Issue 9, 2021

https://doi.org/10.23939/sisn2021.09.131

УДК 004.[852+94]

САМООРГАНІЗАЦІЯ СТРАТЕГІЙ У ГРІ ПЕРЕМІЩЕННЯ АГЕНТІВ

Петро Кравець

Національний університет “Львівська політехніка”,
Petro.O.Kravets@lpnu.ua, ORCID: 0000-0001-8569-423X

Розроблено стохастичну ігрову модель самоорганізації стратегій стохастичної гри мобільних агентів у вигляді циклічних поведінкових патернів, які складаються із узгоджених стратегій переміщення агентів у обмеженому дискретному просторі. Поведінковий патерн багатоагентної системи є візуалізованою формою впорядкованого переміщення агентів, яка виникає із їх початкового хаотичного руху в ході навчання стохастичної гри.

Мобільність агентів багатокрокової стохастичної гри забезпечено тим, що у дискретні моменти часу вони випадково, одночасно і незалежно вибирають власну чисту стратегію переміщення в одному із можливих напрямків.

Поточні платежі гравців визначено як функції програшів, залежні від стратегій сусідніх гравців. Ці функції сформовано зі штрафу за нерівномірність розміщення агентів у обмеженому дискретному просторі та штрафу за зіткнення під час переміщення агентів. Випадковий вибір чистих стратегій гравців спрямовано на мінімізацію їхніх функцій середніх програшів. Генерування послідовностей чистих стратегій виконано за дискретним розподілом, побудованим на основі векторів змішаних стратегій. Елементи векторів змішаних стратегій є умовними імовірностями вибору відповідних чистих стратегій переміщення. Змішані стратегії змінюються у часі, адаптивно враховуючи значення поточних програшів. Цим забезпечено зростання імовірностей вибору тих чистих стратегій, які приводять до зменшення функцій середніх програшів.

Динаміку векторів змішаних стратегій визначено за марковським рекурентним методом, для побудови якого виконано стохастичну апроксимацію модифікованої умови доповняльної нежорсткості, яка справедлива у точках рівноваги за Нешем, та застосовано оператор проєктування на розширюваний одиничний епсилон-симплекс. Збіжність рекурентного ігрового методу забезпечено дотриманням фундаментальних умов та обмежень стохастичної апроксимації.

Стохастична гра розпочинається із ненавчених змішаних стратегій, які задають хаотичну картину переміщення агентів. У ході навчання стохастичної гри вектори змішаних стратегій цілеспрямовано змінюються так, щоб забезпечити впорядковане безконфлікне переміщення агентів.

У результаті комп’ютерного моделювання стохастичної гри отримано циклічні патерни самоорганізації мобільних агентів на поверхні дискретного тора та у межах прямокутної області на площині. Достовірність експериментальних досліджень підтверджено подібністю отриманих патернів переміщення агентів для різних послідовностей випадкових величин.

Результати дослідження запропоновано використати на практиці для побудови розподілених систем із елементами самоорганізації, розв'язування різноманітних потокових і транспортних задач та колективного прийняття рішень в умовах невизначеності.

Ключові слова: самоорганізація; циклічний поведінковий патерн; багатоагентна система; стохастична гра; марковський рекурентний метод.

Вступ

У наш час, в умовах всебічного впровадження інформаційних та кібернетичних систем у важливі для людини сфери діяльності, зростає актуальність розроблення та застосування перевірених часом біонічних методів і засобів розподіленого управління та колективного прийняття рішень [1, 2]. Колективна поведінка тварин – зграй птахів, косяків риб, роїв комах використовується для побудови моделей штучного життя, прийняття рішень та розвитку соціального пізнання. На моделях колективної поведінки відпрацьовують складні динамічні процеси, що виникають в організаційних системах різного призначення – соціально-політичних, військових, інтернет-спільнотах та інших за наявності багатоваріантності можливих дій, конфліктних ситуацій та інформаційної невизначеності.

Колективна поведінка полягає у прийнятті узгоджених рішень усіма членами соціальної групи. Таке узгодження може бути результатом самоорганізації динамічної системи. Самоорганізована поведінка – це процес упорядкування дій активних елементів, зумовлений внутрішніми причинами, без організаторського зовнішнього впливу. Способи самоорганізації будуються на основі локальної взаємодії складових елементів (на мікрорівні) системи і проявляються у вигляді глобально скоординованих дій (на макрорівні) [3].

Проявом колективної самоорганізації є утворення поведінкових патернів [4, 5]. Патерни масової поведінки описують динаміку зграї, натовпу, зародження та розвиток панічних настроїв, революцій, стихійних масових рухів, модних тенденцій, формування суспільної думки, результатів пропаганди. Утворення патернів масової поведінки вивчають у політиці, соціології, психології, маркетингу.

Візуалізація поведінкових патернів може мати лінійні, нелінійні, вихрові, циклічні або інші форми, що ілюструють координацію дій агентів у просторі й часі [4]. Увага дослідників сконцентрована переважно на вивченні механізмів побудови лінійних та нелінійних патернів. Механізми зародження і розвитку наявних у природі динамічних циклічних форм патернів вивчено недостатньо повно.

Циклічні патерни переміщення можуть бути застосовані для моделювання динаміки складних систем, моделювання ризиків та катастроф, процесів самоорганізації, що виникають у потоках рідин або у результаті хімічних реакцій, у динамічних задачах зі збереженням балансу маси або потенціалу, для ефективної організації транспортних та матеріальних потоків, вивчення перенесення речовин у живих клітинах, компактного розміщення компонентів, плоского або об'ємного пакування органічних молекул, у нанокібернетиці.

Застосування циклічних патернів перспективне для дослідження потоків електричних та хімічних сигналів у мозку, вивчення механізмів запам'ятовування та пригадування інформації. Існує гіпотеза, що короткотривала пам'ять мозку формується за рахунок кільцевого збудження активності нейронів (подібно до кільцевого ланцюга Лоренте де Но) [6, 7]. Крім того, вивчення циклічних патернів переміщення має самостійне значення для моделювання динамічної самоорганізації розподілених систем.

Для пояснення колективної поведінки у живій природі запропоновано багато еволюційних теорій, які враховують наявність спільного середовища для навчання, можливість спільного пошуку харчових ресурсів, безпеку перебування у групі, підвищене виживання, економію енергії, ефективніше розмноження тощо [8]. Для перевірки теорій проводять експерименти над соціальними тваринами у природних або штучно створених лабораторних умовах. Організація таких експериментів часто не відповідає принципам біоетики, є незручною і затратною. Крім того, довгий період еволюційного часу, в ході якого сформувалися поведінкові патерни, робить майже неможливим їх природне відтворення упродовж короткотривалого експерименту (на практиці).

Комп'ютерне моделювання надає дослідникам нові, легші способи перевірки гіпотез колективної поведінки. До них належать можливість зміни початкових умов та параметрів моделі, перевірка результатів на великих вибірках даних, повторне відтворення кожного експерименту, якісніша ідентифікація динаміки складних поведінкових явищ.

Продуктивним способом комп'ютерного моделювання патернів самоорганізації є побудова багатоагентних моделей штучного життя [9–11]. Штучне життя – це міждисциплінарний напрям дослідження, метою якого є вивчення процесів функціонування, поведінки та еволюції живих систем за допомогою створених людиною моделей і пристроїв. Комп'ютерне штучне життя черпає ідеї від природи і моделює розвиток популяцій реактивних або когнітивних агентів, реалізованих у вигляді комп'ютерних програм чи роботів.

Агентно-орієнтоване моделювання [12–15] дає змогу вивчити утворення глобальних патернів групової поведінки агентів на основі їх локальної взаємодії подібно до живих природних систем. У груповій поведінці живих істот можна виявити ігровий аспект, який проявляється у вигляді антагонізму, конкуренції або співпраці для досягнення цілей виживання і розвитку. Кожна особина має власну локальну ціль поведінки у поточній ситуації, яка підпорядкована глобальній меті виживання усієї популяції. Наприклад, соціальна істота може отримати необхідні для виживання ресурси за рахунок скоординованої групової роботи усіх членів популяції. Діючи спільно, живі істоти не погіршують своїх шансів на виживання, а популяція загалом покращує такі шанси. Як правило, поведінка популяції реалізується в умовах невизначеності, здебільшого спричиненої впливом зовнішніх чинників. Враховуючи це, як інструмент дослідження моделей штучного життя виберемо стохастичну гру багатоагентної системи [16].

У фаховій науковій літературі питання самоорганізації патернів ігрової поведінки багатоагентних систем вичерпно не висвітлено. Тому дослідження ігрової моделі формування динамічних патернів переміщення агентів є актуальною науково-практичною проблемою.

У цій роботі запропоновано новий спосіб самоорганізації багатоагентної системи на основі моделі стохастичної гри, який полягає в адаптивному формуванні поведінкових патернів безконфліктного циклічного переміщення агентів у обмеженому дискретному просторі.

Вивчення умов та механізмів зародження поведінкових патернів за допомогою моделі стохастичної гри уможливить глибше розуміння ходу природної еволюції та використання її результатів для вирішення проблем колективної співпраці в соціальних, економічних, екологічних, технічних та інших розподілених системах.

Мета роботи

Метою роботи є розв'язування ігрової задачі формування динамічних патернів самоорганізації стратегій, які забезпечують циклічні переміщення агентів у обмеженому дискретному просторі. Для досягнення мети здійснено постановку стохастичної ігрової задачі, вибрано метод розв'язування стохастичної гри, виконано комп'ютерне моделювання еволюції стратегій стохастичної гри, проаналізовано результати ком'ютерного моделювання та сформульовано відповідні висновки і рекомендації.

Постановка ігрової задачі

Розглянемо M -вимірний дискретний (клітинний) простір, у якому розміщена множина D агентів (кількість агентів $|D| \geq 2$). Агенти можуть незалежно переміщуватися у кожному координатному напрямку обмеженої області клітинного простору. Для цього вони мають $N_i \geq 2$ чистих стратегій $U^i = (u^i[1], u^i[2], ..., u^i[N_i])$, які визначають напрямки їхнії можливих переміщень – вперед, назад, наліво, направо тощо. Усі агенти мають рівні права щодо можливості переміщення. Кількість чистих стратегій $N_i = 2 \times M$ за можливості прямого та зворотного переміщення у горизонтальному і вертикальному напрямках.

Переміщення здійснюють синхронно усі агенти у дискретні моменти часу $n = 1, 2, \ldots$. Для цього кожен агент $i \in D$ здійснює незалежний вибір однієї із N_i власних чистих стратегій $u_n^i = u^i \in U^i$.

Вибір чистих стратегій $u_n^i(j), j = 1..N_i$ здійснюється з умовними імовірностями

$$p_n^i(j) = P\left\{u_n^i = u^i(j) \middle| u_t^i, \zeta_t^i \; (t = 1, 2, ..., n-1)\right\}, \; j = 1..N_i,$$

де $\{u_t^i \,|\, t = 1, 2, ..., n-1\}$ – передісторія стратегій, які вибрав гравець із номером i; $\{\zeta_t^i \,|\, t = 1, 2, ..., n-1\}$ – передісторія отриманих за це програшів.

Набір імовірностей $p_n^i = (p_n^i(1), p_n^i(2), ..., p_n^i(N_i)) \; \forall i \in D$ утворює вектори змішаних стратегій гравців, які набувають значення $p_n^i \in S^{N_i}$ на N_i-вимірних одиничних симплексах:

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

Для вибору чистих стратегій $\{u_n^i\}$ використовується випадковий механізм, побудований на динамічному дискретному розподілі, який задано векторами змішаних стратегій p_n^i:

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

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

Завершивши вибір чистих стратегій, усі гравці отримують випадкові поточні програші $\zeta_n^i = \zeta_n^i(u_n^{D_i})$, значення яких є функцією комбінованих стратегій сусідніх гравців $u^{D_i} \in U^{D_i} = \underset{j \in D_i}{\times} U^j$ з підмножин $D_i \subseteq D$, $D_i \neq \varnothing \;\; \forall i \in D$. Припускається, що поточні програші мають постійне математичне сподівання та обмежений другий момент, які не відомі агентам апріорі.

Процес переміщення агентів визначається рекурентною залежністю:

$$c_n^i = c_{n-1}^i + \sum_{j \in D_i} \chi(c_{n-1}^j > 0)\chi(u_n^j \to u_n^i) - \chi(c_{n-1}^i > 0),$$

де c_n^i – поточна кількість агентів, розміщених у місці (клітинці) розташування i-го агента; $\chi(\text{ôî î âà}) \in \{0,1\}$ – індикаторна функція події: якщо умова виконується, то $\chi() = 1$, інакше – $\chi() = 0$; $\to$ – операція спрямування переміщення j-го елемента до i-го елемента.

Переміщення агентів відбувається лише тоді, коли їх кількість у відповідній клітинці дискретного простору більша від нуля.

Переміщення агентів повинні бути збалансованими на кожному кроці гри:

$$\sum_{i \in D} c_n^i = C,$$

де $C = |D|$ – задана кількість агентів стохастичної гри.

Нерівномірність розміщення агентів у межах клітинної області оцінюється штрафом:

$$\xi_n^i[1] = C_i^{-1} \sum_{j \in D_i} |c_n^i - c_n^j|, \tag{2}$$

де $C_i = |D_i|$ – кількість агентів локальної підмножини D_i.

Інтуїтивно зрозуміло: якщо кількість агентів дорівнює кількості клітин дискретної області, то їх безпечне, безконфліктне переміщення повинно бути організоване у вигляді циклів. Для створення патернів у вигляді неперетинних циклів переміщення агентів повинна виконуватися умова безконфліктного руху, коли у напрямку розміщення кожного i-го агента (після того як він звільнить зайняту клітинку) може переміститися не більше ніж один сусідній агент. У цьому сенсі неоднорідність спрямування стратегій у межах локальної підмножини D_i враховується штрафом:

$$\xi_n^i[2] = C_i^{-1} \sum_{j \in D_i} | k_n^i - k_n^j |, \tag{3}$$

де $k_n^i = \sum_{j \in D_i} \chi(u_n^j \to u_n^i)$ – поточна кількість сусідніх агентів, переміщення яких спрямовані у клітинку, де перебуває i-й агент (у напрямку i-го агента).

Поточні програші гравців обчислюються на основі штрафів (2) та (3) після завершення вибору варіантів переміщення $u_n^i \;\; \forall i \in D$:

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

де $\lambda \in [0,1]$ – ваговий коефіцієнт, який визначає частку штрафних критеріїв у формуванні програшів агентів.

Завдяки випадковому вибору чистих стратегій поточні програші $\zeta_n^i \;\; \forall i \in D$ є випадковими величинами з апріорі невідомими стохастичними характеристиками. Для формулювання критеріїв поведінки гравців використовуються усереднені у часі $n = 1, 2, \ldots$ середні програші:

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

Стратегії гравців повинні бути спрямовані на мінімізацію власних функцій середніх програшів:

$$\overline{\lim_{n \to \infty}} Z_n^i \to \min_{u_n^i} \quad \forall i \in D. \tag{6}$$

Стохастична гра переміщення агентів у дискретному просторі полягає у тому, що кожен гравець $i \in D$ повинен навчитися вибирати варіанти переміщення $\{u_n^i\}$ у моменти часу $n = 1, 2, \ldots$ на основі спостереження локально обумовлених поточних програшів $\{\zeta_n^i\}$ так, щоб забезпечити виконання системи критеріїв (6).

Спосіб формування послідовності стратегій $\{u_n^i\} \forall i \in D$ ($n = 1, 2, \ldots$) стохастичної гри визначатиме виконання однієї із умов колективної оптимальності, наприклад, Неша, Парето або іншої [17].

Процес навчання стохастичної гри, який приводить до самоорганізації стратегій переміщення ігрових агентів, оцінюється такими характеристиками:

1) інтегральною функцією середніх програшів:

$$Z_n = C^{-1} \sum_{i \in D} Z_n^i \, ; \tag{7}$$

2) відносною кількістю скоординованих стратегій або коефіцієнтом координації стратегій гравців:

$$K_n = (nC)^{-1} \sum_{t=1}^{n} \sum_{i \in D} \chi\left(\zeta^i = 0\right). \tag{8}$$

На утворення патернів самоорганізації вказуватиме зменшення функції системних програшів $Z_n \ge 0$ та зростання коефіцієнта координації стратегій агентів $K_n \in [0,1]$.

Метод розв'язування стохастичної гри

Для розв'язування стохастичної гри застосуємо адаптивний марковський метод, який, опрацьовуючи поточні програші гравців, забезпечить оптимальний у середньому вибір чистих стратегій. Метод має бути побудований так, щоб збільшувати імовірності вибору тих чистих стратегій, які мінімізують функції середніх програшів гравців (5).

В основу марковського рекурентного методу розв'язування стохастичної гри покладемо умову доповняльної нежорсткості відповідної детермінованої гри [18]. Ця умова справедлива для колективних розв'язків за Нешем для вирівнювальних змішаних стратегій гравців. Виконавши стохастичну апроксимацію умови доповняльної нежорсткості [19, 20], отримаємо такий рекурентний метод:

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

де $\pi_{\varepsilon_{n+1}}^{N_i}$ – проєктор на одиничний ε-симплекс $S_{\varepsilon_{n+1}}^{N_i} \subseteq S^{N_i}$ [21]; $p_n^i \in S_{\varepsilon_n}^{N_i}$ – змішані стратегії i-го агента; $\gamma_n > 0$, $\varepsilon_n > 0$ – монотонно спадні послідовності додатних величин; $\zeta_n^i \in R^1$ – поточний пограш агента; $e(u_n^i)$ – одиничний вектор-індикатор вибору варіанта $u_n^i \in U^i$.

Проєктор $\pi_{\varepsilon_{n+1}}^{N_i}$ на розширювальний ε-симплекс забезпечує виконання умов нормування змішаних стратегій: $p_n^i[j] \ge \varepsilon_n$, $\sum_{j=1}^{N_i} p_n^i[j] = 1$. Крім нормування, він необхідний для забезпечення умов збирання повної статистичної інформації про вибрані чисті стратегії.

Параметр γ_n зменшує величину кроку методу в часі для досягнення оптимальних колективних розв'язків гри: $\left\| p_n^i - p^{i*} \right\| \to 0$, а параметр ε_n розширює ε-симплекс: $S_{\varepsilon_{n+1}}^{N_i} \to S^{N_i}$. Значення цих параметрів можна обчислити так:

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

де $\gamma > 0$; $\alpha \in (0,1]$; $\varepsilon > 0$; $\beta > 0$.

Збіжність стратегій стохастичної гри до колективно-оптимальних значень p^{i*} визначається співвідношеннями параметрів γ_n та ε_n (10), які повинні відповідати фундаментальним умовам стохастичної апроксимації [19–21].

Стохастична гра розпочинається із ненавчених змішаних стратегій агентів: $p_0^i = (1/N_i, ..., 1/N_i)$ $\forall i \in D$. З часом $n = 1, 2, ...$ змішані стратегії динамічно перебудовуються згідно з (9) для адаптивного вибору чистих стратегій.

Один крок повторювальної стохастичної гри полягає у тому, що у момент часу n кожен гравець $i \in D$ вибирає чисту стратегію u_n^i (1) і до моменту часу $n+1$ отримує поточний пограш ζ_n^i (4), який використовується для обчислення нової змішаної стратегії p_{n+1}^i (9).

Стохастична гра агентів реалізує адаптивний метод спроб і помилок, який імітує еволюцію природних систем і потребує значної кількості випробувань. Однак комп'ютерна реалізація стохастичної гри забезпечить стиснення часу моделювання еволюційних процесів, які приводять до самоорганізації багатоагентних систем.

Результати комп'ютерного експерименту

Розглянемо розв'язування ігрової задачі координації переміщення множини агентів на поверхні дискретного тора. Твірну площину тора задано в евклідовому просторі з правилами переміщення агентів:

$$T = \{(x,y) \mid x = x_{\min}..x_{\max}; y = y_{\min}..y_{\max}; x_{\min} - 1 \Rightarrow x_{\max}; x_{\max} + 1 \Rightarrow x_{\min}, y_{\min} - 1 \Rightarrow y_{\max}; y_{\max} + 1 \Rightarrow y_{\min}\},$$

де x, y – цілочислові координати клітин поверхні тора, $x_{\min} = y_{\min} = 1$, $x_{\max} = y_{\max} = 5$ – мінімальне та максимальне значення координат.

На поверхні тора можна переміщуватися без обмежень в усіх можливих напрямках.

Початково у клітинках тора розмістимо $C = |D| = 25$ агентів. За умови випадкового розміщення в одну клітинку тора може потрапити декілька агентів. Далі агенти розігрують повторювальну стохастичну гру, маючи $N_i = 4$ стратегії переміщення: U^i = (*вперед* = 0, *направо* = 1, *назад* = 2, *наліво* = 3).

Результати гри такі:

1) досягнення рівномірного розміщення мобільних агентів на дискретній поверхні (у кожен момент у кожній клітинці повинен перебувати тільки один агент);

2) самоорганізація гри у вигляді таких динамічних патернів переміщення, що траєкторії руху агентів не перетинаються.

Враховуючи, що кількість агентів дорівнює кількості клітин дискретного тора, динамічна самоорганізація агентів можлива за умови їх скоординованого циклічного переміщення. Залежно від початкових умов і параметрів ігрового методу на поверхні тора можуть виникнути декілька автономних циклів або потоків переміщення агентів.

Для розв'язування стохастичної гри використаємо стохастичний ігровий метод (9) з такими параметрами: $\lambda = 0{,}5$, $\gamma = 10$, $\varepsilon = 0{,}999 / N_i$, $\alpha = 0{,}1$, $\beta = 2$. Моделювання еволюційної координації дій агентів обмежується $n_{\max} = 10^5$ кроками стохастичної гри. Для заданих значень параметрів отримаємо розв'язки стохастичної гри у чистих стратегіях.

Функції середніх програшів гравців Z_n (7) та коефіцієнта координації стратегій K_n (8) характеризують хід стохастичної гри переміщення агентів. Графіки цих функцій зображено на рис. 1 у логарифмічному масштабі.

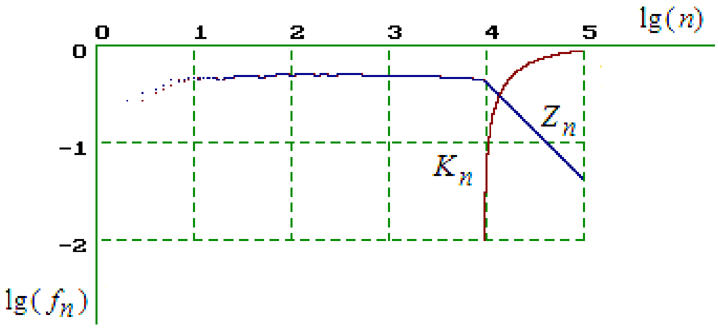
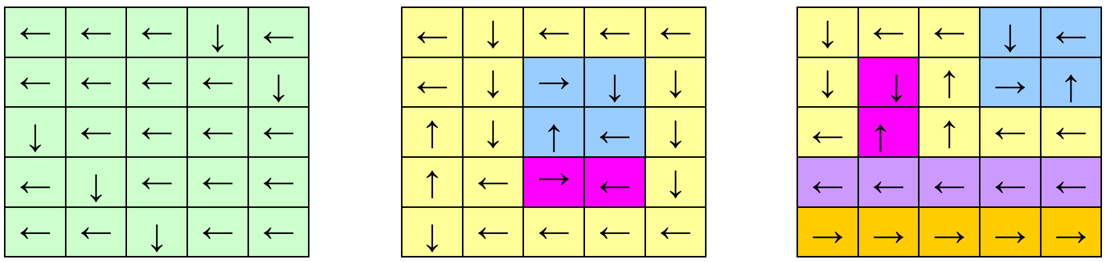


Рис. 1. Показники самоорганізації стохастичної гри

Зменшення функції середніх програшів Z_n свідчить про виконання цільових умов (6) збіжності стохастичної гри. Наближення значення коефіцієнта координації K_n до 1 (тобто до логарифмічного нуля) вказує на самоорганізацію стратегій гравців. Для заданих параметрів стрімке зростання коефіцієнта координації починається за близько $n = 10^4$ кроків навчання стохастичної гри.

Скоординоване, безконфліктне переміщення агентів визначається відповідним підбором параметрів ігрового методу. Сприятливим фактором збіжності стохастичної гри є збалансованість цих параметрів згідно з умовами стохастичної апроксимації [19–21].

На рис. 2 зображено патерни самоорганізації стратегій стохастичної гри переміщення агентів на поверхні тора. Ігрова задача переміщення має багатоваріантний розв'язок.



а *б* *в*

Рис. 2. Патерни самоорганізації стохастичної гри переміщення агентів на торі:
а – один цикл; б – три цикли; в – п'ять циклів

Стрілками позначено напрямки переміщення агентів для навченої стохастичної гри. Вибравши довільну початкову клітинку і рухаючись за напрямком стрілок, побачимо, що переміщення утворює замкнений шлях (цикл). Напрям стрілки за межі прямокутника вказує на те, що рух необхідно продовжити за ходом стрілок на протилежній стороні області. Максимальний за довжиною цикл охоплює усі клітинки дискретної поверхні тора, мінімальний цикл – тільки дві сусідні клітинки із протилежними напрямками переміщення агентів. Для мінімального циклу це не вважається зіткненням, оскільки інші агенти не вливаються у цей потік і одночасно у клітинці перебуває тільки один агент.

Патерн на рис. 2, *а* складається із одного циклу. Патерн на рис. 2, *б* містить три цикли, а на рис. 2, *в* – п'ять циклів переміщення агентів. Виникнення одного або декількох циклів можна регулювати зміною коефіцієнта $\lambda \in [0,1]$, який впливає на баланс штрафних критеріїв у загальній формулі (4) поточних програшів агентів.

З отриманих результатів видно, що проста взаємодія агентів у межах локальних підмножин у ході навчання стохастичної гри приводить до складної, скоординованої поведінки системи у вигляді патернів циклічного переміщення агентів. На початку гри відбуваються зіткнення та скупчення ненавчених агентів. У ході навчання система переходить від початкового хаотичного руху агентів до їх впорядкованого переміщення по замкнених шляхах. Інакше, на поверхні тора виникають узгоджені потоки агентів.

Аналогічні результати отримаємо і для прямокутної дискретної області на площині з правилами переміщення агентів:

$$R = \{(x, y) \mid x = x_{\min}..x_{\max}; y = y_{\min}..y_{\max}; x_{\min} - 1 \Rightarrow x_{\min}; x_{\max} + 1 \Rightarrow x_{\max}, y_{\min} - 1 \Rightarrow y_{\min}; y_{\max} + 1 \Rightarrow y_{\max}\}.$$

На відміну від тора, де допускається переміщення в усіх можливих напрямках, переміщення по площині обмежується краями прямокутної клітинної області. На краях прямокутника стратегій переміщення в агентів буде менше, ніж у його центральній частині. Ще одна, підтверджена експериментально, особливість організації стохастичної гри переміщення агентів у межах прямокутної області полягає у тому, що для можливості формування патернів циклічного переміщення кількість клітин і, відповідно, кількість агентів повинні бути парними, наприклад: $x_{\min} = y_{\min} = 1$, $x_{\max} = 5$, $y_{\max} = 4$, $C = 20$.

Патерни самоорганізації стохастичної гри переміщення агентів на площині зображено на рис. 3. Як і раніше, напрям стрілок визначає стратегії переміщення агентів навченої стохастичної гри.

→	→	→	→	↓
↑	↓	←	←	↓
↑	↓	→	↑	↓
↑	←	↑	←	←

а

→	↓	↓	←	←
↑	←	→	↓	↑
↓	←	←	←	↑
→	→	→	→	↑

б

→	→	→	→	↓
↑	↓	←	↓	←
↑	→	↑	↓	↓
↑	←	←	←	↑

в

Рис. 3. Патерни самоорганізації стохастичної гри переміщення агентів на площині: а – один цикл; б – два цикли; в – три цикли

Як видно на рис. 3, *а,* скоординовані переміщення агентів організовано у вигляді одного циклу, на рис. 3, *б* – двох циклів, а на рис. 3, *в* – трьох циклів.

Отже, самоорганізація розглянутої стохастичної гри полягає в утворенні патернів у вигляді циклів переміщення агентів, що виникають в процесі навчання рекурентного методу (9) на основі локальної взаємодії між агентами, що приводить до глобальної координації усієї розподіленої системи.

Зростання розмірів клітинної області й, відповідно, кількості гравців C приводить до збільшення кількості можливих станів системи за показниковою залежністю N^C, де N – кількість стратегій переміщення кожного агента. Наслідком цього є значне зростання кількості пошукових кроків, необхідних для формування патернів самоорганізації багатоагентної системи. Незважаючи на це, для задач із дослідженою розмірністю ігровий самонавчальний метод (9) зі степеневим порядком середньоквадратичної швидкості збіжності [15] показав прийнятні для комп'ютерної практики результати.

Висновки

1. Запропоновано стохастичний ігровий метод розв'язування складної багатоваріантної задачі організації безпечного колективного переміщення агентів у обмеженому дискретному просторі.

2. Наочно продемонстровано утворення патернів самоорганізації у вигляді циклів переміщення агентів, які зароджуються у результаті цілеспрямованого локально обумовленого вибору безконфліктних стратегій у ході адаптивного навчання стохастичної гри.

3. Патерни самоорганізації стохастичної гри можна застосувати для моделювання колективної поведінки живих організмів. Вивчення правил самоорганізації патернів переміщення агентів дає змогу краще зрозуміти процеси колективної поведінки у природі з метою їх практичного застосування.

4. Стохастичний ігровий метод самоорганізації багатоагентних систем можна використати для розв'язування розподілених потокових і транспортних задач та колективного прийняття рішень в умовах невизначеності.

5. Перспективним дослідженням у цьому напрямі є ігрове моделювання циклічних патернів самоорганізації процесів мозкової активності вищих живих організмів для застосування результатів у системах штучного інтелекту.

Список літератури

1. Gamazine, S., Deneubourg, J.-L., Frank, N. R., Sneyd, J., Theraula, G., Bonabeau, E. (2020). *Self-Organization in Biological Systems*. Princeton University Press.

2. Гийо, А., Мейе, Ж.-А. (2013). *Бионика. Когда наука имитирует природу*. Москва: Техносфера.

3. Zhang, W. J. (editor). (2013). *Self-organization: Theories and Methods*. USA: Nova Science Publishers.

4. Кравець, П. О., Юринець, Р. В., Кісь, Я. П. (2020). Патерни самоорганізації стратегій у грі мобільних агентів. *Вісник Нац. ун-ту "Львівська політехніка". Серія: "Інформаційні системи та мережі", Вип. 7*, 24–34. DOI: 10.23939/sisn2020.07.024.

5. Kelso, J. A., Scott. (1995). Dynamic patterns: The self-organization of brain and behavior. Cambridge, MA: The Mit Press.

6. Josselyn, S. A., Tonegawa, S. (2020). Memory engrams: Recalling imagining the future. Review. *Science. American Association for the Advancement of Science (AAAS), 367, 6473,* 1–14. DOI: 10.1126/science.aaw4325.

7. Самойлов, В., Бигдай, Е. (2020). Физиология человека для технических специальностей: центральная нервная и сенсорная системы: учеб. пособ для вузов. 2-е изд. Москва: Юрайт.

8. Галкін, О. Ю., Тітова Л. О. (2018). Основи еволюційної теорії: навч. посіб. Київ: КПІ ім. І. Сікорського.

9. Adamatzky, A., Komosinski, M. (Eds). (2005). Artificial Life Model in Software. Springer.

10. Adami, C. (1998). Introduction to Artificial Life. Springer.

11. Langton, C. G. (1995). Artificial Life. An Overview. Cambridge: The MIT Press.

12. Sadhu, A. K., Konar, A. (2020). Multi-Agent Coordination. Wiley-IEEE Press.

13. Sun, Z. (2018). Cooperative Coordination and Formation Control for Multi-agent Systems. Springer.

14. Yang, S., Xu, J.-X., Li, X., Shen, D. (2017). Iterative Learning Control for Multi-agent Systems Coordination. John Wiley & Sons.

15. Кравець, П. О. (2015). Ігрова модель самоорганізації мультиагентних систем. *Вісник Нац. ун-ту „Львівська політехніка". Серія: „Інформаційні системи та мережі", 829*, 161–176.

16. Chen, B.-S. (2020). Stochastic Game Strategies and their Applications. CRC Press.

17. Ungureanu, V. (2018). Pareto-Nash-Stackelberg Game and Control Theory: Intelligent Paradigms and Applications. Springer.

18. Neogy, S. K., Bapat, Ravindra, B., Dubey, Dipti. (Eds). (2018). Mathematical Programming and Game Theory. Springer.

19. Kusuoka, S. (2020). Stochastic Analysis. Springer.

20. Kushner, H J., Yin, G. G. (2013). Stochastic Approximation Algorithms and Applications. Springer.

21. Назин, А. В., Позняк, А. С. (1986). Адаптивный выбор вариантов: Рекуррентные алгоритмы. Москва: Наука.

References

1. Gamazine, S., Deneubourg, J.-L., Frank, N. R., Sneyd, J., Theraula, G., Bonabeau, E. (2020). *Self-Organization in Biological Systems*. Princeton University Press.

2. Guillot, A., Meyer, J.-A. (2013). *Bionics. When the Science Imitates the Nature (in russian)*. Moscow: Technosphere.

3. Zhang, W. J. (editor) (2013). *Self-organization: Theories and Methods*. USA: Nova Science Publishers.

4. Kravets, P.A., Jurinets R.V, Kis, Y.P. (2020). Patterns of self-organizing strategies in the game of mobile agents (in ukrainian). *Bulletin of "Lviv polytechnic national university". Series: "Information systems and networks", Issue 7*, 24–34. DOI: 10.23939/sisn2020.07.024.

5. Kelso, J. A., Scott. (1995). *Dynamic patterns: The self-organization of brain and behavior*. Cambridge, MA: The Mit Press.

6. Josselyn, S. A., Tonegawa, S. (2020). Memory engrams: Recalling imagining the future. Review. *Science. American Association for the Advancement of Science (AAAS), 367, 6473,* 1–14. DOI: 10.1126/science.aaw4325.

7. Samoilov, V., Bigdaj, E. (2020). *Human physiology for technical specialities: the central nervous and sensory systems. The manual for high schools. The second edition (in russian)*. Moscow: Urait.

8. Galkin, A. Yu., Titova, L. A. (2018). *Fundamentals of evolutionary theory. Tutorial (in ukrainian).* Kyiv: Igor Sikorsky Kyiv Polytechnic Institute.

9. Adamatzky, A., Komosinski, M. (Eds). (2005). *Artificial Life Model in Software*. Springer.

10. Adami, C. (1998). *Introduction to Artificial Life*. Springer.

11. Langton, C. G. (1995). *Artificial Life. An Overview*. Cambridge: The MIT Press.

12. Sadhu, A. K., Konar, A. (2020). *Multi-Agent Coordination*. Wiley-IEEE Press.

13. Sun, Z. (2018). *Cooperative Coordination and Formation Control for Multi-agent Systems*. Springer.

14. Yang, S., Xu, J.-X., Li, X., Shen, D. (2017). *Iterative Learning Control for Multi-agent Systems Coordination*. John Wiley & Sons.

15. Kravets, P. A. (2015). Game model of self-organizing of multiagent systems (in ukrainian). *Bulletin of "Lviv polytechnic national university". Series: "Information systems and networks", 829*, 161–176.

16. Chen, B.-S. (2020). *Stochastic Game Strategies and their Applications*. CRC Press.

17. Ungureanu, V. (2018). *Pareto-Nash-Stackelberg Game and Control Theory: Intelligent Paradigms and Applications.* Springer.

18. Neogy, S. K., Bapat, Ravindra, B., Dubey, Dipti. (Eds). (2018). *Mathematical Programming and Game Theory*. Springer.

19. Kusuoka, S. (2020). *Stochastic Analysis*. Springer.

20. Kushner, H. J., Yin, G. G. (2013). *Stochastic Approximation Algorithms and Applications*. Springer.

21. Nazin, A. V., Poznyak, A. S. (1986). *Adaptive Choice of Variants: Recurrence Algorithms (in russian).* Moscow: Science.

SELF-ORGANIZING STRATEGIES IN GAME OF AGENT MOVEMENT

Petro Kravets

Lviv Polytechnic National University, Information Systems and Networks Department

E-mail: Petro.O.Kravets@lpnu.ua, ORCID: 0000-0001-8569-423X

In this paper, a stochastic game model of self-organization of strategies of stochastic game of mobile agents in the form of cyclic behavioral patterns, which consist of coordinated strategies for moving agents in a limited discrete space, is developed. The behavioral pattern of a multi-agent system is a visualized form of orderly movement of agents that arises from their initial chaotic movement during the learning of a stochastic game.

The mobility of multi-step stochastic game agents is ensured by the fact that in discrete moments of time they randomly, simultaneously and independently choose their own pure strategy of moving in one of the possible directions.

Current player payments are defined as loss functions that depend on the strategies of neighboring players. These functions are formed from the penalty for irregular spacing of agents in a limited discrete space and the penalty for collisions when moving agents. Random selection of players' strategies aims to minimize their average loss functions. The generation of sequences of pure strategies is performed by a discrete distribution based on the vectors of mixed strategies. The elements of the vectors of mixed strategies are the conditional probabilities of choosing the appropriate pure displacement strategies. Mixed strategies change over time, adaptively taking into account the value of current losses. This provides an increase in the probability of choosing those strategies that lead to a decrease in the functions of average losses.

The dynamics of the vectors of mixed strategies is determined by the Markov recurrent method, for the construction of which a stochastic approximation of the modified condition of complementary non-rigidity, which is valid at Nash equilibrium points, is performed, and a projection operator for expandable unit epsilon simplex is applied. The convergence of the recurrent game method is ensured by compliance with the fundamental conditions and limitations of stochastic approximation.

The stochastic game begins with untrained mixed strategies that set a chaotic picture of moving agents. During the learning of the stochastic game, the vectors of mixed strategies are purposefully changed so as to ensure an orderly, conflict-free movement of agents.

As a result of computer simulation of stochastic game, cyclic patterns of self-organization of mobile agents on the surface of a discrete torus and within a rectangular region on a plane are obtained. The reliability of the experimental studies was confirmed by the similarity of the obtained patterns of moving agents for different sequences of random variables.

The results of the study are proposed to be used in practice for the construction of distributed systems with elements of self-organization, solving various flow and transport problems and collective decision-making in conditions of uncertainty.

Key words: self-organization; cyclic behavioral pattern; multi-agent system; stochastic game; Markov recurrent method.

.