

Лабораторна робота №5. Покриття території роєм і метрики місії

2 вересня 2026 р.

Мета

Навчитися ставити місію обстеження майданчика як багатоагентну задачу, успіх якої вимірюють набором метрик місії, а не сумарною винагородою. Побудувати винагороду за покриття, реалізувати опорну евристику «лети до найближчої невідвіданої комірки» й порівняти з нею навчену політику, у тому числі за обмеження заряду батареї. Навчитися розпізнавати дві патології рою: скупчення апаратів у куті майданчика й простій частини рою.

Склад бригади і розподіл ролей

Робота виконується бригадою з трьох осіб. Кожен учасник відповідає за свою підсистему й захищає її окремо.

- **Учасник А — сітка й облік покриття.** Розбиває майданчик на комірки, веде карту відвіданого з лічильником заходів і власником комірки, будує карту покриття й теплові карти.
- **Учасник В — опорна евристика і винагорода.** Реалізує три різновиди евристики «лети до найближчої невідвіданої комірки» і складену винагороду за покриття з особистою та спільною частинами.
- **Учасник С — навчання, метрики й заряд.** Навчає політику зі спільними вагами, програмує підрахунок семи метрик місії та досліджує поведінку за обмеження заряду батареї.

Інтеграційний крок виконують разом: облік покриття учасника А живить винагороду учасника В, на якій учасник С навчає політику; діагностику двох патологій рою бригада проводить спільно.

Теоретичні відомості

Місію не веде одна нейромережа. Рішення розкладені на рівні, кожен зі своїм темпом. Плутанина між рівнями і є головною причиною неробочих проєктів рою.

Рівень рішення	Що вирішує	Частота
Місія	площа, час, критерій успіху	один раз
Розподіл секторів	кому в який квадрант	0,2–1 Гц
Політика рою	куди летіти, як розійтися	50 Гц
Позиційний контур	утримання траєкторії	100 Гц
Кутові швидкості	стабілізація апарата	500 Гц
Розподіл потужності	сигнали на чотири мотори	1000 Гц

Покриття не читається з датчика. Майданчик розбивають сіткою квадратних комірок і ведуть двійкову карту відвіданого $v_c(t) \in \{0, 1\}$. Базові параметри: площа 10×10 м, комірка 0,5 м ($20 \times 20 = 400$ комірок), 6 апаратів Crazyflie 2.1 на висоті 1,5 м, 1,0 м/с, епізод 180 с при 50 Гц. Критерій успіху місії: 90 % комірок, нуль зіткнень, посадка із зарядом не нижче 15 % — нижче цього рівня місія вважається проваленою. Цільовий рівень метрики «залишок заряду» вищий, понад 30 %: 15–30 % — це вже робота на межі резерву. Винагороду нараховують за *перше* відвідування комірки:

$$r_i(t) = w_p \sum_{c \in A_i(t)} (1 - v_c(t-1)) + \frac{w_s}{N} \sum_{j=1}^N \sum_{c \in A_j(t)} (1 - v_c(t-1)) \quad (1)$$

де $r_i(t)$ — винагорода апарата i на кроці t ; N — число апаратів; $A_i(t)$ — множина комірок під апаратом i (тут одна); $v_c(t-1)$ — 1, якщо комірку вже відвідано до кроку, і 0 інакше; $w_p = 0,7$ і $w_s = 0,3$ — ваги особистої та спільної частин.

Особиста частина розв’язує задачу призначення заслуги (**credit assignment**): за $w_p = 0$ апарат, який відкрив п’ять комірок, і той, який висів, дістають однакове число. Додають штрафи $-10,0$ за зіткнення і $-2,0 \cdot \max(0, 1,0 - d_{\min})$ за близькість.

Одне число «винагорода за епізод» замовникові нічого не каже, тому доповідають набором метрик (у дужках ціль): *покриття*, частка комірок, відвіданих хоч раз ($\geq 90\%$); *час до 90%* (< 150 с); *зіткнення*, зближення пари ближче за 0,30 м на апарато-хвилину ($< 0,05$); *мінімальна дистанція* ($> 0,30$ м); *залишок заряду* ($> 30\%$); *частка часу над відвіданим* ($< 20\%$); *розкид внеску*, відношення комірок найкращого й найгіршого апаратів ($< 2 \times$).

Корисні посилання для ознайомлення:

- [Batra та ін. Decentralized Control of Quadrotor Swarms](#) — орієнтир 0,02 зіткнення на дрон за хвилину;
- [Liu та ін. Energy-Efficient UAV Control for Coverage \(DRL-EC³\)](#) — метрики покриття й енергоефективність;
- [Preiss та ін. Crazyswarm](#) — реальні обмеження польоту рою в приміщенні;
- [Stable-Baselines3](#) — реалізація PPO.

Середовище

Python 3.12, gym-pybullet-drones (MIT) з усіма залежностями: numpy, matplotlib, gymnasium, stable-baselines3, pybullet. Усе встановлюється з PyPI і GitHub.

Одна тонкість встановлення: pybullet не має готових збірок під Python 3.12, тому компілюється з джерел і потребує компілятора C++. На Ubuntu це `sudo apt install build-essential`, на macOS — інструменти командного рядка Xcode. Якщо збірка не проходить, робіть роботу в хмарному ноутбучі з попередньо встановленим компілятором — умови завдання від цього не змінюються.

Порядок виконання

Обчислювальний бюджет. Відеокорта не потрібна: усі запуски розраховані на процесор звичайного ноутбука. Тримайтеся межі **30 хвилин на один запуск навчання**; кількість кроків добийте під цю межу й фіксуйте її в ноутбучі однаковою для всіх порівнюваних запусків. Криву, що не встигла вийти на полицю, теж можна аналізувати — висновок робиться про *однаково* недонавчені політики.

Базові параметри: майданчик 10×10 м, комірка 0,5 м, шість апаратів, ліміт 180 с. Числа в дужках дано як орієнтири для самоперевірки.

1. [A] Розбити майданчик на сітку комірок і реалізувати облік покриття за першим відвідуванням, зберігаючи власника кожної комірки. Перевірити число комірок розрахунком.
2. [B] Реалізувати опорну евристику «лети до найближчої невідвіданої комірки» та щонайменше один її різновид із розведенням цілей. Виміряти метрики обох на кількох зернах.
3. [B] Сформулювати винагороду за покриття з особистою та спільною частинами за формулою (1) й перевірити її на контрольному прикладі з відомою відповіддю.
4. [C] Навчити політику зі спільними вагами та побудувати криву покриття. Якщо винагорода росте, а покриття стоїть — знайти, на чому політика заробляє.
5. [C] Запрограмувати підрахунок семи метрик місії й порівняти евристику з навченою політикою на однакових зернах.
6. [C] Додати модель заряду батареї й знайти мінімальний початковий заряд, за якого місія ще виконує критерій успіху.
7. [разом] Побудувати теплові карти й діаграми внеску, розпізнати дві патології рою та запропонувати зміну винагороди, яка їх прибирає.
8. [разом] **Оформити ноутбук.** Зібрати роботу в один Jupyter-ноутбук: код кожного учасника у своїх комітках із короткими коментарями там, де це доречно. Обов'язково всередині: карти покриття, таблицю метрик місії, результати за обмеженого заряду й діагностику патологій, а також висновки — три-п'ять тверджень із числами. Ноутбук має виконуватися від початку до кінця без ручних правок; у першій комірці вкажіть, хто з учасників за які комірки відповідав.

Контрольні запитання

1. Чому місію не можна керувати однією мережею? Наведіть два рівні прийняття рішень.
2. Чому платять за перше відвідування комірки, а не за кожен крок над нею?
3. Що станеться з поведінкою рою, якщо покласти $w_p = 0$, $w_s = 1$?
4. Обчисліть r_i апарата, який не відкрив комірки, а рій відкрив чотири ($N = 6$).
5. Чому зіткнення нормують на апарато-хвилини, а не на епізод?
6. Чому покриття 100 % не свідчить про добру політику?
7. Як лише за числами логу відрізнити скупчення в куті від простою апарата?
8. Скільки комірок містить сітка при майданчику 12×12 м і комірці 0,4 м?
9. Чому апарат іде на посадку при резерві 15 %, а не при нулі?
10. Політика має винагороду +900 і покриття 30 %. Яка це патологія?