УДК 004.852; 004.94

**П. Кравець**
Національний університет "Львівська політехніка",
кафедра інформаційних систем та мереж

# ІГРОВА КООРДИНАЦІЯ ТА САМООРГАНІЗАЦІЯ СТРАТЕГІЙ У МУЛЬТИАГЕНТНІЙ МОДЕЛІ „ХИЖАК-ЖЕРТВА"

© *Кравець П., 2013*

**Досліджується проблема координації та самоорганізації стратегій мультиагентних систем на основі моделі стохастичної гри виду „хижак – жертва" з локальними зв'язками між агентами. Розроблено ігровий рекурентний метод та алгоритм координації стратегій агентів під час мінімізації функцій середніх програшів. Виконано комп'ютерне моделювання стохастичної гри „хижак – жертва". Досліджено вплив параметрів моделі на збіжність ігрового методу.**

**Ключові слова: координація, самоорганізація, стохастична гра, задача переслідування, модель виду „хижак – жертва".**

**The problem of coordination and self-organization of multi-agent strategies on the basis of predator-prey stochastic game model with local communications between agents is investigated. The game recurrent method and algorithm of formation of the co-ordinated strategies of agents in the course of minimisation of average losses functions are developed. Computer modelling of predator-prey model stochastic game is executed. Influence of parameters of model on convergence of a game method is investigated.**

**Key words: coordination, self-organization, stochastic game, pursuit problem, predator-prey model.**

## Координація та самоорганізація мультиагентних систем

Сучасний розвиток мережних комп'ютерних технологій робить необхідним розроблення та впровадження розподілених інформаційних та програмних систем, складність яких є такою, що будь-яке централізоване керування ними є неефективним або неможливим. Для забезпечення життєздатності таких систем в умовах невизначеності використовують децентралізовані агентно-орієнтовані методи керування [1–5].

Програмний агент – це автономна програмна система з елементами штучного інтелекту, яка може взаємодіяти з іншими агентами та людиною, використовуючи ресурси інформаційної мережі. Агент діє від імені свого власника, виконуючи поставлену перед ним задачу.

Агенти мають такі властивості: 1) автономність: агенти повністю або частково незалежні; 2) інтелектуальність: можливість самостійно приймати рішення щодо своїх подальших дій у мережному інформаційному середовищі на основі аналізу поточної інформації; 3) спеціалізація: як правило, агенти виконують вузькоспеціалізовані функції; 4) децентралізація: кожен з агентів не має повного уявлення про всю систему, і тому відсутні агенти, які керують усією системою.

Мультиагентна система (МАС) – це система, утворена декількома взаємодіючими інтелектуальними агентами. У децентралізованій системі взаємодія агентів є локально обумовленою. Програмні агенти можуть бути статичними, з фіксованим розміщенням на серверах, або мобільними, здатними мігрувати у межах комп'ютерної мережі.

Функціонує МАС, як правило, в умовах апріорної невизначеності, обумовленої зовнішніми або внутрішніми факторами. Так, автономність прийняття рішень агентами призводить до появи внутрішнього випадкового фону задачі. Часткове компенсування невизначеності забезпечується здатністю агентів до самонавчання та адаптивними стратегіями прийняття рішень.



Lviv Polytechnic National University Institutional Repository http://ena.lp.edu.ua

Кожний агент виробляє рішення на основі доступної йому локальної інформації та взаємодії у межах мультиагентної системи.

Витоками теорії МАС є зароджені у другій половині минулого століття моделі, методи та алгоритми кібернетики, теорії автоматів, інформації, еволюції та штучного інтелекту. Системотворчий вплив на агентно-орієнтоване прийняття рішень мали наукові праці з колективної поведінки автоматів, стохастичних ігор, планування розподіленого розв'язання задач, організації комунікації між агентами.

Сучасні дослідження МАС пов'язані з розв'язуванням складних проблем штучного інтелекту: розподілене прийняття рішень, керування знаннями, забезпечення координації та кооперації, структурної організації МАС, планування сценарів колективної поведінки агентів, розроблення методів, мов та засобів комунікації агентів, методів та засобів автоматизованого проектування МАС, забезпечення мобільності агентів, розподіленого вирішення різноманітних задач, багатоагентного навчання, забезпечення пластичності, надійності та стійкості системи до збоїв у роботі. Сучасні підходи та методи проектування МАС регламентуються специфікаціями фахових міжнародних організацій, наприклад, специфікацями FIPA (Foundation for Intelligent Physical Agents).

МАС використовуються для розподіленого розв'язування задач логістики, електронної комерції, військової справи, геоінформаційних систем, подолання наслідків надзвичайних ситуацій, моделювання соціальних подій, керування трудовими ресурсами, дистанційного навчання, складання розкладів, керування ринком цінних паперів та інвестицій, керування мережними потоками, керування енергетичними системами, для побудови стратегічних комп'ютерних ігор, ділових ігор, професійних тренажерів, мережних технологій та розподілених обчислень, пошуку інформації у мережі Internet, організації мобільного зв'язку та інших.

Успішність розподіленого розв'язування спільної задачі залежить від рівня координації дій агентів. Координація – це забезпечення узгодженої, впорядкованої роботи усіх ланок МАС. Координація може бути централізованою або децентралізованою. Із зростанням структурної та функціональної складності розподіленої системи в умовах невизначеності з погляду швидкості розв'язування задачі ефективнішими будуть методи децентралізованої координації агентів.

Координація необхідна для узгодження індивідуальних цілей і варіантів поведінки агентів, за яких кожен агент покращує або не погіршує значення своєї функції корисності, а система загалом покращує якість розв'язування загальної задачі. Методи розв'язування задачі координації ґрунтуються на результатах класичної теорії керування, дослідженні операцій, теорії ігор, планування та на результатах інших областей математики і кібернетики.

В основу більшості відомих методів координації МАС, явно або неявно покладено поняття „спільних зобов'язань" (commitments) та „суспільних угод" (conventions) агентів. „Спільне зобов'язання" означає необхідність виконання агентом його ролі – послідовності дій, які ведуть до досягнення визначеної мети в інтересах співтовариства агентів. Знання про зобов'язаннях інших агентів дозволяють агентові врахувати при плануванні поведінки „суспільний контекст" і обмеження, які він повинен брати до уваги. „Суспільна угода" фіксує умови, за яких зобов'язання виконуються, і обставини, коли агент може або повинен відмовлятися від виконання взятих на себе зобов'язань. У роботі [6] висунуто гіпотезу, за якою всі механізми координації в МАС можуть бути виражені в термінах спільних зобов'язань агентів і відповідних їм угод.

Найпоширенішими підходами до координації групової поведінки агентів є такі [1–5]:

1. Координація за допомогою задоволення загальних правил групової поведінки. Такий підхід використовується у системах із заданою організаційною структурою, у якій правила групової поведінки повинні строго виконуватися. Координація в цьому випадку містить два види діяльності: підтримка організаційної структури співтовариства (забезпечення неможливості порушення суспільних правил) і використання правил групової поведінки для обчислення конкретних дій агента з використанням його локальних знань.

2. Координація поведінки на основі обміну метаінформацією, наприклад, для узгодження зобов'язань і правил вирішення конфліктів, коли агенти інформують один одного про свої локальні плани та ведуть переговори для узгодження своїх зобов'язань.



Lviv Polytechnic National University Institutional Repository http://ena.lp.edu.ua

3. Командна робота, коли агенти співпрацюють для досягнення загальної довгострокової мети, функціонують у динамічному зовнішньому середовищі в умовах невизначеності та протидії з боку суперника (або команди суперників). Кожен агент має обмежену інформацію про власну команду, про зовнішнє середовище та про суперника і реалізує власні наміри за допомогою індивідуальних дій, що виконуються синхронно або асинхронно з діями інших агентів.

4. Координація в умовах конкуренції, коли агенти конкурують між собою. Кожен агент максимізує свою функцію корисності, повністю ігноруючи інтереси інших агентів. У таких МАС взаємодія агентів виражається у формі переговорів, в процесі яких агенти прагнуть досягти взаємовигідної угоди. Правила ведення переговорів повинні бути попередньо встановлені й відомі всім агентам і формалізовані у вигляді протоколу переговорів. Протоколи ґрунтуються або на теоретико-ігровий арбітражної схемі Неша, або на моделі аукціону. Найчастіше використовується стандартний протокол контрактних мереж (FIPA: Contract Net Protocol) або його модифікації.

Децентралізована координація ґрунтується на взаємодії між агентами. Така взаємодія може бути неявною, коли агенти ведуть себе незалежно і впливають один на одного через зміну станів спільного середовища, та явною, комунікативною, коли агенти додатково здійснюють прямий обмін інформацією між собою. Обмін інформації може бути глобальним або у межах локально визначених коаліцій. Агенти можуть обмінюватися знаннями про розвідане середовище, даними про поточний та прогнозований вибір стратегій, значеннями отриманих виграшів тощо. Для обміну знаннями агенти використовують спеціальні мови (наприклад, KQML, ACL) та встановлені протоколи взаємодії.

Можливість комунікації є обов'язковою умовою кооперації агентів для досягнення спільної мети. Кооперація – це координація, налагодження спільних дій у межах коаліцій агентів для оптимізації ними спільних виграшів або програшів. Вступаючи у кооперацію, агенти частково обмежують ступінь власної свободи дій, допускаючи керованість з боку інших агентів коаліції.

Колективні рішення є скоординованими, якщо вони задовольняють вимоги вигідності, стійкості та справедливості для усіх учасників прийняття рішень. На практиці найпоширенішими є критерії рівноваги за Нешем, Слейтером, Джофріоном, Байєсом, корельованої рівноваги, оптимальності за Парето.

Координація є необхідним фактором самоорганізації системи. Самоорганізація – це цілеспрямований процес створення, відтворення, впорядкування або вдосконалення організації (структури та функцій) складної динамічної системи за рахунок внутрішніх факторів, без відповідного зовнішнього впливу.

Термін самоорганізації вперше використав Р. Декарт. У літературі з кібернетики цей термін вперше зустрічається в роботах У. Р. Ешбі [7], який визначає систему із самоорганізацією як таку, що сама змінює свою структуру. На те, що самоорганізація виникає за рахунок внутрішніх взаємодій між елементами системи через загальне середовище (явище стігмергії), вказав П. Грассе [8]. Самоорганізацію у термінах термодинаміки визначив І. Р. Пригожин [9]: відкрита система зменшує свою ентропію (у системі виникає порядок із хаосу), коли в неї надходить енергія за рахунок зовнішнього впливу. Самоорганізацію як явище спонтанного утворення просторових та часових структур у фізичних і хімічних системах, які знаходяться далеко від стану рівноваги [10, 11], визначив Г. Хакен. Механізм самоорганізації як еволюцію циклів та гіперциклів хімічних реакцій у біологічних системах, які мають властивості самовідтворення та здатність до виживання [12], визначив М. Ейген.

За певних умов самоорганізація може проявлятися у розподілених природних (наприклад, фізичних, хімічних, біологічних), організаційних (наприклад, соціальних, екологічних, економічних) або штучних (наприклад, мультиагентних) системах. Класичними прикладами самоорганізації є фізика лазера, хімічна реакція Бєлоусова–Жаботинського, конвективні комірки Бенара, біологічна самоорганізація бактерій, колоній або роїв комах, косяків риб, зграй птахів тощо [13].

Самоорганізація МАС – це здатність колективу агентів з локально обумовленими зв'язками і цілями досягати стійких скоординованих стратегій поведінки в умовах невизначеності за рахунок самонавчання, можливості функціонувати як єдине ціле та забезпечувати виконання глобальної



Lviv Polytechnic National University Institutional Repository http://ena.lp.edu.ua

мети розвитку системи. Самоорганізація та складні форми поведінки МАС можуть проявлятися при реалізації найпростіших дій агентів [14, 15].

МАС мають усі необхідні вимоги для забезпечення самоорганізації: 1) автономність – здатність системи до самостійного прийняття рішень; 2) відкритість – здатність сприймати зовнішній світ за допомогою рецепторів і локально впливати на нього за допомогою ефекторів; 3) наявність програмного або фізичного середовища для розподіленої взаємодії агентів; 4) наявність засобів підтримання взаємодії та кооперації агентів; 5) відсутність зовнішнього або централізованого керування; 6) динамічна оптимізація виконується в процесі роботи системи і не вимагає попереднього настроювання або планування; 7) простота правил локальної взаємодії, яка веде до складної поведінки системи загалом; 8) можливість повторного використання варіантів дій у часі; 9) адаптивність та стійкість стосовно змін – здатність належного реагування на зміни зовнішнього середовища.

Виділяють такі механізми самоорганізації МАС: 1) прямі взаємодії між агентами за допомогою відповідних протоколів обміну даними; 2) непрямі комунікації через середовище; 3) самоорганізація поведінки агентів на основі використання навчання з підкріпленням (заохоченням); 4) кооперативна поведінка індивідуальних агентів; 5) вибір типової архітектури системи (холонічний підхід з елементами централізації).

Вивченням процесів самоорганізації систем різної природи займається теорія систем та синергетика – міждисциплінарна наука про нелінійні відкриті дисипативні (такі, що розсіюють енергію) системи [10].

Мірою оцінювання самоорганізації системи є її ентропія. Зі зменшенням ентропії ступінь самоорганізації системи зростає. Як правило, ефект самоорганізації розподіленої системи спостерігається з боку сторонніх систем з природним або штучним інтелектом. Зовнішніми проявами самоорганізації можуть бути утворення впорядкованих структур, скоординованих дій агентів та властивість емерджентності – набута інтегральна властивість системи, не характерна для її складових. Такі або інші прояви самоорганізації дають можливість розглядати і вивчати МАС як один цілісний організм.

Система проявляє емерджентність (системний ефект), якщо на макрорівні у ній динамічно виникають деякі нові властивості, процеси, поведінка, структури, патерни тощо, які є наслідком локальних взаємодій елементів системи на мікрорівні, причому ці властивості такі, що вони не описуються в термінах властивостей мікрорівня [14, 15].

Необхідними умовами виникнення явища емерджентності є такі: 1) наявність принаймні дворівневої організації: мікрорівень, де відбуваються локальні взаємодії, та макрорівень, де може виявитися емерджентний процес; 2) нелінійність: взаємодія на мікрорівні є нелінійною; 3) наявність зворотного зв'язку: локальні взаємодії на мікрорівні повинні містити зворотний зв'язок; 4) динамічна рівновага: емерджентні процеси проходять у динаміці, тобто вони існують, поки існують взаємодії на мікрорівні; поняття статичної рівноваги для них не існує.

Явища емерджентності мають такі властивості: 1) новизна: невисловлювальність у термінах мікрорівневих компонентів; 2) когерентність: узгоджене проходження хвильових або коливних процесів, яке проявляється тоді, коли в системі виникають відповідні взаємодії; 3) емерджентність проявляється на макрорівні стосовно до породжуючих це явище компонентів; 4) динаміка: емерджентність виникає в процесі взаємодії і не є заданою заздалегідь або ззовні; 5) візуалізація: емерджентність явно демонструє себе тим або іншим способом, наприклад, виникненням впорядкованих структур, фазових переходів тощо.

Емерджентність та самоорганізація є процесами, які можуть виникати окремо (вони виражають різні властивості поведінки), або спільно. Основна подібність емерджентності та самоорганізації в тому, що вони є динамічними процесами, обумовленими локальними взаємодіями складових елементів на мікрорівні, а проявляються на макрорівні.

Обидва процеси є стійкими, але їх відмінність полягає у тому, що емерджентність стійка стосовно множини складових компонентів системи, чиї взаємодії викликають це явище (компоненти можуть з'являтися та зникати), а самоорганізація стійка в тому розумінні, що вона здатна адаптуватися як до внутрішніх змін у системі, так і до змін зовнішнього світу, підтримуючи виниклий порядок.



Lviv Polytechnic National University Institutional Repository http://ena.lp.edu.ua

Враховуючи притаманні колективу агентів фактори конкуренції, взаємодії, кооперації, навчання, самоорганізації, для дослідження МАС в умовах невизначеності використаємо модель стохастичної гри [16], яка забезпечує адаптивний пошук точок рівноваги функцій виграшів на одиничному симплексі комбінованих змішаних стратегій гравців. Під час стохастичної гри гравців вчать вибирати оптимальні у середньому чисті стратегії (дії), перебудовуючи власні вектори динамічних змішаних стратегій (умовних імовірностей варіантів дій). Поряд із характерними індивідуальними особливостями, теорія ігор і теорія МАС мають спільні предмети дослідження. Стохастична гра є одним із засобів моделювання МАС.

З теоретико-ігрового погляду координація – це узгоджений вибір стратегій, що задовольнять умови, накладені на значення функцій виграшів або програшів, отримані під час їх оптимізації колективом агентів. Частковим випадком координації є синхронізація дій агентів.

Ігрова координація стратегій агентів у процесі самоорганізації МАС є актуальною науково-практичною проблемою, ще недостатньо вивченою. З пізнавального погляду важливою є візуалізація отриманих у результаті комп’ютерного експерименту скоординованих стратегій як однієї з ознак самоорганізації МАС.

## Мета роботи

Процеси координації та самоорганізації МАС вивчатимемо на основі ігрової моделі переслідування „хижак – жертва” [17]. Запозичені від природи моделі переслідування з різноманітними сценаріями (наприклад, переслідування акулою косяка риб, атака коршуном зграї птахів тощо) використовуються для відпрацювання стратегій поведінки агентів у розподілених системах прийняття рішень. У моделях переслідування виділяють агентів-хижаків та агентів-жертв, які мають антагоністичні цілі. Кінцевою метою агентів-хижаків є досягнення однієї із жертв під час переслідування, інакше – мінімізація відстані до жертви. Метою агентів-жертв є збереження власного життя, інакше – максимізація чи дотримання безпечної відстані від хижака. Вибір цієї задачі обумовлений її простотою та зручністю візуалізації скоординованих стратегій як ознаками ігрової самоорганізації системи.

Метою цієї роботи є визначення умов та механізмів локальної координації дій агентів, що призводять до самоорганізації МАС на прикладі стохастичної гри „хижак – жертва”. Для досягнення мети необхідно розв’язати такі задачі: побудувати модель стохастичної гри, розробити метод та алгоритм для розв’язування стохастичної гри, виконати програмне комп’ютерне моделювання стохастичної гри для виявлення факторів координації та самоорганізації МАС.

## Постановка ігрової задачі

Враховуючи розподіленість системи та колективний характер формування рішень, антагонізм (між хижаком та жертвою) та узгодженість (у межах групи хижаків або жертв) цілей, розв’яжемо задачу самоорганізації МАС на основі адаптивних методів стохастичних ігор [16, 18].

Стохастична гра $\Gamma = (D,\ \{U^i\}_{i\in D},\ \{P^i(\xi^i)\}_{i\in D})$ задається множиною агентів-гравців $D \neq \varnothing$, векторами чистих стратегій $U^i = (u^i[1], u^i[2], ..., u^i[N])$ та апріорі невідомими розподілами $P^i(\xi^i)$ випадкових програшів $\xi^i$ $\forall i \in D$. Нехай один з агентів є хижаком, а решта – жертвами. Чисті стратегії $u[k]$ , $k = 1..N$ визначають напрямки можливих рухів агентів.

Повторювальна гра розгортається у дискретні моменти часу $n = 1, 2, \ldots$. Для цього кожен агент $i \in D$ здійснює незалежний випадковий вибір однієї з $N \geq 2$ власних чистих стратегій $u_n^i = u^i \in U^i$ .

Структура МАС визначає локально обумовлений механізм формування випадкових програшів $\xi_n^i = \xi_n^i(u_n^{D_i})$, що є функціями спільних стратегій $u^{D_i} \in U^{D_i} = \underset{j\in D_i}{\times} U^j$ з локальних підмножин агентів $D_i \subseteq D$ , $D_i \neq \varnothing$ $\forall i \in D$. Випадкові програші $\{\xi_n^i\}$ є незалежними $\forall u_n \in U$ , $\forall i \in D$, $n = 1, 2, \ldots$, мають постійне математичне сподівання $M\{\xi_n^i(u^{D_i})\} = v(u^{D_i}) = const$ та обмежений другий момент $\sup_n M\{[\xi_n^i(u^{D_i})]^2\} = \sigma^2(u^{D_i}) < \infty$ . Стохастичні характеристики випадкових програшів $\{\xi_n^i\}$ $\forall i \in D$ апріорі не відомі агентам.



Lviv Polytechnic National University Institutional Repository http://ena.lp.edu.ua

Після вибору варіантів $u_n^i$ $\forall i \in D$ гравці отримують поточний програш, який складається з трьох складових:

$$\zeta_n^i = \lambda \xi_n^i[1] + (1-\lambda)\xi_n^i[2] + \mu_n , \tag{1}$$

де $\lambda \in [0,1]$ – ваговий коефіцієнт; $\mu_n \sim Normal(0,d)$ – адитивний білий гауссівський шум, нормально розподілена випадкова величина з нульовим математичним сподіванням та дисперсією $d > 0$.

Перша складова (1) визначає штраф за порушення координації руху сусідніх агентів:

$$\xi_n^i[1] = \sum_{j \in D_i} \chi(u_n^i \neq u_n^j) ,$$

де $\chi() \in \{0,1\}$ – індикаторна функція події, $u_n^i$ – поточна стратегія жертви $i \in D$.

Друга складова (1) визначає штраф за порушення умови реплікації (повторення) лівосторонніх або правосторонніх дій у напрямку руху агента:

$$\xi_n^i[2] = \chi(s \in D_i)\chi(u_n^i \neq u_n^s) + \chi(s \notin D_i)\chi(u_n^i \neq u_n^{Side(i)}) ,$$

де $s$ – ідентифікатор хижака; $u_n^s$ – поточна стратегія хижака; $Side(i)$ – операторна функція для визначення ідентифікатора сусіднього агента, розміщеного ліворуч $Side(i) = Left$ або праворуч $Side(i) = Right$ від напрямку руху жертви.

Якість вибору чистих стратегій у момент часу $n$ визначається середніми програшами

$$Z_n^i = \frac{1}{n}\sum_{t=1}^{n} \zeta_t^i \quad \forall i \in D . \tag{2}$$

Мета гри полягає у мінімізації функцій середніх програшів:

$$\overline{\lim_{n\to\infty}} Z_n^i \to \min_{u_n^i} \quad \forall i \in D . \tag{3}$$

Отже, взаємодіючи з середовищем, на основі спостереження поточних програшів $\{\zeta_n^i\}$ кожен гравець $i \in D$ повинен навчитися вибирати чисті стратегії $\{u_n^i\}$ так, щоби з часом $n = 1,2,...$ забезпечити виконання системи критеріїв (3).

Розв’язки ігрової задачі задовольнятимуть одну з умов колективної рівноваги, наприклад, Неша, Слейтера, Парето, залежно від методу формування послідовностей стратегій $\{u_n^i\} \forall i \in D$.

**Метод розв’язування задачі**

Для вибору варіантів рішень в умовах невизначеності застосуємо марківські адаптивні методи, які на основі опрацювання поточної інформації про систему забезпечують оптимальний у середньому вибір варіантів дій у наступні моменти часу.

Сформуємо послідовність варіантів рішень $\{u_n^i\}$ на основі динамічних векторів змішаних стратегій $p_n^i = (p_n^i(1), p_n^i(2),..., p_n^i(N))$ $\forall i \in D$, елементи $p_n^i(j)$, $j = 1..N$ яких є імовірностями вибору чистих стратегій за умови реалізації передісторії вибору чистих стратегій $\{u_t^i \mid t = 1,2,...,n-1\}$ та отримання відповідних програшів $\{\zeta_t^i \mid t = 1,2,...,n-1\}$. Змішані стратегії набувають значення на $N$-вимірних одиничних симплексах:

$$S^N = \left\{ p \left| \sum_{j=1}^{N} p(j) = 1;\ p(j) \geq 0 \quad (j = 1..N) \right. \right\} .$$

Значення чистих стратегій визначаються з умови:

$$u_n^i = \left\{ u^i(l) \,|\, l = \arg\min_{l} \sum_{k=1}^{l} p_n^i[k] > \omega \ (k,l = 1..N) \right\} , \tag{4}$$

де $\omega \in [0,1]$ – випадкова величина з рівномірним розподілом.

Необхідно визначити метод зміни векторів змішаних стратегій $p_n^i$ $\forall i \in D$, який згідно з (4) забезпечуватиме генерування випадкових чистих стратегій $\{u_n^i\}$ так, щоб в асимптотиці часу забезпечити виконання системи критеріїв (3).



Lviv Polytechnic National University Institutional Repository http://ena.lp.edu.ua

Побудуємо метод розв'язування стохастичної гри на основі стохастичної апроксимації умови доповняльної нежорсткості детермінованої гри, справедливої для змішаних стратегій у точці рівноваги за Нешем [19].

Для цього визначимо полілінійну функцію середніх програшів детермінованої гри:

$$V^i(p^{D_i}) = \sum_{u^{D_i} \in U^{D_i}} v^i(u^{D_i}) \prod_{j \in D_i; u^j \in u^{D_i}} p^j(u^j),$$

де $v(u^{D_i}) = M\{\xi_n^i(u^{D_i})\}$.

Тоді векторна умова доповняльної нежорсткості (CS, Complementary Slackness) матиме вигляд:

$$\vec{\mathrm{CS}} = \nabla_{p^i} V^i(p^{D_i}) - e^{N_i} V^i(p^{D_i}) = 0 \qquad \forall i \in D,$$

де $\nabla_{p^i} V^i(p^{D_i})$ – градієнт функції середніх програшів; $e^N = (1_j \mid j = 1..N)$ – вектор, всі компоненти якого дорівнюють 1; $p^{D_i} \in S^{M_i}$ – комбіновані змішані стратегії гравців з локальних множин $D_i$, задані на опуклому одиничному симплексі $S^{M_i}$ ($M_i = \prod_{j \in D_i} |D_i|$).

Для врахування розв'язків у вершинах одиничного симплексу виконаємо зважування умови доповняльної нежорсткості елементами векторів змішаних стратегій:

$$diag(p^i)(\vec{CS}) = 0 \quad \forall i \in D, \tag{5}$$

де $diag(p^i)$ – квадратна діагональна матриця порядку $N_i$, побудована з елементів вектора $p^i$.

Враховуючи, що $diag(p^i)[\nabla_{p^i} V^i - e^{N_i} V^i] = E\{\zeta_n^i [e(u_n^i) - p_n^i] \mid p_n^i = p^i\}$, де $E\{\}$ – функція математичного сподівання, з (5) на основі методу стохастичної апроксимації отримаємо рекурентну залежність:

$$p_{n+1}^i = \pi_{\varepsilon_{n+1}}^N \left\{ p_n^i - \gamma_n \zeta_n^i \left[ e(u_n^i) - p_n^i \right] \right\}, \tag{6}$$

де $\pi_{\varepsilon_{n+1}}^N$ – проектор на одиничний $\varepsilon$-симплекс $S_{\varepsilon_{n+1}}^N \subseteq S^N$ [18]; $p_n^i \in S_{\varepsilon_n}^N$ – змішані стратегії $i$-го агента; $\gamma_n > 0$ – монотонно спадна послідовність невід'ємних величин, яка регулює величину кроку методу; $\varepsilon_n > 0$ – монотонно спадна послідовність невід'ємних величин, яка регулює швидкість розширення $\varepsilon$-симплексу; $\zeta_n^i \in R^1$ – поточний програш агента; $e(u_n^i)$ – одиничний вектор-індикатор вибору варіанта $u_n^i \in U^i$.

Проектування на розширюваний $\varepsilon_n$-симплекс $S_{\varepsilon_{n+1}}^N$ забезпечує виконання умови $p_n^i[j] \ge \varepsilon_n$, $j = 1..N$, необхідної для повноти статистичної інформації про вибрані чисті стратегії, а параметр $\varepsilon_n \to 0$, $n = 1, 2, ...$ використовується як додатковий елемент керування збіжністю рекурентного методу.

Стохастична гра розпочинається з ненавчених векторів змішаних стратегій зі значеннями елементів $p_0^i(j) = 1/N$, де $j = 1..N$. У наступні моменти часу динаміка векторів змішаних стратегій визначається марківським рекурентним методом (6).

У момент часу $n$ кожен гравець $i \in D$ на основі змішаної стратегії $p_n^i$ вибирає чисту стратегію $u_n^i$, за що до моменту часу $n+1$ отримує поточний програш $\zeta_n^i$, після чого обчислює змішану стратегію $p_{n+1}^i$ згідно з (6).

Завдяки динамічній перебудові змішаних стратегій на основі опрацювання поточних програшів, метод (6) забезпечує адаптивний вибір чистих стратегій у часі.

Параметри $\gamma_n$ та $\varepsilon_n$ визначають умови збіжності стохастичної гри і можуть бути задані так:

$$\gamma_n = \gamma n^{-\alpha}, \ \varepsilon_n = \varepsilon n^{-\beta}, \tag{7}$$

де $\gamma > 0$; $\alpha > 0$; $\varepsilon > 0$; $\beta > 0$.



Lviv Polytechnic National University Institutional Repository http://ena.lp.edu.ua

Збіжність стратегій (6) до оптимальних значень з імовірністю 1 та у середньоквадратичному визначається співвідношеннями параметрів $\gamma_n$ та $\varepsilon_n$, які повинні задовольняти базові умови стохастичної апроксимації [20].

Ефективність прийнятих рішень оцінюється:

1) функцією середніх втрат:

$$Z_n = \frac{1}{L}\sum_{i=1}^{L} Z_n^i , \tag{8}$$

де $L = |D|$ – потужність множини гравців;

2) середньою кількістю скоординованих стратегій гравців:

$$\mathrm{K}_n = \frac{1}{n}\sum_{t=1}^{n}\sum_{i=1}^{L} \chi\left(\xi_n^i[1] + \xi_n^i[2] = 0\right), \tag{9}$$

де $\chi() \in \{0,1\}$ – індикаторна функція події;

**Алгоритм розв'язування стохастичної гри**

1. Задати початкові значення параметрів:

$n = 0$ – початковий момент часу;

$L = |D|$ – кількість гравців;

$N$ – кількість чистих стратегій гравців;

$U^i = \{u^i(1), u^i(2), ..., u^i(N)\}$, $i = 1..L$ – вектори чистих стратегій гравців;

$p_0^i = (1/N, ..., 1/N)$, $i = 1..L$ – початкові змішані стратегії гравців;

$\gamma > 0$ – параметр кроку навчання;

$\alpha \in (0,1]$ – порядок кроку навчання;

$\varepsilon$ – параметр $\varepsilon$-симплекса;

$\beta > 0$ – порядок швидкості розширення $\varepsilon$-симплекса;

$d > 0$ – дисперсія завад;

$n_{\max}$ – максимальна кількість кроків методу.

2. Вибрати варіанти дій $u_n^i \in U^i$, $i = 1..L$ згідно з (4).

3. Отримати значення поточних програшів $\zeta_n^i$, $i = 1..L$ згідно з (1). Поточні значення гауссівського білого шуму обчислюють за формулою:

$$\mu_n = \sqrt{d}\left(\sum_{j=1}^{12} \omega_{j,n} - 6\right),$$

де $\omega \in [0,1]$ – дійсне випадкове число з рівномірним законом розподілу.

4. Обчислити значення параметрів $\gamma_n$, $\varepsilon_n$ згідно з (7).

5. Обчислити елементи векторів змішаних стратегій $p_n^i$, $i = 1..L$ згідно з (6).

6. Обчислити характеристики якості прийняття рішень $Z_n$ (8), $K_n$ (9).

7. Задати наступний момент часу $n := n + 1$ .

8. Якщо $n < n_{\max}$, то перейти на крок 2, інакше – кінець.

**Результати комп'ютерного моделювання**

Розв'язування стохастичної гри переслідування „хижак–жертва" виконаємо за допомогою ігрового методу (6) з параметрами: $D = \{(x, y)\}$, $x = 1..m$, $y = 1..m$, $m = 7$, $s = (4,4)$, $N = 4$, $U = (\text{front=0, right=1, behind=2, left=3})$, $\lambda = 0.5$, $\gamma = 1$, $\varepsilon = 0.999/N$, $\alpha = 0.01$, $\beta = 2$, $n_{\max} = 10^5$.

Метод (6) забезпечує розв'язування стохастичної гри у чистих стратегіях (на вершині одиничного симплексу). Початкові стратегії агентів є нескоординованими, що визначається вибором довільного напрямку руху $u \in U$. У процесі гри агенти навчаються узгоджувати свої дії. Отримані варіанти скоординованих стратегій навченої гри зображено на рис. 1.



Lviv Polytechnic National University Institutional Repository http://ena.lp.edu.ua

| | | | | | | |
|---|---|---|---|---|---|---|
| → | → | → | → | → | → | → |
| ↑ | → | → | → | → | ↓ | ↓ |
| ↑ | ↑ | ↑ | ↑ | → | ↓ | ↓ |
| ↑ | ↑ | ← | • | → | ↓ | ↓ |
| ↑ | ↑ | ← | ↓ | ↓ | ↓ | ↓ |
| ↑ | ← | ← | ← | ↓ | ↓ | ↓ |
| ↑ | ← | ← | ← | ← | ← | ← |

| | | | | | | |
|---|---|---|---|---|---|---|
| ↓ | ← | ← | ← | ← | ← | ← |
| ↓ | ↓ | ← | ← | ← | ← | ← |
| ↓ | ↓ | ← | ↑ | ↑ | ↑ | ↑ |
| ↓ | ↓ | ← | • | → | ↑ | ↑ |
| ↓ | ↓ | ↓ | ↓ | → | → | ↑ |
| ↓ | ↓ | ↓ | ↓ | → | → | ↑ |
| ↓ | ↓ | → | → | → | → | ↑ |

*Рис. 1. Самоорганізація стохастичної гри переслідування:*
*а – лівостороння реплікація дій; б – правостороння реплікація дій*

Позицію хижака зображено точкою у центральній частині рисунка. Стрілками позначено напрямки руху агентів-жертв для навченої стохастичної гри. Відслідковуючи напрямок стрілок, побачимо, що при лівосторонньому орієнтуванні ($Side(i) = Left$) сусідніх агентів має місце правостороння самоорганізація і, навпаки, при правосторонньому орієнтуванні ($Side(i) = Right$) – лівостороння самоорганізація системи.

На рис. 2 у логарифмічному масштабі зображено графіки функцій середньої кількості скоординованих стратегій $K_n$ та середніх програшів гравців $Z_n$, які характеризують ефективність самоорганізації стохастичної гри.

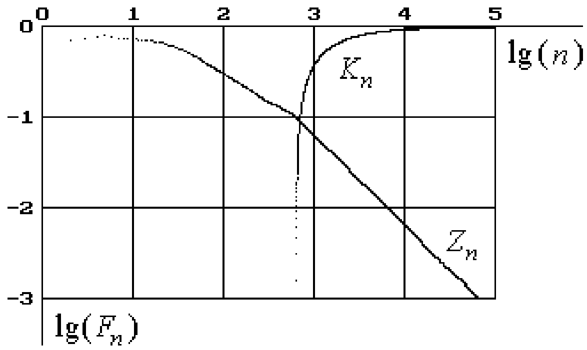


*Рис. 2. Характеристики самоорганізації системи*

Зменшення функції середніх програшів у часі свідчить про збіжність ігрового методу. Функція кількості скоординованих стратегій ілюструє досягнуте у процесі навчання зростання координації дій гравців, починаючи з $\sim 10^3$ кроків стохастичної гри.

Досліджена стійкість координації стохастичної гри при дії завад у вигляді білого шуму. Вплив дисперсії завад $d$ на ефективність ігрового методу (6) зображено на рис. 3. Коефіцієнт координації $K_n$ розраховано на вибірці $n$=100 тис. кроків стохастичної гри для $\alpha = 0.01$ та $\beta = 2$. Для значень $d \in [0; 0,36]$ координація стратегій гравців перевищує рівень у 80%. Зростання інтенсивності завад ($d > 0,36$) призводить до зменшення коефіцієнта координації стратегій.

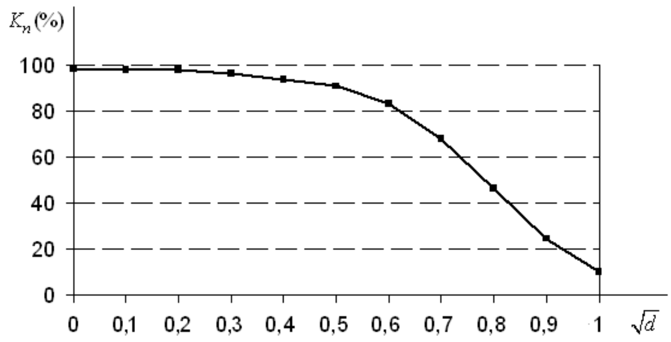


*Рис. 3. Вплив дисперсії на координацію гри*



Lviv Polytechnic National University Institutional Repository http://ena.lp.edu.ua

Збіжність ігрового методу визначається співвідношенням параметрів $\alpha$ та $\beta$, які визначають порядок швидкості збіжності ігрового методу. Значення цих параметрів повинні задовольняти базові умови стохастичної апроксимації [20].

Крім параметрів рекурентного методу, значно впливає на координацію розмірність стохастичної гри. Залежність коефіцієнта координації $K_n$ від кількості $L = m \times m$ гравців зображено на рис. 4. Прийнятною (понад 80 %) є координація стратегій для гри з кількістю агентів $L \leq 11$. Значення коефіцієнта координації розраховано протягом $n_{\max} = 10^5$ кроків стохастичної гри.

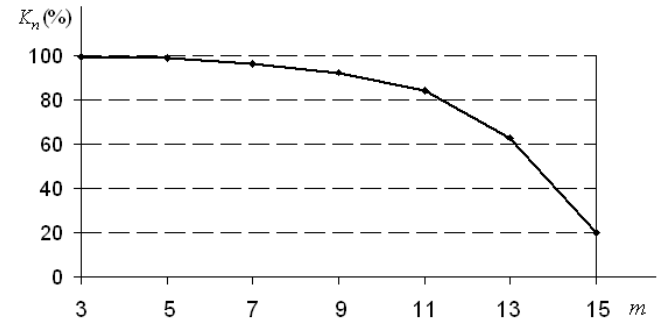


*Рис. 4. Вплив кількості гравців на координацію гри*

Зростання розмірності стохастичної гри призводить до зростання середньої кількості кроків, необхідних для досягнення належного рівня координації стратегій агентів.

Запропонований метод (6) розв’язування стохастичної гри агентів належить до класу реакційних методів (ґрунтується на опрацюванні реакцій середовища на дії агентів) і має відносно невисоку, степеневу швидкість збіжності, що пов’язано з апріорною невизначеністю системи. Інформацію збирають у процесі навчання адаптивною перебудовою векторів змішаних стратегій пропорційно до значень поточних програшів. Цей недолік долається високою швидкодією сучасних засобів обчислювальної техніки та можливістю розпаралелювання задачі або використання мережних засобів розподілених обчислень.

## Висновки

Розроблений ігровий метод забезпечує самоорганізацію МАС „хижак – жертва” через локально обумовлену координацію стратегій агентів. Кожен агент здійснює локальне спостереження за діями сусідніх агентів. Після обчислення відповідних штрафів за порушення умов координації агенти отримують поточні програші, які використовуються ними для формування динамічних векторів змішаних стратегій. Побудований на основі стохастичної апроксимації метод перетворення змішаних стратегій (6) забезпечує мінімізацію функцій середніх програшів на одиничних симплексах. Цілеспрямована динаміка змішаних стратегій перетворює локально скоординовані дії гравців на глобальну координацію (самоорганізацію) стохастичної гри, що проявляється у вигляді схеми правосторонньої або лівосторонньої динаміки МАС, коли колектив агентів-жертв поводиться як цілісний організм.

Координація стратегій агентів досягається під час розв’язування стохастичної гри у реальному масштабі часу на основі поточної інформації та її адаптивного опрацювання. Запропонований метод дозволяє знаходити розв’язки стохастичної гри у чистих стратегіях.

Ефективність координації стратегій системи оцінюється за допомогою характеристичних функцій середніх програшів та коефіцієнта координації. Зменшення функції середніх програшів та зростання коефіцієнта координації свідчать про збіжність ігрового методу.

Швидкість координації чистих стратегій агентів залежить від розмірності стохастичної гри, величини завад та параметрів ігрового методу. При зростанні кількості гравців та інтенсивності завад швидкість та ефективність ігрової координації МАС зменшуються.



Lviv Polytechnic National University Institutional Repository http://ena.lp.edu.ua

Достовірність отриманих результатів підтверджується повторювальністю значень розрахованих характеристик стохастичної гри для різних послідовностей випадкових величин.

Розвиток ігрової моделі „хижак – жертва” можливий у напрямку збільшення кількості агентів-хижаків та застосування інших методів навчання стохастичної гри.

*1. Поспелов Д.А. Многоагентные системы – настоящее и будущее / Д.А. Поспелов // Информационные технологии и вычислительные системы. – 1998. – № 1. – С. 14–21. 2. Городецкий В.И. Информационные технологии и многоагентные системы / В.И. Городецкий // Проблемы информатизации. – 1998. – Вып. 1. – С. 3–14. 3. Тарасов В.Б. От многоагентных систем к интеллектуальным организациям: философия, психология, информатика / В.Б.Тарасов. – М.: Эдиториал УРСС, 2002. – 352 с. 4. Швецов А.Н. Агентно-ориентированные системы: от формальных моделей к промышленным приложениям / А.Н. Швецов. – Вологодский гос. технич. унив. – 101 с.: [Электрон. ресурс]. – http://www.ict.edu.ru/ft/005656/62333e1-st20.pdf. 5. Wooldridge M. An Introduction to Multiagent Systems / M. Wooldridge. – John Wiley & Sons, 2002. – 366 pp. 6. Jennings N. R. Commitments and Conventions: The Foundation of Coordination in Multi-Agent Systems / N. R. Jennings // The Knowledge Engineering Review. – 1993. Vol. 8 (3). – P. 223–250. 7. Ashby W. R. Principles of the Self-Organizing Dynamic System / W. R. Ashby // Journal of General Psychology. – 1947. – Vol. 37. – P. 125–128. 8. Grasse P.P. La reconstruction du nid et les coordinations inter-individuelles chez Bellicositermes natalensis et Cubitermes sp. La theorie de la stigmergie: essai d’interpretation des termites constructeurs / P.P. Grasse // Insect Society. – 1959. – Vol. 6. – P. 41–84. 9. Пригожин И. Самоорганизация в неравновесных системах: От диссипативных структур к упорядоченности через флуктуации / И. Пригожин, Г. Николис. – М.: Мир, 1979. – 512 с. 10. Хакен Г. Синергетика. Иерархия неустойчивостей в самоорганизующихся системах и устройствах / Г. Хакен. – М.: Мир, 1985. 11. Хакен Г. Информация и самоорганизация. Макроскопический подход к сложным системам: Пер. с англ. / Г. Хакен. – М.: КомКнига, 2005. – 248 с. 12. Эйген М. Гиперцикл. Принципы самоорганизации макромолекул / М. Эйген, П. Шустер. – М.: Мир, 1982.- 260 с. 13. Omicini A. Self-Organisation & MAS An Introduction / A. Omicini, L. Gardelli: [Електрон. ресурс]. – http://unibo.lgardelli.com/teaching/2007-selforg-mas.pdf. 14. Самоорганизация и многоагентные системы. I. Модели многоагентной самоорганизации / В. И. Городецкий // Изв. РАН. Теория и системы управления. – 2012. – N 2. – С. 92–120. 15. Самоорганизация и многоагентные системы. II. Приложения и технология разработки / В. И. Городецкий // Изв. РАН . Теория и системы управления. – 2012. – N 3. – С. 55 –75. 16. Доманский В.К. Стохастические игры / В.К. Доманский // Математические вопросы кибернетики. – 1988. – № 1. – С. 26–49. 17. Вольтерра В. Математическая теория борьбы за существование. Пер. с франц. О. Н. Бондаренко / В. Вольтерра. – М.: Наука, 1976. – 287 с. 18. Назин А.В. Адаптивный выбор вариантов: Рекуррентные алгоритмы / А.В. Назин, А.С. Позняк. – М.: Наука, 1986. – 288 с. 19. Мулен Э. Теория игр с примерами из математической экономики / Э. Мулен. – М.: Мир, 1985. – 200 с. 20. Граничин О.Н. Введение в методы стохастической аппроксимации и оценивания: Учеб. пособие / О.Н. Граничин. – СПб.: Издательство С.-Петербургского университета, 2003. – 131 с.*



Lviv Polytechnic National University Institutional Repository http://ena.lp.edu.ua