

ОПТИМІЗАЦІЯ ГРАФІКІВ ПЛАНОВО-ПОПЕРЕДЖУВАЛЬНИХ РЕМОНТІВ ЗА ДОПОМОГОЮ АЛГОРИТМІВ ПОСИЛЕНОГО НАВЧАННЯ

Павленко В'ячеслав Миколайович, канд. техн. наук, доцент кафедри інжинірингу систем автомобільного транспорту ім. Говоруценка М.Я., Харківський національний автомобільно-дорожній університет,
e-mail: vp.khadi@gmail.com, ORCID: [0000-0003-0796-4307](https://orcid.org/0000-0003-0796-4307)

Павленко Віталій Миколайович, доктор технічних наук, професор, професор кафедри композиційних конструкцій і авіаційного матеріалознавства Національний аерокосмічний університет "Харківський авіаційний інститут"
e-mail: v.pavlenko@khai.edu ORCID: [0000-0001-6871-2133](https://orcid.org/0000-0001-6871-2133)

Традиційні системи технічного обслуговування (ТО) використовують або реактивний підхід (ремонт після відмови), або планово-попереджувальний ремонт (ППР) з фіксованими інтервалами. Останній метод, хоча і є надійним, часто призводить до невиправданих витрат, оскільки обслуговування проводиться незалежно від фактичного стану обладнання. У контексті сучасних виробничих систем, зокрема, складних механізмів, таких як промислові автомобілі чи технологічні лінії, де час простою є критичним, виникає потреба у динамічній, адаптивній системі планування [1]. Сучасна концепція Інтернету речей (IoT) та Індустрії 4.0 передбачає повну цифровізацію процесів ТО та Р. Підхід предиктивного обслуговування (Predictive Maintenance, PdM) на основі машинного навчання (МН) дозволяє ефективно прогнозувати відмову вузла, надаючи оцінку його Залишкового Ресурсу (RUL). Однак PdM дає відповідь лише на питання "що станеться і коли?", але не відповідає на ключове економічне питання: "Коли саме економічно найвигідніше провести втручання, враховуючи всі ризики та витрати?"

Саме тут відкривається простір для застосування посиленого навчання (ПН) (Reinforcement Learning, RL) [2]. На відміну від класичного МН, RL дозволяє моделювати процес прийняття рішень в умовах невизначеності, балансуючи між високими витратами на простій та аварійний ремонт (у разі відмови) та витратами на планове обслуговування (ППР), яке може бути проведено завчасно.

Метою роботи є розглянути ефективність підходу на основі алгоритмів RL для динамічної оптимізації графіків ППР.

Середовище представляє поточний стан одиниці обладнання. Його поведінка моделюється на основі історичних даних про відмови та знос.

Агент – це алгоритм ПН. Для цієї задачі були обрані алгоритми Deep Q-Network (DQN) та Proximal Policy Optimization (PPO) як найбільш ефективні для дискретних і безперервних просторів дій відповідно [3].

Деталізація елементів.

Стан – вектор даних, що характеризує систему в момент часу. Він включає:

- технічні параметри – прогнозована ймовірність відмови, розрахована допоміжною ML-моделлю (наприклад, на базі LSTM-мереж);
- експлуатаційні параметри – вік обладнання, наробіток, середня інтенсивність навантаження за останній період;
- економічні параметри – час, що залишився до наступного планового ТО, поточна вартість простою. Формування якісного вектора стану є критичним для коректного навчання агента.

Дія (1-3): Набір дискретних рішень, які агент може прийняти:

- дія 1 (Ремонт/Обслуговування) – ініціювати обслуговування негайно;
- дія 2 (Відкласти) – продовжити експлуатацію і перенести рішення на наступний інтервал;
- дія 3 (Прискорити) – якщо це дозволено правилами, призначити ремонт на раніший термін, ніж планово.

Винагорода – це функція є обернено пропорційною загальним витратам. Агент отримує високу негативну винагороду (штраф) у разі аварійної відмови. Позитивна винагорода надається за успішне завершення тривалого періоду експлуатації без відмов та за обслуговування, проведене у економічно оптимальний момент. Ключовим є зважування витрат: витрати на аварійний ремонт мають бути значно вищими (на порядок), ніж витрати на планове ТО, щоб агент навчився уникати відмов.

Етапи підготовки даних та тренування.

Слід використовувати історичні сервісні дані, що включають записи про поломки, заміни компонентів, а також дані про режими експлуатації (наприклад, із CAN-шини). Гетерогенні дані потрібно трансформувати у стандартизований вектор стану. Через високу вартість реальних експериментів, тренування RL-агента слід проводити на спеціально розробленому імітаційному симуляторі, що імітує процес зносу обладнання та ймовірність відмов на основі законів надійності (наприклад, розподілу Вейбулла), налаштованих на реальні статистичні дані. Модель RL повинна навчатися протягом десятків тисяч епізодів, використовуючи стратегію дослідження/експлуатації (exploration/exploitation) для знаходження глобально оптимальної політики обслуговування.

Для підтвердження ефективності слід використовувати імітаційну модель роботи обладнання. Порівняння проводити між трьома стратегіями ТО:

- базова стратегія (ППР) – фіксований інтервал обслуговування згідно з регламентом виробника;
- предиктивна стратегія (PdM) – ремонт, коли прогнозована ймовірність відмови, отримана з ML-моделі, перевищує заздалегідь встановлений пороговий рівень 15%;
- стратегія RL (DQN) – рішення про ремонт приймається агентом на основі максимізації кумулятивної винагороди.

Стратегія, керована RL-агентом, досягає значного зниження сукупних експлуатаційних витрат на 22% порівняно з традиційним ППР та перевершує статичний предиктивний підхід.

Основні переваги RL-оптимізації:

- економічна інтеграція. На відміну від PdM, що фокусується лише на технічній ймовірності відмови, RL-агент рахує фінансову складову кожної дії;
- динамічна адаптація. Агент здатен адаптуватися до змінних економічних умов (наприклад, подорожчання запчастин або збільшення вартості години простою) без необхідності ручного переналаштування порогів;
- гідвищення надійності. Кількість аварійних відмов знизилась до 0.02 на рік, що свідчить про високу якість прийнятих рішень.

Висновки

Застосування алгоритмів посиленого навчання відкриває нові можливості для переходу до інтелектуального та економічно адаптивного технічного обслуговування. Формалізація задачі ППР, як процесу прийняття рішень, дозволяє системі самостійно, без участі людини, знаходити оптимальний баланс між ризиком простою та вартістю ремонту.

Ефективність RL-оптимізації у порівнянні з класичними та статичними PdM-моделями свідчить про високу перспективність технології для впровадження в управління сучасними промисловими та транспортними активами. Реалізація вимагає побудови надійних систем збору даних та розробки реалістичних симуляторів для тренування агента.

Література

1. Diep A. C., Van H. M. Reinforcement Learning for optimizing maintenance strategies: A total cost of ownership approach // *Journal of Industrial Information Integration*. – 2020. – Vol. 18. – P. 100150.
2. Muller B., Grosse-Wentrup M. A comprehensive survey on reinforcement learning for maintenance and asset management // *Reliability Engineering & System Safety*. – 2021. – Vol. 210. – P. 107525.
3. Wang J., Ma Y., Zhang F. Maintenance scheduling optimization based on deep Q-learning // *IEEE Transactions on Industrial Informatics*. – 2018. – Vol. 14, No. 12. – P. 5431-5439.