

Використання методологій глибокого навчання з підкріпленням в автономних мобільних кіберфізичних системах

Володимир Левицький, аспірант кафедри ІТ¹ (ORCID: 0000-0003-1829-488X), Світлана Білощицька, доктор технічних наук, професор кафедри ІТ¹, професор кафедри Обчислень та науки про дані² (ORCID: 0000-0002-0856-5474)

¹ Київський національний університет будівництва і архітектури, 03037, м. Київ, проспект Повітряних Сил, 31, Україна

² Університет ІТ Астана, проспект Мангилик Ел., 55/11, Астана, Казахстан

АНОТАЦІЯ

У цьому дослідженні пропонується вирішення проблеми, яка виникає під час розробки та експлуатації кіберфізичних систем (КФС) на основі можливостей штучного інтелекту (ШІ). Обговорюється сучасне використання технологій глибокого навчання в автономних мобільних кіберфізичних системах. Представляється узагальнена структура КФС та її проблеми. Результатом дослідження є тестування запропонованих алгоритмів глибокого навчання і глибокого навчання з підкріпленням для розділення мережі і керування ресурсами.

Ключові слова: глибоке навчання з підкріпленням (ГНП); кіберфізична система (КФС); навчання з підкріпленням (НП).

1. ВСТУП

Для ефективного керування автономною мобільною КФС необхідно забезпечити узгодженість часу між фізичними об'єктами та їхніми цифровими аналогами. Математичні моделі створюють цифрові копії, але моделювання цих моделей часто відстає від процесів у реальному часі, що призводить до ненадійних даних.

Іншим ключовим питанням є безпека програмного забезпечення, зокрема захист систем бездротового зв'язку від кібератак. Із збільшенням використання бездротових систем ризик атак зростає [1].

2. ПРИНЦИП ДІЇ КФС

Взаємодія між КФС та її середовищем є надзвичайно важливою. КФС функціонує у двох вимірах: кібервимірі, де існують цифрові елементи, такі як алгоритми та моделі, і фізичному вимірі, де відбуваються реальні процеси. Система повинна безперешкодно працювати в обох вимірах, щоб бути ефективною.

Для функціонування КФС потрібні три ключові пункти:

1. Вхідні дані – дані про навколишнє середовище, такі як температура, тиск або вологість.

2. Команди – інструкції користувача, що вказують на виконання певних завдань.

3. Поточний стан фізичних процесів – дані в реальному часі про фізичні операції, наприклад, положення роботизованої руки або швидкість конвеєра.

Система генерує два виходи:

1. Вихідні дані – інформація про дії системи.

2. Керуючі впливи – рішення та дії на основі вхідних даних і поточного стану фізичних процесів.

Ці входи та виходи дозволяють КФС взаємодіяти з оточенням і виконувати завдання ефективно [2, 3].

3. РІШЕННЯ

Навчання з підкріпленням (НП) — це підхід машинного навчання, при якому автономний агент навчається приймати

оптимальні рішення, взаємодіючи з навколишнім середовищем. Через ітеративний процес прийняття дій та отримання зворотного зв'язку, агент поступово вдосконалює свою стратегію, доки не знайде найкращу послідовність дій для досягнення мети.

Алгоритми НП різняться методами навчання, але всі мають спільну мету — поліпшити взаємодію агента із середовищем через "метод спроб і помилок". Агент коригує свої дії на основі спостережуваного або частково спостережуваного стану середовища.

ГНП виникає в поєднанні НП з глибоким навчанням, використовуючи нейронні мережі для аналізу великих обсягів даних і оптимізації рішень. ГНП успішно застосовується для вирішення складних завдань, таких як розробка ігор, керування роботами, і має потенціал змінити багато галузей [4].

Будуть розглянуті такі алгоритми ГНП: Q-learning, Deep Q-learning, Double Deep Q-learning, Dueling Deep Q-learning. Кожен алгоритм має своє призначення, алгоритм підбирається під задачу.

3.1. Особливості використання алгоритмів

Q-learning: табличний метод із використанням рівняння Беллмана. Використовується зазвичай в більш простих задачах.

DQN: використовує глибокі нейронні мережі для апроксимації Q-значень, що дозволяє масштабувати до великих просторів станів.

Double DQN: усуває упередження завищеної оцінки за допомогою двох мереж для вибору та оцінки дій.

Dueling DQN: відокремлює значення стану від переваги дії, підвищуючи ефективність навчання, коли деякі дії менш актуальні.

4. РЕЗУЛЬТАТИ МОДЕЛЮВАННЯ ТА АНАЛІЗ

4.1. Налаштування для моделювання

Були проведені симуляції для оцінки продуктивності запропонованих алгоритмів. Симуляції виконувалися на ПК з процесором Intel Core i7-7700HQ (2,6 ГГц) та графічним

процесором NVIDIA GeForce GTX 950. Алгоритми були реалізовані з використанням TensorFlow 2.16 та Python 3.12.

Сценарій включає один сервіс. Для моделювання реальних умов і демонстрації переваг алгоритму доступна смуга пропускання (W) на базовій станції (BS) встановлена на рівні 20 МГц, радіус дії BS становить 50 метрів, а кількість користувачів — 500. Порогові значення швидкості та затримки для розширеного мобільного широкосмугового зв'язку (eMBB) базується на угодах рівня обслуговування (SLA) для мереж 5G. Роздільна здатність смуги пропускання становить 0,1 МГц, щоб зберегти контрольований простір дій та покращити вибір дій. Щільність спектральної потужності шуму встановлена на рівні -174 дБм/Гц для врахування ефекту завмирання Релея.

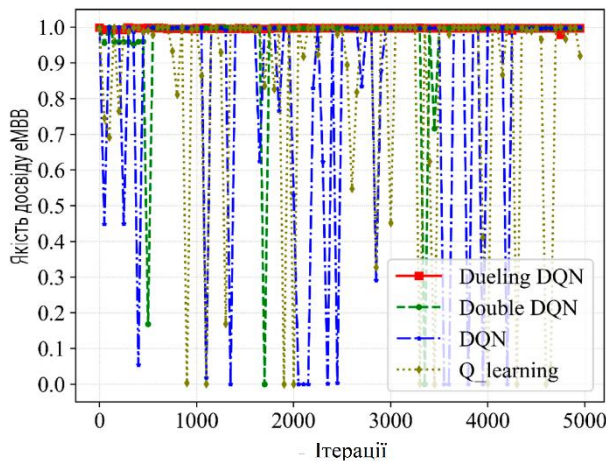


Рис. 1. Порівняльний результат якості досвіду користувачів для сервісу eMBB з різними алгоритмами

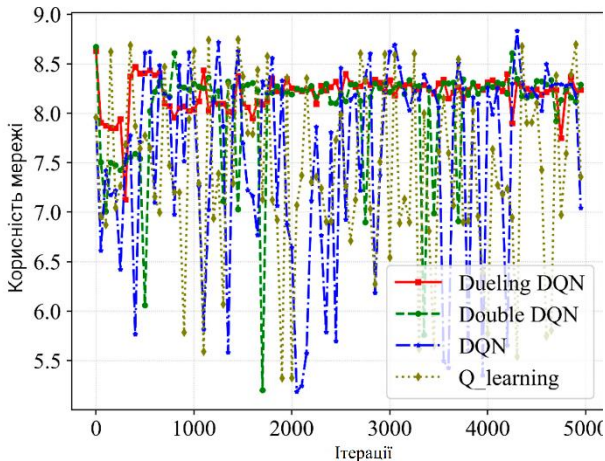


Рис. 2. Порівняльний результат корисності мережі з різними алгоритмами

4.2. Оцінка продуктивності

У цьому розділі представлений аналіз результатів симуляції запропонованих алгоритмів у контексті виділення та управління ресурсами. Q-learning використовується для ділення мережі, тоді як Double DQN вирішує проблеми з виділенням бездротових ресурсів.

Для всіх алгоритмів було застосовано однакові параметри навчання: швидкість навчання 0.01, коефіцієнт дисконтування 0.9 і мінімальний розмір пакету 32. Ваги для оптимізації якості досвіду користувачів (ЯДК) були

відрегульовані для забезпечення порівнюваності, що дозволяє проводити пряме порівняння корисності системи.

На рисунках 1-2 показано результати продуктивності якості досвіду, і корисності мережі для кожного алгоритму.

Після 1000 ітерацій Dueling DQN досягає майже 100% якості досвіду для сервісу eMBB. Double DQN демонструє дещо меншу стабільність, тоді як DQN і Q-learning мають значні коливання. Поліпшення продуктивності Dueling DQN пов'язане з більш ефективним розподілом смуги пропускання, що призводить до кращих результатів ЯДК.

На рисунку 2 показано тенденції корисності мережі зі збільшенням кількості ітерацій, де Dueling DQN покращує корисність на 11%, 8% і 2% у порівнянні з Q-learning, DQN і Double DQN відповідно.

Отже, Dueling DQN є найефективнішим серед досліджуваних алгоритмів, особливо в умовах високого навантаження та зростаючої кількості ітерацій.

5. ВИСНОВОК

Ця стаття надає огляд структури, функціонування та взаємодії КФС із навколишнім середовищем.

Висвітлюються ключові виклики в побудові та експлуатації КФС, показуючи, як методи штучного інтелекту можуть ефективно їх вирішувати. Стаття представляє узагальнену структуру КФС з модулем глибинного навчання та підкріплення. Також був проведений тест продуктивності алгоритмів ГНП.

Загалом, стаття має на меті поглибити розуміння КФС та підкреслити роль ШІ у вирішенні пов'язаних із ними викликів, що робить її цінним ресурсом для дослідників, інженерів і професіоналів у цій галузі.

Список літератури

- [1] Міллер А.О. Кіберфізичні системи: проблеми створення та напрями розвитку, Вісник Львівської Політехніки «Комп'ютерні системи та мережі», 2014, № 806, с. 154-161.
- [2] Wai-xi L., Jun C., Qing C. C., Yu W., DRL-R: Deep reinforcement learning approach for intelligent routing in software-defined data-center networks, Journal of Network and Computer Applications, Volume 177, 2021, sp 102865.
- [3] Arulkumaran, K., Deisenroth, MP, Brundage, M., and Bharath, 2017, Deep reinforcement learning: A brief survey, IEEE Signal Processing magazine, vol. 34 (6), pp. 26-38.
- [4] Timothy R., Yanzhi W., A survey for deep reinforcement learning in markovian cyber-physical systems: Common problems and solutions, Neural Networks, Volume 153, 2022, pp. 13-36.
- [5] Yuan, Y.; Li, Z.; Liu, Z.; Yang, Y.; Guan, X. Double Deep Q-Network Based Distributed Resource Matching Algorithm for D2D Communication. IEEE Trans. Veh. Technol. 2022, 71, 984-993.