

**ДОСЛІДЖЕННЯ ІСНУЮЧИХ ТЕХНОЛОГІЙ АНАЛІЗУ
ЧАСОВИХ РЯДІВ, МОДЕЛЮВАННЯ ПРОГНОЗУ
НА ЇХ БАЗІ ТА ВИЯВЛЕННЯ АНОМАЛІЙ**

Бойко А. М., Мстельов В. О.

*Національний технічний університет
«Харківський політехнічний інститут», м. Харків*

Мета даної роботи – проектування програмного забезпечення задля дослідження часових рядів при вирішенні задачі пошуку і виявлення аномалій в даних, а також побудови прогнозів при роботі з масивами даних, користуючись існуючими методами математичного аналізу.

У ході аналізу основних існуючих технологій аналізу часових рядів, зроблено висновок, що використання алгоритмів паралельних обчислень має перспективу суттєво підвищити швидкість аналізу часових рядів. Швидкість аналізу часових рядів даних, що неперервно генеруються є критично важливим критерієм, адже від нього залежить актуальність і доцільність прийняття рішень щодо керованої системи. Прикладом такої системи може бути керування низкою літальних апаратів на основі низки даних отриманих з сенсорів розташованих у багатьох локаціях нашої планети (прикладом групи таких даних можуть бути метеорологічні показники).

В цілому, у якості вхідних даних розглядаються добірки записів взаємодії користувача програмного забезпечення з продуктами, якими він користується, а також безпосередньо записи самого програмного забезпечення (так звані “логі” користувача і програмного забезпечення). Важливою ознакою таких даних має бути поява значень, що не відповідають природі процесу, що досліджується - аномальних значень.

У ході роботи основний акцент ставиться на проблемі швидкості роботи існуючих алгоритмів прогнозу та пошуку аномалії на основі статистичних даних часових рядів, особливо коли мова йде про обробку великих масивів даних.

Розглядається декілька моделей аналізу даних задля виявлення аномалій і побудови прогнозів, як з точки зору швидкості виконання процесу аналізу без паралелізації обчислень, так і при їх паралелізації.

Виявлено, що далеко не кожна математична модель може бути ефективно описана алгоритмічними методами для виконання паралельних обчислень.

В роботі також розглядається проблема організації середовища для зберігання даних за умови, що дані не можуть бути розміщені на одній обчислювальній машині (локалізовані).

Підсумком виконання роботи є програма, що використовує методи паралельних обчислень для аналізу часових рядів і виявлення аномалій у даних в реальному часі.