УДК 004.415.2; 004.852; 004.942

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

ПРОГРАМНІ ІНСТРУМЕНТАЛЬНІ ЗАСОБИ ІГРОВОГО МОДЕЛЮВАННЯ МУЛЬТИАГЕНТНИХ СИСТЕМ

Описується розроблений інструментальний програмний засіб, призначений для дослідження методів прийняття рішень в мультиагентних системах. Моделювання мультиагентних систем базується на методах теорії ігор, які дозволяють визначити структуру гри, умови взаємодії між гравцями, функції винагород.

Ключові слова – мультиагентна система, моделювання, ООП

This paper describes the developed program tool which is intended for research of decision making methods in the multiagent systems. Multiagent systems which are modeled in this work are based on the game theory methods, which define a structure of game, terms of co-operation between players, functions of rewards.

Keywords – multiagent system, modeling, OOP

Вступ

Розроблення технології штучних агентів та створення мультиагентних систем (МАС) є важливим і перспективним напрямком розвитку нових інформаційних і комунікаційних технологій на основі інтеграції сучасних мережних технологій, методів і засобів штучного інтелекту, баз даних та знань, систем об’єктно-орієнтованого проектування [1].

Агентно-орієнтований підхід знаходить широке застосування для розподіленого розв’язування складних задач і ефективного розв’язування розподілених задач, реінжинірингу бізнесу і побудови віртуальних підприємств, імітаційного моделювання інтегрованих виробничих систем і електронної торгівлі, організації роботи колективів роботів і розподіленої розробки комп'ютерних програм. У найближчому майбутньому він займе центральне місце при розвитку засобів керування інформацією і знаннями, при створенні і впровадженні новітніх систем телекомунікації, розвитку глобальних комп'ютерних мереж, особливо, мережі Інтернет.

Концепція агентів, розроблена в межах мультиагентних технологій, передбачає наявність активності, тобто здатності агента самостійно реагувати на зовнішні події і приймати рішення щодо наступної поведінки. У мультиагентних системах кожен агент будує власну модель поточного рішення, грунтуючись на своїх даних та даних інших агентів.

Згідно з [2], агент – це програмно або апаратно реалізована система, яка діє для досягнення поставлених користувачем цілей. Агент має такі властивості:

– автономність – здатність функціонувати без прямого втручання людей або комп'ютерних засобів і при цьому здійснювати самоконтроль над своїми діями та внутрішніми станами;

– суспільне поводження – здатність взаємодіяти з іншими агентами та людьми, обмінюючись повідомленнями за допомогою мов комунікації;

– реактивність – здатність сприймати стани середовища (фізичного світу, користувача, інших агентів, інформаційної мережі) та реагувати на них через відповідні інтерфейси;

– проактивність – здатність агентів самостійно приймати рішення та здійснювати цілеспрямовану поведінку.

Крім цього, агенти можуть наділятися додатковими ментальними та психологічно вмотивованими властивостями [3].

Основний зміст сучасних розробок МАС пов'язаний і синтезом індивідуальних властивостей агентів, особливостей їх поведінки, враховуючи задану групову динаміку, обумовлену відносинами кооперації або конкуренції, конфлікту або співробітництва, субординації або координації. У науковій літературі з цієї проблеми досліджуються ситуації взаємодії агентів, що визначають структуру МАС в умовах сумісності або антагонізму цілей, комунікації та встановлення локальних просторових і тимчасових відносин між агентами, навчання агентів та колективного прийняття рішень в умовах невизначеності [4 – 7].

З врахуванням цього, очевидною є спорідненість МАС та теорії ігор, які мають спільні предмети дослідження. Для розв’язання задач МАС можна використовувати ігрові моделі та методи, які добре описують ті ситуації, які можуть виникати в процесі роботи МАС. Так, математичні моделі теорії стохастичних ігор дають можливість формального дослідження МАС в умовах невизначеності – навчання, самоорганізації та адаптивної поведінки агентів, координації дій, прийняття рішень у недетермінованих середовищах [8].

Однією із важливих проблем МАС є побудова ефективної координації дій (взаємодії) агентів у процесі розподіленого розв’язування поставленої задачі [1, 9]. У ході взаємодії відбувається обмін поточною інформацією між агентами. У системі з децентралізованим керуванням взаємодія, як правило, має локально-обумовлену структуру, динаміка якої визначається метою та алгоритмом розв’язування задачі і здійснюється на основі протоколів комунікації агентів. Важливим фактором, що впливає на динаміку взаємодії, є також надійність функціонування агентів [10]. Дослідження процесів взаємодії агентів в умовах невизначеності можна виконати за допомогою апарату стохастичних ігор.

Для розв’язання стохастичної гри використовуються рекурентні методи [11, 12], побудовані на основі стохастичної апроксимації [13] детермінованої гри [14]. Працездатність методів розв’язування стохастичної гри визначається умовами їх збіжності до колективних станів рівноваги. Визначення умов збіжності ігрових методів можна виконати аналітично на основі оцінювання випадкових процесів [15]. Однак, точність аналітичного оцінювання у деяких випадках може бути неприйнятною для практичних застосувань. Тому, актуальним є розроблення програмних інструментальних засобів імітаційного моделювання стохастичних ігор.

Метою роботи є розроблення алгоритмів та програмних засобів ігрового моделювання мультиагентних систем в умовах невизначеності з врахуванням взаємодії та відмов агентів.

Постановка ігрової задачі

Стохастична гра задається кортежем $\Gamma=(D,U,\Xi)$, де D – непорожня множина гравців; $U=\underset{i\in D}{\times} U^i$ – дискретний простір комбінованих стратегій, утворений декартовим добутком чистих стратегій гравців U^i; $\Xi=\left[\xi(u)\right]$ – матриця джерел випадкових процесів $\xi(u)=(\xi^i(u)\forall i\in D)$, визначених для кожної комбінованої стратегії $u\in U$. Чисті стратегії U^i $\forall i\in D$ визначають можливі ходи гравців, а матриця Ξ моделює відповідні реакції середовища гри. У загальному випадку випадкові виграші i-го гравця $\xi^i=\xi^i(u^{D_i})$ є функціями комбінованих стратегій $u^{D_i}\in U^{D_i}=\underset{j\in D_i}{\times} U^j$ гравців з локальних підмножин $D_i\subseteq D,\ D_i\neq\varnothing\ \forall i\in D$.

В умовах невизначеності стохастичні характеристики процесів $\xi^i(u^{D_i})$ є апріорі невідомими для гравців. З метою стохастичної ідентифікації середовища, розв’язування гри здійснюється повторенням її ходів.

Нехай кожен гравець $i\in D$ здійснює у дискретні моменти часу $n=1,2,...$ незалежний вибір однієї з власних чистих стратегій $u_n^i=u^i\in U^i=(u^i(1),\ u^i(2),\ ...,\ u^i(N_i))$. Після завершення вибору стратегій гравців з локальних підмножин D_i $\forall i\in D$ гравці спостерігають власний випадковий поточний виграш $\xi_n^i(u_n^{D_i})$. Вважається, що послідовності випадкових величин $\{\xi_n^i\}$ незалежні $\forall u_n^{D_i} \in U^{D_i}$, $\forall i \in D$, $\forall n = 1, 2, ...$, а їхні математичні сподівання $M\{\xi_n^i(u^{D_i})\} = v^i(u^{D_i}) = const$ апріорі не відомі та мають обмежений другий момент $\sup_n M\{[\xi_n^i(u^{D_i})]^2\} = \sigma_i^2(u^{D_i}) < \infty$. Матриці математичних сподівань виграшів $\left[v^i(u^{D_i})\right]$ $\forall i \in D$ назвемо середовищем гри. Якщо $\forall u_n^{D_i} \in U^{D_i}$, $\forall i \in D$ $v(u^{D_i}) > 0$, то середовище є знакододатним, якщо $v(u^{D_i}) < 0$ – знаковід'ємним, інакше – загального виду.

Наступні моделі стохастичної гри будуються залежно від факторів ігрової взаємодії та надійності функціонування гравців. Фактор ігрової взаємодії визначає можливість обміну інформацією між гравцями про реалізовані стратегії або значення поточних виграшів. Фактор надійності функціонування визначає здатність гравців безвідмовно приймати рішення щодо реалізації поточних стратегій поведінки.

У грі без обміну інформацією гравці не повідомляють один одного про реалізовані стратегії та величину отриманого виграшу ξ_n^i. У грі з обміном інформацією кожен і-й гравець повідомляє гравців з локальної підмножини D_i про значення поточного виграшу ξ_n^i. У результаті і-й гравець отримає інформацію від множини гравців $\tilde{D}_i$, виграші яких визначаються його стратегіями. Згортка поточних виграшів $\zeta_n^i = \sum_{k \in \tilde{D}_i} \lambda^i(k)\xi_n^i$, де $\lambda^i = (\lambda^i(k) \mid \lambda^i(k) > 0 \ \forall k \in \tilde{D}_i, \ \sum_{k \in \tilde{D}_i} \lambda^i(k) = 1)$, є оцінкою стратегії, вибраної і-м гравцем у момент часу n.

У грі з відмовами гравці характеризуються імовірностями відмов $\eta^i \in [0,1)$. Нехай $\psi^i \in \{0,1\}$ – ознака участі і-го гравця у грі. Якщо $\psi^i = 0$, то гравець відмовляється від поточного ходу гри з імовірністю η^i, якщо $\psi^i = 1$ – бере участь у грі з імовірністю $1 - \eta^i$. На кожному кроці гри відбувається обмін ознаками поточних станів ψ_n^i між сусідніми гравцями з множин D_i $\forall i \in D$. Відмови гравців призводять до зміни складу множин D_i $\forall i \in D$, а значить, до зміни поточних виграшів. Введемо повні групи подій $\Psi_i = 2^{|D_i|}$, пов'язаних із відмовами гравців з множин D_i. Нехай $D_i(\omega) \subseteq D_i$ – підмножина гравців, які залишаються у грі для події $\omega \in \Psi_i$. Тоді поточні виграші i-го гравця дорівнюють $\xi_n^i = \chi\{\sum_{j \in D_i} \psi_n^j \geq \bar{\psi}\}\xi_n^i(u_n^{D_i(\omega)})$, де $\chi(\cdot) \in \{0,1\}$ – індикаторна функція події, $\bar{\psi} > 0$ – поріг гри, або мінімально допустима кількість гравців, що не відмовили [10].

Послідовності обраних варіантів $\{u_t^{D_i(\omega)} \mid t = \overline{1,n}\}$ оцінюються поточними середніми виграшами

$$\Phi_n^i(\{u_n^{D_i(\omega)}\}) = \frac{1}{n}\sum_{t=1}^{n} \xi_t^i \quad \forall i \in D. \tag{1}$$

Метою кожного гравця є максимізація функції середніх виграшів

$$\underline{\lim}_{n \to \infty} \Phi_n^i(\{u_n^{D_i(\omega)}\}) \to \max \ \ \forall i \in D. \tag{2}$$

Для гри без обміну інформацією пошук розв'язків задачі векторної оптимізації (2) здійснюється у множині точок рівноваги за Нешем

$$\forall i \in D \quad \underline{\lim}_{n \to \infty} \left[\Phi_n^i(\{u_n^{D_i(\omega)}\}) - \Phi_n^i(\{\hat{u}_n^{D_i(\omega)}\})\right] \geq 0. \tag{3}$$

Розв'язки гри з обміном інформацією знаходяться у множині оптимальності за Парето

$$\begin{cases} \forall i \in D \quad \underline{\lim}_{n \to \infty} \left[\Phi_n^i(\{u_n^{D_i(\omega)}\}) - \Phi_n^i(\{\tilde{u}_n^{D_i(\omega)}\})\right] \geq 0, \\ \exists i \in D \quad \underline{\lim}_{n \to \infty} \left[\Phi_n^i(\{u_n^{D_i(\omega)}\}) - \Phi_n^i(\{\tilde{u}_n^{D_i(\omega)}\})\right] > 0. \end{cases} \tag{4}$$

Нерівності (3) та (4) виконуються з імовірністю 1, а $u_n^{D_i(\omega)}, \tilde{u}_n^{D_i(\omega)} \in U^{D_i(\omega)}$; $\hat{u}_n^{D_i(\omega)} = u_n^{D_i(\omega)} \setminus u_n^i + \tilde{u}_n^i \in U^{D_i(\omega)}$; $u_n^i, \tilde{u}_n^i \in U^i$.

Для розв'язування стохастичної гри необхідно сформувати послідовності чистих стратегій $\{u_t^{D_i(\omega)} \mid \forall i \in D, t = \overline{1,n}\}$, які при $n \to \infty$ забезпечать виконання умов (3) або (4).

2. Методи розв'язування ігрової задачі

Формування послідовностей $\{u_t^{D_i(\omega)}\}$ з потрібними властивостями виконаємо випадково, на основі векторів змішаних стратегій p_n^i, елементи яких є умовними імовірностями вибору відповідних чистих стратегій, тобто $p_n^i(u_n^i) = P\{u_n^i \mid u_t^i, \xi_t^i (t = \overline{1, n-1})\}$ $\forall u_n^i \in U^i, \forall i \in D$.

Уточнення поточного розв'язку гри у змішаних стратегіях здійснюється гравцями за допомогою рекурентного методу [12]

$$p_{n+1}^i = \pi_{\varepsilon_{n+1}}^{N_i} \left\{ p_n^i - \gamma_n R(x_n^i, p_n^i, \xi_n^i) \right\}, \tag{5}$$

де γ_n – крок методу; $R(x_n^i, p_n^i, \xi_n^i)$ – вектор руху методу; $\pi_{\varepsilon_{n+1}}^{N_i}$ – проектор на одиничний ε-симплекс [12], необхідний для нормалізації p_n^i та нагромадження повної статистичної інформації про випадкове середовище.

Вектор $R(x_n^i, p_n^i, \xi_n^i)$ забезпечує у середньому переміщення у напрямку до оптимального розв'язку ігрової задачі. Визначення напрямку руху методу (5) виконаємо методом стохастичної апроксимації [13] на основі матричного формулювання асимптотично адекватної безкоаліційної ігрової задачі з функціями середніх виграшів гравців

$$V^i = \sum_{\omega \in \Psi_i} \prod_{j \in D_i(\omega)} \left\{ (1 - \eta^j) \chi(\psi^j = 1) + \eta^j \chi(\psi^j = 0) \right\} V^i(\omega),$$

де $V^i(\omega) = \sum_{u^{D_i(\omega)} \in U^{D_i(\omega)}} v^i(u^{D_i(\omega)}) \prod_{j \in D_i(\omega); u^j \in u^{D_i(\omega)}} p^j(u^j)$ – полілінійна функція середніх виграшів, визначена для однієї із ситуацій гри з відмовами.

Визначення парето-оптимальних розв'язків матричної гри здійснюється методом умовної максимізації системи локальних згорток функцій середніх виграшів $W^i = \sum_{k \in \tilde{D}_i} \lambda^i(k) V^k$ на опуклих одиничних симплексах S^{N_i} $\forall i \in D$. Для диференційованих на S^{N_i} функцій W^i оптимальні змішані стратегії знаходяться з умови доповняльної нежорсткості [14]:

$$\nabla_{p^i} W^i = W^i e^{N_i}, \; p^i \in S^{N_i}, \; \forall i \in D, \tag{6}$$

де $\nabla_{p^i} W^i$ – градієнт функції W^i; e^{N_i} – вектор, що складається з N_i одиниць.

Умова (6) визначає вирівнювальні за Нешем стратегії для коаліцій гравців $\tilde{D}_i$ $\forall i \in D$. У загальному випадку ці стратегії є локальними (ε-оптимальними) розв'язками за Парето базової безкоаліційної гри. Вирівнювальні за Нешем стратегії безкоаліційної гри отримуються з (6) при $W^i = V^i$ $\forall i \in D$.

Виконавши стохастичну апроксимацію (6), отримаємо марківські рекурентні методи:

$$p_{n+1}^i = \pi_{\varepsilon_{n+1}}^{N_i} \left\{ p_n^i - \gamma_n \zeta_n^i \left(g e^{N_i} - \frac{e(u_n^i)}{e^{\mathrm{T}}(u_n^i) p_n^i} \right) \right\}, \tag{7}$$

де $\pi_\varepsilon^{N_i}$ – оператор проектування на ε-симплекс $S_\varepsilon^{N_i} \subseteq S^{N_i}$; $\gamma_n \ge 0$ – параметр, що регулює величину кроку методу; $g \in \{0,1\}$; $e(u_n^i)$ – одиничний вектор-індикатор вибору варіанта u_n^i. При $g = 0$ з (7) отримаємо градієнтний метод [12].

Методи (7) використовуються для пошуку розв'язків гри у повністю змішаних стратегій. Розв'язування гри у неповністю змішаних стратегіях, розміщених на межі одиничного симплексу, здійснюється на основі покомпонентного зважування векторів умови (6):

$$\Upsilon^i = diag(p_n^i)(e^{N_i} W_n^i - \nabla_{p^i} W_n^i), \quad (8)$$

де $\Upsilon^i \in R^{N_i}$; $diag(p_n^i)$ – квадратна діагональна матриця порядку N_i, складена з елементів вектора p_n^i.

На основі стохастичної апроксимації виразу (8), отримаємо рекурентний метод:

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

При $\gamma_n \zeta_n^i \in [0,1]$ з (7) отримаємо безпроекційний метод. Крім цього, з (7) та (9) отримаємо методи без обміну інформацією при $\zeta_n^i = \xi_n^i$, з обміном інформацією при $\zeta_n^i = \sum_{k \in \tilde{D}_i} \lambda^i(k) \xi_n^k$, без відмов гравців при $\xi_n^i = \xi_n^i(u_n^{D_i})$, з відмовами гравців при $\xi_n^i = \chi\{\sum_{j \in D_i} \psi_n^j > \bar{\psi}\} \xi_n^i(u_n^{D_i(\omega)})$, регуляризовані методи при $\zeta_n^i = \xi_n^i + \delta_n e^{\mathrm{T}}(u_n^i) p_n^i$, $\delta_n > 0$ [16].

Теоретичні оцінки умов збіжності рекурентних методів (7) та (9) до асимптотично оптимальних розв'язків у знаковизначених середовищах та середовищах загального виду можна отримати на основі верхніх оцінок умовного математичного сподівання поточної похибки $\Delta_n = \sum_{i \in D} \left\| \Upsilon^i \right\|^2$ виконання умови доповняльної нежорсткості при фіксованій передісторії подій та результатів теорем про рекурентні числові нерівності [15].

Оцінювання асимптотичного порядку швидкості збіжності виконано для послідовностей $\gamma_n = \gamma n^{-\alpha}$; $\varepsilon_n = \varepsilon n^{-\beta}$; $\gamma, \varepsilon, \alpha, \beta > 0$ методом моментів Чжуна [12, 13]:

$$\overline{\lim_{n \to \infty}} n^{\theta} M\{\Delta_n\} \le \vartheta, \quad (10)$$

де θ – параметр порядку, ϑ – величина швидкості збіжності. Більшому θ та меншому ϑ відповідає більша швидкість збіжності ігрового методу.

Планування комп'ютерного експерименту

Для комп'ютерного моделювання гри необхідно задати: 1) кількість гравців та структуру локальних зв'язків між ними; 2) модель середовища та його параметри; 3) модель ігрової взаємодії гравців; 4) цільові умови асимптотичної оптимальності; 5) рекурентний алгоритм зміни векторів змішаних стратегій; 6) початкові значення змішаних стратегій та параметрів алгоритму; 7) довжину вибірки, на якій відбувається навчання ігрового алгоритму.

Структура локальних зв'язків гри задається бінарною матрицею суміжностей $Q = [q]_{L \times L}$ для фіксованої кількості гравців $L = |D|$. Одиничні елементи i-го рядка цієї матриці визначають множину гравців D_i, в базисі векторів змішаних стратегій яких визначаються виграші i-го гравця. Відповідно, одиничні елементи i-го стовпця матриці суміжностей визначають множину гравців $\tilde{D}_i$, виграші яких залежать від стратегій i-го гравця.

Кожен гравець $i \in D$ характеризується ймовірністю відмови η^i від реалізації поточного ходу. Відмова i-го гравця моделюється виконанням умови $\omega \le \eta^i$, де $\omega \in [0,1)$ – дійсне випадкове число з рівномірним законом розподілу. Результати відмов гравців записуються у L-елементний вектор ψ поточного стану активності гравців:

$$\psi[i] = \begin{cases} 1, & \omega > \eta^i, \\ 0, & \omega \le \eta^i. \end{cases}$$

Вважається, що відмови гравців діють тимчасово до закінчення поточного кванту часу, після чого знову визначається стан активностей гравців.

Відмови гравців призводять до зміни структури гри. Оскільки і-й гравець представлений у грі власним вектором змішаних стратегій p_n^i, то при його відмові цей вектор вилучається з базисів гравців множини $\tilde{D}_i$, виграші яких залежать від стратегії і-го гравця. Інакше кажучи, у разі відмови і-го гравця всі елементи і-го стовпця матриці суміжностей приймають нульові значення.

З врахуванням відмов гравців матриця суміжностей перетворюється за правилом:

$$\tilde{Q} = Q * diag(\psi),$$

де $diag(\psi)$ – діагональна квадратна матриця порядку L, складена з елементів вектора ψ ознак активності гравців.

Після коректування матриці суміжностей відбувається сканування її рядків з метою виявлення гравців, які залишаються у грі. Ознакою участі і-го гравця у грі є виконання однієї з таких умов: 1) $\sum_{j=1}^{L} q[i,j] \geq 1$, якщо допускається "гра з природою"; 2) $\sum_{j=1}^{L} q[i,j] \geq 2$, якщо "гра з природою" не допускається.

У результаті відмов гравців тимчасово допускається можливість порушення зв'язності структури гри.

Ознаки участі гравців у грі записуються у бінарний вектор $\mathrm{M} = (\mu[i] \,|\, \mu[i] \in \{0,1\}, i = \overline{1,L})$. Якщо $\mu\ [i] = 1$, то гравець бере участь у грі, інакше – утримується від реалізації поточного ходу.

Для гравців, які беруть участь у грі, обчислюються поточні значення виграшів $\xi_n^i(u^{D_i}, \omega)$, які визначаються у скорегованих базисах множини гравців $D_i = \{ j \,|\, j = \overline{1,L}; q[i,j] = 1; \mu\ [i] = 1 \} \neq \varnothing$. В загальному випадку закон розподілу випадкових величин виграшів $\xi_n^i(u^{D_i}, \omega)$ апріорі не відомий, але для моделювання ігрових алгоритмів його необхідно задати як базову характеристику моделі середовища.

Модель середовища визначається структурою локальних множин $\{D_i\} \forall i \in D$ та законом розподілу поточних виграшів. Досліджуються алгоритми з дискретними бінарними та неперервними обмеженими виграшами.

Бінарні виграші $\xi_n^i(u^{D_i}, \omega) \in \{0,1\}$ формувалися так:

$$\xi_n^i(u^{D_i}, \omega) = \begin{cases} 0, & \omega > v^i(u^{D_i}) \\ 1, & \omega \leq v^i(u^{D_i}) \end{cases},$$

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

Нормально-розподілені випадкові величини (за законом Гаусса) знаходилися як сума дванадцяти рівномірно розподілених на відрізку [0,1] величин:

$$\xi_n^i(u^{D_i}, \omega) = v^i(u^{D_i}) + \sqrt{d^i(u^{D_i})} \left(\sum_{j=1}^{12} \omega_j - 6 \right).$$

Досліджено роботу ігрових алгоритмів у знакододатному середовищі та середовищі загального виду. Матриці математичних сподівань $\left[v^i(u^{D_i}) \right]_{\forall u^{D_i} \in U^{D_i}}$ та дисперсій $\left[d^i(u^{D_i}) \right]_{\forall u^{D_i} \in U^{D_i}}$ виграшів для всіх гравців формуються за допомогою вбудованого генератора випадкових величин, розподілених за рівномірним законом. Для знакододатного середовища математичні сподівання виграшів приймають значення з відрізка $[\delta,\ 1-\delta]$, де $0 < \delta << 1$, а для середовища загального виду – з відрізка [-1,1].

Відносне порівняння ефективності роботи різних ігрових алгоритмів виконано у тотожних середовищах. Для цього передбачено запам'ятовування згенерованих параметрів середовища у файлах даних на зовнішньому носії інформації та відновлення параметрів середовища з файлів.

Вибір чистих стратегій u_n^i гравців здійснюється на основі векторів змішаних стратегій p_n^i. Номер k поточної чистої стратегії $u_n^i = u^i \in U^i$ визначається з умови:

$$k = \min_{j=1}^{N_i}\left(j \left| \sum_{l=1}^{j} p_n^i(l) \geq \omega \right. \right).$$

Після вибору чистих стратегій всіма гравцями з множини D_i $\forall i \in D$ визначаються математичні сподівання $v^i(u^{D_i})$ та дисперсії $d^i(u^{D_i})$ випадкових величин $\xi_n^i(u^{D_i})$, які використовуються як параметри одного з заданих законів розподілу.

Спосіб використання отриманих виграшів $\xi_n^i(u^{D_i})$ залежить від прийнятої моделі ігрової взаємодії гравців.

В іграх без обміну інформацією поточні виграші використовує окремо кожен гравець для перерахунку власних векторів змішаних стратегій згідно з рекурентним методом (7) або (9).

В іграх з обміном інформацією відбувається обмін отриманими виграшами в межах локальних коаліцій гравців D_i, стратегії яких визначають виграші і-го гравця $\forall i \in D$. В результаті обміну і-й гравець отримає поточні виграші від всіх гравців, виграші яких залежать від стратегій і-го гравця. Множина таких гравців утворюють локальну коаліцію $\tilde{D}_i$. Зважені виграші

$$\xi_n^i = \sum_{k \in \tilde{D}_i} \lambda^i(k) \xi_n^k$$

використовує і-й гравець для модифікації власного вектора змішаних стратегій.

Досліджено збіжність ігрових алгоритмів до однієї з вирівнювальних стратегій, для яких виконується умова доповняльної нежорсткості (8).

Значення параметрів γ_n та ε_n проекційних алгоритмів змінюються у часі $n = 1, 2, ...$ за законами: $\gamma_n = \dfrac{\gamma}{n^\alpha}$, $\varepsilon_n = \dfrac{\varepsilon}{n^\beta}$, $\gamma > 0$, $\varepsilon \in (0, \min_{i \in D} N_i^{-1})$, $\alpha > 0$, $\beta > 0$.

Початкові значення параметрів рекурентних алгоритмів вибираються такими, щоб задовольнити відповідні умови збіжності алгоритмів у середньоквадратичному [9, 10].

Якщо спеціально не вказано, то початкові значення елементів векторів змішаних стратегій приймають однаковими

$$p_n^i(j) = 1/N_i, \; j = \overline{1, N_i}, \; \forall i \in D,$$

що моделює ситуацію невизначеності у початковий момент часу, коли гравці не мають інформації про середовище прийняття рішень і будь-який з допустимих варіантів рішення є прийнятним для реалізації.

Для нормування векторів змішаних стратегій p_n^i виконується проектування їх рекурентних перетворень на ε-симплекс $S_{\varepsilon_n}^{N_i}$, яке зводиться до ітераційного алгоритму проектування вектора на одиничну гіперплощину з наступним зануленням його від'ємних компонентів [12].

Обчислення критеріїв. Для кожної моделі гри визначається зміна у часі випадкових процесів:

$$\Delta_n = \sum_{i \in D} \left\| p_n^i - \tilde{p}_n^i \right\|^2 \text{ та } \overline{\Delta}_n = \frac{1}{n} \sum_{t=1}^{n} \Delta_t .$$

Для обчислення $\tilde{p}_n^i = diag(p_n^i) \nabla_{p^i} V_n^i \big/ V_n^i$ виконувалося моделювання повної групи подій, пов'язаної з відмовами гравців. Повна група подій генерувалася послідовним перебором всіх можливих значень L-розрядного двійкового числа $s = (s_i \mid s_i \in \{0,1\}, i = \overline{1, L})$, представленого у вигляді одновимірного масиву. Одиничне значення і-го розряду двійкового числа сигналізує про участь і-го гравця у грі, а нульове значення – про його відмову.

Прийнятність згенерованої ситуації для і-го гравця визначається значенням імовірності $p_{V}^{i}(s)=\prod_{m\in D}\left\{(1-s_m)[\eta^m+(1-\eta^m)(1-p_{im})]+s_m(1-\eta^m)p_{im}\right\}$. Якщо $p_{V}^{i}(s)>0$, то згенерована ситуація є допустимою для і-го гравця.

Додатковим обмеженням на прийнятність згенерованої ситуації для і-го гравця є мінімальна кількість гравців, які можуть брати участь у грі. Якщо допускаються варіанти гри з природою, то необхідно, щоб виконувалась умова $\sum_{m\in D}s_m>0$. Якщо ж варіанти гри з природою не допускаються, то $\sum_{m\in D}s_m>1$.

Одиничні елементи вектора s задають гравців, стратегії яких визначають виграші *і*-го гравця. В базисі цих стратегій визначається значення функції середніх виграшів $V_n^i(s)$ та її градієнта $\nabla_{p^i}V_n^i(s)$.

Для цього методом послідовного перебору генерувалися всі можливі комбінації спільних стратегій гравців, які визначають виграші і-го гравця в ситуації s. Функції середніх виграшів гравців $V_n^i(s)$ для ситуації s визначалися згідно з (1). Для повної групи подій, пов'язаних з відмовами гравців, значення функцій середніх виграшів обчислювалися так:

$$V_n^i=\sum_s p_V^i(s)V_n^i(s).$$

Довжина досліджуваної вибірки становить 10 тис. кроків. На основі оцінки швидкості збіжності (10) поведінка процесу Δ_n у часі апроксимована залежністю $\Delta_n=\vartheta/n^{\theta}$, де $\vartheta>0$; $\theta\in(0,1]$; $n=1,2,...$. Після логарифмування отримується лінійне співвідношення

$$\lg\Delta_n=\lg\vartheta-\theta\lg n.$$

З врахуванням цього будується залежність $\lg\Delta_n=f(\lg n)$. Тоді параметр $\theta=\lg\Delta_n/\lg n$ вказує на порядок швидкості збіжності досліджуваного алгоритму.

Для визначення порядку швидкості збіжності θ виконана апроксимація випадкового процесу $\lg\Delta_n$ лінійною залежністю на відрізку $\lg n\in[3,4]$ з кроком $\delta=0.1$ за методом найменших квадратів.

Для згладжування випадкової складової швидкості збіжності та виділення порядку цієї швидкості виконується усереднення випадкового процесу Δ_n по $K_{\exp}$ реалізаціях:

$$\overline{\overline{\Delta}}_n=\frac{1}{K_{\exp}}\sum_{j=1}^{K_{\exp}}\Delta_{n,j}.$$

Задачі експерименту. Під час експерименту необхідно дослідити:

1) вплив початкових параметрів алгоритмів на порядок та величину швидкості збіжності;
2) вплив законів розподілу випадкових величин на швидкість збіжності ігрових алгоритмів;
3) вплив дисперсії випадкових величин виграшів на величину швидкості збіжності;
4) вплив обміну інформацією на поведінку ігрових алгоритмів;
5) вплив кількості гравців та чистих стратегій на величину швидкості збіжності ігрових алгоритмів;
6) стійкість досягнутих ефективних рішень для гри з локальними зв'язками.

Простір експерименту визначимо у базисі таких параметрів: $\Pi=(Z,v,d,\gamma,\varepsilon,\alpha,\beta)$, де Z – закон розподілу виграшів, v – значення математичних сподівань виграшів, d – значення дисперсій виграшів, $\gamma,\varepsilon,\alpha,\beta$ – параметри алгоритму.

Рекурентні ігрові алгоритми. Шляхом програмного моделювання на комп'ютері досліджено проекційний метод (9) мінімізації похибки доповняльної нежорсткості у знакододатному середовищі з обміном або без обміну нформацією.

Графічна схема ігрового алгоритму зображена на рис. 1. Спочатку задаються вхідні дані – параметри середовища, параметри рекурентного методу та структури мультиагентної системи. Формуються вимоги до вигляду результатів моделювання. Виконується початкова ініціалізація системи – заповнення матриць математичних сподівань виграшів, встановлення початкових значень векторів змішаних стратегій, обчислення початкових значень параметрів рекурентного методу. Далі моделюються відмови агентів та встановлюються поточні зв'язки між ними. На основі векторів змішаних стратегій здійснюється випадковий вибір чистих стратегій. За значеннями чистих стратегій відбувається читання математичних сподівань та дисперсій виграшів, за якими здійснюється генерування поточних випадкових виграшів гравців. Обчислюються поточні значення параметрів ігрового методу і виконується перерахунок векторів змішаних стратегій з їх наступним нормуванням шляхом проектування на одиничний симплекс. Обчислюються та виводяться поточні характеристик гри – функції середніх виграшів, похибка доповняльної нежорсткості тощо. Перевіряється умова завершення гри – за значенням досягнутої точності або за кількістю кроків гри. Якщо умова не виконана, то реалізується наступний крок гри, починаючи від моделювання відмов гравців.

Рис.1. Узагальнена графічна схема ігрового алгоритму

Програмні інструментальні засоби моделювання адаптивних ігрових алгоритмів

Дослідження поведінки рекурентних ігрових методів у стохастичних середовищах здійснено за допомогою розроблених нами програмних інструментальних засобів моделювання стохастичних ігор.

Назва та призначення програми. Програма "Система моделювання стохастичної гри агентів" призначена для моделювання рекурентних ігрових методів в умовах невизначеності.

Мови програмування. Програмний комплекс написано мовою C++ з використанням бібліотеки wxWidgets.

Вхідні дані. Вхідні дані, необхідні для моделювання ігрових методів, задаються за допомогою елементів діалогу інтерактивного середовища програми. Для моделювання необхідно задати: кількість гравців, кількість стратегій гравців, структуру зв'язків між гравцями, закон розподілу виграшів, дисперсії виграшів, вид ігрового методу та початкові значення його параметрів, кількість кроків моделювання, кількість експериментів.

Вихідні дані. Результати моделювання задаються у вигляді графіків, що демонструють зміну у часі функції середніх виграшів та похибки виконання умови доповняльної нежорсткості, а також евклідових норм векторів змішаних стратегій.

Технологічний процес прийняття рішень

Програму побудовано на основі класів, зображених на рис. 2 у вигляді діаграми у нотації UML.

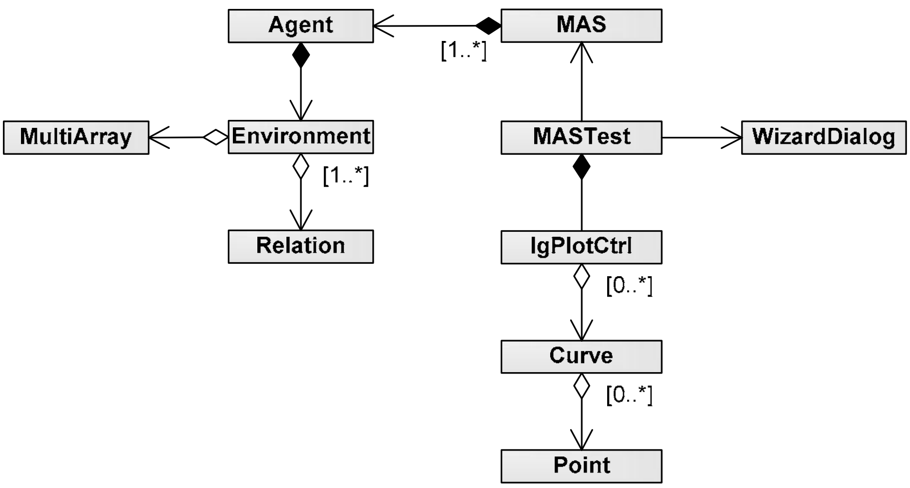

Рис. 2. Діаграма класів програми моделювання стохастичних ігор

Основним класом виступає агент. Клас агента містить вектор чистих та змішаних стратегій та функції ініціалізації (конструктор), вибору стратегії, імітації дії (визначення виграшу, навчання). Кожен агент тісно пов'язаний з класом середовища (Environment), який вміщає класи MultiArray та Relation.

Клас Environment призначений для моделювання середовища агента, він містить структури зв'язків агента, що задаються вектором екземплярів класу Relation, а також багатовимірний масив виграшів, що представлений екземпляром класом MultiArray.

Клас MultiArray призначений для виконання дій над багатовимірними масивами даних:

1) доступу до елементів за допомогою масиву індексів;
2) динамічної зміни кількості вимірів та розмірності масивів;
3) послідовного заповнення масиву даними.

Клас Relation описує зв'язок з іншими агентами, містить ідентифікатор та поточну стратегію агента, а також функції для отримання та встановлення їх значень.

Клас MAS – клас мультиагентної системи відповідає за час життя агентів, ініціалізацію та керування процесом моделювання мультиагентної системи, отримання вихідних даних.

Клас WizardDialog описує вікно діалогового вікна, за допомогою якого користувач задає вхідні дані системи.

Клас Point – клас точки (x,y) для формування графіків функцій.

Клас Curve – клас кривої, що складається з набору точок для запам'ятовування та виведення експерементальних даних.

Клас lgPlotCtrl – клас елемента керування для зображення графіків у логарифмічному масштабі, які представлені вектором типу Curve. Параметри зображення задаються.

Клас MASTest – клас самої програми, відповідає за ініціалізацію та реєстрацію програми в системі, побудову вікон та опрацювання повідомлень головного фреймового вікна. Цей клас отримує вхідні дані від екземпляру класу WizardDialog, передає їх класу MAS. Отримані вихідні дані класу MAS передаються класу lgPlotCtrl для виведення на екран.

Вимоги до технічних та програмних засобів

Для виконання програми використовується комп'ютер стандартної комлектації, наприклад з x86-сумісним CPU, пристроями введення-виведення, блоком оперативної пам'яті і жорстким диском.

Програма моделювання стохастичних ігор може бути виконана під керування однієї з операційних систем: Windows 95/98/ME, Windows NT/2K/XP, Linux/Unix з GTK+ toolkit, X11 або Motif, MacOS.

Необхідно відзначити, що час обчислень залежить від швидкодії комп'ютера, тому для ефективнішої роботи необхідно використовувати комп'ютер, побудований на базі сучасних технологічних рішень, наприклад 2–4 ядерний процесор з тактовою частотою 2–3 GHz та 1–2 Gb оперативної пам'яті.

Інструкція для користування програмою

Після запуску програми на екран виводиться головне вікно програми, у якому розташована робоча область для виведення графіків, а також текстовий елемент (нижня частина вікна) призначений для відображення додаткової інформації про агентів.

Для початку роботи з програмою необхідно вибрати пункт меню Модель $\rightarrow$ Нова Модель, в результаті чого буде виведене діалогове вікно, призначене для введення вхідних даних.

На першій сторінці цього діалогового вікна пропонується задати закон розподілу випадкових виграшів, математичне сподівання та дисперсію виграшів. Далі задаємо алгоритм навчання і його параметри.

Наступне вікно задає стуктуру мультиагентної системи (рис. 3). Необхідно задати кількість агентів, їх стратегії, початкові значення вектора змішаних стратегій, параметри відмов, а також матрицю суміжностей графа мультиагентної системи (взаємозв'язки між агентами).

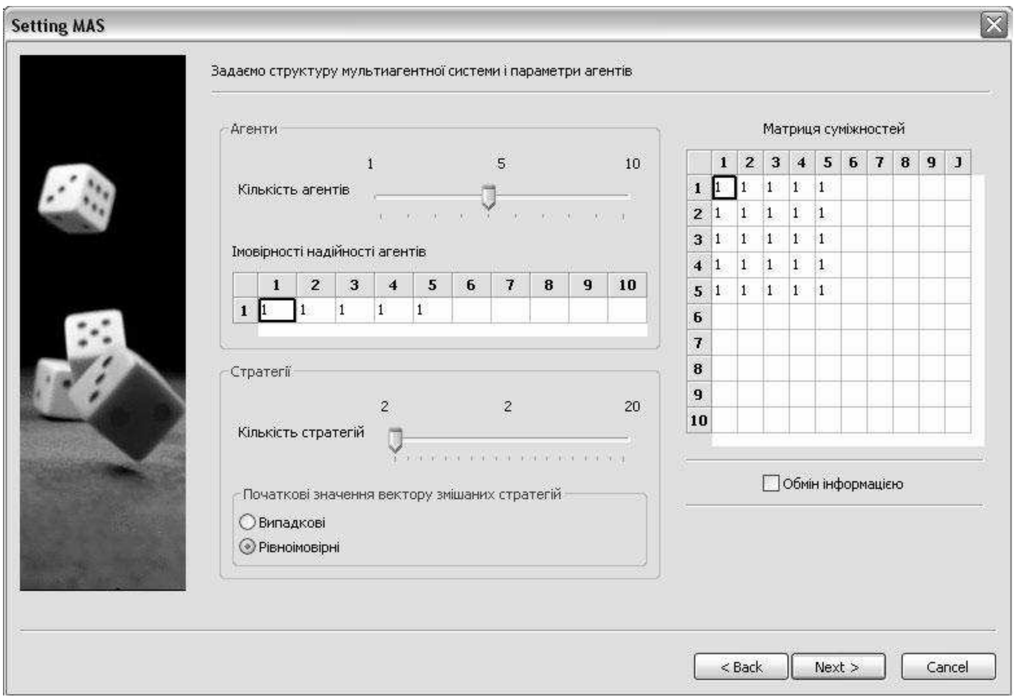

Рис. 3. Формування структури мультиагентної системи

У зображеному на рис. 4 вікні задаються вимоги до результатів – кількість кроків моделювання, види графіків та діапазони значень результатів.

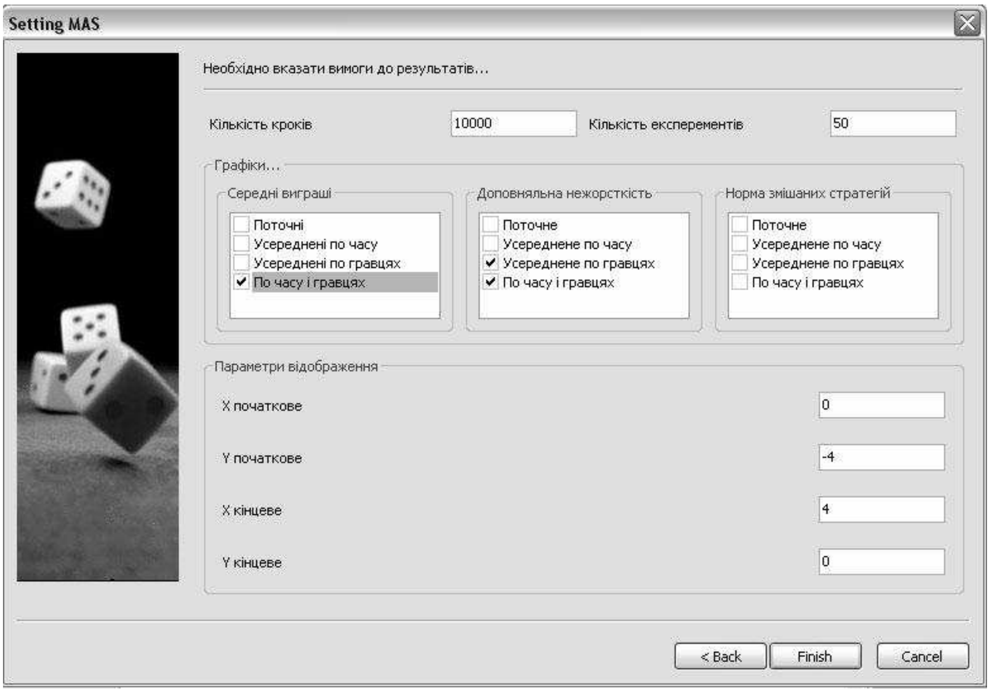

Рис. 4. Формування вимог до результатів моделювання

Після натиснення кнопки Finish розпочинається моделювання і виведення результатів. На рис. 5 зображено графіки функції середніх виграшів та похибки доповняльної нежорсткості, отримані для методу (9) з параметрами $\gamma_0 = 1$, $\beta = 2$, $\alpha = 0.4$. Середовище моделювання задається математичними сподіваннями виграшів $v^i(u^{D_i})$, які приймають значення з діапазону $[0.1; 0.9]$. Графік 1 відповідає функції середніх виграшів, графіки 2 та 3 – усередненій у часі та поточній похибці доповняльної нежорсткості відповідно. Результати подано для бінарних поточних виграшів $\xi_n^i(u^{D_i}) \in \{0,1\}$.

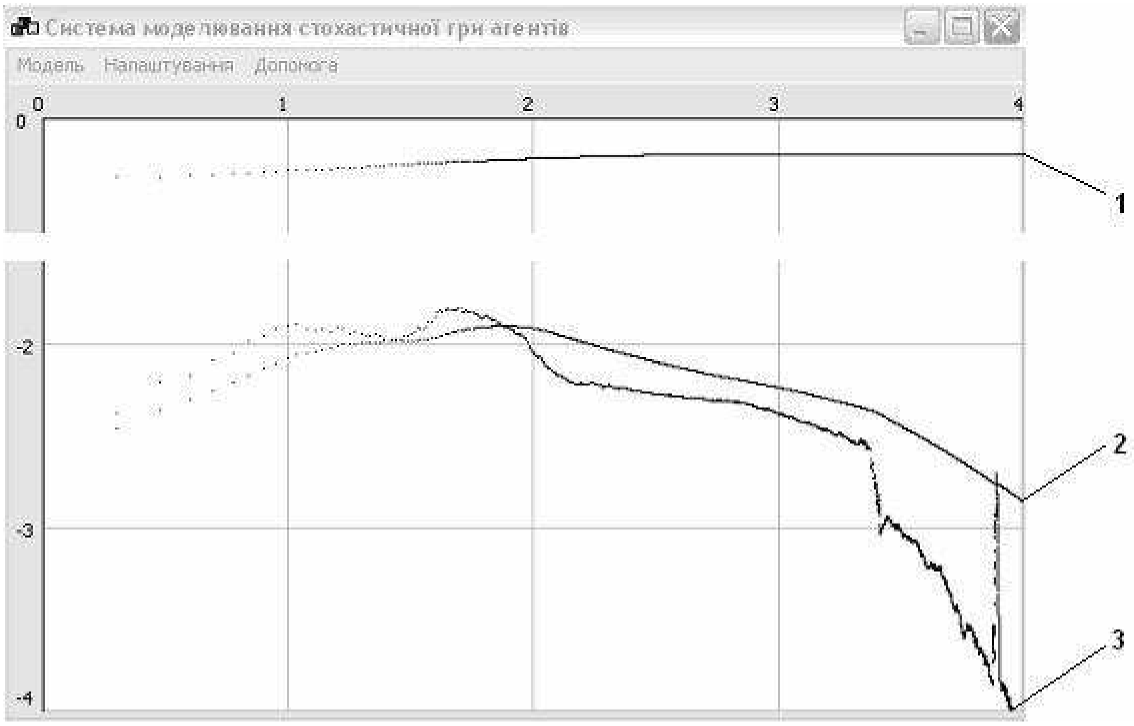

Рис. 5. Вікно результатів

Для завершення роботи необхідно закрити головне вікно програми. Для цього потрібно вибрати пункт меню Модель $\rightarrow$ Вихід або стандартну кнопку закривання вікна, позначену символом ‘$\times$’.

Висновки

У роботі розроблено програмні інструментальні засоби моделювання МАС на основі моделей теорії стохастичних ігор. Такі інструментальні засоби здійснюють імітаційне моделювання ігрових ситуацій, які можуть мати місце в МАС, що дає можливість дослідити динаміку середніх виграшів у часі, співвідношення параметрів, які впливають на збіжність ігрового методу до оптимальних розв’язків за Нешем або Парето, визначити параметри швидкості збіжності.

Результати комп’ютерного моделювання ігрових методів на вибірках великої довжини підтверджують працездатність адаптивних ігрових методів щодо максимізації середніх виграшів гравців у ході досягнення колективних компромісних розв’язків. Достовірність результатів комп’ютерного моделювання ігрових методів забезпечено належною початковою підготовкою та плануванням експерименту і підтверджено повторюваністю розрахункових величин та їх загальною відповідністю теоретичним результатам.

Порядок швидкості збіжності рекурентних ігрових методів, в основному, визначається способом формування регульованих параметрів γ_n та ε_n. Ігрові методи (7) та (9) здійснюють формування послідовності чистих стратегій, необхідної для досягнення асимптотичної мети, методом проб та помилок, що загалом пояснює їх невисоку (степеневу) швидкість збіжності.

Теоретичні результати щодо умов збіжності досліджених ігрових методів є інваріантними до початкових значень векторів змішаних стратегій на одиничному симплексі та законів розподілу поточних виграшів (при однакових значеннях математичного сподівання та дисперсії).

Зростання дисперсії виграшів гравців призводить до сповільнення швидкості збіжності ігрових методів.

З отриманих результатів очевидно, що в умовах невизначеності метод (9) забезпечує вищий порядок швидкості збіжності до множини оптимальних розв’язків, ніж метод (7), як у знаковизначених середовищах, так і у середовищах загального виду. Встановлено, що відмови гравців призводять до ε-оптимальних розв’язків ігрової задачі та до сповільнення швидкості збіжності. При зменшенні порогу гри $\bar{\psi}$ швидкість збіжності ігрових методів зростає.

Методи без обміну та з обміном інформацією мають однаковий асимптотичний порядок швидкості збіжності θ в афінно-еквівалентних середовищах. Обмін інформацією між гравцями призводить до зростання величини швидкості збіжності (зменшення значення ϑ). Регуляризовані методи забезпечують стійкість розв’язків ігрової задачі у змішаних стратегіях.

Співвідношення параметрів α та β визначає порядок швидкості збіжності ігрового алгоритму. Загалом, зростання параметра α призводить до сповільнення, а зростання параметра β – до зростання швидкості збіжності, хоча такі залежності не є лінійними.

Порядки усередненої по реалізаціях похибки доповняльної нежорсткості є більшими від порядків усередненої у часі та по реалізаціях похибки.

Зростання кількості гравців та кількості чистих стратегій призводить до сповільнення швидкості збіжності ігрової задачі. Зростання кількості гравців справляє більший вплив на зменшення швидкості збіжності ігрового алгоритму, ніж зростання кількості чистих стратегій гравців. Останнє пояснюється тим, що при зростанні кількості гравців кількість комбінованих варіантів гри зростає за показниковим, а при зростанні кількості стратегій – за степеневим законом.

Проаналізувавши отримані результати, можна стверджувати, що модель стохастичної гри є ефективним інструментом дослідження МАС і такий підхід заслуговує на подальші наукові пошуки та дослідження.

1. Тарасов В.Б. Агенты, многоагентные системы, виртуальные сообщества: стратегическое направление в информатике и искусственном интеллекте [Текст] / В.Б. Тарасов // Новости искусственного интеллекта. – 1998. – № 2. – С. 5-63. 2. Wooldridge M. Agent theories, architectures and languages: a survey [Текст] / M. Wooldridge, N. Jennings // Intelligent Agents. Proc. of ECAI-94 Workshop on Agent Theories, Architectures and Languages (Amsterdam, The Netherlands, August 8–9, 1994) / Ed. by M.Wooldridge, N.Jennings. – Amsterdam: Springer Verlag, 1995. – P. 1–22. 3. Georgeff M.P. BDI Agents: From Theory to Practice [Текст] / M.P. Georgeff, A.S. Rao // Procedings of the First International Conference on Multi-Agent Systems (ICMAS'95) / Ed. by V.Lesser). – AAAI Press / The MIT Press. – P.312–319. 4. Wooldridge M. An Introduction to Multiagent Systems [Текст] / M. Wooldridge. – John Wiley & Sons (Chichester, England), 2002. – 366 pp. 5. Weiss G. Adaptation and Learning in Multiagent Systems [Текст] / Gerhard Weiss, Sandip Sen, editors. – Berlin: Springer Verlag, 1996. – 585 pp. 6. Stone P. Layered Learning in Multiagent Systems [Текст] / P. Stone. – MIT Press, 2000. – 300 pp. 7. Трухаев Р.И. Модели принятия решений в условиях неопределенности / Р.И. Трухаев. – М.: Наука, 1981. – 257 с. 8. Fudenberg D. The Theory of Learning in Games [Текст] / D. Fudenberg, D. K. Levine. – Cambridge, MA: MIT Press, 1998. – 276 pp. 9. Кравець П.О. Ігрова задача взаємодії елементів мультиагентних систем [Текст] / П.О. Кравець // Комп'ютерні науки та інформаційні технології / Вісник Нац. ун-ту "Львівська політехніка". – 2006. – № 565. – С. 140 – 149. 10. Кравець П.О. Ігрова модель мультиагентної системи з відмовами [Текст] / П.О. Кравець // Комп'ютерні системи та мережі / Вісник Нац. ун-ту "Львівська політехніка". – 2007. – № 603. – С. 77 – 86. 11. Цыпкин, Я.З Рекуррентные алгоритмы оптимизации в условиях неопределенности [Текст] / Я.З., Цыпкин А.С. Позняк // Итоги науки и техники. Сер. Техническая кибернетика. – 1989. – Т. 16. – С. 3 – 70. 12. Назин А.В. Адаптивный выбор вариантов: Рекуррентные алгоритмы [Текст] / А.В. Назин, А.С. Позняк. – М.: Наука, 1986. – 288 с. 13. Вазан М. Стохастическая аппроксимация [Текст] / М. Вазан. – М.: Мир, 1972. – 295 с. 14. Воробьев Н.Н. Основы теории игр: Бескоалиционные игры [Текст] / Н.Н. Воробьев. – М.: Наука, 1984. – 495 с. 15. Невельсон М.Б. Стохастическая оптимизация и рекуррентное оценивание [Текст] / М.Б. Невельсон, Р.З. Хасьминский. – М.: Наука, 1972. – 304 с. 16. Кравець П.О. Регуляризований ігровий метод керування випадковими процесами в умовах невизначеності [Текст] / П.О. Кравець // Комп'ютерна інженерія та інформаційні технології / Вісник Нац. ун-ту "Львівська політехніка". – 2002. – № 468. – С. 101 – 109.