C. L. Webber // Int. J. Bifurcation Chaos. – 2007. – № 17. – P. 10. 5. Marwan N. A historical review of recurrence plots / N. Marwan // The European Physical Journal Special Topics. – 2008. – № 164 (1). – P. 3–12. 6. Zou Y. Inferring direct coupling by means of recurrences / Y. Zou, MC. Romano, M. Thiel, M. Marwan, J. Kurths // International Journal of Bifurcation and Chaos. – 2011. – № 21(4). – P. 1099–1111. 7. Thiel M. Generating surrogates from recurrences / M. Thiel, MC. Romano, J. Kurths, M. Rolfs, R. Kliegl // Philosophical Transactions of the Royal Society A: Mathematical, Physical & Engineering Sciences / В. Б. Киселев. – 2008. – № 366(18650. – P. 545–557. 8. Киселев В. Б. Рекуррентный анализ – теория и практика / В. Б. Киселев // Научно-технический вестник СпбГУ / ИТМО. – 2006. – Вып. 6(29). – С. 118–127. 9. Киселев В. Б. Некоторые методы нелинейного анализа В. Б. Киселев // Научно-технический вестник информационных технологий, механики и оптики СПбГУ ИТМО, 2005, Вып. 4(20). – С. 172–180. 10. Піскун О. В. Особливості застосування рекурентних діаграм і рекурентного кількісного аналізу для дослідження фінансових часових рядів / О. В. Піскун // Фінансовий простір. – 2011. – № 3 (3). – С. 111–118. 11. Васюта К. С. Рекуррентный анализ процессов в телекоммуникационных системах / К. С. Васюта // Наукові записки УНДІЗ, 2008.– № 6 (8). – С. 90–96. 12. Васюта К. С. Розвиток методів виявлення радіосигналів в радіотехнічних системах із використанням рекурентного аналізу / К. С. Васюта, О. Б. Танцюра, О. В. Ревін // Наука і техніка Повітряних сил Збройних сил України. – 2013. – № 2(11). – С. 135–139. 13. Киселев В. Б. Определение стабильности траектории процесса в фазовом пространстве при помощи рекуррентного анализа / В. Б. Киселев // Научно-технический вестник информационных технологий, механики и оптики СПбГУ ИТМО. – 2007. – Вып. 6(40). – С. 121–130. 14. Соловйов В. М. Моделювання складних економічних систем: навч. посіб. / В. М. Соловйов, В. В. Соловйова, Н. А. Хараджян. – Кривий Ріг: Видавничий відділ НметАУ. – 2010. – С. 119. 15. Takens F. Detecting strange attractors in turbulence / F. Takens // Dynamical Systems and Turbulence. Lecture Notes in Mathematics, edited by D.A. Rand L.S. Young. – Heidelberg: Springer-Verlag. – 1981. – P. 366 – 381.

УДК 004.852; 004.94

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

ІГРОВА МОДЕЛЬ САМООРГАНІЗАЦІЇ МУЛЬТИАГЕНТНИХ СИСТЕМ

Розроблено ігрову модель самоорганізації мультиагентних систем в умовах невизначеності. Наведено формулювання стохастичної ігрової задачі, визначено критерії самоорганізації стратегій гравців, розроблено рекурентний метод, алгоритм та програмні засоби, що навчають мультиагентну систему імітувати синхронізоване ритмічне світіння колонії комах-світлячків.

Ключові слова: мультиагентні системи, умови невизначеності, стохастична ігрова модель, самоорганізація.

The game model of multi-agent systems of self-organizing in the conditions of uncertainty is developed. The formulation of a stochastic game problem is carried out, criteria of self-organizing of strategies of players are defined, a recurrent method, algorithm and software of learning of multi-agent system to simulate the synchronised rhythmic luminescence of a colony of fireflies are developed.

Key words: multi-agent systems, uncertainty conditions, stochastic game model, self-organizing.

Мета роботи

Об'єктом дослідження цієї роботи є процеси самоорганізації мультиагентних систем в умовах невизначеності, спрямовані на досягнення скоординованої роботи складових елементів мультиагентних систем за рахунок властивостей самонавчання та адаптації, результатом чого є те, що розподілена система елементів функціонує як цілісний злагоджений організм.

Предметом дослідження є стохастична ігрова модель самоорганізації мультиагентних систем, яка забезпечує баланс значень платіжних функцій команди гравців та виявляється у досягненні скоординованих стратегій агентів.

Метою роботи є побудова ігрової моделі самоорганізації мультиагентних систем для підтримки прийняття рішень в умовах невизначеності. Ця мета досягається розв'язуванням таких задач: розроблення математичної моделі мультиагентної стохастичної гри; розроблення самонавчального методу та алгоритму розв'язування стохастичної гри; розроблення програмних засобів моделювання стохастичної гри; аналіз отриманих результатів та вироблення рекомендацій для їх практичного застосування.

Самоорганізація систем

Згідно з теорією еволюційного розвитку, виокремлення організації (структури та функцій) природних систем здійснюється за рахунок процесів їх самоорганізації.

Самоорганізація – це процес, упродовж якого створюється, відтворюється або вдосконалюється організація складної динамічної системи за рахунок внутрішніх факторів, без відповідного зовнішнього впливу [1]. Внутрішні фактори перебудови системи спрямовані на зменшення ентропії, підвищення скоординованості функціонування її складових елементів. Процеси самоорганізації охоплюють різноманітні потоки обміну ресурсами, енергією та інформацією із зовнішнім середовищем.

В основу дослідження процесів самоорганізації систем покладено міждисциплінарний характер та системний підхід. Хоча термін "самоорганізація" використовується у науковій літературі, починаючи з XX століття, вивченням принципів самоорганізації займалися відомі філософи і вчені східних та західних наукових шкіл від античних часів до наших днів.

Філософські, релігійні та натуралістичні аспекти організації та самоорганізації матерії, природних і суспільних процесів вивчалися у працях Лао-цзи, Сократа, Арістотеля, Платона, Луїса Моліни, Рене Декарта, Френсіса Бекона, Іммануїла Канта, Фрідріха Гегеля, Чарльза Дарвіна, Анрі Пуанкре, Ервіна Шредінгера, Норберта Вінера, Людвіга Берталанфі, Вільяма Ешбі, П'єра Грассе, Іллі Пригожина, Германа Хакена, Ебелінга Вернера, Манфреда Ейгена, Рене Тома, Володимира Арнольда, Олександра Самарського, Сергія Курдюмова та інших.

Припущення про те, що будь-яка система впорядкована за рахунок своєї внутрішньої динаміки, висловлював Р. Декарт [2]. У літературі з кібернетики термін "самоорганізація" вперше вжито в роботах В. Р. Ешбі [3], який визначає систему із самоорганізацією як таку, що сама змінює свою структуру. На те, що самоорганізація виникає за рахунок внутрішніх взаємодій між елементами системи через загальне середовище (явище стігмергії), вказав П. Грассе [4]. Самоорганізацію у термінах термодинаміки означив І. Р. Пригожин [5]: відкрита система зменшує свою ентропію (у системі виникає порядок із хаосу), коли в неї надходить енергія за рахунок зовнішнього впливу. Самоорганізацію як явище спонтанного утворення просторових та часових структур у фізичних і хімічних системах, які далеко від стану рівноваги [6], охарактеризував Г. Хакен. Механізм самоорганізації як еволюцію циклів та гіперциклів хімічних реакцій біологічних систем, які мають властивості самовідтворення та здатність до виживання [7], окреслив М. Ейген. Перспективними напрямами дослідження самоорганізації є теорія катастроф [8], коли плавна зміна внутрішніх параметрів системи призводить до стрибкоподібної зміни її стану (Р. Тома, К. Зіман, В. І. Арнольд), та теорія хаосу [9], коли незначні зміни навколишнього середовища призводять до непередбачених наслідків поведінки динамічних нелінійних систем (А. Пуанкре, А. М. Колмогоров, В. І. Арнольд, Ю. Мозер).

Незважаючи на значний період дослідження та отримані наукові результати, концепція самоорганізації систем все ще у стадії становлення.

За певних умов самоорганізація може проявлятися у розподілених природних (наприклад, фізичних, хімічних, біологічних), організаційних (наприклад, соціальних, екологічних, економічних) або штучних (наприклад, мультиагентних) системах [10]. Класичними прикладами самоорганізації є лазерний промінь, хімічна реакція Бєлоусова–Жаботинського, конвективні комірки Бенара, самоорганізація плазми. Прояви самоорганізації притаманні усім живим організмам, наприклад, штамам бактерій, колоніям або роїям комах, косякам риб, зграй птахів, популяціям хижаків та їх жертв тощо. Ідеї самоорганізації систем використовують для побудови обчислювальних машин, мереж, алгоритмів та програм.

Вивченням процесів самоорганізації систем різної природи займаються теорія динамічних систем, детермінованого хаосу та синергетика – міждисциплінарна наука про нелінійні відкриті дисипативні (такі, що розсіюють енергію) системи [11]. Для дослідження динаміки самоорганізації систем побудовано математичні моделі, наприклад, модель Жаботинського–Корзухіна, брюселятор Пригожина, орегонатор Філда–Нойєса та інші.

Мірою оцінювання самоорганізації системи є її ентропія. Зі зменшенням ентропії ступінь самоорганізації системи зростає. Зменшення ентропії може досягатись надходженням у систему зовнішньої енергії, здатністю до саморегулювання, адаптації та самонавчання компонентів системи.

У розподілених активних системах (здатних самостійно приймати рішення) зменшення ентропії забезпечується зростанням показника інтелектуальності їх компонентів у ході колективного розв'язування поставленої задачі, вмінням аналізувати поточні стани системи, запам'ятовуванням передісторії подій, прогнозуванням поведінки системи у наступні моменти часу, можливістю автономного прийняття рішень, забезпеченням функцій життєздатності, координацією дій, налагодженням комунікацій між складовими елементами для обміну локальною інформацією у ході децентралізованої оптимізації цільових функцій.

Самоорганізація розподіленої системи зумовлена локальними взаємодіями її складових елементів на мікрорівні та проявляється на макрорівні. На макрорівні у системі динамічно виникають деякі нові властивості, процеси, поведінка, структури, патерни тощо, які є наслідком локальних взаємодій елементів системи на мікрорівні, причому ці властивості такі, що вони не описуються в термінах властивостей мікрорівня.

У філософському розумінні самоорганізація є фундаментальною властивістю матерії, коли об'єкти мікрорівня можуть утворювати об'єкти макрорівня. Так, елементарні частинки об'єднуються в атоми, атоми – у молекули, молекули – у супрамолекулярні об'єднання, супрамолекулярні об'єднання – у біологічні системи, біологічні системи – у соціальні системи. Вважається, що супрамолекулярні об'єднання є перехідним мостом між неживою і живою матерією [12]. Супрамолекулярні ансамблі виникають у результаті спонтанного об'єднання мікрокомпонентів за рахунок міжмолекулярних сил у впорядковані високомолекулярні сполуки (наноструктури, кристали, плівки, мембрани, везикули, білки тощо).

Як правило, ефект самоорганізації розподіленої системи спостерігається на макрорівні сторонніх систем з природним або штучним інтелектом. Самоорганізація проявляється у вигляді впорядкованих структур, узгоджених процесів та емерджентності – набутої інтегральної властивості системи, не характерної для її складових елементів.

Емерджентність як прояв самоорганізації систем

Для виникнення явища емерджентності (або системного ефекту) [13] повинні виконуватися такі необхідні умови:

- наявність принаймні дворівневої організації: мікрорівень, де відбуваються локальні взаємодії, та макрорівень, на якому може виявитися емерджентний процес;
- нелінійність: взаємодія на мікрорівні є нелінійною;
- наявність зворотного зв'язку: локальні взаємодії на мікрорівні повинні містити зворотний зв'язок;
- динамічна рівновага: емерджентні процеси відбуваються у динаміці, тобто вони існують, поки існують взаємодії на мікрорівні.

Явища емерджентності мають такі властивості:

- новизна: невисловлювальність у термінах мікрорівневих компонентів;
- когерентність: узгоджений перебіг хвильових або коливних процесів, який проявляється тоді, коли в системі виникають відповідні взаємодії;
- емерджентність проявляється на макрорівні стосовно компонентів, що породжують це явище;
- динаміка: емерджентність виникає в процесі взаємодії і не є заданою заздалегідь або ззовні;
- візуалізація: емерджентність явно демонструє себе тим або іншим способом, наприклад, виникненням впорядкованих структур, фазових переходів тощо.

Емерджентність та самоорганізація є процесами, які можуть виникати окремо (вони виражають різні властивості поведінки системи) або спільно. Основна подібність емерджентності та самоорганізації в тому, що це динамічні процеси, зумовлені локальними взаємодіями складових елементів на мікрорівні, а проявляються на макрорівні.

Обидва процеси є стійкими, але їх відмінність виявляється у тому, що емерджентність стійка стосовно множини складових компонентів системи, чиї взаємодії викликають це явище (компоненти системи можуть з’являтися та зникати, але це істотно не впливає на досягнутий ефект), а самоорганізація стійка в тому розумінні, що вона здатна адаптуватися як до внутрішніх змін у системі, так і до змін зовнішнього світу, підтримуючи виниклий порядок.

Самоорганізація мультиагентних систем

Вивчення умов зародження та динаміки процесів самоорганізації можна здійснювати не тільки проведенням відповідних фізичних або хімічних експериментів, але й за допомогою штучних розподілених систем, наприклад, побудованих на основі інтелектуальних агентів [14–16].

Агент – це активна система штучного інтелекту, призначена для автономного прийняття рішень в умовах невизначеності на основі самонавчання та отримання інформації від керованого середовища, сусідніх агентів та від людини-оператора. Агент та середовище об’єднані ланкою зворотного зв’язку: прийняті рішення надходять на вхід середовища, а відповідні реакції середовища – на вхід агента для корегування його рішень у наступні моменти часу.

Для децентралізованого розв’язування задач керування та прийняття рішень у розподілених системах використовують мультиагентні системи (МАС), утворені декількома інтелектуальними агентами, що взаємодіють. МАС мають такі властивості:

- автономність – дії агентів повністю або частково незалежні;
- спеціалізація – кожен з агентів виконує вузькоспеціалізовані функції;
- обмеженість знань – кожен агент не має цілісного уявлення про систему;
- децентралізація – агенти керують локальними ділянками розподіленої системи.

МАС є розподіленими системами штучного інтелекту зі здатністю самонавчання для адаптації до невизначеностей середовища їх функціонування. Під невизначеністю будемо розуміти відсутність або недостатність інформації в системі прийняття рішень, необхідної для одноетапного вибору таких варіантів рішень, які достовірно забезпечують мінімальне чи максимальне значення цільової функції для задач скалярної оптимізації, або балансу значень цільових функцій для задач векторної оптимізації.

У системі колективного прийняття рішень невизначеність може генеруватися однією із її складових частин – середовищем прийняття рішення або учасниками прийняття рішення.

Факторами невизначеностей МАС можуть бути:

- структурний – невідомий точний склад системи та зв’язки між її елементами;
- алгоритмічний – невідомий алгоритм функціонування системи;
- інформаційний – нечіткість, відсутність повної інформації, необхідної для прийняття рішень;
- лінгвістичний – неоднозначність трактування повідомлень, що передаються між агентами;

- цільовий – невідома глобальна мета функціонування системи;
- соціальний – зумовлений колективною взаємодією агентів, коли дії одного з агентів впливають на вибір рішень іншими агентами;
- стохастичний – вплив на систему неконтрольованих зовнішніх чинників.

У теорії адаптивних систем моделювання невизначеностей найчастіше здійснюється за допомогою випадкових процесів.

Становлення МАС відбулося на основі інтегрування знань в областях колективної поведінки автоматів, штучного інтелекту, технологій “м’яких” обчислень, децентралізованого керування та прийняття рішень, інформаційних та комп’ютерних мереж, паралельного та розподіленого програмування.

Методи проектування МАС регламентуються специфікаціями фахових міжнародних організацій, наприклад, специфікаціями FIPA (Foundation for Intelligent Physical Agents).

Сучасні дослідження у галузі МАС пов’язані зі складними проблемами розподіленого штучного інтелекту:

- планування сценаріїв колективної поведінки агентів;
- колективне вироблення та прийняття рішень;
- забезпечення координації та кооперації у МАС;
- розвідування станів середовища функціонування МАС;
- визначення оптимальної структурної організації МАС;
- керування знаннями;
- розроблення методів і засобів багатоагентного навчання;
- розроблення методів, мов і засобів комунікації агентів;
- розроблення методів і засобів автоматизованого проектування МАС;
- забезпечення мобільності агентів;
- забезпечення пластичності, надійності та стійкості до збоїв у роботі.

МАС забезпечують ефективне розв’язання задач, які не можуть бути розв’язані централізованим способом або за допомогою одного агента, наприклад: мережеві технології та розподілені обчислення, пошук інформації в Internet, мобільний зв’язок, електронна комерція, керування ринком цінних паперів та інвестицій, подолання наслідків надзвичайних ситуацій, військова справа, керування трудовими ресурсами, моделювання соціальних явищ та структур, складання розкладів, дистанційне навчання, логістика, керування транспортними потоками, геоінформаційні системи, тренажери для професійного навчання, стратегічні комп’ютерні ігри та інші.

Крім практичних застосувань, МАС можуть використовуватися в наукових дослідженнях для моделювання процесів самоорганізації систем різного виду та призначення. Особливо добре МАС підходять для моделювання самоорганізації у біологічних, економічних та соціальних системах.

Самоорганізація МАС – це процес перетворення локально скоординованих дій множини агентів на глобально скоординовані дії за рахунок внутрішніх факторів, без відповідного зовнішнього впливу. Це – здатність колективу агентів з локально обумовленими зв’язками і цілями досягати стійких скоординованих стратегій поведінки в умовах невизначеності за рахунок самонавчання. Самонавчальні системи здатні покращувати якість свого функціонування відповідно до заданих критеріїв за рахунок нагромадження та використання інформації для досягнення заданих або оптимальних станів і процесів поведінки системи за наявної невизначеності та змінних зовнішніх умов.

Самоорганізація та складні форми поведінки МАС можуть виникати у разі колективної реалізації найпростіших дій агентів. МАС мають усі вимоги, необхідні для забезпечення самоорганізації:

- автономність – здатність системи до самостійного прийняття рішень;
- відкритість – здатність сприймати зовнішній світ за допомогою рецепторів і локально впливати на нього за допомогою ефекторів;
- наявність програмного або фізичного середовища для розподіленої взаємодії агентів;
- наявність засобів підтримання взаємодії та кооперації агентів;

- відсутність зовнішнього або централізованого керування;
- динамічна оптимізація виконується в процесі роботи системи і не потребує попереднього настроювання або планування;
- простота правил локальної взаємодії, яка веде до складної поведінки системи загалом;
- можливість повторного використання варіантів дій у часі;
- адаптивність та стійкість стосовно змін – здатність належного реагування на зміни зовнішнього середовища.

Виділяють такі механізми самоорганізації МАС:

- прямі взаємодії між агентами за допомогою відповідних протоколів обміну даними;
- непрямі комунікації через середовище;
- самоорганізація поведінки агентів на основі підкріпленого навчання (заохочення);
- кооперативна поведінка агентів;
- вибір типової архітектури системи (холонічний підхід з елементами централізації).

Координація як умова самоорганізації мультиагентних систем

Для виникнення умов самоорганізації МАС необхідна координація дій агентів у процесі розподіленого розв'язування поставленої задачі. Координація – це забезпечення узгодженої, впорядкованої роботи усіх ланок мультиагентної системи.

Координація необхідна для узгодження індивідуальних цілей і варіантів поведінки агентів, за яких кожен агент покращує або не погіршує значення своєї функції корисності, а система загалом покращує якість розв'язування загальної задачі. Методи розв'язування задачі координації ґрунтуються на результатах класичної теорії керування, дослідженні операцій, теорії ігор, планування та на результатах інших галузей математики і кібернетики.

Розрізняють централізовану та децентралізовану координацію МАС. У випадку централізованої координації агенти отримують директивні команди від головного агента, які він враховує, виробляючи власну стратегію поведінки. Механізми децентралізованої координації будуються на обміні інформацією між агентами. Такий обмін може бути глобальним, між усіма агентами, або локальним, у межах визначених підмножин агентів. Агенти обмінюються між собою значеннями вибраних стратегій, поточних реакцій середовища тощо. Для цього агенти використовують спеціальні мови та протоколи комунікації.

Зростання кількості агентів та кількості варіантів прийняття рішень призводить до погіршення техніко-економічних показників МАС з централізованою координацією порівняно з децентралізованою, що зумовлено технічною складністю та зростанням накладних витрат для організації великих централізованих систем.

Координація може бути статичною або динамічною у часі. Якщо у ході розв'язування спільної задачі агенти переходять у стійкі, незмінні у часі стани, то така координація називається статичною. Якщо ж агенти переходять у стійкі, періодичні у часі стани, то маємо динамічну координацію. МАС з динамічною координацією можна використати для генерування складних сигналів для розподілених робототехнічних систем.

Частковим проявом координації є синхронізація стратегій агентів.

Синхронізація як прояв координації систем

Синхронізація – це набуття об'єктами розподіленої системи єдиного ритму роботи. Синхронізація може бути викликана слабкою взаємодією між об'єктами або дією зовнішньої сили.

Ефект синхронізації відкрив Х. Гюйгенс у XVII столітті у результаті спостереження за узгодженим ходом двох маятникових годинників, викликаним слабким впливом власних коливань їх спільної опори.

Явища синхронізації систем поширені у природі, техніці та суспільстві [17]. За певних умов вони можуть бути зумовлені простими фізичними чи хімічними законами або колективною взаємодією активних об'єктів (агентів) зі складною динамікою поведінки. Тут під активними розуміємо автономні об'єкти, здатні контролювати стани зовнішнього середовища, виробляти і реалізовувати керуючі сигнали.

Прикладами синхронізації у природі є ритмічне світіння рою комах-світлячків, сюрчання цвіркунів, рух косяків риб, політ зграї птахів тощо. Узгодження біологічних ритмів живих організмів із зовнішніми природними чинниками, формування серцевих ритмів та електричних ритмів мозку є проявами синхронізації у біології та медицині. Синхронізація є центральним механізмом опрацювання інформації нейронами у різних ділянках мозку та передавання інформації між цими ділянками.

У фізиці прикладами синхронізації є скоординоване коливання декількох маятників, наведене звучання органних труб, осциляція електричних сигналів.

У радіофізиці, радіотехніці, радіолокації, радіовимірах, радіозв'язку синхронізація використовується для стабілізації частоти електричних, електромагнітних та квантових генераторів, синтезу частот, демодуляції сигналів, керування фазованими антенними решітками, у доплерових системах, у системах точного часу, в різних способах передавання інформації.

У механіці явище синхронізації застосовують для побудови різноманітних вібротехнічних пристроїв.

В енергетиці синхронізація використовується для забезпечення точного збігу частот декількох електрогенераторів змінного струму за умови їх паралельної роботи на загальне навантаження, для просторової орієнтації розподілених площин фотоелементів для збирання сонячного світла та його перетворення на електричну енергію.

В астрономії синхронізація виявляється у встановленні співвідношень між середніми кутовими швидкостями обертань небесних тіл навколо своєї осі та швидкостями їх орбітальних рухів.

Синхронізація потоків коду та даних необхідна для забезпечення роботи сучасних розподілених програмних засобів і сховищ даних з мережевою реалізацією, прикладом яких є мультиагентні системи.

У соціальних системах явища синхронізації спостерігаються у русі строю солдат, танцях, хоровому співі, плесканні рук у концертному залі, масових молитвах вірян, злагодженій роботі законодавчих та виконавчих органів влади тощо.

У мультиагентних системах синхронізація необхідна, щоб забезпечити узгоджену роботу їх складових частин, передавання повідомлень між агентами, створення умов самоорганізації, коли розподілена система веде себе як цілісний, штучно сформований організм.

Ігрова модель самоорганізації мультиагентних систем

Враховуючи притаманні колективу агентів фактори децентралізованого вироблення рішень, конкуренції, взаємодії, кооперації, навчання для дослідження самоорганізації МАС в умовах невизначеності, використаємо модель стохастичної гри [18–20], яка забезпечує адаптивний пошук точок рівноваги функцій виграшів на одиничному симплексі комбінованих змішаних стратегій гравців. У ході стохастичної гри гравці навчаються вибирати оптимальні у середньому чисті стратегії (дії), перебудовуючи власні вектори динамічних змішаних стратегій (умовних імовірностей варіантів дій). Поряд із характерними індивідуальними особливостями, теорія ігор і теорія МАС мають спільні предмети дослідження. Стохастична гра є одним із засобів моделювання МАС.

З теоретико-ігрового погляду самоорганізація МАС – це процес скоординованого вибору стратегій агентів, досягнутий за рахунок самонавчання у ході колективної оптимізації платіжних функцій в умовах невизначеності. Колективні рішення є скоординованими, якщо вони задовольняють умови вигідності, справедливості та стійкості для всіх учасників прийняття рішень. Для розв'язування практичних задач найчастіше використовуються критерії колективної рівноваги за Нешем, Слейтером, Джофріоном, Байєсом, корельованої рівноваги, оптимальності за Парето тощо.

Ігрова самоорганізація МАС в умовах невизначеності є актуальною науково-практичною проблемою, яка інтенсивно досліджується у сучасній науковій літературі за напрямами розподіленого штучного інтелекту та прийняття рішень.

Математична модель стохастичної гри

Для дослідження самоорганізації МАС побудуємо модель стохастичної гри агентів $\Gamma = (D,\ \{U^i\}_{i\in D},\ \{Z(x^i)\}_{i\in D})$, де $D \neq \varnothing$ – множина гравців, $U^i = (u^i(1), u^i(2), ..., u^i(N_i))$ – чисті стратегії i-го гравця, $N_i \geq 2$ – кількість чистих стратегій, $Z(x^i)$ – розподіл випадкового програшу x^i $\forall i \in D$.

Чисті стратегії гравців $\{U^i\} \forall i \in D$ надходять на входи стохастичного середовища $E: \{U^i\}_{i\in D} \to \{x^i\}_{i\in D}$, реакції якого доступні гравцям у вигляді випадкових програшів $\{x^i\} \forall i \in D$. В умовах невизначеності модель середовища E задається випадковими розподілами $Z(x^i)$ $\forall i \in D$, які не відомі агентам апріорі.

Нехай динамічна самоорганізація МАС полягає у повторюваності послідовностей чистих стратегій гравців на проміжках дискретних моментів часу $t = N_i$.

Випадкові програші $x^i = x^i(t, u^{D_i}, w)$ є функціями спільних стратегій $u^{D_i} \in U^{D_i} = \underset{j\in D_i}{\times} U^j$ агентів з локальних підмножин $D_i \subseteq D$, $D_i \neq \varnothing$ $\forall i \in D$, відрізку часу t та випадкового шансу $w \in [0,1]$.

Кроки стохастичної гри повторюються у дискретні моменти часу $n = 1, 2, \ldots$. За один крок гри агенти $\forall i \in D$ здійснюють у момент часу n незалежний випадковий вибір власних чистих стратегій $\{u_n^i\}$, за що отримують від середовища до моменту часу $n+1$ випадкові програші $\{x_n^i\}$.

Для формулювання мети гри припустимо, що випадкові програші $\{x_n^i\}$ є незалежними $\forall u_n^{D_i} = u_n^{D_i} \in U^{D_i}$, $\forall i \in D$, $n = 1, 2, \ldots$, мають постійне математичне сподівання $M\{x_n^i\} = v(t, u^{D_i}) = const$ та обмежений другий момент $\sup_n M\{[x_n^i]^2\} = s^2(t, u^{D_i}) < \infty$. Стохастичні характеристики випадкових програшів не відомі агентам апріорі.

Середні програші агентів

$$\Xi_n^i(\{t, u_n^{D_i}\}) = \frac{1}{n}\sum_{t=1}^{n} x_t^i \quad \forall i \in D \tag{1}$$

характеризують якість проведення гри у момент часу n.

Метою кожного агента є мінімізація власної функції середніх програшів:

$$\overline{\lim_{n\to\infty}} \Xi_n^i \to \min_{\{u_n^i\}} \ \forall i \in D. \tag{2}$$

Отже, задача стохастичної гри полягає в тому, що агенти на основі спостереження поточних програшів $\{x_n^i\}$ повинні навчитися вибирати чисті стратегії $\{u_n^i\}$ так, щоб з ходом часу $n = 1, 2, \ldots$ забезпечити виконання системи критеріїв (2).

Розв'язки гри задовольнятимуть одну з умов колективної оптимальності [21], наприклад, Неша, Парето, залежно від способу формування послідовностей чистих стратегій $\{u_n^i\} \forall i \in D$.

Рівновага за Нешем

Оскільки імовірнісні характеристики випадкових величин x_n^i не відомі, то в умовах невизначеності гра повинна складатися з необмеженої (на практиці – з великої) кількості ходів $\{u_n^i\}$ $\forall i \in D$, щоб забезпечити достовірне виконання однієї з умов асимптотичної оптимальності. За відсутності прямого обміну інформацією між гравцями розв'язки стохастичної гри задовольнятимуть умову колективної e-рівноваги за Нешем:

$$\forall i \in D \ \overline{\lim_{n\to\infty}} \left[\Xi_n^i(\{u_n^{D_i}\}) - \Xi_n^i(\{\breve{u}_n^{D_i}\}) \right] \leq e\ , \tag{3}$$

де нерівність (3) виконується з імовірністю 1, $u_n^{D_i}, \tilde{u}_n^{D_i} \in U^{D_i}$; $\breve{u}_n^{D_i} = u_n^{D_i} \setminus u_n^i + \tilde{u}_n^i \in U^{D_i}$; $u_n^i, \tilde{u}_n^i \in U^i$; $e \geq 0$. Якщо $e = 0$, то з (3) отримуємо умову абсолютної рівноваги за Нешем.

Нехай вибір чистих стратегій $u_n^i = u^i \in U^i$ кожен гравець здійснює $i \in D$ у дискретні моменти часу $n = 1,2,...$ з умовними імовірностями:

$$p_n^i(j) = P\{u_n^i = u^i(j) \mid u_t^i, x_t^i(t = 1..(n-1))\}, \ j = 1..N_i .$$

Послідовність векторів $\{p_n^i \mid \forall i \in D; n = 1,2,...\}$ повністю визначає властивості послідовностей чистих стратегій $\{u_n^i \mid \forall i \in D; n = 1,2,...\}$. Нехай ця послідовність при $n \to \infty$ забезпечує з імовірністю 1 досягнення точки $p^* = (p^{i*} \mid \forall i \in D)$ рівноваги за Нешем з середньоквадратичною швидкістю, не меншою за n^{-k}, де $k \in (0,1)$, тобто:

$$\overline{\lim_{n\to\infty}} n^k M\left\{ \left\| p_n^i - p^{i*} \right\|^2 \right\} < \infty \ \ \forall i \in D .$$

Тоді з ймовірністю 1 виконується умова асимптотичної рівноваги за Нешем (3) та значення функції Ξ_n^i середніх програшів гравців наближаються до значень функцій V^i середніх програшів відповідної матричної гри:

$$\forall i \in D \ \ \overline{\lim_{n\to\infty}} n^k M\left\{ \left[\Xi_n^i - V^i(p^{D_i *}) \right]^2 \right\} < \infty . \tag{4}$$

Функції середніх програшів матричної гри мають полілінійний вигляд:

$$V_i(p) = \sum_{u \in U^{D_i}} v_i(u) \prod_{j \in D_i; u_j \in u} p_j(u_j) \quad \forall i \in D , \tag{5}$$

де $p_j \in S^{N_j}$ – змішана стратегія j-го гравця, яка набуває значення на одиничному симплексі $S^{N_j} = \left\{ p \middle| \sum_{k=1}^{N_j} p_j(k) = 1; \ p_j(k) \geq 0 \right\}$; $p \in S^{D_i}$ – колективна змішана стратегія зі значеннями на симплексі $S^{D_i} = \prod_{j \in D_i} S^{N_j}$; $v_i(u) = M\{x_{i,t}(u)\}$ – середній програш i-го гравця для колективної стратегії $u \in U^{D_i}$.

Умова (4) визначає асимптотичну адекватність функції середніх програшів (1) стохастичної гри в умовах невизначеності та функції середніх програшів (5) детермінованої матричної гри. Доведення (4) виконано в [20].

Для функцій середніх програшів (5) рівновага за Нешем описує такі стратегії розв'язування гри, для яких виконується умова [20, 21]:

$$\forall i \in D \quad V_i(p^*) - V_i(p^{D \setminus i*}, p_i) \leq 0 , \tag{6}$$

де $p^* \in S^D$ – оптимальна колективна змішана стратегія гравців; $V_i(p^{D \setminus i*}, p_i)$ – функція середніх програшів, визначена на опуклому симплексі S^D за довільного відхилення змішаної стратегії i-го гравця від точки рівноваги за Нешем у межах одиничного симплексу.

Розв'язок матричної гри з локальними зв'язками є урівноваженим за Нешем у тому розумінні, що будь-яке відхилення від оптимальної стратегії p^{i*} кожного гравця $i \in D$ не може забезпечити йому зменшення функції середнього програшу $V^i(p^{D_i})$, якщо гравці з локальної підмножини D_i дотримуються точки рівноваги $p^{D_i \setminus i*}$. Розв'язок Неша відображає приватну вигоду кожного гравця – він не може власними зусиллями зменшити свій програш, якщо інші гравці дотримуватимуться стану рівноваги.

Оптимальні за Нешем змішані стратегії матричної гри можна отримати з умови доповняльної нежорсткості [21]:

$$\nabla_{p_i} V_i = V_i e^{N_i} \quad \forall i \in D , \tag{7}$$

де $\nabla_{p_i} V_i$ – градієнт полілінійної функції середніх програшів (16); $e^{N_i} = (1_j \mid j = 1..N_i)$ – вектор, всі елементи якого дорівнюють 1.

Змішані стратегії, які задовольняють умову (7), називаються вирівнювальними. Вирівнювальні стратегії є підмножиною розв'язків гри за Нешем.

Умова доповняльної нежорсткості, зважена елементами векторів змішаних стратегій, описує розв'язки гри як у змішаних, так і у чистих стратегіях:

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

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

В умовах апріорної невизначеності матриць програшів оптимальний розв'язок гри не можна отримати явно із системи рівнянь (8). Для цього необхідно застосувати процедуру його цілеспрямованого пошуку на системі одиничних симплексів.

Метод розв'язування стохастичної гри

Для розв'язування задачі (2) необхідно визначити спосіб формування послідовностей чистих стратегій $\{u_n^i\}$ у часі, які забезпечують виконання умови (8) внаслідок асимптотичної адекватності функцій середніх виграшів (4).

Враховуючи величину періоду $t = N_i$ динамічної самоорганізації МАС, формування послідовностей чистих стратегій $\{u_n^i\}$ з потрібними властивостями виконаємо на основі матриць імовірностей переходів між чистими стратегіями агентів:

$$p_n^i = \begin{bmatrix} p_n^i(1,1), & ..., p_n^i(1, N_i) \\ ... & ... \quad ... \\ p_n^i(N_i,1), & ..., p_n^i(N_i, N_i) \end{bmatrix}, \; \forall i \in D. \tag{9}$$

Рядки матриці p_n^i є змішаними стратегіями i-го гравця, якщо він вибрав чисту стратегію $u_n^i \in U^i$. Елементи $p_n^i(j,k)$ рядків є умовними імовірностями вибору чистих стратегій залежно від поточного варіанта дії u_n^i та отриманого програшу x_n^i. Вважатимемо, що вибрані чисті стратегії відповідають станам агента. Тоді p_n^i (9) є матрицею імовірностей зміни станів агента.

Гра розпочинається з ненавчених змішаних стратегій $p_0^i(j,k) = 1/N_i$, де $j,k = 1..N_i$. Для адаптивного формування розподілу випадкових стратегій, який мінімізує функції середніх програшів (1) усіх гравців, імовірність вибору стратегій з меншими програшами повинна зростати у часі $n = 1, 2, ...$.

Побудову методу формування векторів змішаних стратегій виконаємо на основі стохастичної апроксимації [22] умови доповняльної нежорсткості:

$$diag(p^i)[\nabla_{p^i} V^i - V^i e^{N_i}] = M\{x_n^i [e(u_n^i) - p_n^i] \,|\, p_n^i = p^i\}, \tag{10}$$

де $e(u_n^i)$ – визначений у момент часу n одиничний вектор-індикатор переходу в стан $u_{n+1}^i \in U^i$.

Враховуючи (10), отримаємо такий рекурентний метод зміни векторів змішаних стратегій:

$$p_{n+1}^i(u_n) = p_{e_{n+1}} \left\{ p_n^i(u_n) - g_n x_n^i \left[e(u_n^i) - p_n^i(u_n) \right] \right\}, \; \forall i \in D, \tag{11}$$

де $p_n^i(u_n)$ – змішана стратегія i-го гравця у стані $u_n \in U^i$; $p_{e_{n+1}}$ – оператор проектування на одиничний e-симплекс $S_e^{N_i} \subseteq S^{N_i}$ [20], який є підмножиною одиничного симплексу S^{N_i}; $g_n > 0$ – монотонно спадна послідовність додатних величин, яка регулює величину кроку методу; $e_n > 0$ – монотонно спадна послідовність додатних величин, яка регулює швидкість розширення e-симплексу.

Метод (11) забезпечує адаптивний вибір агентами чистих стратегій завдяки динамічній перебудові змішаних стратегій на основі опрацювання поточних програшів.

На підставі поточного розподілу імовірностей $p_n^i(u_n^i)$ агент здійснює випадковий вибір чистої стратегії $\forall i \in D$:

$$u_n^i = \left\{ u^i(l) \,|\, l = \arg\min_l \sum_{k=1}^{l} p_n^i(j,k) > w \;\; (j,l = 1..N_i) \right\} , \qquad (12)$$

де $w \in [0,1]$ – випадкова величина з рівномірним розподілом.

Отже, якщо в момент часу n агент перебуває у стані u_n^i, то на основі змішаної стратегії $p_n^i(u_n)$ він вибирає чисту стратегію $u_n^{i''}$ згідно з (4), за що до моменту часу $n+1$ отримує поточний програш x_n^i, який використовує для обчислення змішаної стратегії $p_{n+1}^i(u_n)$ згідно з (11), після чого переходить у новий стан $u_{n+1}^i = u_n^{i''}$.

Збіжність ігрового методу

Працездатність рекурентного методу (11) у розумінні виконання умови (10) в асимптотиці часу забезпечується умовою псевдоградієнтності [20]

$$r_n = \left\langle diag(p_n^i)(e^{N_i} V_n^i - \nabla V_n^i), p_n^i - \tilde{p}_n^i \right\rangle = V_n^i \Delta_n^i \ge 0 \qquad (13)$$

його вектора руху відносно функції Ляпунова:

$$\Delta_n = \sum_{i \in D} \Delta_n^i = \sum_{i \in D} \left\| p_n^i - \tilde{p}_n^i \right\|^2 , \qquad (14)$$

де $p_n^i = p_n^i(u^i)$; $\tilde{p}_n^i = diag(p_n^i) \nabla_{p_n^i} V_n^i \big/ V_n^i$ – зважена змішана стратегія *i*-го гравця, яка визначає поточне відхилення від виконання умови доповняльної нежорсткості. Умова (13) достовірно виконується у знакододатному середовищі, для якого $v_{\min} = \min_{i \in D} \min_{u^{D_i}} v^i(u^{D_i}) > 0$.

Спрямування Δ_n до нуля, якщо $n = 1, 2, \ldots$, свідчитиме про збіжність ігрового методу до однієї із точок рівноваги за Нешем у змішаних стратегіях, або, враховуючи (8), про досягнення розв'язку гри у чистих стратегіях.

З урахуванням (13) для методу (11) отримана усереднена щодо реалізації подій верхня оцінка:

$$M\{\Delta_{n+1}\} \le (1 - 2 g_n v_{\min}) M\{\Delta_n\} + C(m_n + g_n^2) ,$$

де $C \sim |D| N_{\max} (v_{\max} - v_{\min}) s_{\max}^2 > 0$, $m_n = |g_{n-1} - g_n| + |e_n - e_{n+1}|$, $v_{\max} = \max_{i \in D} \max_{u^{D_i} \in U^{D_i}} v(u^{D_i})$.

Нехай для $n \to \infty$ виконуються умови: $g_n > 0$; $g_{n+1} < g_n$; $\sum_{n \to \infty} g_n = \infty$; $e_n \in (0, \min_{i \in D} N_i^{-1})$; $e_{n+1} < e_n$; $\lim_{n \to \infty} e_n = 0$. Тоді проекційний ігровий метод (11) для будь-якого початкового наближення $p_1^i \in S_{e_1}^{N_i}$; $\forall i \in D$ забезпечує виконання умови доповняльної нежорсткості (8) у знакододатному середовищі $v_{\min} > 0$:

1) з ймовірністю 1, якщо $\sum_{n=1}^{\infty} (|e_n - e_{n-1}| + g_n^2) < \infty$;

2) у середньоквадратичному, якщо $\lim_{n \to \infty} (|e_n - e_{n-1}| g_n^{-1} + g_n) = 0$.

Дослідження збіжності методу (11) виконаємо у класі монотонних послідовностей $\{g_n\}$ та $\{e_n\}$ виду

$$g_n = g(n+a)^{-a}; a > 0; \; e_n = e(n+b)^{-b}; b > 0 . \qquad (15)$$

Збіжність методу (11) спостерігається:

1) з ймовірністю 1, якщо $a \in (0.5;1]$; $b > 0$;

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

Порядок q та величину J швидкості збіжності ігрового методу можна оцінити за допомогою асимптотичного методу моментів Чжуна [22]:

$$\overline{\lim_{n \to \infty}} n^q M\{\Delta_n\} = J . \qquad (16)$$

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

Якщо виконані умови збіжності методу (11) у середньоквадратичному для послідовностей $\{g_n\}$, $\{e_n\}$ виду (15), зокрема, $a \in (0,1]$, $b > 0$, і якщо $a = 1$, виконується нерівність $2g\,v_{\min} > \min(1, b)$, тоді

$$\overline{\lim_{n\to\infty}}\, n^q M\{\Delta_n\} \le \frac{Cm}{2g\,v_{\min} - \min(1,b)\,c\,(a=1)} < \infty\,, \tag{17}$$

де $m = ebc\,(1 + b \le 2a) + gc\,(a = 1 \wedge b \ge 1) + g^2 c\,(1 + b \ge 2a)$; $q = \min(1 + b - a, a)$ – порядок швидкості збіжності, $c\,() \in \{0,1\}$ – індикаторна функція події.

В умові (17) параметр q, який задає порядок швидкості збіжності методу (11), задовольняє нерівність $q \le 1$. Максимальний порядок швидкості збіжності методу (11) становить $q = 1$ і досягається, якщо $a = 1$, $b \ge 1$.

На практиці можна скористатися наближеною оцінкою для визначення q, виконавши логарифмування виразу (16):

$$q \approx \frac{\lg(M\{\Delta_n\})}{\lg(n)}\,. \tag{18}$$

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

Контрольний приклад

Для прикладу розглянемо стохастичну ігрову модель самоорганізації комах-світлячків (fireflies) із родини Lampyridae, які ведуть нічний спосіб життя у тропічних регіонах світу. Самці цих комах для приваблювання самок запускають механізм люмінесцентного випромінювання свого черевця. Самоорганізація проявляється у виникненні явища ритмічного синхронізованого світіння усієї колонії самців.

Моделювання поведінки світлячків виконаємо за допомогою стохастичної гри агентів, кожен з яких може перебувати в одному із двох станів $u_n^i \in \{0,1\}$, де 0 позначає відсутність, а 1 – наявність світіння.

Кожен агент може спостерігати стани сусідніх агентів та змінювати власний стан так, щоб у діях бути максимально подібним на своїх сусідів. Структура зв'язків між агентами зображена на рис. 1.

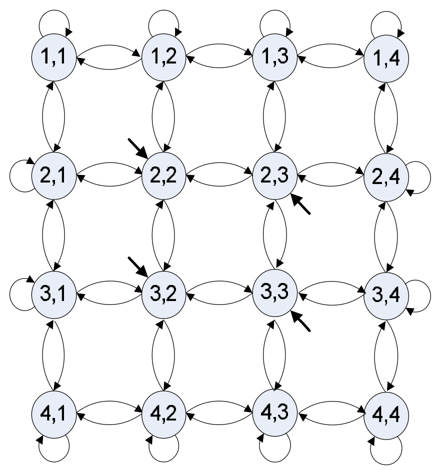

Рис. 1. Структура МАС

Регулярна структура гри задається кількістю агентів $L = m \times m$, $m \ge 2$, підмножинами сусідніх агентів D_i та кількістю чистих стратегій $N_i = N = 2$, $i = 1..L$.

Динаміка процесу самоорганізації складається з просторової та часової координації стратегій агентів. Просторова координація полягає у дотриманні співвідношення стратегій агентів у локально визначених областях D_i $\forall i \in D$ так, як це зображено на рис. 1. Часова координація визначається дотриманням співвідношення стратегій агентів на проміжку часу $t = 2$.

В ігровій термінології просторова координація полягатиме у виборі однакових значень чистих стратегій гравців у фіксовані моменти часу (агенти намагаються повторювати дії один одного), а часова координація – у зміні бінарних стратегій на протилежні значення у два послідовні моменти часу. Результатом самоорганізації агентів є інверсна зміна матриць бінарних чистих стратегій $[0]_{m\times m} - [1]_{m\times m} - [0]_{m\times m} - [1]_{m\times m} - ...$ у часі, що моделює ритмічне світіння колонії світлячків. Така зміна забезпечується матрицями навчених змішаних стратегій $p^{i*}(u_n^i) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ агентів з однаковими початковими станами.

Поточні програші агентів визначимо як штраф за порушення просторової та часової координації стратегій:

$$x_n^i = l \sum_{s \in D_i} | u_n^i - u_n^s | / L_i + (1 - l) | \bar{u}_n^i - u_{n-1}^i | + m_n , \qquad (19)$$

де $x_n^i \in R^1$; $l \in [0,1]$ – ваговий коефіцієнт; D_i – множина сусідніх агентів, що відповідає зображеним на рис. 1 зв'язкам; $L_i = | D_i |$ – кількість сусідніх агентів; u_n^i – чиста стратегія з бінарним значенням; $\bar{u}_n^i$ – інверсне значення чистої стратегії; $m_n \sim Normal(0, d)$ – білий гауссівський шум, нормально розподілена випадкова величина з нульовим математичним сподіванням та дисперсією $d > 0$.

Перша складова виразу (19) визначає штраф за порушення просторової (взаємної) координації стратегій гравців у межах підмножини D_i, друга складова – штраф за порушення часової координації у два послідовні моменти часу, а третя складова визначає дію випадкових завад у вигляді білого шуму.

Оцінювання ефективності ігрової самоорганізації МАС виконаємо за такими показниками:

1) функція середніх втрат або ціна гри:

$$\Xi_n = \frac{1}{L} \sum_{i=1}^{L} \Xi_n^i , \qquad (20)$$

де $L = | D |$ – кількість гравців;

2) коефіцієнт просторової координації стратегій гравців:

$$\mathrm{K}_n = \frac{1}{nL} \sum_{t=1}^{n} \sum_{i=1}^{L} c \left(\sum_{s \in D_i} | u_t^i - u_t^s | = 0 \right), \qquad (21)$$

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

3) коефіцієнт часової координації стратегій гравців:

$$M_n = \frac{1}{n} \sum_{t=1}^{n} \sum_{i=1}^{L} c \left(| \bar{u}_t^i - u_{t-1}^i | = 0 \right); \qquad (22)$$

4) середнє відхилення змішаних стратегій гравців від оптимальних значень $p^{i*}(u_n^i)$:

$$\Delta_n = \frac{1}{nL} \sum_{t=1}^{n} \sum_{i=1}^{L} \sum_{u_t^i \in U^i} \left\| p_t^i(u_t^i) - p^{i*}(u_t^i) \right\| , \qquad (23)$$

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

Результати комп'ютерного моделювання

Характеристики ефективності самоорганізації стохастичної гри Ξ_n, K_n, M_n та Δ_n подано на рис. 2 у логарифмічному масштабі. Графіки отримано для гри $L = 4 \times 4$ агентів з параметрами: $g = 1$, $e = 0.999/2$, $a = 0.01$, $b = 2$, $l = 0.5$, $d = 0.36$. Гра складається з $n_{\max} = 10^4$ кроків.

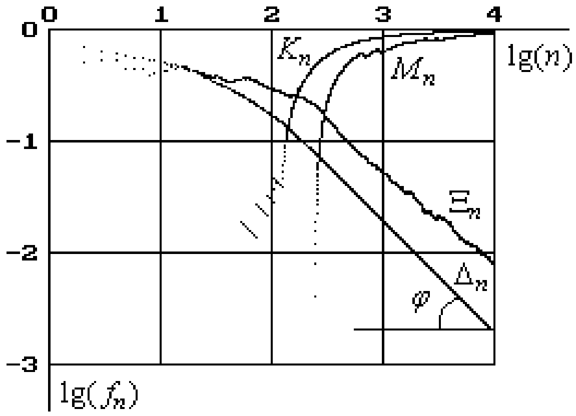

Рис. 2. Характеристики самоорганізації стохастичної гри

Спадання функцій середніх програшів Ξ_n (20) та норми Δ_n (23) відхилення змішаних стратегій від оптимальних значень у часі свідчать про збіжність ігрового методу внаслідок виконання системи критеріїв (2). Графік коефіцієнта просторової координації стратегій K_n (21) ілюструє зміну координації стратегій у межах множини гравців, а графік коефіцієнта часової координації стратегій M_n (22) – ритмічну зміну стратегій (інверсну координацію) гравців у послідовні моменти часу. Зростання цих функцій у часі свідчать про досягнення динамічної самоорганізації стохастичної гри.

Порядок швидкості збіжності гри можна оцінити тангенсом кута j , утвореним між графіком лінійної апроксимації норми відхилення змішаних стратегій гравців та віссю часу для навченої гри. Для заданих параметрів гри забезпечується близький до 1 порядок швидкості збіжності.

Запропонований метод (11) розв'язування стохастичної гри агентів належить до класу реактивних методів (основаних на опрацюванні поточних реакцій середовища на дії агентів) і має порівняно невисоку, степеневу швидкість збіжності. Це пов'язано з тим, що на початку гри агенти не мають ніякої інформації про середовище, з яким вони взаємодіють. Інформацію збирають у процесі навчання, адаптивно перебудовуючи вектори змішаних стратегій пропорційно до значень поточних програшів.

Динаміка вибору чистих стратегій стохастичної гри зображена на рис. 3 та рис. 4. Точками на графіках відмічено середнє значення стратегій гравців:

$$\overline{u}_n = L^{-1}\sum_{i=1}^{L} u_n^i \ .$$

Період графіків відповідає двом послідовним моментам часу. На початковому відрізку часу усереднені стратегії гравців набувають випадкових значень між 0 та 1, що відповідає хаотичному вибору станів ненавченими агентами (рис. 3). Після навчання ($n \sim 10^3$ кроків) відбувається скоординована у часі, ритмічна зміна бінарних стратегій (рис. 4). Уся множина гравців синхронно і почергово переходить зі стану $u_n^i = 0$ у стан $u_{n+1}^i = 1$, де $i = 1..L$.

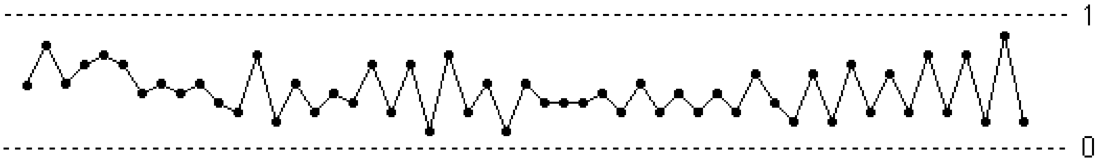

Рис. 3. Нескоординований вибір стратегій ненавченої стохастичної гри

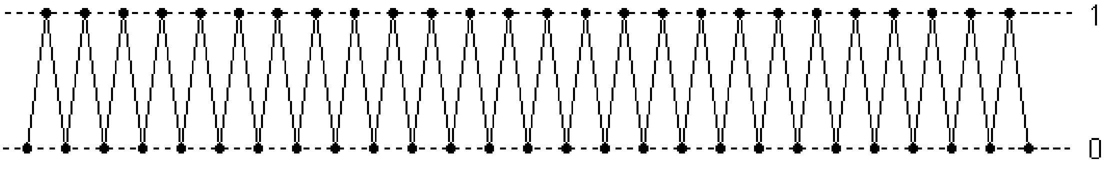

Рис. 4. Динамічно скоординований вибір стратегій навченої стохастичної гри

Ваговий коефіцієнт l визначає пропорцію складників штрафів за порушення просторової та часової координації стратегій у сумарному поточному програші x_n^i (19). Динамічна координація визначається певним балансом між цими штрафами. У разі зростання l до 1 забезпечується просторова координація стратегій, але порушується ритм їх зміни у часі. Вплив параметра l на ефективність ігрового методу (11) зображено на рис. 5 у вигляді графіків функцій K_n (21) та M_n (22), отриманих для $n = n_{\max}$ та $d = 0$. Вісь абсцис визначає значення параметра l , а вісь ординат – відсоток скоординованих гравців.

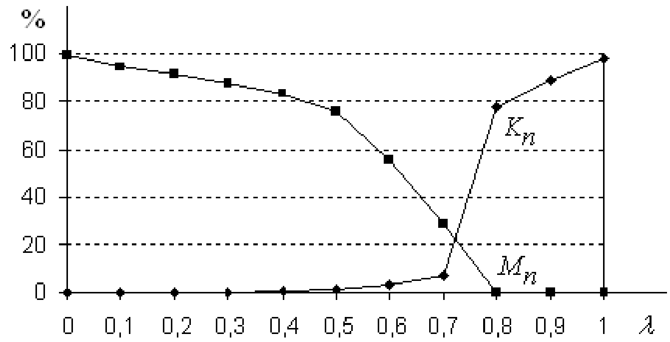

Рис. 5. Вплив вагового коефіцієнта l на ефективність ігрового методу для $d = 0$

Як видно на рис. 5, за допомогою параметра l неможливо одночасно досягти належного рівня просторової координації та ритмічної зміни стратегій гравців. Експериментально встановлено, що потрібного рівня можна досягти впливом білого шуму m_n, який забезпечує випадкове, незміщене коливання значень штрафів за порушення просторової та часової координації стратегій гравців. Відповідні результати, отримані для $a = 0.01$ та $d = 0.36$, подано на рис. 6.

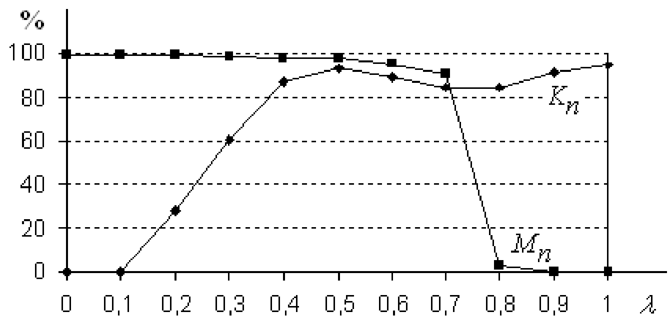

Рис. 6. Вплив вагового коефіцієнта l на ефективність ігрового методу для $d = 0.36$

Дія білого шуму забезпечує динамічну координацію стратегій МАС для значень параметра $l \in [0.4; 0.7]$ на рівні, що перевищує 80 % гравців.

Висновки та перспективи подальших наукових розвідок

Розроблена ігрова модель забезпечує динамічну самоорганізацію МАС, яка проявляється у ритмічній зміні чистих стратегій агентів, що імітує світлові ефекти колонії комах-світлячків. Характерною особливістю розглянутої ігрової самоорганізації є локально обумовлений збір інформації про стратегії поведінки сусідніх агентів, який у результаті навчання приводить до глобальної координації стратегій усіх агентів.

Генерування послідовностей чистих стратегій з потрібними властивостями забезпечується випадковим розподілом, побудованим на динамічних змішаних стратегіях гравців. Обчислення змішаних стратегій здійснюється за адаптивним рекурентним методом, отриманим на основі стохастичної апроксимації умови доповняльної нежорсткості, яка описує колективні розв’язки гри, що задовольняють умову рівноваги за Нешем.

Збіжність ігрового методу визначається обмеженням значень його параметрів згідно з результатами теорії стохастичної апроксимації та рекурентного оцінювання. Швидкість ігрової самоорганізації МАС залежить від кількості гравців, кількості стратегій, інтенсивності завад та параметрів ігрового методу. Належно підібравши параметри ігрового методу, можна досягти близького до 1 степеневого порядку швидкості збіжності.

Крім значень параметрів, самоорганізація стохастичної гри МАС визначається балансом штрафів за порушення просторової координації та часового ритму. Експериментально встановлено, що такий баланс можна забезпечити впливом білого шуму на формування випадкових поточних програшів.

Ефективність ігрової самоорганізації стратегій МАС вивчали за допомогою функцій середніх програшів, коефіцієнтів координації та норми відхилення динамічних змішаних стратегій від оптимальних значень. Спадання функції середніх програшів і функції відхилення змішаних стратегій, зростання коефіцієнтів координації свідчать про збіжність ігрового методу та входження МАС у режим самоорганізації. Повторення значень характеристик гри у різних експериментах з унікальними послідовностями випадкових величин підтверджує достовірність отриманих результатів.

1. Великий тлумачний словник сучасної української мови / уклад. і голов. ред. В.Т. Бусел. – К.; Ірпінь: ВТФ “Перун”, 2001. – 1440 с. 2. Декарт Р. Міркування про метод, щоб правильно спрямовувати свій розум та відшуковувати істину в науках / Рене Декарт; пер. з фр. В. Андрушко, С. М. Гатальська. – К.: Тандем, 2001. – 102 с. 3. Ashby W. R. Principles of the Self-Organizing Dynamic System / W. R. Ashby // Journal of General Psychology. – 1947. – Vol. 37. – P. 125–128. 4. Grasse P. P. La reconstruction du nid et les coordinations inter-individuelles chez Bellicositermes natalensis et Cubitermes sp. La theorie de la stigmergie: essai d’interpretation des termites constructeurs / P. P. Grasse // Insect Society. – 1959. – Vol. 6. – P. 41–84. 5. Пригожин И. Самоорганизация в неравновесных системах: От диссипативных структур к упорядоченности через флуктуации / И. Пригожин, Г. Николис. – М.: Мир, 1979. – 512 с. 6. Хакен Г. Информация и самоорганизация. Макроскопический подход к сложным системам: пер. с англ. / Г. Хакен. – М.: КомКнига, 2005. – 248 с. 7. Эйген М. Гиперцикл. Принципы самоорганизации макромолекул / М. Эйген, П. Шустер. – М.: Мир, 1982. – 260 с. 8. Арнольд В. И. Теория катастроф / В. И. Арнольд. – М.: Эдиториал УРСС, 2004. – 128 с. 9. Рассел Д. Теория хаоса / Джесси Рассел, Рональд Кон. – М.: Книга по требованию. – 2012. – 110 с. 10. Николис Г. Познание сложного. Введение / Г. Николис, И. Пригожин. – М.: Мир, 1990. – 344 с. 11. Хакен Г. Синергетика: пер. с англ. / Г. Хакен. – М.: Мир, 1980. – 405 с. 12. Коновалов А. И. Супрамолекулярные системы – мост между неживой и живой материей / А. И. Коновалов. – М.: РБОФ “Знание” им. С. И. Вавилова, 2010. – 28 с. 13. Горбань О. М. Основи теорії систем і системного аналізу / О. М. Горбань, В .Є. Бахрушин. – Запоріжжя: ДУ “ЗІ–ДМУ”, 2004. – 204 с. 14. Самоорганизация и многоагентные системы. I. Модели многоагентной самоорганизации / В. И. Городецкий // Изв. РАН. Теория и системы управления. – 2012. – № 2. – С. 92–120. 15. Самоорганизация и многоагентные системы. II. Приложения и технология разработки / В. И. Городецкий // Изв. РАН. Теория и системы управления. – 2012. – № 3. – С. 55–75. 16. Wooldridge M. An Introduction to Multiagent Systems / M. Wooldridge. – John Wiley & Sons, 2002. – 366 pp. 17. Блехман И. И. Синхронизация в природе и технике / И. И. Блехман. – М.: Наука, Главная редакция физико-математической литературы, 1981. – 352 с. 18. Fudenberg D. The Theory of Learning in Games / D. Fudenberg, D. K. Levine. – Cambridge, MA: MIT Press, 1998. – 292 p. 19. Доманский В.К. Стохастические игры / В.К. Доманский // Математические вопросы кибернетики. – 1988. – № 1. – С. 26–49. 20. Назин А. В. Адаптивный выбор вариантов: Рекуррентные алгоритмы / А. В. Назин, А. С. Позняк. – М.: Наука, 1986. – 288 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.