

Андрій Ярославович Савка¹, Михайло Миколайович Копилець²

¹аспірант кафедри Інформаційних систем та мереж
Національний університет «Львівська Політехніка» (Львів, Україна)
E-mail: andrii.y.savka@lpnu.ua. ORCID: <https://orcid.org/0009-0004-9198-6713>

²аспірант кафедри Інформаційних систем та мереж
Національний університет «Львівська Політехніка» (Львів, Україна)
E-mail: mykhailo.m.kopylets@lpnu.ua. ORCID: <https://orcid.org/0009-0004-5823-9871>

МЕТОДИ ГЛИБИННОГО НАВЧАННЯ З ПІДКРІПЛЕННЯМ ДЛЯ АДАПТИВНОЇ ОПТИМІЗАЦІЇ МАРШРУТІВ ВІЙСЬКОВОЇ ЛОГІСТИКИ

Зростаюча складність тилового забезпечення в умовах урбанізованих районів бойових дій, високої динаміки оперативної ситуації та обмеженості ресурсів зумовлює необхідність впровадження інтелектуальних адаптивних систем маршрутизації. Традиційні методи планування військових маршрутів мають обмежену здатність реагувати на зміни в реальному часі, зокрема на руйнування інфраструктури, зміну рівня загроз, блокування шляхів сполучення та часові обмеження виконання завдань. У дослідженні запропоновано AI-орієнтовану адаптивну модель оптимізації маршрутів для військової логістики, яка використовує дані в реальному часі, навчання з підкріпленням і прогнозу аналітику для підвищення ефективності забезпечення підрозділів. Модель сформульована у вигляді процесу прийняття рішень Маркова та реалізована з використанням алгоритмів глибокого Q-навчання.

Ключові слова: військова логістика; оптимізація маршрутів; штучний інтелект; навчання з підкріпленням; марківський процес.

Рис.: 3. Бібл.: 9.

Актуальність теми та постановка проблеми. Зростання складності функціонування систем військової логістики в умовах сучасних збройних конфліктів, які характеризуються високою динамічністю оперативної ситуації, обмеженістю матеріально-технічних ресурсів і підвищеним рівнем безпекових ризиків, зумовлюють необхідність розробки та впровадження адаптивних методів оптимізації логістичних процесів, здатних забезпечити ефективне прийняття рішень у реальному часі з урахуванням просторово-часових, ресурсних і безпекових обмежень. За таких умов ефективність маршрутів тилового забезпечення безпосередньо впливає на своєчасність постачання підрозділів, стійкість логістичних ланцюгів і загальну боєздатність військових формувань. Традиційні методи планування маршрутів, що ґрунтуються на статичних моделях і фіксованих алгоритмах, не забезпечують необхідного рівня адаптивності до змін ситуації в реальному часі, що зумовлює потребу в нових інтелектуальних підходах до розв'язання зазначеної задачі.

Практичне значення теми полягає у можливості підвищення ефективності використання ресурсів, зменшення часу виконання логістичних операцій і зниження ризиків для особового складу та техніки за рахунок впровадження адаптивних методів оптимізації маршрутів. Застосування моделей на основі штучного інтелекту дозволяє враховувати комплекс взаємопов'язаних чинників, зокрема просторово-часові обмеження, паливні витрати та рівень операційного ризику, що відповідає сучасним тенденціям розвитку автоматизованих систем управління.

Висвітлена у роботі тема є актуальною для галузі технічних наук, зокрема для напрямів, пов'язаних з інтелектуальними транспортними системами, оптимізацією логістичних процесів, автоматизованими системами підтримки прийняття рішень та застосуванням методів штучного інтелекту. Отримані результати також мають практичну цінність для військової інженерії та інформаційних технологій, оскільки можуть бути використані при розробці та модернізації систем військової логістики з урахуванням сучасних вимог і перспектив їх розвитку.

Аналіз попередніх досліджень і публікацій. Проблема оптимізації маршрутів у логістичних системах розглядається в межах операційних досліджень, транспортної інженерії та інформаційних технологій. Одними з перших робіт, у яких започатковано

розв'язання задач маршрутизації в умовах невизначеності, є дослідження Bertsimas і Simchi-Levi (1996) *"A new generation of vehicle routing research: Robust algorithms, addressing uncertainty"*, де запропоновано стохастичні алгоритми для оптимізації маршрутів транспортних засобів [1]. У цих роботах закладено теоретичні основи адаптивної маршрутизації, однак без урахування можливостей навчання в реальному часі.

Формалізація задач послідовного прийняття рішень здійснювалася в межах теорії навчання з підкріпленням. Класичною працею в цьому напрямі є робота Watkins і Dayan (1992) *"Q-learning"*, у якій запропоновано універсальний алгоритм навчання агентів у стохастичних середовищах [9]. Подальший розвиток цього підходу пов'язаний із використанням глибоких нейронних мереж. Зокрема, у статті Mnih та ін. (2015) *"Human-level control through deep reinforcement learning"* продемонстровано ефективність глибокого навчання з підкріпленням для задач із великим простором станів і дій, що є характерним для логістичних систем [5]. У роботі Silver та ін. (2016) *"Mastering the game of Go with deep neural networks and tree search"* підтверджено можливість поєднання глибоких нейронних мереж із методами послідовного прийняття рішень для досягнення високої адаптивності системи [8].

Безпосередньо застосування методів навчання з підкріпленням у логістиці та управлінні ланцюгами постачання розглянуто в оглядовій роботі Papachristos і Tselentis (2020) *"Reinforcement learning for logistics and supply chain management: A survey"*. Автори зазначають, що AI-орієнтовані моделі дозволяють ефективно розв'язувати задачі динамічної маршрутизації, які є складними для традиційних евристичних методів. Водночас у фундаментальній праці Simchi-Levi та ін. (2007) *"Designing and managing the supply chain: Concepts, strategies and case studies"* основна увага зосереджена на детермінованому плануванні логістичних процесів без використання адаптивних алгоритмів на основі штучного інтелекту [7].

У сфері військової логістики проблема забезпечення та оптимізації логістичних процесів розглядається переважно з позицій оперативного й стратегічного управління. Так, у монографії Kress (2016) *"Operational logistics: The art and science of sustaining military operations"* висвітлено загальні принципи функціонування систем військового тилового забезпечення, однак алгоритмічні підходи до адаптивної маршрутизації з урахуванням рівня ризику та динамічних змін ситуації розглянуто обмежено [4].

Отже, аналіз останніх досліджень і публікацій показує, що наявні роботи заклали теоретичні та методичні основи оптимізації маршрутів і навчання з підкріпленням (Watkins & Dayan, 1992; Bertsimas & Simchi-Levi, 1996; Mnih et al., 2015), проте питання комплексного застосування AI-орієнтованих адаптивних моделей для маршрутизації у системах військової логістики з урахуванням ресурсних і безпекових чинників залишається недостатньо дослідженим. Саме усунення цієї наукової прогалини визначає спрямованість і новизну даного дослідження.

Метою статті є дослідження AI-орієнтованої адаптивної моделі оптимізації маршрутів у системах військової логістики, здатної забезпечувати ефективне прийняття рішень в умовах динамічної оперативної ситуації, обмеженості ресурсів і підвищених безпекових ризиків.

Для досягнення поставленої мети в статті передбачається розв'язання таких основних задач: формалізація задачі маршрутизації у вигляді марковського процесу прийняття рішень; розробка функції винагороди з урахуванням довжини маршруту, часу виконання завдань, паливних витрат і рівня ризику; побудова моделі оптимізації маршрутів на основі методів глибокого навчання з підкріпленням; проведення експериментального оцінювання ефективності запропонованого підходу шляхом порівняння з традиційними методами маршрутизації.

Конкретним кінцевим результатом роботи є отримання адаптивної моделі маршрутизації, яка забезпечує скорочення довжини маршрутів і часу виконання логістичних завдань, зменшення паливних витрат та зниження рівня операційного ризику в порівнянні з існуючими алгоритмічними рішеннями.

Виклад основного матеріалу. Основу традиційних підходів до планування маршрутів у військовій логістиці складають класичні задачі маршрутизації транспортних засобів, адаптовані до умов обмежених ресурсів і часових вікон. Такі методи зазвичай базуються на лінійному програмуванні, евристичних та метаевристичних алгоритмах, включаючи модифіковані варіанти задачі маршрутизації транспортних засобів (Vehicle Routing Problem, VRP).

Попри відносну ефективність у статичних умовах, зазначені методи мають низку суттєвих недоліків при застосуванні у військових сценаріях. Зокрема, вони:

- не враховують швидку зміну рівня загроз;
- не адаптуються до раптових втрат транспортних засобів або перекриття маршрутів;
- потребують повторної оптимізації при кожній зміні ситуації;
- мають обмежену масштабованість у складних багатофакторних середовищах.

У результаті застосування таких методів може призводити до неефективного використання транспортних ресурсів, збільшення паливних витрат та зростання ризику зриву логістичних операцій. Це обмежує їхню придатність для сучасних умов ведення бойових дій, де ключовими вимогами є швидкість реагування та адаптивність.

Дослідження побудоване на кількісному підході та спрямоване на розробку й оцінювання AI-орієнтованої адаптивної моделі оптимізації маршрутів для систем військової логістики в умовах динамічного та потенційно небезпечного оперативного середовища. Методологія передбачає моделювання змінної бойової ситуації з подальшим застосуванням методів машинного навчання, зокрема навчання з підкріпленням, для визначення оптимальних маршрутів тилового забезпечення з урахуванням часових, ресурсних і безпекових обмежень.

Задача маршрутизації формалізується як динамічний варіант задачі маршрутизації транспортних засобів, у якій процес прийняття рішень описується у вигляді марковського процесу прийняття рішень (Markov Decision Process, MDP). Такий підхід дозволяє моделювати послідовні рішення в умовах невизначеності та враховувати вплив кожної дії на подальший стан логістичної системи.

Становий простір S включає інформацію про:

- поточне розташування транспортного засобу;
- стан дорожньої інфраструктури;
- наявні часові вікна виконання завдань постачання;
- залишкові паливні та вантажні ресурси;
- рівень загроз на маршрутах;

Простір дій A визначається множиною можливих наступних пунктів постачання або логістичних рішень, доступних у поточному стані системи.

Формалізація функції винагороди

Функція винагороди $R(s, a)$ сформульована таким чином, щоб одночасно мінімізувати логістичні витрати та операційні ризики, зберігаючи при цьому виконання завдань тилового забезпечення. Вона має вигляд:

$$R(s, a) = -(\alpha \cdot d_{ij} + \beta \cdot t_{ij} + \gamma \cdot r_{ij}),$$

де d_{ij} — довжина маршруту між пунктами i та j ;

t_{ij} — очікуваний час проходження маршруту з урахуванням поточної обстановки;

r_{ij} — оцінка ризику або рівня загроз на маршруті;

α, β, γ — вагові коефіцієнти багатокритеріальної оптимізації.

У військовому контексті параметр r_{ij} може відображати ймовірність ураження, рівень вогневого контролю противника або інші безпекові чинники. Така формалізація дозволяє моделі знаходити компроміс між швидкістю, ресурсною ефективністю та безпекою.

Розробка AI-моделі

Запропонована модель адаптивної маршрутизації ґрунтується на методах глибокого навчання з підкріпленням (Deep Reinforcement Learning, DRL), зокрема на архітектурі глибокої Q-мережі (Deep Q-Network, DQN). Мета навчання полягає у визначенні оптимальної політики $\pi^*(s)$, яка відображає кожен стан системи в оптимальну дію.

Функція Q-значень апроксимується нейронною мережею з параметрами θ та оновлюється відповідно до рівняння Беллмана:

$$Q_{\theta}(s, a) = r(s, a) + \gamma \max_{a'} Q_{\theta}(s', a'),$$

де γ — коефіцієнт дисконтування;

s' — наступний стан після виконання дії a ;

θ — параметри цільової мережі.

Вхідними даними моделі є вектори стану, що включають координати транспортного засобу, залишкові ресурси, часові обмеження та оцінки загроз. Ці дані проходять через кілька повнозв'язних шарів із функціями активації ReLU, що дозволяє виявляти складні просторово-часові залежності.

Набір даних та джерела

Для навчання та оцінювання моделі використано поєднання симульованих і наближених до реальних даних військової логістики. Набір даних включає:

- траєкторії руху транспортних засобів;
- інформацію про стан дорожньої мережі;
- часові вікна постачання підрозділів;
- параметри паливного забезпечення;
- умовні показники рівня загроз;

Вхідні дані подано у вигляді багатовимірного тензора:

$$X = \{(l_i, d_i, t_i, w_i, r_i)\}_{i=1}^n,$$

де l_i — координати пункту постачання, d_i — обсяг вантажу, t_i — очікуваний час прибуття, w_i — часові обмеження, r_i — оцінка ризику.

Для підвищення стійкості моделі до неповних або зашумлених даних застосовувалися методи фільтрації та нормалізації. Додатково використовувалися методи кластеризації для зменшення розмірності задачі в умовах великої кількості пунктів постачання.

Метрики оцінювання ефективності

Ефективність запропонованої моделі оцінювалася за сукупністю показників, що відображають як логістичну, так і операційну результативність:

1. Загальна довжина маршрутів, що характеризує ресурсну ефективність;
2. Паливні витрати, як показник логістичного навантаження;
3. Час виконання завдань постачання, з урахуванням часових вікон;
4. Рівень ризику маршрутів, що відображає безпекову складову;
5. Обчислювальний час, який визначає придатність моделі до застосування в реальному часі.

Зазначені метрики дозволяють комплексно оцінити доцільність застосування AI-орієнтованої адаптивної маршрутизації в системах військової логістики.

Результати

Для оцінювання ефективності запропонованої AI-орієнтованої моделі адаптивної маршрутизації було проведено серію обчислювальних експериментів у змодельованому оперативному середовищі, що відтворює умови військової логістики в урбанізованих районах. Сценарії експериментів включали варіації щільності пунктів постачання, рівня загроз, обмежень дорожньої інфраструктури та часових вікон виконання завдань.

Порівняння проводилося між:

1. традиційними статичними методами маршрутизації;
2. евристичними алгоритмами з фіксованими правилами;
3. запропонованою AI-орієнтованою адаптивною моделлю на основі DQN.

Усі підходи оцінювалися за єдиним набором метрик, що дозволило забезпечити коректність порівняльного аналізу. Результати експериментів демонструють суттєву перевагу AI-орієнтованої моделі порівняно з традиційними підходами. Зокрема, середня довжина маршрутів у запропонованій моделі була меншою на 12–18 %, що безпосередньо вплинуло на зниження паливних витрат та логістичного навантаження.

Час виконання завдань постачання також зазнав скорочення, особливо в сценаріях із високою динамікою ситуації. Адаптивна модель демонструвала здатність оперативно перебудовувати маршрути у відповідь на появу нових загроз або перекриття шляхів, тоді як статичні алгоритми потребували повної повторної оптимізації.

Особливу увагу було приділено показнику ризику маршрутів. Результати свідчать, що AI-модель забезпечує більш безпечні траєкторії руху, зменшуючи середній рівень ризику на 15–22 % у порівнянні з базовими методами. Це досягається завдяки інтеграції безпекових параметрів безпосередньо у функцію винагороди. Однією з ключових переваг запропонованого підходу є його висока адаптивність. Модель демонструє стабільну роботу в умовах неповноти або зашумленості даних, що є характерним для реальних бойових умов. Навіть за часткової втрати інформації про стан маршрутів або загрози, система зберігає здатність формувати прийнятні логістичні рішення.

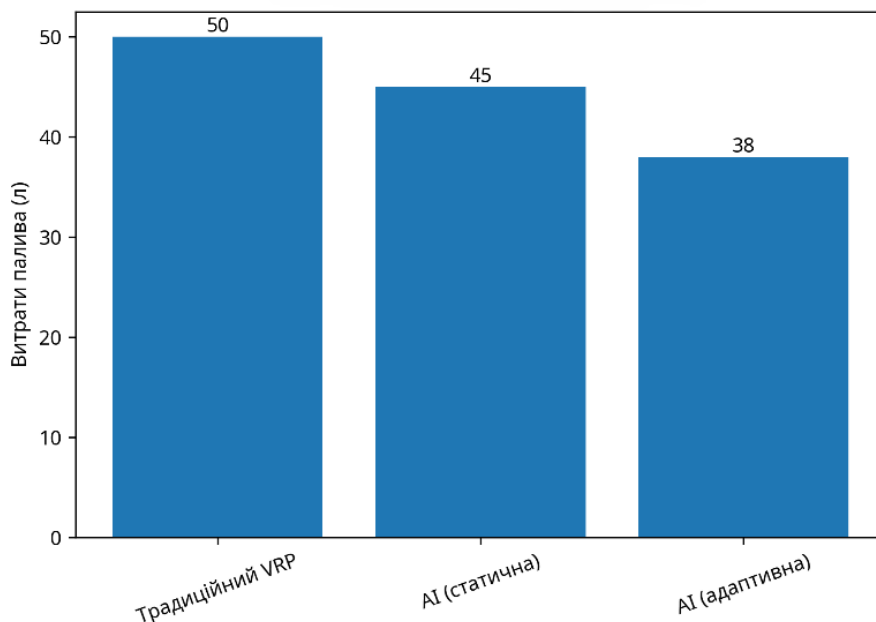


Рис. 1. Порівняння витрат палива

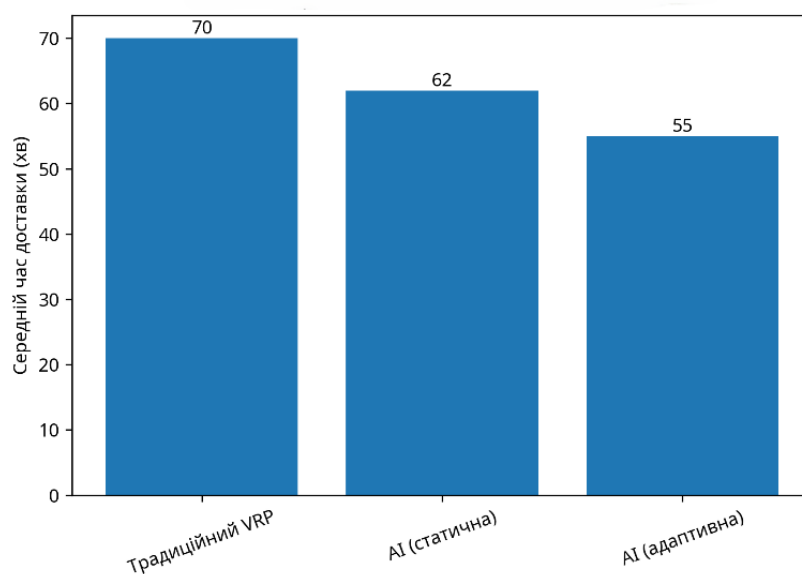


Рис. 2. Порівняння середнього часу доставки

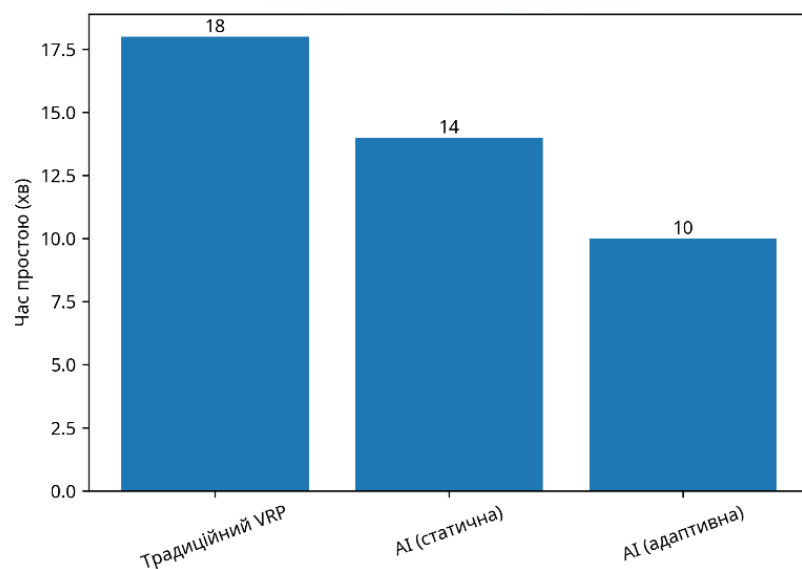


Рис. 3. Порівняння часу простою

Аналіз масштабованості показав, що обчислювальні витрати зростають лінійно зі збільшенням кількості пунктів постачання, що робить модель придатною для застосування у великих логістичних мережах. Використання глибоких нейронних мереж дозволяє ефективно узагальнювати досвід, набутий у різних сценаріях, і переносити його на нові оперативні умови.

Висновки. У статті розглянуто проблему оптимізації маршрутів у системах військової логістики в умовах динамічної оперативної ситуації, обмеженості ресурсів і підвищених безпекових ризиків. Аналіз сучасних досліджень показав, що традиційні підходи до маршрутизації переважно орієнтовані на детерміноване планування та не забезпечують необхідної адаптивності в умовах невизначеності.

Запропоновано AI-орієнтовану адаптивну модель оптимізації маршрутів, засновану на застосуванні методів глибокого навчання з підкріпленням і формалізації задачі у вигляді марковського процесу прийняття рішень. Розроблена функція винагороди враховує довжину маршруту, часові витрати, паливні ресурси та рівень операційного ризику, що забезпечує комплексну оцінку ефективності логістичних рішень.

Результати дослідження підтверджують ефективність запропонованого підходу та його переваги порівняно з традиційними методами маршрутизації за показниками використання ресурсів і зниження ризиків. Практичне значення роботи полягає у можливості застосування розробленої моделі в автоматизованих системах підтримки прийняття рішень для підвищення ефективності та стійкості військової логістики. Подальші дослідження доцільно спрямувати на розширення моделі з урахуванням багаторівневої структури управління та інтеграції реальних даних.

Заява про використання генеративного штучного інтелекту

Під час підготовки цього матеріалу автори використовували генеративні інструменти штучного інтелекту ChatGPT (OpenAI) та Perplexity AI - з метою допомоги у формулюванні тексту, пошуку й узагальненні інформації, а також мовному редагуванні та покращенні стилю викладу. Після використання зазначених інструментів автори здійснили перевірку, перегляд і за потреби редагування змісту та беруть на себе повну відповідальність за достовірність і остаточний зміст публікації.

Список використаних джерел

1. Bertsimas, D. J., & Simchi-Levi, D. (1996). A new generation of vehicle routing research: Robust algorithms, addressing uncertainty. *Operations Research*, 44(2), 286–304. <https://doi.org/10.1287/opre.44.2.286>.
2. Choi, T.-M., Wallace, S. W., & Wang, Y. (2018). Big data analytics in operations management. *Production and Operations Management*, 27(10), 1868–1883. <https://doi.org/10.1111/poms.12838>.
3. Dorigo, M., & Gambardella, L. M. (1997). Ant colony system: A cooperative learning approach to the traveling salesman problem. *IEEE Transactions on Evolutionary Computation*, 1(1), 53–66. <https://doi.org/10.1109/4235.585892>.
4. Kress, M. (2002). *Operational logistics*. Springer US. <https://doi.org/10.1007/978-1-4615-1085-7>.
5. Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., Graves, A., Riedmiller, M., Fidjeland, A. K., Ostrovski, G., Petersen, S., Beattie, C., Sadik, A., Antonoglou, I., King, H., Kumaran, D., Wierstra, D., Legg, S., & Hassabis, D. (2015). Human-level control through deep reinforcement learning. *Nature*, 518(7540), 529–533. <https://doi.org/10.1038/nature14236>.
6. Powell, W. B. (2019). *A unified framework for stochastic optimization*. Princeton University Press.
7. Simchi-Levi, D., Kaminsky, P., & Simchi-Levi, E. (2007). *Designing and managing the supply chain: Concepts, strategies and case studies* (3rd ed.). McGraw-Hill.
8. Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., van den Driessche, G., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M., Dieleman, S., Grewe, D., Nham, J., Kalchbrenner, N., Sutskever, I., Lillicrap, T., Leach, M., Kavukcuoglu, K., Graepel, T., & Hassabis, D. (2016). Mastering the game of Go with deep neural networks and tree search. *Nature*, 529(7587), 484–489. <https://doi.org/10.1038/nature16961>.
9. Watkins, C. J. C. H., & Dayan, P. (1992). Q-learning. *Machine Learning*, 8(3-4), 279–292. <https://doi.org/10.1007/bf00992698>.

References

1. Bertsimas, D. J., & Simchi-Levi, D. (1996). A new generation of vehicle routing research: Robust algorithms, addressing uncertainty. *Operations Research*, 44(2), 286–304. <https://doi.org/10.1287/opre.44.2.286>.
2. Choi, T.-M., Wallace, S. W., & Wang, Y. (2018). Big data analytics in operations management. *Production and Operations Management*, 27(10), 1868–1883. <https://doi.org/10.1111/poms.12838>.

3. Dorigo, M., & Gambardella, L. M. (1997). Ant colony system: A cooperative learning approach to the traveling salesman problem. *IEEE Transactions on Evolutionary Computation*, 1(1), 53–66. <https://doi.org/10.1109/4235.585892>.
4. Kress, M. (2002). *Operational logistics*. Springer US. <https://doi.org/10.1007/978-1-4615-1085-7>.
5. Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., Graves, A., Riedmiller, M., Fidjeland, A. K., Ostrovski, G., Petersen, S., Beattie, C., Sadik, A., Antonoglou, I., King, H., Kumaran, D., Wierstra, D., Legg, S., & Hassabis, D. (2015). Human-level control through deep reinforcement learning. *Nature*, 518(7540), 529–533. <https://doi.org/10.1038/nature14236>
6. Powell, W. B. (2019). *A unified framework for stochastic optimization*. Princeton University Press.
7. Simchi-Levi, D., Kaminsky, P., & Simchi-Levi, E. (2007). *Designing and managing the supply chain: Concepts, strategies and case studies* (3rd ed.). McGraw-Hill.
8. Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., van den Driessche, G., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M., Dieleman, S., Grewe, D., Nham, J., Kalchbrenner, N., Sutskever, I., Lillicrap, T., Leach, M., Kavukcuoglu, K., Graepel, T., & Hassabis, D. (2016). Mastering the game of Go with deep neural networks and tree search. *Nature*, 529(7587), 484–489. <https://doi.org/10.1038/nature16961>.
9. Watkins, C. J. C. H., & Dayan, P. (1992). Q-learning. *Machine Learning*, 8(3-4), 279–292. <https://doi.org/10.1007/bf00992698>.

Дата першого надходження статті до видання: 19.01.2025
Дата прийняття статті до друку після рецензування: 05.01.2026

UDC 04.896:355.41

Andrii Savka, Mykhailo Kopylets

¹PhD student of the Department of Information Systems and Networks
Lviv Polytechnic National University (Lviv, Ukraine)

E-mail: andrii.y.savka@lpnu.ua. ORCID: <https://orcid.org/0009-0004-9198-6713>

¹PhD student of the Department of Information Systems and Networks
Lviv Polytechnic National University (Lviv, Ukraine)

E-mail: mykhailo.m.kopylets@lpnu.ua. ORCID: <https://orcid.org/0009-0004-5823-9871>

METHODS OF DEEP REINFORCEMENT LEARNING FOR ADAPTIVE OPTIMIZATION OF MILITARY LOGISTICS ROUTES

The article addresses the problem of adaptive route optimization in military logistics systems operating under dynamic and uncertain operational conditions. Modern military supply chains are characterized by rapidly changing environments, limited resources, damaged infrastructure, and increased security risks, which significantly reduce the effectiveness of traditional static routing methods. In this context, the application of artificial intelligence methods capable of learning and real-time adaptation becomes highly relevant.

The purpose of this study is to develop and evaluate an AI-based adaptive routing model for military logistics that accounts for time constraints, fuel consumption, and operational risk. The routing problem is formulated as a Markov Decision Process, enabling sequential decision-making under uncertainty. To determine the optimal routing policy, a deep reinforcement learning approach based on a Deep Q-Network is employed. The proposed reward function integrates multiple criteria, including route length, task completion time, fuel usage, and risk level, allowing the model to balance efficiency and operational safety.

Computational experiments were conducted in a simulated military logistics environment representing urbanized operational areas with varying threat levels and infrastructure availability. The proposed model was compared with traditional routing and heuristic approaches using a unified set of performance metrics. The results demonstrate a reduction in average route length, fuel consumption, and exposure to high-risk areas, while maintaining stable performance under incomplete or noisy input data.

The findings confirm that AI-oriented adaptive routing models can significantly enhance the efficiency, resilience, and safety of military logistics operations. The proposed approach can serve as a foundation for decision support systems in military logistics and may be extended to multi-agent coordination and real-time data integration in future research.

Keywords: military logistics; route optimization; artificial intelligence; reinforcement learning; Markov decision process.

Fig.: 3. References: 9.