С. 48 – 55. 3. Савченко Т. Н. Применение методов кластерного анализа для обработки данных психологических исследований // Экспериментальная психология, 2010, том 3, № 2, с. 67–85. 4. Aravind H., C. Rajgopal, K. P. Soman. A Simple Approach to Clustering in Excel // International Journal of Computer Applications (0975 – 8887), Volume 11 – No.7, December 2010. 5. Горчаков А.А. Математический апарат для инвестора // Аудит и финансовій анализ. – III кв. 1997. – С. 1 – 57. 6. Кокс Д.Р., Оукс Д. Анализ данных типа времени жизни / Пер. с англ. О.В. Селезнева. – М.: инансы и статистика, 1988. – 191 с. 7. Вадзинский Р.Н. Справочник по вероятностным распределениям. – Спб.: Наука, 2001. – 295 с. 8. Таблицы по математической статистике / П. Мюллер, П. Нойман, Р. Шторм; пер. с нем. и предисл. В.М. Ивановой. – М.: Финансы и статистика, 1982. – 278 с. 9. Уиллиамс У.Т., Ланс Дж. Н. Методы иерархической классификации. - в кн. Статистические методы для ЭВМ / Под ред. Э. Рэлстона, Г.С. Уилфа: Пер с англ./ Под ред. М.Б. Малютова. – М.: Наука, 1986. – 464 с. 10. Бутвиловский А.В. Основные методы молекулярной эволюции: монография / А.В. Бутвиловский, Е.В. Барковский, В.Э. Бутвиловский, В.В. Давыдов, Е.А. Черноус, В.В. Хрусталев; под общ. ред. проф. Е.В. Барковского. – Мн.: 2009. – 210 с.

004.032.26; 004.852; 004.94

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

МОДЕЛЬ СТОХАСТИЧНОЇ ГРИ НЕЙРОАГЕНТІВ

Розроблено нейроагентну ігрову модель колективного прийняття рішень в умовах невизначеності. Виконано формулювання стохастичної гри та для її розв’язування використано адаптивні методи навчання штучних нейронних мереж без учителя. Розроблено ігровий алгоритм та програмну модель нейроагентного прийняття рішень. Збіжність стохастичної гри нейроагентів підтверджено результатами комп’ютерного експерименту. Досліджено впливи параметрів ігрової моделі на швидкість навчання нейроагентів.

Ключові слова: колективне прийняття рішень, умови невизначеності, стохастична гра нейроагентів, адаптивні методи навчання.

The neuroagent game model of collective decision-making in the conditions of uncertainty is developed. The formulation of stochastic game is executed. Adaptive learning methods of artificial neural networks without the teacher are used for the game solving. The convergence of neuroagent stochastic game is confirmed by results of computer experiment. Influences of parameters of game model on the neuroagent learning rate are investigated.

Key words: collective decision making, uncertainty conditions, neuroagent stochastic game, adaptive learning methods.

Вступ до проблеми колективного прийняття рішень в умовах невизначеності

Для розв’язування задач розподіленого керування та прийняття рішень у технічних, економічних, інформаційних та соціальних системах існує необхідність у колективному виборі варіантів рішень, що задовольняють одну з умов багатокритеріальної оптимальності, наприклад, Неша, Парето тощо [1–7]. Ці умови тією чи іншою мірою визначають ступінь вигідності та справедливості колективно досягнутого рішення.

Формальна модель системи колективного вибору варіантів рішень задається керованим середовищем та керуючими агентами прийняття рішень. Модель середовища визначається структурою та функціями керованої системи. Середовище сприймає керуючі рішення агентів і формує вихідні сигнали, які інтерпретуються агентами як оцінки ефективності реалізованих ними рішень. Агенти отримують оцінки від середовища, вибирають і реалізують варіанти рішень так, щоб максимізувати власні виграші або мінімізувати програші. Такий підхід до побудови системи прийняття рішень називається оптимізаційним [7]. На відміну від ситуаційного підходу, коли рішення отримується одразу на основі вхідних даних, за оптимізаційним підходом рішення уточнюється через ланку зворотного зв'язку.

Агент – це інтелектуальна система прийняття рішень, програмна або технічна реалізація моделі особи, що приймає рішення. Узагальнено агент складається з таких основних підсистем: рецепторної, інтелектуальної та ефекторної. Рецепторна підсистема сприймає сигнали від середовища прийняття рішень, інтелектуальна – виробляє рішення на основі поточної інформації, самонавчання та прогнозування, ефекторна – забезпечує реалізацію вироблених рішень відповідним впливом на середовище.

Колективне вироблення рішень у розподіленому середовищі здійснюється множиною взаємодіючих інтелектуальних агентів, яка називається мультиагентною системою (МАС) [8 – 12]. Кожен агент виробляє рішення на основі доступної йому локальної інформації та взаємодії у межах МАС. Як активний елемент МАС, агент здатний сприймати та аналізувати дані інформаційної мережі, вести переговори та обмінюватися поточною інформацією з іншими агентами, приймати автономні рішення, змінювати стани мережі, інформувати систему та користувача про результати своїх дій.

Інтелектуальні агенти мають такі базові властивості:

1) автономність – переслідуючи поставлену мету, агент може самостійно приймати рішення на основі власних знань, не потребуючи зовнішнього керування;

2) реактивність – агент може сприймати зовнішню інформацію, виробляти та реалізовувати адекватні дії;

3) інтелектуальність – здатність агента навчатися, адаптуватися до змін середовища, опрацьовувати отримані дані методами штучного інтелекту, приймати оптимальні рішення;

4) спеціалізація – як правило, агенти виконують вузькоспеціалізовані функції;

5) мобільність – для досягнення мети програмний агент може переміщуватися у межах інформаційній мережі;

6) координація – розподіл ролей та узгодження дій агентів при розв'язування спільної задачі;

7) інтерактивність – агент взаємодіє з іншими агентами, з інформаційними ресурсами мережі та із користувачем;

8) комунікативність – здатність агентів спілкуватися технічною мовою та розуміти один одного;

9) персональність – унікальні якості агента, що моделюють його психологічні риси характеру, поточні емоції тощо;

10) децентралізація – кожен з агентів не має повного уявлення про всю систему, і тому відсутні агенти, які керують усією системою.

Під час розв'язування поставленої задачі між агентами виникають явища кооперації або конкуренції. Як відомо, ці види взаємодії між агентами в умовах невизначеності вивчає теорія стохастичних ігор [13–15]. Тому для дослідження колективної поведінки агентів під час прийняття рішень доцільно використати математичний апарат теорії стохастичних ігор.

Характерними особливостями ігрового прийняття рішень у МАС є:

1) розподіленість або багатопараметричність середовища прийняття рішень;

2) внутрішня стохастичність середовища;

3) повна або часткова відсутність апріорної інформації (невизначеність) про середовище прийняття рішень;

4) керованість середовища та можливість розподіленої реалізації варіантів керування;

5) багатокритеріальність керування або прийняття рішень;

6) дискретність та скінченність множини варіантів прийняття рішень;

7) стохастична незалежність вибору варіантів рішень у просторі та часі;

8) можливість багатократного повторення реалізацій варіантів рішень у часі;

9) розподілений локально-залежний характер формування та збирання інформації для статистичної ідентифікації середовища прийняття рішень;

10) можливість застосування розподіленого ігрового алгоритму, який забезпечує досягнення області компромісних рішень;

11) реалізація ігрового алгоритму в реальному масштабі часу;

12) визначення моментів зупинки ігрового алгоритму для можливості його практичного застосування.

Функціонування МАС прийняття рішень здійснюється в умовах апріорної невизначеності [16]. Невизначеність може бути спричинена внутрішніми або зовнішніми факторами МАС. Виділимо такі види невизначеностей:

1) структурна – невідомий склад системи та зв'язки між її елементами;

2) алгоритмічна – невідомий алгоритм функціонування системи;

3) інформаційна – нечіткість, відсутність повної інформації, необхідної для прийняття рішень;

4) лінгвістична – неоднозначність висловлювання при обміні повідомленнями між агентами;

5) цільова – невідома глобальна мета функціонування системи;

6) соціальна – зумовлена колективною взаємодією агентів, коли дії одного з агентів впливають на вибір рішень іншими агентами;

7) стохастична – вплив на систему неконтрольованих зовнішніх факторів.

У наукових дослідженнях невизначеність прийняття рішень найчастіше моделюється за допомогою механізму випадкових величин, покладеного в основу стохастичної невизначеності.

Часткове компенсування невизначеності забезпечується здатністю агентів до самонавчання та адаптивними стратегіями прийняття рішень.

У кібернетичній літературі добре дослідженими є ігрові методи самонавчання, що грунтуються на адаптивному формуванні імовірнісних розподілів дискретних варіантів рішень (чистих стратегій гравців) [14]. Після завершення вибору та реалізації чистих стратегій усіма гравцями кожен з них отримує від середовища поточний виграш, який використовується для зміни динамічних змішаних стратегій (векторів умовних імовірностей вибору варіантів рішень), покладених в основу механізму генерування випадкових чистих стратегій. Ця зміна полягає у тому, що імовірність вибору чистої стратегії зростає пропорційно до значення отриманого поточного виграшу. Побудований на основі стохастичної апроксимації метод перебудови змішаних стратегій з часом забезпечує максимізацію функцій середніх виграшів на одиничних симплексах. Інтелектуальні можливості агентів такої гри є обмеженими, оскільки вони моделюють лише рефлексорну поведінку біологічних систем.

Для побудови інтелектуальних систем прийняття рішень можна використати штучні нейронні мережі (ШНМ) як моделі процесів нервової діяльності біологічних систем зі здатністю до запам'ятовування, аналізу та прогнозування поведінки [17–20]. Нейронні мережі реалізують „м'які" обчислення за зразком процесів, що проходять у мозку людини і використовуються як моделі об'єктів з невідомими характеристиками. ШНМ складаються з множини нейронів та зв'язків між ними. Навчання ШНМ полягає у корегуванні синаптичних зв'язків між нейронами на основі інформації, яка надходить до нейромережі з навколишнього середовища. Для отримання необхідної структури зв'язків одні зв'язки між нейронами підсилюються, а інші – послаблюються.

Система прийняття рішень, основана на ШНМ, називається нейроагентом. Нейроагентні ігрові моделі прийняття рішень в умовах невизначеності є недостатньо вивченими у сучасній фаховій науковій літературі.

Застосування нейроагентних ігрових моделей є перспективним напрямом підвищення ефективності процесів колективного вироблення та прийняття рішень в умовах невизначеності завдяки таким особливостям:

1) нелінійність – нейронні мережі дають змогу отримувати нелінійну залежність вихідного сигналу від вхідного;

2) адаптивність – нейронні мережі мають здатність пристосовувати свої синаптичні ваги до змін навколишнього середовища;

3) пластичність та стійкість до відмов – нейронні мережі зберігають інформацію у розподіленому по всіх зв'язках нейронної мережі вигляді, і вихід з ладу одного або декількох нейронів не призводить до відмови всієї системи;

4) універсальність – нейронні мережі не потребують спеціального програмування, оскільки вони дають змогу розв'язувати різні задачі опрацювання інформації за однаковими алгоритмами навчання нейронів.

Метою цієї роботи є розроблення моделі стохастичної гри нейроагентів для колективного прийняття рішень в умовах невизначеності. Для досягнення мети необхідно: виконати постановку задачі ігрового прийняття рішень в умовах невизначеності; задати середовище колективного прийняття рішень; розробити структуру нейроагентів; вибрати метод навчання нейроагентів для розв'язування сформульованої задачі; розробити алгоритм та програмні засоби моделювання стохастичної гри нейроагентів; проаналізувати отримані результати та виробити рекомендації для їх практичного використання.

Постановка ігрової задачі прийняття рішень в умовах невизначеності

Матрична стохастична гра $\Gamma = (I,\ \{U_i\}_{i\in I},\ [v_i]_{i\in I})$ прийняття рішень задається:

1) множиною агентів $I = \{i \mid i = 1..L\}$, де $L = |I|$ – потужність множини або кількість гравців;
2) множинами чистих стратегій агентів $U_i = \{u_i(1), u_i(2), ..., u_i(N_i)\}$, де N_i – кількість чистих стратегій гравця з номером i;
3) матрицями середніх виграшів агентів $[v_i(u)]$ $\forall i \in I$, $\forall u \in U$, де $u \in U = \underset{i\in I}{\times} U_i$ – колективні стратегії гравців.

Стохастична гра розгортається у дискретні моменти часу $t = 1,2,...$. Після реалізації колективної стратегії $u_t = u$ кожен агент отримує випадковий поточний виграш $x_{i,t}(u) \in R^1$ з апріорі невідомим математичним сподіванням $M\{x_{i,t}(u)\} = v_i(u)$ та обмеженою дисперсією $d_i(u) < \infty$. Тут і далі індекс i позначає номер гравця, а індекс t – поточний момент часу.

Отримані поточні виграші агентів усереднюються у часі для оцінювання ефективності процесу прийняття рішень кожним агентом:

$$\Xi_{i,t}(\{u_t\}) = t^{-1} \sum_{t=1}^{t} x_{i,t} \quad \forall i \in I . \tag{1}$$

Метою агентів є максимізація їх середніх платіжних функцій:

$$\varliminf_{t\to\infty} \Xi_{i,t} \to \max_{u_i} \quad \forall i \in I . \tag{2}$$

Отже, на основі спостережень поточних виграшів $x_{i,t}$ кожен агент $i \in I$ повинен вибирати поточні рішення $u_{i,t} = u_i \in U_i$ так, щоб з часом $t = 1,2,...$ забезпечити максимізацію системи цільових функцій (1).

Розв'язок багатокритеріальної задачі (2) слід шукати у множині точок колективної рівноваги (наприклад, Неша) або оптимальності (наприклад, Парето) залежно від способу вибору агентами послідовностей варіантів рішень.

Ігровий нейроагентний метод розв'язування задачі

Відомі адаптивні методи генерування послідовностей $\{u_{i,t}\}\ \forall i \in I$, $t = 1,2,...$ на основі динамічних розподілів дискретних випадкових величин, у ролі яких використовуються змішані стратегії гравців [14]. На відміну від цього, розглянемо нейроагентний метод розв'язування матричної стохастичної гри, схему якої зображено на рис. 1.

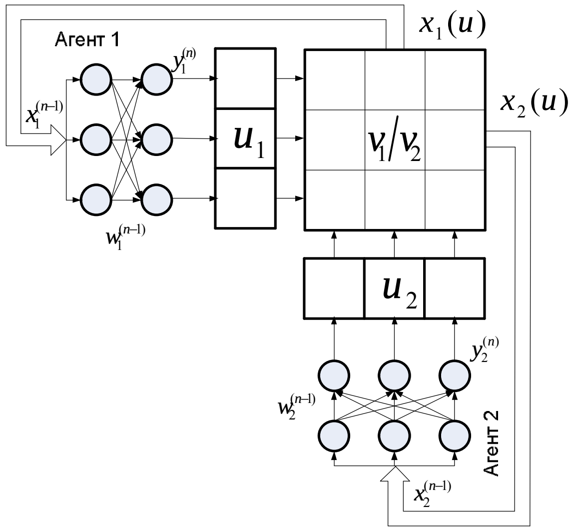

Рис. 1. Структура матричної гри нейроагентів

Модель середовища прийняття рішень задається матрицями математичних сподівань випадкових виграшів $v_i\ \ \forall i \in I$. На вхід середовища подаються значення варіантів рішень $u_i \in U_i$. Виходом середовища є відповідні значення поточних виграшів $x_i(u)$.

Кожен нейроагент задається штучною мережею з $n = 2$ шарами нейронів. Кількість елементів кожного шару є однаковою і дорівнює кількості варіантів рішень $N_i = |U_i|$. На входи нейроагентів подаються вектори параметрів $x_i^{(n-1)}$, обчислені на основі виходів середовища $x_i(u)\ \ \forall i \in I$. Виходами нейроагентів є вектори параметрів $y_i^{(n)}$, на основі яких визначаються варіанти рішень $u_i \in U_i\ \ \forall i \in I$. Ваги $w_i^{(n-1)}$ позначають величину синаптичних зв'язків між нейронами i-го агента. Додатні значення ваг зв'язків відповідають збуджувальним, а від'ємні – гальмівним синапсам. Нульове значення ваг означає відсутність зв'язку між нейронами.

Функціонування нейроагента здійснюється за одним із адаптивних алгоритмів навчання без учителя, наприклад, Хебба, Кохонена або за іншим [19]. Навчання без учителя, або самонавчання, за своєю природою найближче до свого біологічного прототипу – мозку. Самонавчання не орієнтується на наявність правильних виходів нейромережі. Алгоритм самонавчання самостійно виявляє внутрішню структуру вхідних даних, перебудовуючи ваги синаптичних зв'язків так, щоб близькі (за деякою метрикою) набори вхідних сигналів викликали достатньо близькі вихідні набори сигналів. Фактично процес самонавчання нейромережі розв'язує задачу кластеризації даних, виявляючи статистичні властивості навчальних множин та групуючи подібні вихідні множини у кластери. Подаючи на вхід навченої нейромережі вектор із заданого класу, отримаємо характерний для цього класу вихідний вектор. Вихідний вектор наперед не відомий. Його формування зумовлено структурою навчальної вибірки, випадковим розподілом початкових значень ваг зв'язків між нейронами та комбінацією збуджених нейронів вихідного шару нейромережі.

Нейроагенти здійснюють випадковий вибір варіантів рішень $u_{i,t} \in U_i$ незалежно між собою $\forall i \in I$ та у часі $t = 1,2,...$. Для цього кожен нейроагент будує вектор умовних імовірностей $p_{i,t}$

вибору варіантів рішень $u_{i,t}$ проектуванням вектора виходів $y_{i,t}^{(n)}$ на одиничний N_i-вимірний e-симплекс:

$$p_{i,t}(u_{i,t} \mid u_{i,t}, x_{i,t}, t = 1,2,...,t-1) = p_{e_t}^{N_i}\{y_{i,t}^{(n)}\} \quad \forall i \in I, \tag{3}$$

де $p_{e_t}^{N_i}$ – проектор на одиничний e-симплекс $S_{e_t}^{N_i} \subseteq S^{N_i}$ [14]. Параметр e_t регулює швидкість розширення e-симплексу $S_{e_t}^{N_i}$ до одиничного симплексу S^{N_i} і може бути використаний як додатковий фактор керування збіжністю ігрового нейроагентного методу прийняття рішень.

Отриманий вектор імовірностей використовується для побудови емпіричного розподілу дискретних випадкових величин, на основі якого вибираються варіанти рішень:

$$u_{i,t} = \left\{ u_i[k] \middle| k = \arg\left(\min_k \sum_{j=1}^{k} p_{i,t}(u_{i,t}[j]) > w \right), k = 1..N_i \right\} \quad \forall i \in I, \tag{4}$$

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

Реакцією середовища прийняття рішень на вибраний варіант є значення випадкової величини з невідомим розподілом Z, яка інтерпретується як поточний виграш агента:

$$x_i(u_t) \sim Z(v_i(u_t), d_i(u_t)),$$

де $v_i(u_t)$ – математичне сподівання, $d_i(u_t)$ – дисперсія.

Отримані поточні виграші $x_{i,t}(u_t)$ подаються на входи нейроагентів

$$x_{i,t}^{(n-1)} = e x_i(u_t) \quad \forall i \in I, \tag{5}$$

де $e = (1 \mid \forall u_i \in U_i)$ – вектор, усі елементи якого дорівнюють одиниці.

За потреби здійснюється нормалізація елементів вектора $x_{i,t}^{(n-1)}$, наприклад, така:

$$x_{i,t}^{(n-1)} = e x_i(u_t) / |x_{\max}| \quad \forall i \in I,$$

де $x_{\max}$ – максимальне значення поточних виграшів. Нормалізація може призвести до зменшення кількості кроків, необхідних для навчання нейроагента.

Сумарні входи $\overline{x}_{i,t}^{(n)}$ нейронів n-го шару обчислюються на основі виходів $y_{i,t}^{(n-1)}$ нейронів $(n-1)$-го шару:

$$\overline{x}_{i,t}^{(n)}[k] = \sum_{j=1}^{N_i} w_{i,t}^{(n-1)}[j,k] y_{i,t}^{(n-1)}[j] \quad \forall i \in I, \; k = 1..N_i, \tag{6}$$

де $w_{i,t}^{(n-1)}[N_i, N_i]$ – матриця ваг зв'язків між вузлами нейромежі, обчислених у момент часу t. Тут $w_{i,t}^{(n-1)}[j,k]$ позначає вагу зв'язку між j-м вузлом $(n-1)$-го шару та k-м вузлом n-го шару.

Для обчислення виходів $y_{i,t}^{(n)}$ нейроагента використовується передатна функція $j()$ нейрона:

$$y_{i,t}^{(n)}[k] = j(\overline{x}_{i,t}^{(n)}[k]), \tag{7}$$

де $k = 1..N_i$.

Залежно від розв'язуваної задачі та виду нейромережі передатна функція $j()$ може бути пороговою, лінійною з насиченням, сигмоїдальною, синусоїдальною, радіально-симетричною тощо.

Найчастіше для моделювання штучної нейронної мережі використовують лінійну

$$y_{i,t}^{(n)}[k] = \begin{cases} 0, & if \;\; \overline{x}_{i,t}^{(n)}[k] \le q, \\ h(\overline{x}_{i,t}^{(n)}[k] - q), & if \;\; \overline{x}_{i,t}^{(n)}[k] > q \end{cases} \tag{8}$$

та біполярну сигмоїдальну

$$y_{i,t}^{(n)}[k] = -0.5 + 1/(1 + \exp(-h(\overline{x}_{i,t}^{(n)}[k] - q))) \tag{9}$$

передатні функції. Параметр $h > 0$ задає тангенс кута нахилу прямої для лінійної передатної функції та рівень крутості для сигмоїдальної передатної функції. Параметр $q \geq 0$ визначає поріг активації нейрона.

Навчання нейроагента здійснюється зміною ваг $w_{i,t}^{(n-1)}$ синаптичних зв'язків між нейронами. Перерахунок ваг зв'язків здійснюється за сигнальним методом Хебба, методом Кохонена або за іншим методом навчання без учителя.

Навчання за методом Хебба підсилює зв'язки між збудженими нейронами:

$$w_{i,t}^{(n-1)}[j,k] = w_{i,t-1}^{(n-1)}[j,k] + g_t(y_{i,t-1}^{(n-1)}[j] * y_{i,t-1}^{(n-1)}[k]) , \quad j = 1..N_i , \; k = 1..N_i , \qquad (10)$$

де g_t – параметр кроку навчання нейроагента.

Збудженими називаються нейрони, для яких значення сумарного входу $x_i^{(n)}$ перевищує поріг активації q .

Навчання за диференційним методом Хебба призводить до підсилення зв'язків між тими нейронами, виходи яких змінилися найбільше:

$$w_{i,t}^{(n-1)}[j,k] = w_{i,t-1}^{(n-1)}[j,k] + g_t(y_{i,t}^{(n-1)}[j] - y_{i,t-1}^{(n-1)}[j]) * (y_{i,t}^{(n-1)}[k] - y_{i,t-1}^{(n-1)}[k]) , \; j = 1..N_i , \; k = 1..N_i . \qquad (11)$$

Навчання за методом Кохонена основується на механізмі конкуренції, суть якого полягає у мінімізації різниці між вхідними сигналами нейрона-переможця, що надходять з виходів нейронів попереднього шару, та ваговими коефіцієнтами його синапсів:

$$w_{i,t}^{(n-1)}[j,k_i^*] = w_{i,t-1}^{(n-1)}[j,k_i^*] + g_t(y_{i,t-1}^{(n-1)}[j] - w_{i,t-1}^{(n-1)}[j,k_i^*]) , \; j = 1..N_i , \qquad (12)$$

де k_i^* – індекс нейрона-переможця i -го агента.

На відміну від методу Хебба, за якого одночасно можуть збуджуватися декілька нейронів одного шару, за методом Кохонена нейрони одного шару змагаються між собою за право активації. Це правило відоме у літературі з машинного навчання за назвою "переможець забирає все".

За методом Кохонена перебудова ваг зв'язків здійснюється тільки для нейрона-переможця. Переможцем є той нейрон, значення синапсів якого максимально подібні на вхідний образ.

Визначення нейрона-переможця здійснюється обчисленням відстані між векторами $y_{i,t-1}^{(n)}$ та $w_{i,t-1}^{(n-1)}$:

$$D_{i,t-1}[k] = \sqrt{\sum_{j=1}^{N_i} \left(y_{i,t-1}^{(n-1)}[j] - w_{i,t-1}^{(n-1)}[j,k] \right)^2} , \; j = 1..N_i .$$

Перемагає нейрон з найменшою відстанню:

$$k_i^* = \boldsymbol{index} \min_{k=1..N_i} (D_{i,t-1}[k]) .$$

Інший спосіб визначення нейрона-переможця полягає у максимізації виходів $y_{i,t-1}^{(n)}$ нейронів n -го шару:

$$u_{i,t-1} = \left(u[k] \middle| k = \arg \max_{j=1..N_i} y_{i,t-1}^{(n)}[j] \right).$$

У цьому випадку індекс нейрона-переможця є порядковим номером вибраного варіанта рішення $u_{i,t-1}$:

$$k_i^* = \boldsymbol{index}(u_i[k] \mid c\,(u_i[k] = u_{i,t-1}), \, k = 1..N_i) \quad \forall i \in I , \qquad (13)$$

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

Навколо нейрона-переможця може задаватися радіус навчання R у просторі векторів ваг нейронів:

$$r_{i,t}[k] = \left\| w_{i,t}^{(n-1)}[k_i^*] - w_{i,t}^{(n-1)}[k] \right\| , \; k = 1..N_i ,$$

де $w_{i,t}^{(n-1)}[k_i^*]$ – вектор ваг нейрона-переможця; $\| * \|$ – евклідова норма вектора.

Кожен нейрон, відстань від вектора ваг якого до вектора ваг нейрона-переможця є меншою ніж радіус навчання ($r_{i,t}[k] < R$), бере участь у перерахунку ваг синапсів. Ваги нейронів, які знаходяться за межами радіуса навчання, не змінюються. Радіус навчання зменшується у часі так, щоб у кінці процесу навчання корегувати ваги зв'язків міг тільки один нейрон-переможець.

Параметри g_t у (10) – (12) та e_t у (3) визначають швидкість навчання нейроагентів. Для забезпечення збіжності процесу навчання нейроагентів ці параметри задаються як додатні монотонно спадні величини:

$$g_t = g_0 / t^a \, , \; e_t = e_0 / t^b \, , \tag{14}$$

де $g_0, a > 0$; $e_0, b > 0$.

Вибір варіантів рішень продовжується до заданої кількості кроків $t \le t_{\max}$, або до виконання умови точності навчання:

$$d_t = |I|^{-1} \sum_{i \in I} \left\| w_{i,t}^{(n-1)} - w_{i,t-1}^{(n-1)} \right\| < e \quad \forall i \in I \, , \tag{15}$$

де e – точність навчання нейроагентів, яка визначається евклідовою нормою зміни ваг зв'язків між нейронами за два послідовні моменти часу.

Показники ефективності стохастичної гри

Для практичних застосувань необхідно визначити показники ефективності гри, за якими можна оцінити процес збіжності ігрового методу. За відсутності безпосереднього обміну інформацією між гравцями такі показники будуються на основі умови колективної рівноваги за Нешем [1, 14, 21].

Для цього визначимо функції середніх виграшів гравців:

$$V_i(p) = \sum_{u \in U} v_i(u) \prod_{j \in I; u_j \in u} p_j(u_j) \, , \tag{16}$$

де $p \in S^I$, $S^I = \prod_{j \in I} S^{N_j}$, $v_i(u) = M\{x_{i,t}(u)\}$, а значення $p_j \in S^{N_j}$ визначаються згідно із (3).

Рівновага за Нешем визначає такі стратегії розв'язування гри, для яких виконується умова:

$$\forall i \in I \quad V_i(p^*) - V_i(p^{I \backslash i*}, p_i) \ge 0 \, , \tag{17}$$

де $p^* \in S^I$ – оптимальна колективна змішана стратегія гравців; $V_i(p^{I \backslash i*}, p_i)$ – функція середніх виграшів, визначена на симплексі S^I при довільному відхиленні змішаної стратегії i-го гравця від точки рівноваги за Нешем у межах одиничного симплексу.

Оптимальні за Нешем вирівнювальні змішані стратегії можуть бути отримані з умови доповняльної нежорсткості [21]:

$$\nabla_{p_i} V_i = V_i e^{N_i} \quad \forall i \in I \, , \tag{18}$$

де $p_i \in S^{N_i}$ – змішана стратегія i-го гравця; $\nabla_{p_i} V_i$ – градієнт полілінійної функції середніх виграшів (16); $e^{N_i} = (1_j \mid j = 1..N_i)$ – вектор, всі елементи якого дорівнюють 1.

Для врахування розв'язків на межі одиничного симплексу виконаємо зважування векторної умови доповняльної нежорсткості елементами векторів змішаних стратегій:

$$diag(p_i)(\nabla_{p_i} V_i(p) - e^{N_i} V_i(p)) = 0 \;\; \forall i \in I \, , \tag{19}$$

де $diag(p_i)$ – квадратна діагональна матриця порядку N_i , побудована з елементів вектора p_i ; $p \in S^I$ комбіновані змішані стратегії гравців, задані на опуклих симплексах S^I .

Визначимо функцію Ляпунова як сумарну поточну похибку гравців під час пошуку точки рівноваги за Нешем:

$$\Delta_t = |I|^{-1} \sum_{i \in I} \Delta_{i,t} \, , \tag{20}$$

де $\Delta_{i,t} = \left\| p_{i,t} - \tilde{p}_{i,t} \right\|^2$, $p_{i,t}, \tilde{p}_{i,t} \in S^{N_i}$. Значення $\Delta_t \ge 0 \quad \forall t = 1, 2, ...$ і перетворюється на нуль в точках рівноваги за Нешем, які можуть досягатися як всередині так і на вершинах одиничних симплексів.

Вектори $p_{i,t}$ визначаються згідно з (3), а $\tilde{p}_{i,t}$ обчислюється так:

$$\tilde{p}_{i,t}(j) = V_{i,t}(j) / V_{i,t}, \quad j = 1..N_i,$$

де $V_{i,t}(j) = p_{i,t}(j) \nabla V_{i,t}(j)$.

Спрямування Δ_t до нуля при $t = 1, 2, ...$ свідчитиме про збіжність ігрового методу до однієї із точок рівноваги за Нешем у змішаних стратегіях, або, враховуючи (19), про досягнення розв’язку гри у чистих стратегіях.

Порядок q та величина J швидкості збіжності ігрового методу можуть бути оцінені за допомогою асимптотичного методу моментів Чжуна [22]:

$$\overline{\lim_{t \to \infty}} n^q M\{\Delta_t\} = J. \tag{21}$$

Процедура оптимізації буде тим краща, чим більше q та менше J.

На практиці можна скористатися наближеною оцінкою для визначення q, виконавши логарифмування виразу (21):

$$q \approx \frac{\lg(M\{\Delta_t\})}{\lg(t)}. \tag{22}$$

Тобто параметр q можна визначити як тангенс кута f нахилу лінійної апроксимації функції $M\{\Delta_t\}$ у логарифмічній системі координат.

Процес стохастичної неантагоністичної гри можна простежити також за зміною функції середніх виграшів гравців:

$$\Xi_t = |I|^{-1} \sum_{i \in I} \Xi_{i,t}. \tag{23}$$

Алгоритм Кохонена функціонування нейроагентів

1. Задати початкові значення параметрів:

$t = 0$ – початковий момент часу;

I – множина гравців;

N_i $\forall i \in I$ – кількості варіантів рішень;

$U_i = \{u_i[1], u_i[2], ..., u_i[N]\}$ $\forall i \in I$ – множини варіантів рішень;

$[v_i]$ $\forall i \in I$ – матриці математичних сподівань виграшів;

$[d_i]$ $\forall i \in I$ – матриці дисперсій виграшів;

$w_{i,0}^{(n-1)}[N_i, N_i]$ $\forall i \in I$ – матриці початкових ваг зв’язків між вузлами нейромеж;

h, q – параметри передатної функції нейронів;

g_0 – параметр кроку навчання;

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

e_0 – параметр e -симплекса;

b – порядок швидкості розширення e -симплекса;

$t_{\max}$ – максимальна кількість кроків методу;

e – точність навчання.

2. Вибрати варіанти рішень $u_{i,t}$ $\forall i \in I$ згідно з (3) – (4).

3. Отримати значення поточних виграшів нейроагентів як випадкові величини з нормальним розподілом $x_i(u_t) \sim Normal(v_i(u_t), d_i(u_t))$ $\forall i \in I$.

4. Обчислити входи нейроагентів $x_{i,t}^{(n-1)}$ $\forall i \in I$ згідно з (5) та відповідні виходи $y_{i,t}^{(n-1)}$ $\forall i \in I$ нейронів $(n-1)$-го шару згідно з (7).

5. Обчислити сумарні входи нейронів $\bar{x}_{i,t}^{(n)}$ $\forall i \in I$ (6) та відповідні виходи $y_{i,t}^{(n)}$ $\forall i \in I$ (7) для нейронів n-го шару.

6. Обчислити значення параметра g_t згідно з (14).

7. Визначити індекси k_i^* нейронів-переможців $\forall i \in I$ згідно з (13).

8. Обчислити ваги зв'язків для нейронів-переможців $w_{i,t}^{(n-1)}[j,k_i^*]$ $\forall i \in I$, $j = 1..N_i$ згідно з (12).

9. Обчислити характеристики якості колективного прийняття рішень Ξ_t (23), Δ_t (20), q (22).

10. Задати наступний момент часу $t := t+1$.

11. Якщо умова (15) завершення гри не виконується, то перейти на крок 2, інакше – кінець.

Результати комп'ютерного моделювання

Виконаємо розв'язування стохастичної гри двох нейроагентів $|I| = 2$ з двома чистими стратегіями $N_i = 2$, $i = 1..2$. Матриці середніх виграшів $[v_i]_{2\times 2}$ $i = 1..2$ такої гри подано у таблиці.

Матриці виграшів гравців

Стратегії	Перший гравець		Другий гравець	
$\times$	$p_2(u_2[1])$	$p_2(u_2[2])$	$p_2(u_2[1])$	$p_2(u_2[2])$
$p_1(u_1[1])$	0.1	0.6	0.1	0.9
$p_1(u_1[2])$	0.9	0.4	0.9	0.1

Відповідно до таблиці вигляд зрізів функцій середніх виграшів нейроагентів (17) зі значеннями на одиничному симплексі зображено на рис. 2.

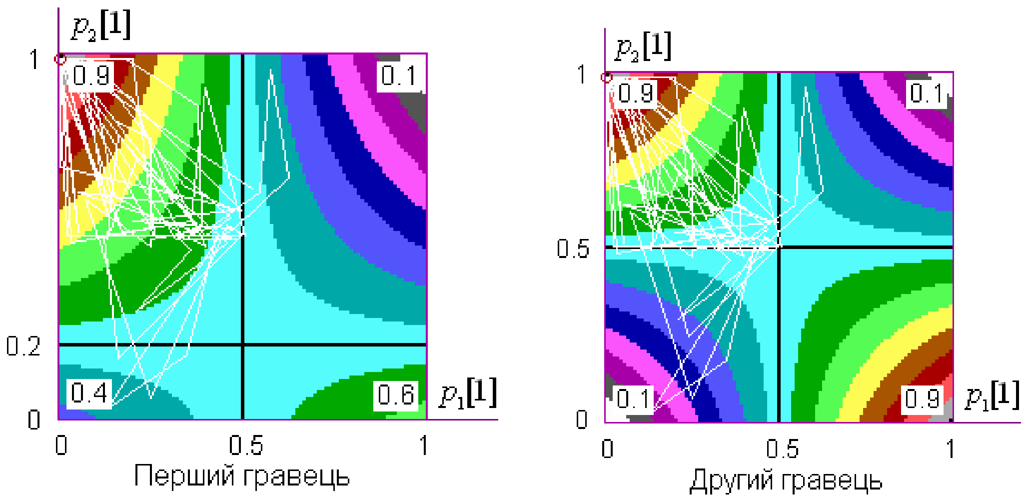

Рис. 2. Функції середніх виграшів нейроагентів

Аналіз функцій середніх виграшів показує, що гра має один розв'язок Неша у змішаних $(p_1[1], p_2[1]) = (0.5, 0.2)$ та два розв'язки у чистих стратегіях $(p_1[1], p_2[1]) = (0,1)$, $(p_1[1], p_2[1]) = (1,0)$.

В умовах невизначеності елементи матриць середніх виграшів $\left[v^i(u)\right]_{\forall u\in U}$ апріорі невідомі і доступні для спостереження у вигляді випадкових поточних значень $x_i(u)$. Для моделювання випадкових виграшів виберемо нормальний закон розподілу:

$$x_i(u) = Normal(v_i(u), d_i(u)),$$

де $v^i(u)$ – математичне сподівання; $d_i(u)$ – дисперсія. Нормально розподілені випадкові величини обчислимо за допомогою суми рівномірно розподілених випадкових дійсних чисел $w\in[0,1]$:

$$x_i(u_t) = v_i(u_t) + \sqrt{d_i(u_t)}\left(\sum_{j=1}^{12} w_{i,t}[j] - 6\right).$$

Початкові значення ваг зв'язків $w_0^{(n-1)}$ між нейронами є випадковими величинами, рівномірно розподіленими в інтервалі [0, 1]. Як передатну функцію нейронів вибрано лінійну залежність виходу від сумарних входів (8).

Збіжність ігрового нейроагентного методу визначається точним співвідношенням параметрів, які у загальному випадку повинні задовольняти базові умови стохастичної оптимізації [22–24]. Параметри методу навчання нейроагентів набувають таких значень: $g=1$; $a=0{,}1$; $N_i = N = 2$; $e = 0.999/N$; $b=1$.

На рис. 3 у логарифмічному масштабі зображено графіки функцій середніх виграшів $\Xi_{i,t}$ та похибок досягнення оптимального колективного рішення в одній із точок рівноваги за Нешем $\Delta_{i,t}$.

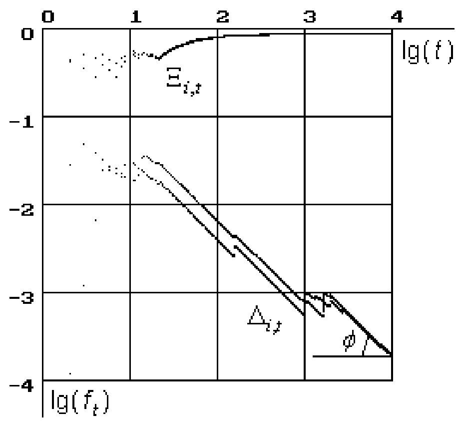

Рис. 3. Характеристики навчання нейроагентів

Зростання $\Xi_{i,t}$ та зменшення $\Delta_{i,t}$ у часі свідчать про збіжність нейроагентного методу прийняття рішень у сенсі виконання зваженої умови доповняльної нежорсткості (19).

Траєкторії зміни умовних імовірностей колективного вибору варіантів рішень у межах одиничного симплексу зображено на рис. 2. З отриманих даних видно, що для заданих значень параметрів ігровий нейроагентний метод (12) забезпечує розв'язування стохастичної гри на вершині одиничного симплексу $(p_1[1], p_2[1]) = (0,1)$ і має близький до 1 порядок $q \approx tg(f)$ швидкості збіжності.

Вивчимо залежність часу навчання стохастичної гри нейроагентів від основних параметрів алгоритму. Час навчання визначимо як мінімальну кількість кроків, необхідних для навчання нейроагентів із заданою точністю $e > 0$:

$$t_{out} = (t = t_{\min} \mid d_t < e),$$

де поточна точність навчання d_t обчислюється згідно з (15).

У зв'язку з випадковим вибором варіантів рішень необхідно виконати усереднення часу навчання нейроагентів для різних послідовностей випадкових величин:

$$\overline{t} = \frac{1}{k_{\exp}} \sum_{j=1}^{k_{\exp}} t_{out}[j],$$

де $k_{\exp}$ – кількість експериментів.

Середня кількість кроків навчання $\overline{t}$ залежить від параметрів алгоритму навчання нейроагентів та параметрів середовища прийняття рішень.

Графік залежності середнього часу $\overline{t}$ навчання нейроагентів від параметра α зображено на рис. 4. Результати отримано для значення дисперсії поточних виграшів $d = 0.01$. Параметр $\alpha \in (0,1]$ визначає порядок монотонного спадання величини $g_t > 0$ (14), яка регулює швидкість навчання нейроагентів. Із збільшенням α значення g_t зменшується.

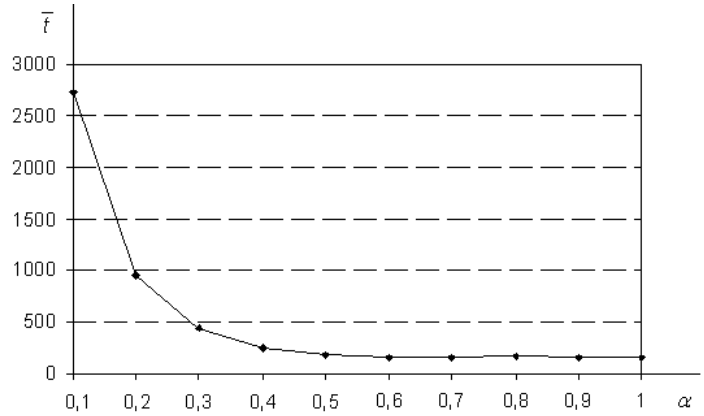

Рис. 4. Вплив параметра α на час навчання нейроагентів

Точність навчання гри нейроагентів дорівнює $e = 10^{-3}$. Дані усереднено за $k_{\exp} = 100$ експериментами. В усіх експериментах забезпечується з більшими імовірностями вибір тих чистих стратегій рішень, для яких виконується умова (19).

Як видно із результатів моделювання, збільшення параметра α призводить до зменшення середньої кількості кроків $\overline{t}$ навчання нейроагентів з точністю e.

Залежність середньої кількості кроків $\overline{t}$ навчання нейроагентів від дисперсії d стохастичного середовища (дисперсії оцінок варіантів рішень) подано на рис. 5. Результати отримано для значення порядку кроку навчання нейроагента $\alpha = 0.5$.

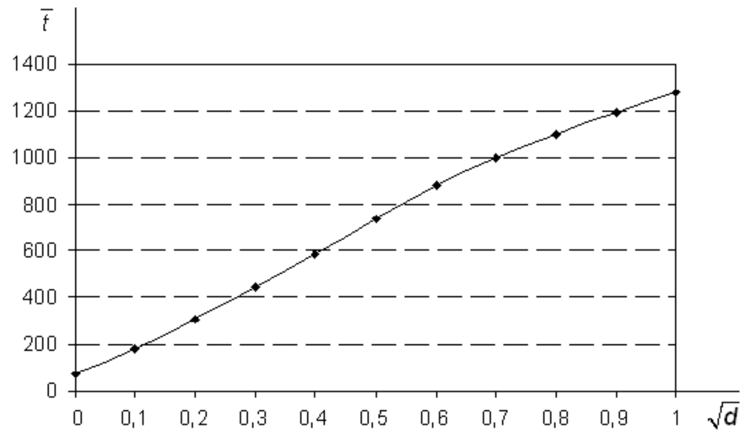

Рис. 5. Вплив дисперсії оцінок варіантів рішень
на час навчання нейроагентів

Із зростанням дисперсії зростає середня кількість кроків, необхідних для навчання нейроагентів. Значне зростання дисперсії може призвести до неправильного визначення оптимального варіанта рішення.

Моделювання стохастичної гри з іншими матрицями середніх виграшів нейроагентів може дати інший результат, зберігаючи отримані залежності між параметрами, які необхідні для оптимізації платіжних функцій гравців.

Висновки

У цій статті запропоновано нейроагентну ігрову модель колективного прийняття рішень в умовах стохастичної невизначеності. Нейроагенти побудовано на основі штучних нейронних мереж зі зворотними зв’язками та з навчанням без учителя. Поточне колективне рішення отримується після незалежного вибору власних варіантів рішень усіма гравцями. Кожен гравець формує поточний варіант рішення за значеннями виходів нейромережі. Вибір варіантів рішень здійснюється нейроагентами випадковим способом, незалежно від часу та від інших агентів. Випадковий вибір передбачає обчислення імовірностей вибору варіантів рішень оптимальним проектуванням виходів нейромережі на одиничний симплекс. Після вибору колективного варіанта рішення визначається реакція середовища прийняття рішень як сукупність значень поточних виграшів нейроагентів. Поточний виграш кожного нейроагента передається на входи відповідної двошарової нейромережі. Далі відбувається навчання нейроагентів шляхом зміни ваг зв’язків між нейронами за одним із алгоритмів навчання без учителя. Процес навчання повторюється у часі до стабілізації ваг зв’язків між нейронами із заданою точністю. Навчання спрямоване на максимізацію середніх виграшів нейроагентів. Розв’язок гри досягається в одній із точок колективної оптимальності або рівноваги залежно від значень параметрів вибраного методу навчання нейроагентів.

Розроблена програмна модель підтверджує збіжність ігрового нейроагентного методу (12) прийняття рішень. Ефективність методу оцінено за допомогою характеристичних функцій середніх виграшів та похибок колективного вибору оптимального варіанта прийняття рішення. Збіжність нейроагентного ігрового методу залежить від кількості гравців, кількості варіантів рішень та співвідношень параметрів методу і параметрів середовища прийняття рішень.

Достовірність отриманих результатів підтверджено повторюваністю значень розрахованих характеристик ігрового нейроагентного методу прийняття рішень для різних послідовностей випадкових величин.

Результати цієї роботи можна використати для побудови розподілених систем керування та прийняття рішень в умовах невизначеності.

Проведені дослідження можна продовжити у напрямку застосування іншої конфігурації нейроагентів та інших методів їх навчання, обміну інформацією між агентами стохастичної гри, зростання кількості гравців та кількості їх чистих стратегій, визначення теоретичних умов збіжності ігрового нейроагентного методу.

1. Опойцев В.И. Равновесие и устойчивость в моделях коллективного поведения / В.И. Опойцев. – М.: Наука, 1977. – 245 с. 2. Кини Р.Л. Принятие решений при многих критериях: предпочтения и замещения: Пер. с англ. под ред. И. Ф. Шахнова / Р.Л. Кини, Х. Райфа. – М.: Радио и связь, 1981. – 560 с. 3. Подиновский В.В. Парето-оптимальные решения многокритериальных задач / В.В. Подиновский, В.Д. Ногин. – М.: Наука, 1982. – 256 с. 4. Ларичев О. И. Теория и методы принятия решений: Учеб. / Ларичев О.И. – М. : Логос, 2000. – 296 с. 5. Ситник В. Ф. Системи підтримки прийняття рішень: навч. посіб. / В.Ф. Ситник. – К.: КНЕУ, 2004. – 614 с. 6. Орлов А. И. Теория принятия решений: учебник / А. И. Орлов. – М.: Экзамен, 2006. – 573 с. 7. Дубовой В. М. Моделі прийняття рішень в управлінні розподіленими динамічними системами. Монографія / В. М. Дубовой, О. О. Ковалюк. – Вінниця:

УНІВЕРСУМ-Вінниця, 2008. – 185 с. 8. Weiss G. Multiagent Systems. A Modern Approach to Distributed Artificial Intelligence / G. Weiss, editor. – Springer Verlag, Berlin, 1996. – 643 pp. 9. Городецкий В.И. Информационные технологии и многоагентные системы / В.И. Городецкий // Проблемы информатизации. – 1998. – Вып. 1. – С. 3 – 14. 10. Wooldridge M. An Introduction to Multiagent Systems / M. Wooldridge. – John Wiley & Sons, 2002. – 366 p. 11. Тарасов В.Б. От многоагентных систем к интеллектуальным организациям: философия, психология, информатика / В.Б. Тарасов. – М.: Эдиториал УРСС, 2002. – 352 с. 12. Швецов А.Н. Агентно-ориентированные системы: от формальных моделей к промышленным приложениям / А.Н. Швецов. – Вологодский гос. технич. унив. – 101 с.: [Электрон. ресурс]. – http://www.ict.edu.ru/ft/005656/62333e1-st20.pdf. 13. Доманский В.К. Стохастические игры // Математические вопросы кибернетики / В.К. Доманский. – 1988. – № 1. – С. 26-49. 14. Назин А.В. Адаптивный выбор вариантов: Рекуррентные алгоритмы / А.В. Назин, А.С. Позняк. – М.: Наука, 1986. – 288 с. 15.Fudenberg D. The Theory of Learning in Games / D. Fudenberg, D.K. Levine. – Cambridge, MA: MIT Press, 1998. – 292 pp. 16. Трухаев Р.И. Модели принятия решений в условиях неопределенности / Р.И. Трухаев. – М.: Наука, 1981. – 257 с. 17. Уоссермен Ф. Нейрокомпьютерная техника / Ф. Уоссермен. – М. : Мир, 1992 г. – 240 с. 18. Барский А. Б. Нейронные сети: распознавание, управление, принятие решений / А. Б. Барский. – М.: Финансы и статистика, 2004. – 176 с. 19. Хайкин С. Нейронные сети: полный курс / С. Хайкин. – М.: Вильямс, 2006. – 1104 с. 20. Злобин В. К. Нейросети и нейрокомпьютеры: учеб. пособие для студ. вузов / В. К. Злобин, В. Н. Ручкин. – СПб. : БХВ-Петербург, 2011. – 256 с. 21. Мулен Э. Теория игр с примерами из математической экономики / Э. Мулен. – М.: Мир, 1985. – 200 с. 22. Вазан М. Стохастическая аппроксимация / М. Вазан. – М.: Мир, 1972. – 295 с. 23. Невельсон М.Б. Стохастическая оптимизация и рекуррентное оценивание / М.Б. Невельсон, Р.З. Хасьминский. – М.: Наука, 1972. – 304 с. 24. Граничин О.Н. Введение в методы стохастической аппроксимации и оценивания: Учеб. пособие / О.Н. Граничин. – СПб.: Издательство С.-Петербургского университета, 2003. – 131 с.