УДК 004.56

А. В. Притула
Л. М. Куперштейн

МЕТОДИ КООРДИНАЦІЇ ТА ВЗАЄМОДІЇ АГЕНТІВ У МУЛЬТИАГЕНТНІЙ СИСТЕМІ ТЕСТУВАННЯ НА ПРОНИКНЕННЯ З ВИКОРИСТАННЯМ НАВЧАННЯ З ПІДКРІПЛЕННЯМ

Вінницький національний технічний університет.

Анотація

Запропоновано підхід до координації агентів у мультиагентній системі тестування на проникнення з використанням навчання з підкріпленням. Досліджено доцільність використання моделі CTDE та алгоритмів IQL, MADDPG і QMIX для узгодження дій агентів. Розглянуто роль спільного середовища знань у підвищенні ефективності системи.

Ключові слова: мультиагентна система, навчання з підкріпленням, координація агентів, CTDE, тестування на проникнення, QMIX, blackboard model.

Abstract

A coordination approach for agents in a multi-agent penetration testing system using reinforcement learning is proposed. The expediency of using CTDE models and IQL, MADDPG, and QMIX algorithms for creating agent actions has been investigated. The role of shared knowledge in improving system efficiency is examined.

Keywords: multi-agent system, reinforcement learning, agent coordination, CTDE, penetration testing, QMIX, blackboard model.

Вступ

У сучасних умовах гібридної війни кіберпростір перетворився на ключове поле бою з використанням високотехнологічних методів атак, що підкреслює необхідність посилення кібербезпеки та впровадження проактивних заходів захисту. Тестування на проникнення є одним із основних методів оцінки захищеності інформаційних систем, що дозволяє виявити вразливості та перевірити ефективність засобів захисту. Застосування штучного інтелекту та глибокого навчання з підкріпленням у цьому процесі сприяє автоматизації та підвищенню точності виявлення загроз [1]. Однак для досягнення більшої гнучкості та адаптивності в умовах нестабільного інформаційного середовища доцільним є впровадження мультиагентних систем, здатних діяти автономно, координовано та колективно вирішувати завдання кіберзахисту. Але через проблеми координації та узгодженості дій між автономними агентами важливо розробляти моделі взаємодії агентів, які забезпечують ефективну координацію та адаптацію до змін у середовищі.

Результати дослідження

У мультиагентних системах (MAC) координація дій між автономними агентами є критичною для досягнення спільних цілей, але в складних та динамічних середовищах виникають проблеми узгодженості дій, що можуть призвести до конфліктів стратегій (наприклад, паралельне втручання в одну ціль або конкуренції за обмежені ресурси) [2]. Для ефективного функціонування таких систем необхідно впроваджувати формалізовані протоколи комунікації та механізми пріоритетності, які дозволяють агентам обмінюватися інформацією, узгоджувати дії та вирішувати конфлікти. Такі підходи забезпечують адаптивність та стійкість MAC у постійно змінюваних умовах середовища.

Підхід Centralized training with decentralized execution (CTDE) [3] у мультиагентному навчанні з підкріпленням поєднує централізоване навчання з децентралізованим виконанням, що дозволяє агентам навчатися спільно, використовуючи глобальну інформацію, але діяти автономно з обмеженим

доступом до даних. Це особливо актуально для систем тестування на проникнення, де агенти повинні адаптуватися до динамічного середовища та взаємодіяти без централізованого контролю. При цьому CTDE забезпечує:

– масштабованість, оскільки дозволяє ефективно навчати велику кількість агентів;
– адаптивність, оскільки агенти можуть швидко реагувати на зміни в середовищі, використовуючи локальну інформацію.

Такий підхід також зменшує ризик конфліктів між агентами та підвищує загальну ефективність системи.

У мультиагентному навчанні з підкріпленням особливе значення мають алгоритми, здатні забезпечити ефективну координацію дій між агентами. В таблиці 1 наведено порівняння трьох популярних підходів — Independent Q-Learning (IQL) [4], Multi-Agent Deep Deterministic Policy Gradient (MADDPG) [5] та QMIX [6], які відрізняються архітектурою, вимогами до інформації та здатністю до масштабування в умовах складних середовищ.

Таблиця 1 - Порівняння алгоритмів навчання з підкріпленням

Алгоритм	Короткий опис	Переваги	Недоліки
IQL	Кожен агент навчається незалежно (інші агенти - частина середовища).	Простота реалізації; не потребує обміну інформацією між агентами.	Може призводити до нестабільності та неузгодженості стратегій.
MADDPG	Використовує централізованого критика під час навчання та децентралізованих акторів для виконання.	Покращує координацію; ефективний у змішаних кооперативно-конкурентних середовищах.	Вимагає повної інформації про стан під час навчання; складніший у реалізації.
QMIX	Факторизує спільну функцію цінн на основі локальних значень агентів із монотонним обмеженням.	Забезпечує узгодженість між централізованим навчанням і децентралізованим виконанням; добре масштабується.	Монотонність може обмежувати виразність моделі в деяких сценаріях.

Для забезпечення ефективної взаємодії між агентами в процесі тестування на проникнення запропоновано архітектуру MAC з використанням спільного середовища знань (blackboard model). Центральний агент-менеджер координує дії інших агентів, забезпечуючи розподіл завдань, збір результатів і синхронізацію стратегій. Комунікація реалізується через спільну «дошку знань», що дозволяє уникати конфліктів, дублювання дій і забезпечує адаптацію до змін середовища (рис. 1).

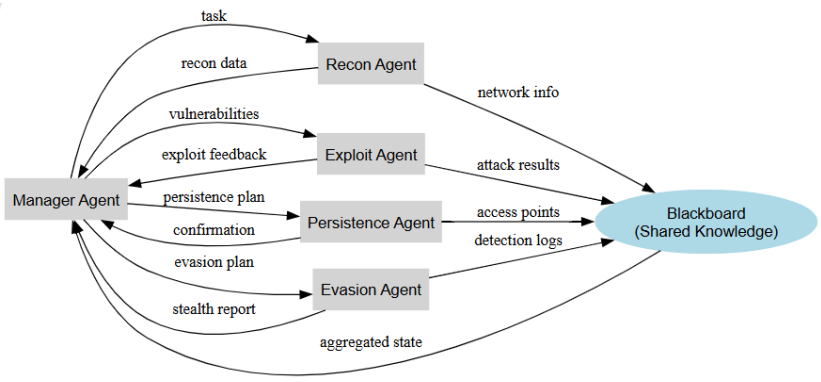

Рисунок 1 - Архітектура MAC тестування на проникнення

Адаптована схема заснована на попередніх дослідженнях, представлених у [7], із доповненням компонентів, пов’язаних з координацією та обміном знаннями між агентами.

Уявімо ситуацію, коли MAC виконує тестування безпеки корпоративної мережі. Агент Exploit Agent виявляє вразливість у веб-додатку і готується її експлуатувати. Водночас агент Persistence Agent, не маючи повної інформації про стан цілі, вже починає встановлювати бекдор у ту саму систему. Без належної координації це може призвести до наступних наслідків:

– дублювання дій;
– порушення логіки атаки;
– провалу тесту, якщо бекдор спрацює до завершення експлуатації.

Завдяки використанню моделі CTDE [3] або алгоритму QMIX [6], система централізовано

навчається на множині таких ситуацій і формує узгоджену стратегію. У момент виконання кожен агент діє автономно, але орієнтуючись на спільну цільову функцію або інформацію зі спільного середовища знань (наприклад, через blackboard). Таким чином, Exploit Agent може отримати сигнал, що Persistence Agent вже діє, і змінити свою стратегію — наприклад, переключитися на іншу ціль або відкласти атаку. Це дозволяє зменшити конфлікти дій і забезпечити ефективну послідовність атак.

Blackboard model відіграє ключову роль у забезпеченні координації агентів у MAC тестування на проникнення, оскільки це спільне середовище знань, доступне всім агентам, яке дозволяє централізовано зберігати та поширювати критичну інформацію про поточний стан системи, результати дій агентів та спільні цілі. Blackboard model забезпечує наступні можливості:

– Recon Agent може викладати результати розвідки (відкриті порти, сервіси, мережеві структури);

– Exploit Agent читає з дошки знайдені вразливості, вже використані експлойти, щоб уникати дублювання;

– Persistence Agent фіксує створені бекдори, що інші агенти могли зважити на вже наявні точки доступу;

– Evasion Agent враховує звіти про активність систем виявлення атак (IDS/EDR) для підбору менш помітних дій;

– Manager Agent бачить загальну картину та може динамічно перегруповувати стратегії агентів.

Таким чином, спільна «дошка знань» дозволяє агентам приймати більш обґрунтовані рішення, узгоджувати стратегії та динамічно реагувати на зміни в середовищі, що особливо важливо в умовах обмеженої видимості та асинхронного виконання дій.

Висновки

Координація дій між агентами є критично важливою для ефективного функціонування мультиагентної системи тестування на проникнення, особливо в умовах динамічного та частково спостережуваного середовища. Використання підходу CTDE у поєднанні з сучасними алгоритмами навчання з підкріпленням, такими як IQL, MADDPG та QMIX, дозволить досягти високої адаптивності, узгодженості стратегій та масштабованості системи.

Запровадження спільного середовища знань забезпечує ефективний обмін інформацією між агентами, зменшуючи ризик конфліктів і підвищуючи загальну продуктивність. Перспективними напрямами розвитку є впровадження генеративних агентів, здатних формувати нові сценарії атак, а також навчання системи на основі реальних даних із використанням симуляційних середовищ. Це дозволить ще більше підвищити точність, автономність і гнучкість системи в умовах реальних загроз.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Толкачова А. Методи для тестування безпеки веб-застосунків [Електронний ресурс] / Анастасія Толкачова, Андріян Піскозуб // Кібербезпека: освіта, наука, техніка. – 2024. – Т. 2, № 26. – С. 115–122. – Режим доступу: https://doi.org/10.28925/2663-4023.2024.26.668 (дата звернення: 13.04.2025). – Назва з екрана.

2. Sharma R. Multi-Agent system: enhancing collaboration in AI [Електронний ресурс] / Rajeev Sharma // Markovate. – Режим доступу: https://markovate.com/multi-agent-system/ (дата звернення: 13.04.2025). – Назва з екрана.

3. Ikeda T. Centralized training with decentralized execution reinforcement learning for cooperative multi-agent systems with communication delay [Електронний ресурс] / Takuma Ikeda, Takeshi Shibuya // 2022 61st annual conference of the society of instrument and control engineers (SICE), Kumamoto, Japan, 6–9 верес. 2022 р. – [Б. м.], 2022. – Режим доступу: https://doi.org/10.23919/sice56594.2022.9905866 (дата звернення: 13.04.2025). – Назва з екрана.

4. Lee K. M. Investigation of independent reinforcement learning algorithms in multi-agent environments [Електронний ресурс] / Ken Ming Lee, Sriram Ganapathi Subramanian, Mark Crowley // Frontiers in artificial intelligence. – 2022. – Т. 5. – Режим доступу: https://doi.org/10.3389/frai.2022.805823 (дата звернення: 14.04.2025). – Назва з екрана.

5. Multi-Agent Actor-Critic for Mixed Cooperative-Competitive Environments [Електронний ресурс] / Ryan Lowe [та ін.] // arXiv.org. – Режим доступу: https://arxiv.org/abs/1706.02275 (дата звернення: 15.04.2025). – Назва з екрана.

6. QMIX: monotonic value function factorisation for deep multi-agent reinforcement learning [Електронний ресурс] / Tabish Rashid [та ін.] // Proceedings of the 35th international conference on machine learning. – [Б. м.], 2018. – С. 4295–4304. – Режим доступу: https://doi.org/10.48550/arXiv.1803.11485. (дата звернення: 15.04.2025) – Назва з екрана.

7. Притула А. В. Архітектура мультиагентної системи для тестування на проникнення [Електронний ресурс] / Андрій Вікторович Притула, Леонід Михайлович Куперштейн // Вінниця, 24–27 берез. 2025 р. – Вінниця, 2025. – Режим доступу: https://conferences.vntu.edu.ua/index.php/all-fitki/all-fitki-2025/paper/view/24211 (дата звернення: 15.04.2025). – Назва з екрана.

Притула Андрій Вікторович – студент групи 125-23а, факультет інформаційних технологій та комп'ютерної інженерії, Вінницький національний технічний університет, Вінниця, e-mail: andrik.pritula@gmail.com.

Купершейн Леонід Михайлович – к.т.н., доцент кафедри захисту інформації, Вінницький національний технічний університет, м. Вінниця email: kupershtein.lm@gmail.com

Prytula Andrii V. – Student of Faculty of Information Technologies and Computer Engineering, Vinnytsia National Technical University, e-mail: andrik.pritula@gmail.com.

Kupershtein Leonid M. – PhD, Associated Professor of Information Protection Chair, Vinnytsia National Technical University, Vinnytsia, email: kupershtein.lm@gmail.com