

**ПРОГНОЗУВАННЯ НЕЛІНІЙНИХ НЕСТАЦІОНАРНИХ
ПРОЦЕСІВ ЗА ДОПОМОГОЮ ЛІНГВІСТИЧНОГО МОДЕЛЮВАННЯ ТА
ПРИХОВАНИХ МАРКОВСЬКИХ МОДЕЛЕЙ**

Селін Ю.М. к.т.н., Селін О.М. к.т.н., доцент,
Жданова Е.Г. к.т.н., доцент, Богданенко Н.О. бакалавр

*Національний технічний університет України
«КПІ ім. Ігоря Сікорського», Україна*

Abstract. The article describes a mathematical apparatus that can be used to analyze data of various nature for the analysis and forecasting of non-linear non-stationary processes of different nature. The object of the study is non-linear non-stationary processes in ecology, economics and finance. The methods of hidden Markov models and linguistic modeling are described. The proposed mathematical tool can use in solving the problems of data analysis of various nature for nonlinear non-stationary processes forecasting. The results of numerical experiments on the use of the mathematical apparatus are presented. The analysis showed the high adequacy of the results obtained with the help of the developed information technology and high quality of the received forecasts.

Ключові слова: НЕЛІНІЙНІ НЕСТАЦІОНАРНІ ПРОЦЕСИ, ПРОГНОЗУВАННЯ, ЛІНГВІСТИЧНЕ МОДЕЛЮВАННЯ, ПРИХОВАНІ МАРКОВСЬКІ МОДЕЛІ.

Інтелектуальні методи аналізу даних по визначенню - це процес пошуку в «сирих» даних раніше невідомих, нетривіальних, практично корисних і доступних інтерпретації знань, необхідних для прийняття рішень в різних сферах людської діяльності [1]. Інтелектуальний аналіз даних - термін, який застосовується для опису отримання знань в базах даних, дослідження даних, обробки зразків даних, очищення та збору даних. Це процес виявлення кореляції, тенденцій, шаблонів, зв'язків і категорій [2] Інтелектуальний аналіз даних розвивається на базі таких наук, як прикладна статистика, розпізнавання образів, штучний інтелект, теорія баз даних тощо.

Процес автоматичного пошуку прихованих закономірностей або взаємозв'язків між змінними в інтелектуальному аналізі даних ділиться на завдання класифікації, моделювання і прогнозування з використанням статистичних і математичних методів.

Складні завдання на макрорівні, такі як завдання прогнозування якості життя, характеризуються аналізом і прогнозуванням нелінійних нестационарних процесів. У зв'язку з цим пропонуються до розгляду деякі методи, можна застосовувати для використання в подібних завданнях. До них можна віднести наступні методи: Приховані марковские моделі (ПММ); Лінгвістичний моделювання (ЛМ).

Метод ПММ вирішує наступну задачу. Якщо ви задаєте послідовності O_i спостереження та моделі $\lambda = (A, B, P)$, то, як ефективно вивісити $R(O / \lambda)$ - вірогідність такої послідовності при заданих моделях?

Очевидно, що рішення першої задачі може зробити ефективно розташування елементів з декотрого обмеженого набору, що мають готові, вже вбудовані моделі цих елементів. У цьому випадку існує елемент дискретизується і після сегментації на певних примітивах розглядається як набір послідовників нагляду. Замість безпосередньо графічного електронного сигналу системи зазвичай працює з яким-то набором примітивів (меншим за об'єктом), що відповідає процесам спостереження. Цей етап називається попередньою обробкою. Потім кожна послідовність спостережень, яка поки є невідомим набором спостережень (словом), перевіряється на відповідність усім наявним в словнику моделям (слів). Для цього обчислюється ймовірність генерації такої тривалості наближення для кожного з них, що знаходяться в слові моделей. Слово, модель якого з найбільшою ймовірністю генерує таку послідовність спостереження, і є шуканим результатом. [3].

При лінгвістичному моделюванні відбувається заміна тимчасового чмслового ряду спостережень на символний ряд:

$$L: F(t_i) \rightarrow \langle e_i \rangle;$$

де $F(t_i)$ - часовий числовий ряд,

$\langle e_i \rangle$ - символний ряд.

У загальному випадку лінгвістична модель динамічного процесу складається з наступних елементів

$$\langle D, I, L, G \rangle,$$

де D - сукупність тимчасових рядів динамічного процесу і рядів, похідних від вхідних даних;

I - спосіб і правила інтервалізації;

L - морфізм відображення інтервального представлення низки на певний алфавіт;

G - відновлена граматики динамічного процесу.

Згідно етапів побудови лінгвістичної моделі вихідна задача буде розбита на такі підзадачі: отримання різницевих рядів; інтервалізації; лінгвістизації; побудови матриці переходів.

Отриману послідовність символів аналізують на наявність граматичних конструкцій. На виході отримуємо список граматичних конструкцій з можливостями їх наявності процесі, а також матрицю ймовірностей переходу з символу в символ. Прогноз представляється у вигляді спектра ймовірностей.

Наведені методи є універсальними як з боку виду отриманої інформації так і з боку наявності в цій інформації нелінійностей і нестационарностей. Але мають загальний недолік всіх статистичних методів - відсутність історичної інформації.

Автори в ході проведених досліджень провели безліч чисельних експериментів з багатьма числовими рядами і отримали практичне підтвердження придатності описаних методів для аналізу і прогнозування нелінійних нестационарних процесів різної природи [4].

Література

1. U. Fayyad, G. Piatetsky-Shapiro, P. Smyth. From Data Mining to KnowledgeDiscovery in Databases. AI Magazine. 1996 №17(3).
2. В. Л. Плєскач. Інформаційні системи і технології на підприємствах / В. Л. Плєскач, Т. Г. Затонацька. // підручник - К. 2011 : Знання.
3. Шулькевич Т.В. Прогнозирование нелинейных нестационарных процессов разной природы / Шулькевич Т.В., Нестеренко А.В., Селин Ю.Н. // Информационные технологии и специальная безопасность. Научный журнал № 1(003) 2018. – С. 56-62..
4. Shulkevich T., Selin Y., Savchenko V. Data Mining and Nonlinear Non-stationary Processes Forecasting by Using Linguistic Modeling Method. In: Hu Z., Petoukhov S., Dychka I., He M. (eds) Advances in Computer Science for Engineering and Education II. ICCSEEA 2019. Advances in Intelligent Systems and Computing, vol 938, pp 409-418. Springer, Cham.

References

1. U. Fayyad, G. Piatetsky-Shapiro, P. Smyth. From Data Mining to KnowledgeDiscovery in Databases. AI Magazine. 1996 №17(3).

2. V. L. Pleskach. Information systems and technologies at enterprises / VL Plesachach, TG Zatonatskaya. // textbook - K. 2011: Znannia.

3. Shulkevich T.V. Prediction of nonlinear non-stationary processes of various nature / Shulkevich T.V., Nesterenko A.V., Selin Yu.N. // Information technology and special security. Scientific journal No. 1 (003) 2018. - P. 56-62 ..

4. Shulkevich T., Selin Y., Savchenko V. Data Mining and Nonlinear Non-stationary Processes Forecasting by Using Linguistic Modeling Method. In: Hu Z., Petoukhov S., Dychka I., He M. (eds) Advances in Computer Science for Engineering and Education II. ICCSEEA 2019. Advances in Intelligent Systems and Computing, vol 938, pp 409-418. Springer, Cham.