

<https://doi.org/10.31891/2219-9365-2024-80-29>

УДК 004

ФОРКУН Юрій

Хмельницький національний університет

<https://orcid.org/0000-0002-7906-4191>

forkun@ridne.net

БАДЬОРА Ярослав

Хмельницький національний університет

y.badyora@gmail.com

МЕТОД РОЗРОБКИ ПРОГРАМНИХ ЗАСТОСУНКІВ ДЛЯ СИМУЛЯЦІЇ ТА УТВОРЕННЯ СТРАТЕГІЙ «ДИЛЕМИ В'ЯЗНЯ» НА ОСНОВІ НЕЙРОННИХ МЕРЕЖ

Робота присвячена розробці методу розробки програмних застосунків для симуляції та утворення стратегій «дилеми в'язня» на основі нейронних методом динамічних стратегій вирішення «дилеми в'язня» за допомогою нейронних мереж і генетичних алгоритмів.

Ключові слова: метод динамічних стратегій, «дилема в'язня», симуляція, нейронна мережа, нейрон, стратегія, генетичний алгоритм.

FORKUN Yurii, BADYORA Yaroslav

Khmelnytskyi National University

THE METHOD OF DEVELOPING SOFTWARE APPLICATIONS FOR THE SIMULATION AND CREATION OF "PRISONER'S DILEMMA" STRATEGIES BASED ON NEURAL NETWORKS

The research examines strategies for resolving the Prisoner's Dilemma through neural networks and genetic algorithms. In particular, dynamic strategies as opposed to traditional static strategies — allow for behavioral adaptation based on a variety of input data and can be developed using neural networks.

Through the use of genetic algorithms, the neural network architecture and learning parameters are optimized by simulating the evolution of strategies within a competitive dynamic. This study shows how these kinds of dynamic strategies can be used to better understand the behavior of individual agents operating in economies and societies where people's selfish interests frequently collide with the well-being of the group as a whole.

By taking this approach, the Prisoner's Dilemma can be modeled to a new level and more realistic behavioral strategies can be constructed in repeated dilemma scenarios. It also makes it easier to investigate the different factors that influence the evolution of strategies.

Keywords: prisoner's dilemma, game theory, simulation, neural network, neuron, strategy, genetic algorithm.

ПОСТАНОВКА ПРОБЛЕМИ У ЗАГАЛЬНОМУ ВИГЛЯДІ ТА ЇЇ ЗВ'ЯЗОК ІЗ ВАЖЛИВИМИ НАУКОВИМИ ЧИ ПРАКТИЧНИМИ ЗАВДАННЯМИ

«Дилема в'язня» є однією з класичних моделей теорії ігор, яка підкреслює боротьбу між самомаксимізацією та добробутом спільноти. У цій дилемі гравці стикаються з вибором: співпрацювати або втекти, де втеча приносить найбільшу винагороду, незалежно від вибору іншого гравця. Однак, якби обидва гравці навіть вирішили втекти, результат для обох був би меншим, ніж результат співпраці. Цей парадокс описує міркування про явище прийняття рішень у короткостроковій перспективі часто через відсутність раціональності, яка переважає серед членів групи в довгостроковій перспективі, дуже поширене явище в більшості соціально-економічних і біологічних систем.

Класичні розв'язання дилеми «в'язня» свідчать про прийняття статичних стратегій, таких як «око за око», або «ніколи не співпрацювати, завжди відмовлятися». Однак вони демонструють свою застосовність у статичних середовищах, без урахування плинності середовища та можливих моделей поведінки гравців з часом [1]. Це робить звичайні стратегії непридатними для моделювання ситуацій реального світу, де стратегії гравців є міжчасовими та розвиваються у відповідь на дії інших. Наприклад, дослідження [3] показують, що в складних системах, які включають довгострокові взаємодії, статичні стратегії з самого початку не обіцяють бажаних рішень, отже, причина, чому існує потреба в нових підходах до стратегічного моделювання.

Застосування нейронних мереж у модифікації стратегій «дилеми в'язня» є новим підходом, оскільки воно формує такі типи стратегій, які можуть навчатися та змінювати свою поведінку відповідно до попередніх взаємодій. Вони також можуть з'ясувати складні та складні нелінійні зв'язки між рішеннями,

прийнятими гравцями, та результатами таких рішень – отже, краща поведінка, яка буде використовуватися під час навчання в змагальному середовищі [1]. Подібним чином, результати дослідження в статті [4] демонструють, що нейронні мережі здатні досягти більшої гнучкості та адаптивних реакцій у поведінкових стратегіях гравців, ніж звичайні методи, що є вирішальним в умовах високої динаміки та високої невизначеності системи.

Інтеграція нейронних мереж і генетичних алгоритмів видається ефективною з точки зору оптимізації стратегії в дилемі ув'язненого. Генетичні алгоритми дозволяють шукати оптимальні структури нейронних мереж і параметри їх навчання, а також розвивати стратегії в конкурентному середовищі. Наприклад, у дослідженні [4] наголошується на застосуванні генетичних алгоритмів для еволюції кооперативних стратегій у повторюваній дилемі: «всі або жодна». Це дозволяє знайти більш надійні стратегії, які можуть пристосовуватися до змін середовища або інших гравців. Вона дає також гарну можливість вивчити еволюцію довгострокової стратегії, особливо вплив на неї оточуючих та інших агентів.

Ще однією важливою перевагою використання нейронних мереж є їх здатність моделювати складні соціальні та економічні взаємодії, де поведінка агентів залежить не лише від поточних дій, але й від минулого досвіду та очікуваних майбутніх змін. Наприклад, [7] вивчає алгоритми онлайн-навчання, які дозволяють моделям адаптуватися до мінливих умов під час гри, змінюючи свої стратегії у відповідь на дії інших агентів і зміни середовища. Це дозволяє створювати точніші моделі взаємодії в соціальних системах, де гравці можуть змінювати свої стратегії в залежності від умов гри.

Дослідження також підкреслюють важливість адаптивних стратегій моделювання задачі «дилемі в'язня», особливо в ситуаціях, коли правила гри змінюються, або коли інформація про дії інших гравців обмежена [6]. Нейронні мережі можна навчити та адаптувати до нових умов, враховуючи не лише поточні, але й майбутні події, що дозволяє їм приймати більш оптимальні рішення у складних та невизначених середовищах.

У практичному контексті використання нейронних мереж для моделювання адаптивної поведінки в «дилемі в'язня» може бути корисним для вирішення проблем колективного управління ресурсами, моделювання економічного співробітництва та прийняття рішень в умовах невизначеності. Дослідження [5] вказує на те, що адаптивні стратегії, засновані на нейронних мережах, можуть підвищити ефективність колективного прийняття рішень за таких обставин, при цьому враховуючи зміни в моделях поведінки та міжагентній взаємодії. Це дозволяє розробляти нові способи вирішення задач у складних системах, у яких зміна поведінки гравців може критично вплинути на результат взаємодії.

Важливим аспектом цього дослідження є розуміння різних факторів, таких як соціальні норми чи винагороди, які можуть вплинути на еволюційні стратегії в динамічних середовищах. Дослідження [2] показує, що моделювання адаптивної поведінки в таких системах може допомогти зрозуміти, як різні стратегії співпраці та зради розвиваються в умовах конкуренції; що є важливим для розробки ефективних моделей управління в складних соціально-економічних системах. Зокрема, врахування змін у винагороді за співпрацю чи відмову дозволяє точніше моделювати поведінку агента та ідентифікувати, які стратегії є найбільш надійними в довгостроковій перспективі.

Важливим аспектом розробки систем штучного інтелекту (ШІ) для вирішення повторюваної дилемі в'язня є розробка нових стратегій, які враховують минулі взаємодії. Наприклад, концепція пам'яті-один, яка використовується при моделюванні взаємодії людини з комп'ютером, дозволяє агентам враховувати минулі кроки під час прийняття рішень. Це дає їм змогу будувати системи, які передбачають поведінку опонента на основі минулого досвіду, тим самим підвищуючи ефективність кооперативних стратегій у довгостроковій перспективі. Такий підхід демонструє значні переваги порівняно зі звичайними моделями без урахування історії попередніх дій [9].

Поєднання нейронних мереж з еволюційними алгоритмами, такими як генетичні, є ще одним перспективним підходом. Наприклад, дослідження показали, що генетичні алгоритми разом з нейронною мережею дозволяють розробляти адаптивні стратегії в мінливому середовищі, а також здатні розвиватися, що обов'язково важливо для стабільної співпраці між агентами [3]. Крім того, такий підхід дозволяє агентам швидко знаходити оптимальні стратегії в нових умовах, реагуючи на зміни в поведінці опонента.

Адаптивність стратегій також забезпечується завдяки онлайн-навчанню, що дозволяє агентам змінювати свої дії в реальному часі. Використання підкріплювального навчання у поєднанні з нейронними мережами відкриває нові можливості для створення агентів, які можуть не лише ефективно імітувати людську поведінку, але й адаптуватися до її змін у динамічних умовах гри [7]. Це особливо важливо у ситуаціях, де поведінка інших гравців може бути непередбачуваною, що вимагає швидкої реакції та корекції власної стратегії.

Online-навчання забезпечує адаптацію стратегій, що дозволяє агентам змінювати свої дії в режимі реального часу. Використання навчання з підкріпленням у поєднанні з нейронними мережами відкриває

нові шляхи для створення агентів, які можуть не тільки ефективно моделювати поведінку людини, але й адаптуватися до змін у динамічному ігровому середовищі [7]. Це особливо важливо в ситуаціях, коли поведінка інших гравців непередбачувана і вимагає швидкої реакції та коригування власної стратегії.

Використовуючи нейронні мережі для моделювання їхніх стратегій, можна вивести приховані стимули, що стоять за цими діями, а також дозволити їм швидко адаптуватися в новому середовищі. Ці моделі здатні використовувати великі історичні набори даних взаємодії завдяки методам штучного інтелекту машинного навчання, що потенційно дає змогу прогнозувати майбутні дії опонентів на основі попередніх взаємодій між кількома агентами [1]. Це допомагає вдосконалювати розроблені стратегії, які можуть витримати будь-які мінливі обставини гри.

Таким чином, останні дослідження показують, що нейронні мережі, інтегровані з еволюційними методами, такими як генетичні алгоритми та клітинні автомати, забезпечують потужні засоби вирішення дилеми ув'язненого. Вони створюють адаптивні моделі, які ефективно реагують на зміни в поведінці інших акторів і забезпечують стабільну співпрацю в динамічних середовищах[7][9]. Цей підхід відкриває нові горизонти для досліджень теорії ігор і ШІ, спрямованих на розробку надійних і гнучких стратегій для вирішення складних проблем у багатоагентних системах.

ФОРМУЛЮВАННЯ ЦІЛЕЙ СТАТТІ

Мета роботи – удосконалення досліджень методів вирішення “дилеми в'язня”, шляхом приближення методології симуляції стратегій до поведінки у реальному світі.

ВИКЛАД ОСНОВНОГО МАТЕРІАЛУ

Більшість стратегій для розв'язання “дилеми в'язня”, які наразі використовуються, ґрунтуються на статичних стратегічних підходах, згідно з аналізом цих підходів. Стабільні умови сприяють ефективності цих стратегій. Однак умови взаємодії агентів у реальному світі варіюються через різноманітні фактори, особливо в соціальних та економічних системах, що робить статичні стратегії більш обмеженими та менш успішними. Необхідні нові методи стратегічного моделювання через динамічні зміни в поведінці агентів, викликані непередбачуваними зовнішніми умовами. Необхідно впровадити техніки, які можуть реагувати на змінні умови навколишнього середовища та відповідно коригувати стратегії, щоб вирішити цю проблему.

Проаналізувавши існуючі методи формування стратегій вирішення дилеми, можна зрозуміти що існуючі стратегії максимально ефективні тільки в сталих, та малозмінних середовищах. Такі математично прості стратегії важко відобразити аналогією з реального світу, а значить, важко знайти прикладне рішення для їх використання.

Щоб вирішити цю проблему, на основі існуючих досліджень, нами було запропоновано метод, що дозволить зробити генерацію та навчання таких стратегій більш гнучкою, а стратегії що вийдуть у результаті навчання, будуть розумними та поводити себе аналогічно мінливим змінам свого середовища.

Таким чином, ми пропонуємо змодельовати стратегію з вирішення “дилеми в'язня”, за допомогою нейронних мереж, які дозволяють агентам мати складну модель поведінки, таким чином ефективно коригуючи свої дії відповідно до різноманітних варіацій середовища. Ці нейронні мережі допоможуть створювати оптимальні стратегії у великих наборах даних і залучати попередні історії взаємодій, таким чином сприяючи стабільній співпраці гравців. Такий підхід назвемо методом динамічних стратегій.

Цей метод використовує інформацію про стратегію суперника та поточний стан гри як вхідні дані нейронної мережі, що репрезентує стратегію гри. Ці методи дозволяють агентам адаптуватися до змін у поведінці супротивника протягом тривалих ігрових сесій, що є критично важливим у динамічних середовищах, де взаємодії агентів непередбачувані. Ці моделі можуть навчатися в реальному часі через зворотне поширення помилки, що робить цей метод особливо ефективним для швидкої адаптації під час гри [7]. На відміну від традиційних статичних стратегій, які не змінюються в залежності від поведінки супротивників, нейронні мережі дозволяють адаптуватися. Вони мають здатність оцінювати попередні дії, робити з них висновки та миттєво коригувати свій підхід.

Нейронні мережі були обрані для розробки динамічних стратегій через ряд факторів. По-перше, складні, нелінійні зв'язки між вхідними даними та виходами можуть бути змодельовані нейронними мережами. Це особливо важливо в контексті «дилеми в'язня», де гравці повинні постійно адаптуватися, оскільки рішення, які вони приймають, часто мають непередбачувані наслідки. Нейронні мережі дуже точно прогнозують поведінку супротивника, оскільки вони обробляють великі обсяги даних і навчаються на основі взаємодій. По-друге, вони дозволяють розробляти гнучкі стратегії, які можна коригувати під час гри. Для повторюваних взаємодій у грі, наприклад, здатність мереж «пам'ятати» минулі стани та дії можуть забезпечуватися використанням моделей LSTM та GRU, що і є одним з ключових елементів нашого методу.

Важливо врахувати ряд факторів, пов'язаних з грою, при розробці нейронної мережі, яка може ефективно моделювати стратегії «дилеми в'язня». Оскільки взаємодії між агентами є вирішальними для прийняття рішень, структура мережі повинна враховувати як поточний стан гри, так і їхні минулі взаємодії.

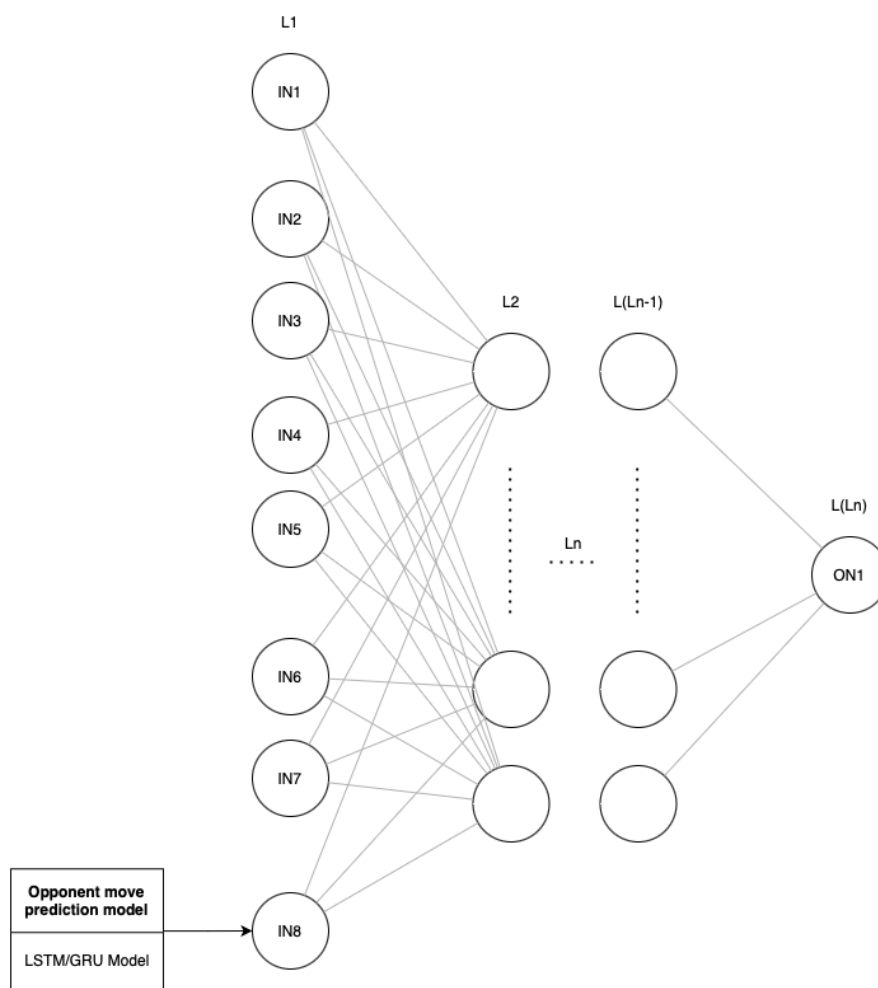


Рис. 1. Схема нейронної мережі стратегії

Щоб забезпечити найточнішу оцінку поведінки супротивника, пропонується дотримуватись наступної схеми побудови нейронної мережі агента (Рисунок 1):

Нейрон IN1 повинен виконувати функцію лічильника ходів, допомагаючи агенту визначити, чи знаходиться він на ранній, середній чи пізній стадії гри. Це важливо, оскільки поведінка агентів може змінюватися в залежності від етапу гри: гравці можуть бути більш співпрацюючими на початку, але пізніше можуть виявитися більш агресивними. В результаті модель може змінювати свою поведінку у відповідь на зміни стратегії суперника завдяки цій інформації.

Нейрони IN2 та IN3 повинні контролювати бали агента та його суперника. Це надасть агенту можливість оцінити стан гри на той момент і вирішити, чи залишитися на кооперативних тактиках, чи перейти до більш конкурентних. Більш точну оцінку поточного стану гри можна зробити, якщо нормалізація балів неможлива під час симуляції. Це можна досягти, представляючи різницю між оцінками агентів за допомогою одного нейрона.

Нейрони IN4 та IN5 запроєктовані таким чином, щоб аналізувати минулі дії як супротивника, так і агента, що дозволяє нейронній мережі вчитися на попередніх взаємодіях і коригувати свою стратегію на основі цієї інформації. Це особливо важливо для розробки адаптивних стратегій, які можуть ефективно реагувати на зміни в поведінці супротивника під час гри. Дослідження, що використовують стратегії, ґрунтовані на пам'яті, продемонстрували, що врахування минулих дій супротивників може суттєво підвищити ефективність стратегії.

Проектування нейронів IN6 та IN7, які оцінюють, наскільки подібні стратегії агентів одна до одної, є ще однією важливою особливістю цієї моделі. Агенти можуть використовувати цю інформацію, щоб визначити, які гравці найімовірніше співпрацюватимуть, а які можуть цього не зробити. Окрім сприяння співпраці між подібними стратегіями, цей підхід також допомагає запобігти конфлікту з тими, чії стратегії несумісні. Ця здатність ідентифікувати подібні стратегії покращує як кооперативну, так і конкурентну динаміку в складних багатокористувацьких сценаріях.

Нейрон IN8 повинен бути використаний для захоплення виходу моделі LSTM/GRU під час повторних взаємодій з одним і тим же супротивником. Нейрон IN8 потрібно використовувати для захоплення виходу моделі LSTM/GRU під час повторних взаємодій з одним і тим же супротивником. Протягом гри цей нейрон допомагає виявляти поведінкові патерни суперника. Ця здатність не лише дозволяє прогнозувати наступний хід супротивника, але й полегшує внесення змін у власну стратегію ефективним чином на основі зібраної інформації.

Ця структура репрезентації стратегії особливо добре працює під час тривалих ігрових сесій, коли стратегія супротивника може змінюватися у відповідь на нову інформацію. Додаткові дослідження підтвердили ефективність цієї методології в моделюванні адаптивної поведінки, коли агенти ефективно підвищували свій рівень співпраці, прогнозуючи дії супротивників за допомогою обмежених тактик пам'яті. Такі моделі можна навчати в реальному часі за допомогою зворотного поширення, що підвищує адаптивність і гнучкість стратегії.



Рис. 2. Діаграма алгоритму процесу навчання динамічних стратегій

Данна структура зберігає алгоритм дій стратегії у різних ситуаціях в достатньо компактному режимі. Проте, навчання даних нейронних мереж для їх подальшого використання для вирішення «дилеми в'язня», може зайняти багато ресурсів, бо навчання методом генетичних потребує багато ресурсів для роботи. Щоб уникнути зайвих витрат, та прискорити навчання, нами сформовано поетапний алгоритм для підготовки та навчання стратегій. Цей процес включає в себе такі етапи (Рисунок 2):

1. *Програмна реалізація ефективних статичних стратегій.* На цьому етапі впроваджуються стратегії, які вже довели свою ефективність у попередніх дослідженнях, такі як «око за око» або «безумовна співпраця». Цей крок є вирішальним, оскільки закладає основу для порівняння з динамічними стратегіями. Впровадження встановлених стратегій забезпечує базу, що дозволяє чіткіше оцінити ефективність динамічних стратегій, які будуть розроблені пізніше.

2. *Розробка симуляційного середовища.* Створено симуляційне середовище для Дилеми ув'язненого з змінними параметрами, що дозволяє гнучко налаштовувати умови гри та тестувати різні стратегії. Ця гнучкість є необхідною для дослідження того, як різні стратегії працюють у різних сценаріях.

Маніпулюючи екологічними змінними, ми можемо краще зрозуміти динаміку взаємодій агентів та умови, які сприяють або співпраці, або конкуренції.

3. *Програмна реалізація динамічних стратегій.* Цей етап передбачає інтеграцію моделей нейронних мереж з динамічним оновленням стратегій у реальному часі. Впровадження динамічних стратегій є необхідним для того, щоб агенти могли навчатися та адаптуватися під час гри. Ця здатність є життєво важливою в умовах, де супротивники можуть змінювати свої стратегії, оскільки вона дозволяє агентам ефективно реагувати на нові виклики та підвищувати свої шанси на успіх.

4. *Генерація агентів.* На основі випадкових початкових значень ваг нейронних мереж генеруються динамічні стратегії для змагання зі статичними агентами. Цей процес генерації є вирішальним, оскільки він вводить різноманітність у поведінці агентів, що призводить до більш надійних симуляцій. Створюючи різноманітних агентів з різними початковими стратегіями, ми можемо спостерігати, як ці агенти взаємодіють і еволюціонують з часом, що надає цінні відомості про ефективність різних стратегічних підходів.

5. *Наближення динамічних стратегій до роботи статичних.* Використання глибокого навчання дозволяє динамічним стратегіям тісно узгоджуватися з поведінкою статичних стратегій, що полегшує початкову фазу навчання та оптимізації. Ця узгодженість є важливою, оскільки вона дозволяє динамічним стратегіям використовувати переваги перевірених статичних стратегій, зберігаючи при цьому гнучкість для адаптації до змін. Цей гібридний підхід дозволить покращити ефективність проходження наступних етапів навчання.

6. *Симуляція ігор та навчання методом штучного відбору.* Під час симуляцій використовується метод клітинних автоматів для організації взаємодій агентів, моделюючи складні середовища з численними учасниками. Цей метод є важливим для оцінки того, як агенти з подібними або різними стратегіями взаємодіють у різних ситуаціях і як їхні стратегії еволюціонують з часом. Процес моделювання дозволяє досліджувати нові поведінкові патерни, які виникають внаслідок взаємодії агентів, що сприяє глибшому розумінню співпраці та конкуренції в багатоагентних системах. Змінюючи покоління, методом штучного відбору ми можемо вивести цілі сімейства динамічних стратегій, що за допомогою відповідних нейронів зможуть відрізнити своїх від чужих, навіть без аналізу поведінки.

7. *Збір та аналіз результатів.* Останній етап передбачає аналіз отриманих результатів симуляцій для визначення найбільш ефективних динамічних стратегій для різних ігрових умов. Цей аналіз є критично важливим для валідації запропонованих методів та розуміння, які стратегії дають найкращі результати в конкретних ситуаціях. Оцінюючи ефективність різних стратегій, ми можемо вдосконалити наші моделі та визначити напрямки майбутніх досліджень.

Таким чином, запропонований метод динамічних стратегій, заснований на нейронних мережах та генетичних алгоритмах, суттєво підвищує адаптивність стратегій у порівнянні зі статичними підходами. Це створює основу для розробки більш ефективних моделей поведінки в динамічних середовищах, роблячи цей підхід цінним для вирішення широкого спектра викликів у багатоагентних системах. Оскільки ця галузь продовжує розвиватися, інтеграція цих технік може призвести до більш складних симуляцій, які краще відображають реальні взаємодії, що в кінцевому підсумку покращить наше розуміння співпраці та конкуренції в різних контекстах.

ВИСНОВКИ З ДАНОГО ДОСЛІДЖЕННЯ І ПЕРСПЕКТИВИ ПОДАЛЬШИХ РОЗВІДОК У ДАНОМУ НАПРЯМІ

У цьому дослідженні подано метод розробки програмних застосунків для симуляції та утворення стратегій «дилеми в'язня» на основі нейронних мереж та описано реалізацію динамічних стратегій, застосованих через нейронні мережі, для вирішення «дилеми в'язня». Головною інновацією є використання нейронних мереж, таких як LSTM та GRU, в комбінації з вхідними показниками, що дозволяє агентам змінювати свої стратегії в режимі реального часу відповідно до змінюваних умов гри та попередніх взаємодій. Протягом тривалих сесій цей динамічний підхід покращує реакцію та гнучкість агентів, дозволяючи їм точно передбачати та протидіяти поведінці супротивників.

Ці нейронні мережі додатково оптимізуються шляхом використання генетичних алгоритмів, що підвищує здатність агентів адаптуватися та покращувати свої тактики. Використовуючи гібридний підхід, створюється потужна, гнучка система, яка може як адаптуватися до конкурентних умов, так і підтримувати співпрацю.

Подальші дослідження можуть зосередитися на оптимізації цих адаптивних нейронних тактик, вивченні дедалі складніших ігрових ситуацій та розширенні структур нейронних мереж для підвищення гнучкості. Поєднання нейронних мереж та динамічних стратегій відкриває нові перспективи для розвитку адаптивної поведінки в багатоагентних системах.

References

1. GRASP like algorithm. José Barahona da Fonseca // Computer Aided Chemical Engineering, Volume 24, 2017, Pages 279-284. [https://doi.org/10.1016/S1570-7946\(07\)80070-8](https://doi.org/10.1016/S1570-7946(07)80070-8)
2. Optimization of computer programming based on mathematical models of artificial intelligence algorithms. Yuhui Zheng// Computers and Electrical Engineering, Volume 110, September 2023, 108834. <https://doi.org/10.1016/j.compeleceng.2023.108834>
3. Nikoleta E. Glynatsi, Vincent Knight and Marc Harper. Properties of Winning Iterated Prisoner's Dilemma Strategies, 2024.
4. Domajnko, Martin. Iterated prisoner's dilemma and survival of the fittest from an ecological perspective. Proceedings of the 11th Student Computing Research Symposium (SCORES'22) 2022.
5. Ferraz, V., Pitz, T. Analyzing the Impact of Strategic Behavior in an Evolutionary Learning Model Using a Genetic Algorithm. Comput Econ 63, 437–475 (2024). <https://doi.org/10.1007/s10614-022-10348-1>
6. Rui Dong, Xinghong Jia, Xianjia Wang, Yonggang Chen. Optimal Tag-Based Cooperation Control for the “Prisoner's Dilemma”, 2020. <https://doi.org/10.1155/2020/8498613>
7. Montero-Porras, E., Grujić, J., Fernández Domingos, E. et al. Inferring strategies from observations in long iterated Prisoner's dilemma experiments. Sci Rep 12, 7589 (2022). <https://doi.org/10.1038/s41598-022-11654-2>
8. Lin, B., Bouneffouf, D., Cecchi, G. (2022). Online Learning in Iterated Prisoner's Dilemma to Mimic Human Behavior. In: Khanna, S., Cao, J., Bai, Q., Xu, G. (eds) PRICAI 2022: Trends in Artificial Intelligence. PRICAI 2022. Lecture Notes in Computer Science, vol 13631. Springer, Cham. https://doi.org/10.1007/978-3-031-20868-3_10
9. de Paulo, K.P., Estombelo-Montesco, C.A. & Tejada, J. New memory-one strategies of the Iterated Prisoner's Dilemma: a new framework to programmed human-AI interaction. Discov Psychol 4, 20 (2024). <https://doi.org/10.1007/s44202-024-00133-6>
10. Umberto Cerruti, Simone Dutto, Nadir Murru, A symbiosis between cellular automata and genetic algorithms, Chaos, Solitons & Fractals, Volume 134, 2020, 109719, ISSN 0960-0779, <https://doi.org/10.1016/j.chaos.2020.109719>