

Використання технологій глибинного навчання з підкріпленням у автономних мобільних кіберфізичних системах

© Кіцила В. З., Бочкарьов О. Ю., 2020

Розглянуто проблему використання технологій глибинного навчання з підкріпленням у автономних мобільних кіберфізичних системах. Проаналізовані технології глибинного навчання використовуваних у автономних мобільних кіберфізичних системах. Запропоновано варіанти використання технологій глибинного навчання з підкріпленням у автономних мобільних кіберфізичних системах.

Ключові слова: глибинне навчання з підкріпленням, автономна мобільна кіберфізична система.

The problem of applying deep reinforcement learning technologies to autonomous mobile cyber-physical systems. Analyzed the technologies of deep learning used in autonomous mobile cyber physical systems. Options for the use of deep reinforcement learning technologies in autonomous mobile cyberphysical systems are proposed.

Keywords: deep reinforcement learning, autonomous mobile cyber-physical system.

Вступ

Інтенсивний розвиток інформаційних технологій та їх впровадження в усіх сферах людської діяльності супроводжується збільшенням автономності та складності кіберфізичних систем (КФС). Ріст популярності смартфонів та бездротового зв'язку підвищив розвиток такої області, як мобільні кіберфізичні системи [1-2].

Важливу роль у сфері інформаційних технологій відіграє швидкість обчислень та кількість ресурсів доступних для завдань обробки [3]. Один загальний механізм для швидкої реалізації вузлів автономної мобільної кіберфізичної системи на базі смартфонів використовує підключення до мережі, для зв'язку мобільної системи з сервером або хмарним середовищем, дозволяючи виконувати складні завдання обробки, які не є можливими в умовах обмежених локальних ресурсів.

Останнім часом, для поліпшення роботи всіх складових частин КФС і в тому числі мобільних КФС, все частіше використовуються методи штучного інтелекту (ШІ), зокрема алгоритми глибокого навчання (deep learning) [4].

Стан проблеми

В процесі проектування і експлуатації автономних мобільних кіберфізичних систем виникає ряд проблем. Однією з них є проблема узгодженості часу між фізичними об'єктами і їх цифровими двійниками [5]. Для управління об'єктами в реальному світі і передбачення їх

поведінки створюються цифрові копії в віртуальному просторі, на основі математичних моделей [6-8].

Час, необхідний для симуляції такої моделі, може відрізнятись від часу, з яким протікають реальні фізичні процеси в об'єкті що моделюється. Часто трапляється так, що математична модель настільки складна, що час комп'ютерної симуляції перевищує реальний.

Для прискорення обчислень, при збереженні якості симуляції, пропонується використання методів глибинного навчання або глибинного навчання з підкріпленням.

Ще однією важливою проблемою є забезпечення безпеки КФС від кібератак. Кількість систем, що використовують бездротовий зв'язок значно зросли, і ці системи піддаються кібератакам. Складність та динаміка кібератак вимагає, щоб захисні механізми були чуйними та адаптивними. Для вирішення цих проблем широко пропонується методи машинного навчання, а точніше глибинного навчання з підкріпленням [9].

Для кращої якості функціонування мобільними КФС необхідно мати якомога більше інформації про стан середовища, в якому функціонує така система, і точніше визначати процеси, що протікають в її складових частинах [10]. Сенсори і датчики служать органами почуттів для CPS.

Алгоритми глибинного навчання та глибинного навчання з підкріпленням підвищують органи чуття КФС, прикладом може виступати обробка жестів з допомогою wi-fi модуля.

Постановка задачі

Розглянути та проаналізувати використання технологій глибинного навчання у автономних мобільних кіберфізичних системах. Запропонувати варіанти використання технологій глибинного навчання з підкріпленням у автономних мобільних кіберфізичних системах.

Розв'язання задачі

Навчання з підкріпленням (reinforcement learning) – область машинного навчання, розглядає питання про те, як автономна інтелектуальна програма (названа агентом) спостерігає та діє в певному середовищі, навчаючись вибирати оптимальні дії для досягнення певної мети, визначеної на початку навчання. Здатність до навчання базується на багаторазовій взаємодії з оточенням [11]. Фактично агент і середовище утворюють систему зі зворотним зв'язком.

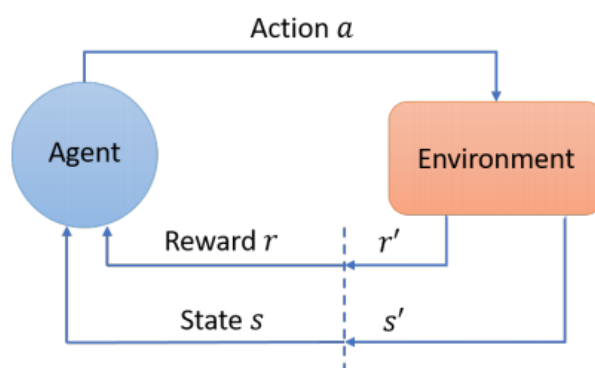


Рис. 1. Загальна структура взаємодії агента та середовища

Незважаючи на існування багатьох різних алгоритмів Навчання з підкріпленням, які різняться у конкретній реалізації навчальної функціональності, вони виконують ті самі дії у взаємодії агент-середовище, показані на рис. 1. Навчання з підкріпленням прийнято вважати “системою спроб та помилок”, агент намагається придумати найкращі дії з врахуванням стану середовища.

Глибинне навчання з підкріпленням (deep reinforcement learning) – це поєднання алгоритмів навчання з підкріпленням та методів глибинного навчання з використанням нейронних мереж. Цей тип навчання включає в себе комп’ютери, які діють на складних моделях та розглядають великі обсяги вхідних даних, щоб визначити оптимізований шлях чи дію.

Загальну класифікацію алгоритмів глибинного навчання з підкріпленням можна подати, як на рис. 2.

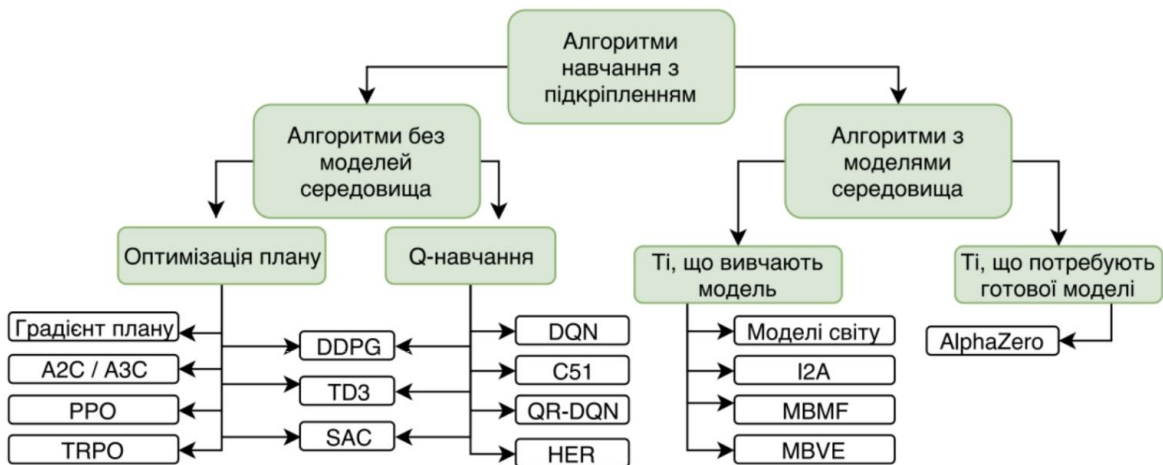


Рис. 2. Загальна класифікація алгоритмів DLR

Майже усі алгоритми DLR складаються з трьох послідовних дій для досягнення мети і мають таку структуру, рис. 3.



Рис. 3. Загальна структура алгоритмів DLR

Порівняння популярних методів RL та DLR

Популярний метод RL : q-learning метою якого є максимізувати сукупну винагороду зі знижкою на основі Рівняння Беллмана [12]:

$$Q(s_t, a_t) = E[r_{t+1} + \gamma r_{t+2} + \gamma^2 r_{t+3} + \dots | s_t, a_t] \quad (1)$$

Коефіцієнт знижки $\gamma \in [0, 1]$ керує рівнями важливості майбутніх винагород. Q-навчання потребує використання таблиці пошуку або Q-таблиці для зберігання очікуваної винагороди (Q-значення) при успішній дії з урахуванням набору станів. Це вимагає великої кількості пам'яті, а отже, Q-навчання не є ефективним. Альтернативою RL методу q-learning в DRL є використання DQN (deep Q – network) і АЗС алгоритму (actor-critic algorithm).

Asynchronous Advantage Actor Critic (АЗС) визначає, що є сервер, який збирає результати від безлічі акторів. І відновлює ваги як тільки набирається batch потрібного розміру, і не чекає результатів кожного актора. Це прибирає з прикладів непотрібну кореляцію, що покращує навчання.

Запропоновані варіанти використання технологій DRL в автономних кіберфізичних системах:

1. **Алгоритм Q-learning з LSTM** (нейронна мережа з довгою короткотривалою пам'яттю) – мобільна КФС моніторингу безпеки в автономних системах транспортних засобів [13].
2. **Алгоритм АЗС з RNN** (рекурентна нейронна мережа) – мобільна КФС виявлення аномалій в каналах зв'язку між ПЛК (програмованим логічним контролером), керуючим технологічними процесами на заводі з переробки, і системою інженерного моніторингу (SCADA).
3. **Алгоритм АЗС з Dyna-Q** -- мобільна КФС розрахування можливих маневрів Безпілотного літаючого апарату (БПЛА) з можливістю приймати рішення по ухиленню від зіткнення.
4. **Алгоритм Double dueling DQN** – Розумна міська мережа розподілу ресурсів.
5. **Алгоритм Q-learning з LSTM і DQN** – мобільна КФС управління мережею електроенергетичних станцій.
6. **Алгоритм DQN з LSTM** – мобільна КФС моніторингу дорожнього стану та виявлення дорожніх нерівностей.
7. **Алгоритм Q-learning and Dyna-Q** – мобільна КФС моніторингу бездротових мереж [14].

На основі наведених вище варіантів використання методів глибокого навчання з підкріпленням можна навести узагальнену структурну схему кіберфізичної системи з модулем глибокого навчання з підкріпленням. Але для початку слід зазначити як відбувається взаємодія КФС з оточенням і згадати основні компоненти КФС.

Взаємодія автономної кіберфізичної системи з оточенням відбувається одночасно у двох просторах: 1) у кіберпросторі, під яким розуміють штучні кібернетичні засоби і 2) фізичному просторі (середовищі), в якому знаходяться фізичні об'єкти та відбуваються відповідні фізичні процеси [15].

На вхід КФС надходять:

1. вхідні дані d_x з множини всіх можливих варіантів вхідних даних D_x та команди c_x з множини команд C_x ;
2. поточний стан фізичних процесів s_p з множини станів S_p .

На виході КФС отримують:

1. вихідні дані d_y з множини всіх можливих варіантів вихідних даних D_y ;
2. керівні впливи a_p з множини керівних впливів A_p .

На рис. 4. наведено узагальнену структуру КФС та її взаємодію з оточенням [1, 15].

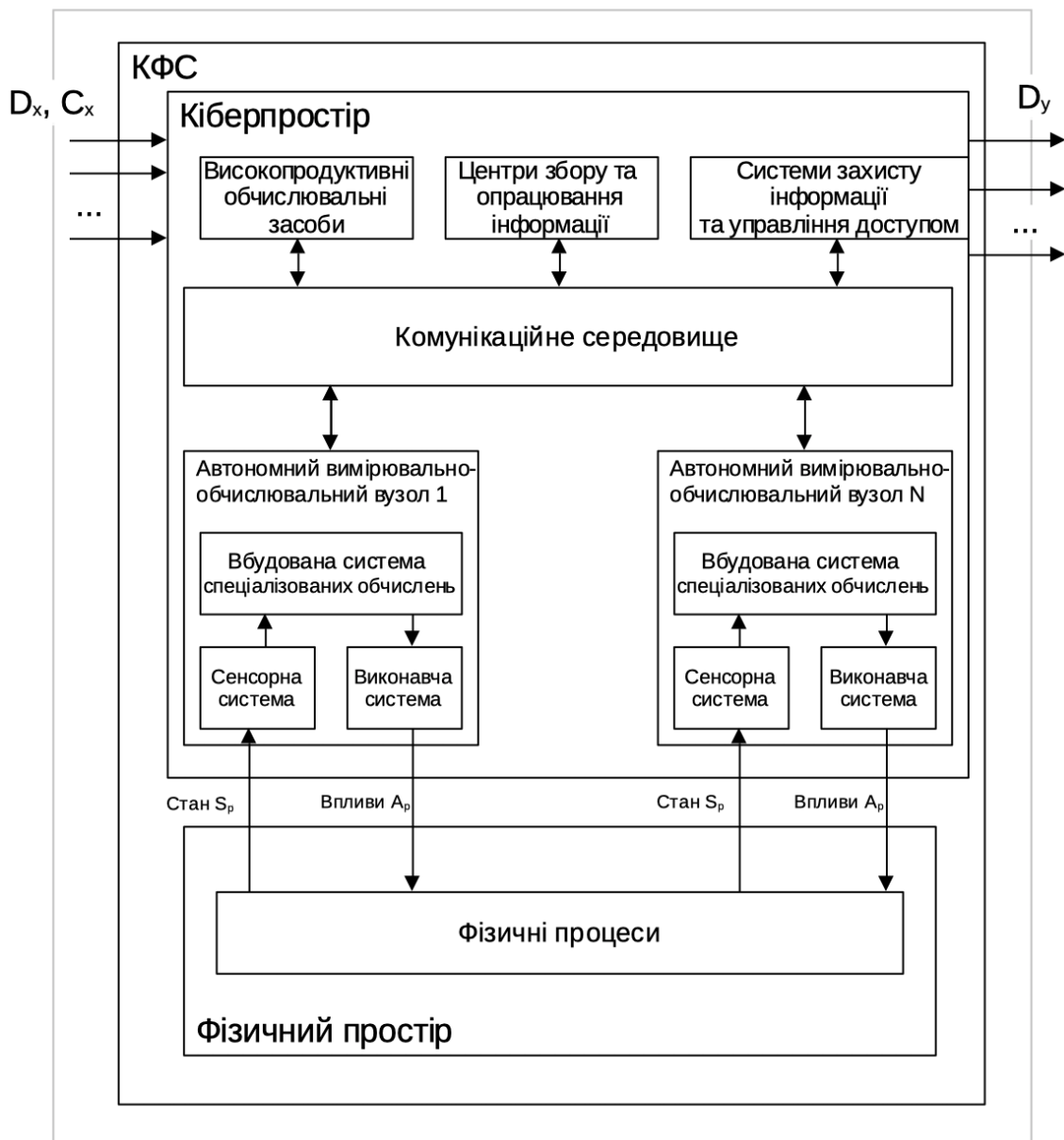


Рис. 4. Узагальнена структура кіберфізичної системи та її взаємодія з навколишнім середовищем

Тепер ми можемо навести узагальнену структурну схему КФС з модулем DRL. Наведемо її на рис. 5.

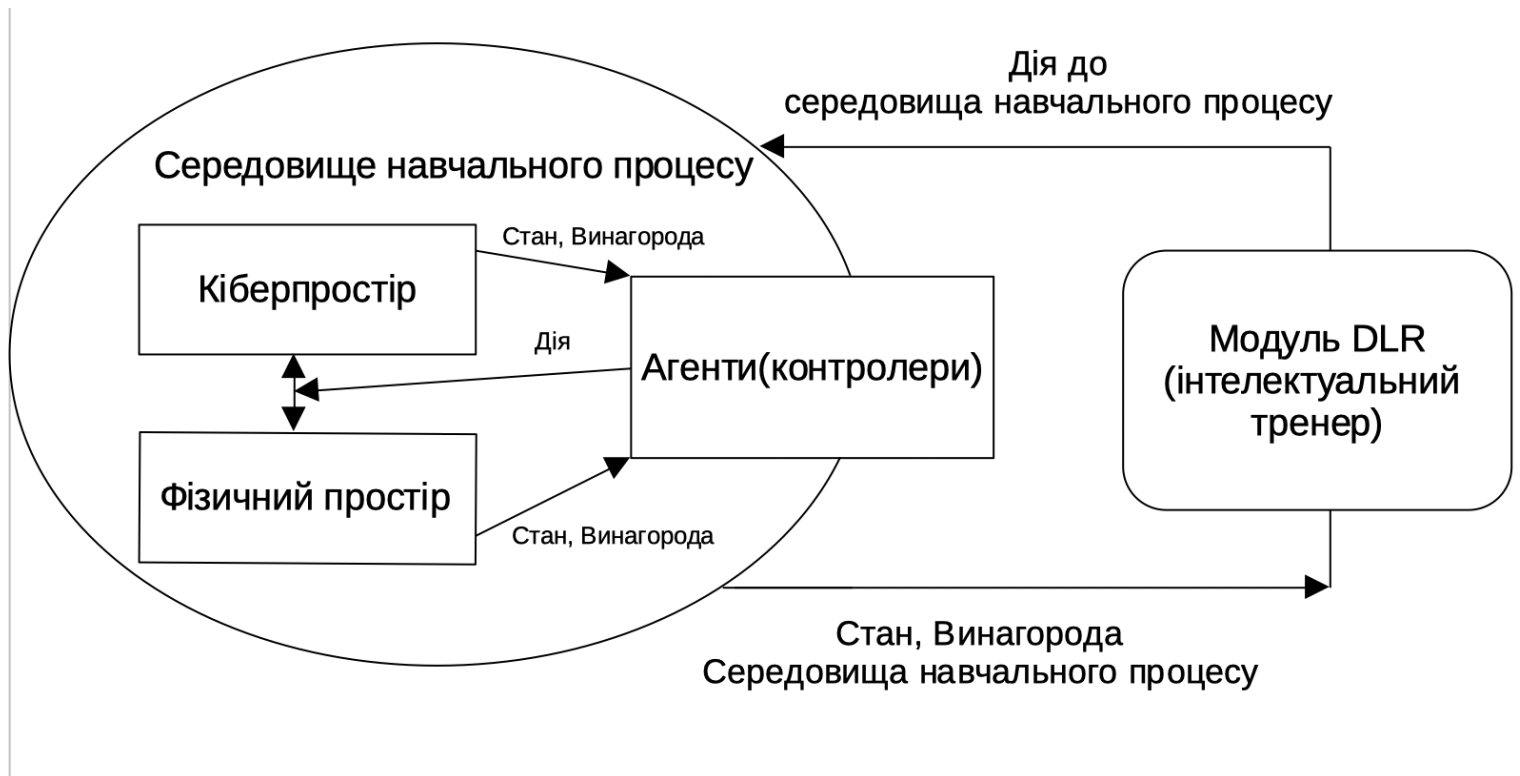


Рис. 5. Узагальнена структурна схема КФС з модулем DLR

Висновки

У даній роботі наведено узагальнену структуру кіберфізичної системи та її взаємодію з навколишнім середовищем. Наведено загальну класифікацію алгоритмів DRL. Розглянуто використання технологій глибокого навчання у автономних мобільних кіберфізичних системах. Розглянуто проблеми які виникають при побудові та функціонуванні КФС та показано як ці проблеми можна вирішити з використанням штучного інтелекту. Запропоновано варіанти використання технологій глибокого навчання з підкріпленням у автономних мобільних КФС. А також наведено узагальнену структурну схему КФС з модулем глибокого навчання з підкріпленням.

Література

1. Мельник А.О. Кіберфізичні системи: проблеми створення та напрямки розвитку // Вісник НУ «Львівська політехніка» «Комп'ютерні системи та мережі», 2014, №806. - С.154-161.
2. Jules White et al. R&D challenges and solutions for mobile cyber-physical applications and supporting Internet services // Journal of Internet Services and Applications, Volume 1, Number 1, May 2010. – P. 45–56.
3. Мельник А.О. Інтеграція рівнів кіберфізичної системи // Вісник Національного університету "Львівська політехніка". Комп'ютерні системи та мережі. - 2015. - № 830. - С. 61-67.

4. Кіберфізичні системи: багаторівнева організація та проектування / А. О. Мельник, В. А. Мельник, В. С. Глухов, А. М. Сало, за ред. А. О. Мельника. – Львів: "Магнолія 2006", 2019. – 237 с.
5. Lee, Edward A. "The past, present and future of cyber-physical systems: a focus on models" *Sensors (Basel, Switzerland)* vol. 15, 3 4837–69. 26 Feb. 2015
6. R. G. Sanfelice. Analysis and Design of Cyber-Physical Systems. A Hybrid Control Systems Approach // *Cyber-Physical Systems: From Theory to Practice* / D. Rawat, J. Rodrigues, I. Stojmenovic. — CRC Press, 2016
7. Shao Z., Liu J. (2013) Spatio-temporal Hybrid Automata for Cyber-Physical Systems. In: Liu Z., Woodcock J., Zhu H. (eds) *Theoretical Aspects of Computing — ICTAC 2013. ICTAC 2013. Lecture Notes in Computer Science*, vol 8049. Springer, Berlin, Heidelberg
8. E. A. Lee and S. A. Seshia, *Introduction to Embedded Systems — A Cyber-Physical Systems Approach*, <http://LeeSeshia.org>, 2011.
9. T. T. Nguyen and V. J. Reddi, "Deep Reinforcement Learning for Cyber Security" ArXiv. 2019.
10. Кіберфізичні системи: технології збору даних / О.Ю. Бочкар'юв, В.А. Голембо, Я.С. Парамуд, В.О. Яцук. За ред. А.О. Мельника, Львів: "Магнолія 2006", 2019. - 176 с.
11. Sutton RS, Barto AG (1998) *Reinforcement Learning: An Introduction*. MIT press, Cambridge.
12. Arulkumaran, K., Deisenroth, M. P., Brundage, M., and Bharath, A. A. (2017). Deep reinforcement learning: A brief survey. *IEEE Signal Processing Magazine*, 34 (6), 26-38.
13. Chong Li, Meikang Qiu, *Reinforcement Learning for Cyber-Physical Systems*, Chapman and Hall/CRC, 2019. - 256 p.
14. Li M, Li H (2020) Application of deep neural network and deep reinforcement learning in wireless communication. *PLoS ONE* 15 (7): e0235447.
15. Голембо В. А., Бочкар'юв О. Ю. Підходи до побудови концептуальних моделей кіберфізичних систем // *Вісник Національного університету "Львівська політехніка". Комп'ютерні системи та мережі*. – 2018. – № 881. – С. 15–24.