

УДК 681.3.015

Н. Р. Кондратенко, к. т. н., доц.;

О. В. Чеборака, асп.

НЕЧІТКІ МНОЖИНИ В ЗАДАЧАХ ПРОГНОЗУВАННЯ ЧАСОВИХ ПОСЛІДОВНОСТЕЙ

Розглянуто використання інтервальних функцій належності в задачах прогнозування часових послідовностей. Розроблено метод побудови інтервальних нечітких моделей прогнозування часових послідовностей. Побудовано інтервальну нечітку модель прогнозування хаотичної часової послідовності Макей—Гласа, яка продемонструвала високу якість прийняття рішень.

Вступ

Сучасний напрямок використання прогресивних інформаційних технологій передбачає створення систем підтримки прийняття рішень, ключовим моментом яких є задачі прогнозування часових послідовностей, які виникають в різних сферах людської діяльності (наука, техніка, економіка, медицина, сільське господарство, промисловість та ін.). Прогнозування є важливим елементом організації управління як окремими господарюючими суб'єктами, так і економікою в цілому. Можливість прогнозувати некеровані аспекти подій перед прийняттям кінцевого рішення дозволяє зробити найкращий вибір. Однак дуже часто необхідно приймати рішення, маючи суперечливі дані, що обумовлено присутністю певного рівня «шуму». Для виходу з подібної складної ситуації використовуються інтервальні нечіткі множини, коли в якості значень функцій належності виступає не число, а інтервал. Це дозволяє розширити можливість опису предметної галузі.

Нечітка модель з використанням інтервальних функцій належності

В побудові інтервальної нечіткої моделі прогнозування часових послідовностей вихідною інформацією є зашумлена послідовність даних $x(t) = s(t) + n(t)$, де $x(t)$, $t = \overline{1, N}$ — часова послідовність і $n(t)$ — помилка вимірювання («шум»). Задача прогнозування часової послідовності полягає у визначенні значення $s(t+l)$ (де l — крок прогнозування), враховуючи p попередніх значень послідовності $x(t)$, $t = \overline{1, N}$, а саме: $x(t-p+1)$, $x(t-p+2)$, ..., $x(t)$.

Для побудови та настроювання інтервальної нечіткої моделі прогнозування з вибірки даних послідовності $x(t)$, $t = \overline{1, N}$ формуються $(D-p-l+1)$ навчальних пар $x^{(1)}, x^{(2)}, \dots, x^{(i)}, \dots, x^{(D-p-l+1)}$, де D — розмір послідовності даних, взятих для формування навчальної вибірки.

Структура нечіткої моделі прогнозування складається з нечіткої бази знань (набір правил ЯК-ЩО-ТО), операції фазифікації, нечіткого логічного висновку та операції дефазифікації.

Нечітка база знань формується з навчальної вибірки таким чином: дані визначають центри нечітких множин, які з'являються у наборі правил [1].

Для l -крокового прогнозування буде $(D-p-l+1)$ правил, які отримуються з навчальних пар $x^{(1)}, x^{(2)}, \dots, x^{(D-p-l+1)}$:

$$R^l: \text{ЯКЩО } x_1 = F_1^l \text{ I } x_2 = F_2^l \text{ I } \dots \text{ I } x_p = F_p^l, \text{ ТО } y = G^l,$$

де $F_1^l, F_2^l, \dots, F_p^l, G^l$ — нечіткі множини, чії функції належності зосереджені відповідно в $x(1), x(2), \dots, x(p), x(p+l)$.

.....
 R^i : ЯКЩО $x_1 = F_1^i \text{ I } x_2 = F_2^i \text{ I } \dots \text{ I } x_p = F_p^i$, ТО $y = G^i$,

де $F_1^i, F_2^i, \dots, F_p^i, G^i$ – нечіткі множини, чії функції належності зосереджені відповідно в $x(i), x(i+1), \dots, x(i+p-1), x(i+p-1+D)$.

.....
 $R^{D-p-l+1}$: ЯКЩО $x_1 = F_1^{D-p-l+1} \text{ I } x_2 = F_2^{D-p-l+1} \text{ I } \dots \text{ I } x_p = F_p^{D-p-l+1}$, ТО $y = G^{D-p-l+1}$,

де $F_1^{D-p-l+1}, F_2^{D-p-l+1}, \dots, F_p^{D-p-l+1}, G^{D-p-l+1}$ – нечіткі множини, чії функції належності зосереджені відповідно в $x(D-p-l+1), x(D-p-l+2), \dots, x(D-l), x(D)$.

Отримана нечітка база знань містить $(D-p-l+1)$ правил, серед яких є суперечливі та надлишкові правила. При виключенні таких правил з бази знань якість прийняття рішень нечіткої моделі прогнозування підвищується.

Для опису інтервальних функцій належності обрано модифіковану гаусову форму

$$\mu(x) = e^{-\left(\frac{x-m}{[\delta^{\min}, \delta^{\max}]}\right)^2}, \tag{1}$$

де m — координата максимуму інтервальної функцій належності; δ^{\min} і δ^{\max} — відповідно коефіцієнт концентрації (розтягнення) нижньої та верхньої границі інтервальної функцій належності.

Значення параметра m інтервальної функцій належності визначаються даними експериментальної вибірки.

Інтервальна функція належності вихідної змінної визначається за формулою (нечіткий логічний висновок)

$$\mu_{G^i}(x_1, x_2, \dots, x_p) = [\underline{\mu}_{G^i}(x_1, x_2, \dots, x_p), \bar{\mu}_{G^i}(x_1, x_2, \dots, x_p)], \tag{2}$$

де $i = \overline{1, M}$, M — кількість правил.

Оскільки на виході інтервальної нечіткої моделі прогнозування буде інтервал $y = [y^l, y^u]$, причому $y^l \leq y^u$, то значення нижньої y^l та верхньої y^u границь інтервалу визначаються за допомогою формул [1]

$$y^l = \frac{\sum_{i=1}^M \mu_{G^i}^l w_i^l}{\sum_{i=1}^M \mu_{G^i}^l}, \tag{3}$$

$$y^u = \frac{\sum_{i=1}^M \mu_{G^i}^u w_i^u}{\sum_{i=1}^M \mu_{G^i}^u}, \tag{4}$$

де w_i^l і w_i^u — відповідно ваги i -го правила для визначення нижньої та верхньої границь інтервалу.

Для деяких значень параметрів інтервальних функцій належності значення нижньої границі вихідного інтервалу може бути більшим за значення верхньої границі $y^l > y^u$. Для уникнення даного результату використовується спеціальний алгоритм розрахунку нижньої та верхньої границь вихідного інтервалу [1].

Для того, щоб інтервальна нечітка модель прогнозування часових послідовностей працювала адекватно, потрібно провести її настроювання. Процедура настроювання інтервальної нечіткої моделі прогнозування складається з п'яти етапів:

I. Побудова нечіткої моделі прогнозування з використанням традиційних функцій належності.

II. Вилучення з нечіткої бази знань суперечливих та надлишкових правил.

III. Настроювання нечіткої моделі прогнозування з використанням традиційних функцій належності за допомогою генетичного алгоритму.

В основі генетичного алгоритму лежать 2 операції: схрещування – генерація хромосом-нащадків з хромосом-батьків, та мутація – випадкове змінення елементів хромосоми [2, 3]. Операція схрещування полягає в обміні частин хромосом в векторі параметрів функції належності та векторі ваг правил. Операція мутації передбачає зміну з деякою ймовірністю елементів хромосоми.

Структуру хромосоми, яка описує нечітку модель з використанням традиційних функцій належності $F(X, W, \Delta)$, де $X = (x_1, x_2, x_3, \dots, x_p)$ — вхідний вектор, $W = (w_1, w_2, \dots, w_M)$ — вектор ваг правил, $\Delta = (\delta_1, \delta_2, \dots, \delta_{p \cdot M})$ — вектор параметрів концентрації традиційних функцій належності, зображено на рис. 1.

δ_{11}	δ_{21}	...	δ_{p1}	δ_{12}	δ_{22}	...	δ_{p2}	...	δ_{1M}	δ_{2M}	...	δ_{pM}	w_1	...	w_M
---------------	---------------	-----	---------------	---------------	---------------	-----	---------------	-----	---------------	---------------	-----	---------------	-------	-----	-------

Рис. 1. Код нечіткої моделі прогнозування – хромосома

Функція відповідності рішення хромосоми оцінюється за формулою

$$FF = \frac{1}{\sqrt{Z}} \sqrt{\sum_{i=1}^Z (y_i - y_i^e)^2}, \tag{5}$$

де y_i^e — експериментальне значення; Z — кількість вхідних векторів (розмір навчальної вибірки).

IV. Побудова нечіткої моделі з використанням інтервальних функцій належності $F(X, W^u, W^l, \Delta^{\max}, \Delta^{\min})$ шляхом розтягнення параметрів нечіткої моделі з використанням традиційних функцій належності $F(X, W, \Delta)$ [4].

$W^u = (w_1^u, w_2^u, \dots, w_M^u)$ і $W^l = (w_1^l, w_2^l, \dots, w_M^l)$ — вектори ваг правил, які визначаються за формулами

$$w_j^u = w_j k_w^u; \tag{6}$$

$$w_j^l = w_j k_w^l, \tag{7}$$

де k_w^u і k_w^l – коефіцієнти розтягнення ваг правил w_j нечіткої моделі $F(X, W, \Delta)$; $j = \overline{1, M}$.

$\Delta^{\max} = (\delta_1^{\max}, \delta_2^{\max}, \dots, \delta_{p \cdot M}^{\max})$ і $\Delta^{\min} = (\delta_1^{\min}, \delta_2^{\min}, \dots, \delta_{p \cdot M}^{\min})$ — вектори параметрів концентрації інтервальних функцій належності, які визначаються за формулами

$$\delta_{ij}^{\max} = \delta_{ij} k_{\delta}^{\max}; \tag{8}$$

$$\delta_{ij}^{\min} = \delta_{ij} k_{\delta}^{\min}, \tag{9}$$

де k_{δ}^{\max} і k_{δ}^{\min} — коефіцієнти розтягнення параметрів концентрації функцій належності δ_{ij} нечіткої моделі $F(X, W, \Delta)$; $i = \overline{1, p}$, $j = \overline{1, M}$.

Коефіцієнти розтягнення k_W^u, k_W^l, k_S^{\max} і k_S^{\min} задає експериментатор, причому $k_W^u \geq k_W^l, k_S^{\max} \geq k_S^{\min}$.

V. Настроювання нечіткої моделі прогнозування з використанням інтервальних функцій належності за допомогою генетичного алгоритму [3].

δ_{11}^{\max}	δ_{11}^{\min}	...	δ_{p1}^{\max}	δ_{p1}^{\min}	...	δ_{1M}^{\max}	δ_{1M}^{\min}	...	δ_{pM}^{\max}	δ_{pM}^{\min}	w_1^u	w_1^l	...	w_M^u	w_M^l
----------------------	----------------------	-----	----------------------	----------------------	-----	----------------------	----------------------	-----	----------------------	----------------------	---------	---------	-----	---------	---------

Рис. 2. Код інтервальної нечіткої моделі прогнозування – хромосома

Структура хромосоми, яка описує інтервальну нечітку модель прогнозування $F(X, W^u, W^l, \Delta^{\max}, \Delta^{\min})$, зображена на рис. 2.

Функцією відповідності рішення хромосоми виступає показник якості інтервальної нечіткої моделі $FF (FF_1/FF_2)$, який оцінюється за допомогою формул

$$FF_1 = \sum_{i=1}^Z k_i; \tag{10}$$

$$FF_2 = \frac{1}{\sqrt{Z}} \sqrt{\sum_{i=1}^Z (y_i^u - y_i^l)^2}, \tag{11}$$

де $k_i = \begin{cases} 0, y_i^l \leq y_i^e \leq y_i^u; \\ 1, y_i^e < y_i^l \vee y_i^e > y_i^u. \end{cases}$

FF_1 визначає кількість експериментальних прикладів, що не потрапили у верифікований вихідний інтервал.

Задача настроювання нечіткої моделі прогнозування з використанням інтервальних функцій належності полягає в тому, щоб знайти такий вектор $(W^u, W^l, \Delta^{\max}, \Delta^{\min})$, за якого $FF_1 \rightarrow 0, FF_2 \rightarrow \min$.

Тестова задача

Метод побудови інтервальних нечітких моделей апробовано на тестовій задачі: прогнозування хаотичної часової послідовності Макей—Гласа, яка представлена нелінійним диференціальним рівнянням, один із видів якого

$$\frac{ds(t)}{dt} = \frac{0,2s(t - \tau)}{1 + s^{10}(t - \tau)} - 0,1s(t). \tag{12}$$

Результати розв’язання диференціального рівняння у вигляді послідовності значень (1201 точка) використано для формування навчальної (перші 600 точок) та тестової вибірки (решта точок).

Для дослідження впливу кількості входів на якість прийняття рішень нечіткою моделлю прогнозування часової послідовності Макей—Гласа побудовано декілька моделей з різною кількістю входів ($p = 2...30$). Результати дослідження показали, що найкращу якість прийняття рішень демонструє нечітка модель прогнозування з 30 входами. База знань тридцятивходової нечіткої моделі після вилучення суперечливих та надлишкових правил містить 378 правил.

Для підвищення якості прийняття рішень тридцятивходовою нечіткою моделлю прогнозування з використанням традиційних функцій належності проводиться її настроювання за допомогою генетичного алгоритму. На рис. 3 зображено графік залежності показників якості на навчальній та тестовій вибірках від тривалості настроювання.

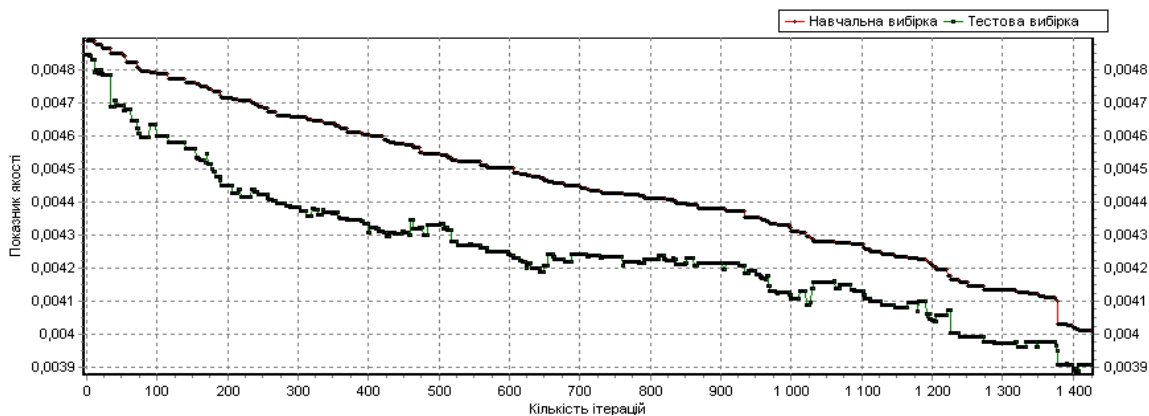


Рис. 3. Графік залежності показників якості від тривалості настроювання

На рис. 4 зображено результати прогнозування тридцятивходової нечіткої моделі з використанням традиційних функцій належності, настроєної генетичним алгоритмом. Точність точкового прогнозу побудованої моделі є достатньо високою, а саме: середня похибка прогнозування на тестовій вибірці становить 0,27 %, а максимальна похибка не перевищує 2 %

На рис. 5 показано графік похибки прогнозування.

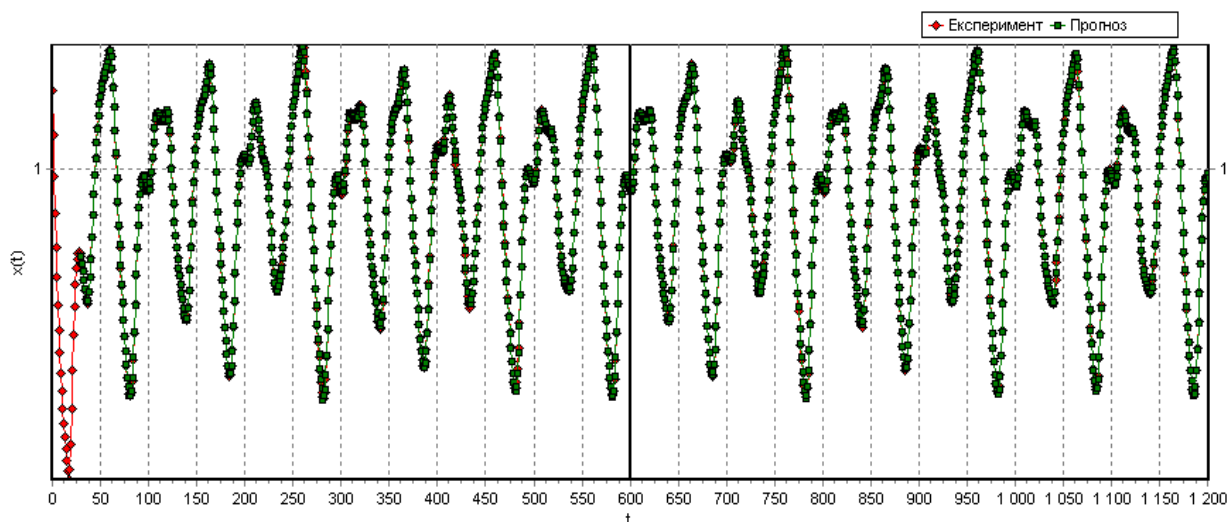


Рис. 4. Результати прогнозування з використанням традиційних функцій належності

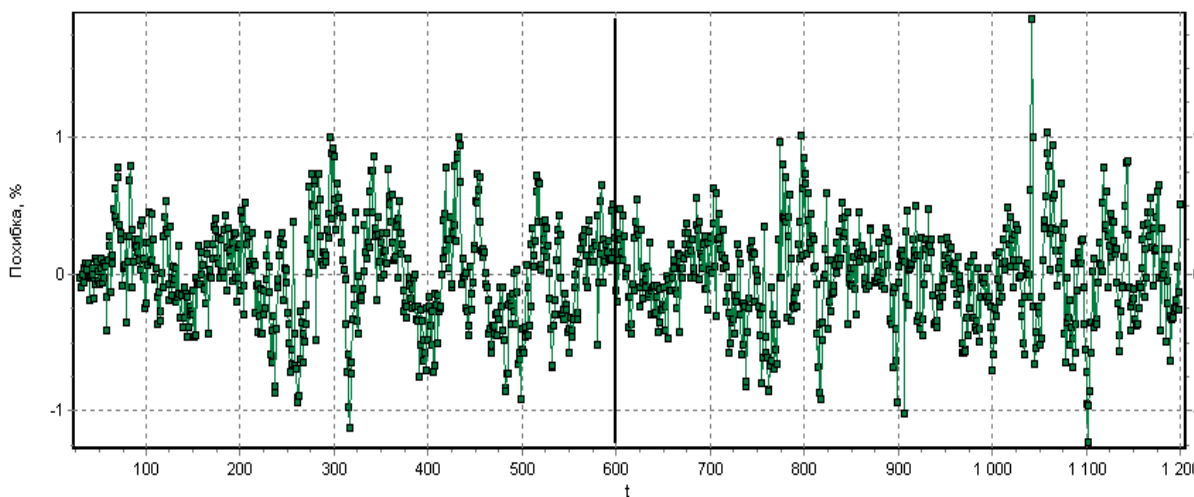


Рис. 5. Графік похибки прогнозування послідовності Макей—Гласа

На основі тридцятивходової нечіткої моделі з використанням традиційних функцій належності побудовано інтервальну нечітку модель. На рис. 6 зображено результати прогнозування інтервальної нечіткої моделі. В інтервал попадають усі експериментальні значення, тобто похибка прогнозування становить 0 %.

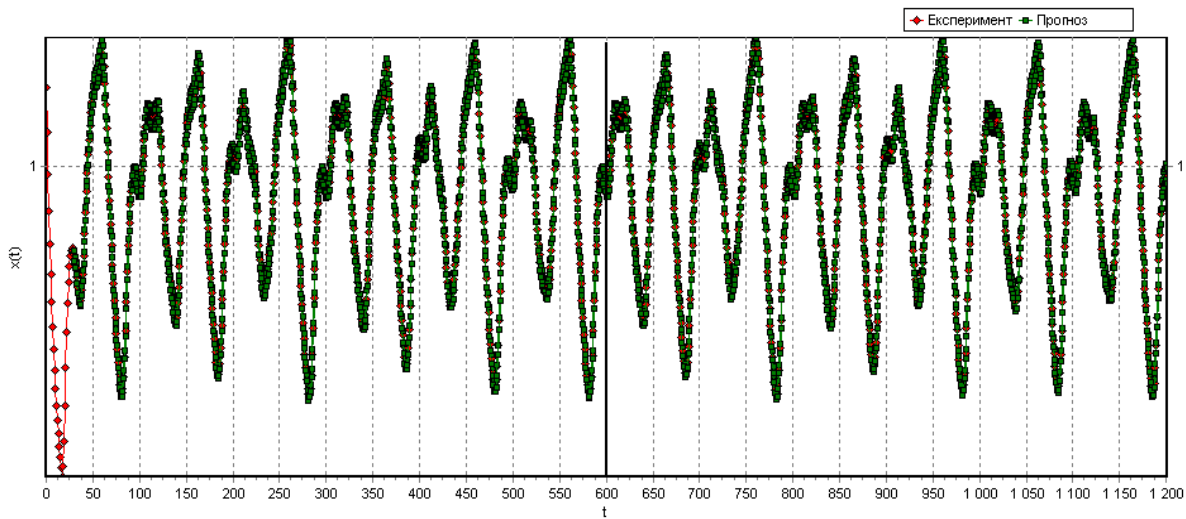


Рис. 6. Графічні результати інтервального прогнозування послідовності Макей—Гласа

Висновки

Розроблено метод побудови нечітких моделей з використанням інтервальних функцій належності для розв'язання задач прогнозування часових послідовностей. Побудовано інтервальну нечітку модель прогнозування хаотичної часової послідовності Макей—Гласа, яка продемонструвала високу якість прийняття рішень.

СПИСОК ЛІТЕРАТУРИ

1. Mendel J. M. Uncertain Rule-Based Fuzzy Logic Systems: Introduction and New Directions. — NJ: Prentice Hall, 2001.
2. Ротштейн А. П. Интеллектуальные технологии идентификации: нечеткие множества, генетические алгоритмы, нейронные сети. — Вінниця: УНІВЕРСУМ-Вінниця, 1999. — 320 с.
3. Кондратенко Н. Р., Куземко С. М., Чеборака О. В. Використання інтервальних функцій належності в задачах класифікації ендокринних захворювань // Інформаційні технології та комп'ютерна інженерія. — 2005. — № 3. — С. 85—90.
4. Кондратенко Н. Р., Зелінська Н. Б., Куземко С. М. Діагностика гіпотиреозу на основі нечіткої логіки з використанням інтервальних функцій належності // Наукові вісті (Національний технічний університет України «Київський політехнічний інститут»). — 2003. — № 4. — С. 52—58.

Кондратенко Наталія Романівна — доцент, **Чеборака Олександр Валерійович** — аспірант.

Кафедра обчислювальної техніки, Вінницький національний технічний університет