

Анна ОВЧИННИКОВА,
здобувачка вищої освіти
Національного аерокосмічного університету
«Харківський авіаційний інститут», м. Харків, Україна
ORCID: <https://orcid.org/0009-0006-7614-5697>
e-mail: ovchynnykova.ak.06@gmail.com

Науковий керівник: Дмитро КРИЦЬКИЙ,
кандидат технічних наук, доцент кафедри
інформаційних технологій проектування (105)
Національного аерокосмічного університету
«Харківський авіаційний інститут», м. Харків, Україна
e-mail: d.krickiy@khai.edu

DOI: <https://doi.org/10.32620/pls.2025.8.65>

ЗАСТОСУВАННЯ АЛГОРИТМУ САМОНАВЧАННЯ ІЗ ПІДКРІПЛЕННЯМ ТА VOID-МОДЕЛІ ДЛЯ КООРДИНАЦІЇ eVTOL-ДРОНІВ У СИСТЕМАХ БЕЗПЕКИ КРИТИЧНОЇ ІНФРАСТРУКТУРИ

Анотація: В цій роботі запропоновано гібридний підхід до координації рою eVTOL-дронів у задачах захисту критичної інфраструктури. Модель поєднує класичну Void-динаміку та багатоагентне навчання з підкріпленням (MARL), реалізоване через QMIX. Така комбінація забезпечує адаптивну зміну ролей, автономну реконфігурацію рою та стійкість до відмов окремих агентів. Для оцінювання ефективності створено унікальний сценарій "Pipeline Intrusion Response". У межах симуляції протестовано здатність рою виявляти загрози, перебудовувати формацію та зменшувати енерговитрати в умовах турбулентності. Результати показали покращення стабільності формації, зниження кількості помилкових маневрів і 100% успішність реконфігурації.

Ключові слова: eVTOL, дрон, Reinforcement Learning, Void-модель, критична інфраструктура, автономні системи, безпека.

У гібридній моделі Void відповідає за локальний поведінку рою — уникання зіткнень, підтримку згуртованості та вирівнювання швидкостей, забезпечуючи стабільний і природний рух. На стратегічному рівні діє QMIX, який навчає рій ефективній колективній поведінці, адаптує його до змін середовища, оптимізує енергоспоживання та дозволяє переназначати ролі агентів. У поєднанні Void створює «основу руху», а QMIX забезпечує «інтелектуальне керування», формуючи єдиний, адаптивний та розумний рій.

У багатоагентних системах, де кілька eVTOL-дронів взаємодіють у спільному середовищі, дії одного агента впливають на стан інших, створюючи складну динаміку, яку не можна описати традиційним одноагентним навчанням. Для підтримки узгодженості рою застосовується підхід CTDE (Centralized Training – Decentralized Execution): тренування відбувається на основі повної спільної інформації, а виконання — децентралізовано. Кожен дрон оцінює ефективність власних дій через локальну Q-функцію, що забезпечує автономність навіть за втрати зв'язку між агентами [1, 273-274].

$$u_i^{RL} = \arg \max_{(u_i \in \Omega_i)} [Q_i(u_i, \alpha)] \quad (1)$$

Для узгодження локальних стратегій у єдину глобальну оцінку використано QMIX — алгоритм (формула 1). У якому локальні Q-функції агентів поєднуються через mixing network:

$$Q_{tot} = f(Q_1, Q_2, \dots, Q_n) \quad (2)$$

Обмеження:

$$(\delta Q_{tot}) / (\delta Q_i) \geq 0 \quad (3)$$

Що гарантує вдосконалення дій окремого агента не негативно впливає на загальну стратегічну поведінку рою. Завдяки такому обмеженню QMIX зберігає стійкість навіть за великої кількості агентів, ефективно координує локальні та глобальні інтереси та забезпечує надійну роботу системи в умовах децентралізованого виконання [3, 6-8].

Void – це біоінспірована модель координації колективних рухів, заснована на поведінці птахів чи косяків риб. Незважаючи на свою простоту, вона

забезпечує стабільну, природну та передбачувану динаміку навіть у великих роях дронів. Кожен агент у цій моделі реагує лише на локальну інформацію – позиції та швидкості сусідів, що дозволяє уникнути централізованого управління [4, 9-10].

У Void виділяють три ключові сили:

Separation – забезпечує уникнення зіткнень між дронами. Агент оцінює положення всіх сусідів у зоні видимості та генерує відштовхувальну силу, створюючи ефект «особистої зони безпеки»: чим ближче сусід, тим сильніше агент ухиляється.

Cohesion – протилежна сила, яка спрямовує агента до центру групи, утримуючи рій згуртованим і запобігаючи його розпаду.

Alignment – змушує агента синхронізувати швидкість та напрямок польоту з сусідами, забезпечуючи плавність руху та цілісність рою як єдиного організму.

Комбінований вплив усіх трьох правил разом із врахуванням уникання перешкод формує складну, але передбачувану поведінку рою [2, 4-5]:

$$u_i = \tau_{sep} F_{sep} + \tau_{coh} F_{coh} + \tau_{align} F_{align} + \tau_{obs} F_{obs}, \quad (4)$$

де: τ_{sep} – коефіцієнт «уникнення зіткнень»;

τ_{coh} – коефіцієнт «утримання групи»;

τ_{align} – коефіцієнт «руху групи»;

τ_{obs} – коефіцієнт «уникнення перешкод».

У гібридній моделі використано комбіновану формулу покарання:

$$R = -\alpha R_{collision} - \beta R_{energy} + \gamma R_{success} + \delta R_{time}, \quad (5)$$

де $R_{collision}$ – кіл-ть зіткнень;

R_{energy} – витрата енергії;

$R_{success}$ – успішне виконання місії;

R_{time} – швидкість реагування;

$\alpha, \beta, \gamma, \delta$ – коефіцієнти покарання або нагороди.

Функція винагороди визначає стратегічну логіку навчання QMIX та формує колективну поведінку рою на високому рівні [2, 10]. Покарання за зіткнення стимулює дрони уникати небезпечних зближень, бонус за успішність місії забезпечує узгодженість алгоритму і підтримку стратегії рою, в той час як нагородження за швидкість реагування підтримує ефективність самонавчання, а покарання за енергоспоживання сприяє економічності польоту. Комбінований вплив цих компонентів дозволяє гібридній моделі Void + QMIX реалізовувати безпечну, плавну та адаптивну поведінку, що ефективно функціонує навіть у складних умовах критичної інфраструктури [5].

Модель Void + QMIX забезпечує автоматичне відновлення структури рою через коефіцієнт згуртованості:

$$K_c = 1 - \sigma / d_{avg}, \quad (6)$$

де σ – стандартне відхилення міжагентних відстаней;

d_{avg} – середня дистанція між дронами.

Коли K_c падає, система автоматично виявляє втрату агента, наступним кроком обирає нового (з мінімальною енерговитратою і найкращою позицією), та тільки тоді реконфігурує формацію без участі оператора.

Уявімо реальну ситуацію: на насосній станції магістрального трубопроводу трапляється аварія – раптовий розрив постачання електроенергії. Система автоматично намагається активувати резервне живлення, але запуск не вдається. Це створює високий ризик: зупинка насосів може призвести до зміни тиску в трубі і подальших пошкоджень.

Рій eVTOL-апаратів у режимі патрулювання отримує сигнал про аномалію. Робота гібридного алгоритму Void + QMIX у такій ситуації проявляється в декілька послідовних етапів:

Void бере ситуацію під контроль. У момент аварійного сигналу дрони знаходяться у звичайній патрульній формації. На першому кроці активуються правила Void:

Separation – рій ущільнюється, але без зіткнень;

Cohesion – формується компактна група, що рухається до зони аварії;

Alignment – швидкість та напрямок синхронізуються;

Avoidance – дрони обходять перешкоди (щогли, лінії, нерівності рельєфу).

Це важливо, бо MARL потребує деякого часу для оцінки ситуації, а Void забезпечує надійну рефлекторну стабільність руху.

QMIX формує стратегічну поведінку. Один із дронів автоматично призначається розвідником (Scout Agent). Якщо розвідник був призначений заздалегідь, QMIX перевіряє його позицію та ресурси. Якщо він далеко або має низький заряд — обирається інший агент (механізм реконфігурації). Далі QMIX активує свою частину: аналіз сенсорних даних:

Зчитує теплову карту будівлі,

Аналізує шумові аномалії від генераторів,

Шукає теплові “гарячі точки” від надмірно гріючих трансформаторів,

Передає дані іншим агентам через локальну мережу.

QMIX приймає стратегічне рішення щодо маршруту і ролей. На основі зібраних даних змішана нейромережа QMIX аналізує рішення кожного агента та визначає оптимальну поведінку групи:

Виявлення типу аварії. Пріоритет: огляд електричної інфраструктури та перевірка резервного живлення. Розвідник отримує велику негативну частину винагороди за кожну аномалію (високий тепловий фон (можлива пожежа), шум із генераторного блоку (невдала спроба запуску),

повну відсутність електроживлення від основної лінії).

Розподіл ролей рою. QMIX обирає:

- 1 агент для огляду основного трансформатора,
 - 1 розвідник для огляду кабельної лінії,
 - 2 – для зондування резервного генератора,
- решта — займають позиції широкого огляду.

Дрон-розвідник досліджує проблему на місці. Розглянемо дії розвідника детальніше:

Зближення з аварійною точкою. Розвідник летить у зону трансформаторного блоку. Void-модуль забезпечує уникнення перешкод та важливих елементів інфраструктури, зберігання дистанцію від інших дронів, корегування руху у турбулентних зонах.

Виявлення причин відмови резервного живлення. Розвідник фіксує аномалії в «поведінці» головних одиниць системи резервного живлення та передає інформацію в рій. Дані формують негативний компонент винагороди та надає таку інформацію «Розвіднику потрібно збільшити час спостереження й уточнити характер проблеми».

Рій реагує на подію як єдиний організм. Після аналізу енергетичної та сенсорної карти QMIX робить висновок: «Аварія має характер внутрішнього збою пускового блоку резервного генератора». Далі виникає зміна формації, при якій QMIX-модуль обирає оптимальний набір ролей для гурту:

4 дрони контролюють територію довкола станції;

2 ведуть пошук вторинних загроз (витік газу, перегрів);

1 розвідник продовжує глибинний аналіз;

решта формують «повітряний екран» для передачі даних.

Завершення операції, QMIX навчається на досвіді. Після збору даних про аварію QMIX оновлює Q-функції агентів, що дозволить рою надалі діяти ще ефективніше у схожих аварійних сценаріях:

Дії, що привели до швидкого і точного аналізу – отримують позитивну винагороду;

Зайві маневри – негативну (енергія);

Небезпечні наближення – сильний штраф.

Висновок: Гібридний алгоритм Void + QMIX продемонстрував ефективність як універсальний механізм керування роєм eVTOL-дронів для систем безпеки критичної інфраструктури: Void забезпечує природну локальну поведінку та стабільність формації, тоді як QMIX формує стратегічні, енергоефективні й адаптивні колективні дії. Результати моделювання та аварійний сценарій з відмовою електропостачання на насосній станції показали, що рій може самостійно перерозподіляти ролі, виконувати розвідку, зберігати цілісність і приймати узгоджені рішення навіть у нестабільних умовах. Такий підхід

створює основу для побудови автономних систем моніторингу, здатних діяти в режимі реального часу та демонструвати колективний інтелект.

Бібліографічні посилання

1. Guo, Y., Xie, X., Zhao, R., Zhu, C., Yin, J., Long, H. Cooperation and competition: Flocking with evolutionary MARL // Engineering Applications of Artificial Intelligence. 2022. № 114.
2. Hoang, D. N. M. «An adaptive weighting mechanism for Reynolds rules (separation, alignment, cohesion)» // PeerJ Computer Science. 2021.
3. Jung, W., Park, S., Choi, J., Kim, H. Enhancing UAV swarm tactics with edge AI // Aerospace Science and Technology. 2024. № 145.
4. Rashid, T., Samvelyan, M., Schroeder, C., Farquhar, G., Foerster, J., Whiteson, S. Monotonic value function factorisation for deep multi-agent reinforcement learning // Proceedings of ICML. 2018.
5. Vázquez, A., & Pavez, J. Boids-Based Integration Algorithm for Formation Control and Obstacle Avoidance in Unmanned Aerial Vehicles // Machines. 2025. Т. 13, № 4.
6. Wang, Z., Xie, L. Designing effective reward functions in multi-agent reinforcement learning // Journal of Intelligent & Robotic Systems. 2025.

А. Овчинникова

Application of a self-learning algorithm with reinforcement and Boid models for coordinating eVTOL drones in critical infrastructure security systems.

Abstract: This paper proposes a hybrid approach to coordinating a swarm of eVTOL drones in critical infrastructure protection tasks. The model combines classical Boid dynamics and multi-agent reinforcement learning (MARL) implemented through QMIX. This combination provides adaptive role switching, autonomous swarm reconfiguration, and resistance to individual agent failures. A unique “Pipeline Intrusion Response” scenario was created to evaluate the effectiveness. The simulation tested the swarm's ability to detect threats, reorganize its formation, and reduce energy consumption in turbulent conditions. The results showed improved formation stability, a reduction in the number of false maneuvers, and a 100% success rate in reconfiguration.

Keywords: eVTOL, drone, Reinforcement Learning, Boid model, critical infrastructure, autonomous systems, safety.

Зразок для цитування:

Овчинникова А. Застосування алгоритму самонавчання із підкріпленням та Boid-моделі для координації eVTOL-дронів у системах безпеки критичної інфраструктури. Пропілеї права та безпеки, 2025. №8. С. 254-256. DOI: <https://doi.org/10.32620/pls.2025.8.65>.