

Д. В. Марчук
І. А. Дудатьєв

ОБРОБКА BIG DATA ДАНИХ, ЩО НАДХОДЯТЬ З СЕНСОРІВ АНАЛІЗУ ЯКОСТІ ВОДИ

Вінницький національний технічний університет

Анотація

У роботі розглянуто підходи до обробки великих даних (Big Data), що надходять із сенсорних систем моніторингу якості води в режимі реального часу. Представлено основні фізико-хімічні параметри води, які контролюються сенсорами, зокрема температуру, рН, концентрацію розчиненого кисню, залишкового хлору та рівень мінералізації. Описано принципи фільтрації даних за допомогою фільтра Калмана, а також використання методів машинного навчання (SVM, CNN, LSTM) для виявлення аномалій і прогнозування змін показників якості води. Підкреслено важливість інтеграції IoT-пристроїв, хмарних сервісів і аналітики Big Data для формування єдиної інтелектуальної системи моніторингу стану водопровідної мережі.

Ключові слова: Big Data, сенсорні мережі, IoT, якість води, фільтр Калмана, машинне навчання, моніторинг у реальному часі, прогнозування.

Abstract

The paper examines approaches to processing Big Data obtained from real-time water quality monitoring sensor systems. It presents key physico-chemical parameters of water controlled by sensors, including temperature, pH, dissolved oxygen, residual chlorine, and total mineralization level. The study describes data filtering using the Kalman filter and the application of machine learning models (SVM, CNN, LSTM) for anomaly detection and water quality prediction. The importance of integrating IoT devices, cloud services, and Big Data analytics is emphasized to create an intelligent monitoring system for urban water supply networks.

Keywords: Big Data, sensor networks, IoT, water quality, Kalman filter, machine learning, real-time monitoring, prediction.

Вступ

Сучасні системи моніторингу стану довкілля дедалі частіше використовують інтелектуальні сенсорні технології, які дозволяють здійснювати безперервний контроль параметрів середовища у режимі реального часу. Особливої актуальності такі системи набувають у сфері водопостачання, де якість води є критично важливим показником для здоров'я населення та екологічної безпеки. Традиційні лабораторні методи аналізу не забезпечують достатньої оперативності, тоді як інтеграція сенсорних мереж і технологій Інтернету речей (IoT) створює можливість постійного збору великого обсягу даних — Big Data.

Метою даного дослідження є аналіз підходів до обробки даних, отриманих із сенсорів контролю якості води, а також оцінка ефективності використання методів машинного навчання та фільтра Калмана для підвищення точності моніторингу.

Big Data дані з сенсорів аналізу якості води

Системи моніторингу якості води в реальному часі стають ключовим компонентом сучасних інтелектуальних інфраструктур водопостачання. Застосування сенсорних мереж дозволяє отримувати безперервні потоки даних про фізико-хімічні параметри води, що є критично важливим для забезпечення відповідності нормативним стандартам безпеки. На відміну від традиційних періодичних лабораторних аналізів, децентралізовані IoT-системи забезпечують багатоканальне вимірювання таких параметрів, як температура, концентрація розчиненого кисню (O₂), рН, рівень залишкового хлору та значення загальної мінералізації (PPM) у режимі реального часу, що формує значні масиви даних типу Big Data [1].

Температура (T) є фундаментальним параметром, вона впливає одразу на декілька чинників: на розчинність газів, на біологічну активність та швидкість хімічних реакцій у воді. Її зміна корелює з концентрацією кисню за законом Генрі:

$$C_{O_2} = k_H * P_{O_2} * e^{-aT},$$

де C_{O_2} – концентрація розчиненого кисню, k_H – константа Генрі, P_{O_2} – парціальний тиск кисню, a – температурний коефіцієнт.

Контроль рівня рН дозволяє оцінювати кислотно-лужний баланс. Його вимірювання базується на рівнянні Нернста:

$$E = E_0 - \frac{RT}{nF} \ln(a),$$

де E – потенціал електрода, E_0 – стандартний електродний потенціал, R – універсальна газова стала, T – температура, F – число Фарадея.

Концентрація хлору (Cl_2) є ключовим санітарним показником для дезінфекції водопровідної води. Узагальнений показник мінералізації (PPM) визначається як:

$$PPM = \sum_{i=1}^n C_i,$$

де C_i – концентрація розчинених іонів.

Обробка даних із сенсорів водопровідної води базується на багаторівневих інформаційних моделях. Для забезпечення достовірності їх вимірювань застосовуються фільтри згладжування (наприклад, фільтр Калмана):

$$x_{k|k} = x_{k|k-1} + K_k(z_k - Hx_{k|k-1}),$$

де $x_{k|k}$ – оцінка стану, K_k – коефіцієнт Калмана, z_k – вимірювання, H – матриця спостережень.

Для виявлення аномалій, прогнозування параметрів якості води та класифікації подій забруднення також застосовують засоби штучного інтелекту. Часто використовуються моделі SVM, згорткові мережі й рекурентні LSTM-архітектури для роботи з часовими рядами, що демонструють високу точність при аналізі нелінійних профілів якості води [2,3].

Системи реального часу повинні не лише реєструвати значення параметрів, але й виконувати інтелектуальну оцінку відхилень. Функція виявлення аномалій може записуватися як:

$$A(t) = \begin{cases} 1, & |x(t) - \hat{x}(t)| > \epsilon \\ 0, & \text{else} \end{cases},$$

де $x(t)$ – виміряне значення, $\hat{x}(t)$ – прогноз, ϵ – допустиме відхилення.

Прогнозування параметрів є критичним для своєчасного реагування на можливе погіршення якості води.

Застосування технологій Big Data у поєднанні з інтелектуальними сенсорними платформами забезпечує підвищену точність контролю параметрів водопровідної води в реальному часі. Можлива інтеграція IoT-сенсорів, сервісно-орієнтованих хмарних модулів та методів машинного навчання забезпечить формування єдиного цифрового середовища для безперервного збору даних, їх фільтрації, виявлення відхилень та формування прогнозів. Реалізація таких підходів дозволяє знизити ризики затримок у реагуванні на зміни якості води та забезпечити оперативну підтримку прийняття рішень на рівні міських систем водопостачання.

Подальший розвиток систем моніторингу орієнтований на підвищення метрологічної стабільності сенсорних модулів, удосконалення алгоритмів адаптивної фільтрації (зокрема методів на основі фільтра Калмана) та інтеграцію моделей прогнозової аналітики для оцінки стану водних систем у довготерміновій перспективі [4].

Висновки

У роботі проаналізовано, що застосування сенсорних систем моніторингу якості води у поєднанні з технологіями Big Data дозволяє забезпечити безперервний контроль ключових фізико-хімічних параметрів водного середовища. Використання методів фільтрації даних, зокрема фільтра Калмана, підвищує точність обробки вимірювань та зменшує вплив шумів сенсорних сигналів. Водночас застосування алгоритмів машинного навчання, таких як SVM, CNN та LSTM, створює можливість ефективного виявлення аномалій та прогнозування змін показників якості води. Інтеграція IoT-пристроїв, хмарних сервісів і аналітичних інструментів формує основу інтелектуальних систем моніторингу водопровідних мереж, що дозволяє підвищити оперативність реагування на можливі відхилення.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Durgun Y. Real-time water quality monitoring using AI-based sensors: pollutant detection and UV disinfection analysis in smart urban water systems. *Journal of King Saud University - Science*. 2024, V. 36. P. 1–8.
2. S. Kimothi, A. Thapliyal, S. V. Akram, R. Singh, A. Gehlot, H. G. Mohamed, D. Anand, M. Ibrahim, and I. D. Noya. Big Data Analysis Framework for Water Quality Indicators with Assimilation of IoT and ML. *Electronics*. 2022, V. 13(11), P. 1927.
3. Chen J., Chen S., Fu R., Li D., Jiang H., Wang C., Peng Y., Jia K., Hicks B. J. Remote sensing big data for water environment monitoring: Current status, challenges, and future prospects. *Earth's Future*. 2022, V. 10(2), P. 12-28
4. Zhang, J., Sheng, Y., Chen, W., Lin, H., Sun, G., & Guo, P. (2021). Design and Analysis of a Water Quality Monitoring Data Service Platform. *Computers, Materials & Continua*. 2021. V. 66(1). P. 389-405.

Марчук Данило Віталійович — аспірант групи G6-25a, факультет інформаційних електронних систем, Вінницький національний технічний університет, Вінниця, e-mail: danchi0120@gmail.com.

Науковий керівник: **Дудат'єв Ігор Андрійович** — канд. техн. наук, доцент кафедри IPTC, Вінницький національний технічний університет, м. Вінниця

Marchuk Danylo V. — Department of Information Electronic Systems, Vinnytsia National Technical University, Vinnytsia, e-mail: danchi0120@gmail.com.

Supervisor: **Dudatiev Ihor A.** — PhD in Technical Sciences, Associate Professor of the Department of IRTS, Vinnytsia National Technical University, Vinnytsia