УДК 004.852; 004.942

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

ІГРОВА САМООРГАНІЗАЦІЯ СИСТЕМИ АГЕНТІВ З ІНДИВІДУАЛЬНИМ ОЦІНЮВАННЯМ СТРАТЕГІЙ

Досліджено проблему ігрової самоорганізації мультиагентних систем для розв'язування задач стохастичної оптимізації в умовах невизначеності. Побудувано імітаційну модель, рекурентні методи, алгоритм та виконано комп'ютерне розв'язування ігрової задачі. Результати роботи можуть бути використані для побудови елементів розподілених інтелектуальних систем.

The problem of game self-organizing multiagent system for the tasks solution of stochastic optimization in conditions of uncertainty is investigated. The simulation model, a recurrent methods, an algorithm are constructed, and a game task solution on the computer is executed. The results of work can be used for construction of the elements of distributed intelligent systems.

Вступ. Компоненти глобальних динамічних систем, якими є інформаційна мережа Internet, мобільний зв'язок, ринкова економіка, електронна комерція тощо, повинні в процесі їхнього використання розвиватися та адаптуватися до потреб користувачів, оскільки на стадії їхнього проектування неможливо передбачити усі необхідні функції. Така задача не під силу централізованим методам керування у зв'язку із багатофакторністю, невизначеністю даних та високою вартістю реалізації. Для її розв'язування необхідно використати розподілені системи автономних, пов'язаних між собою агентів [1–3].

Агент – це автономна апаратна або програмна сутність, яка діє в інтересах досягнення сформульованих користувачем цілей, із урахуванням її внутрішніх уявлень про середовище. Агент існує у середовищі, взаємодіє з середовищем та іншими агентами і може змінювати середовище. Мультиагентна система (МАС) – це множина автономних агентів, що взаємодіють. МАС об'єднують результати багатьох дисциплін: штучний інтелект, розподілене опрацювання інформації, соціологію, керування, біологію, психологію, філософію. МАС будують за однією із архітектур: 1) реактивною, коли агенти працюють за принципом ситуація–дія без наявності моделі середовища; 2) делібертивною, або ментальною, коли агенти приймають рішення на основі моделі середовища та мають здатність до "роздумів" – аналізу та виведення варіантів рішень; 3) змішаною – найперспективніша, яка має переваги реактивної (простота організації та мінімум наявної інформації про середовище) та делібертивної (порівняно легко забезпечують цілеспрямовану або "розумну" поведінку у складних середовищах) архітектур.

Агенти є новою моделлю обчислень для розподілених, великих, відкритих та гетерогенних технічних систем опрацювання інформації. На етапі проектування агентів таких систем у їхні функції необхідно закласти можливості самоорганізації та саморегулювання [4–6].

Самоорганізація – це здатність системи змінювати свої функції та будову залежно від завдань та зовнішніх чинників, підтримуючи життєдіяльність або оптимальні характеристики роботи [7].

Самоорганізація МАС черпає ідеї з біології, фізики, хімії та соціальних систем. Самоорганізація реалізується через взаємодію, координацію, навчання та колективні знання агентів. Типовим прикладом самоорганізації є системи, які моделюють поведінку популяцій живих організмів, так зване "штучне життя". Актуальною проблемою є вивчення умов та властивостей самоорганізації природних систем для перенесення їх на кібернетичні системи з елементами штучного інтелекту.

Самоорганізація є базовою властивістю МАС через автономність та локальну взаємодію складових компонентів. Глобальна поведінка складних систем є результатом такої взаємодії між різними агентами, що утворюють цілісну систему. Головна характеристика цілісної системи – це властивість знаходити розв'язки складних задач за допомогою колективів агентів з порівняно простою структурою та індивідуальною поведінкою без центрального керування або підпорядкування.

Однак у реальних системах вплив довкілля, локальні взаємодії агентів та розподілені керуючі впливи на середовище можуть призвести до непередбачуваної або небажаної поведінки системи загалом. Визначення структурної організації та методів навчання автономних агентів, які приводили б до їхньої колективної самоорганізації, є ключовим для проектування МАС. Враховуючи, що МАС функціонують в умовах невизначеності, знайти універсальне вирішення цієї проблеми надзвичайно важко. Розробляючи реальні МАС, можливість самоорганізації необхідно досліджувати на передпроектній стадії за допомогою аналітичного та імітаційного моделювання.

Перспективними засобами для дослідження умов самоорганізації є розроблення та застосування моделей та методів стохастичних ігор, які мають багато спільного з МАС [8–10]. В термінах теорії ігор агенти – це гравці, що взаємодіють у середовищі і досягають колективних станів оптимальності. Вид середовища визначається конкретною предметною галуззю. Гравці мають усі властивості агентів: 1) невизначеність даних – гравці володіють неповною інформацією про середовище або глобальну мету розвитку системи; 2) реактивність – можливість впливати на стан середовища та оперативно реагувати на його зміни; 3) адаптивність – здатність навчатися для реалізації цілеспрямованих дій; 4) автономність – можливість приймати самостійні рішення залежно від поточного стану середовища; 5) колаборативність – можливість співпраці та утворення коаліцій з іншими гравцями для досягнення власних цілей; 6) інтелектуальність – наявність часткових знань або механізмів виведення у конкретній предметній галузі; 7) комунікативність – можливість спілкування з іншими гравцями; 8) мобільність – здатність до переміщення у середовищі.

В ігрових ситуаціях важливими є питання організації взаємодії між агентами, обміну інформацією в ході навчання та оцінювання стратегій. Останнє може бути колективним, коли оцінку повідомляють усім гравцям або їхній локальній коаліції, індивідуальним, коли формується оцінка для кожного гравця окремо, комбінованим, яке об'єднує перші два варіанти.

Використання методології стохастичних ігор як базової для дослідження МАС розглядають багато сучасних дослідників. Основна увага приділяється моделюванню ситуативних ігор агентів з евристичними правилами поведінки [2, 11, 12], що не завжди гарантує самоорганізацію МАС. Недостатньо вивченими є питання побудови структури та методів навчання агентів, які гарантовано забезпечували б колективну самоорганізацію МАС, необхідну для розв'язання конкретної оптимізаційної задачі.

Мета роботи. Метою роботи є побудова математичної та імітаційної моделей ігрової самоорганізації мультиагентної системи для розв'язування задачі розподіленої стохастичної оптимізації в умовах невизначеності з незалежним індивідуальним формуванням оцінок якості функціонування агентів.

Для досягнення мети необхідно розв'язати такі задачі: 1) задати середовище моделювання; 2) виконати формальну постановку задачі – визначити вхідні та вихідні дані, мету розв'язування та обмеження; 3) визначити структуру та функції агентів; 4) побудувати ігрові методи та алгоритми розв'язування задачі; 5) розробити програмну імітаційну модель ігрового розв'язування сформульованої задачі; 6) отримати та проаналізувати результати комп'ютерного моделювання; 7) оцінити ефективність ігрової самоорганізації системи агентів.

Формулювання задачі. Визначимо середовище у вигляді системи $L \ge 1$ незалежних випадкових величин

$$V = \{V^i \mid V^i = \zeta(m_i, d_i); V^i \in R^1;\ i = 1,2,\ldots,L\},$$

розподілених за законом ζ зі стаціонарними математичним сподіванням m_i та дисперсією d_i. Нехай кожен процес контролює автономний агент з внутрішнім скалярним значенням $z_i \in R^1$ на осі дійсних чисел R^1. Тоді елементи МАС задають множиною $Z = \{z_i \mid z_i \in R^1;\ i = 1,2,\ldots,L\}$. Для кожного агента визначимо підмножину $Z_i \subset Z$ топологічно або функціонально близьких (сусідніх) агентів. Загалом така множина може бути утворена випадково:

$$Z_i = \{z_j \mid j = \lfloor \omega_k * L \rfloor + 1; \omega_k \in [0,1); k = 1,2,...,K_i; K_i \leq L\},$$

де ω_k – розподілене за рівномірним законом дійсне випадкове число; $\lfloor x \rfloor$ – ціле, що не перевищує значення дійсного x. Один і той самий агент може входити у декілька множин: $Z_i \cap Z_j \neq \varnothing$.

Кожному агенту поставимо у відповідність вихідну функцію–характеристику $F(Z_i) \in R^1$, яка визначає технічні або економічні показники його роботи.

Нехай задачею агентів є відстеження (прогнозування) контрольованого ними значення випадкової величини V^i, $i = 1,2,...,L$ з еталонним середнім значення m_i. Тоді якість роботи агента визначимо відхиленням його вихідної характеристики від поточного значення випадкового процесу:

$$\xi^i = \left| F(Z_i) - V^i \right|. \qquad (1)$$

Загалом закон розподілу ζ та його параметри невідомі агентам. Можливість досягнення еталонних характеристик середовища забезпечується виконанням умов:

$$\min F(Z_i) + d_i \leq m_i \leq \max F(Z_i) - d_i.$$

Індивідуальна оцінка ξ^i відома тільки для i-го агента. Задачею кожного агента є незалежний вибір значення z_i так, щоб забезпечити виконання умови:

$$M\xi^i \to \min,$$

де M – символ математичного сподівання.

Оскільки функція випадкового розподілу $\zeta(m_i, d_i)$ є невідомою, то для розв'язування сформульованої задачі в умовах апріорної невизначеності необхідно використати моделі та методи адаптивного пошуку [13–15], різновидністю яких є моделі та методи повторюваних стохастичних ігор.

Ігрова модель взаємодії агентів. Нехай $Z = \{[a_i, b_i] \subset R^1; i = 1,2,...,L\}$ – множина агентів-гравців з внутрішніми станами, які набувають значення на інтервалах $z_i \in [a_i, b_i]$. Кожен гравець має $N_i \geq 2$ чистих стратегій

$$z_i = a_i + k_i * h_i - h_i / 2, \qquad (2)$$

де $k_i = 1,2,...,N_i$; $h_i = (b_i - a_i) / N_i$.

Чисті стратегії вибирають у моменти часу $n = 1,2,...$ з імовірностями $p_n^i = \left(p_n^i(j) \mid j = 1,2,...,N_i; \sum_{j=1}^{N_i} p_n^i(j) = 1 \right)$, $i = 1,2,...,L$. Нехай елементи векторів p_n^i є умовними імовірностями $p_n^i(j) = P\left\{ z_n^i = z^i(j) \middle| \xi_t^i, z_t^i, t = 1,2,...,n-1 \right\}$, значення яких залежать від передісторії зміни випадкових процесів ξ_t^i та z_t^i. Вектори p_n^i визначають динамічні змішані стратегії гравців, які набувають значення на одиничних ε–симплексах:

$$S_\varepsilon^{N_i} = \left\{ p^i \mid p^i \in S^{N_i}; p^i(j) \geq \varepsilon \ (j = 1,2,...,N_i) \right\}, \ \varepsilon \in (0, \min_i N_i^{-1}). \qquad (3)$$

Номер чистої стратегії k_i визначають з умови

$$k_i = \left(J \middle| \min_J \sum_{j=1}^{J} p_n^i(j) > \omega \right), \ J = 1,2,...N_i, \qquad (4)$$

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

Після реалізації чистих стратегій z_n^i усіма гравцями кожен з них зчитує значення випадкового процесу V_n^i і, згідно із (1), обчислює поточні втрати ξ_n^i. Ефективність поточної ігрової ситуації оцінюють функцією середніх втрат

$$\Phi_n^i = \frac{1}{n}\sum_{t=1}^{n}\xi_t^i . \tag{5}$$

Отримані поточні втрати використовують гравці для перерахунку власних векторів змішаних стратегій p_n^i, $i = 1,2,...,L$. Зміна векторів p_n^i повинна бути такою, щоб в асимптотиці часу забезпечити мінімум функції середніх втрат:

$$\overline{\lim_{n\to\infty}}\,\Phi_n^i \to \min .$$

Це досягається побудовою відповідних методів навчання агентів, які забезпечують самоорганізацію системи загалом.

Методи розв'язування задачі. Для задачі мінімізації середніх втрат побудуємо метод зміни елементів вектора змішаних стратегій так, щоб при виборі стратегії $z_n^i(j)$ меншим значенням ξ_n^i відповідали більші прирости елемента $p_n^i(j)$. Для забезпечення належності вектора p_n^i одиничному симплексу, якщо необхідно, виконаємо нормування його елементів.

Загальне рекурентне перетворення векторів змішаних стратегій матиме вигляд [16]:

$$p_{n+1}^i = \pi_{\varepsilon_{n+1}}^{N_i}\left\{p_n^i - \gamma_n R(p_n^i, z_n^i, \xi_n^i)\right\}, \tag{6}$$

де $\pi_{\varepsilon_{n+1}}^{N_i}$ – проектор на одиничний ε-симплекс (3), що забезпечує виконання умови $p_{n+1}^i \in S_\varepsilon^{N_i}$; γ_n – монотонно спадна послідовність невід'ємних величин, яка регулює крок методу; $\varepsilon_n \in (0, \min_i N_i^{-1})$ – монотонно спадна послідовність величин, що регулює розширення одиничного ε-симплексу; R – крок методу.

Проектування на одиничний ε-симплекс при фіксованому значенні n полягає у розв'язуванні задачі умовної мінімізації функції

$$\left\|q_n^i - p_n^i\right\|^2 \to \min_{p_n^i}$$

з обмеженнями $\sum_{j=1}^{N_i} p_n^i(j) = 1$, $p_n^i(j) \ge \varepsilon_n$, де $q_n^i \in R^{N_i}$. Ця задача може бути розв'язана, наприклад, методом множників Лагранжа [17].

Раціональне переміщення на одиничному симплексі визначається умовою псевдоградієнтності стосовно функції Ляпунова $\Delta(p)$:

$$\rho_n(p^i) = \left\langle M\{R(p_n^i, z_n^i, \xi_n^i) \mid p_n^i = p^i\}, \nabla_{p^i}\Delta(p)\right\rangle \ge 0 , \tag{7}$$

де $\langle\cdot,\cdot\rangle$ – скалярний добуток векторів у евклідовому просторі; $p^i \in S_0^{N_i}$; $p \in S$; $S = \prod_{i=1}^{m} S_0^{N_i}$.

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

Загалом цим умовам відповідає функція

$$\Delta_n = \sum_{i=1}^{L} \Delta_n^i = \sum_{i=1}^{L} \left\| p_n^i - p^{i*} \right\|^2 ,$$

де p^{i*} – оптимальна стратегія і-го гравця. Інакше, умова псевдоградієнтності означає, що для досягнення мети вектор руху методу повинен у середньому утворювати гострий кут з напрямком на оптимальний розв'язок ігрової задачі.

Умова псевдоградієнтності може бути забезпечена цілим класом рекурентних методів виду (6). Для синтезу рекурентних методів використаємо метод стохастичної апроксимації [18]. Для цього допустимо, що математичні сподівання випадкових величин відомі $M\{\xi_n^i(Z_i)\} = v^i(Z_i)$ для всіх i та Z_i. Тоді функцію середніх виграшів i–го гравця обчислюють так:

$$V^i(p) = \sum_{k_1,\dots,k_l} v^i(k_1,\dots,k_l) \prod_{j=1}^{l} p^j(k_j),$$

де $l = | Z_i |$, $i = 1,2,\dots,L$.

Побудову ігрових рекурентних методів виконаємо на основі умови доповняльної нежорсткості, яка виконується для точок рівноваги за Нешем [19] у змішаних стратегіях

$$\nabla_{p^i} V^i = g V^i e^{N_i}, \tag{8}$$

де e^{N_i} – вектор, що складається з N_i одиниць; $g \in \{0,1\}$ – допоміжний множник. Система (8) отримана розв'язуванням задачі умовної мінімізації диференційованих функцій $V^i(p)$ на одиничних симплексах $S_0^{N_i}$. Для умови доповняльної нежорсткості значення $g = 1$. Якщо $g = 0$, то з (8) отримаємо градієнтний метод.

Якщо $M\{R(p_n^i, z_n^i, \xi_n^i)\} = \nabla_{p^i} V^i(p) - g V^i e^{N_i}$, то нерівність (7) перетворюється до вигляду

$$V^i(p) - V^i(p \setminus p^i + p^{i*}) \geq 0 \tag{9}$$

що справедливо для всіх $p \in S$.

Оскільки $\nabla_{p^i} V^i(p) - g V^i e^{N_i} = M\left\{ \xi_n^i \left[\frac{e(z_n^i)}{e^{\mathrm{T}}(z_n^i) p_n^i} - g e^{N_i} \right] \middle| p_n^i = p^i \right\}$, то на основі стохастичної апроксимації отримуємо метод

$$p_{n+1}^i = \pi_{\varepsilon_{n+1}}^{N_i} \left\{ p_n^i - \gamma_n \xi_n^i \left[\frac{e(z_n^i)}{e^{\mathrm{T}}(z_n^i) p_n^i} - g e^{N_i} \right] \right\}. \tag{10}$$

Ще один метод можна одержати на основі поелементного зважування векторної умови доповняльної нежорсткості (8), якщо $g = 1$. Нехай $M\{R_n \mid p_n^i = p^i\} = diag(p^i)[V^i e^{N_i} - \nabla_{p^i} V^i]$. Тоді умова псевдоградієнтності (7) матиме вигляд

$$V^i(p) \left\| p^i - \tilde{p}^i \right\|^2 \geq 0, \tag{11}$$

де $\tilde{p}^i = \dfrac{diag(p^i) \nabla V^i}{V^i}$. Умова (11) виконується для знакододатних середовищ ($V^i(p) \geq 0$ для всіх $p \in S$), якщо $p^{i*} = \tilde{p}^i$.

Враховуючи, що $diag(p^i)(V^i e^{N_i} - \nabla_{p^i} V^i) = M\{\xi_n^i [p_n^i - e(z_n^i)] \mid p_n^i = p^i\}$, на основі стохастичної апроксимації отримаємо такий метод:

$$p_{n+1}^i = \pi_{\varepsilon_{n+1}}^{N_i} \left\{ p_n^i - \gamma_n \xi_n^i \left[e(z_n^i) - p_n^i \right] \right\}. \qquad (12)$$

Параметри методів γ_n та ε_n визначають умови та швидкість збіжності ігрового методу. Для комп'ютерного моделювання можна прийняти, що

$$\gamma_n = \gamma_0 n^{-\alpha}, \ \varepsilon_n = \varepsilon_0 n^{-\beta}, \qquad (13)$$

де $\gamma_0, \alpha, \beta > 0$, $\varepsilon_0 \in (0, \min_i N_i^{-1})$. Конкретні значення параметрів α та β визначають з умов збіжності ігрового методу.

На основі верхніх оцінок умовного математичного сподівання поточної похибки $\Delta_n = \sum_{i \in D} \left\| \Delta_n^i \right\|^2$, де $\Delta_n^i = diag(p_n^i)(e^{N_i} V_n^i - \nabla_{p^i} V_n^i)$, виконання умови доповняльної нежорсткості при фіксованій передісторії подій та наслідків теореми Роббінса–Сігмунда [16] автором отримано умови збіжності з ймовірністю 1. На основі усереднення одержаних оцінок за реалізаціями подій та результатів теореми про рекурентні числові нерівності [20] отримано умови збіжності у середньоквадратичному.

Середньоквадратичний порядок швидкості збіжності ігрових методів дорівнює $n^{-\theta}$, де $\theta = \min(1 + \beta - \alpha, \alpha - \beta)$, $\alpha \in (0,1]$; $0 < \beta < \alpha$ для методів (10) та $\theta = \min(1 + \beta - \alpha, \alpha)$, $\alpha \in (0,1]$; $\beta > 0$ для методу (12).

Оцінювання асимптотичного порядку швидкості збіжності виконано для послідовностей величин (13) методом моментів Чжуна [16]:

$$\overline{\lim_{n \to \infty}} n^{\theta} M\{\Delta_n\} \le \vartheta$$

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

Максимальний асимптотичний порядок швидкості збіжності ігрових методів у знакододатних середовищах дорівнює: $\theta_{\max} = 0.5$ при $\alpha \in (1/2, 1]$, $\beta = \alpha - 1/2$ – для методів (10); $\theta_{\max} = 1$ при $\alpha = 1$, $\beta \ge 1$ – для методу (12). Можливість досягнення максимального порядку визначається належним підбором початкових значень параметрів методів γ_0 та ε_0 (13).

Алгоритм розв'язування задачі.

Ігровий алгоритм складається з таких кроків.

Крок 1. Ініціалізація методу. Задати кількість гравців L, початкові значення параметрів ігрового методу $\gamma_0; \alpha; \varepsilon_0; \beta$, які задовольняють умови асимптотичної збіжності, та початкові значення векторів змішаних стратегій p_0^i, $i = 1, 2, \ldots, L$. Вектори змішаних стратегій можуть набувати довільні значення на одиничному ε–симплексі ($\varepsilon > 0$). В початковий момент часу доцільно задати $p_0^i(j) = 1/N_i$; $j = 1, 2, \ldots, N_i$. Задати момент часу $n = 1$.

Крок 2. Вибір чистих стратегій гравців та визначення внутрішніх станів агента. Для кожного гравця $i = 1, 2, \ldots, L$ згенерувати випадкове число ω, розподілене за рівномірним законом на відрізку [0,1], та визначити номер k_i чистої стратегії з виконання умови (4). За номером чистої стратегії визначити внутрішнє значення i–го агента з інтервалу $[a_i, b_i]$ згідно з (2).

Крок 3. Визначення поточного значення шуканого випадкового процесу. Значення V^i визначають як випадкові величини, розподілені за нормальним законом з математичним сподіванням m_i та дисперсією d_i. Нормально-розподілені випадкові величини можна обчислити через суму дванадцяти рівномірно розподілених величин:

$$V^i = m_i + \sqrt{d_i}\left(\sum_{j=1}^{12}\omega_j - 6\right),$$

де $\omega \in [0,1)$ – випадкове число, розподілене за рівномірним законом.

Крок 4. Визначення поточних втрат гравців. Поточні втрати гравців ξ_n^i визначають згідно з (1).

Крок 5. Зміна регульованих параметрів алгоритму. Обчислити значення параметрів γ_n та ε_n у момент часу n згідно з (13).

Крок 6. Перерахунок елементів векторів змішаних стратегій. Обчислення нових векторів змішаних стратегій здійснюють за одним із рекурентних перетворень (10), (12), після чого виконують їхнє проектування на одиничний ε-симплекс, яке зводиться до ітераційного алгоритму проектування вектора на одиничну гіперплощину з подальшим зануленням його від'ємних компонентів [16].

Крок 7. Перевірка умови закінчення гри. Момент закінчення гри визначається однією з таких умов: 1) коли поточна (або середня) відстань між фронтами випадкових процесів є меншою від заданої величини, тобто $L^{-1}\sum_{i=1}^{L}\xi_n^i \le r; n = 1,2,...$ (або $L^{-1}\sum_{i=1}^{L}\Phi_n^i \le r$), де $r > 0$; 2) при незначній зміні векторів змішаних стратегій за два послідовні моменти часу, тобто коли $L^{-1}\sum_{i=1}^{L}\left\|p_{n+1}^i - p_n^i\right\| \le \varepsilon$, де $\varepsilon > 0$; 3) у разі досягнення заданої кількості кроків, тобто якщо $n \ge n^*$.

Якщо умова не виконана, то виконати присвоєння $n = n+1$ і перейти на крок 2, інакше – на крок 8.

Крок 8. Виведення значень внутрішніх станів агентів z_i. Кінець.

Результати комп'ютерного моделювання. Допустимо, що множина агентів Z задається цілочисловими дискретними координатами (або індексами) та значеннями на площині $Z = \{z_{i,j} \mid z_{i,j} \in R^1; i,j = 1,2,...,L\}$. Розділимо Z на підмножини X_i та Y_j такі, що $Z_{i,j} = X_i \cup Y_j$, $Z_{i,j} \ne \varnothing$, $X_i \cap Y_j = z_{i,j}$. Нехай вихідна функція–характеристика елемента (i,j) має вигляд:

$$F(Z_{i,j}) = F(X_i) + F(Y_j),$$

де $F(X_i) = \sum_{j=1}^{L} z_{i,j}$, $F(Y_j) = \sum_{i=1}^{L} z_{i,j}$.

Для формування поточних втрат $\xi_n^{(i,j)}$ значення $F(Z_{i,j})$ порівнюється з $V_{i,j} = V_i + V_j$, де $V_i = Normal(m_i, d_i)$ – нормально розподілена випадкова величина з математичним сподіванням m_i та дисперсією d_i. З (1) отримаємо

$$\xi^{i,j} = \frac{1}{2}\left|F(Z_{i,j}) - V_{i,j}\right|. \qquad (14)$$

Нехай усі гравці мають однакову кількість чистих стратегій $N_i = N$, які є цілочисловими значеннями з інтервалу $[a,b]$. Для забезпечення можливості розв'язання пошукової задачі значення математичного сподівання повинно потрапляти в область значень функції $F(Z_{i,j})$:

$$2La + (d_i + d_j) \le m_i + m_j \le 2Lb - (d_i + d_j).$$

Якщо $m_i = m_j = m$ та $d_i = d_j = d$ для усіх i та j, то у разі мінімізації (14) отримаємо стохастичний варіант задачі побудови магічного квадрата. У магічному квадраті суми елементів кожного рядка та стовпчика є однаковими.

З постановки задачі випливає, що пошукова задача має неоднозначний розв'язок. Часткові розв'язки, одержані за допомогою ігрового методу (12), якщо *m=100, d=0* , $N = 30$, *L=6,* наведено на рис. 1 .

16	*13*	*17*	*18*	*25*	*11*	*100*
6	*16*	*18*	*25*	*19*	*16*	*100*
14	*19*	*27*	*14*	*4*	*22*	*100*
19	*14*	*13*	*15*	*18*	*21*	*100*
21	*20*	*12*	*16*	*16*	*15*	*100*
24	*18*	*13*	*12*	*18*	*15*	*100*
100	*100*	*100*	*100*	*100*	*100*	

24	*13*	*15*	*25*	*7*	*16*	*100*
15	*10*	*9*	*22*	*17*	*27*	*100*
14	*20*	*10*	*15*	*19*	*22*	*100*
27	*15*	*8*	*15*	*13*	*22*	*100*
11	*14*	*30*	*16*	*26*	*3*	*100*
9	*28*	*28*	*7*	*18*	*10*	*100*
100	*100*	*100*	*100*	*100*	*100*	

Рис. 1. Варіанти магічних квадратів

Динаміка збіжності ігрового методу (12) для значень параметрів $\gamma_0 = 1; \alpha = 0.1; \varepsilon_0 = 0.999 / N_i; \beta = 2$ зображена на рис. 2 у вигляді графіків функції середніх втрат Φ_n^i (5) для різних значень дисперсії d. При зростанні дисперсії швидкість збіжності ігрового методу зменшується.

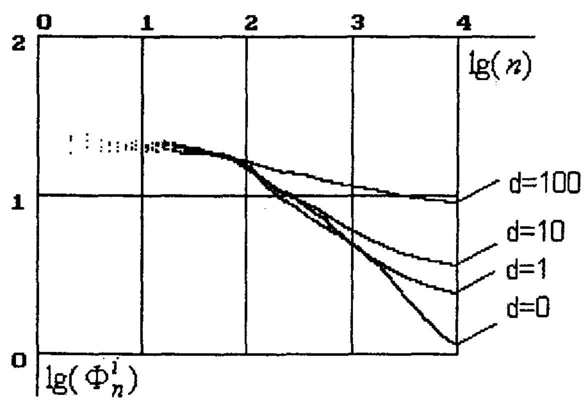

Рис. 2. Зміна функції середніх втрат у часі

Ігровими методами (10) та (12) формують послідовності чистих стратегій методом проб та помилок, що загалом пояснює їхню невисоку (степеневу) швидкість збіжності.

Момент закінчення гри визначається інтегральною оцінкою внутрішніх станів агентів

$$n_{out} = (n \mid L^{-2} \sum_{(i,j)} \xi_n^{(i,j)} \le r; n = 1, 2, ...).$$

Оскільки n_{out} є випадковою величиною, то момент закінчення гри визначимо значенням, усередненим за $T = 100$ реалізаціями цієї випадкової величин:

$$\overline{n} = \frac{1}{T} \sum_{k=1}^{T} n_{out}(k).$$

На рис. 3 наведено залежність середньої кількості кроків $\overline{n}$ до закінчення гри від значення дисперсії d для таких значень параметрів ігрового методу: *m=100; r=1;* $N = 30$; *L=6.* Для відстеження тренду виконано лінійну середньоквадратичну апроксимацію результатів експерименту.

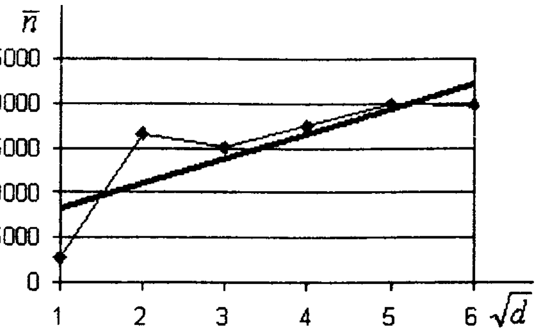

Рис. 3. Залежність моменту закінчення гри від дисперсії

З рис. 3 видно, що при збільшенні значення дисперсії середня кількість кроків, необхідна для розв'язування пошукової задачі, має тенденцію до зростання.

Залежність моменту закінчення гри від кількості агентів L^2 зображено на рис. 4 для значень параметрів $m=100$; $d=0$; $r=1$; $N=30$.

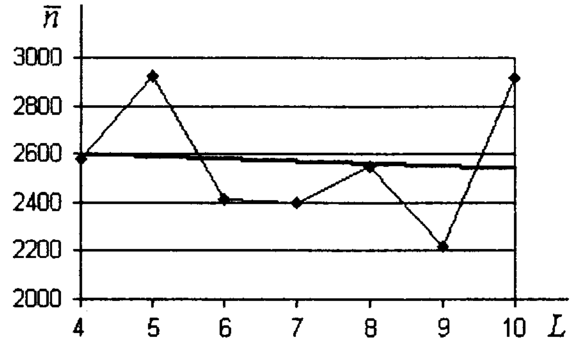

Рис. 4. Залежність моменту завершення гри від кількості агентів

У разі збільшення кількості агентів L загальна кількість станів гри зростає за показниковим законом N^{L^2}. Однак втрати гравця (i, j) мають локальну залежність тільки від стратегій гравців, розміщених в i-му рядку та j-му стовпчику магічного квадрата. Кількість локальних станів гри становить $2N^L$. Хоча така кількість станів теж визначається показниковим законом, але з рис. 4 видно, що зростання кількості гравців не має значного впливу на ефективність ігрового методу – середня кількість кроків закінчення гри знаходиться майже на одному рівні. Це пояснюється багатозначністю розв'язків ігрової задачі. У разі зростання кількості гравців кількість можливих розв'язків задачі побудови магічного квадрата теж збільшується, що компенсує зростання кількості комбінованих станів гри.

Вплив кількості чистих стратегій гравців на ефективність ігрового методу зображено у вигляді графіка на рис. 5 для значень параметрів $m=100$; $d=0$; $r=1$; $L=6$.

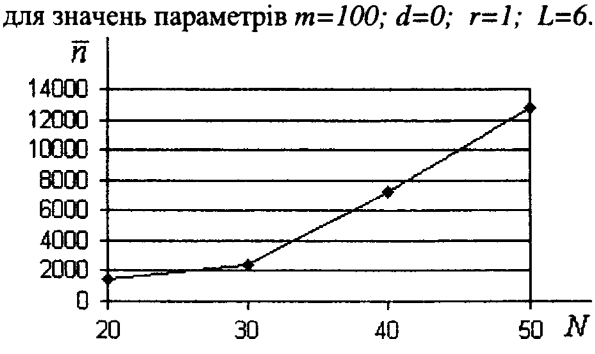

Рис. 5. Залежність моменту закінчення гри від кількості чистих стратегій

При зростанні кількості чистих стратегій пошуковий простір розширюється за степеневим законом, але кількість можливих розв'язків задачі побудови магічного квадрата практично не змінюється. В результаті зростає кількість непродуктивних варіантів пошуку, що відповідно призводить до зростання кількості кроків розв'язування ігрової задачі.

Замість очікуваного більшого впливу на збіжність ігрового методу кількості гравців (показниковий закон), ніж кількості чистих стратегій (степеневий закон), у зв'язку із специфікою розв'язуваної задачі маємо зворотний ефект.

Як видно з рис. 4 та 5, ігрові методи є ефективнішими від методу повного перебору варіантів у детермінованих середовищах при $d=0$, або методу рівноімовірного вибору варіантів у стохастичних середовищах при $d>0$, оскільки потребують значно меншої кількості кроків, ніж N^{L^2}. У стохастичних середовищах перевага ігрових методів безсумнівна, оскільки методи послідовного перебору варіантів не працюють. Ефективність розглянутих ігрових методів стосовно інших методів стохастичної оптимізації вимагає окремого дослідження.

Така особливість роботи ігрових методів пояснюється їхніми адаптивними властивостями на основі самонавчання. Ігрові методи добре справляються із ситуаціями невизначеності стохастичних середовищ і, як показують результати додаткових досліджень, практично не залежать від виду закону розподілу випадкових величин.

Висновки. Ігрова модель стохастичної гри є корисним та потужним інструментом для вивчення процесів самоорганізації МАС. В реактивних середовищах самоорганізація забезпечується цілим класом рекурентних ігрових методів, побудованих на основі стохастичної апроксимації векторної функції приросту змішаних стратегій, яка задовольняє умову псевдоградієнтності з напрямком на точку оптимального розв'язку задачі.

На практиці самоорганізація гри проявляється у селективному виборі станів, які забезпечують колективну мінімізацію функцій середніх втрат агентів. Властивість самоорганізації гри незалежних агентів з індивідуальним оцінюванням стратегій при усій строгості математичних перетворень містить у собі елемент загадковості. Гравці з однаковими початковими умовами за відсутності обміну інформації та стохастичної невизначеності середовища з часом навчаються вибирати такі (у загальному різні) стани, які задовольняють сформульовану умову оптимальності. Причина цього полягає в механізмах адаптивного ігрового пошуку на основі динамічних векторів змішаних стратегій. Властивість самоорганізації гри МАС перевірено та підтверджено результатами комп'ютерного моделювання, які можуть бути використані для стохастичної оптимізації розподілених систем в умовах невизначеності або побудови елементів розподілених інтелектуальних систем.

Відкритими для дослідження є проблеми організації обміну даними між ігровими агентами, асинхронної взаємодії агентів та моделювання гри інтелектуальних агентів з елементами когнітивних підсистем формування та прийняття рішень.

1. Gerhard Weiss and Sandip Sen, editors. Adaptation and Learning in Multiagent Systems. Springer Verlag, Berlin, 1996. 2. Stone P. Layered Learning in Multiagent Systems. MIT Press, 2000. 3. Wooldridge M. An Introduction to Multiagent Systems. John Wiley & Sons (Chichester, England), 2002. 4. Vittikh V.A., Skobelev P.O. Multy-agent systems for modeling of self-organization and cooperation processes. // Proceedings of the XIII International Conference on the Application of Artificial Intelligence in Engineering. Galway. 1998, pp. 91 – 96. 5. Хакен Г. Информация и самоорганизация. Макроскопический подход к сложным системам. М.: Мир, 1991. 6. Бурков В.Н., Новиков Д.А. Теория активных систем: состояние и перспективы. – М.: СИНТЕГ, 1999. 7. Великий тлумачний словник сучасної української мови/ Уклад. і гов. ред. В.Т.Бусел. – К.; Ірпінь: ВТФ "Перун", 2001. 8. Доманский В.К. Стохастические игры // Математические вопросы кибернетики. – 1988. – № 1. – С. 26–49. 9. Fudenberg, D., Levine, D.K.: The Theory of Learning in Games. MIT Press, 1998. 10. Gintis H. Game Theory Evolving. Princeton University Press, 2000. 11. Korf R. E. A Simple Solution to Pursuit Games // Proceedings of the Eleventh International Workshop on Distributed Artificial Intelligence, Glen Arbor, Michigan, February 1992, pp. 183–194. 12. Stephens L. M., Merx M. The Effect of Agent Control

Strategy on the Performance of a DAI Pursuit Problem // Proceedings of the Tenth International Workshop on Distributed Artificial Intelligence, Bandera, Texas, October 1990, pp. 263–292. 13. Растригин Л.А., Рипа К.К., Тарасенко Г.С. Адаптация случайного поиска. – Рига, 1973. 14. Цыпкин Я.З. Основы теории обучающихся систем. – М., 1970. 15. Срагович В.Г. Адаптивное управление. – М., 1981. 16. Назин А.В., Позняк А.С. Адаптивный выбор вариантов: Рекуррентные алгоритмы. – М., 1986. 17. Катренко А.В. Дослідження операцій. Підручник. – Львів, 2004. 18. Вазан М. Стохастическая аппроксимация. – М., 1972. 19. Мулен Э. Теория игр с примерами из математической экономики. – М., 1985. 20. Невельсон М.Б., Хасьминский Р.З. Стохастическая оптимизация и рекуррентное оценивание. – М., 1972.

УДК 621

В.Т. Кремінь, А.С. Сметана
Національний університет "Львівська політехніка",
кафедра електронних обчислювальних машин

МЕТОД ПРЯМОЇ КОРЕКЦІЇ ПОМИЛОК ДЛЯ РАДІОМОДЕМІВ WIRELESS USB ФІРМИ CYPRESS SEMICONDUCTOR

Запропонована реалізація завадостійкого кодування для радіомодемів WirelessUSB виконана на мікроконтролері PSoC. Показано, як можна використовувати апаратні можливості цих радіомодемів для виправлення помилок прийому, спричинених завадами у радіоканалі. Описано розроблений метод прямого корегування помилок, а також наведено його алгоритмічне та апаратне виконання.

The noise resistant coding implementation for radiomodems CYPRESS WirelessUSB is proposed using PSoC microcontrollers. It is shown, like the hardware abilities of these radio modems can be used to correct receiving errors, caused by interference in radio channel. The developed forward error correction code was described and its algorithmic and hardware implementations were given.

Вступ. Під час розробки систем передачі даних особлива увага приділяється завадостійкості лінії і достовірності прийнятих даних. Достовірність прийнятих даних характеризує ймовірність спотворення для кожного переданого біта (середній час напрацювання машини на один збій – Bit Error Rate, BER). Величина BER для каналів зв'язку без додаткових засобів захисту від помилок (наприклад, самокоригувальних кодів) становить, як правило, 10^{-4} – 10^{-6}. Завадостійкість лінії визначає її здатність зменшувати рівень завад, які створюються навколишнім середовищем, на внутрішніх провідниках. Завадостійкість лінії залежить від типу використовуваного фізичного середовища, а також від екранувальних і тих, що придушують завади, засобів самої лінії [7]. Найменш завадостійкими є радіолінії. Тому, передаючи дані по радіоканалу, необхідно вживати додаткові заходи для підвищення достовірності прийнятих даних.

Одним із можливих виходів із цієї ситуації є розробка і використання методів завадостійкого кодування (самокоригування кодування), які дають змогу виправляти прийняті дані на боці приймача без запитів повторної передачі. Такий підхід особливо зручний при симплексному режимі зв'язку. Використання кодування з виправленням помилок, звичайно ж, веде до надлишковості