УДК 004.852; 004.94

П.О. Кравець
Національний університет “Львівська політехніка”,
кафедра інформаційних систем та мереж

ДИНАМІЧНА КООРДИНАЦІЯ СТРАТЕГІЙ МУЛЬТИАГЕНТНИХ СИСТЕМ

Розв’язано задачу динамічної координації стратегій мультиагентних систем в умовах невизначеності на основі моделі стохастичної гри. Динамічна координація полягає у навчанні системи генерувати просторово-розподілені періодичні сигнали. Побудовано модель стохастичної гри, визначено критерії динамічної координації стратегій гравців, розроблено рекурентний метод, алгоритмічне та програмне забезпечення розв’язування стохастичної гри.

Ключові слова: мультиагентна система, стохастична гра, динамічна координація стратегій.

The problem of dynamic coordination of strategies of multiagent systems in the conditions of uncertainty on the basis of stochastic game model is solved. Dynamic coordination consists in system learning to generate the spatially-distributed periodic signals. The model of stochastic game is constructed, criteria of dynamic coordination of strategies are defined, a recurrent method, algorithmic and software support for the stochastic game solving are developed.

Key words: multiagent system, stochastic game, dynamic coordination of strategies.

Вступ

Зростання продуктивності сучасних обчислювальних систем, розвиток інформаційних мережевих технологій, засобів паралельного та розподіленого програмування дали новий поштовх розпочатим у середині XX ст. дослідженням з колективної поведінки автоматів, технологій „м’яких” обчислень, розподіленого штучного інтелекту, децентралізованого керування та прийняття рішень. У результаті реконструкції та інтеграції знань цих областей сформувався новий агентно-орієнтований науковий напрямок.

Агент – це автономна система вироблення та прийняття рішень з елементами штучного інтелекту, яка для розв’язування поставленої задачі може взаємодіяти з іншими агентами та людиною, використовуючи ресурси інформаційної мережі. Агент функціонує згідно з цілями, закладеними його розробником або власником.

Мультиагентна система (МАС) – це система, утворена декількома інтелектуальними агентами, що взаємодіють, з властивостями: автономності – агенти повністю або частково незалежні, спеціалізації та обмеженості знань – кожен з агентів виконує вузькоспеціалізовані функції і не має цілісного уявлення про систему, децентралізації – агенти керують локальними ділянками розподіленої системи.

Сучасні дослідження МАС пов’язані з розв’язуванням складних проблем розподіленого штучного інтелекту: керування знаннями, забезпечення координації та кооперації, визначення оптимальної структурної організації МАС, планування сценарів колективної поведінки агентів, розроблення методів, мов та засобів комунікації агентів, методів та засобів автоматизованого проектування МАС, багатоагентного навчання, забезпечення мобільності агентів, забезпечення пластичності, надійності та стійкості до збоїв у роботі [1 – 4].

МАС використовуються для задач, які не можна розв’язати за допомогою одного агента або централізованим способом, наприклад: електронна комерція, військова справа, подолання наслідків надзвичайних ситуацій, моделювання соціальних структур та явищ, керування трудовими ресурсами, дистанційне навчання, складання розкладів, керування ринком цінних паперів та інвестицій, керування транспортом, логістика, геоінформаційні системи, стратегічні комп'ютерні ігри, тренажери, мережеві технології та розподілені обчислення, керування мережевими потоками, пошук інформації в Internet, мобільний зв'язок тощо.

Мультиагентна система виконує розподілене розв'язування поставленої задачі. Успішність мультиагентного розв'язування спільної задачі залежить від рівня координації дій агентів. Координація – це забезпечення узгодженої, впорядкованої роботи усіх ланок мультиагентної системи.

Координація може бути централізованою та децентралізованою. За централізованої координації існує головний агент (або коаліція таких агентів), який надсилає рішення усім іншим агентам. Агенти можуть враховувати рекомендоване рішення, виробляючи стратегії власної поведінки. Децентралізована координація базується на взаємодії між агентами. Така взаємодія може бути неявною, коли агенти поводяться незалежно і впливають один на одного через зміну станів спільного середовища, та явною, комунікативною, коли агенти додатково прямо обмінюються інформацією. Взаємодіючи із середовищем, агенти збирають та опрацьовують інформацію про розвідані стани, будують модель середовища у просторі стан-дія, наприклад, у вигляді когнітивної карти правил прийняття рішень та обмінюються знаннями 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 спеціальну мову та встановлені протоколи взаємодії.

Якщо зростають розміри МАС, що діють у межах розподілених комп'ютерно-інформаційних систем, методи централізованої координації поступаються за ефективністю методам децентралізованої координації у зв'язку з локально обумовленими властивостями взаємодії агентів, спеціалізацією та обмеженістю їхніх знань про кероване середовище.

Враховуючи притаманні колективу агентів фактори децентралізованого вироблення рішень, конкуренції, взаємодії, кооперації, навчання, самоорганізації, для дослідження МАС в умовах невизначеності використаємо модель стохастичної гри [5, 6]. Ігрова координація стратегій агентів у процесі самоорганізації МАС є актуальною науково-практичною проблемою, поки що недостатньо вивченою.

З теоретико-ігрового погляду координація – це узгоджений вибір стратегій, що задовольнять умови, накладені на значення функцій виграшів або програшів, отримані у ході їх оптимізації колективом агентів.

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

Ігрова координація може бути просторовою та динамічною у часі. Просторова координація – це досягнення стійкої карти стратегій гравців у ході розв'язування стохастичної гри. Карта стратегій – це просторово розподілені значення стратегій гравців, які задовольняють критерії функціонування агентів МАС. Динамічна координація – це досягнення стійкої ритмічної зміни карти стратегій гравців у ході розв'язування стохастичної гри.

Треба розрізняти динамічну координацію стратегій гравців та динамічну повторювальність ходів стохастичної гри. Динамічна координація розглядається як фінальний розв'язок стохастичної гри, її динамічна рівновага. Динамічна повторювальність є способом розв'язування стохастичної гри в умовах невизначеності.

Практичне застосування динамічної координації полягає у здатності МАС генерувати складні сигнали, які можуть бути використані, наприклад, у робототехнічних системах для організації просторового орієнтування та переміщення агентів.

Координація призводить до виникнення умов самоорганізації системи. Самоорганізація – це процес перетворення локально скоординованих дій агентів на глобально скоординовані дії у межах

МАС за рахунок внутрішніх факторів, без відповідного зовнішнього впливу. Це – здатність колективу агентів з локально обумовленими зв'язками і цілями досягати стійких скоординованих стратегій поведінки в умовах невизначеності за рахунок самонавчання. Самоорганізація та складні форми поведінки МАС можуть виникати при колективній реалізації найпростіших дій агентів. Дослідженням процесів самоорганізації систем різної природи займається синергетика – міждисциплінарна наука, що вивчає нелінійні відкриті дисипативні (такі, що розсіюють енергію) системи [7].

Метою роботи є розроблення ігрової моделі, методу, алгоритму та програмних засобів для забезпечення умов динамічної координації та самоорганізації МАС.

Модель стохастичної гри

Динамічну координацію стратегій виконаємо на основі моделі стохастичної гри агентів $\Gamma = (D,\ \{U^i\}_{i\in D},\ \{Z(x^i)\}_{i\in D})$, де $D \neq \varnothing$ – множина гравців, $U^i = (u^i(1), u^i(2), ..., u^i(N_i))$ – вектор чистих стратегій i-го гравця, $N_i \geq 2$ – кількість чистих стратегій, $Z(x^i)$ – розподіл випадкового програшу x^i $\forall i \in D$.

Агенти взаємодіють зі стохастичним середовищем $E : \{U^i\}_{i\in D} \to \{x^i\}_{i\in D}$, заданим входами $\{U^i\} \forall i \in D$ та випадковими значеннями виходів $\{x^i\} \forall i \in D$. Входи середовища визначаються чистими стратегіями гравців, а виходи задають поточні програші гравців. Модель середовища E у вигляді випадкових розподілів $Z(x^i)$ $\forall i \in D$ не відома агентам апріорі.

Нехай динамічна координація визначається повторювальністю реалізацій чистих стратегій гравців на проміжку моментів часу $t = N_i$. Структура МАС визначає локально обумовлений механізм формування випадкових програшів $x^i = x^i(t, u^{D_i}, w)$ агентів, що є функціями спільних стратегій $u^{D_i} \in U^{D_i} = \underset{j\in D_i}{\times} U^j$ з локальних підмножин $D_i \subseteq D$, $D_i \neq \varnothing$ $\forall i \in D$ на відрізку часу t та випадкового шансу $w \in [0,1]$.

Повторювальна стохастична гра розгортається у дискретні моменти часу $n = 1, 2, ...$. Поточні програші $x_n^i = x_n^i(t, u^{D_i}, w)$ знаходять для значень $n \geq t$. Вважається, що випадкові програші $\{x_n^i\}$ є незалежними $\forall u_n = u \in U$, $\forall i \in D$, $n = 1, 2, ...$, мають постійне математичне сподівання $M\{x_n^i\} = v(t, u^{D_i}) = const$ та обмежений другий момент $\sup_n M\{[x_n^i]^2\} = s^2(t, u^{D_i}) < \infty$. Стохастичні характеристики випадкових програшів апріорі не відомі агентам.

На момент часу n середні програші агентів набувають значення:

$$\Xi_n^i(\{t, u_n^{D_i}\}) = \frac{1}{n}\sum_{t=1}^{n} x_t^i \quad \forall i \in D. \tag{1}$$

Метою агентів є мінімізація власних функцій середніх програшів

$$\overline{\lim_{n\to\infty}} \Xi_n^i \to \min_{p_n^i} \quad \forall i \in D. \tag{2}$$

Отже, взаємодіючи із середовищем, на основі спостереження поточних програшів $\{x_n^i\}$ гравці повинні навчитися вибирати чисті стратегії $\{u_n^i\}$ так, щоб з ходом часу $n = 1, 2, ...$ забезпечити виконання системи критеріїв (2).

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

Для розв'язування задачі (2) необхідно визначити правило формування послідовностей $\{u_n^i\}$ варіантів рішень у часі.

Метод розв'язування стохастичної гри

Враховуючи, що $t = N_i$, формування послідовності варіантів рішень $\{u_n^i\}$ з потрібними властивостями виконаємо на основі матриць імовірностей

$$p_n^i = \begin{bmatrix} p_n^i(1,1), & ..., p_n^i(1,N_i) \\ ... & ... \quad ... \\ p_n^i(N_i,1), & ..., p_n^i(N_i,N_i) \end{bmatrix}, \ \forall i \in D.$$

Елементи $p_n^i(j,k)$ векторів змішаних стратегій є умовними імовірностями вибору чистих стратегій залежно від поточного варіанта дії u_n^i та отриманого програшу x_n^i. Рядки матриці p_n^i є змішаними стратегіями i-го гравця, який перебуває у стані $u_n^i \in U^i$. Змішані стратегії набувають значення на одиничному симплексі

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

Гра розпочинається з ненавчених стратегій з однаковими імовірностями їх вибору $p_0^i(j,k) = 1/N_i$, де $j,k = 1..N_i$. Для забезпечення адаптивних властивостей імовірність вибору стратегій з меншими програшами повинна збільшуватися у часі $n = 1,2,...$. У результаті повинен сформуватися розподіл імовірностей, який мінімізує функції середніх програшів (1) усіх гравців. Для забезпечення цього застосуємо рекурентний метод модифікації векторів змішаних стратегій:

$$p_{n+1}^i(u_n) = p_{e_{n+1}} \left\{ p_n^i(u_n) - g_n x_n^i \left[e(u_{n+1}^i) - p_n^i(u_n) \right] \right\}, \tag{3}$$

де $p_n^i(u_n)$ – змішана стратегія i-го гравця у стані $u_n \in U^i$; $p_{e_{n+1}}$ – проектор на одиничний e-симплекс [8], який є підмножиною одиничного симплексу; $g_n > 0$ – монотонно спадна послідовність невід'ємних величин, яка регулює величину кроку методу; $e_n > 0$ – монотонно спадна послідовність невід'ємних величин, яка регулює швидкість розширення e-симплексу; $e(u_{n+1}^i)$ – визначений у момент часу n одиничний вектор-індикатор переходу в стан $u_{n+1}^i \in U^i$.

Значення чистої стратегії визначається випадково на основі поточного розподілу імовірностей $p_n^i(u_n^i)$:

$$u_n^i = \left\{ u^i(l) \,|\, l = \arg\min_l \sum_{k=1}^{l} p_n^i(j,k) > w \ \ (j,l = 1..N_i) \right\}, \tag{4}$$

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

Перебуваючи в момент часу t_n у стані u_n^i, i-й гравець на основі змішаної стратегії $p_n^i(u_n)$ вибирає дію u_n^i згідно з (4) і до моменту часу t_{n+1} отримує поточний програш x_n^i, після чого обчислює змішану стратегію $p_{n+1}^i(u_n)$ згідно з (3) та переходить у новий стан u_{n+1}^i.

Метод (3) одержано на основі стохастичної апроксимації умови доповняльної нежорсткості, справедливої для змішаних стратегій у точці рівноваги за Нешем [9]. Для побудови методу (3) умова доповняльної нежорсткості зважена елементами змішаних стратегій, що дає змогу знаходити розв'язки гри у чистих стратегіях. Завдяки динамічній перебудові змішаних стратегій на основі опрацювання поточних програшів метод (3) забезпечує адаптивний вибір стратегій агентів у часі.

Параметри g_n та e_n можна обчислити так:

$$g_n = g n^{-a}, \ e_n = e n^{-b}, \tag{5}$$

де $g > 0$; $a > 0$, $e > 0$; $b > 0$. У загальному випадку параметри g_n, e_n можуть змінюватися у залежному від поточного стану u_n^i темпі часу $n(u_n^i)$.

Збіжність $\lim_{n\to\infty} \left\| p_n^i(u_n^i) - p^{i*}(u_n^i) \right\| \to 0$, $\forall i \in D$ методу (3) до оптимальних стратегій $p^{i*}(u_n^i)$ $\forall u_n^i \in U^i$, які мінімізують середні програші (1), забезпечується умовами стохастичної апроксимації [10].

Контрольний приклад

Виконаємо динамічну координацію стратегій для стохастичної гри з регулярною структурою зв'язків між агентами. Структура гри задається кількістю агентів $L = m \times m$, $m \geq 2$, підмножинами сусідніх агентів D_i, кількістю чистих стратегій $N_i = N = 2$, $i = 1..L$. Динамічна координація складається з просторової та часової коодинації стратегій. Просторова коодинація полягає у дотриманні співвідношення стратегій гравців у локально визначених областях D_i $\forall i \in D$ так, як це зображено на рис. 1. Часова координація визначається дотриманням співвідношення стратегій гравців на проміжку часу $t = 2$.

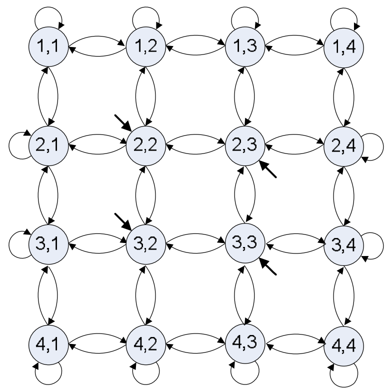

Рис. 1. Структура МАС

Нехай чисті стратегії гравців набувають бінарних значень $u_n^i \in \{0,1\}$, що спрощує візуальний аналіз динамічно скоординованих стратегій. Просторову координацію визначимо у вигляді однакових значень стратегій гравців (гравці намагаються повторювати дії один одного), а часову координацію – у вигляді ритмічної зміни інверсних значень стратегій у два послідовні моменти часу. У результаті динамічна координація полягатиме у повторенні матриць стратегій $[0]_{m\times m} - [1]_{m\times m} - [0]_{m\times m} - [1]_{m\times m} - \ldots$.

Для забезпечення динамічно скоординованих стратегій поточні програші гравців визначимо за допомогою різницевих функцій:

$$x_n^i = l \sum_{s \in D_i} | u_n^i - u_n^s | / L_i + (1 - l) | \overline{u}_n^i - u_{n-1}^i | + m_n , \tag{6}$$

де $x_n^i \in R^1$; $l \in [0,1]$ – ваговий коефіцієнт; D_i – множина сусідніх гравців, що відповідає зображеним на рис. 1 зв'язкам; $L_i = | D_i |$ – кількість сусідніх гравців; u_n^i – значення чистої стратегії; $\overline{u}_n^i$ – інверсне значення чистої стратегії; $m_n \sim Normal(0, d)$ – адитивний білий гауссівський шум, нормально розподілена випадкова величина з нульовим математичним сподіванням та дисперсією $d > 0$.

Перша складова виразу (6) визначає штраф за порушення просторової координації, друга складова – штраф за порушення часової координації стратегій гравців, а третя складова – дію випадкових завад у вигляді білого шуму.

Поточні значення білого шуму можна обчислити за формулою:

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

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

Ефективність динамічної координації оцінюється:

1) функцією середніх втрат:

$$\Xi_n = \frac{1}{L} \sum_{i=1}^{L} \Xi_n^i , \tag{7}$$

де $L = |D|$ – кількість гравців;

2) середньою кількістю просторово скоординованих стратегій гравців:

$$K_n = \frac{1}{n}\sum_{t=1}^{n}\sum_{i=1}^{L} c\left(\sum_{s\in D_i} |u_t^i - u_t^s| = 0\right), \qquad (8)$$

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

3) середньою кількістю скоординованих у часі стратегій гравців, які визначають ритмічну зміну послідовностей стратегій:

$$M_n = \frac{1}{n}\sum_{t=1}^{n}\sum_{i=1}^{L} c\left(|\overline{u}_t^i - u_{t-1}^i| = 0\right); \qquad (9)$$

4) середнім відхиленням змішаних стратегій гравців від оптимальних значень $p^{i*}(u_n^i)$:

$$\Delta_n = \frac{1}{Ln}\sum_{t=1}^{n}\sum_{i=1}^{L}\sum_{u_t^i\in U^i} \left\| p_t^i(u_t^i) - p^{i*}(u_t^i) \right\|, \qquad (10)$$

де $\|\cdot\|$ – евклідова норма вектора.

Алгоритм навчання стохастичної гри

1. Задати початкові значення параметрів:

$n = 0$ – початковий момент часу;

$L = |D|$ – кількість гравців;

$D_i \ \forall i \in D$ – множини сусідніх гравців, які визначають структуру гри;

N_i, $i = 1..L$ – кількості чистих стратегій гравців;

$t = N_i$ – період динамічної координації стратегій гравців;

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

$p_0^i = [1/N_i]_{N_i \times N_i}$, $i = 1..L$ – матриці змішаних стратегій гравців;

$g > 0$ – параметр кроку навчання;

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

e – параметр e-симплексу;

$b > 0$ – порядок швидкості розширення e-симплексу;

$d > 0$ – дисперсія завад;

$m_n \sim Z(0, d)$ – закон розподілу завад;

$l \in [0,1]$ – ваговий коефіцієнт поточних втрат;

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 x_n^i, $i = 1..L$ згідно з (6).

4. Знайти значення параметрів g_n, e_n згідно з (5).

5. Обчислити елементи векторів змішаних стратегій p_n^i, $i = 1..L$ згідно з (3).

6. Знайти характеристики якості прийняття рішень Ξ_n (7), K_n (8), M_n (9) та Δ_n (10).

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

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

Результати комп'ютерного моделювання

Результати моделювання n_{max} = 10 тис. кроків стохастичної гри подано на рис. 2 – 11.

На рис. 2 у логарифмічному масштабі зображено графіки функцій, які характеризують ефективність динамічної координації стохастичної гри $L = 4\times 4$ агентів для значень параметрів $g = 1$, $e = 0.999/2$, $a = 0.01$, $b = 2$, $l = 0.5$, $d = 0.36$. Графік 1 зображає функцію середньої кількості просторово скоординованих стратегій K_n, графік 2 – середню кількість ритмічних змін стратегій гравців M_n, графік 3 – середні поточні програші гравців Ξ_n.

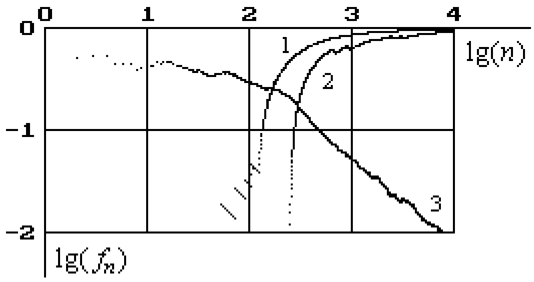

Рис. 2. Характеристики самоорганізації стохастичної гри

Зменшення функції середніх програшів у часі свідчить про збіжність ігрового методу. Функція кількості просторово скоординованих стратегій ілюструє координацію стратегій гравців, а функція кількості ритмічних змін стратегій – інверсну координацію стратегій гравців у послідовні моменти часу. Зростання цих функцій у часі свідчить про самоорганізацію стохастичної гри.

Динаміка вибору чистих стратегій стохастичної гри зображена на рис. 3, 4. Точками на графіках позначено стратегії гравців $u^i = 1$, де $i = 1..L$. Період графіків відповідає двом послідовним моментам часу. На початковому відрізку часу спостерігається хаотичний вибір стратегій ненавченими агентами (рис. 3). Після навчання ($n \sim 10^3$ кроків) відбувається скоординована у часі, ритмічна зміна бінарних стратегій (рис. 4).

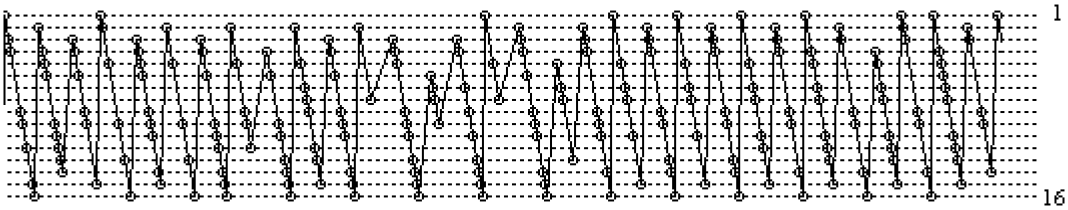

Рис. 3. Нескоординований вибір стратегій ненавченої стохастичної гри

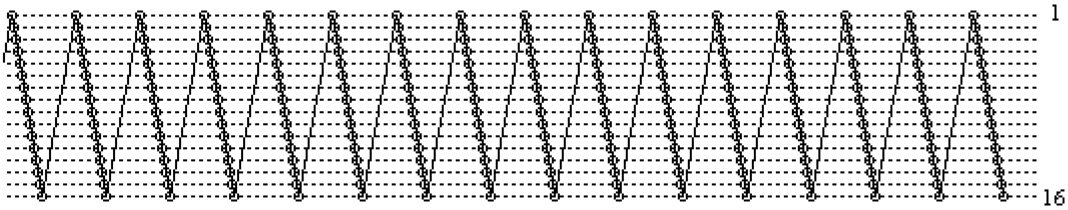

Рис. 4. Динамічно скоординований вибір стратегій навченої стохастичної гри

Метод (3) забезпечує розв'язування стохастичної гри у чистих стратегіях (на вершині одиничного симплексу). Зображена на рис. 4 ритмічна зміна стратегій згенерована навченими матрицями імовірностей переходів $p_n^{i*} = \begin{array}{c} \rightarrow \\ \end{array}\!\begin{bmatrix} 0,1 \\ 1,0 \end{bmatrix}$ $\forall i \in D$. Стрілкою позначено стани гравців $\forall i \in D$ у поточний момент часу n. Згідно зі сформованими матрицями імовірностей p_n^{i*} у момент часу $n+1$ гравці переходять в інверсні стани.

Ваговий коефіцієнт l визначає рівень впливу на поточний програш x_n^i (6) просторово-скоординованих стратегій та ритмічної зміни стратегій гравців у послідовні моменти часу. Динамічна координація визначається певним балансом між цими впливами. Якщо зростає l, забезпечується просторова координація, але порушується ритм їхньої зміни з бігом часу. Результати впливу параметра l на ефективність ігрового методу (3) зображено на рис. 5 у вигляді графіків функцій K_n (8) та M_n (9), отриманих при $n = n_{\max}$ та $d = 0$.

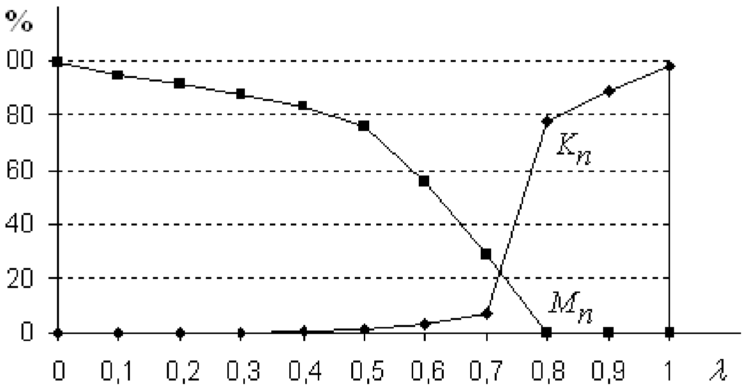

Рис. 5. Вплив вагового коефіцієнта l *на ефективність ігрового методу, якщо* $d = 0$

Як видно на рис. 5, за допомогою параметра l неможливо досягти належного балансу взаємної координації та ритмічної зміни стратегій гравців. Експериментально встановлено, що такого балансу можна добитись впливом білого шуму, який забезпечує випадкове, незміщене коливання рівня балансу між різними видами штрафів. Відповідні результати, отримані при $a = 0.01$, $d = 0.36$, подано на рис. 6.

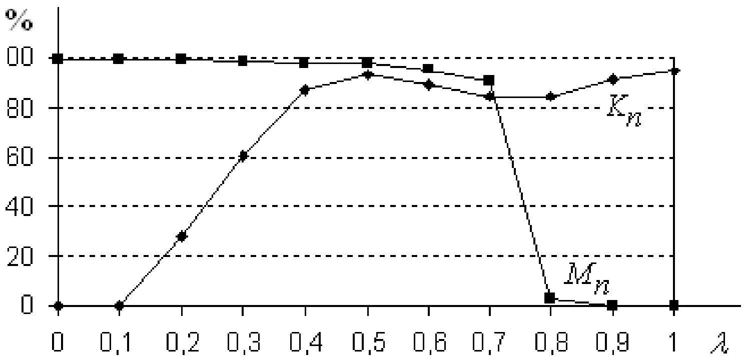

Рис. 6. Вплив вагового коефіцієнта l *на ефективність ігрового методу, якщо* $d = 0.36$

У разі дії білого шуму умови скоординованого ритмічного вибору стратегій виконуються для значень параметра $l \in [0.4; 0.7]$ на рівні, що перевищує 80 %.

Вплив дисперсії завад d на ефективність ігрового методу (3) зображено на рис. 7. Результати отримано при $a = 0.01$, $l = 0.5$. Якщо $\sqrt{d} \in [0.5; 0.7]$ характеристики ефективності методу перевищують рівень 80 %. За межами цього відрізка один з параметрів ефективності погіршується.

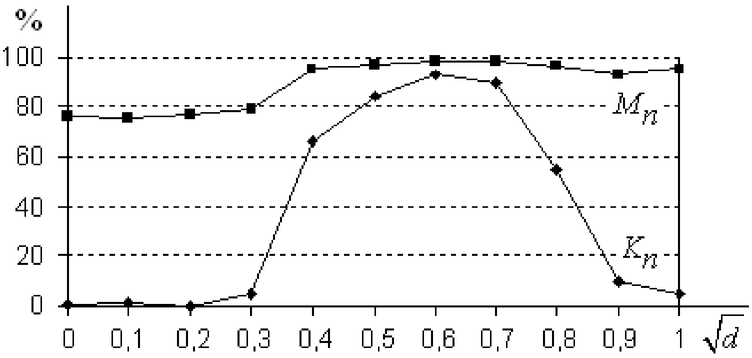

Рис. 7. Вплив дисперсії d *на ефективність ігрового методу*

Ефективність ігрового методу великою мірою залежить від параметрів a та b, які визначають швидкість збіжності ігрового методу. Значення цих параметрів повинні задовольняти базові умови стохастичної апроксимації. Залежність параметрів ефективності стохастичної гри від a наведена у вигляді графіків на рис. 8. Дані отримано для параметрів $d = 0.36$ та $l = 0.5$.

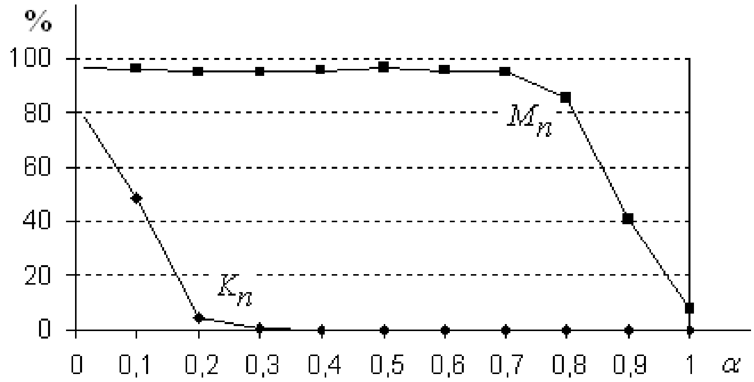

Рис.8. Вплив параметра a на ефективність ігрового методу

Зростання a призводить до погіршення просторової координації гравців. Однак зберігається високий ступінь (при $a \le 0.8$) часового ритму їхніх стратегій. Скоординовані та ритмічні дії гравців досягаються за малих значень $a \le 0.01$.

Залежність коефіцієнтів просторової та часової координації від кількості $L = m \times m$ гравців зображено на рис. 9 для параметрів $a = 0.01$, $d = 0.36$, $l = 0.5$. Значення коефіцієнтів отримано на момент часу $n = 10^4$ кроків гри. Прийнятна (більша за 80 %) координація стратегій відзначається для гри з максимальною кількістю $L = 5 \times 5$ агентів, кожен з яких має по $N = 2$ чисті стратегії.

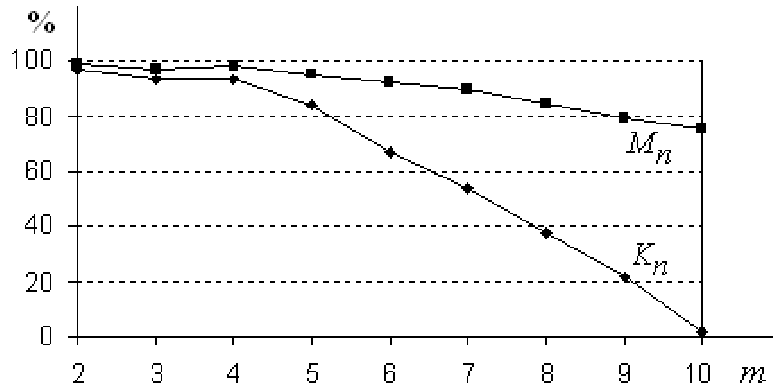

Рис.9. Залежність параметрів координації МАС від кількості гравців

Якщо зростає розмірність стохастичної гри, збільшується середня кількість кроків, потрібних для досягнення належного рівня координації стратегій агентів. На рис. 10 зображено залежність середньої кількості кроків гри $\overline{n}$, необхідних для досягнення коефіцієнта координації $K_n = 80\%$, від кількості гравців $L = m \times m$, кожен з яких має по $N = 2$ чисті стратегії.

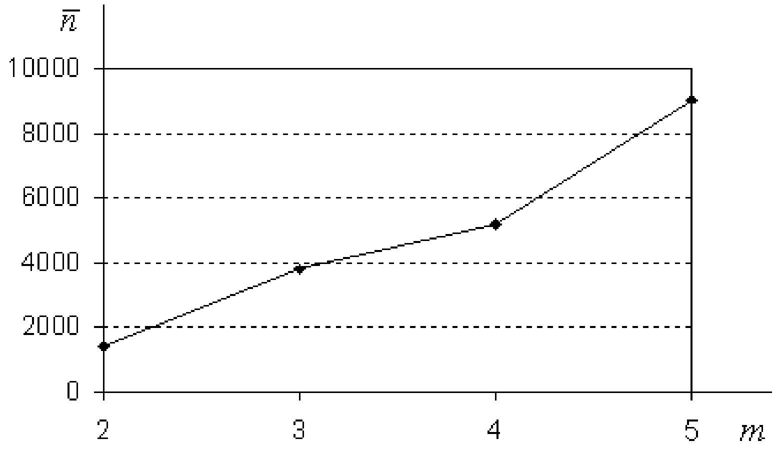

Рис.10. Залежність середньої кількості кроків навчання гри від кількості гравців

Запропонований метод (3) розв'язування стохастичної гри агентів належить до класу реактивних методів і має порівняно невисоку, степеневу швидкість збіжності. Це пов'язано з тим, що на початку гри агенти не мають ніякої інформації про середовище, з яким вони взаємодіють. Збір інформації здійснюється у процесі навчання за допомогою адаптивної перебудови векторів змішаних стратегій пропорційно до значень поточних програшів.

У загальному швидкість збіжності ігрового методу залежить від усіх параметрів моделі. На рис. 8 зображено графік середньої норми (10) відхилення змішаних стратегій від оптимальних значень для параметрів $g=1$, $e=0.999/2$, $a=0.01$, $b=2$, $l=0.5$, $d=0.36$. Порядок швидкості збіжності можна оцінити тангенсом кута j, утвореним між графіком лінійної апроксимації норми відхилення змішаних стратегій гравців від оптимальних значень та віссю часу.

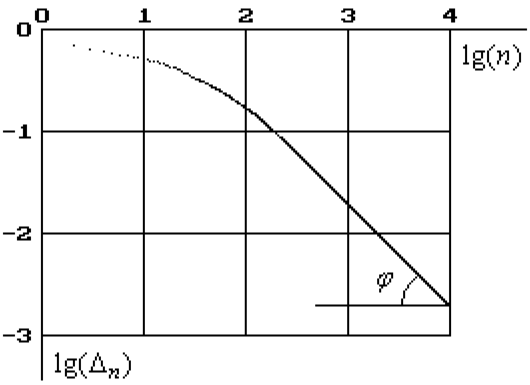

Рис.11. Динаміка середньої похибки розв'язування стохастичної гри

Підібравши параметри, можна досягти близького до 1 порядку швидкості збіжності розглянутої стохастичної гри.

Висновки

Розроблені модель, метод та алгоритм розв'язування стохастичної гри забезпечують динамічну координацію МАС, яка проявляється у локально обумовленому просторовому та часовому узгодженні стратегій гравців.

Динамічна координація забезпечується адаптивним пошуковим методом розв'язування стохастичної гри з урахуванням поточних штрафів за відповідні порушення просторової координації та часового ритму. Встановлено, що ефективність навчання гравців виконанню скоординованих ритмічних дій визначається балансом між цими штрафами, якого можна буде досягти впливом білого шуму.

Динамічна координація стратегій агентів досягається у ході розв'язування стохастичної гри у реальному масштабі часу на основі збору поточної інформації та її адаптивного опрацювання. Розглянутий метод належить до класу реактивних методів і моделює рефлексивну поведінку живих організмів. Метод дає змогу знаходити розв'язки стохастичної гри у чистих стратегіях.

Ефективність динамічної координації стратегій системи контролюється за допомогою характеристичних функцій середніх програшів, коефіцієнтів просторової і часової координації та евклідової норми відхилення динамічних змішаних стратегій від оптимальних значень. Зменшення функції середніх програшів, зростання коефіцієнтів координації та зменшення норми відхилення змішаних стратегій свідчить про збіжність ігрового методу.

Швидкість динамічної координації чистих стратегій агентів залежить від розмірності стохастичної гри, величини завад та параметрів ігрового методу. Якщо зростає кількість гравців та інтенсивність завад, швидкість та ефективність ігрової координації МАС зменшується. Особливістю запропонованого рекурентного методу розв'язування стохастичної гри є порівняно невисока, степенева швидкість збіжності, що пов'язано з апріорною невизначеністю системи. Цей недолік долається високою швидкодією сучасних засобів обчислювальної техніки та можливістю розпаралелювання задачі або використання мережевих засобів розподілених обчислень.

Збіжність методу забезпечується дотриманням базових умов стохастичної апроксимації. Параметри збіжності визначаються теоретично та уточнюються експериментально. Належно підібравши параметри ігрового методу, значення яких повинні відповідати умовам стохастичної апроксимації, можна досягти близького до 1 степеневого порядку швидкості збіжності.

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

Одержані результати можна використати для генерування просторово розподілених періодичних сигналів у мультиагентних та робототехнічних системах.

Розвиток запропонованого ігрового методу можливий у напрямі підвищення інтелектуального рівня агентів.

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. Wooldridge M. An Introduction to Multiagent Systems / M. Wooldridge. – John Wiley & Sons, 2002. – 366 pp. 5. Fudenberg D. The Theory of Learning in Games / D. Fudenberg, D.K. Levine. – Cambridge, MA: MIT Press, 1998. – 292 pp. 6. Доманский В.К. Стохастические игры / В.К. Доманский // Математические вопросы кибернетики. – 1988. – № 1. – С. 26 – 49. 7. Хакен Г. Информация и самоорганизация. Макроскопический подход к сложным системам: Пер. с англ. / Г. Хакен. – М.: КомКнига, 2005. – 248 с. 8. Назин А.В. Адаптивный выбор вариантов: Рекуррентные алгоритмы / А.В. Назин, А.С. Позняк. – М.: Наука, 1986. – 288 с. 9. Мулен Э. Теория игр с примерами из математической экономики / Э. Мулен. – М.: Мир, 1985. – 200 с. 10. Граничин О.Н. Введение в методы стохастической аппроксимации и оценивания: Учеб. пособие / О.Н. Граничин. – СПб.: Изд-во С.-Петербургского университета, 2003. – 131 с.

УДК 004.031.2

І.І. Кушнірецька, Л.В. Чирун,
Національний університет “Львівська політехніка”,
кафедра інформаційних систем та мереж

СИСТЕМА ПІДТРИМКИ ПРИЙНЯТТЯ РІШЕНЬ ДЛЯ ВИБОРЦЯ З ВИКОРИСТАННЯМ ФАСЕТНОГО МЕТОДУ КЛАСИФІКАЦІЇ

Розглянуто використання можливостей фасетної класифікації для знаходження результату системи підтримки прийняття рішень виборця. Запропонована методика дає можливість виборцеві вибрати політичну партію для голосування згідно з попередньо визначеними класицікаційними ознаками.

Ключові слова: фасета, фасетна класифікація, СППР.

This paper reviews the opportunities for finding faceted classification results decision support system for the voter. The technique enables the voter to choose a political party to vote, according to predefined signs for classification.

Key words: facet, faceted classification, features selection, DSS.

Вступ. Загальна постановка проблеми

Під час опрацювання великих масивів даних та складання різних зведень виникає проблема, що полягає у швидкості знаходження потрібної інформації, відповідно до заданих критеріїв і