

ISSN 2074-5893



ПИТАННЯ

ПРИКЛАДНОЇ

МАТЕМАТИКИ

і

МАТЕМАТИЧНОГО

МОДЕЛЮВАННЯ

2011

УДК 664.123.4:004.67

О.М. Зігунов, В.Д. Кишенько

Національний університет харчових технологій

ПАТЕРН-АНАЛІЗ ВХІДНОЇ ІНФОРМАЦІЇ У ПІДСИСТЕМІ ТЕХНОЛОГІЧНОГО МОНІТОРИНГУ ДИФУЗІЙНОГО ВІДДІЛЕННЯ ЦУКРОВОГО ЗАВОДУ

Розглядаються методи обробки і аналізу даних, отриманих підсистемою технологічного моніторингу дифузійного відділення цукрового заводу. Показана можливість ефективного розпізнавання патернів типу «пік» і «повільна хвиля» за вейвлет-спектрами часових рядів.

Рассматриваются методы обработки и анализа данных, полученных подсистемой технологического мониторинга диффузионного отделения сахарного завода. Показана возможность эффективного распознавания паттернов типа «пик» и «медленная волна» по вейвлет-спектрам временных рядов.

Methods of processing and data analysis obtained by technological monitoring subsystem of diffusion department of sugar factory are considered. The possibility of effective recognition of patterns such as «peak» and «slow wave» by wavelet spectra of time series is shown.

Ключові слова: патерн-аналіз, система технологічного моніторингу, вейвлет-спектр, технологічний процес.

Вступ. У сучасних умовах одним з найважливіших факторів науково-технічної революції є широка інформатизація і автоматизація людської діяльності, створення і розвиток відповідних автоматизованих і інформаційних систем. Під автоматизованою системою розуміється [3] система, що складається з персоналу і комплексу засобів автоматизації його діяльності, що реалізує інформаційну технологію виконання встановлених функцій.

Сучасний рівень розвитку технологій породжує не тільки складні виробничі цикли взаємозалежних технологічних процесів (ТП), але і значне ускладнення самих ТП, ефективне керування якими неможливо з викорис-

танням стандартних SCADA-систем. Підвищення ефективності моніторингу стану складних промислових об'єктів можливо на основі розширення аналітичних інформаційних систем на рівень технологічного керування підприємством і введення додатково до SCADA-системи нової аналітичної інформаційної системи, що перетворює загальний технологічний інформаційний потік до оптимального вигляду для ситуаційного аналізу.

Постановка задачі. Система технологічного моніторингу (СТМ) для забезпечення функціонування системи управління складними технологічними комплексами може виділити зміни, спричинені в результаті функціонування технологічного процесу, коли потрібна детальна інформація про різні коливання і зміни. Моніторинг включає до себе такі основні напрямки діяльності:

1. Спостереження за факторами, які впливають на довілля та на стан технологічного процесу.

2. Оцінку фактичного стану технологічного процесу.

3. Прогноз стану технологічного процесу та оцінка цього стану.

Отже, технологічний моніторинг є багатоцільовою інформаційною системою. Його основні задачі: спостереження за станом об'єкта, оцінка та прогноз його стану; визначення інтенсивності різних впливів, виявлення факторів та джерел таких впливів.

У підсистемі технологічного моніторингу дифузійного відділення цукрового заводу використовуються датчики температури, маси, витрати, тиску, рівня, властивостей речовин, струмового навантаження, об'єднані в єдину мережу.

Варто також врахувати можливість впливу зовнішніх факторів на показання датчиків, що приводить до перекрученого представлення про стан об'єкта.

Метод розв'язування. Використання в СТМ мережі різнотипних датчиків змушує створювати складну систему збору, обробки і аналізу даних (рис. 1), що повинна:

- здійснювати прийом і форматування сигналів датчиків;
- формувати вектор характеристик стану об'єкта з використанням вектора сигналів датчиків. З метою усунення надмірності вектор характеристик оптимізується;
- зберігати історію об'єкта (дані, зібрані за весь період спостереження за об'єктом) у вигляді «знімків» (кадрів) поточного стану протягом декількох сезонів роботи цукрового заводу із використанням мето-

- дів компактного представлення даних при збереженні можливості ретроспективного аналізу.
- «учитися» розпізнавати риси нормальної поведінки (зміни у стані) об'єкта і аномальних змін. Процес «навчання» повинен охоплювати природні життєві цикли об'єкта, пов'язані зі змінами зовнішніх впливів (температура, вологість і ін.) та режимів експлуатації. Навчання і аналіз змін – це два паралельних взаємозалежних процеси. На початковому етапі експлуатації системи неминучі фіктивні тривоги;
- аналізувати зміни з метою локалізації відхилень і оцінки їхньої значущості. Для класифікації необхідно створити і вести базу знань, що враховує досвід експлуатації аналогічних об'єктів, результати моделювання, думки експертів та ін.



Рис. 1. Підсистема технологічного моніторингу дифузійного відділення цукрового заводу

Унаслідок недосконалості датчиків у їхніх показаннях є шумові складові, тому дані, що надходять до системи, повинні сприйматися в статистичному плані. Виходячи з того, що параметри розподілів сигналів заздалегідь невідомі, то варто намагатись використовувати універсальні оцінки, одержувані, наприклад, у вигляді декомпозиції вейвлетів (wavelets decomposition) [1] або ієрархії кластерів [2].

Для скорочення обсягу оброблюваних і збережених даних як характеристики стану об'єкта використовують агреговані, «сплавлені» показання декількох датчиків. Підставою для агрегування можуть служити функціональні зв'язки, виявлені при моделюванні роботи об'єкта, кореляційні зв'язки, зафіксовані у процесі спостережень за поведінням об'єкта та ін. Мета агрегування – скорочення розмірностей у багатомірному просторі даних моніторингу. Відомими прийомами скорочення розмірностей є PCA (Principal Component Analysis) [4], MDS (Multidimensional Scaling) [5], SOM (Self Organizing Maps) [6], процес візуальної інтерактивної редукції розмірностей VHDR (Visual Hierarchical Dimension Reduction).

Історія об'єкта потрібна для виявлення змін у його стані і навчання СТМ. Вона може бути представлена у вигляді «знімків» – записів бази даних, що містять оцінки поточного стану об'єкта. Наприклад, це може бути набір мікрокластерів, модифікованих під ієрархічну кластеризацію. «Знімок» можна представити у вигляді кортежу $S = (D, H, T)$, де D – дата формування «знімка», H – інформація про стан об'єкта на час формування «знімка», T – додаткова інформація.

Відповідно до завдання стиснення масиву даних, що зберігаються, знімки можна зібрати в «піраміду» (pyramidal time frame), у якій пізніші знімки частіше заповнюють часову вісь, чим ранні. Можна використовувати логарифмічну шкалу, де часові інтервали між знімками змінюються за законом $\Delta t = a^i$, де a – деяке ціле число $a \geq 1$; i – ранг інтервалу.

Якщо інтервал часу T від початку процесу спостереження до поточного моменту t_c виражений у дискретах тактового сигналу, то він буде населений знімками рангів від 0 до $\log_a(T)$.

Якщо зберігати тільки $(a^m + 1)$ останніх знімків кожного рангу, то загальна кількість збережених знімків складе не більше $(a^m + 1)\log_a(T)$. Для цілей аналізу важлива точність апроксимації стану об'єкта в момент часу, що цікавить. Найбільш розрідженими є інтервали на початку процесу. Помилка апроксимації тут буде найбільшою. Оцінки можна одержати з наступних міркувань.

Серед збережених кадрів знайдуться знімки рангу i , які задовольняють співвідношення

$$\text{int}(\log_a T) - m = r_i,$$

де $\text{int}(\log_a T)$ – найближче найбільше ціле.

Це означає, що інтервал T розбитий на a^m інтервалів розміром не більше a^r тактів, і максимальна похибка при апроксимації буде в межах

$$\delta = 0,5(a^r/T).$$

Можливо подальше скорочення кількості знімків, що зберігаються. Так, недоцільно зберігати знімки, що відносяться до перехідних процесів, пов'язаних зі швидкою зміною температури навколишнього середовища. У найпростішому випадку можна відбирати знімки за мінімумом варіації показань датчиків у підбраному змінному вікні, однак можлива процедура, за якою аналізується мінливість мікрокластерів, що становлять знімок. Це «промивання» більш високого, глобального рівня (макроаналіз).

Вейвлет-спектр містить повну інформацію про сигнал. Характерною рисою вейвлет-спектра є те, що він дозволяє виявляти не тільки частотні характеристики сигналу, але і їхню локалізацію за часом. При цьому за зовнішнім виглядом вейвлет-спектра можна судити про подібність ділянок сигналу материнському вейвлету.

Вейвлет-аналіз знайшов широке застосування при діагностуванні станів об'єктів керування, оскільки його використання дозволяє адекватним чином досліджувати масштабно-інваріантну динаміку складних технічних систем. Перевагою вейвлет-аналізу також є можливість локальної оцінки різномасштабних частотних характеристик часових рядів, що особливо цінно при вирішенні задач розпізнавання порушень роботи об'єкта.

Діагностика на основі аналізу вхідної інформації у підсистемі технологічного моніторингу вимагає виявлення певних характерних рис – патернів. Їхня кількість і вид дозволяють судити про наявність або відсутність відхилень. Кожний патерн визначається набором ознак: частотою, амплітудою, тривалістю, повторюваністю.

Аналіз одержаних результатів. Розглянемо виявлення патернів типу «пік» і «повільна хвиля» за вейвлет-спектрами.

В якості вихідних даних використовувалися дані від датчиків системи автоматизації дифузійного відділення цукрового заводу, до складу якої входить підсистема технологічного моніторингу.

Розрахунок вейвлет-спектрів виконувався з використанням пакету Interactive Wavelets, розташованого в мережі Internet за адресою <http://ion.researchsystems.com/IONScript/wavelet/>. Оцінювалася можливість виявлення особливостей вейвлет-спектрів, що відповідають різним патернам.

Характерним для ряду станів є поява в часовому ряді комбінацій «пик» – «повільна хвиля». Типовий запис патерну «пик» - «повільна хвиля» представлений на рис. 2. Також на рис. 2 наведений вейвлет-спектр цієї ділянки, обчислений для вейвлета Морле (1), який має аналітичний вигляд [1]:

$$\psi(t) = e^{-\frac{t^2}{2}} e^{i2\pi t} \quad (1)$$

Розрахунок вейвлет-коефіцієнтів сигналу $\{x(t), t \in \overline{1, T}\}$, відбувається за формулою (2)

$$W(a, b) = \frac{1}{\sqrt{a}} \sum_{t=0}^{T-1} x(t) \int_{t\Delta T}^{(t+1)\Delta T} \psi^* \left(\frac{\tau - b}{a} \right) d\tau, \quad (2)$$

де ΔT – інтервал дискретності, a, b – параметри масштабу і часового зсуву відповідно, $\psi(t)$ – материнська вейвлет-функція, оператор * позначає комплексне спряження.

На рис. 2 виділені ділянки сигналу з характерною рисою типу «пик» і ділянки вейвлет-спектра, що їм відповідають. Патерни типу «пик» мають високу частоту і тому відповідні їм ділянки вейвлет-спектра розташовані в області малих значень масштабного коефіцієнта a , який визначає величину розтягання вейвлета.

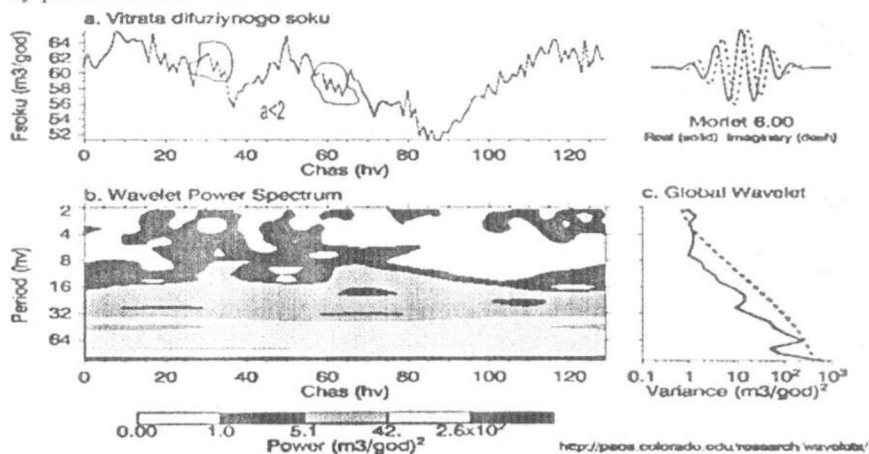


Рис. 2. Аналізований сигнал витрати дифузійного соку та його вейвлет-спектр: а) вихідні дані; б) вейвлет-спектр; с) глобальний вейвлет-спектр

Очевидно, що використання вейвлет-спектра може істотно спростити пошук піків у вхідних даних.

Підвищення наочності може бути досягнуте шляхом побудови ліній рівня вейвлет-спектра. Математично це означає, що при певному рівні здійснюється перетин поверхні в тривимірному просторі, що являє вейвлет-перетворення, площиною. У результаті формуються лінії рівня – ізолінії.

Іншим найважливішим типом патерна, який необхідно визначати в досліджуваних часових рядах, є повільна хвиля. Використання різних материнських вейвлетів показує, що найкращі результати по виявленню повільних хвиль досягаються при використанні Симлет-вейвлета 4-го порядку.

На рис. 3 наведено результати визначення повільних хвиль.

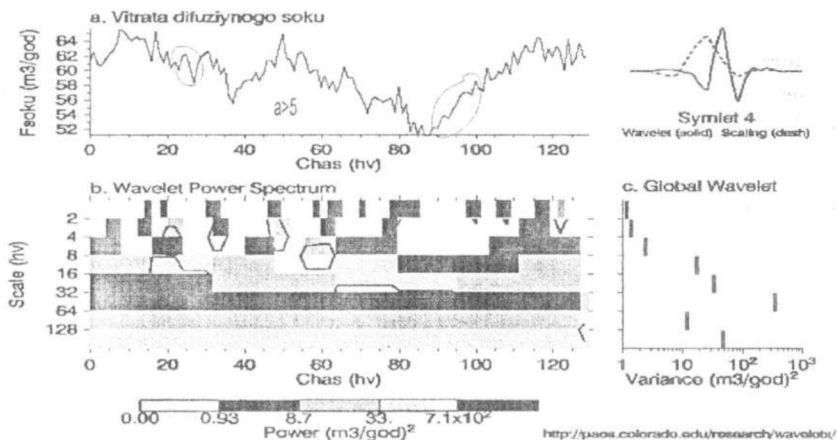


Рис. 3. Виявлення патернів типу «повільна хвиля» за допомогою Симлет-вейвлета: а) вихідні дані; б) вейвлет-спектр; в) глобальний вейвлет-спектр

Як видно з рис. 3, запропонований у роботі підхід до виявлення патернів і певні параметри вейвлет-спектрів дозволяють здійснювати часову локалізацію патернів і обчислювати їхню кількість на заданому часовому інтервалі.

Висновки. У даній роботі показана принципова можливість ефективного розпізнавання патернів типу «пік» і «повільна хвиля» за вейвлет-спектрами часових рядів. Також показано, що перехід від вейвлет-спектра

до ізоліній, які представляють собою лінії перетину поверхні в тривимірному просторі, що зображує вейвлет-спектр, площинами, дозволяє виявити зазначені патерни, здійснити їхню часову локалізацію і визначити кількість. Крім того, були визначені чисельні значення масштабних коефіцієнтів, необхідні для автоматизації розпізнавання кожного з аналізованих патернів. Подальшим напрямком роботи є аналіз інших патернів, таких як «гострі хвилі», а також розробка спеціального програмного забезпечення для обробки часових рядів на основі вейвлет-перетворення.

Бібліографічні посилання

1. **Витязев В.В.** Вейвлет-анализ временных рядов : учеб. пос. / В.В. Витязев. – СПб., 2001. – 58 с.
2. **Мандель И.Д.** Кластерный анализ. / И. Д. Мандель – М., 1988. – 176 с.
3. **Рябинин. И. А.** Надёжность и безопасности структурно-сложных систем / И. А. Рябинин. – СПб., 2000. – 351 с.
4. **Jolliffe J.** Principal Component Analysis. –New York: Springer-Verlag, 1986.– 487 p.
5. **Mead A.** Review of the development of multidimensional scaling methods // The Statistician. – San Francisco: Wiley, 1992. – № 33. – P. 27–39.
6. **Kohonen T.** Self-Organizing Maps. – New York: Springer-Verlag, 1997. – 378 p.

Надійшла до редколегії 10.11.10