УДК 519.6
© І. Казимира[1], П. Клочко[1], 2012

СИСТЕМА ДОВГОСТРОКОВОГО ПРОГНОЗУВАННЯ ЧАСОВИХ РЯДІВ

Запропоновано систему довгострокового прогнозування часових рядів з автоматичним вибором базису вейвлет перетворення. Наведено результати тестування системи.

We propose system for forecasting long-term time series with automatic choice of wavelet transform basis. The results of system testing are presented.

1. ВСТУП

Пошук максимально точних методів прогнозування часових рядів є актуальною задачею. Зазвичай, процес розробки математичної моделі, яка описує поведінку системи є надто складним і довготривалим. Серед багатьох методів, які дозволяють отримати прогноз поведінки системи при відсутності її математичної моделі, популярними є нейромережеві методи [1,2]. В роботі [3] запропонований метод, який базуються на нечітких моделях Такагі–Сугено–Канга [4, 5] з використанням попереднього вейвлет-розкладу [6-9], який дозволяє виявити особливості (моменти виникнення тих чи інших частот) сигналу. Метод передбачає вейвлет-аналіз часового ряду, прогноз кожної складової окремо і об’єднання отриманих прогнозів за допомогою вейвлет-синтезу. Однак, в [3] невирішеним питанням лишається вибір оптимального базису вейвлет-перетворення, що суттєво впливає на результат прогнозування та час навчання мережі. В роботі запропонована система прогнозування з автоматичним вибором базису попереднього вейвлет-перетворення. Вибір множини базисів здійснюється на основі критеріїв запропонованих в [10]. Подальше визначенням найкращого базису здійснюється за критерієм мінімальної відносної похибки.

2. ПОСТАНОВКА ЗАДАЧІ

Метою роботи є розробка системи прогнозування з автоматичним вибором базису попереднього вейвлет-розкладу, та її перевірка прогнозуванням цін на різні види палива.

Основні задачі роботи наступні:

Розробка алгоритму прогнозування часових рядів. Алгоритм базується на таких основних процедурах: виборі множини базисів вейвлет-

[1] Національний університет «Львівська політехніка»

розкладу, вейвлет-розкладі навчальної вибірки, прогнозуванні коефіцієнтів розкладу нейро-нечіткими мережами Такагі–Сугено–Канга, вейвлет-синтезі найкращих результатів.

Реалізація алгоритму у вигляді програмної системи.

Оскільки обчислювальна складність алгоритму є досить високою, при реалізації алгоритму мають бути використані програмні інструменти паралелізму.

Перевірка роботи системи шляхом аналізу результатів прогнозування цін на різні види палива.

3. ВЕЙВЛЕТ-РОЗКЛАД ВХІДНОГО РЯДУ

Доцільність використання вейвлет-розкладу при прогнозуванні доведена в [11]. Схема розкладу вхідного ряду довжиною 2^n (n = 1,2...) наведена на рис. 1.

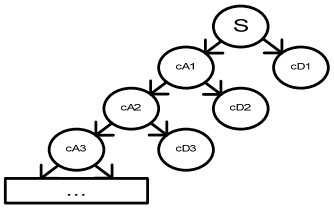

Рис. 1. Схема вейвлет-розкладу сигналу

cA – коефіцієнти апроксимації. Отримуються шляхом пропускання сигналу через низькочастотний фільтр з імпульсним відгуком g

cD – коефіцієнти деталізації. Отримуються шляхом пропускання сигналу через високочастотний фільтр з імпульсним відгуком h.

Ці два фільтра є квадратурнимо-дзеркальними і матриці, які їм відповідають, залежать від конкретного базису вейвлет-розкладу.

4. НЕЙРО-НЕЧІТКІ МЕРЕЖІ ТАКАГІ-СУГЕНО-КАНГА

Схема виводу в моделі Такагі-Сугено-Канга при використанні M правил і N змінних x_j має вигляд:

$$\begin{gathered} if\ (x_1\ is\ A_1^{(i)})(x_2\ is\ A_2^{(i)})\ldots(x_N\ is\ A_N^{(i)}) \\ then\ y_i = p_{i0} + \sum_{j=1}^{N} p_{ij} x_j, \end{gathered} \tag{1}$$

де умова $(x_1\ is\ A_1^{(i)})$ реалізується функцією належності:

$$\mu_A(x_i) = 1 \Big/ \left(1 + \left(\left(x_i - c_i\right)/a_i\right)^{b_i}\right), \tag{2}$$

де a_i, c_i, b_i- параметри функції належності, що підлягають адаптації в процесі навчання.

Варто відмітити, що в модулях ANFIS програмного пакету MATLAB реалізовано 8 функцій належності:

"trimf" – трикутна функція;

"trapmf" – трапецієподібна функція;

"dsigmf" – різниця між двома сигмоїдними функціями;

"psigmf" – добуток двох сигмоїдних функцій;

"gauss2mf" – двостороння гаусова функція належності;

"gaussmf" – симетрична гаусова функція належності;

"gbellmf" – узагальнена колоколоподібна функція;

"pimf" – пі-подібна функція;

При M правилах вихідний результат має вигляд:

$$y(x) = \sum_{i=1}^{M} \omega_i y_i(x) \Big/ \sum_{i=1}^{M} \omega_i, \qquad (3)$$

де ω_i – ваги входів

Отже, формулі (1) можна співставити відповідну багатошарову мережу зображену на рис. 2.

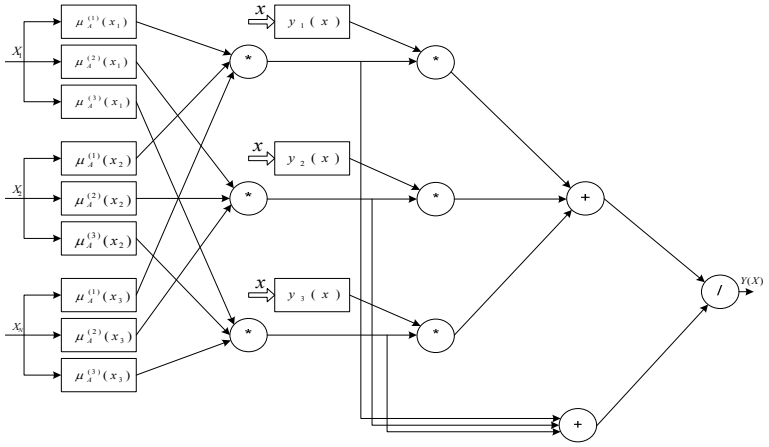

Рис.2. Багатошарова нейро-нечітка мережа

Необхідними функціями модуля ANFIS для прогнозування є:

genfis1(args) – генерує з даних систему нечіткого логічного висновку типу Сугено;

anfis(args) – основна функція налаштування систем нечіткого логічного виведення типу Сугено;

evalfis(args) – виконує нечіткий логічний висновок;

Варто відмітити, що при прогнозуванні за допомогою вищенаведених функцій, вхідний ряд треба перетворити у матрицю, кількість стовбців якої було б на одиницю більше кількості входів нейромережі.

Останній стовбець матриці буде відповідати виходу нейромережі в процесі навчання.

5. АЛГОРИТМ ПРОГНОЗУВАННЯ

В даній роботі запропонований наступний алгоритм вибору базису вейвлет-розкладу при прогнозуванні часових рядів:

– визначення та виділення необхідної кількості точок навчальної вибірки з вхідної послідовності із заокругленням до степеня двійки (в меншу сторону);

– вибір множини базисів за критеріями описаними в [10].

– визначення та виділення необхідної кількості точок прогнозу(перевірки) для кожного коефіцієнту розкладу;

– розклад вхідної послідовності за вибраним базисом;

– створення нейро-нечітких мереж для кожного з коефіцієнтів розкладу з усіма можливими функціями належності (fuzzification) та їх подальше навчання;

– вибір найкращого з прогнозів для кожного коефіцієнту розкладу за критерієм відносної помилки.

– вейвлет-синтез найкращих прогнозованих показників

Кроки 5-7 повторюється для всіх вибраних базисів вейвлет-перетворення, після чого визначається найкращий базис за критерієм відносної помилки.

6. ОПИС СИСТЕМИ

Вищенаведений алгоритм реалізований у вигляді двох незалежних процесів, які обмінюються повідомленнями за допомогою іменованого каналу (англ. *named pipe*) [12-14]:

Основного процесу.

Основний процес виконує розрахункову частину, яка складається з: модулю вейвлет-розкладу (рис. 3), модулю нейро-нечіткого прогнозу вейвлет–коефіцієнтів (рис. 4.), модулю вейвлет синтезу та порівняння результатів (рис. 5). Розрахункова частина підключає бібліотеки C++, сформовані в MATLAB, які реалізують функціональні можливості нейро-нечітких мереж Такагі-Суггено-Канга.

Для пришвидшення процесу прогнозування створення нейро-нечітких мереж з усіма можливими функціями належності (fuzzification) та їх подальше навчання, реалізовано в окремих потоках(Posix threads).

Інтерфейсу користувача – налаштування вхідних даних та візуалізація результату

В реалізованій системі можливе конфігурування таких вхідних даних за допомогою графічного інтерфейсу (рис. 4) або конфігураційного файлу:

```
#Приклад конфігураційного файлу
# Шлях до файлу, який містить часовий ряд
series_file = /home/student/MATLAB/fs_main_forecast/series
# Кількість навчальних точок
training_data_size = auto
# Кількість точок прогнозу
forecast_data_size = 8
# Стартова точка
start_point = 1
# Перелік назв вейвлет-базисів
wave_names={"db1","db2","db3","db4","db5","db6","db7","db8",
"db9","db10","db11","db12","db13","db14","db15","db16","db17",
"db18","db19","db20","db21","db22","db23","db24","db25","db26",
"db27","db28","db29","db30","db31","db32","db33","db34","db35",
"db36","db37","db38","db39","db40","db41","db42","db43","db44",
"db45","coif1","coif2","coif3","coif4","coif5","bior1.1","bior1.3",
"bior1.5","bior2.2","bior2.4","bior2.6","bior2.8","bior3.1","bior3.3",
"bior3.5","bior3.7","bior3.9","bior4.4","bior5.5","bior6.8","rbio1.1",
"rbio1.3","rbio2.2","rbio2.4","rbio2.6","rbio2.8","rbio3.1","rbio3.3",
"rbio3.5","rbio3.7","rbio3.9","rbio4.4","rbio5.5",  "rbio6.8" }
# Глибина розкладу вейвле- перетворення
depth_dec = default
# Кількість стовбців при сортуванні вхідного ряду
sort_column_count = 5
# Перелік функцій належності
in_mf_types = {'gbellmf','gaussmf','gauss2mf','pimf','dsigmf','psigmf'}
```

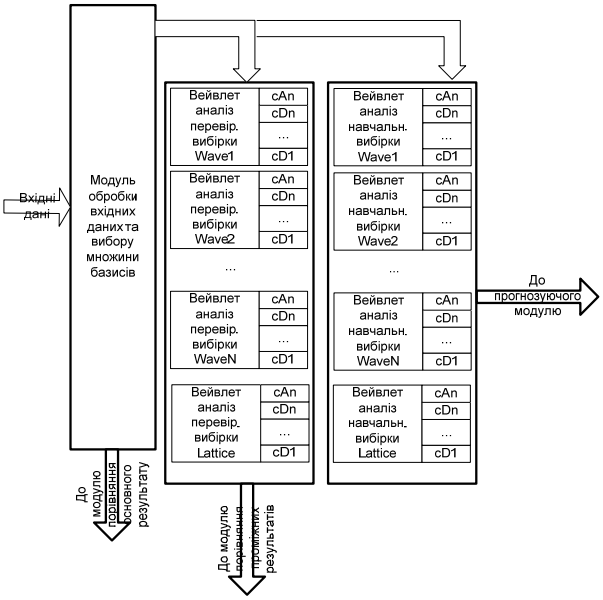

Рис. 3. Схема модулю вейвлет-розкладу

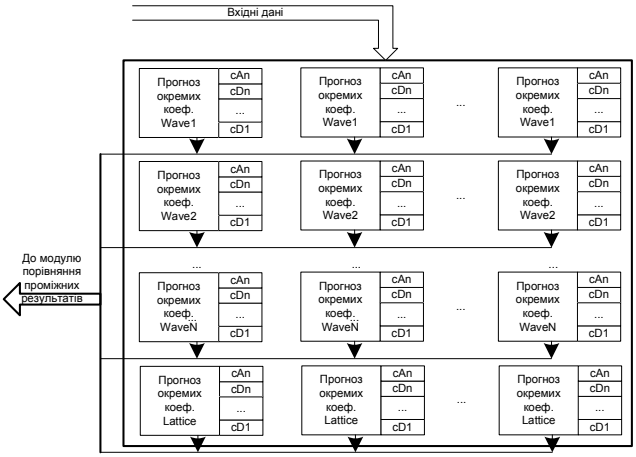

Рис. 4. Схема модулю нейро-нечіткого прогнозу вейвлет–коефіцієнтів

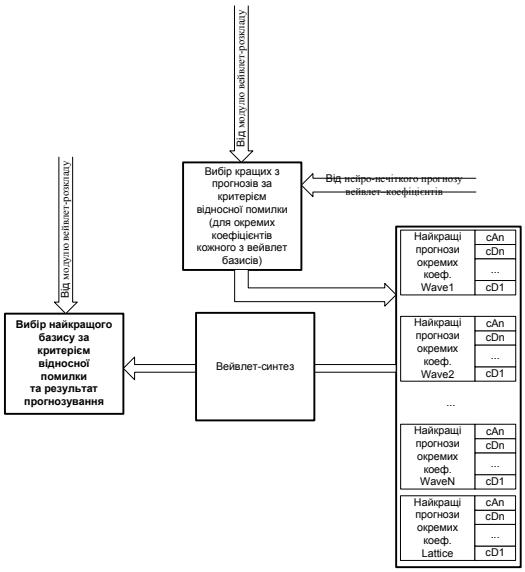

Рис. 5. Схема модулю вейвлет синтезу та порівняння результатів

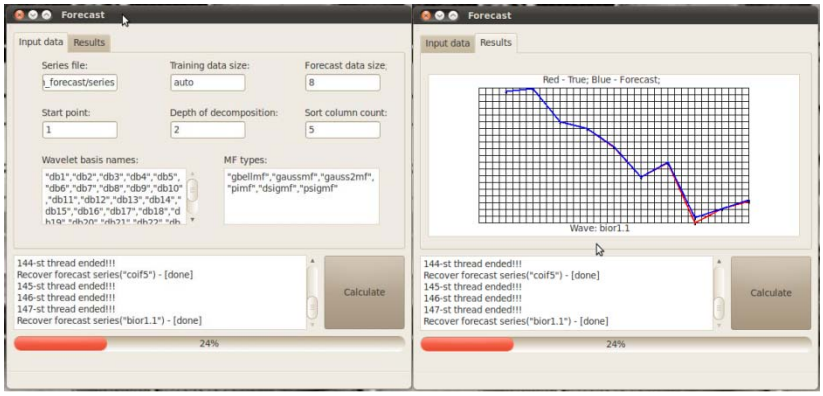

Рис. 6. Інтерфейс користувача

7. РЕЗУЛЬТАТИ ПРОГНОЗУВАННЯ

Розроблена система протестована на п’ятьох часових рядах (цінах на паливо за 21 рік). Відносні помилки результатів прогнозування вказані в табл. 1.

Таблиця 1

Результати прогнозування цін на паливо з 1990 по 2011 роки

Тип палива / Базис	Неетил. прем. (роздр. продаж)	Неетил. прем. (без ПДВ)	Дизель (роздр. продаж)	Дизель (без ПДВ)	ПДВ 2005=100
db1	11,58	31,73	9,59	31,89	0,5
db2	10,71	27,48	3,00	20,56	1,94
db3	5,22	29,02	7,13	33,06	5,07
db4	13,36	19,00	10,62	12,46	0,94
coif1	16,11	25,04	5,77	26,97	1,18
coif2	10,37	31,074	5,04	66,52	34,86
bior1.1	11,59	31,74	9,54	31,89	0,45
bior1.3	13,69	25,72	4,57	29,01	0,57
bior1.5	14,74	19,21	5,54	31,49	1,00
bior2.2	13,30	23,31	15,82	101,67	1,00
bior2.4	6,45	28,63	18,49	35,02	0,87
rbio1.1	11,58	31,74	9,54	31,89	0,50
rbio1.3	8,91	27,03	8,06	77,77	0,93
rbio2.2	12,05	29,87	4,51	36,74	0,62
rbio2.4	11,04	70,02	3,18	19,87	1,06
ВП кращого прогнозу:	5,22%	19,00%	3,00%	12,46%	0,50%
Кращий базис	db3, bior2.4	db4, bior1.5	db2, rbio2.4	db4	db1, rbio1.1

8. ВИСНОВКИ

Розроблено та реалізовано систему прогнозування часових рядів з автоматичним вибором базису вейвлет-перетворення при прогнозуванні. Система протестована на сімох рядах цін на різні види палива. Різноманітність результуючих базисних функцій (базисів вейвлет-розкладу вхідного ряду, при використанні яких результат прогнозу буде мати найменше відхилення від істинного) підтверджує актуальність реалізованої системи, а також необхідність побудови адаптивних базисних функцій.

1. Круглов В. В., Борисов В. В., Искусственные нейронные сети. Теория и практика. – 1-е. – М.: Горячая линия – Телеком, 2001. – 382 с. 2. Хайкин С. Нейронные сети: полный курс – 2-е. – М.: «Вильямс», 2006. – 1104 с. 3. Анушина Е.С. Прогнозирование временных рядов на основе интеллектуальных вычеслительных технологий//Интеллектуальные системы 2009. №1(19) – С.84-90.

4. Кофман А. Введение в теорию нечетких множеств. М.: Радио и связь, 1982. – 432 с. 5. Zadeh L.A. Fuzzy sets. – Information and Control, 1965, vol.8, N 3,pp.338-353. 6. Смоленцев Н. К. Основы теории вейвлетов. Вейвлеты в MATLAB. – М.: ДМК Пресс,2008. – 448 с.:ил. 7. Daubechies I., Ten Lectures on Wavelets // CBMS-NSF Lecture Notes nr. 61, SIAM, 1992. – 377 p. 8. Малла С. Вэйвлеты в обработке сигналов / Малла С.; пер. с англ. – М.: Мир, 2005. – 671 с., ил. 9. Charles K. Chui, An Introduction to Wavelets, (1992), Academic Press, San Diego 10. Мандрикова О. В., Полозов Ю. А.Критерии выбора базисной вейвлет-функции в задачах аппроксимации и выделения локальных особенностей сложных временных рядов // 5-ая междунар. конф. «Проектирование инженерных и научных приложений в среде MATLAB», Харьков, 11–13 мая 2011 г. 11. Серая О.В. Прогнозирование вейвлет-аппроксимации временного ряда // Восточно-Европейский журнал передовых технологий. – 4/4 (52) 2011 – 49–51с. 12. Воеводин В. В., Воеводин Вл. В. Параллельные вычисления. – СПб: БХВ-Петербург, 2002. – 608 с. 13. Антонов А. С. Параллельное программирование с использованием технологии MPI: Учебное пособие. – М.: Изд-во МГУ, 2004. – 71 с. 14. Чекалов О. Л. Основи функціонування операційних систем. Практикум. Навчальний посібник. – Суми: Вид-во СумГУ, 2010. – 85 с.