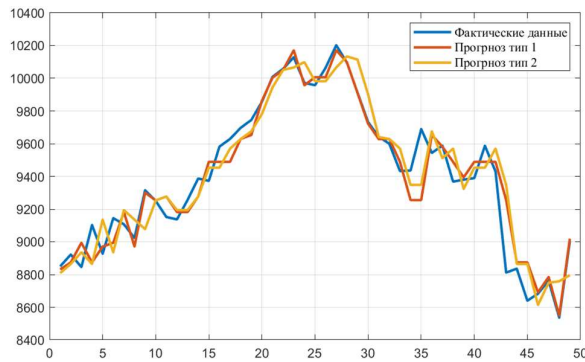


Рис. 4. Фактические и прогнозируемые значения при $n = 50$.

В случае с $n = 50$, модель прогнозирования дает среднюю процентную ошибку MAPE примерно 1,3%. Данный показатель говорит о том, что ошибка прогноза составила 1,3% от фактических значений. Среднеквадратичная разница между прогнозируемыми значениями и фактическими значениями (по метрике MSE) составляет $MSE = 27342$ и квадратный корень из средней квадратичной разницы между прогнозируемыми значениями и фактическими значениями (по метрике RMSE) составляет $RMSE = 165.35$.

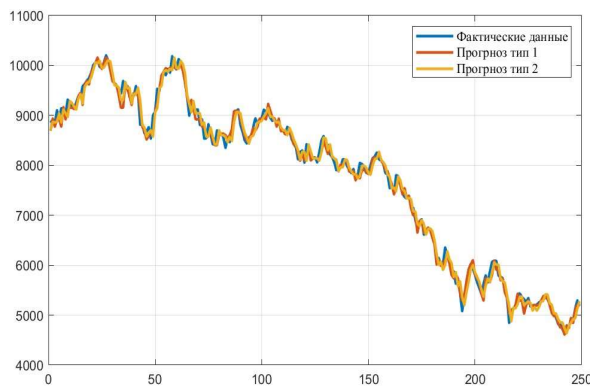


Рис 5. Фактические и прогнозируемые значения при $n = 250$

Для случая с $n = 250$ результаты прогнозирования показали, что средняя процентная ошибка MAPE составляет примерно 1,65%, среднеквадратичная разница между прогнозируемыми значениями и фактическими значениями равна $MSE = 25481$ и квадратный корень из средней квадратичной разницы между прогнозируемыми значениями и фактическими значениями составляет $RMSE = 159.63$.

V. ЗАКЛЮЧЕНИЕ

В данной работе предложена модель прогнозирования временных рядов второго типа с применением хедж-алгебры для прогнозирования временного ряда. Результаты прогнозирования оптимизируются с помощью генетического алгоритма путем минимизации значения меры ошибки MAPE при разных количествах наблюдений. Практическое применение модели показано для прогнозирования временных рядов ежедневных значений акций ТАИEX с использованием значений закрытия ТАИEX в качестве целевого временного ряда, а также максимальных и минимальных значений ТАИEX за тот же период для нечеткой логики второго типа.

Предложенная модель прогнозирования больше подходит для коротких временных рядов. При увеличении времени увеличивается и ошибка. Результаты прогнозирования оценены путем сравнения прогнозируемых значений и фактических значений временного ряда.

БЛАГОДАРНОСТИ

Выражаем благодарность учёным и исследователям в рассматриваемой области, предложившим эффективное направление прогнозирования временных рядов, которое является основой для разработки новых улучшений, предложенных в этой работе.

БИБЛИОГРАФИЯ

- [1] L. A. Zadeh, "Fuzzy Sets," *Information and control*, vol. 8, pp. 338–353, 1965.
- [2] Q. Song and B. S. Chissom, "Fuzzy time series and its models," *Fuzzy Sets Syst*, vol. 54, pp. 269–277, 1993.
- [3] T. T. D. Nguyen and L. V. Chernenkaya, "Forecasting model of intuitionistic fuzzy time series using ratio distribution," *International Journal of Open Information Technologies*, vol. 11, no. 11, pp. 35–44, 2023.
- [4] Т. Т. З. Нгуен and Л. В. Черненкокая, "Фаззификация в моделях прогнозирования нечетких временных рядов," *Журнал Известия Тульского государственного университета – Технические науки (ТулГУ, г. Тула)*, vol. 8, no. Системный анализ, Управление и обработка информации, pp. 337–346, 2023.
- [5] Нгуен Тхи Тху Зунг and Л. В. Черненкокая, "Дискретизация в моделях прогнозирования нечетких временных рядов," *Журнал Известия Тульского государственного университета – Технические науки (ТулГУ, г. Тула)*, vol. 8, no. Системный анализ, Управление и обработка информации, pp. 296–304, 2023, doi: 10.24412/2071-6168-2023-8-296-297.
- [6] K. Huarng and H. K. Yu, "A type 2 fuzzy time series model for stock index forecasting," *Physica A: Statistical Mechanics and its Applications*, vol. 353, no. 1–4, pp. 445–462, 2005, doi: 10.1016/j.physa.2004.11.070.
- [7] O. Castillo and P. Melin, "Comparison of hybrid intelligent systems, neural networks and interval type-2 fuzzy logic for time series prediction," *IEEE International Conference on Neural Networks - Conference Proceedings*, pp. 3086–3091, 2007, doi: 10.1109/IJCNN.2007.4371453.
- [8] N. S. Bajestani and A. Zare, "Forecasting TAIEX using improved type 2 fuzzy time series," *Expert Syst Appl*, vol. 38, no. 5, pp. 5816–5821, 2011, doi: 10.1016/j.eswa.2010.10.049.
- [9] F. Gaxiola, P. Melin, F. Valdez, and O. Castillo, "Interval type-2 fuzzy weight adjustment for backpropagation neural networks with application in time series prediction," *Inf Sci (N Y)*, vol. 260, pp. 1–14, 2014, doi: 10.1016/j.ins.2013.11.006.
- [10] O. Castillo, J. R. Castro, P. Melin, and A. Rodriguez-Diaz, "Application of interval type-2 fuzzy neural networks in non-linear identification and time series prediction," *Soft comput*, vol. 18, no. 6, pp. 1213–1224, 2014, doi: 10.1007/s00500-013-1139-y.

- [11] Abhishekh, S. S. Gautam, and S. R. Singh, “A refined weighted method for forecasting based on type 2 fuzzy time series,” *International Journal of Modelling and Simulation*, vol. 38, no. 3, pp. 180–188, 2018, doi: 10.1080/02286203.2017.1408948.
- [12] J. A. Jiang, C. H. Syue, C. H. Wang, J. C. Wang, and J. S. Shieh, “An Interval Type-2 Fuzzy Logic System for Stock Index Forecasting Based on Fuzzy Time Series and a Fuzzy Logical Relationship Map,” *IEEE Access*, vol. 6, pp. 69107–69119, 2018, doi: 10.1109/ACCESS.2018.2879962.
- [13] S. S. Pal and S. Kar, “A Hybridized Forecasting Method Based on Weight Adjustment of Neural Network Using Generalized Type-2 Fuzzy Set,” *International Journal of Fuzzy Systems*, vol. 21, no. 1, pp. 308–320, 2019, doi: 10.1007/s40815-018-0534-z.
- [14] N. F. Rahim, M. Othman, R. Sokkalingam, and E. A. Kadir, “Type 2 fuzzy inference-based time series model,” *MDPI, Symmetry*, vol. 11, no. 11, pp. 1–13, 2019, doi: 10.3390/sym11111340.
- [15] A. C. V. Pinto, T. E. Fernandes, P. C. L. Silva, F. G. Guimarães, C. Wagner, and E. Pestana de Aguiar, “Interval type-2 fuzzy set-based time series forecasting using a data-driven partitioning approach,” *Evolving Systems*, vol. 13, no. 5, pp. 703–721, 2022, doi: 10.1007/s12530-022-09452-2.
- [16] Y. Yin, Y. Sheng, and J. Qin, “Interval type-2 fuzzy C-means forecasting model for fuzzy time series,” *Appl Soft Comput*, vol. 129, p. 109574, 2022, doi: 10.1016/j.asoc.2022.109574.
- [17] D. N. Thi Thu and L. V. Chernenkaya, “A Forecasting Model Intuitionistic Fuzzy Time Series Using Distribution Ratio-Based,” in *2023 International Russian Automation Conference (RusAutoCon)*, Sochi, Russian Federation: IEEE, Sep. 2023, pp. 392–397. doi: 10.1109/RusAutoCon58002.2023.10272755.
- [18] D. N. T. Thu and L. V. Chernenkaya, “A High-Order Heuristic Fuzzy Time Series Forecasting Model Based on Hedge Algebras Approach,” in *2023 International Russian Automation Conference (RusAutoCon)*, Sochi, Russian Federation: IEEE, Sep. 2023, pp. 721–728. doi: 10.1109/RusAutoCon58002.2023.10272750.
- [19] Васильев Б. Ю., Нгуен Т. Х. Анализ влияния полупроводниковых преобразователей на батарею и двигатель асинхронного привода шахтных горнотранспортных машин // Горный информационно-аналитический бюллетень. – 2023. – No 9-1. – С. 299–318. DOI: 10.25018/0236_1493_2023_91_0_299.

A forecasting model fuzzy time series type 2 with hedge algebraic and general optimization algorithm

Nguyen Thi Thu Dung, L. V. Chernenkaya

Abstract—In order to keep with the evolution of socio-economic problems, the development of forecasting models increasingly needs improvement. Existing fuzzy time series (FTS) forecasting models are based on the first type fuzzy logic theory, but the second type fuzzy logic theory shows greater coverage and more accurate modeling of reality in many cases. This is suitable because in reality, the degree of membership of an element to a set cannot be determined specifically, but only within a range. In this paper, a fuzzy time series forecasting model is proposed based on type two fuzzy logic theory and Hedge algebra structure. The parameters of the proposed model are optimized using genetic algorithm. The proposed model is tested by forecasting the daily values of TAIEX data and the forecasting performance is evaluated by RMSE, MAPE and MSE metrics.

Keywords—Type II fuzzy logic, type II fuzzy time series forecasting model, hedge algebra, genetic algorithm, forecasting model optimization, fuzzification, type II defuzzification.

REFERENCES

- [1] L. A. Zadeh, "Fuzzy Sets," *Information and control*, vol. 8, pp. 338–353, 1965.
- [2] Q. Song and B. S. Chissom, "Fuzzy time series and its models," *Fuzzy Sets Syst*, vol. 54, pp. 269–277, 1993.
- [3] T. T. D. Nguyen and L. V. Chernenkaya, "Forecasting model of intuitionistic fuzzy time series using ratio distribution," *International Journal of Open Information Technologies*, vol. 11, no. 11, pp. 35–44, 2023.
- [4] T. T. D. Nguyen and L. V. Chernenkaya, "Fuzzification in forecasting models of fuzzy time series," *Journal of Tula State University - Technical Sciences (Tula State University, Tula)*, vol. 8, System analysis, Management and information processing, pp. 337–346, 2023.
- [5] Nguyen Thi Thu Dung and L.V. Chernenkaya, "Discretization in forecasting models of fuzzy time series," *Journal of Tula State University - Technical Sciences (Tula State University, Tula)*, vol. 8, no. System analysis, Management and information processing, pp. 296–304, 2023, doi: 10.24412/2071-6168-2023-8-296-297.
- [6] K. Huarng and H. K. Yu, "A type 2 fuzzy time series model for stock index forecasting," *Physica A: Statistical Mechanics and its Applications*, vol. 353, no. 1–4, pp. 445–462, 2005, doi: 10.1016/j.physa.2004.11.070.
- [7] O. Castillo and P. Melin, "Comparison of hybrid intelligent systems, neural networks and interval type-2 fuzzy logic for time series prediction," *IEEE International Conference on Neural Networks - Conference Proceedings*, pp. 3086–3091, 2007, doi: 10.1109/IJCNN.2007.4371453.
- [8] N. S. Bajestani and A. Zare, "Forecasting TAIEX using improved type 2 fuzzy time series," *Expert Syst Appl*, vol. 38, no. 5, pp. 5816–5821, 2011, doi: 10.1016/j.eswa.2010.10.049.
- [9] F. Gaxiola, P. Melin, F. Valdez, and O. Castillo, "Interval type-2 fuzzy weight adjustment for backpropagation neural networks with application in time series prediction," *Inf Sci (NY)*, vol. 260, pp. 1–14, 2014, doi: 10.1016/j.ins.2013.11.006.
- [10] O. Castillo, J. R. Castro, P. Melin, and A. Rodriguez-Diaz, "Application of interval type-2 fuzzy neural networks in non-linear identification and time series prediction," *Soft comput*, vol. 18, no. 6, pp. 1213–1224, 2014, doi: 10.1007/s00500-013-1139-y.
- [11] Abhishekh, S. S. Gautam, and S. R. Singh, "A refined weighted method for forecasting based on type 2 fuzzy time series," *International Journal of Modelling and Simulation*, vol. 38, no. 3, pp. 180–188, 2018, doi: 10.1080/02286203.2017.1408948.
- [12] J. A. Jiang, C. H. Syue, C. H. Wang, J. C. Wang, and J. S. Shieh, "An Interval Type-2 Fuzzy Logic System for Stock Index Forecasting Based on Fuzzy Time Series and a Fuzzy Logical Relationship Map," *IEEE Access*, vol. 6, pp. 69107–69119, 2018, doi: 10.1109/ACCESS.2018.2879962.
- [13] S. S. Pal and S. Kar, "A Hybridized Forecasting Method Based on Weight Adjustment of Neural Network Using Generalized Type-2 Fuzzy Set," *International Journal of Fuzzy Systems*, vol. 21, no. 1, pp. 308–320, 2019, doi: 10.1007/s40815-018-0534-z.
- [14] N. F. Rahim, M. Othman, R. Sokkalingam, and E. A. Kadir, "Type 2 fuzzy inference-based time series model," *MDPI, Symmetry*, vol. 11, no. 11, pp. 1–13, 2019, doi: 10.3390/sym11111340.
- [15] A. C. V. Pinto, T. E. Fernandes, P. C. L. Silva, F. G. Guimarães, C. Wagner, and E. Pestana de Aguiar, "Interval type-2 fuzzy set-based time series forecasting using a data-driven partitioning approach," *Evolving Systems*, vol. 13, no. 5, pp. 703–721, 2022, doi: 10.1007/s12530-022-09452-2.
- [16] Y. Yin, Y. Sheng, and J. Qin, "Interval type-2 fuzzy C-means forecasting model for fuzzy time series," *Appl Soft Comput*, vol. 129, p. 109574, 2022, doi: 10.1016/j.asoc.2022.109574.
- [17] D. N. Thi Thu and L. V. Chernenkaya, "A Forecasting Model Intuitionistic Fuzzy Time Series Using Distribution Ratio-Based," in *2023 International Russian Automation Conference (RusAutoCon)*, Sochi, Russian Federation: IEEE, Sep. 2023, pp. 392–397. doi: 10.1109/RusAutoCon58002.2023.10272755.
- [18] D. N. T. Thu and L. V. Chernenkaya, "A High-Order Heuristic Fuzzy Time Series Forecasting Model Based on Hedge Algebras Approach," in *2023 International*

Russian Automation Conference (RusAutoCon), Sochi, Russian Federation: IEEE, Sep. 2023, pp. 721–728. doi: 10.1109/RusAutoCon58002.2023.10272750.

- [19] Vasilev B. U., Nguyen T. H. Influence of semiconductor converters on asyn-chronous drive battery and motor in mining machines. *MIAB. Mining Inf. Anal. Bull.* 2023; (9-1):299-318. [In Russ]. DOI: 10.25018/0236_1493_2023_91_0_299