УДК 004.852; 004.94
DOI: 10.20535/SRIT.2308-8893.2019.4.11

ІГРОВІ СТРАТЕГІЇ ПРИЙНЯТТЯ РІШЕНЬ В ІЄРАРХІЧНИХ СИСТЕМАХ. ІІ. КОМП'ЮТЕРНЕ МОДЕЛЮВАННЯ СТОХАСТИЧНОЇ ГРИ

П.О. КРАВЕЦЬ

Анотація. Розроблено алгоритм розв'язування стохастичної гри для прийняття рішень в ієрархічних системах в умовах невизначеності. Виконано аналіз результатів комп'ютерного моделювання стохастичної гри для автократичної, анархічної та демократичної систем прийняття рішень зі структурою бінарного дерева. Установлено, що найменший період навчання для досягнення близького до консенсусного рішення мають автократично-центричні ієрархічні системи. Вивчено вплив параметрів на збіжність ігрового методу у процесі пошуку консенсусного або мажоритарного колективного рішення.

Ключові слова: прийняття рішень, ієрархічна система, умови невизначеності, стохастична гра, комп'ютерне моделювання.

ВСТУП

Ієрархічна система прийняття рішень — це багаторівнева система з розміщенням елементів від вищого до нижчого рангу [1–4]. У такій організаційній структурі можна виокремити елементи вищого, середнього та нижчого рівнів. Елементи одного рівня зазвичай мають однаковий ранг або однакові можливості щодо прийняття рішень.

Елементи найвищого рівня генерують керувальні рішення для елементів нижчого рівня. Елементи середнього рівня на основі рішень, надісланих від елементів вищого рівня, виробляють власні рішення для елементів нижчого рівня. Рішення елементів найнижчого рівня спрямовані на виконання конкретних робіт.

Для дослідження ієрархічних систем прийняття рішень в основному використовують методи теорії ігор — безкоаліційних або коаліційних, детермінованих або стохастичних, дискретних або диференціальних, у нормальній формі (з одночасним вибором варіантів дій) або ієрархічних ігор (з правом першого ходу) [3–8]. Залежно від постановки завдання досліджуються стратегії прийняття рішень, які забезпечують досягнення класичних колективних розв'язків, оптимальних за Нешем, Парето, Слейтером, Джофріоном, Байєсом, Штакельбергом або інших [9, 10]. Серед різноманіття ігрових задач виділимо підклас стохастичних ігор у нормальній формі, які дають можливість знаходити колективні розв'язки гри в умовах стохастичної невизначеності матриць платежів [8].

Гравець є активною інформаційною сутністю, здатною взаємодіяти з іншими гравцями, навчатися і приймати самостійні рішення на основі опрацювання поточних даних. Використовуючи термінологію мультиагентних систем [11–15], таку активну сутність назвемо агентом прийняття рішень.

Оскільки прийняття рішень в ієрархічній системі спрямоване на колективне розв'язування конкретного завдання, то дії агентів повинні бути скоординованими. Координація — це узгодження дій агентів системи прийняття рішень з метою досягнення поставлених цілей. Координація здійснюється шляхом виконання керівних рішень або через домовленості між елементами системи [13, 14]. У цій роботі розглядаються обидва варіанти: 1) рішення вважається сформованим, якщо у ході стохастичної гри більшість агентів вибрали стратегії, що збігаються з наперед заданою керівною стратегією; 2) агенти приймають скоординоване колективне рішення у ході взаємодії і самонавчання.

Поряд з теоретичним дослідженням систем прийняття рішень важливим є розроблення комп'ютерних засобів моделювання для апробації та доповнення теоретичних результатів на практиці.

Комп'ютерне моделювання ієрархічних систем прийняття рішень є актуальною прикладною проблемою, вирішення якої дозволяє уточнити результати теоретичного оцінювання умов збіжності методів прийняття рішень у конкретних практичних ситуаціях.

Практичні проблеми прийняття рішень в ієрархічних системах досліджуються спеціалістами з багатьох предметних галузей, пов'язаних з колективною поведінкою та колективною взаємодією. Для цього використовують або готові інструментальні засоби, або спеціально розроблені програмні продукти, коли необхідно врахувати певні особливості постановки завдання колективного прийняття рішень [15–17].

Мета роботи — розроблення алгоритмічних та програмних засобів моделювання стохастичних ігор у нормальній формі для координації рішень в ієрархічних системах та визначення факторів, які впливають на збіжність ігрового методу в умовах стохастичної невизначеності.

Результати роботи можуть бути використані для підтримання прийняття скоординованих колективних рішень у системах з ієрархічною організацією.

ІЄРАРХІЧНІ СИСТЕМИ ПРИЙНЯТТЯ РІШЕНЬ

Нехай платежі ігрових агентів ієрархічної системи є функцією власних стратегій, стратегій агентів вищого і нижчого рівнів, зважених коефіцієнтом $\lambda \in [0,1]$. Значення коефіцієнта визначає ступінь впливу стратегій сусідніх агентів на формування платежів кожного агента і, відповідно, напрямки потоків інформації між рівнями ієрархії системи прийняття рішень. Відхилення від стратегії агента-керівника враховується з коефіцієнтом λ, а відхилення від стратегій агентів-виконавців — з коефіцієнтом $1-\lambda$. Залежно від значень цього коефіцієнта отримаємо різні види ієрархічних систем:

$\lambda = 1$ — автократична;

$\lambda = 0$ — анархічна;

$0 < \lambda < 1$ — множина демократичних систем.

Назви ієрархічних систем є умовними, а їх інтерпретація у цій роботі не замінює їх відомі трактування.

В автократичній системі агенти нижчого рівня беруть до виконання вказівні рішення агента вищого рівня. Потоки керування спрямовані згори вниз. Ієрархію анархічної системи слід розуміти як делегування повноважень на вищі рівні від самоорганізованих спільнот агентів найнижчого рівня. Потоки інформації в анархічній системі спрямовані знизу вгору. У демократичних системах на вироблення рішень впливають потоки інформації як згори вниз, так і знизу вгору.

Експериментальне дослідження ігрового методу прийняття рішень виконаємо для стохастичної гри зі структурою бінарного дерева. Симетричне бінарне дерево складається з $L = 2^m - 1$ вузлів, де m — кількість рівнів ієрархічної системи прийняття рішень ($m = 1, 2, ...$). Вузли дерева позначають гравців, а дуги – їх стратегії, які впливають на формування поточних програшів гравців. Нехай поточні втрати гравців визначаються власними стратегіями та стратегіями сусідніх гравців так, як це показано на рис. 1, що відповідає структурі демократичної системи з параметром $0 < \lambda < 1$.

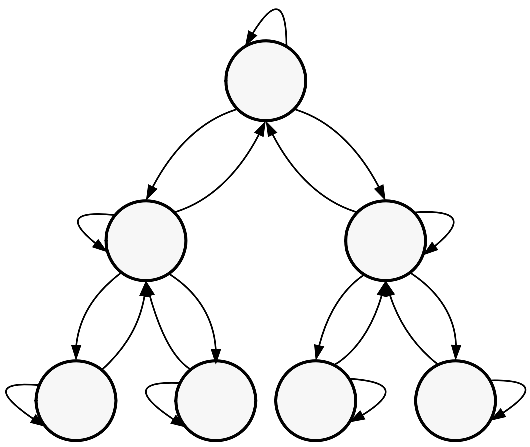

Рис. 1. Структура демократичної системи

Для граничних значень параметра $\lambda = 1$ та $\lambda = 0$ зі структури демократичної системи відповідно отримаємо структури автократичної та анархічної систем прийняття рішень, зображені на рис. 2–5.

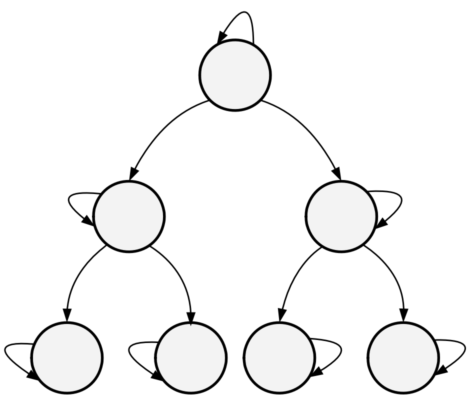

Рис. 2. Структура автократичної системи

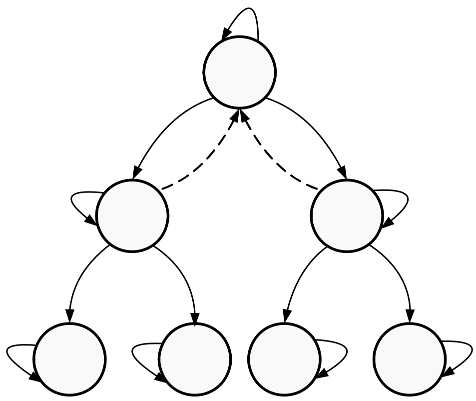

Рис. 3 Структура автократичної системи з колегіальним центром

Поведінка гравця найвищого рівня в автократичній системі, зображеній на рис. 2, є вільною від впливу стратегій інших гравців. Якщо кореневий гравець вибирає власні стратегії $u_n^{\text{root}} \in U^{\text{root}}$ за деяким дискретним розподілом $p^{\text{root}} \in S^{N_{\text{root}}}$, то його стратегія змінюється у часі: $u_n^{\text{root}} = \text{var}$, $n = 1, 2, \ldots$. Чим ближче розподіл p^{root} до рівноймовірного вибору чистих стратегій, тим більше кореневий гравець дезорієнтує гравців нижчих рівнів. У результаті вихід системи на узгоджене колективне рішення є проблематичним або неможливим.

Можливі такі два працездатні способи формування чистих стратегій цього гравця, які забезпечують координацію автократичної системи. Перший спосіб полягає в тому, що стратегії кореневого гравця є постійними $u_n^{\text{root}} = u^{\text{root}} = \text{const}$ у часі $n = 1, 2, \ldots$. У цьому випадку маємо *автократичну систему з догматичним центром*. Другий спосіб відповідає структурі *автократичної системи з колегіальним центром,* як це показано на рис. 3. У цьому випадку на формування втрат кореневого гравця впливають власні стратегії та стратегії хоча б одного із гравців нижчого рівня.

Аналогічно, якщо гравці найнижчого рівня анархічної системи обирають власні стратегії $u_n^i \in U^i$ з дискретними розподілами $p^i \in S^{N_i}$ (рис. 4), то вони генеруватимуть змінні стратегії $u_n^i = \text{var}$ поведінки у часі і з близькою до 1 імовірністю система не зможе досягнути скоординованого консенсусного (вирівнювального) рішення.

Для того щоб анархічна система вийшла на узгоджене рішення, можна скористатися одним із двох способів. Перший полягає у тому, що всі гравці найнижчого рівня, попередньо домовившись, дотримуються однакових стратегій поведінки $u_n^i = u = \text{const}$ у часі. У цьому випадку маємо *анархічну систему з догматичною спільнотою*. Другий спосіб полягає у тому, що гравці найнижчого рівня об'єднуються у спільноту і їх поточні втрати визначаються як власними стратегіями, так і стратегіями інших гравців спільноти. У цьому випадку (рис. 5) маємо *анархічну систему з горизонтальними зв'язками* на найнижчому рівні (з *організованою спільнотою*). Для працездатності такої анархічної системи повинен існувати циклічний шлях залежностей платіжних функцій від стратегій гравців, яким можна обійти усі елементи найнижчого рівня. Тоді однакові стратегії дерева агентів будуть формуватися динамічно в часі, у процесі навчання гравців.

Розглянуті ієрархічні системи є *активними* системами прийняття рішень, оскільки агенти не просто транслюють рішення між рівнями ієрархії, а мають власні стратегії поведінки.

Крім того, ієрархічні системи є системами з *розподіленим* (або колективним) прийняттям рішень, виходячи з того, що вони складаються з організованої за рангами множини автономних агентів (ураховано структурний аспект).

Якщо у процесі навчання (вироблення рішень) розподілена система прагне вийти на рішення центру, то вона є *централізованою*. Рішення центру відоме до початку гри і не змінюється у ході гри. Центром може бути один агент або група агентів з наперед узгодженим рішенням. Якщо ж у процесі навчання розподілена система формує узгоджене колективне рішення, яке не відоме до початку гри, то вона є *децентралізованою* (враховано функціональний аспект).

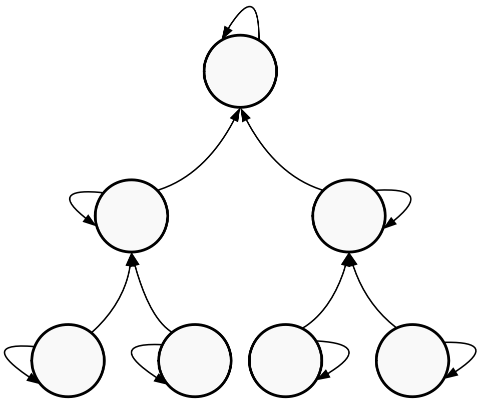

Рис. 4. Структура анархічної системи

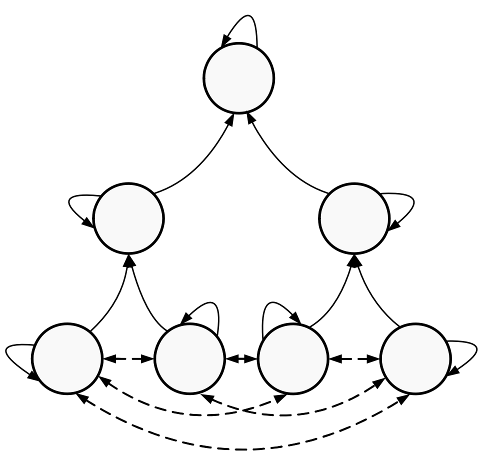

Рис. 5 Структура анархічної системи з горизонтальними зв'язками

Із цих міркувань догматична автократична система є централізованою (з потоками керівних рішень згори вниз), автократична система з колегіальним центром — децентралізованою, анархічна система з догматичною спільнотою — централізованою (знизу вгору), анархічна система з горизонтальними зв'язками — децентралізованою, демократичні системи є децентралізованими.

АЛГОРИТМ РОЗВ'ЯЗУВАННЯ СТОХАСТИЧНОЇ ГРИ

Крок 1. Задати початкові значення параметрів гри: $n = 0$ — початковий момент часу; D — множина (дерево) гравців; $L = |D|$ — потужність множини (кількість) гравців; N_i — кількості чистих стратегій гравців $\forall i \in D$; $U^i = \{u^i(1), u^i(2), ..., u^i(N_i)\}$ $\forall i \in D$ — вектори чистих стратегій гравців; $p_0^i = (1/N_i, ..., 1/N_i)$ $\forall i \in D$ — змішані стратегії гравців; $\gamma > 0$ — параметр кроку навчання; $\alpha \in (0,1]$ — порядок кроку навчання; $\varepsilon > 0$ — параметр ε-симплексу; $\beta > 0$ — порядок швидкості розширення ε-симплексу; $\lambda \in [0,1]$ — ваговий коефіцієнт поточних програшів; $n_{\max}$ — максимальна кількість кроків стохастичної гри.

Крок 2. Вибрати поточні чисті стратегії:

$$u_n^i = \left\{ u^i(j) \,|\, j = \arg\min_j \sum_{k=1}^{j} p_n^i(k) > \omega \ (j = 1..N_i) \right\} \ \forall i \in D,$$

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

Крок 3. Отримати значення поточних програшів:

$$\xi_n^i(u_n^{D_i}) = |D_i|^{-1}\left(\lambda\,|u_n^i - u_n^k| + (1-\lambda)\sum_{j\in D_i\setminus\{k\}}|u_n^i - u_n^j|\right) + \mu_n \quad \forall i \in D,$$

де D_i — множина сусідніх гравців i-го гравця; $u_n^{D_i} \in U^{D_i} = \underset{j\in D_i}{\times} U^j$ — колективна стратегія множини гравців D_i; $u_n^i \in R^1$ — числовий еквівалент варіанта рішення; k — гравець вищого рівня (керівник); i — гравець середнього рівня; j — гравець нижчого рівня; $\mu_n \sim Normal(0,d)$ — білий гауссівський шум з дисперсією $d \geq 0$; $\lambda \in [0,1]$ — ваговий коефіцієнт. Для кореневого гравця $\lambda = 0$, а для гравців найнижчого рівня $\lambda = 1$.

Крок 4. Обчислити значення параметрів:

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

де $\gamma_n > 0$ — монотонно спадна послідовність додатних величин, яка регулює величину кроку методу; $\varepsilon_n > 0$ — монотонно спадна послідовність додатних величин, яка регулює швидкість розширення ε-симплексу.

Крок 5. Визначити нові значення векторів змішаних стратегій:

$$p_{n+1}^i = \pi_{\varepsilon_{n+1}}^{N_i}\{p_n^i - \gamma_n \xi_n^i [e(u_n^i) - p_n^i]\} \quad \forall i \in D,$$

де $\pi_{\varepsilon_{n+1}}^{N_i}$ — проектор на одиничний ε-симплекс $S_{\varepsilon_{n+1}}^{N_i} \subseteq S^{N_i}$ [18]; $p_n^i \in S_{\varepsilon_n}^{N_i}$ — змішані стратегії i-го гравця; $\xi_n^i \in R^1$ — поточний програш гравця; $e(u_n^i)$ – одиничний вектор-індикатор вибору варіанта $u_n^i \in U^i$.

Крок 6. Розрахувати характеристики якості ігрового прийняття рішень:

1) функція усереднених у часі втрат системи:

$$\Xi_n = (nL)^{-1}\sum_{t=1}^{n}\sum_{i\in D}\xi_t^i;$$

2) коефіцієнт узгоджених рішень гравців:

$$\mathrm{K}_n = (nL)^{-1}\sum_{t=1}^{n}\sum_{i\in D}\chi\left(\sum_{s\in D_i}|u_t^i - u_t^s| = 0\right),$$

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

3) середня похибка відхилення змішаних стратегій гравців від змішаної стратегії гравця-керівника найвищого рівня:

$$\Delta_n = (nL)^{-1}\sum_{t=1}^{n}\sum_{i\in D}\left\|p_{i,t} - p_{\mathrm{root},t}\right\|,$$

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

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

Крок 8. Якщо $n < n_{\max}$, то перейти до кроку 2, інакше — кінець гри.

РЕЗУЛЬТАТИ ЕКСПЕРИМЕНТАЛЬНИХ ДОСЛІДЖЕНЬ

Результати комп'ютерного моделювання процесу ігрового прийняття рішень в ієрархічних системах зі структурою симетричного бінарного дерева подано у вигляді графіків на рис. 6 – рис. 15. Результати отримано для таких значень параметрів гри: $L = 7$ (для $m = 3$ рівнів ієрархічної системи прийняття рішень), $N_i = N = 4$, $\gamma = 1$, $\varepsilon = 0{,}999/N$, $\alpha = 0{,}01$, $\beta = 1$, крім випадків, коли вивчається вплив цих параметрів на збіжність ігрового методу. Для компактного зображення результатів моделювання з великим діапазоном значень використано логарифмічну шкалу.

Ураховуючи те, що метою гравців ієрархічної системи є вирівнювання чистих стратегій, їх матриці платежів (програші) задаються так: $[v^i] = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ $\forall i \in D$. Якщо гравцями вибрано однакові стратегії, то їх програш дорівнює 0, інакше — 1.

Для таких матриць платіжні функції

$$V^i = p_1^1 p_1^2 v_{1,1} + p_1^1 p_2^2 v_{1,2} + p_2^1 p_1^2 v_{2,1} + p_2^1 p_2^2 v_{2,2} =$$

$$= p_1^1 (1 - p_1^2) + (1 - p_1^1) p_1^2 = p_1^1 + p_1^2 - 2 p_1^1 p_1^2$$

мають вигляд сідлової поверхні у просторі змішаних стратегій гравців.

На рис. 6 зображені зрізи значень сідлової платіжної функції, отримані з кроком $h = 1/16$, і траєкторія навчання ланки ієрархічної стохастичної гри «керівник – підлеглий», кожен з яких має по два варіанти прийняття рішень (гра 2×2). Показано зрізи платіжної функції без дії завад ($d = 0$) та в умовах дії завад ($d = 0{,}01$). У вершинах та в центрі квадрата позначено значення платіжної функції для комбінованих стратегій двох гравців.

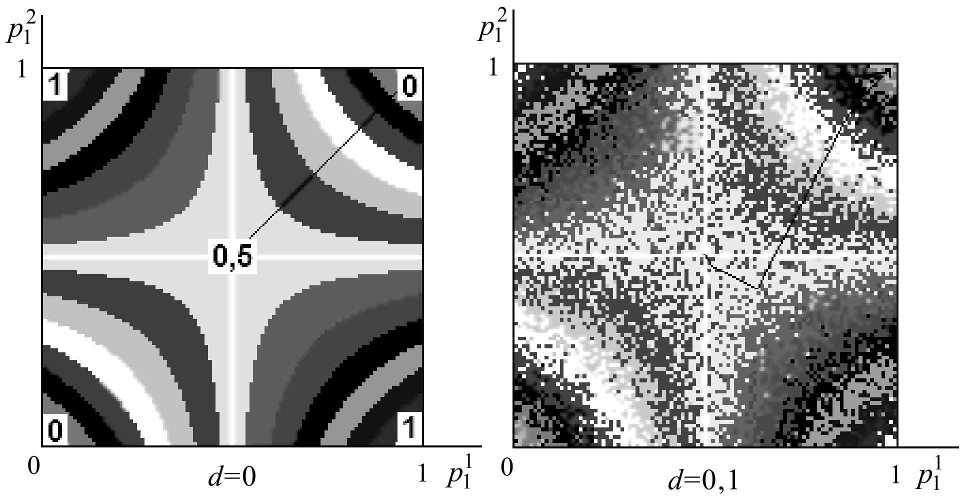

Рис. 6. Проекція зрізів платіжної функції на площину

Гра з вирівнювання стратегій має два стійкі рівноцінні розв'язки у чистих стратегіях, розміщені у вершинах (0, 0) та (1, 1) зі значеннями платіжних функцій $V^i = 0$, та нестійкий розв'язок у змішаних стратегіях у точці (0,5; 0,5) зі значеннями платіжних функцій $V^i = 0{,}5$. Для розв'язків у чистих стратегіях справедливі умови рівноваги за Нешем та оптимальності за Парето для гри у нормальній формі, а також умова оптимальності за Штакельбергом для послідовної гри. Розв'язок у змішаних стратегіях задовольняє умову рівноваги за Нешем.

Якщо гравці дотримуються оптимальних стратегій у точці рівноваги за Нешем, то зі зміною стратегії одного з них він не зможе зменшити свій програш (для задачі мінімізації середніх втрат). Оптимальність за Парето випливає з тотожності матриць платежів гравців. У точці оптимальності за Парето не існує спільної стратегії, яка дозволить зменшити середній програш одночасно для всіх гравців. Оптимальність за Штакельбергом визначається за деревом можливих комбінацій чистих стратегій, які мінімізують втрати гравців, з урахуванням попередніх ходів гравців вищих рівнів ієрархії, виходячи із оптимістичної гіпотези про те, що гравці сприяють один одному у виборі ходів гри.

За відсутності дії завад ($d = 0$) пошукова траєкторія стохастичної гри прямує до однієї з точок рівноваги за найкоротшим, близьким до лінійного, шляхом. Дія завад ($d > 0$) призводить до викривлення траєкторії пошуку.

Графіки однієї з реалізацій функцій коефіцієнта скоординованих стратегій K_n, середніх програшів гравців Ξ_n та похибки змішаних стратегій Δ_n, які характеризують ефективність стохастичної гри демократичної ($\lambda = 0{,}5$) системи прийняття рішень, зображено на рис. 7. Результати отримано для системи без завад ($d = 0$) та в умовах дії завад ($d = 1$). Загальний вигляд та взаємне розташування графіків функцій зберігається у разі повторення експерименту для різних послідовностей випадкових величин. Для автократичної ($\lambda = 1$, з догматичним або колегіальним центром) та анархічної ($\lambda = 0$, з догматичною або організованою спільнотою) систем отримаємо подібні залежності [3, 4].

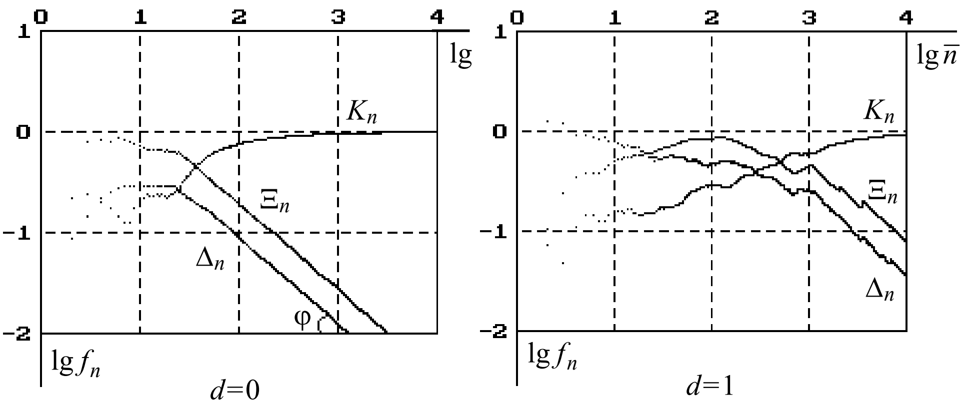

Рис. 7. Характеристики ігрового прийняття рішень в демократичній системі

Спадання функцій Ξ_n, Δ_n та зростання коефіцієнта координації K_n у часі свідчать про збіжність ігрового методу згідно із сформульованою метою. Порядок швидкості збіжності може бути оцінений тангенсом кута φ лінійної апроксимації графіка норми Δ_n відхилення змішаних стратегій гравців від контрольних значень та віссю часу. Як видно з рис. 7, за відповідного підбору параметрів досягається близький до 1 порядок степеневої швидкості збіжності розглянутої стохастичної гри.

У системах із завадами для збіжності ігрового методу необхідна більша кількість пошукових кроків для досягнення оптимального колективного розв'язку.

Запропонований метод розв'язування ієрархічної стохастичної гри прийняття рішень належить до класу реактивних методів і має відносно невелику швидкість збіжності. Це зумовлено тим, що на початок гри немає ніякої інформації про середовище, з яким взаємодіють гравці. Інформація збирається у процесі навчання шляхом адаптивної перебудови векторів змішаних стратегій пропорційно до значень поточних програшів.

Графіки на рис. 8–15 усереднено за 100 реалізаціями стохастичної гри з різними послідовностями випадкових величин.

Графік залежності періоду $\overline{n}$ навчання гравців вибирати скоординовані рішення з коефіцієнтом $K_n = 0{,}9$ показано на рис. 8, а залежності коефіцієнта координації K_n на момент $n_{\max} = 10^4$ кроків гри від параметра λ ієрархічної системи — на рис. 9. Значення параметра $0 < \lambda < 1$ відповідає одній з демократичних (див. рис. 1), $\lambda = 1$ — автократичній (рис. 2) і $\lambda = 0$ — анархічній (рис. 3) системам прийняття рішень.

Як видно з рис. 9, близьке до консенсусного рішення отримується для діапазону значень параметра $\lambda \in [0{,}3; 0{,}9]$. У цей діапазон потрапляють анархічно-центричні ($0 < \lambda < 0{,}5$), нейтральна ($\lambda = 0{,}5$) та автократично-центричні ($0{,}5 < \lambda < 1$) демократичні системи прийняття рішень.

Аналіз графіка на рис. 8 показує, що анархічно-центричним системам необхідний більший період навчання для прийняття скоординованих рішень, ніж автократично-центричним. Для заданих початкових даних найменший середній час навчання має автократично-центрична система з параметром $\lambda = 0{,}8$, значення якого вказує на те, що у ході формування колективного рішення стратегії гравців середнього рівня ієрархії на 80% визначаються стратегіями керівника і на 20% стратегіями підлеглих працівників.

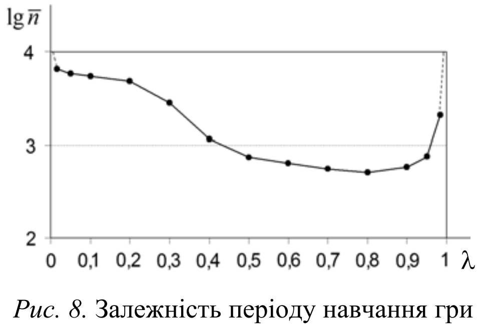

Рис. 8. Залежність періоду навчання гри від параметра λ

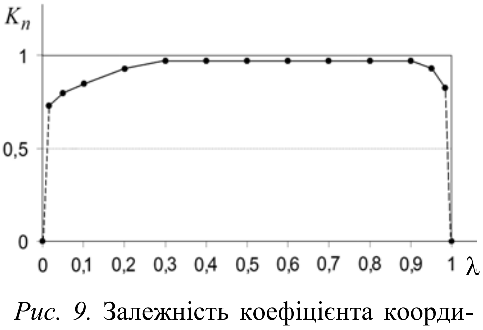

Рис. 9. Залежність коефіцієнта координації від параметра λ

Ігрові моделі автократичної та анархічної систем, отримані з демократичної системи для крайніх значень параметра λ, не забезпечують координації стратегій гравців. Для $\lambda = 1$ кореневий гравець, а для $\lambda = 0$ гравці найнижчого рівня є ізольованими від інших гравців і вибирають змінні стратегії: $u_n^{\text{root}} = \text{var}$, $u_n^i = \text{var}$, $n = 1, 2, \ldots$. У результаті це дезорганізує роботу інших агентів відповідно автократичної або анархічної систем прийняття рішень.

Для координації автократичної системи кореневий гравець повинен дотримуватися постійної стратегії $u_n^{\text{root}} = u = \text{const}$ у часі $n = 1, 2, \ldots$. Така стратегія задається ще до початку гри (догматичний центр), або може бути сформована за взаємодії з агентами нижчого рівня у процесі навчання (колегіальний центр).

Аналогічно для координації анархічної системи потрібна попередня домовленість між усіма агентами спільноти найнижчого рівня ієрархії $u_n^i = u = \text{const}$, $n = 1, 2, \ldots$ (догматична спільнота), або наявність горизонтальних зв'язків між агентами спільноти з можливістю навчання агентів у ході гри.

Значний вплив на координацію рішень гравців має розмірність стохастичної гри, яка визначається кількістю гравців та кількістю їх чистих стратегій.

Залежність періоду $\overline{n}$ навчання (середньої кількості кроків гри), необхідного для досягнення коефіцієнта узгоджених рішень $K_n = 0{,}9$, від кількості агентів $L = 2^m - 1$, організованих у вигляді симетричного бінарного дерева прийняття рішень з $m = 1, 2, \ldots$ рівнями зображено на рис. 10, а від кількості чистих стратегій N для усіх вище розглянутих видів ієрархічних систем прийняття рішень за відсутності дії завад — на рис. 11. Номери графіків відповідають таким ієрархічним системам: 1 — демократична ($\lambda = 0{,}5$), 2 — догматична автократична, 3 — автократична з колегіальним центром, 4 — анархічна з догматичною спільнотою, 5 — анархічна з горизонтальними зв'язками на найнижчому рівні ієрархії у вигляді повнозв'язного мультиграфу.

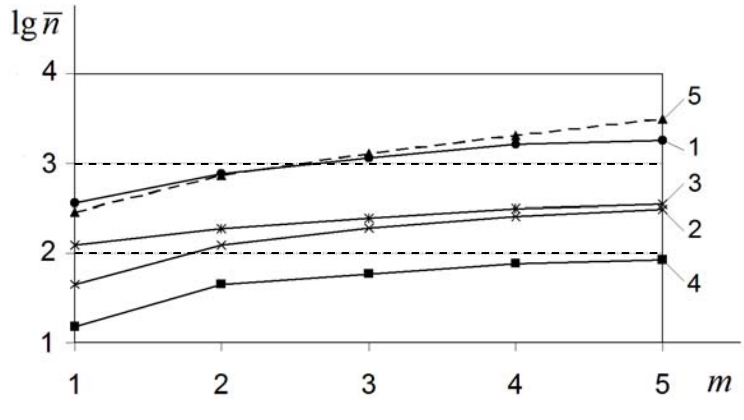

Рис. 10. Залежність періоду навчання гри від кількості агентів

Як видно з рис. 10 та 11, зростання кількості гравців та їх чистих стратегій призводить до збільшення періоду навчання усіх видів ієрархічних стохастичних ігор.

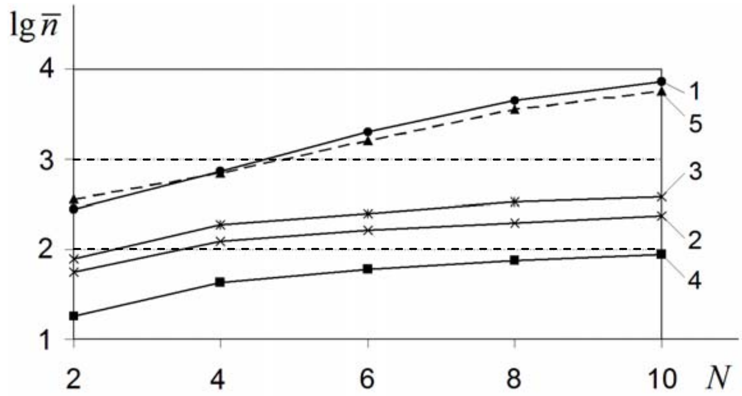

Рис. 11. Залежність періоду навчання гри від кількості стратегій агентів

Час збіжності ігрового процесу прийняття рішень в ієрархічних системах залежить від того, які домовленості прийнято в системі перед початком гри. Чим більше агентів мають початкове узгоджене рішення, тим менший час навчання системи. Те, що анархічна система з догматичною спільнотою (графік 4) має найменший час вироблення рішення (час навчання) визначається початковими умовами — усі 2^{m-1} агентів найнижчого рівня (половина від загальної кількості агентів) мають узгоджені (однакові) рішення ще до початку гри.

Зі збільшенням кількості зв'язків між агентами зростає кількість станів ієрархічної системи, які використовуються для випадкового пошуку оптимальних рішень і, відповідно, зростає час такого пошуку. У зв'язку з цим автократична система (графіки 2 та 3) є ефективнішою від демократичної системи (графік 1), догматична автократична система (графік 2) буде ефективнішою від автократичної системи з колегіальним центром (графік 3), а анархічна система з догматичною спільнотою (графік 4) — ефективнішою від анархічної системи з горизонтальними зв'язками (графік 5) між агентами найнижчого рівня ієрархії.

Період навчання стохастичної гри залежить від коефіцієнта координації $K_n \in (0,1]$, який задає потрібну частку гравців з узгодженими рішеннями. Залежність періоду навчання від коефіцієнта координації для різних видів ієрархічних систем за відсутності дії завад зображено на рис. 12.

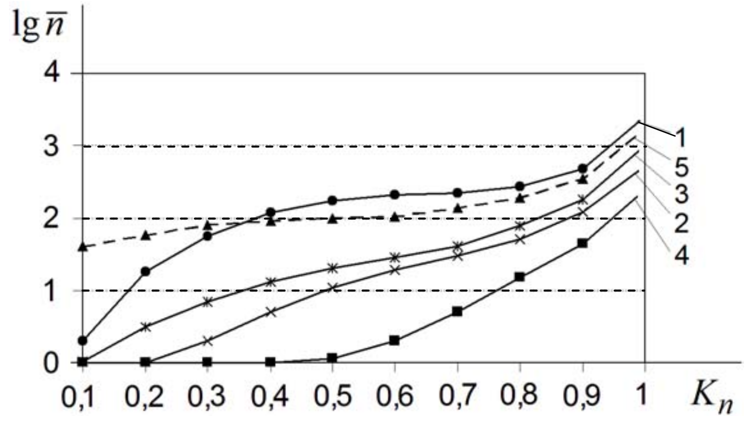

Рис. 12. Залежність періоду навчання гри від коефіцієнта координації

Як видно з рис. 12, зі зростанням вимог до ступеня координації ієрархічних систем зростає період їх навчання, причому зберігається порядок ранжування систем за швидкістю збіжності стохастичної гри так, як це зображено на рис. 10 та 11.

Крім розмірності гри, на швидкість збіжності ігрового методу впливають його параметри α та β, співвідношення яких повинні задовольняти фундаментальні умови стохастичної апроксимації [18–20]. Залежність періоду навчання гри $\overline{n}$ від параметрів α та β зображено на рис. 13, а залежність коефіцієнта координації від цих параметрів — на рис. 14. Результати отримано для демократичної системи з параметром $\lambda = 0{,}5$ за відсутності завад.

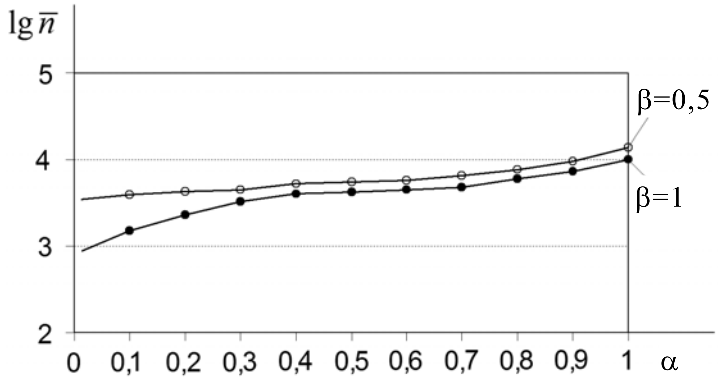

Рис. 13. Залежність періоду навчання гри від параметрів ігрового методу

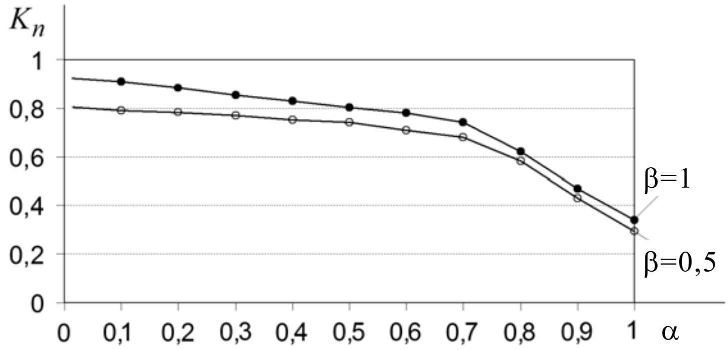

Рис. 14. Залежність коефіцієнта координації від параметрів ігрового методу

Зростання параметра $\alpha \in (0; 1]$ призводить до зменшення кроку навчання стохастичної гри (кроку зміни ймовірностей вибору чистих стратегій). Зменшення параметра $\beta \in (0;1]$ сповільнює швидкість розширення ε-симплексу, що є додатковим фактором обмеження кроку навчання. Результатом цього є зменшення коефіцієнта координації та збільшення періоду навчання стохастичної гри.

Для заданих параметрів ігрового методу мінімальна кількість кроків навчання стохастичної гри, яка забезпечує координацію понад 90% гравців, досягається для $\alpha \in (0; 0{,}1]$ та $\beta = 1$.

Вплив випадкових завад у вигляді білого шуму з дисперсією d на період $\overline{n}$ навчання стохастичної гри для різних видів ієрархічних систем зображено на рис. 15.

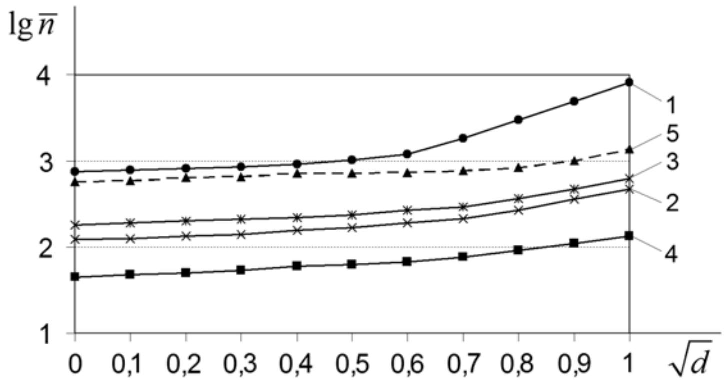

Рис. 15. Вплив завад на період збіжності стохастичної гри

Дія завад призводить до спотворення потоків даних між рівнями ієрархії системи прийняття рішень. Відповідно зі зростанням дисперсії завади збільшується період вироблення скоординованого колективного рішення в ієрархічній системі.

Позитивним фактором практичної реалізації розглянутого ігрового алгоритму внаслідок дії завад є забезпечення координації автократичної системи прийняття рішень з недогматичним центром ($\lambda = 1$, $u_n^{\text{root}} = \text{var}$, $n = 1, 2, \ldots$). Дія білого шуму зумовлює диференціацію елементів вектора змішаних стратегій гравців і за рахунок підсилювальної властивості рекурентного методу кореневий гравець навчиться вибирати одну із чистих стратегій ($u_n^{\text{root}} = u^{\text{root}} = \text{const}$, $n = 1, 2, \ldots$) з близькою до 1 імовірністю. У результаті проведення самонавчальної стохастичної гри автократична система з недогматичним центром вийде на консенсусне рішення.

ВИСНОВКИ

Проведені експериментальні дослідження підтверджують збіжність стохастичної гри для координації рішень в ієрархічних системах. Аналіз результатів комп'ютерного моделювання дозволяє зробити такі узагальнення:

1. Розроблений алгоритм та виконане на його основі комп'ютерне моделювання стохастичної гри забезпечують узгоджене прийняття рішень в ієрархічних системах за рахунок вирівнювання стратегій гравців на основі збирання поточної інформації та її адаптивного опрацювання.

2. Час збіжності стохастичної гри визначається розмірністю ієрархічної системи (кількістю учасників прийняття рішення і кількістю чистих стратегій гравців) та співвідношенням параметрів ігрового методу. Оптимізація параметрів ігрового методу за обмежувальної дії фундаментальних умов стохастичної апроксимації забезпечує близький до 1 степеневий порядок швидкості збіжності.

3. Достовірність результатів забезпечується належним плануванням комп'ютерного експерименту та підтверджується повторювальністю значень характеристик стохастичної гри, отриманих у ході імітаційного моделювання для різних послідовностей випадкових величин.

4. Виконані у цій роботі дослідження можна розвинути в напрямі моделювання ресурсних обмежень на інформаційні потоки між рівнями

ієрархії і врахування відновлювальних відмов агентів для вивчення надійності та гнучкості ієрархічних систем прийняття рішень.

ЛІТЕРАТУРА

1. *Hierarchies* in Distributed Decision Making / Christoph Schneeweiss. — Springer, 2013. — 341 p.
2. *Недашківська Н.І.* Методологія та інструментарій підтримки прийняття рішень на основі ієрархічних та мережевих моделей: дис. … доктора техн. наук: 01.05.04 / Н. І. Недашківська. — К.: ННК «ІПСА» НТУУ «КПІ», 2018. — 407 с. — [Електрон. ресурс]. — Режим доступу: http://ela.kpi.ua/bitstream/123456789/25119/1/Nedashkivska_diss.pdf.
3. *Кравець П.О.* Ігрова модель прийняття рішень в ієрархічних системах / П.О. Кравець // Інформаційні системи та мережі: вісн. НУ «Львівська політехніка». — 2017. — № 872. — С. 111–120.
4. *Кравець П.О.* Ігрова модель системи з авторитарним прийняттям рішень / П.О. Кравець // Інформаційні системи та мережі: вісн. НУ «Львівська політехніка». — 2018. — № 901. — С. 61–67.
5. *Гермейер Ю.Б.* Игры с непротивоположными інтересами / Ю.Б. Гермейер. — М.: Наука, 1976. — 328 с.
6. *Harrington J. E., Jr.* Games, Strategies, and Decision Making / J.E. Harrington, Jr. — Worth Publishers, 2014. — 540 p.
7. *Grabisch M.* Set Functions, Games and Capacities in Decision Making / M. Grabisch. — Springer, 2016. — 473 p. — DOI: 10.1007/978-3-319-30690-2.
8. *Ummels M.* Stochastic Multiplayer Games: Theory and Algorithms / M. Ummels. — Amsterdam University Press, 2014. — 174 p.
9. *Petrosjan L.A.* Game Theory and Application / L.A. Petrosjan, V.V. Mazalov. — Nova Science Publishers, 2002. — 295 p.
10. *Ungureanu V.* Pareto-Nesh-Stackelberg Game and Control Theory: Intelligent Paradigms and Applications / V. Ungureanu. — Springer, 2018. — 343 p.
11. *Wooldridge M.* An Introduction to Multiagent Systems / M. Wooldridge. — John Wiley & Sons, 2009. — 461 p.
12. *Radley N.* Multi-Agent Systems – Modeling, Control, Programming, Simulations and Applications / N. Radley. — Scitus Academics LLC, 2017. — 284 p.
13. *Iterative* Learning Control for Multi-agent Systems Coordination / S. Yang, J.-X. Xu, X. Li, D. Shen. — John Wiley & Sons, 2017. — 272 p.
14. *Sun Z.* Cooperative Coordination and Formation Control for Multi-agent Systems / Zhiyong Sun. — Springer, 2018. — 179 p.
15. *Agent* for Games and Simulations: Trends in Techniques, Concepts and Design / F. Dignum, J.Bradshaw, B. G. Silverman, W. van Doesburg. — Springer, 2009. — 237 p.
16. *Bekker K.* The Guide to Computer Simulation and Games / K. Bekker, J.R. Parker. — John Wiley and Sons, 2011. — 456 p.
17. *Simulation* of Decision-Making as Active Learning Tools: Design and Effects of Political Science Simulations / P.Bursens, V. Donche, D. Gijbels, P. Spooren (Editors). — Springer, 2018. — 206 p.
18. *Назин А.В.* Адаптивный выбор вариантов / А.В. Назин, А.С. Позняк. — М.: Наука, 1986. — 288 с.
19. *Kushner H.* Stochastic Approximation and Recursive Algorithms and Applications / H. Kushner, G. G. Yin. — Springer Science & Business Media, 2013. — 417 p.
20. *Benveniste A.* Adaptive Algorithms and Stochastic Approximations / A. Benveniste, M. Metivier, P. Priouret. — Springer Science & Business Media, 2012. — 365 p.

Надійшла 07.06.2019