

ПАТЕРНИ САМООРГАНІЗАЦІЇ СТРАТЕГІЙ У ГРІ МОБІЛЬНИХ АГЕНТІВ

Петро Кравець¹, Ростислав Юринець², Ярослав Кісь³

Національний університет “Львівська політехніка”

¹ Petro.O. Kravets@lpnu.ua, ORCID 0000-0001-8569-423X

² Rostyslav.V. Yurynets@lpnu.ua, ORCID 0000-0003-3231-8059

³ Yaroslav.P. Kis@lpnu.ua, ORCID 0000-0003-3421-2725

© Кравець П., Юринець Р., Кісь Я., 2020

Розглянуто актуальну проблему самоорганізації стратегій стохастичної гри багатоагентної системи. Проявом самоорганізації є формування скоординованих поведінкових патернів групи мобільних агентів, наділених здатністю переміщуватися в обмеженому дискретному просторі.

Агент – це автономний об’єкт, який може взаємодіяти із навколишнім середовищем, іншими агентами і людиною для вибору варіантів рішень. Багатоагентна система складається із групи агентів, які виконують спільну роботу, співпрацюючи між собою у межах локальних підмножин агентів. Поведінковий патерн багатоагентної системи – це візуалізована форма впорядкованого переміщення агентів, яка виникає із їх початкового хаотичного руху під час навчання стохастичної гри.

Повторювальна стохастична гра полягає у реалізації керованого випадкового процесу вибору варіантів рішень. Для цього ігрові агенти випадково, одночасно і незалежно вибирають одну із власних чистих стратегій у дискретні моменти часу. Чисті стратегії гравців визначають напрямки переміщення у двовимірному просторі: вперед, назад, направо, наліво. Після завершення вибору усіх стратегій обчислюють поточні програші гравців. Для формування впорядкованого переміщення кожен агент повинен повторювати дії сусідніх агентів. Тоді поточні програші визначаються індикаторною функцією подібності стратегій сусідніх гравців. Обчислені поточні програші використовують для адаптивного перерахунку змішаних стратегій гравців. Імовірність вибору чистої стратегії збільшується, якщо її реалізація призвела до зменшення поточного програшу. Під час повторювальної гри агенти сформують вектори змішаних стратегій, які мінімізують функції середніх програшів гравців.

Для розв’язування ігрової задачі побудови патернів самоорганізації багатоагентної системи використано адаптивний марківський рекурентний метод, побудований на основі стохастичної апроксимації модифікованої умови доповняльної нежорсткості, яка справедлива у точках рівноваги за Нешем. Для нормування елементів векторів змішаних стратегій застосовано операцію їх проектування на одиничний розширюваний епсілон-симплекс. Збіжність ігрового методу забезпечується дотриманням фундаментальних умов та обмежень стохастичної оптимізації.

Комп’ютерне моделювання підтвердило можливість застосування моделі стохастичної гри для побудови патернів самоорганізації багатоагентної системи. Форма отриманих патернів залежить від способу локального орієнтування мобільних агентів. Під час комп’ютерного експерименту отримано вихрові та лінійні патерни переміщення агентів. Достовірність експериментальних досліджень підтверджується подібністю отриманих результатів для різних послідовностей випадкових величин.

Результати цієї роботи доцільно застосувати для вивчення патернів колективної поведінки агентів для глибшого розуміння процесів самоорганізації природних систем та для побудови розподілених систем прийняття рішень.

Ключові слова: самоорганізація, патерн поведінки, багатоагентна система, стохастична гра, марківський рекурентний метод.

Вступ

Зародження у середині ХХ ст. кібернетики як науки про загальні закони управління та опрацювання інформації у складних системах історично обумовлено розвитком електроніки, автоматики, комп'ютерної техніки, телекомунікації та інформаційних систем. Прогрес кібернетичної науки у наш час пов'язують з інформаційними мережами з мобільними засобами комунікації та інтелектуальним аналізом даних, робототехнікою, нанотехнологіями, квантовими обчислювальними системами. Натхненником наукових теорій про загальні закони управління, збирання, зберігання, перероблення та передавання інформації в організованих системах була і є жива природа. Так, ще у 1948 році американський математик Норберт Вінер у роботі „Кібернетика, або керування і зв'язок в живих організмах і машинах” вказав на подібність процесів керування, що проходять у живих організмах та у складних технічних системах.

Запозичення методів живої природи для розроблення інженерних пристроїв та систем призвело до становлення нової науки за назвою “біоніка” (1960 р.), яка виникла на стику кібернетики, біофізики, біохімії, інженерної психології та інших наук [1]. Задачею біоніки є вивчення принципів та закономірностей функціонування живих організмів з метою застосування у штучно організованих системах, необхідність чого зумовлена надзвичайною ефективністю та досконалістю природних форм, відшліфованих у ході еволюції на космічній шкалі часу. Вважається, що тренд еволюції живих організмів визначається адаптивним методом проб і помилок з селекцією кращих, стійкіших до виживання форм. Найскладнішою проблемою біоніки є моделювання роботи людського мозку для побудови систем штучного інтелекту та вдосконалення засобів обчислювальної техніки.

Більшість із практичних застосувань біоніки пов'язано із використанням фізико-хімічних властивостей або поведінкових моделей окремих живих організмів. До успішного застосування біонічних методів у техніці можна віднести: ультразвукові, теплові та радіолокатори, прилади нічного бачення, пристрої з ефектом котячого ока, пристрої селективного (вибіркового) зору, засоби камуфляжу, електронний ніс, біонічні імплантати, екзоскелети, передбачення погоди і землетрусів, реактивні гідродвигуни, обтічні форми плавальних та літальних апаратів, ефективно використання потоків повітря у планерній авіації, застібки-липучки, засоби з присосками для пересування по ковзких поверхнях, гідрофобні матеріали, каркасні конструкції в архітектурі, виготовлення надміцних матеріалів, засоби буріння та дренажні системи, спеціальні ріжучі та колючі інструменти, енергоефективні технології, системи кодування інформації, розроблення інтелектуальних систем, людиноподібні роботи (андроїди), багатонігі комахоподібні роботи, промислові роботи тощо. Багато таємниць природи ще не розкрито і, відповідно, існує чимало можливостей у розвитку біоніки.

Важливим, але значно складнішим, є вивчення аспектів самоорганізованої поведінки, яка виникає у популяціях живих істот та їх перенесення для керування або прийняття рішень у розподілених технічних та інформаційних системах.

Самоорганізація – це цілеспрямований просторово-часовий процес створення, упорядкування або вдосконалення структури та функцій складної відкритої динамічної системи за рахунок її внутрішніх факторів (без організуючого зовнішнього впливу). Конструктивним механізмом самоорганізації складних систем є багатоваріантність, альтернативність, випадковість та селективність можливих дій активних елементів.

Зовнішніми проявами самоорганізації є утворення впорядкованих структур у вигляді патернів, які проявляються на макрорівні та є наслідком локальних взаємодій елементів системи на

мікрорівні. Означення самоорганізації як формування патерну за допомогою простих локальних взаємодій компонентів системи, виконано у роботі [2].

Патерни зустрічаються як у живій, так і неживій природі. У загальному розумінні патерни візуалізують регулярність форм природного світу (симетрії, дерева, фрактали, спіралі, хвилі, вихори, меандри, дюни, бульбашки, піни, мозаїки, кристали, плями, смужки тощо), які виникають у процесі еволюції. У роботі [3] визначено, що патерн самоорганізації – це регулярний (невипадковий) розподіл конструктивних елементів систем у часі або просторі, що дає змогу цілісно описати систему, не здійснюючи характеристику усіх її елементів.

Крім впорядкованого розподілу конструктивних елементів, у природі часто зустрічаються патерни колективної поведінки живих організмів: ссавців, птахів, риб, комах, бактерій. Колективна самоорганізація групи живих організмів, як правило, формується на основі їх локальної взаємодії і проявляється у вигляді сформованих глобальних поведінкових патернів, наприклад, скоординованого переміщення стада тварин, зграї птахів, косяка риб, переслідування жертв групою хижаків, синхронізованого світіння колонії комах тощо. Поведінкові патерни визначають форму впорядкованих дій агентів, яка за певних умов виникає із хаосу. На практиці поведінкові патерни самоорганізації знаходять застосування у політичних системах, соціології, психології, системах колективного прийняття рішень.

У [4] вказується, що поведінкові патерни самоорганізації також виникають у штучних колективних системах, наприклад, виникнення соціальної поведінки групи роботів, у якій під час взаємодії виявляється лідер та ведені роботи. Незважаючи на гуманні закони розвитку робототехніки, сформульовані А. Азімовим в оповіданні “Я, робот” (1941 р.), як зазначено у [5, 6], соціальна поведінка агентів-роботів з розвиненим штучним інтелектом у майбутньому може бути небезпечною для людини.

Для побудови поведінкових патернів самоорганізації використовують генетичні, імунні, зграйні, ройові та мурашині алгоритми. Крім того, для дослідження поведінкових патернів можна використати методи підсиленого Q-навчання, мереж автоматів, штучних нейронних мереж, прихованих марківських мереж, детермінованих або стохастичних ігор.

Комп’ютерне моделювання побудови патернів поведінки популяції живих істот можна виконати за допомогою штучних багатоагентних систем [7]. Агент – це автономна інформаційно-програмна сутність, яка може опрацьовувати сигнали від зовнішнього середовища, інших агентів та від людини для вибору варіантів можливих дій або прийняття рішень для досягнення поставлених цілей. У багатоагентній системі агенти не можуть досягнути мети поодиночі, для цього вони повинні вступати у діалог між собою.

Агенти можуть бути реалізовані у вигляді машин (роботів), або у вигляді програм. Мобільний агент – це: 1) робот, стратегії дій якого визначають вектори можливих переміщень у заданому просторі; 2) програма, яка може перемістити своє виконання на інші процесори у межах комп’ютерної мережі. У цій роботі розглядається перший варіант реалізації мобільних агентів.

Враховуючи, що у природних популяціях має місце співпраця або конкуренція у ході вирішення системної задачі, відповідну багатоагентну модель можна розглядати як гру, спрямовану на відтворення патернів колективної поведінки агентів. Для імітації цілеспрямованого випадкового вибору варіантів дій в умовах невизначеності застосовують стохастичні ігрові моделі [8–11].

Патерн поведінки агента – це відтворювана реакція агента (з точністю до ізоморфізму дій) на зовнішні подразники навколишнього середовища, спрямована на досягнення конкретної мети. Для багатоагентної системи патерн – це системна, глобально скоординована поведінка групи агентів, яка виникає на основі їх локальної взаємодії під час багатокрокового адаптивного навчання.

Вивченню механізмів адаптивної координації та самоорганізації багатоагентних систем приділяється значна увага у фаховій науковій літературі [9–16]. Однак, у відомих наукових працях недостатньо висвітлено питання формування поведінкових патернів у природних і, особливо, у штучно створених розподілених системах. Тому актуальним є дослідження умов та механізмів зародження патернів самоорганізації штучних багатоагентних систем, що виникають як результат адаптивного навчання стохастичної гри.

Відтворення патернів самоорганізації за допомогою багатоагентної стохастичної гри дозволить виявити причини та механізми їх утворення в природі, глибше зрозуміти закони природної еволюції, застосувати еволюційні алгоритми на практиці для побудови різноманітних систем на основі розподіленого штучного інтелекту.

Мета роботи

Метою цієї роботи є побудова математичної моделі та методу розв'язування стохастичної гри, яка у ході навчання на основі локальної координації дій агентів забезпечує глобальну самоорганізацію системи у вигляді стійких патернів колективних стратегій агентів.

Математична модель стохастичної гри

Нехай задано множину агентів $D \neq \emptyset$, кожен з яких має $N_i \geq 2$ стратегій переміщення $U^i = (u^i[1], u^i[2], \dots, u^i[N_i])$ в обмеженій області дискретного (клітинного) простору. Чисті стратегії $u[k]$, $k = 1..N_i$ визначають напрямки (вектори) можливих рухів агентів – вперед, назад, наліво, направо тощо. За наявності стратегій прямого і зворотного переміщень у кожному напрямку клітинний простір матиме $2^{-1} \max_{i \in D} N_i$ вимірів.

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

Варіанти переміщення $\{u_n^i\}$ вибираємо на основі динамічних векторів змішаних стратегій $p_n^i = (p_n^i(1), p_n^i(2), \dots, p_n^i(N_i)) \forall i \in D$, елементи $p_n^i(j)$, $j = 1..N_i$ яких є умовними імовірностями вибору чистих стратегій, які визначаються передісторією чистих стратегій $\{u_t^i | t = 1, 2, \dots, n-1\}$ та отриманими за це відповідними програшами $\{\zeta_t^i | t = 1, 2, \dots, n-1\}$. Змішані стратегії набувають значення на N_i -вимірних одиничних симплексах:

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

Значення чистих стратегій визначають за умови:

$$u_n^i = \left\{ u^i(l) \mid l = \arg \min_l \sum_{k=1}^{N_i} p_n^i[k] > \omega \quad (k, l = 1..N_i) \right\}, \quad (1)$$

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

За вибір варіанта u_n^i кожен i -й агент отримує випадковий програш $\zeta_n^i = \zeta_n^i(u_n^{D_i})$, який визначається комбінованими стратегіями сусідніх гравців $u^{D_i} \in U^{D_i} = \times_{j \in D_i} U^j$ з множини $D_i \subseteq D$, $D_i \neq \emptyset \quad \forall i \in D$. Вважають, що поточні програші мають постійне математичне сподівання та обмежений другий момент, які не відомі агентам апріорі.

Агенти мають можливість орієнтуватися в навколишньому середовищі, визначаючи стратегії дій сусідніх агентів, розміщених зліва, справа, попереду, позаду, згори, знизу. Серед можливих способів визначення напрямку руху кожного i -го агента виберемо лівостороннє або правостороннє орієнтування, яке полягає у повторенні дій сусіднього агента, розміщеного відповідно зліва або справа від вектора руху i -го агента. Причому, дії сусідніх агентів є рекомендаційними, кожен агент має свободу незалежного випадкового вибору власних стратегій поведінки згідно із дискретним розподілом, що задається його змішаною стратегією.

Відхилення власної стратегії від стратегії сусіднього агента оцінюють штрафом

$$\xi_n^i[1] = \chi(\omega \leq p_{left}^i) \chi(u_n^i \neq u_n^{Left(i)}) + \chi(\omega > p_{left}^i) \chi(u_n^i \neq u_n^{Right(i)}), \quad (2)$$

де $\chi() \in \{0,1\}$ – індикаторна функція події; $p_{left} \in [0,1]$ – імовірність лівостороннього орієнтування ($p_{left} + p_{right} = 1$); $Left(i)$ та $Right(i)$ – операторні функції визначення ідентифікатора сусіднього агента з множини D_i , розміщеного відповідно зліва або справа від напрямку u_n^i руху i -го агента.

Під час випадкового переміщення агентів можливі колізії їх зіткнення, коли сусідні агенти вибрали зустрічний напрям руху. У випадку зіткнення агентів нараховується штраф:

$$\xi_n^i[2] = \chi(u_n^i \uparrow \downarrow u_n^{Front(i)}), \quad (3)$$

де $\uparrow \downarrow$ – ситуація зіткнення; $Front(i)$ – операторна функція визначення ідентифікатора сусіднього агента, розміщеного у напрямку u_n^i руху i -го агента.

Після завершення вибору варіантів $u_n^i \quad \forall i \in D$ гравці отримують поточний програш, сформований зі складових (2) та (3):

$$\zeta_n^i = \lambda \xi_n^i[1] + (1 - \lambda) \xi_n^i[2], \quad (4)$$

де $\lambda \in [0,1]$ – ваговий коефіцієнт, який задає важливість одного із штрафних критеріїв.

Під час повторювальної гри з переміщення агентів у момент часу n середні програші гравців досягнуть значення:

$$Z_n^i = \frac{1}{n} \sum_{t=1}^n \zeta_t^i \quad \forall i \in D. \quad (5)$$

Кожен гравець змінює стратегії поведінки так, щоб мінімізувати власну функцію середніх програшів (5):

$$\overline{\lim}_{n \rightarrow \infty} Z_n^i \rightarrow \min_{u_n^i} \quad \forall i \in D. \quad (6)$$

Отже, стохастична гра переміщення агентів полягає у тому, що на основі спостереження поточних програшів $\{\zeta_n^i\}$ (4) кожен гравець $i \in D$ повинен навчитися вибирати варіанти переміщення (чисті стратегії) $\{u_n^i\}$ (1) так, щоб забезпечити виконання системи критеріїв (6) у часі $n = 1, 2, \dots$.

Залежно від способу формування послідовності стратегій $\{u_n^i\} \forall i \in D$ ($n = 1, 2, \dots$) розв'язки гри задовольнятимуть одну з умов колективної оптимальності, наприклад, Неша, Парето або іншу [17].

Ефективність самоорганізації стохастичної гри оцінюється:

1) системною функцією середніх програшів:

$$Z_n = |D|^{-1} \sum_{i \in D} Z_n^i; \quad (7)$$

2) коефіцієнтом координації стратегій гравців (відносно кількістю скоординованих стратегій гравців):

$$K_n = \frac{1}{n |D|} \sum_{t=1}^n \sum_{i \in D} \chi(\zeta_t^i = 0). \quad (8)$$

Зменшення значення функції $Z_n \geq 0$ та зростання функції $K_n \in [0,1]$ свідчатимуть про самоорганізацію стратегій стохастичної гри.

Метод розв'язування стохастичної гри

Цілеспрямований випадковий вибір чистих стратегій, що визначається цільовими умовами (6), виконаємо застосуванням адаптивних марківських методів, які забезпечують оптимальний у середньому вибір варіантів дій на основі опрацювання поточної інформації про систему.

Марківський рекурентний метод розв'язування стохастичної гри побудуємо методом стохастичної апроксимації умови доповняльної нежорсткості [9–11, 17, 18] відповідної детермінованої гри, справедливої для змішаних стратегій у точці рівноваги за Нешем:

$$p_{n+1}^i = \pi_{\varepsilon_{n+1}}^{N_i} \left\{ p_n^i - \gamma_n \zeta_n^i \left[e(u_n^i) - p_n^i \right] \right\}, \quad (9)$$

де $\pi_{\varepsilon_{n+1}}^{N_i}$ – проектор на одиничний ε -симплекс $S_{\varepsilon_{n+1}}^{N_i} \subseteq S^{N_i}$ [19]; $p_n^i \in S_{\varepsilon_n}^{N_i}$ – змішані стратегії i -го агента; $\gamma_n > 0$, $\varepsilon_n > 0$ – монотонно спадні послідовності додатних величин; $\zeta_n^i \in R^1$ – поточний програш агента; $e(u_n^i)$ – одиничний вектор-індикатор вибору варіанта $u_n^i \in U^i$.

Оператор проектування на розширюваний ε_n -симплекс $S_{\varepsilon_{n+1}}^N$ забезпечує виконання умов $p_n^i[j] \geq \varepsilon_n$, $\sum_{j=1}^{N_i} p_n^i[j] = 1$ і є необхідним для повноти статистичної інформації про вибрані чисті стратегії.

Параметр γ_n регулює величину кроку рекурентного методу, а параметр ε_n – швидкість розширення ε -симплекса. Значення цих параметрів обчислюють так:

$$\gamma_n = \gamma n^{-\alpha}, \quad \varepsilon_n = \varepsilon n^{-\beta}, \quad (10)$$

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

Співвідношення параметрів (10) визначатимуть умови збіжності стохастичної гри. Збіжність стратегій (9) до колективно-оптимальних значень з імовірністю 1 та у середньоквадратичному визначається з фундаментальних умов стохастичної апроксимації [20].

На початку гри змішані стратегії забезпечують рівноімовірний вибір варіантів переміщення агентів: $p_0^i = (1/N_i, \dots, 1/N_i) \quad \forall i \in D$. Потім з часом $n=1, 2, \dots$ значення змішаних стратегій змінюється згідно з (9). Динамічна перебудова змішаних стратегій методом (9) забезпечує адаптивний вибір чистих стратегій у часі. Метод підсилює імовірності вибору тих чистих стратегій, які у середньому призводять до зменшення величини штрафів за вибрані напрямки переміщення агентів.

Отже, у момент часу n кожен гравець $i \in D$ на основі змішаної стратегії p_n^i вибирає чисту стратегію u_n^i згідно з (1), за що отримує поточний програш ζ_n^i (4), після чого обчислює змішану стратегію p_{n+1}^i згідно з (9).

Результати комп'ютерного моделювання

Для прикладу розглянемо ігрову задачу побудови патернів самоорганізації агентів у двовимірному клітинному просторі: $D = \{(x, y) | x=1..m, y=1..m\}$, $m=10$. У кожну клітину розмістимо агента з $N_i=4$ стратегіями переміщення: $U^i = (\text{вперед}=0, \text{направо}=1, \text{назад}=2, \text{наліво}=3)$. Агенти, розміщені на краях клітинного простору, матимуть меншу кількість можливих стратегій переміщення.

Розв'язування стохастичної гри переміщення агентів виконаємо ігровим методом (9) із такими параметрами: $\lambda = 0,5$, $\gamma = 1$, $\varepsilon = 0,999/N_i$, $\alpha = 0,01$, $\beta = 2$. Період моделювання задається $n_{\max} = 10^4$ кроками стохастичної гри.

В основу стохастичного ігрового методу покладено адаптивний метод проб і помилок, який імітує природну еволюцію самонавчання просторовій координації дій множини агентів. Природна еволюція розгортається на космічній шкалі часу, тому і навчання стохастичної гри вимагає значної кількості випробувань, але завдяки комп'ютерній реалізації забезпечується значне прискорення часу моделювання процесів самоорганізації колективної поведінки агентів.

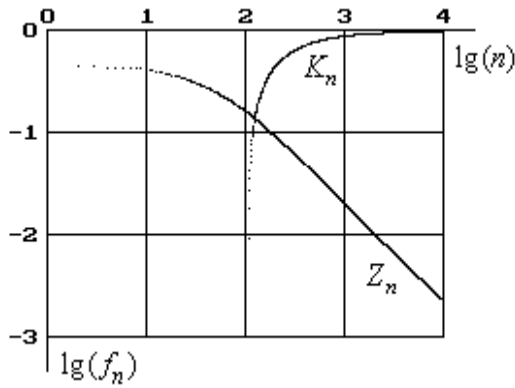
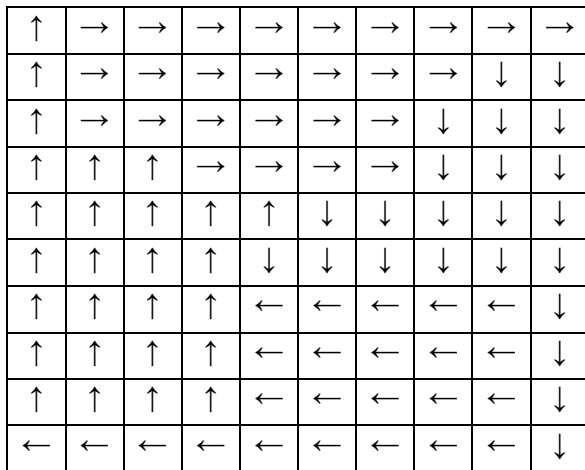


Рис. 1. Характеристики самоорганізації стохастичної гри

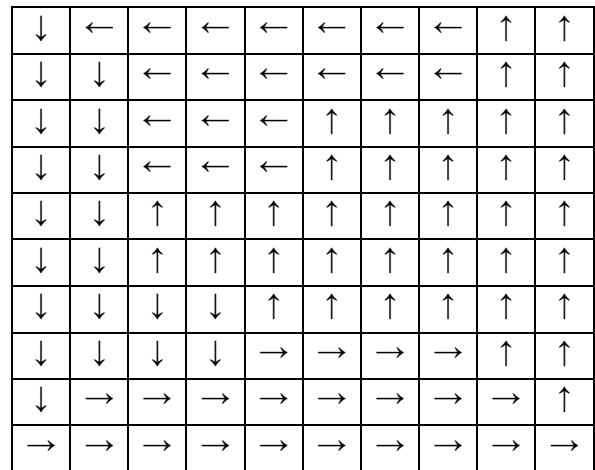
На рис. 1 у логарифмічному масштабі зображено графіки функцій середніх програшів гравців Z_n (7) та коефіцієнта координації стратегій K_n (8), які характеризують ефективність самоорганізації стохастичної гри переміщення агентів.

Спадання графіка функції середніх програшів та зростання коефіцієнта координації дій агентів у часі вказують на збіжність ігрового методу. Для заданих значень параметрів стохастичної гри за умови детермінованого орієнтування на дії сусідніх гравців координація стратегій досягається, починаючи з $\sim 10^2$ кроків навчання стохастичної гри. Із зростанням кількості агентів та кількості чистих стратегій збільшується кількість пошукових кроків, необхідних для самоорганізації стохастичної гри.

Метод (9) забезпечує розв’язки стохастичної гри у чистих стратегіях. На початку гри стратегії переміщення гравців в одному із чотирьох напрямків є хаотичними. Одну із реалізацій патернів скоординованих стратегій для навченої гри зображено на рис. 2 окремо для лівостороннього ($p_{left} = 1, p_{right} = 1 - p_{left} = 0$) та правостороннього ($p_{right} = 1$) орієнтування агентів.



а) лівостороннє орієнтування



б) правостороннє орієнтування

Рис. 2. Вихрові патерни самоорганізації стохастичної гри переміщення агентів

Стрілки позначають напрямки переміщення агентів для навченої стохастичної гри. Отримані вихрові патерни переміщення агентів подібні на траєкторії кружляння зграї птахів чи косяків риб. Йдучи за напрямками стрілок, побачимо, що лівостороннє орієнтування агентів призводить до формування правостороннього патерну (рух за годинниковою стрілкою), а правостороннє орієнтування – до лівостороннього патерну (рух проти годинникової стрілки). Починаючи з хаотичного переміщення, агенти у ході гри навчилися діяти злагоджено, що проявилось у вигляді регулярних патернів колективної поведінки. Повторення експерименту з іншою послідовністю випадкових величин призведе до утворення ізоморфних поведінкових патернів самоорганізації багатоагентної стохастичної гри.

Якщо агенти здійснюють лівостороннє та правостороннє орієнтування з однаковою імовірністю $p_{left} = p_{right} = 0,5$, то вони навчатимуться здійснювати скоординоване лінійне переміщення в одному із чотирьох напрямків, і патерни самоорганізації матимуть вигляд, як на рис. 3.

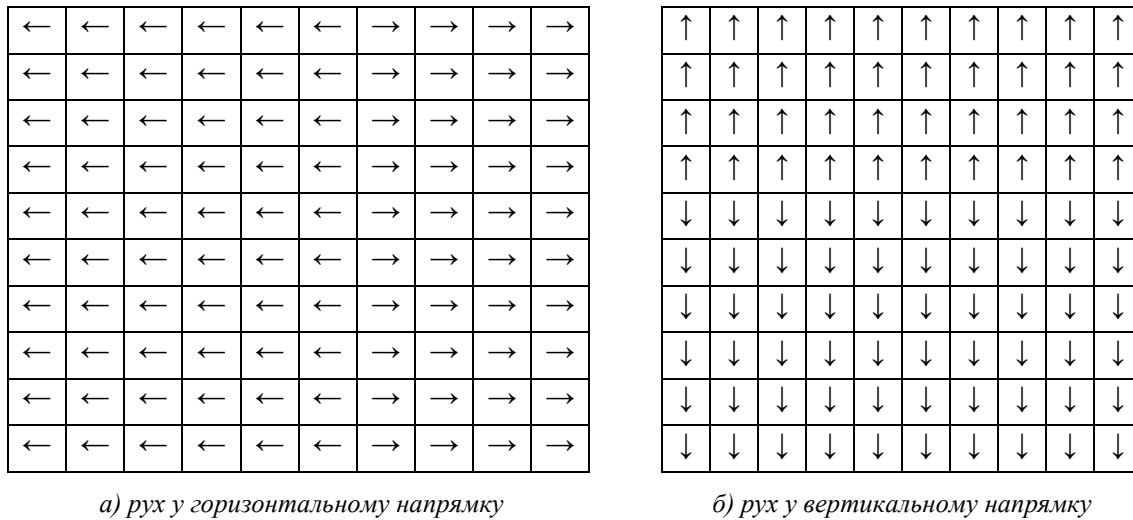


Рис. 3. Лінійні патерни самоорганізації стохастичної гри переміщення агентів

Як видно на рис. 3, переміщення агентів скоординовані з агентами, розміщеними як зліва, так і справа від напрямків їхнього руху. Усі агенти рухаються в одному горизонтальному або вертикальному напрямку. Частина агентів рухається у протилежному напрямку із збереженням одночасного лівостороннього і правостороннього орієнтування. Щоб усі агенти переміщувалися тільки в одному напрямку, необхідно додатково ввести критерій їх орієнтування на дії розміщених попереду агентів.

Для інших значень імовірностей лівостороннього та правостороннього орієнтування отримаємо проміжні патерни самоорганізації стохастичної гри, які міститимуть фрагменти скоординованого лінійного та вихрового руху.

Застосування імовірнісного орієнтування на дії сусідніх агентів призводить до значного зростання часу навчання стохастичної гри. Для цього випадку зростання коефіцієнта координації починається з $\sim 10^4$ кроків гри.

На рис. 2 та рис. 3 наочно продемонстровано синергетичний ефект емерджентності – поява інтегральної властивості системи агентів, яка не характерна для її складових. Локальна координація дій агентів під час навчання стохастичної гри призводить до глобальної самоорганізації системи у вигляді сформованих патернів поведінки, коли колектив ігрових агентів поводить як цілісний організм.

Висновки

1. Патерни самоорганізації можуть виникати не тільки у природних, але і у штучно створених розподілених системах зі зворотними зв'язками завдяки механізмам саморегулювання, координації дій, навчання та адаптації до невизначеностей.

2. Концептуальний аналіз процесів самоорганізації живої та неживої природи дав змогу запропонувати стохастичний ігровий метод для вивчення самоорганізації штучних багатоагентних систем.

3. Вигляд патернів самоорганізації стохастичної гри у просторі стратегій визначається способом формування поточних платежів та механізмом вибору стратегій агентів.

4. Побудова правил локальної взаємодії між агентами, яка під час стохастичної гри призводить до появи патернів самоорганізації, дає змогу краще вивчити явища колективної поведінки у природі з метою їх практичного застосування.

5. Ігрове моделювання процесів колективної поведінки агентів можна використати для побудови розподілених технічних та інформаційних систем підтримки прийняття рішень, синхронізації руху, управління групами роботів, управління складними проектами тощо.

6. Нові результати дослідження у цьому напрямку можна очікувати від запровадження інших критеріїв формування локально-обумовлених платежів гравців та інших методів навчання стохастичної гри.

Список літератури

1. Гийо, А., Мейє, Ж.-А. (2013). *Бионика. Когда наука имитирует природу*. Москва: Техносфера.
2. Ball, Ph. (1999). *The Self-Made Tapestry: Pattern Formation in Nature*. Oxford: Oxford University Press.
3. Шаталкин, А. И. (2012). *Таксономия. Основания, принципы и правила*. Москва: Товарищество научных изданий КМК.
4. Уорвик, К. (1999). *Наступление машин. Почему миром будет править новое поколение роботов*. Москва: Наука / Интерпериодика.
5. Баррат, Дж. (1915). *Последнее изобретение человечества: искусственный интеллект и конец эры Homo sapiens*. Москва: Альпина Нон-Фикшн.
6. Кайку, М. (2017). *Фізика майбутнього*. Львів: Літопис.
7. Weiss, G. (2013). *Multiagent Systems. Second Edition*. The MIT Press.
8. Chen, B.-S. (2019). *Stochastic Game Strategies and their Applications*. CRC Press.
9. Кравець, П.О. (2013). Ігрова координація та самоорганізація стратегій в мультиагентній моделі “хижак-жертва”. *Вісник Нац. ун-ту “Львівська політехніка”*. Серія: “Комп’ютерні науки та інформаційні технології”, 751, 136–146.
10. Кравець, П. О. (2015). Ігрова модель самоорганізації мультиагентних систем. *Вісник НУ “Львівська політехніка”*. Серія: “Інформаційні системи та мережі”, 829, 161–176.
11. Kravets, P., Burov, Y., Lytvyn, V., Vysotska, V. (2019). Gaming Method of Ontology Clusterization. *Webology*, 16 (1), 55–76.
12. Иванов, Д. Я. (2016). Формирование строя в большой группе мобильных роботов с использованием метода паттернов. *Робототехника и техническая кибернетика*, 11 (2), 39–44.
13. Кучеров, Д. П., Козуб, А. М., Костина, О. М. (2017). Управління мультиагентною системою в потенціальному полі. *Озброєння та військова техніка*, 14 (2), 55–61.
14. Кузнецов, А. В. (2018). Модель движения и взаимодействия системы интеллектуальных агентов. *Вестник ВГУ: Серия: Системный анализ и информационные технологии*, 2, 130–138.
15. Коноваленко, О. С., Брусенцев, В. О. (2019). Мультиагентні системи управління та підтримки прийняття рішень. *Вісник Національного технічного університету «ХПИ»*. Серія: *Машинознавство та САПР*, 1, 18 – 27. DOI:10.20998/2079-0775.2019.1.03.
16. Ткачев, В. Н., Токарев, В. В., Чурюмов, Г. И. (2019). Разработка алгоритма мультиагентного управления группой мобильных “s-bot”. *Київ: Реєстрація, зберігання і обробка даних*, 21 (1), 46–56.
17. Ungureanu V. (2018). *Pareto-Nash-Stackelberg Game and Control Theory: Intelligent Paradigms and Applications*. Springer.
18. Neogy, S. K., Vapat, R. B., Dubey Dipti. (2018). *Mathematical Programming and Game Theory*. Springer.
19. Назин, А. В., Позняк, А. С. (1986). *Адаптивный выбор вариантов: Рекуррентные алгоритмы*. Москва: Наука.
20. Kushner, H., Yin, G. G. (2013). *Stochastic Approximation and Recursive Algorithms and Applications*. Springer Science & Business Media.

References

1. Guillot, A., Meyer, J.-A. (2013). *Bionics. When the Science Imitates the Nature*. Moscow: Technosphere.
2. Ball, Ph. (1999). *The Self-Made Tapestry: Pattern Formation in Nature*. Oxford: Oxford University Press.
3. Shatalkin, A. I. (2012). *Taxonomy. The bases, principles and rules*. Moscow: KMK Scientific Press Ltd.
4. Warwick, K. (1999). *March of the machines. Why the new race of robots will rule the world*. Moscow: Science / Interperiodicals.
5. Barrat, J. (2015). *Our Final Invention: Artificial Intelligence and the End of the Human Era*. Alpina Non-Fiction.
6. Kaku, M. (2017). *Physics of the Future*. Lviv: Litopys.
7. Weiss, G. (2013). *Multiagent Systems. Second Edition*. The MIT Press.
8. Chen, B.-S. (2019). *Stochastic Game Strategies and their Applications*. CRC Press.
9. Kravets, P. A. (2013). Game coordination and self-organizing of strategies in multiagent “predator-prey” models. *Bulletin of “Lviv polytechnic”*. Series: “Computer sciences and information technologies”, 751, 136–146.

10. Kravets, P. A. (2015). Game model of self-organizing of multiagent systems. *Bulletin of "Lviv polytechnic". Series: "Information systems and networks"*, 829, 161–176.
11. Kravets, P., Burov, Y., Lytvyn, V., Vysotska, V. (2019). Gaming Method of Ontology Clusterization. *Webology*, 16 (1), 55–76.
12. Ivanov, D. Y. (2016). System formation in the big group of mobile robots with use of a method of patterns. *Robotics and technical cybernetics*, 11 (2), 39–44.
13. Kuchеров, D. P., Kozub, A. M., Kostyna, O. M. (2017). Movement of the multi agent system in potential field. *Weapons and Military Equipment*, 14 (2), 55–61.
14. Kuznetsov, A. V. (2018). Model of movement and interaction of system of intellectual agents. *Bulletin VSU: Series: System analysis and information technology*, 2, 130–138.
15. Konovalenko, O. E., Brusentsev, V. O. (2019). Multiagent control systems and decision-making supports. *Bulletin of National technical university "KhPI". Series: Machine science and CAD*, 1, 18 – 27. DOI:10.20998/2079-0775.2019.1.03.
16. Tkachov, V. M., Tokarev, V. V., Churyumov, G. I. (2019). Algorithm design of multiagent managements of group mobile "s-bot". *Kyiv: Data Recording, Storage & Processing*, 21 (1), 46–56.
17. Ungureanu, V. (2018). *Pareto-Nash-Stackelberg Game and Control Theory: Intelligent Paradigms and Applications*. Springer.
18. Neogy, S. K., Bapat, R. B., Dubey Dipti. (2018). *Mathematical Programming and Game Theory*. Springer.
19. Nazin, A. V., Poznyak, A. S. (1986). *Adaptive Choice of Variants: Recurrence Algorithms*. Moscow: Science.
20. Kushner, H., Yin, G. G. (2013). *Stochastic Approximation and Recursive Algorithms and Applications*. Springer Science & Business Media.

PATTERNS OF SELF-ORGANIZING STRATEGIES IN THE GAME OF MOBILE AGENTS

Petro Kravets¹, Rostyslav Yurynets², Yaroslav Kis³

Lviv Polytechnic National University, Information Systems and Networks Department

¹ Petro.O. Kravets@lpnu.ua, ORCID 0000-0001-8569-423X

² Rostyslav.V. Yurynets@lpnu.ua, ORCID 0000-0003-3231-8059

³ Yaroslav.P. Kis@lpnu.ua, ORCID 0000-0003-3421-2725

In this research the actual problem of self-organizing of strategies of stochastic game of multiagent systems is considered. Self-organizing display are formations of the co-ordinated behavioural patterns of group of the mobile agents endowed with the ability to move within a limited discrete space.

The agent is an independent object which can interact with environment, other agents and the person for a choice of variants of decisions. The multiagent system consists of group of agents which perform the general work, co-operating among themselves within local subsets of agents. The behavioural pattern of multiagent systems is the visualised form of purposeful moving of agents which arises from their initial chaotic movement during training of stochastic game.

A repetitive stochastic game is to implement a controlled random process of selecting decision options. To do this, game agents randomly, simultaneously and independently choose one of their own pure strategies at discrete times. Pure player strategies determine the direction of movement in two-dimensional space: forward, back, to the right, to the left. When all strategies have been selected, the current player losses are calculated. To form an orderly move, each agent must repeat the actions of neighboring agents. The current losers are then determined by the indicator function of the similarity of the strategies of the neighboring players. The calculated current losses are used to adaptively recalculate mixed player strategies. The probability of a pure strategy selection increases, if its realisation has led to reduction of current loss. In the course of a recurring game, agents will form vectors of mixed strategies that will minimize the functions of the players' average losses.

To solve the game problem of constructing patterns of self-organization of a multiagent system it is used the markovian adaptive recurrent method constructed on the basis of stochastic approximation

of the modified complementary slackness condition which is correct in balance to Nash points. For a normalization of elements of vectors of the mixed strategies operation of their projecting on a unit expanded epsilon-simplex is applied. Convergence of a game method is provided with observance of fundamental conditions and restrictions of stochastic optimisation.

Computer simulation confirmed the possibility of using a stochastic game model to build patterns of self-organization of a multi-agent system. The form of the received patterns depends on a way of local orientation of mobile agents. During computer experiment vortical and linear patterns moving of agents are received. The reliability of the experimental studies is confirmed by the similarity of the results obtained for different sequences of random variables.

The results of this work are expedient for applying to studying of patterns of collective behaviour of agents for deeper understanding of the processes of self-organizing of natural systems and for the construction of distributed decision-making systems.

Key words: self-organization, behaviour pattern, multi-agent system, stochastic game, markovian recurrent method.