

УДК 004.8:004.75:519.216

DOI: 10.31673/2412-9070.2026.318107

Ю. В. КРАВЧЕНКО, д-р техн. наук, професор;

ORCID ID: 0000-0002-0281-4396

Є. М. МЄЗЄНЦЕВ, магістр,

ORCID ID: 0009-0004-2228-0303

Київський національний університет імені Тараса Шевченка

МОДЕЛЬ ДЕЦЕНТРАЛІЗОВАНОГО КЕРУВАННЯ ГРУПОЮ АГЕНТІВ У ВИГЛЯДІ DEC-POMDP З ДИНАМІЧНОЮ ТОПОЛОГІЄЮ ВЗАЄМОДІЇ ТА СТОХАСТИЧНИМИ ПОРУШЕННЯМИ КОМУНІКАЦІЇ

Стаття присвячена теоретичному обґрунтуванню та розробці математичної моделі децентралізованого керування динамічними групами автономних агентів, зокрема роями безпілотних літальних апаратів, в умовах фундаментальної невизначеності та ворожого операційного середовища. Авторами запропоновано удосконалення класичної моделі децентралізованого частково спостережуваного процесу прийняття рішень Маркова (Decentralized Partially Observable Markov Decision Process – Dec-POMDP) в модель, яка враховує динамічну топологію взаємодії між агентами (Dynamic Topology Decentralized Partially Observable Markov Decision Process – DT-Dec-POMDP). Ключова відмінність полягає в інтеграції динамічної топології взаємодії та експліцитному моделюванні стохастичних порушень каналів зв'язку як ендегенних факторів, що визначають інформаційний простір кожного агента. Доведено, що ігнорування зазначених аспектів призводить до втрати стійкості теоретично оптимальних політик у реальних сценаріях, зокрема під час розвідки, електронного придушення та пошуково-рятувальних операцій. Для подолання інформаційних лакун, спричинених втратою пакетів, запропоновано архітектурний блок Comm-Disturbance Filter, що реалізує механізм предиктивної компенсації на основі рекурсивної екстраполяції станів сусідніх агентів. Для обробки стабілізованого потоку спостережень та контекстно-залежної агрегації інформації запроваджено Multi-agent Transformer, здатного адаптивно зв'язувати вхідні дані в умовах мінливої мережевої структури. Результати дослідження виявляють критичні пороги стійкості системи залежно від інтенсивності збоїв зв'язку та демонструють здатність запропонованої архітектури підтримувати функціональність у зоні компенсації. Робота закладає теоретичні підстави для створення нового класу стійких, масштабованих та адаптивних автономних систем, здатних до ефективного координування в умовах екстремальної невизначеності та протидії.

Ключові слова: рій БпЛА, динамічна топологія, стохастичне порушення зв'язку, трансформер, мульти-агентні системи, кооперативна автономія.

Вступ

Ефективна координація групи автономних агентів, таких як рої безпілотних літальних апаратів (далі БпЛА) у невизначених та динамічних середовищах залишається критичною науково-технічною проблемою сучасної теорії керування та інтелектуальних систем. Ключовим математичним формалізмом у цій статті для опису таких кооперативних систем є децентралізований частково спостережуваний марківський процес прийняття рішень (далі Dec-POMDP – Decentralized Partially Observable Markov Decision Process) [1]. Визначена модель дозволяє формалізувати спільну поведінку агентів в умовах неповної інформації про стан середовища та відсутності єдиного центру управління.

Однак класичні моделі Dec-POMDP базуються на припущенні щодо фіксованої топології взаємодії між агентами та наявності надійних, безперебійних каналів зв'язку [2], [3]. Ці припу-

щення суттєво розбігаються з умовами застосування мультиагентних систем в реальному невизначеному середовищі за відсутністю гарантованої інфраструктури зв'язку, при динамічному русі агентів, наявності перешкод, електронному придушенні (jamming), втраті пакетів, часових затримках та випадкових розривах зв'язку. У реальних сценаріях застосування роїв дронів (наприклад, у військових операціях, пошуково-рятувальних місіях, моніторингу великих територій чи автономному маневруванні в умовах протидії) топологія комунікаційної мережі постійно змінюється через:

1. Відносний рух агентів (БпЛА наближаються/віддаляються, змінюють висоту, входять у зони тіні).
2. Стохастичні порушення каналів (пакетна втрата з ймовірністю p_{loss} – проста стохастична модель втрат пакетів у комунікаційних каналах, коли кожен пакет втрачається незалежно від інших з фіксованою ймовірністю, затримки з експоненціальним розподілом, повна тимчасова втрата зв'язку з окремими сусідами).
3. Зовнішні фактори (перешкоди рельєфу, погодні умови, активне радіоелектронне придушення).

Постановка задачі

Такі умови призводять до того, що класична Dec-POMDP-модель стає неадекватною: фіксована топологія не відображає реальної динаміки графу взаємодії $G(t) = (V, E(t))$, а припущення про надійну комунікацію робить неможливим точне оновлення belief-стану агентів на основі неповних або спотворених спостережень. В результаті оптимальні політики, отримані в ідеалізованій постановці, демонструють різке падіння ефективності (або повну нестійкість) при переході до реальних умов.

Додатковим ускладнюючим фактором є те, що канали зв'язку піддаються стохастичним порушенням – випадковим затримкам, втраті пакетів та обмеженням пропускної здатності. У поєднанні з динамічною топологією ці фактори призводять до деградації якості прийняття колективних рішень, знижують відмовостійкість рою та критично обмежують його масштабованість. Існуючі підходи часто абстрагуються від одного з цих аспектів, розглядаючи або динамічні графи за ідеального зв'язку, або ненадійну комунікацію у статичних мережах.

Відсутність єдиної моделі, яка б інтегрувала децентралізоване прийняття рішень (POMDP), динамічно змінну структуру взаємодії та стохастичну комунікаційних збоїв, створює значний розрив між теоретичними розробками та практичними вимогами до стійких автономних систем. Отже, актуальною науковим питанням є розробка моделі керування групою агентів у формі Dec-POMDP з динамічною топологією та стохастичними порушеннями зв'язку. Розв'язання цього питання дозволить синтезувати стратегії координації, здатні компенсувати інформаційну невизначеність і забезпечити баланс між обчислювальною складністю та ефективністю системи в умовах реального часу.

Аналіз останніх досліджень і публікацій

Аналіз актуальних досліджень у галузі функціонування груп агентів (БпЛА) свідчить про інтенсивний розвиток окремих технологічних аспектів, проте цілісний підхід до децентралізованого керування в умовах комплексної невизначеності залишається несформованим. Питання безпеки, автентифікації та динамічного управління довірою в мережах БпЛА детально розглядаються у працях І. Гу та ін. [4] і Л. Ге та ін. [5]. Доповнюючи цей напрям, А. Сисоєв та ін. [6] пропонують легковагові схеми групового підпису для захисту комунікацій у мультиагентних системах. Хоча ці дослідження є фундаментом для побудови захищених мереж, вони фокусуються переважно на цілісності даних, залишаючи поза увагою механізми адаптивного прийняття рішень при фізичній зміні топології зв'язків. Проблема координації та динамічного розподілу завдань (MRTA) є центральною темою праці С. Алкефарі та М. Менаї [7]. Запропонований дослідниками гібридний метод оптимізації демонструє високу ефективність, проте він базується на неявному припущенні про надійний обмін інформацією. Питання прикладного застосування інтелектуальних агентів для моніторингу довілля розвиваються у дослідженні

Х. Лю та ін. [8]. Вченими підкреслюється складність управління групами у реальних ландшафтах.

Практичний аспект експлуатації роїв БпЛА та аналіз їхньої ефективності у сучасних конфліктах висвітлено у роботах О. Малькова та ін. [9] та С. Мосова [10]. Дослідниками сфокусовано увагу на критичну залежність успіху місії від стійкості зв'язку та автономності кожного вузла. Концептуальні засади використання цифрових двійників для вдосконалення децентралізованого контролю запропоновано у праці Н. Панкратової та ін. [11]. Розвиток супутніх технологій, зокрема методів локалізації агентів, представлено у роботі М. Норбельт та ін. [12], а архітектурні принципи побудови сучасних систем керування обґрунтовано В. Посвістак та Д. Мірошниченко [13].

Незважаючи на значну кількість напрацювань, спостерігається суттєва фрагментарність досліджень. Існуючі моделі здебільшого розглядають або питання інформаційної безпеки [4]–[6], або алгоритми оптимізації у стабільних мережевих структурах [7], [13]. Науковий дефіцит полягає у відсутності інтегрованої моделі децентралізованого керування, яка б одночасно враховувала динамічну трансформацію топології через рух агентів та стохастичний характер збоїв зв'язку (втрата пакетів, затримки), що є невід'ємною рисою реальних умов експлуатації [9], [10]. Відтак, розробка моделі на базі Dec-POMDP, що поєднує децентралізоване прийняття рішень із формалізацією нестабільності комунікаційного середовища, є необхідним кроком для створення стійких автономних систем.

Мета і задачі дослідження

Метою дослідження є розроблення математичної моделі децентралізованого керування групою агентів на основі формалізму Dec-POMDP, яка інтегрує динамічну топологію взаємодії та стохастичні порушення комунікації як ендогенні фактори. Така модель спрямована на підвищення стійкості та ефективності координації рою БпЛА в умовах невизначеності. Основна ідея полягає в експліцитному моделюванні стохастичних порушень каналів зв'язку та інтеграції динамічної топології взаємодії, що визначають інформаційний простір кожного агента.

Результати дослідження

Перехід від теоретичних моделей кооперативного керування до практичної експлуатації роїв БпЛА в динамічних середовищах потребує перегляду базових припущень класичного формалізму Dec-POMDP. Традиційна постановка задачі, що спирається на фіксовану топологію та гарантовану доставку інформаційних пакетів, виявляється неадекватною в умовах інтенсивного маневрування агентів та активного радіоелектронного придушення. Центральною задачею тут постає необхідність підтримання цілісності колективного рішення при постійній трансформації графа взаємодії $G(t)$ з одного боку та стохастичному випадінні критично важливих спостережень з іншого. Для її вирішення пропонується модифікована модель, у якій процес прийняття рішень інтегрований із механізмами імовірнісної оцінки стану каналів зв'язку та динамічної адаптації структури взаємодії, що дозволяє формалізувати «гнучку» поведінку рою в умовах граничної невизначеності.

Формалізація задачі децентралізованого керування роєм БпЛА в умовах динамічної топології. Як фундаментальну основу для моделювання кооперативних систем в умовах інформаційної невизначеності у цьому дослідженні прийнято Dec-POMDP. Вибір зумовлений тим, що саме цей процес дозволяє формалізувати спільне прийняття рішень групою агентів, що оперують на основі неповних локальних спостережень. Відмова від припущення про статичну топологію та надійний зв'язок є критичною для застосування в реальних сценаріях, таких як динамічний розподіл завдань (MRTA) у відкритому середовищі [4] або бойове застосування, де зв'язок та конфігурація рою постійно змінюються під впливом противника та оперативної обстановки [6], [7]. Однак стандартна модель Dec-POMDP [2], [3], задана кортежем $(I, S, \{A_i\}, P, R, \{\Omega_i\}, O)$, базується на імпліцитному припущенні про стаціонарність структури взаємодії. Тут канали обміну даними вважаються або нескінченно надійними, або їх динаміка вважається несуттєвою порівняно з еволюцією стану середовища S .

Для авіаційних мультиагентних систем, зокрема роїв БпЛА, це припущення входить у пряме протиріччя з фізикою процесу. У реальних місіях (пошук цілей, патрулювання, подолання засобів ППО) топологія комунікаційної мережі є ендегенним параметром, який змінюється безперервно в часі. Динаміка графа зв'язності, у свою чергу, визначається не лише зовнішніми завадами, а й безпосередньо діями самих агентів. Таким чином, ігнорування топологічної нестационарності призводить до розриву між обчисленою оптимальною політикою та її реальною стійкістю (robustness). Для подолання вказаних недоліків у роботі запропоновано розширення моделі до вигляду Dynamic Topology Dec-POMDP (DT-Dec-POMDP). Головна концептуальна зміна такого розширення полягає у переході від статичного вектора станів до розширеного простору станів, що включає топологічну компоненту.

Вводиться динамічний граф взаємодії:

$$G_t = (V, E_t), \quad (1)$$

де V – множина агентів, а $E_t \subseteq V \times V$ – множина ребер у момент часу t , що відображає фізичну та інформаційну можливість зв'язку між агентами.

Глобальний стан системи в момент часу t визначається як розширений стан:

$$\tilde{s}_t = (s_t, G_t) \in \mathcal{S} \times \mathcal{G}, \quad (2)$$

де: $s_t \in \mathcal{S}$ – сукупність фізичних координат та внутрішніх станів агентів, $G_t \in \mathcal{G}$ – відповідна топологічна структура взаємодії.

Вираз (2) описує розширення простору станів від класичного простору \mathcal{S} до декартового добутку $\mathcal{S} \times \mathcal{G}$. Формально це означає, що топологія графа взаємодії включається до марковського стану системи.

Таким чином, марковська властивість формулюється відносно розширеного стану:

$$P(\tilde{s}_{t+1} | \tilde{s}_t, a_t) = P(s_{t+1}, G_{t+1} | s_t, G_t, a_t), \quad (3)$$

що забезпечує врахування еволюції як фізичної динаміки агентів, так і структури їх інформаційної взаємодії.

Ключовим теоретичним внеском дослідження є перевизначення функції переходу P . У класичній моделі Dec-POMDP дії агентів впливають лише на фізичний стан системи s' , тоді як структура взаємодії агентів вважається фіксованою або неявною. У запропонованій моделі вводиться топологічно-обізнана функція імовірності переходу:

$$P(s', G' | s, G, a) = P(\tilde{s}_{t+1} = (s', G') | \tilde{s}_t = (s, G), a_t = a), \quad (4)$$

де розширений стан системи визначається як описано у (2).

На відміну від класичної постановки, у запропонованій моделі дії агентів впливають не лише на фізичні координати та внутрішні стани s' , але й опосередковано змінюють конфігурацію графа взаємодії G' . Таким чином, формується причинно-наслідковий ланцюг $a_t \rightarrow s_{t+1} \rightarrow G_{t+1} \rightarrow o_{t+1}$, де топологія мережі впливає на процес спостереження та доступну агентам інформацію.

Отже, на відміну від стандартної моделі Dec-POMDP, у якій граф взаємодії є статичним або неявним, запропонована модель передбачає динамічну еволюцію топології як функцію керуючих дій та стохастичних збурень. Це створює замкнений контур взаємодії між керуванням, фізичною динамікою та інформаційною структурою системи й забезпечує можливість прогнозування втрати зв'язку як наслідку певного маневру.

Параметризація зв'язності та стохастика ребер. Множина ребер $E(t)$ не є детермінованою. Кожне ребро $e_{ij} \in E(t)$ характеризується ваговим коефіцієнтом $a_{ij}(t)$, що визначає імовірність успішної передачі інформаційного пакета:

$$a_{ij}(t) = f(d_{ij}(t), PL(d_{ij}), \eta), \quad (5)$$

де $d_{ij}(t)$ – евклідова відстань між агентами, PL – модель загасання сигналу (Path Loss), а η – вектор стохастичних завад середовища.

Представлена формалізація дозволяє розглядати топологію не просто як набір ліній зв'язку, а як складне імовірнісне поле, де кожна дія агента (наприклад, збільшення дистанції для

розширення зони сканування) неминуче призводить до трансформації інформаційної структури рою. На рисунку 1 представлено динамічну еволюцію графа зв'язності рою БПЛА у часових зрізах t_0 , t_1 , t_2 .

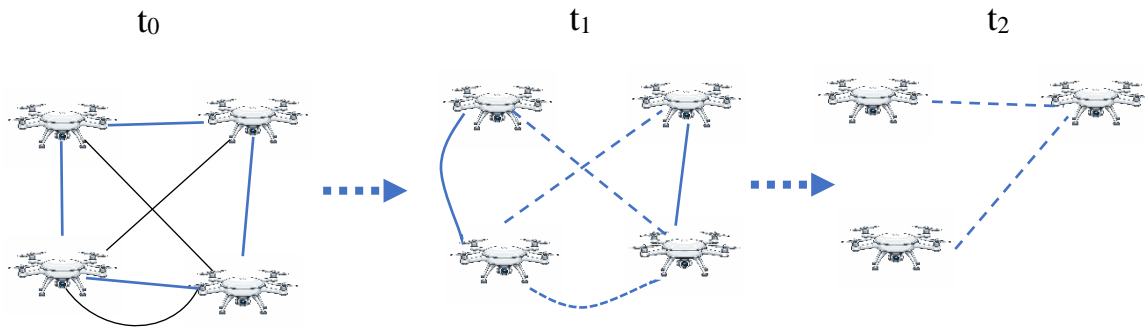


Рис. 1. Динамічна топологія рою агентів (БПЛА)

У момент t_0 рій знаходиться у компактній формації, що відповідає повнозв'язному графу з високими значеннями a_{ij} (товсті лінії). У моменти t_1 та t_2 зафіксовано процес просторового розосередження. Візуально це відображається через витончення ребер (зниження ймовірності зв'язку) та їх повне зникнення для периферійних агентів.

Перехід від «щільної» до «розрідженої» топології є критичним моментом, що демонструє, яким чином зміна $G(t)$ здатна модифікувати функцію спостереження $O(o | s', G', a)$. У цьому випадку агенти будуть змушені переходити від прямої координації до стратегій, що базуються на екстраполяції та прогнозуванні дій сусідів, з якими втрачено прямий контакт. Описаний механізм адаптації повинен стати ґрунтом для подальшої побудови архітектури нейронного трансформера.

Моделювання стохастичних порушень комунікації та механізми компенсації інформаційних втрат. У децентралізованих мультиагентних системах (рої БПЛА) стабільність колективної координації є функцією якості міжагентного обміну даними. У реальних операційних середовищах навіть за умови прямої радіовидимості канал зв'язку піддається впливу багатопробного поширення сигналу, ефекту Доплера та навмисних радіоелектронних перешкод. Вони здатні детермінувати появу «інформаційних розривів», які руйнують цілісність загального уявлення про стан системи.

У межах запропонованої моделі DT-Dec-POMDP порушення комунікації інтерпретуються не як зовнішній шум, а як ендегенний процес формування спостережень. Таким чином, динаміка графа зв'язності $G(t)$ безпосередньо визначає множину доступних сусідів $N_i(t)$, а якість зв'язку з кожним із них – ймовірність успішної передачі даних. Підхід дозволяє інтегрувати метрики якості зв'язку безпосередньо у структуру belief-стану агента, що формалізується наступним чином.

Таким чином, динаміка графа зв'язності $G(t)$ безпосередньо визначає множину доступних сусідів $N_i(t)$, а якість зв'язку з кожним із них – ймовірність успішної передачі даних. Стохастична природа збоїв зв'язку ($\xi_{ij}(t)$) створює не тільки інформаційні лакуни, а й уразливості для безпеки системи, відкриваючи вектори для атак [4]. Отже, запропонована модель $p_{\text{loss}} = \Phi(d_{ij}, \text{SIR}, \rho)$ може слугувати входом для динамічних систем управління довірою [5] та легковагових криптографічних механізмів [6]. Передача інформаційного пакета між парою агентів (i, j) у дискретний момент часу t описується індикаторною випадковою змінною $\xi_{ij}(t) \in \{0, 1\}$. Процес успішної доставки повідомлення моделюється через розподіл Бернуллі:

$$\xi_{ij}(t) \sim \text{Bernoulli}(1 - p_{\text{loss}}), \quad (6)$$

де p_{loss} – інтегральна ймовірність втрати пакета. На відміну від спрощених моделей, у даному дослідженні p_{loss} розглядається як динамічна функція, що залежить від поточної топології $G(t)$ та операційного контексту:

$$p_{\text{loss}} = \Phi(d_{ij}, \text{SIR}, \rho), \quad (7)$$

де d_{ij} – евклідова відстань, SIR – відношення сигналу до інтерференції, а ρ – щільність вузлів у локальному кластері. Таким чином, ефективне спостереження агента $o_i^{\text{eff}(t)}$ стає композитним вектором:

$$o_i^{\text{eff}(t)} = (o_i^{\text{sens}(t)}, \{\xi_{ij}(t) \cdot m_j(t)\}_{j \in N_i(t)}), \quad (8)$$

де $o_i^{\text{sens}(t)}$ – власне сенсорне спостереження агента i в момент часу t , $N_i(t)$ – множина сусідів агента i в динамічній топології на момент t , $m_j(t)$ – повідомлення, отримане від агента j (якщо зв'язок існує), $\xi_{ij}(t) \in \{0, 1\}$ – бінарний коефіцієнт успішної доставки повідомлення від j до i (моделює стохастичну пакетну втрату): $\xi_{ij}(t) = 1$ з ймовірністю $1 - p_{\text{loss}}$ (пакет доставлено), $\xi_{ij}(t) = 0$ з ймовірністю p_{loss} (пакет втрачено, внесок $m_j(t)$ обнуляється).

При $\xi_{ij}(t) = 0$ відповідний компонент вектора стає нульовим (або відсутнім), що створює лауну (прогалину) в інформаційному полі агента i , посилюючи часткову спостережуваність та невизначеність у Dec-POMDP. Ця форма дозволяє явно враховувати як динамічну топологію (через змінну множину $N_i(t)$), так і стохастичні порушення комунікації (через незалежні бернуллівські коефіцієнти $\xi_{ij}(t)$) у моделі спостережень.

Архітектура Comm-Disturbance Filter: концепція предиктивної компенсації. Для нівелювання деструктивного впливу інформаційних лакун, спричинених стохастичною втратою пакетів, у роботі розроблено блок Comm-Disturbance Filter (фільтр комунікаційних збурень). Концептуальна роль цього модуля полягає у математичній реконструкції втрачених даних шляхом їх короткострокового прогнозування. Це дозволяє підтримувати безперервність *belief-стану* агента та забезпечувати цілісність процесу прийняття рішень навіть за умов розриву зв'язку. Процес компенсації реалізується за триетапним алгоритмом:

1. Етап інтелектуальної детекції та синхронізації. На цьому рівні здійснюється моніторинг вхідного потоку даних за допомогою аналізу часових міток (*timestamping*) та верифікації ідентифікаторів послідовності. Алгоритм ідентифікує факт втрати повідомлення у режимі реального часу – до моменту активації циклу оновлення стратегії. Це дозволяє системі миттєво переключитися з режиму прямого спостереження на режим предиктивної генерації даних.

2. Етап рекурсивної екстраполяції стану. У разі підтвердження втрати пакета активується предиктивна модель, яка виконує рекурсивну екстраполяцію вектора стану сусіда. На основі апріорних знань про динамічну модель польоту та ретроспективного аналізу траєкторії сусіда j обчислюється його ймовірне поточне положення:

$$\hat{s} = f(s_j(t-1), a_j(t-1)) + \epsilon(t), \quad (9)$$

де $f(\cdot)$ – оператор переходу, що враховує інерційність та обмеження маневреності БпЛА. Такий підхід забезпечує високу точність прогнозування вектора швидкості та просторових координат при короткочасних (1,0–1,5 с.) розривах комунікації, що критично для збереження щільних формацій. Ця рекурсивна екстраполяція, за своєю суттю, реалізує легковаговий «цифровий двійник» сусіднього агента, основним призначенням якого є не симуляція, а оперативне відновлення інформаційної цілісності *belief-стану* в реальному часі [11].

3. Етап синтезу та стабілізації спостереження. На фінальному етапі формується компенсований вектор спостереження, де прогнозовані дані адаптивно заміщують інформаційні лакуни. Цей відновлений інформаційний потік подається на вхід нейронного трансформера як стабільний сигнал. Завдяки такій «інформаційній безперервності» вдається уникнути різкої деградації значень градієнта в процесі навчання та запобігти хаотичним змінам керуючих сигналів, які зазвичай виникають через нестабільність вхідних даних.

Для верифікації моделі розглянуто сценарій з інтенсивністю втрат $p_{\text{loss}} = 0,2$, що відповідає умовам активного радіоелектронного маневрування. Результати моделювання вказують на існування критичного порогу стійкості p_{loss}^* . Експериментально виділено три режими функціонування рою:

1. Зона прозорості ($p_{\text{loss}} < 0,1$). Вплив втрат нівелюється за рахунок надлишковості інформації у щільних графах.

2. Зона компенсації ($0,1 \leq p_{\text{loss}} \leq 0,35$). Використання Comm-Disturbance Filter дозволяє зберегти відхилення траєкторій у межах 5-8% від номіналу.

3. Зона фрагментації ($p_{\text{loss}} > 0,35$). Екстраполяційні похибки накопичуються експоненціально, що призводить до розпаду єдиного belief-простору рою та переходу агентів до суто індивідуальних стратегій виживання.

Інтеграція фільтра компенсації безпосередньо у формалізм DT-Dec-POMDP трансформує природу навчання. Замість того, щоб навчати політики на «брудних» даних, система навчається на стабілізованому інформаційному потоці. Таким чином, фільтр реалізує функцію прогнозного буфера, що стабілізує вхідний інформаційний потік, наближуючи його характеристики до квазістаціонарних, що критично важливо для збіжності градієнтних методів навчання з підкріпленням.

Архітектура керування на базі Multi-agent Transformer. Таким чином, у межах парадигми DT-Dec-POMDP ключовим викликом для синтезу керуючих впливів вбачається нестационарність інформаційного простору. На відміну від статичних конфігурацій, топологія рою БПЛА еволюціонує як часовий граф $G(t) = (V, E(t))$, у якому склад і потужність множини ребер $E(t)$ безпосередньо корелюють із фізичною геометрією системи та динамікою заводової обстановки. За таких умов традиційні рекурентні архітектури (RNN/LSTM) виявляються недостатньо ефективними через їхню чутливість до фіксованої розмірності входу та послідовну природу обробки, що не дозволяє миттєво адаптуватися до раптової появи або зникнення інформаційних вузлів.

Для розв'язання цієї суперечності запропоновано використання архітектури Multi-agent Transformer (MAT). Стабілізований вектор спостереження $\delta_i(t)$, крім екстрапольованих даних про сусідів, може містити й високоточні локальні дані, отримані за допомогою сучасних методів автономної локалізації, наприклад, на базі VSLAM [9]. Архітектура MAT, запропонована для обробки цього вектора, забезпечує рівень абстракції та гнучкості, необхідний для побудови повноцінних автономних систем управління нового покоління [10]. Фундаментальною перевагою є механізм (*self-attention*), який трансформує процес агрегації даних у процес семантичного зважування. Замість лінійного підсумовування повідомлень, агент обчислює відносну релевантність кожного сусіда:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V, \quad (10)$$

де запит (Q) формується на основі власного стану агента, а ключі (K) та значення (V) – на основі повідомлень від доступних у момент t сусідів.

Практична реалізація MAT у рої БПЛА дозволяє забезпечити контекстну пріоритезацію: під час щільного маневрування найбільшу вагу отримують дані від найближчих апаратів (для забезпечення безаварійності), тоді як у режимі пошуку цілі пріоритет зміщується на агентів, що мають підтверджений сенсорний контакт із об'єктом. Потрібно відмітити, що блок Comm-Disturbance Filter, описаний вище, готує для Трансформера «стабілізований» вхід, де замість фізичних пропусків даних використовуються їхні екстрапольовані оцінки. Таким чином досягається забезпечення природної масштабованості системи: архітектура MAT інваріантна до кількості агентів, що дозволяє вільно масштабувати рій від одиниць до десятків апаратів без структурної перебудови контролера.

Стратегія навчання в умовах нестационарності: Centralized Training – Decentralized Execution. Можна констатувати, що процес оптимізації політик у середовищі DT-Dec-POMDP ускладнюється явищем інформаційної нестационарності. Для локального агента динамічна зміна топології графа $G(t)$ часто не відрізняється від зміни стану самого середовища, що і створює ефект «рухомої мішені» під час навчання. Для стабілізації цього процесу запропоновано парадигму Centralized Training – Decentralized Execution (CTDE). Методологія цієї парадигми базується на функціональному розподілі ролей під час життєвого циклу системи:

1. Централізоване навчання (ЦН). На етапі тренування (в симуляційному середовищі) доступний «глобальний критик», який володіє повною інформацією про стан усіх агентів та

істинну конфігурацію графа $G(t)$. Таким чином можна коректно оцінювати спільну функцію цінності, нівелюючи локальну невизначеність окремих вузлів.

2. Децентралізоване виконання (ДВ). Під час реальної місії «критика» відкидається, і кожен БпЛА використовує лише локальну політику (актора), адаптовану до роботи з частковими та зашумленими спостереженнями через МАТ-контролер.

Такий підхід дозволяє явно інтегрувати топологічну динаміку в процес навчання. Моделювання підтвердило, що включення характеристик графа зв'язності у функцію цінності дозволяє формувати політики, стійкі до фрагментації рою. Наприклад, при критичному рівні втрат пакетів ($p_{\text{loss}} > 0,25$) система демонструє поведінку, спрямовану на мінімізацію ризику, обмежуючи активні маневри та зберігаючи стабільну траєкторію до моменту відновлення стійкого зв'язку. Запропонована архітектура націлена на забезпечення високої ресурсної стійкості системи. У сценаріях пошуково-рятувальних місій додавання нових груп БпЛА до вже функціонуючого рою відбувається без додаткового перенавчання всієї системи. Нові учасники інтегруються у спільний *attention*-простір, адаптуючи свою поведінку під загальний контекст місії без необхідності повторного навчання всієї мережі.

Висновки та перспективи подальших досліджень

У результаті проведеного дослідження розроблено цілісну математичну модель та архітектуру децентралізованого керування роєм БпЛА, що адаптована до функціонування в екстремальних умовах нестабільного зв'язку. Наукова новизна роботи полягає у розширенні класичного формалізму Dec-POMDP до моделі DT-Dec-POMDP, яка дозволила вперше інтегрувати динаміку графа взаємодії $G(t)$ та стохастичні порушення комунікації як ендогенні фактори самого процесу прийняття рішень, а не як зовнішній шум.

Для безпосереднього подолання деструктивного впливу втрати пакетів у структуру системи інтегровано блок Comm-Disturbance Filter. Завдяки реалізації механізму предиктивної компенсації через рекурсивну екстраполяцію станів, цей модуль забезпечує стабілізацію потоку спостережень та підтримує цілісність інформаційного простору кожного агента. Наступним ієрархічним рівнем архітектури, що працює вже зі стабілізованими даними, виступає Multi-agent Transformer. Використання механізму уваги (self-attention) дозволило реалізувати контекстно-залежну агрегацію інформації, що забезпечує стійку координацію рою навіть за умов миттєвої трансформації мережевої структури.

Ефективність запропонованого підходу підтверджується застосуванням стратегії навчання Centralized Training – Decentralized Execution (CTDE). Такий підхід дозволив синтезувати стійкі групові політики у симуляційному середовищі з їх подальшою успішною імплементацією в реальних децентралізованих сценаріях. Експериментально встановлені критичні пороги стійкості системи надали можливість чітко детермінувати зони її ефективного функціонування залежно від інтенсивності комунікаційних збоїв. Отже, запропонована інтегрована модель створює надійне теоретичне підґрунтя для розробки нового класу стійких та адаптивних автономних систем, здатних до координації в умовах граничної невизначеності. Подальші дослідження будуть спрямовані на практичну валідацію отриманих результатів, оптимізацію обчислювальних витрат алгоритмів та посилення механізмів кібербезпеки в ієрархії управління.

Внесок авторів

Юрій КРАВЧЕНКО – визначення загальної проблематики, наукової новизни, мети, завдань та формулювання висновків; Єгор МСЗЄНЦЕВ – аналіз літератури, розроблення математичної моделі DT-Dec-POMDP, формалізація динамічної топології взаємодії та стохастичних порушень комунікації, обґрунтування архітектурних компонентів Comm-Disturbance Filter і Multi-agent Transformer, підготовка основного тексту статті.

Декларація про штучний інтелект

Штучний інтелект не використовувався.

Конфлікт інтересів

Автори заявляють про відсутність конфлікту інтересів та підтверджують, що під час підготовки цієї роботи не існувало жодних комерційних, фінансових чи інших взаємовідносин, які могли б бути розцінені як такі, що здатні вплинути на результати дослідження або їх інтерпретацію. Робота виконана відповідно до принципів академічної доброчесності, етичних норм проведення наукових досліджень та вимог редакційної політики щодо запобігання конфлікту інтересів.

Список використаної літератури

1. Abdo, E., Brafman, R. I., Shani, G., & Soffair, N. (2023). *Team-imitate-synchronize for solving Dec-POMDPs*. In M.-R. Amini, S. Canu, A. Fischer, T. Guns, P. Kralj Novak, & G. Tsoumakas (Eds.), *Machine learning and knowledge discovery in databases: ECML PKDD 2022 (Lecture Notes in Computer Science, Vol. 13716)*. Springer. https://doi.org/10.1007/978-3-031-26412-2_14
2. Pey, J. J. J., Samarakoon, S. M. B. P., Muthugala, M. A. V. J., & Elara, M. R. (2025). *A decentralized partially observable Markov decision process for complete coverage onboard multiple shape changing reconfigurable robots*. *Expert Systems with Applications*, 271, Article 126565. <https://doi.org/10.1016/j.eswa.2025.126565>
3. Keller, R., & Brafman, R. I. (2026). *Solving Dec-POMDPs as POMDPs using imitation learning*. In C. Dima, A. Ferrando, & V. Malvone (Eds.), *PRIMA 2025: Principles and practice of multi-agent systems (Lecture Notes in Computer Science, Vol. 16366)*. Springer. https://doi.org/10.1007/978-3-032-13562-9_9
4. Hu, Y., Gan, Y., Wu, Y., Wang, H., Ma, C., & Xiong, C. (2025). *GDM-DTM: A group decision-making-enabled dynamic trust management method for malicious node detection in low-altitude UAV networks*. *Sensors*, 25(13), Article 3982. <https://doi.org/10.3390/s25133982>
5. He, L., Zhao, M., Wang, X. A., Wang, J., Wang, Z., & Liu, S. (2025). *A post-quantum authentication and key agreement scheme for drone swarms*. *Electronics*, 14(17), Article 3364. <https://doi.org/10.3390/electronics14173364>
6. Sysoyev, A., Nauruzov, K., Karati, A., Abramkina, O., Vitulyova, Y., Yeskendirova, D., & Abdoldina, F. (2025). *Lightweight group signature scheme based on PUF for UAV communication security*. *Drones*, 9(10), Article 693. <https://doi.org/10.3390/drones9100693>
7. Alqefari, S., & Menai, M. E. B. (2025). *A hybrid method to solve the multi-UAV dynamic task assignment problem*. *Sensors*, 25(8), Article 2502. <https://doi.org/10.3390/s25082502>
8. Liu, X., Li, G., Li, L., & Lu, D. (2025). *Modeling Moso bamboo tree density and aboveground biomass using multi-site UAV-LiDAR data*. *Remote Sensing*, 18(1), Article 115. <https://doi.org/10.3390/rs18010115>
9. Malkov, O., Pokhnatiuk, S., Kudymov, S., & Sherykhov, I. (2025). *Проблемні питання авіаційної підтримки військ армійською авіацією за досвідом російсько-української війни та можливі шляхи їх вирішення [Problematic issues of aviation support of troops by army aviation based on the experience of the Russian-Ukrainian war and possible ways to solve them]*. *Військово-технічний збірник*, 33(2), 127–140. <https://doi.org/10.33577/2312-4458.33.2025.127-140>
10. Mosov, S. (2024). *Роїння дронів військового призначення: реалії та перспективи [Swarming of military-purpose drones: realities and prospects]*. *Збірник наукових праць Центру воєнно-стратегічних досліджень Національного університету оборони України*, 1(80), 77–86. <https://doi.org/10.33099/2304-2745/2024-1-80/77-86>
11. Pankratova, N., Pankratov, V., & Golinko, I. (2025). *Стратегія застосування цифрового двійника за умов децентралізованого керування роєм безпілотних літальних апаратів [Strategy for using a digital twin in conditions of UAV swarm decentralized control]*. *Reports of the National Academy of Sciences of Ukraine*, 6, 23–34. <https://doi.org/10.15407/dopovid2025.06.023>
12. Norbelt, M., Luo, X., Sun, J., & Claude, U. (2025). *UAV localization in urban area mobility environment based on monocular VSLAM with deep learning*. *Drones*, 9(3), Article 171. <https://doi.org/10.3390/drones9030171>

13. Posvistak, V., & Miroschnychenko, D. (2024). Architecture of autonomous control system for FPV drones. *Herald of Khmelnytskyi National University. Technical Sciences*, 337(3), 223–230. <https://doi.org/10.31891/2307-5732-2024-337-3-33>

Yu. Kravchenko, Ye. Mieziientsev

A DEC-POMDP MODEL FOR DECENTRALIZED CONTROL OF MULTI-AGENT GROUPS WITH DYNAMIC INTERACTION TOPOLOGY AND STOCHASTIC COMMUNICATION DISTURBANCES

The article is devoted to the theoretical justification and development of a mathematical model for decentralized control of dynamic groups of autonomous agents, in particular swarms of unmanned aerial vehicles, in conditions of fundamental uncertainty and a hostile operating environment. The authors propose an extension of the classical model of decentralized partially observable Markov decision process into a model that considers the dynamic topology of interaction between agents (Dynamic Topology Decentralized Partially Observable Markov Decision Process – DT-Dec-POMDP). The key difference lies in the integration of dynamic interaction topology and explicit modeling of stochastic communication channel disturbances as internal factors that determine the information environment of each agent.

It has been proven that ignoring these aspects leads to the loss of stability of theoretically optimal strategies in real-world scenarios, particularly during reconnaissance, electronic suppression, and search and rescue operations. To address information gaps caused by packet loss, we propose the Comm-Disturbance Filter architectural block, which implements a predictive compensation mechanism based on recursive extrapolation of neighboring agents' states. To process a stabilized stream of observations and context-dependent information aggregation, a Multi-agent Transformer has been introduced, capable of adaptively weighing input data in conditions of a changing network structure. The research results reveal critical thresholds of system stability depending on the intensity of communication failures and demonstrate the ability of the proposed architecture to maintain functionality in the compensation zone. The work provides a theoretical basis for the creation of a new class of stable, scalable, and adaptive autonomous systems capable of effective coordination in conditions of extreme uncertainty and counteraction.

Keywords: UAV swarm, dynamic topology, stochastic communication disturbances, transformer, multi-agent systems, cooperative autonomy.

Надійшла до редакції: 07.04.2026

Прийнята до друку: 26.05.2026

Опубліковано: 29.06.2026

© 2026 Ю. В. Кравченко, Є. М. Мезенцев.

Цей матеріал ліцензовано за умовами CC BY 4.0. <https://creativecommons.org/licenses/by/4.0/>