

*Нгуен Тхи Тху Зунг*¹,

аспирант;

*Черненко Людмила Васильевна*²,

профессор, д-р техн. наук, ст. науч. сотр.

ЭВРИСТИЧЕСКАЯ НЕЧЕТКАЯ МОДЕЛЬ ПРОГНОЗИРОВАНИЯ ВЫСОКОГО ПОРЯДКА С ХЕДЖ-АЛГЕБРОЙ

^{1,2} Россия, Санкт-Петербург, Санкт-Петербургский политехнический
университет Петра Великого;

¹ thudung.mta.tb@gmail.com, ² ludmila@qmd.spbstu.ru

Аннотация. В статье предлагается новая эвристическая модель для нечеткого прогнозирования временных рядов с использованием нечетких отношений высокого порядка в сочетании с подходом хедж-алгебры. Модель была протестирована на практических данных. Прогнозируемые результаты оценены и сравнены с существующими моделями. Анализ показал, что предлагаемая модель имеет более высокую эффективность.

Ключевые слова: прогнозирование временных рядов, модели прогнозирования, нечеткие модели высокого порядка, эвристические модели, хедж-алгебра, нечеткие группы отношений, лингвистические переменные, фазсификация, дефазсификация.

*Nguyen Thi Thu Dung*¹,

Postgraduate (PhD) Student;

*Liudmila V. Chernenkaya*²,

Doctor of Technical Science, Professor

HIGH-ORDER HEURISTIC FUZZY FORECASTING MODEL WITH HEDGE ALGEBRA

^{1,2} Peter the Great St. Petersburg Polytechnic University, St. Petersburg, Russia;

¹ thudung.mta.tb@gmail.com, ² ludmila@qmd.spbstu.ru

Abstract. This paper proposes a new heuristic model for fuzzy time series forecasting using high-order fuzzy relations combined with a hedge algebra approach. The model was tested on practical data. The forecasted results are evaluated and compared with existing models and it is shown that the proposed model has higher performance.

Keywords: time series forecasting, forecasting models, higher order fuzzy models, heuristic models, hedge algebra, fuzzy relation groups, linguistic variables, fuzzification, defuzzification.

Введение

Современное состояние социально-экономических проблем требует, чтобы совершенствовалась разработка моделей прогнозирования. Одной из моделей прогнозирования, которая, как считается, имеет хорошую

производительность прогнозирования, является модель прогнозирования нечетких временных рядов. Суть модели прогнозирования нечетких временных рядов заключается в построении нечетких связей между наблюдениями в разные рассматриваемые моменты времени на основе правила: «наблюдения в определенный момент времени — это накопленные результаты наблюдений в предыдущие моменты времени». Базовая модель прогнозирования нечетких временных рядов состоит из таких основных частей, как: дискретизация временных рядов, фаззификация элементов, построение нечетких отношений и дефаззификация для получения результатов прогнозирования (см. рис 1).

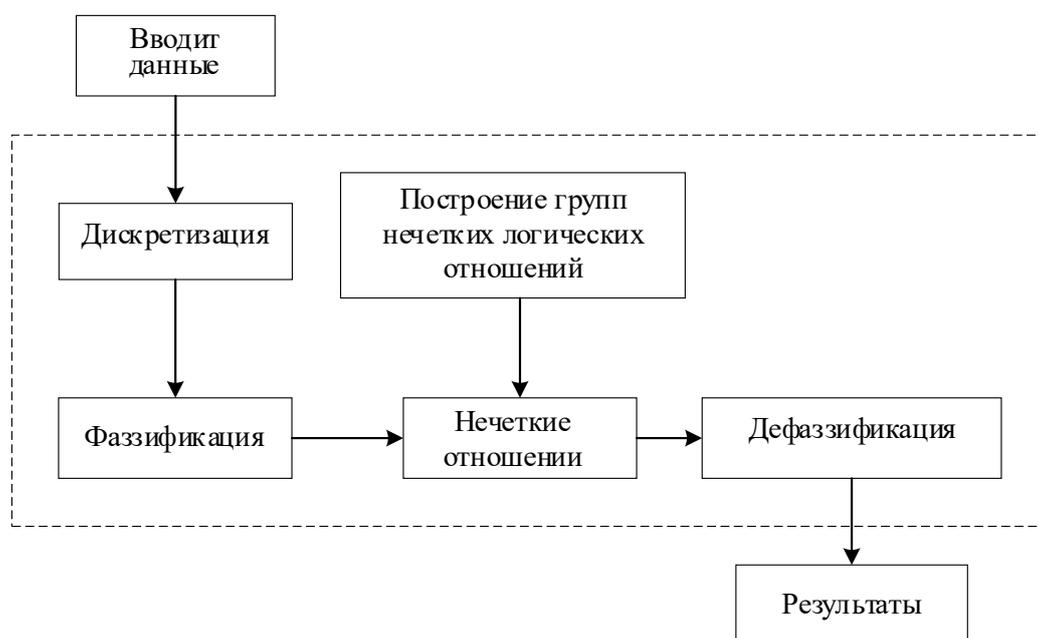


Рис. 1. Процесс внедрения предлагаемой модели

В настоящее время модели прогнозирования нечетких временных рядов все больше совершенствуются, чтобы обеспечить наилучшие показатели прогнозирования. Усовершенствования в основном направлены на повышение производительности каждой части модели. Доказано, что построение нечетких отношений высокого порядка обеспечивает лучшую производительность прогнозирования. Модель нечетких временных рядов высокого порядка определяется следующим образом: пусть $F(t)$ — нечеткий временной ряд. Если $F(t)$ выбрано из $F(t-1), F(t-2), \dots, F(t-n)$, то нечеткое логическое отношение выражено следующим образом: $F(t-n), \dots, F(t-2), F(t-1) \rightarrow F(t)$, и представляет собой модель прогнозирования нечетких временных рядов n -го порядка. Эвристическая модель прогнозирования находит широкое применение

в прогностических моделях на практике. Авторы показывают, что эвристическая модель дает лучше результаты по сравнению с другими моделями. Кроме того, эвристическая модель лишена некоторых недостатков классической модели, таких как: сильная зависимость от нечетких данных высшего порядка и интенсивное использование памяти.

С другой стороны, одной из характеристик нечеткой теории является то, что лингвистические переменные связаны с нечеткими множествами. Одна из проблем здесь заключается в том, что еще не определены соотношения между нечеткими множествами и соответствующими лингвистическими переменными, поэтому целесообразно применить подход хедж-алгебры. Хедж-алгебраическая структура — это новый подход, разработанный авторами Н. Хо и В. Вехлер. Это семантическая количественная алгебраическая структура лингвистических переменных, обозначаемая $AH = (X, G, C, H, \leq)$, где $G = \{c^+, c^-\}$ - множество первичных образующих, c^+ и c^- - отрицательный и положительный первичный термин лингвистической переменной X ; при $C = \{0, W, 1\}$ представляет собой набор констант, которые различаются элементами в X ; H - это набор хеджа, где $H^- = h_{-1}, h_{-2}, \dots, h_{-q}$ и $H^+ = h_{+1}, h_{+2}, \dots, h_{+p}$; где $h_{-1} < h_{-2} < \dots < h_{-q}$ и $h_{+1} < h_{+2} < \dots < h_{+p}$; " \leq " является семантически упорядочивающим отношением на X . Эксплуатация хедж-алгебраического подхода в модели прогнозирования нечетких временных рядов дает хорошие результаты по сравнению с обычными моделями прогнозирования нечетких временных рядов.

В этом исследовании предлагается эвристическая модель прогнозирования нечетких временных рядов высокого порядка, основанная на подходе алгебраического хеджа, при котором в предложенной модели применяются отношения эвристической нечеткой логики высокого порядка и анализируется теория эвристической алгебры для определения семантики нечетких множеств в фаззификации, а затем используется теория эвристической алгебры для дефаззификации.

2. Построение модели

Предлагаемая модель использует процесс семантизации и десемантизации для выполнения фаззификации и дефаззификации в модели прогнозирования нечетких временных рядов сочетания с эвристическими группами нечетких отношений высокого порядка. Фаззификация и дефаззификация тесно связаны с процессом семантизации и десемантизации.

Разработанная структура и последовательность выполнения предлагаемой модели представлена на рисунке 2.



Рис. 2. Процесс внедрения предлагаемой модели

2. Процесс внедрения предлагаемой модели

Алгоритм предложенной модели реализуется в соответствии со следующими этапами:

Шаг 1. Определить универсум дискурса и интервалов. Пусть универсум дискурса $U = D_{\min} - D_1, D_{\max} + D_2$, где D_{\min} и D_{\max} — минимальное и максимальное значение исторических данных соответственно, где D_1 и D_2 — два собственных положительных действительных значения. Универсум дискурса U в этом шаге разделен на n равные промежутки: $U = u_i, i = 1, n$.

Шаг 2. Фаззификация и семантизация.

Шаг 2.1. Определить лингвистические термины A_1, A_2, \dots, A_n . Требуется преобразование к структуре хеджа алгебраической: $A_1 = hA_1c, A_2 = hA_2c, \dots, A_n = hA_nc$, где hA_i представляет собой хедж-ряд, действующий c с $c = c^-, c^+$.

Шаг 2.2. Количественная семантика лингвистических переменных. Семантическая квантификация лингвистических переменных зависит от параметров α и θ , тогда: $SA_1 < SA_2 < \dots < SA_n$, где $SA = Semantization A$ — количественное семантическое значение по смысловой метке A .

Семантические отношения по смысловым меткам соответствуют нечетким отношениям:

$$SA_k \rightarrow SA_m \text{ или } Semantization(A_k) \rightarrow Semantization(A_m)$$

Шаг 3. Построить нечеткие группы отношений высокого порядка. На этом шаге необходимо выбрать максимальный порядок нечетких отношений w .

Шаг 3.1. Понадобится построить нечеткие отношения первого порядка, второго порядка, ..., w -порядка, а затем перегруппировать в группы нечетких отношений.

Шаг 3.2. Фаза обучения эвристической модели.

Сначала необходимо выбрать данные для обучения. Для обучения на первый момент временного ряда выбираются порядки 1 и 2 и значение флага $n = 1$. Выполняются принципы эвристической модели Бая для фазы обучения.

Шаг 3.3. Фаза тестирования эвристической модели.

На основании результатов шага 3.2 выполняются принципы эвристической модели Бая, чтобы найти подходящие отношения для каждого объекта прогнозирования на этапе тестирования.

Шаг 4. Дефаззификация и десемантизация.

Прогнозные значения рассчитываются на основе последствий нелинейной десемантизации. Выбраны значения параметра семантизации sp и параметра десемантизации dp .

3. Результат эксперимента и оценка эффективности модели

Прогнозируемые значения по предлагаемой модели и фактические значения представлены на линейном графике (рис. 3). Видно, что фактические значения и прогнозируемые значения довольно близко соответствуют, а это означает, что эффективность прогнозирования модели достаточно высокая.

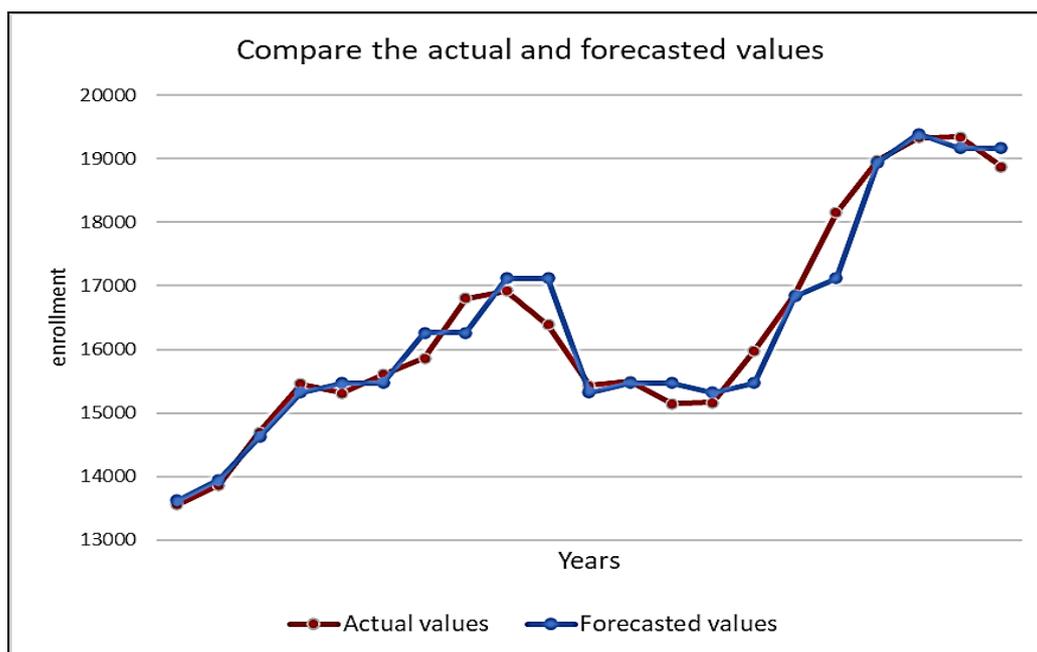


Рис. 3. Процесс внедрения предлагаемой модели

Для уточнения эффективности прогноза необходимо использовать методы исследования, точность которых определяется разностными методами. Существует несколько способов решения этой проблемы. В этой предлагаемой модели используется RMSE — метрика, симметричная относительно знака ошибки, чувствительная к большим отклонениям от среднего значения фактической величины продажи, которая определяется следующей формулой:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_i)^2}$$

Метрика RMSE предлагаемого метода составляет 424,3.

Таблица 1

Сравнение результатов

Метод	RMSE
Метод Сонга и Чиссома (1993)	642,3
Метод Сонга и Чиссома (1994)	880,7
Метод Салливан и Вудалл (1994)	621,3
Метод Чена (1996)	638,7
Метод Ченга, Ченга, Хоанга (2008)	478,5
Метод Ченга, Чена, Тэоха, Чианга (2008)	438,3
Метод Хвана, Ченга, Лея (1998)	528,1
Метод Джилани (2007)	476,6

Проведено сравнение результатов предложенной модели прогнозирования с существующими моделями (см. табл. 1). Анализ показывает, что предлагаемая модель прогнозирования имеет более высокую эффективность.

Заключение

В данной работе рассмотрено построение новой нечеткой эвристической модели для прогнозирования временных рядов высокого порядка с использованием подхода хедж-алгебры, который поддерживает семантическую количественную оценку лингвистических переменных, увеличивая связность между лингвистическими переменными и значениями нечетких множеств. Показано применение предложенной модели на реальных данных. Проведено сравнение результатов прогнозирования с целевыми данными, которое показало довольно хорошую эффективность предложенной модели прогнозирования. Также проведено сравнение предложенной модели с существующими моделями прогнозирования с использованием показателя RMSE, которое показало, что предлагаемая модель имеет высокую производительность. Эта статья является основой для разработки будущих моделей прогнозирования временных рядов.

Список литературы

1. Нгуен Т.Т.З., Черненькая Л.В. Системный анализ в управлении развитием территориальных комплексов Вьетнама // В сборнике: Системный анализ в проектировании и управлении. Сборник научных трудов XXV Международной научной и учебно-практической конференции: в 3 ч. Санкт-Петербург, 2021. – С. 346–352.
2. Нгуен Тхи Тху Зунг, Черненькая Л.В. Дискретизация в моделях прогнозирования нечетких временных рядов // Журнал Известия Тульского государственного университета – Технические науки (ТулГУ, г. Тула), Выпуск 8, Системный анализ, управление и обработка информации. – 2023. – С. 296–304.
3. Нгуен Тхи Тху Зунг, Черненькая Л. В. Фазсификация в моделях прогнозирования нечетких временных рядов // Журнал Известия Тульского государственного университета – Технические науки (ТулГУ, г. Тула), Выпуск 8, Системный анализ, Управление и обработка информации. – 2023. – С. 337–346.
4. Zadeh L.A. Fuzzy Sets // Information and control. – 1965. – Vol. 8. – Pp. 338–353.
5. Нгуен Тхи Тху Зунг, Черненькая Л.В. Модель анализа факторов на основе нечеткой кластеризации с-средних // Журнал Известия Тульского государственного университета – Технические науки (ТулГУ, г. Тула). – 2023.
6. Ho N.C., Wechler W. Extended hedge algebras and their application to fuzzy logic // Fuzzy Sets and Systems. – 1992. – Vol. 52. – Pp. 259–281.
7. Cat N.H., Wechler W. Hedge algebras: an algebraic approach to structures of sets of linguistic domains of linguistic truth variable // Fuzzy Sets and Systems. – 1990. – Vol. 35. – Pp. 281–293.
8. Bai E., Wong W.K., Chu W.C., Xia M., Pan F. A heuristic time-invariant model for fuzzy time series forecasting // Expert Syst Appl. – 2011. – Vol. 38, no. 3 (Mar. 2011). – Pp. 2701–2707. – DOI: 10.1016/j.eswa.2010.08.059.