

УДК 001.891.573

Т.В. Афанасьева, М.С. Тонерян**ПРИМЕНЕНИЕ НЕЧЕТКОГО МОДЕЛИРОВАНИЯ
К ПРОГНОЗИРОВАНИЮ КОЛИЧЕСТВА АБИТУРИЕНТОВ**

Работа посвящена моделям нечетких процессов для прогнозирования количества абитуриентов. Данная проблема актуальна и имеет важное значение, так как от ее адекватного решения зависит объем финансирования вуза. Существует много подходов прогнозирования количества абитуриентов. Однако, временные ряды, содержащие сведения о поступающих в вуз, как правило, характеризуются небольшой длиной, нестационарностью поведения, что затрудняет построение статистических моделей и моделей на основе искусственных нейронных сетей для целей прогноза. Также, традиционные модели не могут быть применены, когда исторические данные являются лингвистическими значениями. Нечеткие временные ряды являются эффективным инструментом для решения таких проблем. В качестве приложения нечетких временных рядов приводится пример прогнозирования абитуриентов из Университета Алабамы.

Нечеткие временные ряды; прогнозирование; нечеткие процессы с нечеткими приращениями.

T.V. Afanasjeva, M.S. Toneryan**APPLICATION OF FUZZY MODEL TO FORECASTING UNIVERSITY
OF APPLICANTS**

The paper is devoted to models of fuzzy processes for forecasting of number of entrants. This problem is actual and is important, as the university financing depends on its adequate decision. There have been a good many methods to forecast university enrollments in the literature. However, the enrollment's time series, as a usually, are short and nonstationary, that complicates creation of statistical models and models on the basis of artificial neural networks. Also, traditional models can't be applied, when historical data are linguistic values. Fuzzy time series is an effective tool to deal with such problems. In this paper, as an application of fuzzy time series in educational research, the forecast of the enrollments of the University of Alabama is carried out.

Fuzzy time series; forecasting; fuzzy processes with fuzzy differences.

Введение. Очень важно сделать достаточно точные оценки количества поступающих в университеты, так как многие решения в университете будут зависеть от них. Когда исторические данные о количестве поступающих состоят из лингвистических значений, представленных в условиях недостаточности данных, например, экспертами, единственным формальным инструментом являются нечеткие временные ряды (НВР), модели которых для описания нечетких процессов предложены в работе [1] для прогнозирования количества учащихся университета Алабамы и развиты в работах, обзор которых приведен в [2]. В приложениях НВР есть две возможности: когда исторические данные находятся в лингвистической форме и когда исторические данные представлены в числовых значениях. Во втором случае данные сначала должны быть фаззифицированы. В нашем исследовании исторические данные образуют числовой ВР из 22 наблюдений, который в дальнейшем будем обозначать для краткости «Алабама».

Алгоритм прогнозирования количества абитуриентов университета на основе модели нечеткого временного ряда

Шаг 1. Определим универсальное множество U , в пределах которого изменяются исторические данные и на которых нечеткие множества будут определены ($U = [13\ 000, 20\ 000]$).

Шаг 2. Разделим универсум U на 7 интервалов с равными длинами.

Шаг 3. Определим нечеткие множества на универсуме U с треугольными функциями принадлежности A_i $i=(1...7)$.

Шаг 4. Фаззифицируем исторические данные.

Шаг 5. По историческим данным построим модель НВР

$$A_i = A_{i-1} \circ R, \text{ где } R(t, t-1) = R = \bigcup_{i=1}^{10} R_i, \quad (1)$$

$$R_1 = A_1^T \times A_1, R_2 = A_1^T \times A_2, R_3 = A_2^T \times A_3, R_4 = A_3^T \times A_3, R_5 = A_3^T \times A_4,$$

$$R_6 = A_4^T \times A_4, R_7 = A_4^T \times A_3, R_8 = A_4^T \times A_6, R_9 = A_6^T \times A_6, R_{10} = A_6^T \times A_7,$$

$$A_1 \rightarrow A_1, A_1 \rightarrow A_2, A_2 \rightarrow A_3, A_3 \rightarrow A_3, A_3 \rightarrow A_4,$$

$$A_4 \rightarrow A_4, A_4 \rightarrow A_3, A_4 \rightarrow A_6, A_6 \rightarrow A_6, A_6 \rightarrow A_7,$$

где A_{i-1} является регистрацией года $i-1$ и A_i предсказанная регистрация года i , $R(t, t-1)$ – система нечетких отношений, ' \circ ' – операция *max-min* композиции.

Шаг 6. Вычисляем прогнозируемые результаты. Предположим, что регистрация года t известна, чтобы предсказать регистрацию года $t+1$. Пусть A_{i-1} в (1) регистрация в год t , и применим формулу (1). Тогда A_i и будет прогнозированной регистрацией года $t+1$. Для 1972–1991 гг. результаты показаны на рис. 1.

Шаг 7. Дефаззифицируем полученные результаты. Подробное изложение моделей НВР и процедуры прогнозирования ВР «Алабама» приведено в [1]. Ошибки прогноза по критерию MAPE колеблются от 0,1 до 8,7 % со средней ошибкой, составляющей 3,18 %. На 1991 г. ошибка прогноза составляет 1,7 %. Для временной прогнозируемой модели средняя ошибка в 3,18 % является довольно удовлетворительной.

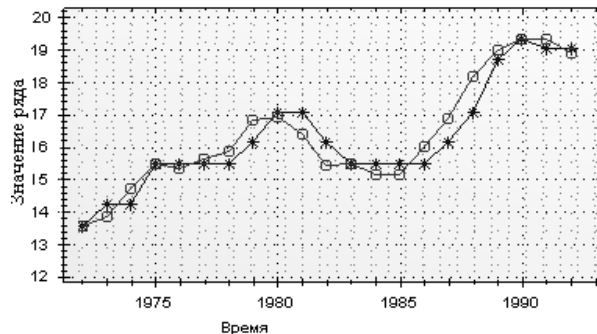


Рис. 1. Прогнозирование ВР «Алабама» на основе модели НВР
(\circ – фактическая регистрация, * – предсказанная регистрация)

Алгоритм прогнозирования количества абитуриентов университета на основе модели нечеткого процесса с нечеткими приращениями. Пусть $X_t, (t=1,2,...) \subset R^1$ – универсальное множество, на котором определены нечеткие множества $\tilde{x}_t^i, (i=1,2,...), \tilde{v}_t^j, (j=1,2,...), \tilde{\alpha}_t^s, (s=1,2,...)$, и \tilde{X}_t – коллекция $\tilde{x}_t^i, (i=1,2,...), \tilde{V}_t$ – коллекция $\tilde{v}_t^j, (j=1,2,...), \tilde{A}_t$ – коллекция $\tilde{\alpha}_t^s, (s=1,2,...)$. И пусть существуют отношения $R_v: \tilde{X} \times \tilde{X} \rightarrow \tilde{V}, R_A: \tilde{X} \times \tilde{X} \rightarrow \tilde{A}$, тогда определим модель нечеткого процесса с нечеткими приращениями (*T-модель*) в виде

$$\tilde{X}_t = (\tilde{X}_{t-1} \times \tilde{V}_t \times \tilde{A}_t) \circ R(t, t-1), \quad (3)$$

где

$$\tilde{V}_t = \tilde{V}_{t-1} \times \tilde{V}_{t-2} \times \dots \times \tilde{V}_{t-p} \circ R_v(t, t-p), \quad \tilde{A}_t = \tilde{A}_{t-1} \times \tilde{A}_{t-2} \times \dots \times \tilde{A}_{t-q} \circ R_a(t, t-q).$$

Здесь система отношений $R_v(t, t-p)$, $R_a(t, t-q)$ определяет модель поведения нечетких приращений, получивших название нечетких элементарных тенденций [3]. Алгоритм прогнозирования количества абитуриентов на основе модели нечеткого процесса с нечеткими приращениями включает адаптацию алгоритма, рассмотренного выше и представленного в работе [1].

Шаги с 1 по 4 выполняются аналогично.

Шаг 5. По историческим данным построим модель (3).

Шаг 6. Применим модель (3) для прогноза значений (рис. 2).

Шаг 7. Дефазифицируем полученные результаты. Помимо прогноза значений в этом алгоритме идентифицирована общая нечеткая тенденция «Рост», а также более корректно спрогнозированы нечеткие элементарные тенденции (нечеткие приращения).

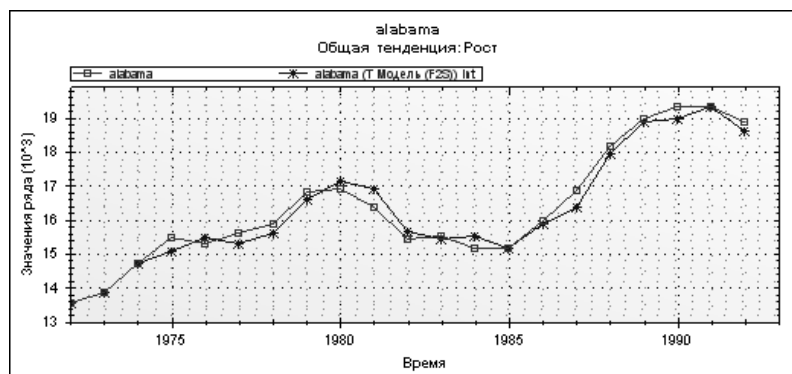


Рис. 2. Прогнозирование ВР «Алабама» на основе T-модели
(о – фактическая регистрация, * – предсказанная регистрация)

В табл. 1 приводятся сравнительные результаты тестирования моделей – нечетких и статистических – по внешним ошибкам для краткосрочного прогноза ВР «Алабама» (на 1 интервал). Анализ табл. 1 показывает, что нечеткие модели по сравнению с базовыми статистическими моделями при краткосрочном прогнозе короткого нестационарного ВР «Алабама» демонстрируют потенциал и конкурентоспособность.

Таблица 1

Сравнение моделей для прогноза ВР «Алабама» на 1 интервал

Модель	MAPE %
T-модель [3]	0,4
Jilani&Bruneu,2007 [5]	1,0
Chen, 2002 [4]	1,5
ARIMA (0,1,1)(0,1,2)	2,1
S-модель [1]	2,4
Метод Брауна	4,2

Заключение. В настоящей статье были рассмотрены алгоритмы прогнозирования количества абитуриентов в Университете штата Алабама на основе модели нечеткого ВР и на основе нечеткого процесса с нечеткими приращениями. Конечно, есть серьезная проблема с этими методами: прогнозные значения во многом зависят от интерпретации результатов модели прогнозирования в шаге 7. Различные методы дефаззификации могут привести к различным прогнозируемым результатам. Это делает процесс довольно субъективным. Чтобы преодолеть этот недостаток, должен быть применен объективный метод. Несмотря на проблемы, преимущества делают нечеткие модели ВР очень конкурентоспособными в прогнозировании, в частности количества абитуриентов университетов. Это также означает, что теории нечетких множеств являются успешным инструментом для решения сложных практических задач.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Song Q., Chissom B.* Fuzzy time series and its models // *Fuzzy Sets and Systems*. – 1993. – № 54. – P. 269-277.
2. *Афанасьева Т.В., Ярушклина Н.Г.* Нечеткое моделирование временных рядов и анализ нечетких тенденций. – Ульяновск: УлГТУ, 2009. – 299 с.
3. *Афанасьева Т.В., Ярушклина Н.Г.* Нечеткий динамический процесс с нечеткими тенденциями в анализе временных рядов // *Вестник РГУПС*. – 2011. – № 3. – С. 6-15.
4. *Chen S.M.* A new method to forecast enrollments using fuzzy time series // *International Journal of Applied Sciences and Engineering*. – 2004. – № 2 (3). – P. 234-244.
5. *Jilani T.A.* Fuzzy Metric Approach for Fuzzy Time Series Forecasting based on Frequency Density Based Partitioning / *Tahseen Ahmed Jilani, Syed Muhammad Aqil Burney, Cemal Ardil // Proceedings of World Academy of Science, Engineering and Technology*. – 2007. – Vol. 23.

Статью рекомендовал к опубликованию д.т.н., профессор А.А. Стецко.

Афанасьева Татьяна Васильевна – Федеральное государственное бюджетное образовательное учреждение высшего профессионального образования «Ульяновский государственный технический университет»; e-mail: tv.afanasjeva@gmail.com; 432027, г. Ульяновск, ул. Северный Венец, 32; тел.: +79063918568; кафедра информационных систем; к.т.н.; доцент.

Тонерян Мкртыч Саркисович – e-mail: mkr73@yandex.ru; тел.: +79084739746; кафедра прикладной математики и информатики; ассистент; аспирант.

Afanasjeva Tatyana Vasilievna – Federal State Budget Institution of Higher Professional Education; e-mail: tv.afanasjeva@gmail.com; 32, Severny Venetz street, Ulyanovsk, 432027, Russia; phone: +79063918568; the department of information systems; cand. of eng. sc.; associate professor.

Toneryan Mkrtych Sarkisovich – e-mail: mkr73@yandex.ru; phone: +79084739746; the department of the higher mathematics; assistant; postgraduate student.