

АВТОМАТИЗАЦІЯ ТЕХНОЛОГІЧНИХ ПРОЦЕСІВ

УДК 004.896

НОСОВ А. О., КОРЖИК М. В.*
Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»

ПРОАКТИВНЕ КЕРУВАННЯ МОБІЛЬНИМИ РОБОТАМИ У ДИНАМІЧНОМУ ВИРОБНИЧОМУ СЕРЕДОВИЩІ

Сучасне виробництво дедалі більше спирається на використання мобільних роботів для автоматизації логістичних і технологічних операцій. Проте динамічний характер виробничого середовища створює складні наукові проблеми для систем керування мобільними платформами, оскільки існуючі алгоритми навігації часто не забезпечують належного балансу між швидкістю виконання завдань, енергоефективністю та безпекою руху. Метою дослідження є узагальнення сучасних методів проактивного керування мобільними роботами на основі формалізації задачі через багатовимірний POMDP, впровадження генеративних моделей для прогнозування динаміки та багатокритеріальної оптимізації з адаптивним балансуванням цілей. Проведено комплексний аналіз наукових публікацій 2016–2025 років та використано методи системного аналізу для синтезу нової архітектури керування.

Виявлено, що оптимальне проактивне керування досягається через інтеграцію трьох ключових компонентів: прогнозування майбутніх станів середовища, багатокритеріального планування з урахуванням продуктивності, безпеки та енергоефективності, а також адаптивного балансування цілей у реальному часі. Запропоновано трирівневу архітектуру системи керування та розроблено математичну модель, яка формалізує процес адаптивного проактивного керування з урахуванням прогнозованої динаміки середовища. Запропонована архітектура створює теоретичну основу для розробки нового покоління автономних робототехнічних систем, здатних ефективно функціонувати у складних виробничих умовах.

Ключові слова: проактивне керування, мобільні роботи, POMDP, багатокритеріальна оптимізація, програмовані логічні контролери, динамічне виробниче середовище

DOI: 10.20535/2617-9741.4.2025.348822

* Corresponding author: korzhyk@kpi.ua

Received 05 June 2025; Accepted 04 September 2025

Постановка проблеми. Сучасне виробництво дедалі більше спирається на використання мобільних роботів для автоматизації логістичних і технологічних операцій. Проте динамічний характер виробничого середовища, де одночасно діють люди, інші роботи, транспортні засоби та змінюються завдання, створює низку складних наукових і практичних проблем для систем керування мобільними платформами.

Об'єктом дослідження є процеси автоматизованого керування промисловими робототехнічними системами в динамічному виробничому середовищі, де існуючі алгоритми часто не забезпечують належного балансу між швидкістю виконання завдань, енергоефективністю та безпекою руху. У результаті мобільні роботи витрачають надмірну кількість енергії, виконують навігаційні завдання занадто повільно і залишаються вразливими до небезпечних зближень чи зіткнень із рухомими об'єктами. Удосконалення алгоритми навігації та планування руху мобільних роботів у зазначеному сенсі є актуальною науково-прикладною проблемою.

Вирішення цієї проблеми має важливе значення для підвищення продуктивності та безпечності сучасних виробничих процесів, а також для розроблення інтелектуальних робототехнічних систем, здатних автономно працювати у складних і невизначених середовищах. Проблема ускладнюється тим, що традиційні однокритеріальні підходи не дозволяють одночасно оптимізувати всі критично важливі показники (час, енерговитрати, безпека), а реактивні алгоритми реагують на зміни вже постфактум, що підвищує ризик неефективних або навіть аварійних ситуацій.

Відтак, нагальною задачею є розроблення нових теоретичних і практичних підходів до проактивного керування мобільними роботами, які дозволять передбачати майбутні зміни у середовищі та заздалегідь коригувати поведінку роботів. Особливої уваги потребує формалізація таких задач через частково спостережувані марковські процеси прийняття рішень (POMDP) з урахуванням множинних критеріїв, а також інтеграція генеративних моделей і глибинного навчання з підкріпленням для прогнозу динаміки середовища та адаптивного вибору дій.

Таким чином, постає важливе наукове завдання – знайти методи, які б дозволили знизити енерговитрати, скоротити загальний час виконання завдань і мінімізувати ризики небезпечних ситуацій у динамічному виробничому середовищі, забезпечуючи водночас високу точність і автономність навігації мобільних роботів.

Аналіз попередніх досліджень свідчить, що загальною науковою проблемою є розробка надійних та адаптивних методів проактивного керування мобільними роботами у складному, динамічному виробничому середовищі з багатьма цілями та жорсткими обмеженнями безпеки. Відповідно до цього, у літературі сформовано кілька напрямів вирішення задач.

Одна з базових стратегій – формалізація задачі планування через багатокритеріальні POMDP та MDP. Так, у [1] запропоновано підхід Multi-Objective POMDPs для забезпечення робастного автономного керування, однак не вирішеною частиною наукової проблеми є врахування складних реалій динамічного виробництва та можливих неузгодженостей між різними критеріями. Провідні сучасні огляди [2, 3] демонструють розширення методів підкріплення для мультицільових задач, однак у більшості робіт бракує детального механізму адаптивної ваги критеріїв залежно від контексту та реального часу, що є необхідним для практичного впровадження у гнучкому виробництві.

Значна увага приділяється генеративним моделям та трансформерам для прогнозування станів середовища і врахування невизначеностей [3, 4]. Проте не вирішеною частиною наукової проблеми є масштабованість таких підходів, а також питання ресурсозатратності обчислень у реальних промислових системах, де існують обмеження на потужність обчислювальної техніки й затримки у керуванні.

Архітектурні рішення для промислових мобільних роботів також активно розвиваються. В роботі [5] розглядається впровадження програмованих логічних контролерів (PLC) як надійної основи для фізичного виконання, однак загальною науковою проблемою є забезпечення ефективного двостороннього зв'язку та безпечного поєднання високорівневих алгоритмів AI/ML із жорсткими вимогами індустріальних стандартів. Дотепер не повністю розв'язано питання модульної інтеграції, стандартизації протоколів обміну даними та формальної верифікації поведінки мобільних роботів у небезпечних сценаріях.

Вивчення безпечного навчання з підкріпленням (safe RL) [6] показало, що хоча існують механізми для додавання обмежень безпеки безпосередньо у процес навчання, не вирішеною частиною проблеми залишається формалізована гарантія дотримання цих обмежень при перенесенні агентів у реальні системи Simulation-to-Reality (Sim2Real).

Таким чином, загальною науковою проблемою є створення гнучких, безпечних, проактивних систем керування мобільними роботами з підтримкою багатокритеріальних цілей та інтеграцією з апаратними промисловими контролерами. Невирішеною частиною наукової проблеми є розробка комплексних архітектур, які забезпечують і високий рівень автономності, і сувору відповідність стандартам безпеки, з урахуванням реальних обмежень виробничого середовища.

Метою статті є дослідження й узагальнення сучасних методів проактивного керування мобільними роботами у динамічному виробничому середовищі на основі формалізації задачі через багатовимірний POMDP, впровадження генеративних і трансформерних моделей для прогнозування динаміки, а також багатокритеріальної оптимізації з адаптивним балансуванням цілей. Також приділено увагу програмованим логічним контролерам (PLC), як фундаменту для забезпечення надійності, функціональної безпеки та ефективної взаємодії із високорівневими інтелектуальними агентами.

Формалізація задачі

Проактивне керування мобільними роботами в умовах динамічного виробничого середовища потребує адекватної математичної постановки, яка враховує часткову невизначеність та множинність критеріїв. Задачу доцільно формалізувати як частково спостережуваний Марковський процес прийняття рішень (POMDP), розширений до багатокритеріального випадку. POMDP дозволяє моделювати ситуації, коли робот володіє неповною інформацією про стан середовища, а динаміка середовища носить імовірнісний характер. Включення множинних цілей (наприклад, продуктивність, безпека, енергоефективність) приводить до постановки багатоцільового Марковського процесу прийняття рішень MDP/POMDP, де функція винагороди є векторною. Це означає, що робот повинен балансувати кілька критеріїв одночасно під час вибору дій, досягаючи компромісу між ними. Зокрема, нещодавні роботи запропонували модель topological POMDP, яка

дозволяє включати додаткові цільові показники (наприклад, максимізацію безпеки чи плавності руху) разом із типовими метриками продуктивності (мінімізація часу або відстані) [1]. Така формалізація забезпечує більш робастне прийняття рішень в умовах невизначеності, оскільки реальні робототехнічні системи мають працювати в середовищах, де стан не повністю спостережуваний, а цілі можуть конфліктувати між собою. Відзначимо, що традиційні однокритеріальні підходи (з одним скалярним показником) часто неадекватні для складних задач – вони спрощують багатовимірну сутність проблеми і можуть давати субоптимальні результати [2]. Натомість багатокритеріальні POMDP/MDP моделі дозволяють врахувати компроміси між конфліктними цілями на етапі планування і залучити апарат навчання з підкріпленням для пошуку стратегій, оптимальних за [2].

Прогнозування динаміки середовища (генеративні моделі, трансформери)

Ключовою особливістю проактивного керування є передбачення змін у середовищі до того, як вони відбудуться, і відповідне завчасне коригування планів робота. У динамічному виробничому середовищі це може означати прогнозування траєкторій рухомих об'єктів (наприклад, людей, інших роботів, транспортних засобів), появи нових завдань або перепланування маршрутів при зміні обстановки. Сучасні підходи спираються на генеративні моделі для моделювання динаміки середовища. Зокрема, використовуються неймережеві моделі – світові моделі (world models), які навчаються представляти перехідні ймовірності середовища. Такі моделі можуть генерувати можливі майбутні стани або цілі послідовності станів на основі поточного спостереження, що дозволяє роботу передбачати наслідки своїх дій та еволюцію оточення [3]. Наприклад, у робототехнічних системах продемонстровано, що генерація послідовностей проміжних станів (явно у вигляді зображень або приховано в латентному просторі) дає змогу прогнозувати розвиток ситуації і планувати дії робота з урахуванням очікуваних змін [3–5].

Окремо варто виділити використання архітектур типу Transformer для прогнозування динаміки. Трансформери, відомі своїм механізмом самоуваги, відмінно захоплюють довгострокові залежності в послідовностях даних. Це знайшло відображення в плануванні траєкторій і моделюванні динаміки середовища на великому горизонті часу. Зокрема, нещодавні дослідження інтегрують трансформери як модель динаміки середовища у контур планування: трансформер навчається передбачати наступні стани середовища на основі історії, що дозволяє роботу планувати багатоходові дії [4, 6]. Оскільки чисто трансформерні моделі можуть бути обчислювально витратними на дуже довгих горизонтах, пропонуються гібридні підходи – наприклад, поєднання короткострокового планування з трансформерною динамічною моделлю та довгостроковою оцінкою через окрему модель цільової функції [4]. Експериментально показано, що такий підхід підвищує продуктивність і вибірковість планування у безперервних задачах керування в реальному часі.

На практиці прогнозування середовища також може включати домен-специфічні моделі. Наприклад, в задачах навігації роботів серед людей використовують спеціальні модулі прогнозування траєкторій людей (з урахуванням їх інтентів і намірів). Одне з досліджень для автономних мобільних роботів на виробництві пропонує враховувати майбутнє переміщення операторів: інтеграція модуля прогнозу руху людини дала змогу роботам здійснювати випереджувальні та безпечні маневри у співпраці з людьми на конвеєрі [6]. Таким чином, компонента прогнозування в архітектурі проактивного керування може бути реалізована через треновані генеративні моделі – варіаційний аутоенкодери (VAE), генеративні змагальні мережі (GAN), дифузійні моделі або через гібридні моделі на основі трансформерів, що прогнозують розвиток обстановки [3, 4]. Це забезпечує превентивну інформацію для планувальника робота, дозволяючи йому оптимізувати стратегію наперед, а не реагувати постфактум.

Багатокритеріальна оптимізація в системах навчання з підкріпленням

У проактивному керуванні мобільними роботами часто виникає потреба оптимізувати кілька показників одночасно – наприклад, мінімізувати час виконання завдання, забезпечуючи безпечну дистанцію до людей і економне споживання батареї. У результаті постає задача багатокритеріальної оптимізації, яка може бути вирішена методами навчання з підкріпленням (RL) через відповідне розширення постановки. Існує кілька принципових підходів до багатокритеріальної RL:

– скаляризація цілей: зведення векторної винагороди до однієї скалярної шляхом лінійної комбінації або іншої агрегації критеріїв. Робот оптимізує цю зведену мету як звичайну винагороду. Перевага методу – простота впровадження, але він вимагає апріорі заданих ваг для критеріїв. Неправильний вибір ваг може змістити рішення у небажану сторону, а фіксовані ваги не дозволяють охопити весь спектр компромісів [2, 7]. Таким чином, скаляризація часто спрощує постановку і може втрачати рішення, оптимальні за Парето, якщо ваги обрані негнучко;

– обмежене RL (Constrained RL): формулювання задачі як оптимізації основної мети за наявності жорстких обмежень на допоміжні показники. У такому підході вводять розділення на функцію винагороди (наприклад, продуктивність) та функцію «вартості» або штрафу (наприклад, небезпека/ризик), і вимагають, щоб очікуване значення вартості не перевищувало заданий поріг. Фактично, це постановка з обмеженнями для MDP, де шукається політика, що максимізує основний виграш при виконанні обмежень на середні або навіть найгірші Value at Risk (VaR), Conditional Value at Risk (CVaR) значення штрафів [6, 8]. Розв'язання таких задач часто здійснюється методом множників Лагранжа: алгоритм динамічно оновлює множник (вагу) при штрафі, намагаючись виконати обмеження під час навчання [6]. Це дозволяє автоматично перебалансовувати пріоритети між ціллю та обмеженням: якщо обмеження порушується, вага штрафу зростає, змушуючи політику бути обережнішою. Обмежене RL отримало розвиток у контексті безпечного навчання (safe RL), де критично важливо уникати небезпечних дій; воно знімає проблему ручного підбору коефіцієнтів в єдиній винагороді, розділяючи завдання на максимізацію нагороди та контроль ризику [6, 8];

– навчання Парето-оптимальних політик: даний підхід прагне знайти або наблизити множину політик, оптимальних за Парето, замість того щоб обирати одну конкретну компромісну політику. Існують методи, що дозволяють навчити набір стратегій – кожна для різних пріоритетів між цілями – за одне проходження навчання. Наприклад, сучасні алгоритми Multi-Objective RL можуть використовувати гіпермережі або багатоголові нейромережі, які на вході отримують параметри преференцій і генерують відповідну політику. Таким чином, одна узагальнена модель здатна представляти ціле сімейство рішень для різних комбінацій ваг [7, 9]. Як показано в роботі [7] методика Pareto Set Learning на основі гіпермереж змогла згенерувати широкий спектр політик, що щільно покривають Парето-фронт і перевершують попередні алгоритми за показниками різноманітності рішень та якості компромісів [7, 9]. Перевага підходу – гнучкість після навчання: можна обрати потрібну політику в залежності від актуальних вимог, не перенавчаючи агент щоразу наново.

У контексті проактивного керування мобільними роботами багатокритеріальні методи RL дозволяють системі знаходити збалансовані стратегії. Наприклад, політика може одночасно враховувати швидкість виконання завдання й ризик зіткнення: навчання з множинними критеріями унеможливорює досягнення максимальної продуктивності за рахунок неприйняттого підвищення ризику з боку робота. Сучасні дослідження підтверджують, що явне багатокритеріальне планування сприяє кращій пояснюваності та адаптивності рішень системи, оскільки інженери можуть налаштовувати пріоритети або обмеження у зрозумілих термінах (наприклад, «мінімізувати час, але щоб ризик аварії $\leq 1\%$ »), а агент – відповідно навчатися цим вимогам [2, 6, 7, 9].

Динамічне балансування критеріїв

У реальних виробничих сценаріях відносна важливість критеріїв може змінюватися з часом. Проактивна система керування повинна вміти підлаштовувати свою політику під такі зміни без повної зупинки і перенавчання. Динамічне зважування критеріїв (adaptive weighting) означає, що вагові коефіцієнти при цілях або обмеженнях можуть автоматично коригуватися на основі ситуаційного контексту або зовнішніх пріоритетів.

Один підхід до цього – вбудувати параметр преференцій безпосередньо у процес прийняття рішень агента. Зокрема, можна розширити простір станів або входів нейромережі політики, додаючи до них вектор ваг поточних цілей. У цьому разі агент навчається набору політик під різні ваги і може перемикаватися між ними, змінюючи вектор пріоритетів на вході. Наприклад, автори [9] запропонували розширений Марковський процес прийняття рішень зі змішаною спостережуваністю (MOMDP), де разом із станом враховується стан преференцій; агент досліджує не лише фізичний простір, а й простір можливих налаштувань ваг, що дозволяє йому адаптивно реагувати на зміну вимог до критеріїв. Такий підхід фактично реалізує єдину політику, параметризовану преференціями, яка покриває множину компромісів [8, 9].

Альтернативний або суміжний підхід – використання вторинного оптимізатора для ваг. Це означає, що сам процес визначення ваг при критеріях може розглядатися як задача ухвалення рішень. Методи на основі подвійного контру (як у згаданому методі множників Лагранжа) автоматично підлаштовують вагу штрафу, щоб дотримуватися обмежень [8]. У ширшому сенсі, можна використовувати окремий агент або правило, яке спостерігає за поточним виконанням кількох цілей і коригує вагові коефіцієнти у режимі реального часу. Наприклад, якщо робот почав надто ризикувати безпекою заради швидкості, система збільшить вагу критерію безпеки до відновлення балансу. Свіжі дослідження демонструють ефективність стратегій, де критик оцінює досягнення по кожному з завдань і рекомендує зміну ваг для уникнення домінування одних цілей над іншими [8]. Завдяки такому динамічному балансуванню, проактивна система керування залишається стійкою до зміни умов: наприклад, при аварійній ситуації пріоритет автоматично зміститься на безпеку, а після нормалізації обстановки – повернеться до продуктивності.

Математична модель адаптивного проактивного керування

Узагальнюючи розглянуті вище концепції багатокритеріальної оптимізації, прогнозування динаміки середовища та адаптивного балансування, пропонується комплексна математична модель проактивного керування мобільними роботами. Ця модель вперше об'єднує в єдиному математичному виразі проактивне планування, адаптивне багатокритеріальне прийняття рішень та прогнозування динаміки середовища, створюючи теоретичну основу для нового покоління систем керування мобільними роботами.

Формалізація оптимальної проактивної політики

Оптимальна проактивна політика керування мобільним роботом визначається як розв'язок наступної оптимізаційної задачі:

$$\pi_{pa}^*(s_t, \omega_t) = \arg \max_{\pi} \mathbb{E} \left[\sum_{k=0}^H \gamma^k \left(\sum_{i=1}^N \omega_{i,t+k}(\xi_t) \times r_i(s_{t+k}, a_{t+k}) \right) \middle| \hat{\mathbf{P}}(s_{t+1:t+H} | s_t, \theta) \right],$$

де π_{pa}^* – оптимальна проактивна політика, яка визначає розподіл ймовірностей вибору дій робота; s_t – поточний стан системи; ω_t – вектор адаптивних ваг критеріїв; H – горизонт планування; γ – коефіцієнт дисконтування; N – кількість критеріїв оптимізації; $\omega_{i,t+k}$ – адаптивна вага i -го критерію, що залежить від контексту ξ_t ; r_i – функція винагороди для i -го критерію; $\hat{\mathbf{P}}$ – прогнозований розподіл майбутніх станів середовища; θ – параметри генеративної моделі, що використовуються для прогнозування майбутніх станів середовища.

Запропонована модель представляє собою розширення класичної теорії марковських процесів прийняття рішень на випадок проактивного багатокритеріального керування в динамічному середовищі. На відміну від традиційних підходів, де робот реагує на поточний стан середовища, наша модель дозволяє роботу приймати рішення на основі прогнозу майбутнього розвитку ситуації. Ключовою особливістю є те, що політика π_{pa}^* не просто відображає стани в дії, а враховує два додаткові фактори: вектор адаптивних ваг ω_t та прогнозовану динаміку середовища $\hat{\mathbf{P}}$. Це дозволяє роботу не лише реагувати на поточну ситуацію, але й передбачати майбутні зміни та адаптувати свою поведінку відповідно до змінних пріоритетів. Математичне очікування $\mathbb{E}$ обчислюється за всіма можливими траєкторіями розвитку середовища, що робить систему робастною до невизначеності. Робот розглядає не одну детерміновану послідовність майбутніх станів, а цілий розподіл можливих сценаріїв, вибираючи дії, які забезпечують найкращий очікуваний результат.

Інтеграція проактивності в оптимізаційну задачу

Принципова новизна полягає у включенні прогнозованого розподілу майбутніх станів $\hat{\mathbf{P}}(s_{t+1:t+H} | s_t, \theta)$ безпосередньо в цільову функцію. Класичні формулювання MDP та POMDP використовують фіксовані ймовірності переходів $\mathbf{P}(s' | s, a)$, які описують лише наступний крок (тут a – дія, яку виконує робот через k кроків). Натомість наша модель прогнозує цілу траєкторію станів на горизонті H , що дозволяє роботу «заглянути в майбутнє» та оптимізувати свої дії з урахуванням довгострокових наслідків. Це особливо важливо в динамічних виробничих середовищах, де зміни відбуваються не миттєво, а розвиваються протягом часу. Наприклад, робот може передбачити, що через 10 секунд на його шляху з'явиться інший робот або людина, і завчасно скоригувати траєкторію, уникнувши необхідності екстреної зупинки.

Динамічна адаптація вагових коефіцієнтів

Унікальною особливістю моделі є функція $\omega_{i,t+k}(\xi_t)$, яка дозволяє вагам критеріїв змінюватися не тільки між різними епізодами роботи, але й всередині одного горизонту планування. Це революційний підхід порівняно з традиційними методами, де ваги або фіксовані, або змінюються лише між запусками системи. Контекстна залежність від ξ_t означає, що система може автоматично перебалансовувати пріоритети в реальному часі. Якщо контекст вказує на підвищену небезпеку (наприклад, велика кількість людей в робочій зоні), вага критерію безпеки $\omega_{2,t+k}$ автоматично зростає, а вага продуктивності $\omega_{1,t+k}$ зменшується. При цьому важливо, що ваги можуть змінюватися по-різному для різних моментів часу в межах горизонту планування:

$$\omega_{i,t}(\xi_t) \neq \omega_{i,t+1}(\xi_t) \neq \dots \neq \omega_{i,t+H}(\xi_t).$$

Це дозволяє моделювати складні сценарії, коли пріоритети змінюються нелінійно. Наприклад, робот може планувати швидкий рух на початку траєкторії (високий ω_1), але уповільнитися ближче до зони з людьми (високий ω_2), а потім знову прискоритися після проходження небезпечної ділянки.

Багатомасштабна оптимізація

Модель реалізує багатомасштабний підхід до оптимізації, поєднуючи тактичний рівень (адаптація ваг $\omega_{i,t+k}$ на кожному кроці горизонту), стратегічний рівень (оптимізація на всьому горизонті H з урахуванням довгострокових наслідків) та оперативний рівень (врахування поточного контексту ξ_t для негайної адаптації). Така багаторівнева структура дозволяє системі бути одночасно реактивною (швидко реагувати на зміни) та проактивною (планувати наперед). Математично це виражається через динамічну функцію адаптації ваг:

$$\omega_{i,t+k}(\xi_t) = \frac{\exp(\beta_i \cdot f_i(\xi_t, k))}{\sum_{j=1}^N \exp(\beta_j \cdot f_j(\xi_t, k))},$$

де $f_i(\xi_t, k)$ – функція важливості i -го критерію, яка залежить від контексту та прогнозованого часу; β – параметри чутливості для кожного критерію.

Ключова перевага – обчислювальна ефективність: замість однієї надскладної задачі система вирішує кілька простіших задач, адаптованих до свого масштабу. Це забезпечує баланс між швидкістю реакції та якістю планування.

Формалізація для гібридних систем керування

Концепція гібридної архітектури

Особливістю запропонованої моделі є її спеціальна розробка для реалізації в гібридних системах керування, які поєднують переваги штучного інтелекту з надійністю промислових контролерів. На відміну від суто теоретичних моделей, наша формалізація враховує практичні аспекти інтеграції високорівневого інтелектуального планування з низькорівневим детермінованим виконанням. Така архітектура дозволяє використовувати сильні сторони обох підходів: здатність AI до навчання, прогнозування та оптимізації складних багатокритеріальних задач поєднується з гарантованою надійністю, передбачуваністю та швидкодією промислових контролерів.

Гібридність системи проявляється не просто у паралельному існуванні двох підсистем, а в їх глибокій інтеграції з синергетичним ефектом. AI-компонент може приймати складні рішення на основі неповної інформації та прогнозів майбутнього, але його рішення завжди проходять через фільтр безпеки PLC. З іншого боку, PLC не просто виконує команди AI, а активно бере участь у процесі керування, надаючи зворотний зв'язок про фізичні обмеження та поточний стан виконавчих механізмів.

Переваги гібридної формалізації

Запропонована гібридна архітектура забезпечує унікальне поєднання властивостей, які важко або неможливо досягти в чисто централізованих або розподілених системах. Гарантована безпека досягається через те, що PLC завжди має можливість втрутитися та скоригувати або скасувати потенційно небезпечні дії. Це створює надійний захисний бар'єр, який неможливо обійти навіть при збоях в AI-модулі.

Інтелектуальна адаптивність забезпечується здатністю AI оптимізувати продуктивність в межах безпечного простору дій. AI може експериментувати з різними стратегіями, навчатися на досвіді та покращувати свою роботу з часом, але завжди в рамках, визначених PLC. Це дозволяє системі еволюціонувати та адаптуватися до нових ситуацій без компромісу в безпеці.

Масштабованість архітектури проявляється в можливості додавання нових AI-модулів або заміни існуючих без необхідності змінювати базовий рівень безпеки. Це критично важливо для промислових систем, де оновлення повинні відбуватися поступово та без зупинки виробництва. Сертифікованість PLC-рівня за промисловими стандартами забезпечує можливість використання системи в регульованих галузях.

Ефективність досягається через оптимальний розподіл обчислювальних ресурсів: прості операції виконуються швидким PLC, а складні обчислення делегуються потужним AI-серверам. Це дозволяє досягти високої продуктивності без надмірних витрат на обладнання.

Ця формалізація створює математичну основу для практичної реалізації проактивних систем керування, які поєднують найкращі властивості обох підходів: інтелектуальність та здатність до навчання від AI з надійністю та передбачуваністю промислових контролерів. Такий синтез відкриває нові можливості для створення по-справжньому автономних та безпечних робототехнічних систем для виробництва майбутнього.

Архітектура системи проактивного керування

Об'єднання розглянутих компонентів – прогнозування середовища, багатокритеріального планування та адаптивного балансування – потребує цілісної архітектури керування мобільним роботом. Оптимальна архітектура має бути ієрархічною та модульною, щоб забезпечити як високорівневе проактивне планування, так і надійне низькорівневе виконання команд. Можна виділити три основні рівні (підсистеми) такої архітектури:

– модуль прогнозування. Цей компонент відповідає за збір інформації про поточний стан середовища (через датчики робота, Internet of Things (IoT) пристрої тощо) та прогнозування майбутніх станів. Реалізація включає навчений генеративний модельний блок, що на основі історичних даних передбачає рух перешкод, зміну оточення або появу нових цілей. Прогнозування може здійснюватись з різним горизонтом: короткостроково для запобігання неминучим зіткненням, довгостроково – для передбачення заторів чи необхідності перепланування маршрутів. Важливо, що модуль надає планувальнику оцінку ймовірних сценаріїв майбутнього, часто у вигляді розподілу ймовірностей або кількох альтернативних траєкторій об'єктів [5]. Це дозволяє врахувати невизначеність: наприклад, якщо людина може піти вліво чи вправо, прогноз надасть обидва варіанти з певними ймовірностями;

– модуль планування і ухвалення рішень. Отримуючи як вхід поточний стан (або розподіл станів POMDP) і передбачені сценарії, планувальник обчислює оптимальну послідовність дій робота. На цьому рівні застосовуються алгоритми навчання з підкріпленням або пошуку в просторі станів, що враховують множинні критерії. Планувальник прагне знайти політику дій, яка максимізує сумарну вигоду за заданою багатоцільовою метою (або забезпечує компроміс згідно з поточними вагами критеріїв), зважаючи на передбачену еволюцію середовища [2, 7, 9];

– модуль виконання. Низовий рівень, відповідальний за безпосереднє керування рухом робота (або групи роботів) відповідно до спланованої траєкторії чи послідовності дій. Сюди входять контури керування приводами, датчики низького рівня, механізми запобігання зіткненням на апаратному рівні (наприклад, аварійна зупинка при непередбаченій перешкоді). Модуль виконання також забезпечує зворотний зв'язок: він повідомляє планувальнику про відхилення фактичного руху від плану, про неочікувані події (які могли не бути передбачені) та про завершення проміжних підзадач. Таким чином, виконання відбувається у замкненому циклі з плануванням: якщо реальність відрізняється від прогнозу, цикл Sense-Predict-Plan-Act повторюється, щоб скоригувати подальші дії [5];

У рамках цієї архітектури забезпечується багаторівнева адаптація: високорівневий планувальник проактивно оновлює стратегію, а низькорівневий контролер гарантує надійність фізичного пересування. Подібні ідеї закладені в сучасних промислових рішеннях. Зокрема, [5] запропонували ієрархічну розподілену архітектуру для автономних мобільних роботів у логістиці, що включає зовнішній рівень (координація і спостереження), рівень навігації і контролю робота, і спеціальні модулі на кшталт людино-орієнтованої навігації та динамічного вибору профілів руху. Ключовою особливістю їх системи є здатність працювати в середовищах з участю людей і інших роботів, перебудовуючи поведінку в режимі реального часу в залежності від ситуації. Це узгоджується із загальною концепцією: оптимальна система проактивного керування повинна бути модульною (для гнучкого оновлення компонентів, наприклад, підміни моделі прогнозу на кращу) та масштабованою – щоб керувати не лише одним роботом, а й флотом роботів, які обмінюються прогнозованою інформацією про середовище.

Висновки. Проактивне керування мобільними роботами в динамічному середовищі досягається через тісну інтеграцію прогнозування, багатокритеріального планування і адаптивного балансування цілей. Формалізація задачі як багатовимірного POMDP забезпечує теоретичний каркас для врахування невизначеності та конфліктних цілей. Генеративні моделі й трансформери надають можливість заглянути вперед у майбутні стани, зменшуючи реактивні запізнення системи. Багатокритеріальні алгоритми RL та механізми динамічного зважування гарантують, що робот оптимізує не один, а відразу декілька показників своєї роботи, залишаючись гнучким до зміни пріоритетів. Зрештою, правильно спроектована архітектура об'єднує ці компоненти в єдине ціле, де кожен рівень – від прогнозу до виконання – сприяє випереджальному, безпечному і ефективному функціонуванню робота в складних умовах сучасного виробництва.

Особливо важливим чинником ефективності сучасних мобільних роботів у виробничих умовах є впровадження промислових контролерів, насамперед програмованих логічних контролерів (PLC). Інтеграція PLC забезпечує не лише надійний зв'язок із приводами й датчиками, а й гарантує реалізацію локальної логіки безпеки та детермінованих реакцій на події у реальному часі, що є критичним для виробничих і динамічних середовищ. Дослідження останніх років демонструють, що PLC відіграють ключову роль у побудові стійкої архітектури керування: саме вони забезпечують гарантовану безпеку, обробку аварійних ситуацій, а також гнучкість інтеграції з інтелектуальними модулями планування й навчання з підкріпленням.

Зокрема, впровадження підходів, коли PLC взаємодіють із високорівневими AI-агентами, дозволяє реалізувати комбіновані стратегії, що поєднують швидкість і гнучкість машинного навчання із сертифікованою апаратною надійністю. Така синергія особливо ефективна для систем, які повинні одночасно реагувати на зовнішні зміни, гарантувати безпечну поведінку та виконувати виробничі завдання з мінімальними затримками. Сучасні дослідження й практичні впровадження свідчать, що архітектури з використанням PLC у ролі «мосту» між AI/ML-планувальниками та фізичними пристроями відкривають шлях

до створення дійсно гнучких і безпечних робототехнічних систем, готових до застосування у виробничих процесах Industry 4.0.

Такий підхід, підтверджений актуальними науковими публікаціями 2016–2025 рр. та останніми дослідженнями з надійності, безпеки й програмування мобільних роботів, лягає в основу оновленої науково-дослідної тези і формує напрям для подальших експериментальних розробок у сфері промислової робототехніки.

Перспективи подальших досліджень. Перспективи подальших досліджень у сфері проактивного керування мобільними роботами пов'язані з удосконаленням архітектури взаємодії між інтелектуальними алгоритмами планування, прогнозування та прийняття рішень і апаратною інфраструктурою на основі сучасних промислових контролерів та програмованих логічних контролерів (PLC). У майбутньому значну увагу слід приділяти глибокій інтеграції багатокритеріальних та адаптивних AI-методів із функціями апаратної безпеки та реального часу, які забезпечують PLC.

Окремого значення набуває розвиток методів формалізованої перевірки та верифікації політик навчання з підкріпленням, щоб гарантувати не лише ефективність, а й відповідність усім вимогам функціональної безпеки у промислових умовах. Очікується, що в найближчі роки ключовими стануть дослідження з побудови уніфікованих інтерфейсів та стандартів для ефективної та безпечної взаємодії між AI-агентами та контролерами, а також створення універсальних модульних платформ, здатних швидко масштабуватися для потреб різних виробничих систем і флотів робототехніки.

Важливим напрямом залишатиметься впровадження цифрових двійників (digital twins) і використання хмарних технологій для моделювання, навчання та тестування стратегій керування у віртуальному середовищі до їхнього застосування на реальних об'єктах. Це дозволить мінімізувати ризики, пов'язані з експериментами у виробничих умовах, та скоротити час виведення нових роботизованих рішень на ринок.

Крім того, перспективним є створення спеціалізованих мов програмування й інструментальних засобів, орієнтованих на підвищення рівня безпеки та надійності програм для PLC, що дозволить стандартизувати розробку і підвищити довіру до автоматизованих рішень у критичних галузях.

У цілому, майбутні дослідження повинні бути спрямовані на забезпечення максимальної автономності, гнучкості та безпеки мобільних роботів у складних і динамічних виробничих середовищах, поєднуючи найновіші досягнення в галузі штучного інтелекту з перевагами промислових контролерів.

Список використаної літератури

1. Wray K. H., Witwicki S. J., Zilberstein S. Multi-Objective POMDPs for Robust Autonomy. *International Joint Conference on Artificial Intelligence (IJCAI)*. 2017. URL: <https://openreview.net/forum?id=xewBjsJEKkR>
2. Roijers D. M., Vamplew P., Dazeley R., Whiteson S., Van Moffaert K., Heintz F., Mannion P. A practical guide to multi-objective reinforcement learning and planning. *Autonomous Agents and Multi-Agent Systems*. 2022. Vol. 36. Article 95. URL: <https://link.springer.com/article/10.1007/s10458-022-09552-y>
3. Zhang K., Yun P., Cen J., Cai J., Zhu D., Yuan H., Zhao C., Feng T., Wang M. Y., Chen Q., Pan J., Zhang W., Yang B., Chen H. Generative Artificial Intelligence in Robotic Manipulation: A Survey. *arXiv preprint arXiv:2503.03464*. 2025. URL: <https://arxiv.org/abs/2503.03464>
4. Kotb M., Weber C., Hafez M. B., Wermter S. QT-TDM: Planning With Transformer Dynamics Model and Autoregressive Q-Learning. *arXiv preprint arXiv:2407.18841*. 2024. URL: <https://arxiv.org/abs/2407.18841>
5. Taranta D., Marques F., Lourenço A., Prates P. A., Souto A., Pinto E., Barata J. *An Autonomous Mobile Robot Navigation Architecture for Dynamic Intralogistics. Proceedings of the 2021 IEEE 19th International Conference on Industrial Informatics (INDIN)*, Mallorca, Spain, 21–23 July 2021. IEEE, 2021. DOI: 10.1109/INDIN45523.2021.9557419. URL: <https://oecd-opsi.org/wp-content/uploads/2023/07/An-Autonomous-Mobile-Robot-Navigation.pdf>
6. Yang Q., Simão T. D., Tindemans S. H., Spaan M. T. J. Safety-constrained reinforcement learning with a distributional safety critic. *Machine Learning*. 2023. Vol. 112. P. 859–887. URL: <https://link.springer.com/article/10.1007/s10994-022-06187-8>
7. Liu E., Wu Y.-C., Huang X., Gao C., Wang R.-J., Xue K., Qian C. Pareto Set Learning for Multi-Objective Reinforcement Learning. *arXiv preprint arXiv:2501.06773*. 2025. URL: <https://arxiv.org/abs/2501.06773>
8. Mao L., Ma Z., Li X. A Multi-Task Dynamic Weight Optimization Framework Based on Deep Reinforcement Learning. *Applied Sciences*. 2025. Vol. 15, № 5. Article 2473. URL: <https://www.mdpi.com/2076-3417/15/5/2473>
9. Buet-Golfouse F., Pahwa P. Robust multi-objective reinforcement learning with dynamic preferences. *Proceedings of the 14th Asian Conference on Machine Learning. Proceedings of Machine Learning Research*. 2023. Vol. 189. P. 96–111. URL: <https://proceedings.mlr.press/v189/buet-golfouse23a.html>

Andrii Nosov, Mykhailo Korzhuk

PROACTIVE CONTROL OF MOBILE ROBOTS IN DYNAMIC MANUFACTURING ENVIRONMENTS

Modern manufacturing increasingly relies on the use of mobile robots for automation of logistics and technological operations. However, the dynamic nature of manufacturing environments creates complex scientific challenges for mobile platform control systems. Existing navigation algorithms often fail to provide an adequate balance between task execution speed, energy efficiency, and movement safety. The aim of this research is to synthesize modern methods of proactive control of mobile robots in dynamic manufacturing environments based on problem formalization through multidimensional POMDP, implementation of generative models for dynamics prediction, and multi-criteria optimization with adaptive goal balancing.

A comprehensive analysis of scientific publications from 2016–2025 was conducted on the topics of autonomous navigation, multi-criteria reinforcement learning, and integration of artificial intelligence systems with industrial controllers. Systems analysis methods were used to identify unresolved scientific problems and synthesize a new control architecture. Theoretical approaches were applied including formalization through partially observable Markov decision processes, prediction methods based on generative models, and principles of multi-criteria optimization.

It was found that optimal proactive control is achieved through the integration of three key components: prediction of future environmental states, multi-criteria planning considering productivity, safety, and energy efficiency, as well as adaptive goal balancing in real-time. A three-level control system architecture is proposed, which includes a prediction module based on generative models, a planning module using reinforcement learning algorithms, and an execution module based on programmable logic controllers. It is shown that the integration of high-level AI-algorithms with industrial PLCs provides a unique combination of adaptability and guaranteed safety. A mathematical model has been developed that formalizes the process of adaptive proactive control considering predicted environment dynamics and variable criteria priorities.

Proactive control of mobile robots in dynamic environments requires a comprehensive approach that combines prediction, multi-criteria optimization, and reliable hardware execution. The proposed architecture creates a theoretical and practical foundation for developing a new generation of autonomous robotic systems capable of functioning effectively in complex manufacturing conditions with safety guarantees.

Keywords: proactive control, mobile robots, POMDP, multi-criteria optimization, programmable logic controllers, dynamic manufacturing environment

References

1. Wray, K.H., Witwicki, S.J. and Zilberstein, S. (2017) "Multi-Objective POMDPs for Robust Autonomy", *International Joint Conference on Artificial Intelligence (IJCAI)*. URL: <https://openreview.net/forum?id=xewBjsJEKkR>
2. Roijers, D.M., Vamplew, P., Dazeley, R., Whiteson, S., Van Moffaert, K., Heintz, F. and Mannion, P. (2022) "A practical guide to multi-objective reinforcement learning and planning", *Autonomous Agents and Multi-Agent Systems*, 36, Article 95. URL: <https://link.springer.com/article/10.1007/s10458-022-09552-y>
3. Zhang, K., Yun, P., Cen, J., Cai, J., Zhu, D., Yuan, H., Zhao, C., Feng, T., Wang, M.Y., Chen, Q., Pan, J., Zhang, W., Yang, B. and Chen, H. (2025) "Generative Artificial Intelligence in Robotic Manipulation: A Survey", *arXiv preprint arXiv:2503.03464*. URL: <https://arxiv.org/abs/2503.03464>
4. Kotb, M., Weber, C., Hafez, M.B. and Wermter, S. (2024) "QT-TDM: Planning With Transformer Dynamics Model and Autoregressive Q-Learning", *arXiv preprint arXiv:2407.18841*. URL: <https://arxiv.org/abs/2407.18841>
5. Taranta, D., Marques, F., Lourenço, A., Prates, P.A., Souto, A., Pinto, E. and Barata, J. (2021) *An Autonomous Mobile Robot Navigation Architecture for Dynamic Intralogistics*, in *Proceedings of the 2021 IEEE 19th International Conference on Industrial Informatics (INDIN)*, Mallorca, Spain, 21–23 July 2021. IEEE. DOI:10.1109/INDIN45523.2021.9557419. URL: <https://oecd-opsi.org/wp-content/uploads/2023/07/An-Autonomous-Mobile-Robot-Navigation.pdf>
6. Yang, Q., Simão, T.D., Tindemans, S.H. and Spaan, M.T.J. (2023) "Safety-constrained reinforcement learning with a distributional safety critic", *Machine Learning*, 112, pp. 859–887. URL: <https://link.springer.com/article/10.1007/s10994-022-06187-8>
7. Liu, E., Wu, Y.-C., Huang, X., Gao, C., Wang, R.-J., Xue, K. and Qian, C. (2025) "Pareto Set Learning for Multi-Objective Reinforcement Learning", *arXiv preprint arXiv:2501.06773*. URL: <https://arxiv.org/abs/2501.06773>
8. Mao, L., Ma, Z. and Li, X. (2025) "A Multi-Task Dynamic Weight Optimization Framework Based on Deep Reinforcement Learning", *Applied Sciences*, 15(5), Article 2473. URL: <https://www.mdpi.com/2076-3417/15/5/2473>
9. Buet-Golfouse, F. and Pahwa, P. (2023) "Robust multi-objective reinforcement learning with dynamic preferences", *Proceedings of the 14th Asian Conference on Machine Learning, Proceedings of Machine Learning Research*, 189, pp. 96–111. URL: <https://proceedings.mlr.press/v189/buet-golfouse23a.html>