УДК 004.852, 004.942

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

ІГРОВА МОДЕЛЬ МУЛЬТИАГЕНТНОЇ СИСТЕМИ З ВІДМОВАМИ

Побудовано стохастичну ігрову модель мультиагентної системи з відмовами. На основі методу стохастичної апроксимації умови доповняльної нежорсткості розроблено рекурентний метод для розв'язування стохастичної гри. Визначено умови збіжності ігрового методу до вирівнювальних стратегій. Досліджено вплив відмов агентів на збіжність ігрового методу в умовах невизначеності.

The stochastic game model of multiagent system with failures is constructed. The recurrence method for the solution of stochastic game by a stochastic approximation method of a complementary slackness condition is developed. The conditions of a game method convergence to equilizing strategy are determined. The influence of failures of the agents on convergence of a game method in conditions of uncertainty is investigated.

Вступ. Сучасні дослідження розподілених систем різної природи (біологічних, соціальних, економічних, комп'ютерних тощо) ґрунтуються на моделях мультиагентних систем (МАС) [1]. Ключовим питанням моделювання МАС є з'ясування того, як досягається системна функціональність множини їх складових елементів. Під цим розуміють таку організацію МАС, яка забезпечує результат роботи, який не може бути зведений до суми результатів роботи її окремих елементів. Завдяки системній функціональності, як правило, досягається більша ефективність роботи колективу агентів порівняно з роботою множини автономних агентів. Системна функціональність може бути досягнута за рахунок переговорів, укладання договорів, компромісів, співпраці та координації дій агентів [2].

В умовах апріорної невизначеності забезпечення системної функціональності здійснюється на основі самонавчання та адаптації до середовища і до інших агентів. Робота МАС в умовах невизначеності породжує багатофакторні динамічні процеси, оптимізація яких є складною науково-практичною проблемою [1 – 3].

Враховуючи колективний характер організації МАС, продуктивним напрямком їх дослідження є застосування моделей та методів стохастичних ігор [4, 5]. Тоді системну функціональність МАС можена вивчати як спосіб досягнення компромісних розв'язків гри, для яких виконується умова рівноваги за Нешем, оптимальності за Парето та ін. [6]. Рівновага за Нешем визначає оптимальні приватні інтереси кожного гравця, які можна висловити твердженням: "не існує індивідуальної стратегії, яка дозволить покращити індивідуальний виграш, якщо решта гравців притримується стану рівноваги". Оптимальний за Парето розв'язок відображає колективні інтереси гравців: "не існує стратегії, яка дозволить підвищити виграші одночасно для всіх гравців".

Дослідження адаптивних властивостей стохастичних ігор в умовах невизначеності, методів та умов їх збіжності до оптимальних станів колективної рівноваги виконано у працях багатьох дослідників цієї проблеми, наприклад [4, 5, 7 – 9]. Основна увага приділяється побудові математичних моделей та методів стохастичних ігор, синтезу адаптивних ігрових методів, визначенню умов їх збіжності в ситуаціях без обміну та з обміном інформацією між гравцями.

Для практичних застосувань МАС важливим є досягнення стійких станів оптимальної рівноваги. Одним із факторів, який послаблює стійкість системи, є відмови агентів [10, 11]. Сьогодні результати досліджень ігрових моделей МАС з відмовами є недостатньо висвітленими у літературних джерелах.

Метою цієї роботи є вивчення впливу відмов агентів на досягнення стану їх колективної рівноваги та на швидкість збіжності ігрових методів.

1. Математична модель стохастичної гри з відмовами

Розглянемо модель функціонування МАС, яка складається з множини $D \neq \varnothing$ агентів, які розігрують повторювальну гру по максимізації середніх виграшів у розподіленому стохастичниму середовищі Ξ. Будемо вважати, що модель середовища апріорі не відома агентам. Інформація про середовище доступна агентам тільки у вигляді відповідної на їх вплив реакції. Для цього кожен з агентів $i \in D$ має набір чистих стратегій $U^i = (u^i(1), u^i(2), ..., u^i(N_i))$, $N_i \geq 2$, які використовуються як варіанти керівних дій. У дискретні моменти часу *n=1,2,…* агенти здійснюють незалежний вибір та реалізацію однієї із власних чистих стратегій $u_n^i = u^i \in U^i$. Після завершення вибору стратегій всіма агентами до моменту часу *n+1* кожен з них спостерігає випадкові поточні виграші $\xi_n^i = \xi_n^i(u_n^{D_i}, \omega)$, які є функцією спільних стратегій $u^{D_i} \in U^{D_i} = \underset{j \in D_i}{\otimes} U^j$ агентів з локальних підмножин $D_i \neq \varnothing$ та елементарних випробувань $\omega \in \Omega$.

Вважається, що послідовності випадкових величин $\{\xi_n^i\}$ є незалежними $\forall u_n^{D_i} \in U^{D_i}$, $\forall i \in D$, $\forall n = 1,2,...$, а їх математичні сподівання $M\{\xi_n^i(u^{D_i}, \omega)\} = v^i(u^{D_i}) = const$ апріорі не відомі та мають обмежений другий момент $\sup_n M\{[\xi_n^i(u^{D_i}, \omega)]^2\} = \sigma_i^2 < \infty$. Обмін інформацією між агентами про вибрані стратегії або отримані поточні виграші не здійснюється.

Агенти характеризуються імовірностями відмов $\eta^i \in [0,1)$. Нехай $\psi^i \in \{0,1\}$ – ознака участі *i*-го агента у грі. Якщо $\psi^i = 0$, то агент відмовляється від поточного ходу гри з імовірністю η^i, якщо $\psi^i = 1$ – бере участь у грі з імовірністю $1 - \eta^i$. Будемо вважати, що функціональність агентів відновлюється до наступного моменту часу. На кожному кроці гри відбувається обмін ознаками поточних станів ψ_n^i між сусідніми агентами з множин D_i $\forall i \in D$. Відмови агентів 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 агентів з множин D_i. Нехай $D_i(\omega) \subseteq D_i$ – підмножина агентів, які залишаються у грі для події $\omega \in \Psi_i$. Тоді поточні виграші i-го агента дорівнюють $\xi_n^i = \chi\{\sum_{j \in D_i} \psi_n^j \geq \overline{\psi}\} \xi_n^i(u_n^{D_i(\omega)})$, де $\chi(\cdot) \in \{0,1\}$ – індикаторна функція події, $\overline{\psi} > 0$ – поріг гри, або мінімально допустима кількість агентів, що не відмовили. Серед можливих значень порогу гри розглянемо два: якщо $\overline{\psi} = 1$, то є можливими варіанти гри з природою; якщо $\overline{\psi} > 1$, то гра з природою не допускається.

З врахуванням відмов агентів 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 агента є максимізація функції середніх виграшів

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

За відсутності обміну інформацією між агентами пошук розв'язків задачі векторної оптимізації (2) виконаємо у множині точок рівноваги за Нешем:

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

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

Отже, спостерігаючи власні поточні виграші ξ_n^i та отримуючи інформацію про відмови з локальної підмножини D_i, агенти-гравці повинні вибирати чисті стратегії u_n^i так, щоб сформована послідовність $\{u_n^i\}$ варіантів дій при $n \to \infty$ задовольняла умову рівноваги (3).

2. Метод розв'язування стохастичної гри

Генерування чистих стратегій u_n^i здійснюється агентами на основі динамічних векторів змішаних стратегій p_n^i, елементи яких визначаються як умовні імовірності вибору відповідних чистих стратегій $p_n^i(u_n^i) = P\{u_n^i \mid u_t^i, \xi_t^i \, (t = \overline{1, n-1})\} \quad \forall u_n^i \in U^i, \forall i \in D$. Для розв'язування стохастичної гри необхідно розробити метод зміни векторів змішаних стратегій, який забезпечить виконання умови (3). Такий метод повинен бути адаптивним до невизначеностей середовища прийняття рішень [12].

Пошук розв'язків ігрової задачі виконаємо за допомогою рекурентних методів

$$p_{n+1}^i = p_{n+1}^{N_i}\left\{p_n^i - \gamma_n R_n(u_n^i, p_n^i, o_n^i)\right\} \in S^{N_i} \quad \forall i \in D, \tag{4}$$

де $\pi_{n+1}^{N_i}$ – перетворення, яке забезпечує належність вектора p_{n+1}^i одиничному симплексу $S^{N_i} = \{p^i \in R^{N_i} \mid p^i(j) \ge 0 (j = \overline{1, N_i}), \sum_{j=1}^{N_i} p^i(j) = 1\}$; $\gamma_n \ge 0$ – крок методу; $R_n(\cdot) \in R^{N_i}$ – вектор руху методу.

Для забезпечення умови (3) вектори змішаних стратегій, що породжуються методом (4), повинні задовольняти умову псевдоградієнтності:

$$\rho_n(p^i) = \left\langle M\{R(p_n^i, u_n^i, \xi_n^i) \mid p_n^i = p^i\}, \nabla_{p^i} \Delta(p) \right\rangle \ge 0, \tag{5}$$

де $p \in S = \prod_{i=1}^{L} S^{N_i}$; $L = |D|$ – кількість агентів; $\Delta(p)$ – функція Ляпунова.

Умова псевдоградієнтності означає, що для досягнення мети вектор руху методу повинен у середньому утворювати гострий кут з напрямком на оптимальний розв'язок ігрової задачі p^*.

Функція Ляпунова задовольняє такі умови: 1) $\Delta(p)$ – диференційована на симплексі *S*; 2) $\Delta(p) > 0$ для всіх точок симплексу *S*, крім точок $p^* \in S$; 3) $\Delta(p^*) = 0$ у точках оптимального розв'язку.

В загальному цим умовам відповідає функція

$$\Delta = \sum_{i=1}^{L} \Delta^i = \sum_{i=1}^{L} \left\| p^i - p^{i*} \right\|^2, \tag{6}$$

де p^{i*} – оптимальна стратегія *i*-го агента, $\|\circ\|$ – евклідова норма.

Для побудови методу з потрібними властивостями розглядається матричне формулювання асимптотично адекватної безкоаліційної ігрової задачі з функціями середніх виграшів агентів

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

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; $D_i(\omega) \subseteq D_i$ – підмножина агентів, які залишаються у грі для події ω; $v^i(u^{D_i(\omega)})$ – математичне сподівання виграшів для спільної стратегії $u^{D_i(\omega)}$.

Для диференційованих на S^{N_i} функцій V^i оптимальні змішані стратегії знаходяться з умови доповняльної нежорсткості [13]

$$\nabla_{p^i} V^i = V^i e^{N_i}, \ p^i \in S^{N_i}, \forall i \in D, \tag{7}$$

де $\nabla_{p^i} V^i$ – градієнт функції V^i; e^{N_i} – вектор, що складається з N_i одиниць. Значення $p^i \ \forall i \in D$, для яких виконується умова (7), називаються вирівнювальними.

Для знаходження можливих розв'язків гри на межі одиничного симплексу виконаємо зважування виразу (7) елементами векторів змішаних стратегій p^i:

$$diag(p^i)(V^i e^{N_i} - \nabla_{p^i} V^i) = 0, \ \forall i \in D, \tag{8}$$

де $diag(p^i)$ – квадратна діагональна матриця порядку N_i, побудована з елементів вектора p^i.

Враховуючи, що $diag(p^i)(V^i e^{N_i} - \nabla_{p^i} V^i) = M\{\xi_n^i [p_n^i - e(u_n^i)] \mid p_n^i = p^i\}$, методом стохастичної апроксимації [14] зваженої умови доповняльної нежорсткості (8), отримаємо такий рекурентний метод:

$$p_{n+1}^i = \pi_{\varepsilon_{n+1}}^{N_i} \left\{ p_n^i - \gamma_n \chi\{\sum_{j \in D_i} \psi_n^j \geq \overline{\psi}\} \xi_n^i (u_n^{D_i(\omega)}) \left[p_n^i - e(u_n^i) \right] \right\}, \tag{9}$$

де $\pi_{\varepsilon_{n+1}}^{N_i}$ – проектор на одиничний ε-симплекс [7]; $e(u_n^i)$ – одиничний вектор-індикатор вибору чистої стратегії $u_n^i \in U^i$.

Умова псевдоградієнтності (5)

$$V^i(p) \left\| p^i - \tilde{p}^i \right\|^2 \geq 0,$$

виконується для методу (9) у знакододатних середовищах при $p^{i*} = \tilde{p}^i$, де $\tilde{p}^i = diag(p^i) \nabla V^i / V^i$. Середовище є знакододатним, якщо $V^i(p) > 0$ для всіх $p \in S$.

3. Умови збіжності ігрового методу

Для виконання умови (3) необхідно накласти обмеження на послідовності величин $\{\gamma_n\}$ та $\{\varepsilon_n\}$.

Дослідження умов збіжності методу (9) виконаємо на основі поточного значення функції Ляпунова Δ_n, яка має вигляд (6) при $p^i = p_n^i$ та $p^{i*} = \tilde{p}_n^i$ і визначає поточну похибку виконання умови доповняльної нежорсткості (8).

Для методу (9) отримано верхню оцінку умовного математичного сподівання функції Δ_{n+1} по σ-алгебрі подій $\mathsf{F}_n = \sigma\{u_t^i, \xi_t^i \mid \forall i \in D; \ t = \overline{1, n}\}$, яка визначає передісторію гри:

$$M\{\Delta_{n+1} | \mathsf{F}_n\} \leq (1 - 2\gamma_n Y_{\min} v_{\min}) \Delta_n + K(\mu_n + \gamma_n^2), \tag{10}$$

де $Y_{\min} = \min_{D_i \subseteq D} \left(1 - \prod_{j \in D_i} \eta^j \right)$, якщо допускаються варіанти гри з природою; $Y_{\min} = \min_{D_i \subseteq D} \left(1 - \prod_{j \in D_i} \eta^j - \sum_{j \in D_i} (1 - \eta^j) \prod_{k \in D_i \setminus j} \eta^k \right)$, якщо варіанти гри з природою не допускаються; $v_{\min} = \min_{i \in D} \min_{u^{D_i} \in U^{D_i}} v^i(u^{D_i}) > 0$; $K \sim |D| (v_{\max} - v_{\min}) \sigma_{\max}^2 > 0$; $\mu_n = |\gamma_{n-1} - \gamma_n| + |\varepsilon_n - \varepsilon_{n+1}|$.

Якщо імовірності відмов η є однаковими для всіх агентів, то $Y_{\min} = 1 - \eta$ для варіантів гри з природою, та $Y_{\min} = (1-\eta)^2$ за відсутності варіантів гри з природою.

Усереднюючи (10) по F_n, отримаємо

$$M\{\Delta_{n+1}\} \le (1 - 2\gamma_n Y_{\min} v_{\min}) M\{\Delta_n\} + K(\mu_n + \gamma_n^2) . \qquad (11)$$

Твердження 1. В умовах попарної незалежності послідовностей випадкових величин $\{\xi_n^i\}$, $\{u_n^i\}$ $\forall i \in D$ $\forall u_n^{D_i} \in U^{D_i}$, $\forall i \in D$, $\forall n = 1,2,...$ проекційний ігровий метод (9) з відмовами агентів при $n \to \infty$; $\gamma_n > 0$; $\gamma_{n+1} < \gamma_n$; $\lim\limits_{n\to\infty} \gamma_n = \infty$, $\varepsilon_n \in (0, \max\limits_{i\in D} N_i^{-1})$; $\varepsilon_{n+1} < \varepsilon_n$; $\lim\limits_{n\to\infty} \varepsilon_n = 0$ для довільного початкового наближення на одиничному симплексі $p_1^i \in S_{\varepsilon_1}^{N_i}$, $\forall i \in D$ забезпечують пошук вирівнювальних стратегій, для яких виконується умова доповняльної нежорсткості (8) у знакододатному середовищі $v_{\min} > 0$:

1) з імовірністю 1, якщо $\sum\limits_{n=1}^{\infty} (\mu_n + \gamma_n^2) < \infty$;

2) у середньоквадратичному, якщо $\lim\limits_{n\to\infty} (\eta_n \gamma_n^{-1} + \gamma_n) = 0$.

Метод (9) забезпечує в асимптотиці часу виконання умови рівноваги за Нешем у грі з відмовами з деякою похибкою відносно умови рівноваги у грі без відмов агентів. Інакше, розв'язки гри з відмовами є ε-рівноважними за Нешем.

Доведення твердження 1 ґрунтується на оцінках (10), (11), застосуванні теореми Робінса-Сігмунда та теорем про рекурентні числові нерівності [7, 12, 15].

У класі монотонно спадних невід'ємних послідовностей $\{\gamma_n\}, \{\varepsilon_n\}$ вигляду

$$\gamma_n = \gamma(n+a)^{-\alpha}; a > 0; \ \varepsilon_n = \varepsilon(n+b)^{-\beta}; b > 0 \qquad (12)$$

збіжність методу (9) існує:

1) з імовірністю 1, якщо $\alpha \in (0.5;1]$; $\beta > 0$;

2) у середньоквадратичному, якщо $\alpha \in (0,1]$; $\beta > 0$.

З врахуванням (12) оцінку (11) запишемо у вигляді

$$M\{\Delta_{n+1}\} \le \left(1 - 2v_{\min} Y_{\min} \frac{\gamma}{n^{\alpha}}\right) M\{\Delta_n\} + K\left(\frac{\gamma\alpha}{n^{1+\alpha}} + \frac{\varepsilon\beta}{n^{1+\beta}} + \frac{\gamma^2}{n^{2\alpha}}\right) \le$$

$$\le \left(1 - 2v_{\min} Y_{\min} \frac{\gamma}{n^{\alpha}}\right) M\{\Delta_n\} + K \frac{m}{n^{\rho}}, \qquad (13)$$

де $m = \varepsilon\beta\chi(1+\beta \le 2\alpha) + \gamma\chi(\alpha = 1 \wedge \beta \ge 1) + \gamma^2 \chi(1+\beta \ge 2\alpha)$; $\rho = \min(1+\beta, 2\alpha)$.

Швидкість збіжності методу (9) до вирівнювальних стратегій визначається методом моментів Чжуна [7, 12]

Твердження 2. Нехай виконані умови збіжності методу (9) у середньоквадратичному для послідовностей $\{\gamma_n\}$, $\{\varepsilon_n\}$ вигляду (12), зокрема, $\alpha \in (0,1]$, $\beta > 0$, і при $\alpha = 1$ виконується нерівність $2\gamma Y_{\min} v_{\min} > \min(1, \beta)$, тоді

$$\overline{\lim_{n\to\infty}} \, n^{\theta} M\{\Delta_n\} \le \frac{Km}{2\gamma Y_{\min} v_{\min} - \min(1,\beta)\chi(\alpha = 1)} = \vartheta < \infty, \qquad (14)$$

де $\theta = \min(1 + \beta - \alpha, \alpha)$ – порядок швидкості збіжності.

Доведення твердження 2 ґрунтується на оцінці (13) та застосуванні теорем про рекурентні числові нерівності [15].

З (14) слідує, що асимптотична швидкість збіжності методу (9) становить $\vartheta n^{-\theta}$. Параметр ϑ визначає величину, а параметр θ – порядок швидкості збіжності.

Наслідок 1. В умовах твердження (2) параметр θ, який визначає порядок швидкості збіжності методу (9), задовольняє нерівності $\theta \le 1$. Максимальний порядок швидкості збіжності становить n^{-1} і досягається при $\theta = 1$, $\alpha = 1$, $\beta \ge 1$.

Наслідок 2. В умовах твердження 2 для $\theta = 1$ при $\alpha = 1$ оптимальні початкові значення параметрів методу (9) становлять:

$\varepsilon \in (0, \min_{i \in D} N_i^{-1})$; $\gamma = (1 + \sqrt{1 + 4(Y_{\min} v_{\min})^2 \varepsilon + 2Y_{\min} v_{\min}})/(2Y_{\min} v_{\min})$, якщо $\beta = 1$; $\gamma = (1 + \sqrt{1 + 2Y_{\min} v_{\min}})/(2Y_{\min} v_{\min})$, якщо $\beta > 1$.

З (14) слідує, що відмови агентів приводять до сповільнення швидкості збіжності ігрового методу. Метод (9), який допускає варіанти "гри з природою" в середньому забезпечує більшу величину швидкості збіжності, ніж метод, який не дозволяє таких варіантів.

Відмови агентів дають можливість дослідити вплив динаміки локальних зв'язків між ними на поведінку ігрового методу. З оцінки (14) випливає, що повнозв'язна гра має більшу швидкість збіжності до рівноважного стану ε-оптимальності, ніж гра з локальними зв'язками. При зменшенні потужності локальної множини D_i швидкість збіжності ігрового методу зменшується, але зростає стійкість рівноваги досягнутого рішення, оскільки умова рівноваги за Нешем (3) виконується для i-го гравця і тоді, коли агенти з доповняльної множини $D \setminus D_i$ не притримуються оптимальних стратегій поведінки. Ігрові методи з відмовами забезпечують в асимптотиці виконання умови ε-рівноваги за Нешем, визначеної для функції середніх виграшів для гри з абсолютно надійними агентами. При зменшенні імовірностей відмов агентів до нуля ε-рівноважні розв'язки прямують до рівноважних за Нешем розв'язків гри.

4. Результати комп'ютерного моделювання

Результати комп'ютерного моделювання гри з відмовами показано на рис. 1 – 7 у вигляді графіків $\lg(f)$ усереднених по кількості агентів функцій середніх виграшів $f = |D|^{-1} \sum_{i \in D} \Phi_n^i$ (1), поточної $f = \Delta_n$ та усередненої у часі $f = n^{-1} \sum_{t=1}^{n} \Delta_t$ похибки доповняльної нежорсткості. Результати отримано для гри $|D| = |D_i| = 5$ агентів, кожен із яких має по $N_i = 2$ чистих стратегії. Досліджено роботу ігрового методу (9) з параметрами $\gamma_n = 0.5n^{-0.5}$, $\varepsilon_n = 0.45n^{-1}$. Знакододатне середовище задано імовірностями виграшів $v^i(u^{D_i}) \in [0.1; 0.9]$. Поточні виграші агентів визначаються бінарним розподілом:

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

де $\omega \in [0,1]$ – дійсне випадкове число.

Кількість ітерацій моделювання однієї реалізації ігрового методу прийнята такою, що дорівнює 10 тис. кроків. Узагальнення результатів отримано усередненням 50 реалізацій методу для фіксованих початкових даних.

Експериментально встановлено, що збіжність досліджуваних у роботі адаптивних ігрових методів не залежить від початкового наближення на одиничному симплексі та від закону розподілу випадкових виграшів. При зростанні розмірності ігрової задачі (кількості агентів, кількості чистих стратегій, потужності множин D_i) швидкість збіжності ігрових методів зменшується.

Порядок швидкості збіжності ігрових методів визначається на основі логарифмування виразу (14) та оцінки $\theta \approx tg(\varphi)$, де φ – кут нахилу прямої лінійної апроксимації функції Δ_n з координатним напрямком часу на відрізку [10^3, 10^4]. Більшим значенням θ відповідає більша швидкість збіжності ігрового методу (9).

Вплив відмов агентів на швидкість збіжності методів. Визначимо як впливають відмови агентів на швидкість збіжності ігрового методу (9). Для цього зафіксуємо параметри $\gamma_0 = 0.5$, $\alpha = 0.5$,

$\beta = 1$, а імовірності відмов η задамо з інтервалу [0,1) однаковими для всіх агентів. Результати моделювання гри з відмовами для методу (9) з бінарними виграшами показано на рис. 1. Графік з номером $1 \le k \le 5$ відповідає значенню $\eta = 0.9 - 0.2 * (k - 1)$, графік з номером 6 – значенню $\eta = 0$.

Найбільша швидкість збіжності забезпечується ігровим методом без відмов агентів (графік 6). Із зростанням імовірностей відмов швидкість збіжності зменшується.

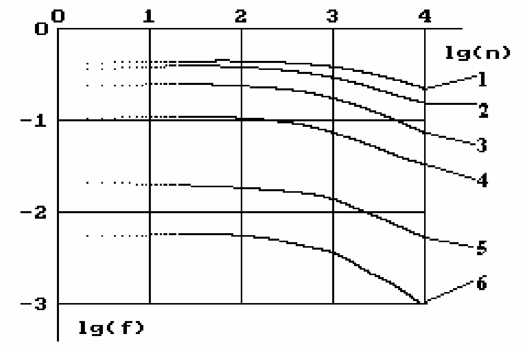
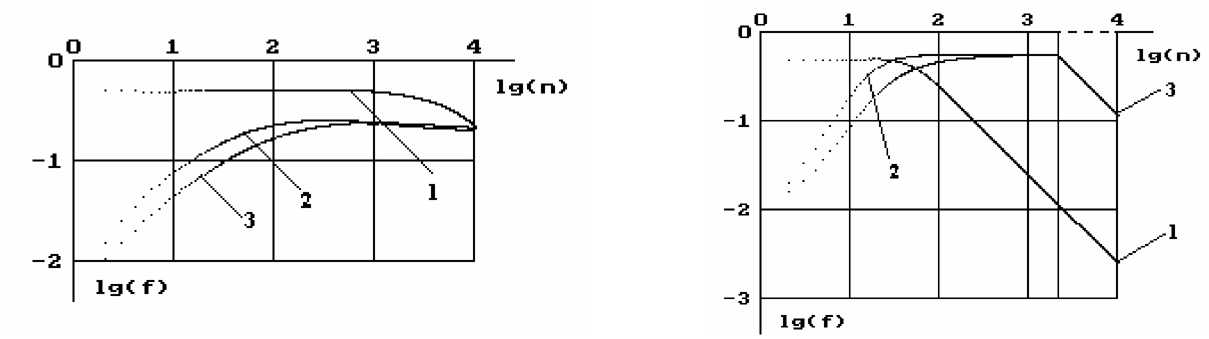


Рис. 1. Вплив імовірностей відмов агентів на збіжність ігрового методу

Моделювання динаміки відмов агентів. **Зростання імовірності відмов у часі.** На відміну від попередньої моделі, де імовірності відмов агентів були фіксованими для всіх кроків експерименту, будемо змінювати їх значення у часі. Нехай імовірності відмов агентів задаються зростаючими у часі функціями, які приймають значення з інтервалу [0,1]. Зростання імовірностей відмов моделює зростання "апатії" агентів або послаблення їх "зацікавленості" до ситуації, що склалася у грі.

На рис. 2а показані результати моделювання методу (9), коли імовірності відмов агентів зростають у часі за лінійним законом, а на рис. 2б – за експоненційним законом. Графік 1 відповідає функції середніх виграшів, 2 – поточній похибці доповняльної нежорсткості, 3 – усередненій у часі похибці доповняльної нежорсткості.

Відмови агентів призводять до погіршення показників гри, що проявляється у зменшенні їх середніх виграшів (графік 1). Порівняно з лінійним законом, зростання імовірностей відмов за експоненційним законом призводить до швидшого зменшення середніх виграшів агентів.



а *б*

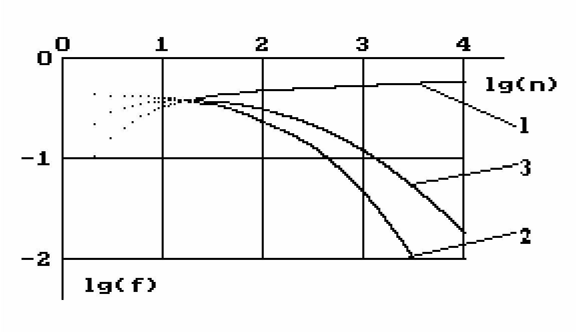
Рис. 2. Реакція ігрового методу на зростання імовірностей відмов агентів у часі:
а – за лінійним законом $\eta = 0.25 \lg n$ *; б – за експоненційним законом* $\eta = 1 - e^{-0.05n}$

При зростанні імовірностей відмов агентів спостерігається відхилення від ситуації рівноваги, що проявляється у зростанні похибки доповняльної нежорсткості (графіки 2 та 3). Гра із зростаючими відмовами продовжується до моменту часу, коли імовірності відмов досягають свого максимального (одиничного) значення (графік 3 рис. 2б).

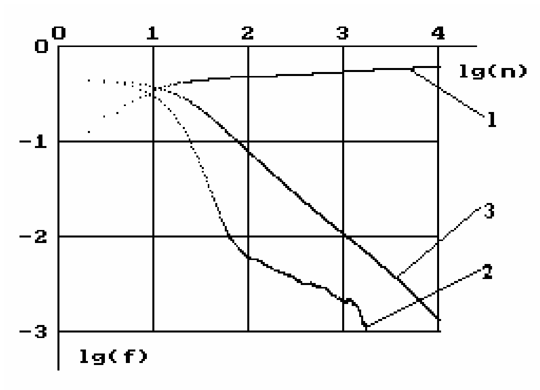
Моделювання зменшення імовірності відмов агентів у часі. Нехай імовірності відмов агентів є спадною функцією часу. Це означає, що з плином часу агенти беруть все активнішу участь у грі. Інакше кажучи, "азарт" або "зацікавлення" агентів зростає у процесі гри.

Будемо зменшувати імовірності відмов агентів від максимального (одиничного) значення до нуля за лінійним та за експоненційним законами. Характеристики поведінки методу (9) з лінійно спадною імовірністю відмов показані на рис. 3а, а за спадним експоненційним законом – на рис. 3б. Графік 1 демонструє зміну функції середніх виграшів, 2 – поточної похибки доповняльної нежорсткості, 3 – усередненої у часі похибки доповняльної нежорсткості.

Зменшення імовірностей відмов агентів призводить до зростання функцій середніх виграшів (графік 1). Швидше, порівняно з лінійним, експоненційне зменшення імовірностей відмов агентів забезпечує більше значення швидкості збіжності ігрового методу (графіки 2 та 3).



а



б

Рис. 3. Реакція ігрового методу на зменшення імовірностей відмов агентів у часі:
а – за лінійним законом $\eta = 1 - 0.25\lg n$ *; б – за експоненційним законом* $\eta = e^{-0.05n}$

Порівняно з методом без відмов агентів (графік 6 рис. 1) метод із спадною імовірністю відмов забезпечує на початковому відрізку часу більший порядок швидкості збіжності, що можна пояснити наявністю швидко затухаючих часових затримок перерахунку векторів змішаних стратегій. Результатом цього є повільнішою зміна елементів векторів змішаних стратегій, ніж для гри з абсолютно надійними агентами, які не пропускають поточних ходів.

Моделювання періодичної зміни імовірностей відмов агентів. Для практичних задач реальнішими є ситуації, коли активність агентів змінюється періодично у часі. Для моделювання цього будемо змінювати імовірності відмов агентів за гармонійним законом з фіксованим періодом часу T:

$$\eta = 0.5(1 + \sin(2\pi n/T)).$$

Результати моделювання перехідних станів показані на рис. 4. Як і раніше, графік 1 зображає функцію середніх виграшів, 2 – поточну похибку доповняльної нежорсткості, 3 – усереднену в часі похибку доповняльної нежорсткості.

Синусоїдальна зміна імовірностей відмов накладає відбиток на всі характеристики поведінки ігрового методу. Функції середніх виграшів агентів та функції похибок доповняльної нежорсткості мають чітко виражений періодичний характер. Порівняно з методом без відмов агентів, метод з періодичною зміною імовірностей відмов призводить до сповільнення швидкості збіжності гри.

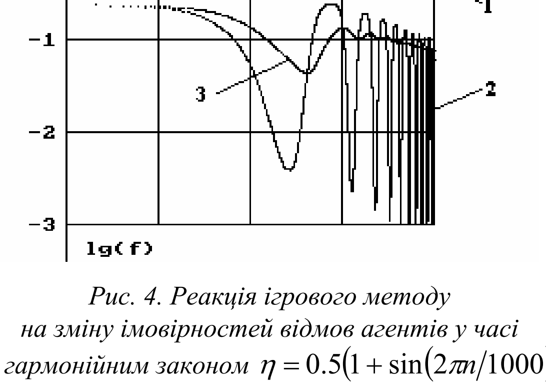


Рис. 4. Реакція ігрового методу на зміну імовірностей відмов агентів у часі за гармонійним законом $\eta = 0.5(1 + \sin(2\pi n/1000))$

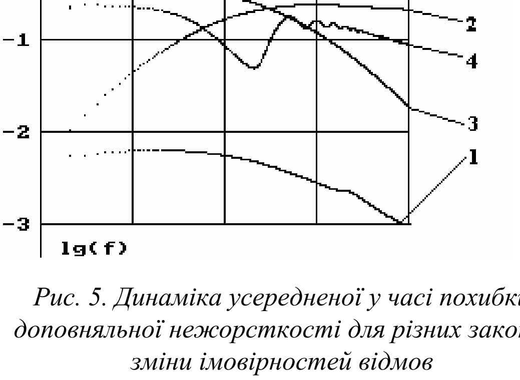


Рис. 5. Динаміка усередненої у часі похибки доповняльної нежорсткості для різних законів зміни імовірностей відмов

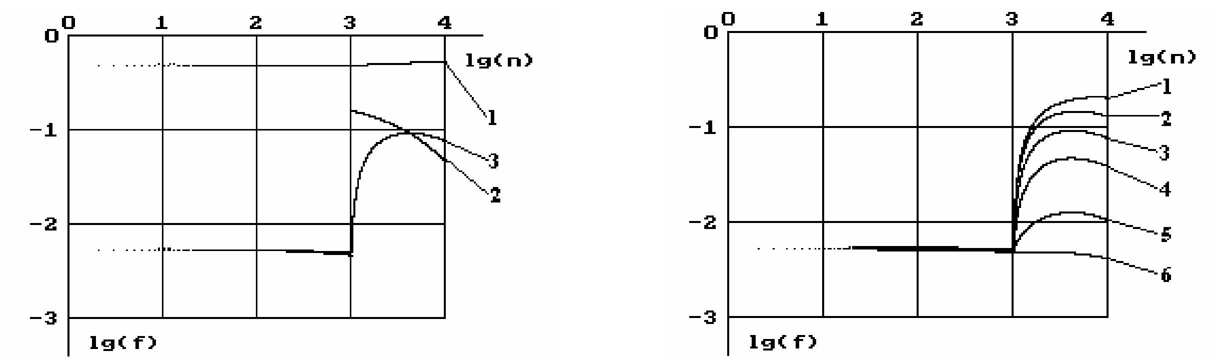
Для зручності відносного порівняння на рис. 5 зображені графіки зміни усередненої у часі похибки доповняльної нежорсткості для розглянутих вище методів з відмовами та без відмов агентів. Графік 1 відповідає варіанту гри без відмов агентів, графік 2 – з лінійно зростаючими імовірностями відмов $\eta = 0.25\lg n$, графік 3 – з лінійно спадними імовірностями відмов $\eta = 1 - 0.25\lg n$, графік 4 – з періодичною зміною імовірностей відмов $\eta = 0.5\left(1 + \sin\left(2\pi n/500\right)\right)$.

Метод без відмов агентів забезпечує мінімальну початкову похибку умови доповняльної нежорсткості. Порядок швидкості збіжності методу з лінійним зменшенням імовірностей відмов дещо перевищує порядок швидкості збіжності методу без відмов агентів. Монотонне зростання у часі імовірностей відмов агентів призводить до погіршення характеристик роботи ігрового методу, що проявляється у зростанні похибки доповняльної нежорсткості.

Вплив порогової зміни надійності агентів на поведінку ігрового методу. Розглянемо модель гри, у якій агенти є абсолютно надійними лише до моменту часу *n=1000* кроків, а далі їх імовірність відмов змінюється стрибком до певного значення і залишається такою до кінця експерименту (рис. 6). Виявимо вплив величини стрибкової зміни імовірностей відмов агентів на стійкість стану рівноваги ігрового методу (9).

Дані отримано для значень параметрів $\gamma_0 = 0.1$, $\alpha = 0.5$, $\beta = 1$ та бінарних виграшів агентів.

На рис. 6а подано графіки характеристик ігрового методу для імовірності відмов $\eta = 0.5$. Графік 1 відповідає функції середніх виграшів, 2 – поточній похибці умови доповняльної нежорсткості, 3 – усередненій у часі похибці умови доповняльної нежорсткості. На рис. 6б зображено графіки зміни усередненої у часі похибки доповняльної нежорсткості для різних значень імовірностей відмов агентів. Графік з номером $1 \le k \le 5$ відповідає значенню $\eta = 0.9 - 0.2*(k-1)$, графік з номером 6 – значенню $\eta = 0$.



а *б*

Рис. 6. Стійкість стану рівноваги гри для відновлювальних відмов

З рисунків видно, що стрибкове зростання імовірності відновлювальних відмов агентів призводить до стрибкового зростання похибки доповняльної нежорсткості (рис. 6б). В ході подальшого розгортання гри спостерігається поступове зменшення похибки доповняльної нежорсткості, що засвідчує адаптаційні властивості ігрового методу та стійкість досягнутого ним стану рівноваги.

Гра з невідновлювальними відмовами агентів. Нехай до фіксованого моменту часу, який визначається досягненням стану рівноваги з певною точністю, імовірності відмов агентів дорівнюють нулю, після чого підмножина агентів залишає гру. Агенти, які відмовили, не беруть участь у грі до моменту її закінчення. Визначимо вплив потужності множини агентів, що відмовили, на стійкість стану рівноваги.

Результати дослідження зображені у вигляді графіків усередненої у часі похибки доповняльної нежорсткості на рис. 7. До моменту часу *n=5000* кроків моделюється повнозв’язна гра без відмов агентів. До цього часу вектори змішаних стратегій агентів будуть достатньо сформованими (див. графік 1 на рис. 5), тобто метод (9) можна вважати навченим з певною точністю. Далі частина агентів завершує гру і не повертається до неї до моменту її завершення. Графік з номером *k* ($k = \overline{1,5}$) демонструє хід гри для випадку, коли відмовило *k-1* агентів.

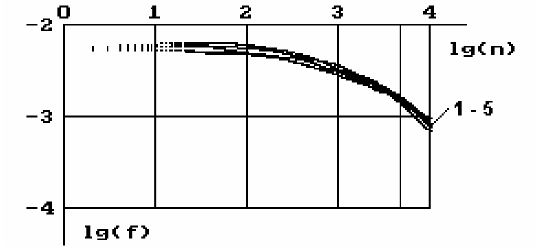


Рис. 7. Стійкість стану рівноваги гри для невідновлювальних відмов

Невідновлювальні відмови агентів призводять до зменшення розмірності ігрової задачі і, в результаті, – до зміни функцій середніх виграшів. Із рис. 7 видно, що вихід агентів з гри після завершення першої, основної, фази її навчання істотно не впливає на зміну похибки доповняльної нежорсткості. Це свідчить про стійкість стану гри, який досягається під час навчання ігрового методу (9).

Висновки. 1. Порядок швидкості збіжності ігрового методу, в основному, визначається способом формування його регульованих параметрів та імовірностями відмов агентів.

2. Відмови агентів призводять до відхилення асимптотичного розв'язку ігрової задачі від оптимального значення та до зменшення величини швидкості збіжності ігрового методу.

3. Результати комп'ютерного моделювання стохастичної гри на вибірках достатньо великої довжини (декілька тисяч кроків) підтверджують результати теоретичних досліджень. Незначні відмінності експериментальних даних від теоретичних оцінок асимптотичної швидкості збіжності ігрового методу обумовлені переважаючою дією на початковому відрізку часу моделювання складових випадкових процесів вищих порядків, які не були враховані у теоретичних оцінках.

4. Достовірність результатів комп'ютерного моделювання ігрового методу забезпечено підготовкою та плануванням експерименту і підтверджено повторюваністю розрахункових величин та їх загальною відповідністю теоретичним результатам.

5. Окремого дослідження вимагає проблема впливу відмов агентів на збіжність гри з обміном інформацією.

1. Gerhard Weiss and Sandip Sen, editors. Adaptation and Learning in Multiagent Systems. Springer Verlag, Berlin, 1996. 2. Wooldridge M. An Introduction to Multiagent Systems. John Wiley & Sons (Chichester, England), 2002. 3. Stone, P. Layered Learning in Multiagent Systems. MIT Press,2000. 4. Доманский В.К. Стохастические игры // Математические вопросы кибернетики. – 1988. – № 1. – С. 26 – 49. 5. Fudenberg, D., Levine, D.K.: The Theory of Learning in Games. MIT Press, 1998. 6. Воробьев Н.Н. Основы теории игр: Бескоалиционные игры. – М.: Наука, 1984. 7. Назин А.В., Позняк А.С. Адаптивный выбор вариантов: Рекуррентные алгоритмы. – М., 1986. 8. Кравець П.О. Ігрові методи випадкового пошуку в умовах невизначеності // Інформаційні системи та мережі: Вісник НУ "Львівська політехніка". – 2005. – № 549. – С. 105 – 117. 9. Кравець П.О. Ігрова задача взаємодії елементів мультиагентних систем // Комп'ютерні науки та інформаційні технології: Вісник НУ "Львівська політехніка". – 2006. – № 565. – С. 140 – 149. 10. Mailath, George J. & Morris, Stephen. Coordination failure in repeated games with almost-public monitoring // Theoretical Economics, Society for Economic Theory, vol. 1(3), 2006 – P. 311 – 340. 11. Ulidowski Irek. Refusal Simulation and Interactive Games // Algebraic Methodology and Software Technology: 9th International Conference, AMAST 2002, Saint-Gilles-les-Bains, Reunion Island, France, September 9-13, 2002. Proceedings. – P. 208. 12. Цыпкин Я.З., Позняк А.С. Рекуррентные алгоритмы оптимизации в условиях неопределенности // Итоги науки и техники. Сер. Техническая кибернетика. – 1989. – Т. 16. – С. 3 – 70. 13. Мулен Э. Теория игр с примерами из математической экономики. – М.: Мир, 1985. 14. Вазан М. Стохастическая аппроксимация. – М.: Мир, 1972. 15. Невельсон М.Б., Хасьминский Р.З. Стохастическая аппроксимация и рекуррентное оценивание. – М.: Наука, 1972.