

АЛГОРИТМ СЕГМЕНТАЦІЇ ДАНИХ ПРИ МОНІТОРИНГУ РІВНЯ ПАЛИВА

Іващев Д. В.¹, Герасимов В. В.²

¹Дніпровський національний університет імені Олеся Гончара, Україна

²Дніпровський національний університет імені Олеся Гончара, к. т. н., доцент,
Україна

Анотація. У статті представлено алгоритм автоматичної сегментації сигналів від датчиків рівня палива транспортних засобів, що базується на аналізі похідних та адаптивному визначенні моментів змін у поведінці сигналу. Метод дозволяє з достатньою точністю виділяти однорідні ділянки сигналу, що є основою для подальшого аналізу шумових характеристик. Алгоритм забезпечує прийнятний рівень точності та демонструє стійкість до короточасних аномалій і типових перешкод, які часто виникають у реальних умовах експлуатації. Проведене порівняння результатів сегментації з емпіричними оцінками підтвердило ефективність і надійність розробленого підходу. Запропоноване рішення є достатньо універсальним, що дозволяє застосовувати його до різних типів систем збору даних, включаючи умови підвищеної складності, характерні для роботи сучасних транспортних засобів.

Ключові слова: сегментація сигналу, рівень палива, похідна, аналіз даних, обробка сигналів.

Вступ

Сучасні системи контролю використовують різноманітні датчики для моніторингу даних та діагностики стану обладнання. Проте отримані дані часто зашумлені, що ускладнює їх аналіз. Тому важливо розробляти ефективні алгоритми попередньої обробки, зокрема сегментації та фільтрації сигналів. В нашому випадку сегментація необхідна для поділу масиву даних на ділянки роботи транспортного засобу, його простою та екіпірування (поповнення палива). У роботі запропоновано алгоритм сегментації даних з метою виділення різних режимів роботи системи для подальшого очищення сигналів і аналізу шуму. Це сприяє підвищенню ефективності моніторингу й надійності роботи паливного обладнання.

Постановка задачі

Перед початком аналізу даних, отриманих від дачика рівня палива транспортного засобу з певною періодичністю, важливо розділити ці дані на сегменти, що відрізняються рівнем шумового забруднення, тобто:

- Виділити сегменти, пов'язані з заправкою, які вважаються нешумними через стандартизовані умови експлуатації (двигун не працює).
- Визначити шумні сегменти, де високе динамічне навантаження та зовнішні фактори призводять до появи аномалій, викидів та шумів.

Огляд існуючих методів для розбиття даних на сегменти:

1. К-середніх (K-means): Розділяє дані на кластери шляхом мінімізації дисперсії всередині груп.
 - Плюси: Простота реалізації, ефективність на великих даних.
 - Мінуси: Чутливий до шуму та потребує попереднього завдання числа кластерів (K) [1].
2. Ієрархічна кластеризація: Будує древоподібну структуру кластерів, об'єднуючи або розділяючи групи.
 - Плюси: Не вимагає вказівки числа кластерів наперед, візуалізація через дендрограму.
 - Мінуси: Висока обчислювальна складність для великих наборів даних [2].

Через перелічені обмеження вказаних кластеризаційних алгоритмів, було прийнято рішення розробити власний алгоритм, орієнтований на аналіз структурних змін у сигналі рівня пального.

Було розроблено програму з комбінованим методом аналізу та виділення сегментів.

Етапи обробки даних в програмі

- 1) З масиву завантажених даних відкидаються рядки, де рівень пального менше межі, визначеної конструктивними особливостями паливного баку. Відфільтровані таким чином дані сортується часом.
- 2) Гауссове згладжування. Початковий сигнал рівня пального обробляється за допомогою гауссового фільтра. Це дозволяє усунути короточасний шум, забезпечуючи більш точне виявлення змін тренду [3].
- 3) Сегментація за зміною приросту. Сигнал розбивається на вікна фіксованого розміру. Для кожного вікна обчислюється середня різниця між сусідніми значеннями. На основі локального стандартного відхилення цих різниць встановлюється адаптивний поріг, перевищення якого сигналізує про ймовірну зміну тренду.
- 4) Аналіз третьої похідної. Для уточнення меж сегментів виконується аналіз третьої похідної сигналу поблизу точок, де виявлено порогове перевищення. Це дає змогу виявити точку перегину — місце зміни нахилу.

- 5) Визначення напрямку сегмента. Після знаходження точки перегину порівнюються кути нахилу до та після неї. Якщо різниця між ними перевищує заданий поріг, сегмент класифікується як перехід від спадного до висхідного тренду або навпаки.

Результати

Для аналізу роботи програми було обрано масив даних (11 наборів) в якості вхідних даних. З кожного набору було взято фіксовано кількість точок, що відповідає двом місяцям вимірювань (89280 точок). Вихідні дані та межі сегментів після обробки за вказаним алгоритмом одного з транспортних засобів наведено на рис. 1.

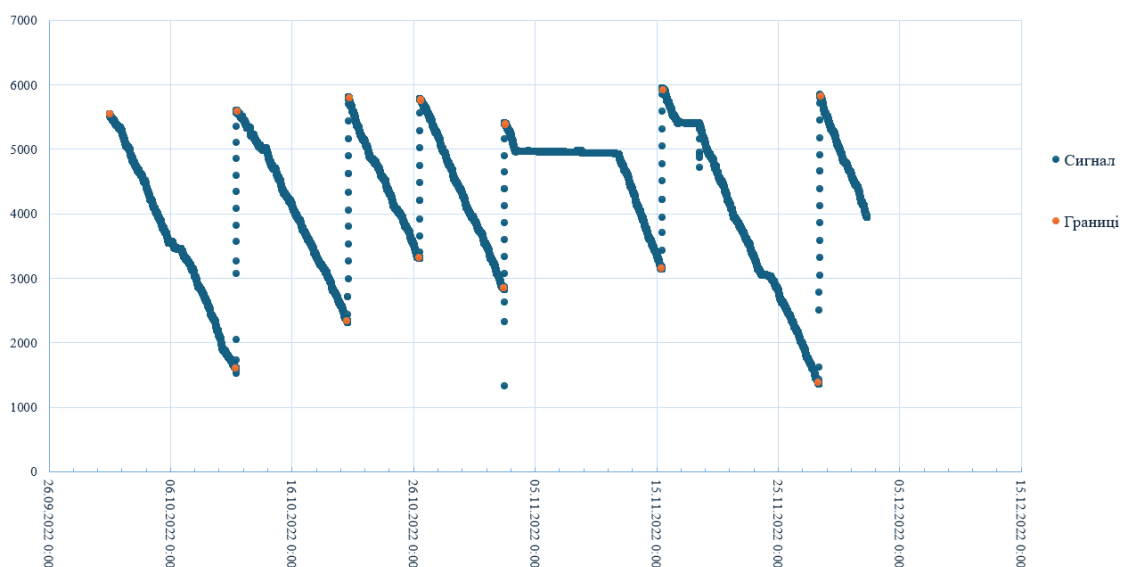


Рисунок 1 - Графік вихідного сигналу та границь сегментів для першого транспортного засобу

Після чого емпірично, шляхом візуального спостереження, масиви даних від всіх транспортних засобів теж були розбиті на сегменти.

Надалі було проведено порівняння сегментів, отриманих за допомогою запропонованого алгоритму, та вирахованих емпірично. В більшості випадків сегменти, які відповідають заправці паливом, визначаються "із запасом" на початку і в кінці, що для дослідження характеристик має позитивний ефект, так як дозволяє уникнути перехідних зон.

Висновки. Розроблений алгоритм сегментації масиву даних, отриманих від датчиків рівня палива транспортних засобів ефективно виконує сегментацію, використовуючи адаптивний поріг та третю похідну для точного

виявлення змін тренду. Алгоритм стійкий до більшості шумів, хоча незначна частина сегментів були класифіковані неправильно через різкі аномалії.

Загалом, розроблений алгоритм дозволяє досить ефективно розділяти дані на сегменти з шумовим забрудненням (стоянка та рух транспортного засобу) та без (заправка) для подальших досліджень.

ЛІТЕРАТУРА

- 1.MathWorks Documentation. k-Means Clustering. URL: https://de.mathworks.com/help/stats/k-means-clustering.html?spm=a2ty_o01.29997173.0.0.b7825c31fP2dip (дата звернення: 24.03.2025)
- 2.MathWorks Documentation. Hierarchical Clustering. URL: https://de.mathworks.com/help/stats/hierarchical-clustering.html?spm=a2ty_o01.29997173.0.0.b7825c31fP2dip (дата звернення: 25.03.2025)
- 3.MathWorks Documentation. Filtering and Smoothing Data. URL: https://de.mathworks.com/help/curvefit/smoothing-data.html?spm=a2ty_o01.29997173.0.0.b7825c31fP2dip (дата звернення: 25.03.2025)

DATA SEGMENTATION ALGORITHM FOR FUEL LEVEL MONITORING

Ivashchev D. V., Herasymov V. V.

Abstract. *The article presents an algorithm for the automatic segmentation of signals from vehicle fuel level sensors based on derivative analysis and adaptive detection of changes in signal behavior. The method allows for sufficiently accurate identification of homogeneous segments, which serves as a basis for further noise characteristic analysis. The algorithm ensures an acceptable level of accuracy and demonstrates robustness against short-term anomalies and typical interferences commonly encountered in real operating conditions. A comparison of the segmentation results with empirical estimates confirmed the effectiveness and reliability of the proposed approach. The solution is sufficiently universal, making it suitable for application across various types of data acquisition systems, including those operating under complex and demanding conditions typical of modern vehicles.*

Keywords: *signal segmentation, fuel level, derivative, data analysis, signal processing.*

REFERENCES

1. MathWorks Documentation. k-Means Clustering. Available at: <https://de.mathworks.com/help/stats/k-means-clustering.html> (accessed March 24, 2025).
2. MathWorks Documentation. Hierarchical Clustering. Available at: <https://de.mathworks.com/help/stats/hierarchical-clustering.html> (accessed March 25, 2025).
3. MathWorks Documentation. Filtering and Smoothing Data. Available at: <https://de.mathworks.com/help/curvefit/smoothing-data.html> (accessed March 25, 2025).