

Міністерство освіти і науки України

Національний університет «Києво-Могилянська академія»

Факультет інформатики

Кафедра математики

Магістерська робота

освітній ступінь – магістр

на тему: «**СТОХАСТИЧНІ ІГРИ НА ГРАФАХ**»

Виконав: студент 2-го року навчання
освітньо-наукової програми

«Системний аналіз»,

спеціальності 124 Системний аналіз

Бутовський Владислав Валерійович

Керівник: Чорней Р. К.,

кандидат фіз.-мат. наук, доцент

Рецензент _____

(прізвище та ініціали)

Магістерська робота захищена

з оцінкою _____

Секретар ЕК _____

« ____ » _____ 20 ____ р.

Київ – 2021

ЗМІСТ

	Стор.
Анотація	3
<i>ВСТУП</i>	4
РОЗДІЛ 1: Постановка задачі	6
1.1 Вступ.	6
1.2 Граф стохастичної гри.	6
1.3 Функція прибутку.	7
1.4 Механізм переходу.	8
1.5 Сідлова точка.	9
РОЗДІЛ 2: Застосунок для пошуку оптимальної стратегії стохастичної гри	11
2.1 Чисельний приклад	11
2.1.1 Опис гри.	11
2.2 Алгоритм пошуку оптимальної стратегії.	13
2.2.1 Теоретичний матеріал для пошуку оптимальної стратегії.	13
2.2.2 Програмний застосунок	14
2.2.3 Алгоритм пошуку оптимальної стратегії.	17
2.3 Висновок.	20
РОЗДІЛ 3: Аналіз функції виграшу та задачі максимізації	21
3.1 Вступ.	21
3.2 Дослідження функції середнього виграшу для скінченої гри.	21
3.3 Задача максимізації виграшу в умовах неповної інформації.	22
3.4 Висновок.	25
Висновок	26
Література.	27

АНОТАЦІЯ

Мета роботи – знаходження оптимальних стратегій в стохастичних іграх на графі з двома коаліціями та нульовою сумою.

Об'єктом дослідження є багатокрокова стохастична гра на графі, яка складається зі скінченної кількості гравців та двох коаліцій.

Предметом дослідження є стратегії керування коаліційною грою на графі.

Методами дослідження є методи теорії керування, теорії випадкових процесів, метод простих ітерацій.

Вступ

Актуальність даної роботи пов'язана зі збільшенням зацікавленості до використання розділу прикладної математики – теорії ігор, у сучасних економічних [6] і соціальних областях[7].

Стохастичні ігри – один з розвинутих розділів теорії ігор. Грає ключову роль в конструюванні моделей в багатьох областях: теорія контрактів[5],[10], теорія корпоративних фінансів[9], мікро та макроекономіка [5],[8] тощо.

Кожна стохастична гра має правила, які заздалегідь повинні бути визначеними, як з кількістю учасників гри, так і з обмеженнями щодо дій гравців. Після кожного етапу гри за допомогою функції виграшу кожен гравець отримує свій прибуток. Прибуток може бути позитивним і від'ємним. Ключова ціль використання теорії ігор – це знаходження найліпшого варіанту вибору дій кожного з гравців та оптимальних стратегій. Поділяють два види стохастичних ігор. До першого виду відносяться однокрокові ігри, які, після першого етапу, одразу закінчуються. До даного виду відносяться «Дилема в'язня», «Дилема мандрівників». До другого – багатокрокові ігри, які тривають більш, ніж один етап.

Концепція рішення ключового поняття теорії ігор є рівновага за Нешем або рівновага Неша. Рівновага Неша виникає лише за умовою, коли жоден учасник не може збільшити виграш, змінивши свою стратегію, якщо інші учасники своїх стратегій не змінюють. Рівновага також виникає за умов, коли кожна компанія, яка бере участь у вирішенні питання, прагне максимізувати свій прибуток.

Кооперативні ситуації виникають під час гри, коли більш ніж один гравець утворює коаліції. Кооперативна теорія ігор в класичній постановці – це створення групи задля максимізації виграшу. Створення будь-якої коаліції зазвичай вимагає координаційної групної діяльності учасників, що не може не вплинути на певні додаткові витрати.

Мета роботи – знаходження оптимальних стратегій в стохастичних іграх на графі з двома коаліціями та нульовою сумою.

Методами дослідження є методи теорії керування, теорії випадкових процесів, метод простих ітерацій.

Дипломна робота складається з трьох розділів.

У першому розділі міститься постановка задачі і теоретичні відомості.

У другому розділі надана інформація про чисельний приклад для визначеної стохастичної гри, алгоритм пошуку оптимальної стратегії та програмний застосунок.

У третьому розділі міститься аналіз функції середнього виграшу для однокрокової та багатокрокової гри, досліджена задача максимізації виграшу в умовах неповної інформації.

1 Постановка задачі

1.1 Вступ

В розділі розглянемо постановку задачі стохастичної гри, визначеної на графі.

1.2 Граф стохастичної гри

Нехай дано скінченний неорієнтований граф $G = (V, E)$, де V – множина вершин графу, а E – множина ребер. Згідно з означенням [1, с.282] пара суміжних вершин (a, b) – інтерпретує ребро графа.

Сусідами вершини a є множина

$$N(a) = \{b : (a, b) \in E\} \quad (1.2.1)$$

Повним околom вершини a вважатимемо:

$$\tilde{N}(a) = N(a) \cup \{a\}, \quad (1.2.2)$$

тобто дана множина включає вершину a та множину $N(a)$.

У кожен момент часу t гравці обох коаліцій приймають рішення на основі повної інформації станів сусідів та попередніх особистих станів. Після спільного рішення всіх гравців i -й гравець (у i -й вершині) отримує від j -го гравця (позитивну чи негативну) виплату (винагороду).

Введемо параметр

$$\xi = (\xi_t, t \in \mathbb{N}), \quad (1.2.3)$$

що буде характеризувати залежність стану системи від часу t . Припускаємо, що t приймає дискретні значення $t = 0, 1, 2, \dots$

ξ_k^t – граничний розподіл за час t для вершини k .

У грі ξ приймає рандомізовані значення.

Для кожної вершини $i \in E$ X_i метричний простір зі зліченим базисом. А метричний простір станів гри позначимо за

$$X := \prod_{i \in V} X_i \quad (1.2.4)$$

Гравці розташовані у вершинах графу. Ребра демонструють зв'язки між гравцями. Гравці поділені на дві коаліції K_1 та K_2 . Кожна коаліція може складатись з скінченної кількості гравців. Виграш одної коаліції є програшом іншої, таким чином розглядаємо гру з нульовою сумою.

Нехай дано набір допустимих дій для кожного гравця i : A_i , згідно з означенням 2.4 [1, с.284]:

$$A := \prod_{i \in V} A_i, \quad (1.2.5)$$

множина усіх дій системи.

A_i^t – множина дій для i -того гравця у час t . Позначимо за α_i^t – дію, яку обрав i -тий гравець у час t .

З означення 2.5 [1, с.285], якщо π – стратегія K_1 коаліції, що складається з послідовності можливих рішень гравців, які містяться у коаліції K_1 :

$$\pi = (\pi_i, i \in V_1),$$

$$\text{де } \forall i \pi_i = \{\pi_i^0, \dots, \pi_i^t, \dots\} : \pi_i^t = \pi_i^t(\cdot / x^0, a^0, \dots, x^{t-1}, a^{t-1}, x^t). \quad (1.2.6)$$

Якщо γ – множина стратегій, що складається з послідовності можливих рішень гравців, які містяться у коаліції K_2 :

$$\gamma = (\gamma_j, j \in V_2),$$

$$\text{де } \forall j \gamma_j = \{\gamma_j^0, \dots, \gamma_j^t, \dots\} : \gamma_j^t = \gamma_j^t(\cdot / x^0, a^0, \dots, x^{t-1}, a^{t-1}, x^t) \quad (1.2.7)$$

В кінці вибору стратегій кожної з коаліцій K_1 та K_2 , система переходить у наступний стан. Закон переходу буде описано пізніше.

1.3 Функція прибутку

Нехай дано функцію виграшу r , яка визначена [1, с.288] і характеризується за допомогою чотирьох параметрів

$$r_{ij}(x_i, x_j, a_i, a_j), \quad (1.3.1)$$

де x_i – стан гравця i -того,

x_j – стан гравця j -того,

a_i – дія обрана гравцем i ,

a_j – дія обрана гравцем j відповідно.

Під час кожного часового параметру t та стану системи, гравці отримують винагороди. Коаліції K_1 та K_2 отримують винагороду, яка обчислюється як сума функції винагород гравців коаліції.

Згідно з рівності [1, с.282] функції виграшу, винагорода для коаліції K_1 буде:

$$r_1(x, a) = \sum_{i \in V_1} \sum_{j \in V_2} r_{ij}(x_i, x_j, a_i, a_j) \quad (1.3.2)$$

і відповідно для коаліції K_2 :

$$r_2(x, a) = \sum_{j \in V_2} \sum_{i \in V_1} r_{ji}(x_j, x_i, a_j, a_i) \quad (1.3.3)$$

Розглядаємо гру з нульовою сумою, де виграш однієї коаліції тягне програш іншої. З [1, с.288] Має місце рівність:

$$r_{ij} = -r_{ji} \quad (1.3.4)$$

де i та j відповідно гравці-сусіди. Тоді для гри з двома коаліціями K_1 та K_2 і функціями виграшу r_1 та r_2 відповідно буде мати місце рівність:

$$r_1(x, a) = -r_2(x, a) \quad (1.3.5)$$

1.4 Механізм переходу

Кожен час t гравці обох коаліцій приймають рішення, опираючись на повну інформацію станів їх сусідів та тих рішень, які приймали раніше.

Згідно з означенням 2.2 [1, с.283], наведемо означення ймовірності переходу для стану системи:

Означення 1.4.1

Нехай $\xi = (\xi_t, t \in \mathbb{N})$ – буде деяким дискретним Марковським процесом на просторі X . Тоді ймовірності переходів будуть:

$$P\{\xi_k^{t+1} \in C_k \mid \xi^t = x^t, \dots, \xi^0 = x^0\} = P\{\xi_k^{t+1} \in C_k \mid \xi_{\bar{N}(k)}^{t+1} = x_{\bar{N}(k)}^t\} \quad (1.4.1)$$

Гравці приймають рішення одночасно, ймовірність переходу системи можна записати у вигляді:

$$P\{\xi_K^{t+1} \in C_k \mid \xi^t = x^t\} = \prod_{k \in K} P\{\xi_k^{t+1} \in C_k \mid \xi^t = x^t\}, \quad (1.4.2)$$

Для коаліцій K_1 та K_2 опишемо ймовірності, використовуючи стратегії π та γ .

Ймовірність вибору дій згідно з означенням 2.5 [1, с.285] для коаліції K_1 буде виглядати наступним чином:

$$\begin{aligned} P(\alpha_i^t \in B_i \mid \xi^0 = x^0, \alpha^0 = a^0, \dots, \xi^{t-1} = x^{t-1}, \alpha^{t-1} = a^{t-1}, \xi^t = x^t) = \\ = \pi_i^t(B_i \mid x^0, a^0, \dots, x^{t-1}, a^{t-1}, x^t) \forall i \in V_1 \end{aligned} \quad (1.4.3)$$

для коаліції K_2 та стратегій γ :

$$\begin{aligned} P(\alpha_j^t \in B_j \mid \xi^0 = x^0, \alpha^0 = a^0, \dots, \xi^{t-1} = x^{t-1}, \alpha^{t-1} = a^{t-1}, \xi^t = x^t) = \\ = \gamma_j^t(B_j \mid x^0, a^0, \dots, x^{t-1}, a^{t-1}, x^t) \forall j \in V_2 \end{aligned} \quad (1.4.4)$$

Згідно з означенням 2.5 [1, с.286] загальна ймовірність набуде наступного вигляду:

$$\begin{aligned} P\left(\alpha^t \in \times_{i \in V} B_i \mid \xi^0 = x^0, \alpha^0 = a^0, \dots, \xi^{t-1} = x^{t-1}, \alpha^{t-1} = a^{t-1}, \xi^t = x^t\right) = \\ = \prod_{i \in V} P\left(\alpha_i^t \in B_i \mid \xi^0 = x^0, \alpha^0 = a^0, \dots, \xi^{t-1} = x^{t-1}, \alpha^{t-1} = a^{t-1}, \xi^t = x^t\right) = \\ = \prod_{i \in V_1} \pi_i^t(B_{i=1} \mid x^0, a^0, \dots, x^t) \times \prod_{j \in V_2} \gamma_j^t(B_{j=1} \mid x^0, a^0, \dots, x^t) \end{aligned} \quad (1.4.5)$$

1.5 Сідлова точка

Позначимо множину локальних стратегій для гри як LS .

$\forall \pi = (\pi_i, i \in V_1), \gamma = (\gamma_j, j \in V_2), t \in \mathbb{N}$ позначимо асимптотичну середню винагороду [1, с.288] для початкового стану системи $\xi_V^0 = x$:

$$\phi(x, \pi, \gamma) = \liminf_{t \rightarrow \infty} \frac{1}{t+1} E_{\pi, \gamma}^x \sum_{k=0}^t r(\xi^k, \alpha^k) \quad (1.5.1)$$

Позначимо за β – коефіцієнт дисконтування, тоді з формули (1.5.1) отримаємо наступний вигляд:

$$\mathcal{J}(x, \pi, \gamma) = \liminf_{t \rightarrow \infty} \frac{1}{t+1} \text{Exp}_{\pi, \gamma}^x \sum_{k=0}^t r(\xi^k, \alpha^k) \beta^t \quad (1.5.2)$$

Тоді згідно з означенням 2.8 [1, с. 288] стратегія π^* буде оптимальною для коаліції K_1 , якщо для будь-якої стратегії γ' , яку оберуть гравці коаліції K_2 та усіх станів x системи буде мати місце нерівність

$$\inf_{\gamma \in LS} \sup_{\pi \in LS} \phi(x, \pi, \gamma) \leq \phi(x, \pi^*, \gamma'). \quad (1.5.3)$$

Аналогічно і для гравців коаліції K_2 . Якщо для будь-якої стратегії π' , яку оберуть гравці з коаліції K_1 , та для будь-яких станів системи x буде мати місце нерівність

$$\sup_{\pi \in LS} \inf_{\gamma \in LS} \phi(x, \pi, \gamma) \geq \phi(x, \pi', \gamma^*), \quad (1.5.4)$$

то стратегія γ^* є оптимальною стратегією для гравців коаліції K_2 .

Згідно з другою частиною означення 2.8 [1, с. 288] ціна гри c буде існувати за умови:

$$\sup_{\pi \in LS} \inf_{\gamma \in LS} \phi(x, \pi, \gamma) = \inf_{\gamma \in LS} \sup_{\pi \in LS} \phi(x, \pi, \gamma) = c \quad (1.5.5)$$

для стохастичної гри.

Для існування ціна гри є необхідність існування сідлової точки, отже, якщо справдовується рівність (3.2.6), то сідлова точка існує.

2 Застосунок для пошуку оптимальної стратегії стохастичної гри

2.1 Чисельний приклад

Розглянемо інший граф G , який зображений на рис 2.3.1:

$$G = (V, E), V = \{v_1, v_2, v_3, v_4, v_5\}, E = \{e_1, e_2, e_3, e_4, e_5\}, \\ K_1 = \{k_1, k_2\}, K_2 = \{k_3, k_4, k_5\}$$

Оскільки, вершина 1 та 3 є суміжними, то гравцю 1 відома інформація про стани гравця 3, і навпаки. Аналогічно для усіх суміжних вершин.

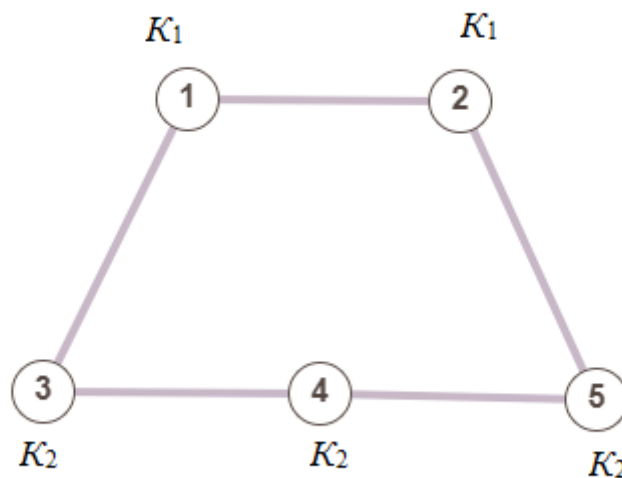


Рис. 2.1.1

2.1.1 Опис гри

Визначимо стратегії для кожної коаліції:

$$\pi = (\pi_1, \pi_2) \tag{2.1.1.1}$$

$$\gamma = (\gamma_3, \gamma_4, \gamma_5),$$

Тоді стратегіями гравців згідно з (1.2.6) та (1.2.7):

$$\pi_1 = \{\pi_1^n\}$$

$$\pi_2 = \{\pi_2^n\}$$

$$\gamma_3 = \{\gamma_3^n\}$$

$$\gamma_4 = \{\gamma_4^n\}$$

$$\gamma_5 = \{\gamma_5^n\}$$

(2.1.1.2)

Нехай кожен гравець має по дві можливі дії кожного етапу гри:

$$A_i \in \{a_{i1}, a_{i2}\}, i \in V \quad (2.1.1.3)$$

та по три можливі стани:

$$X_i \in \{x_{i1}, x_{i2}, x_{i3}\}, i \in V \quad (2.1.1.4)$$

Визначимо ймовірності вибору дій для обох коаліцій з (1.4.3) та (1.4.4):

$$\begin{aligned} P(\alpha_1^{n+1} \in B_1 \mid \xi^0 = x^0, \alpha^0 = a^0, \xi^1 = x^1, \dots, \xi^n = x^n) &= \pi_1^1(B_1 \mid x_1^n, x_2^n, x_3^n) \\ P(\alpha_2^{n+1} \in B_2 \mid \xi^0 = x^0, \alpha^0 = a^0, \xi^1 = x^1, \dots, \xi^n = x^n) &= \pi_2^1(B_2 \mid x_1^n, x_2^n, x_5^n) \\ P(\alpha_3^{n+1} \in B_3 \mid \xi^0 = x^0, \alpha^0 = a^0, \xi^1 = x^1, \dots, \xi^n = x^n) &= \gamma_3^1(B_3 \mid x_1^n, x_3^n, x_4^n) \cdot \\ P(\alpha_4^{n+1} \in B_4 \mid \xi^0 = x^0, \alpha^0 = a^0, \xi^1 = x^1, \dots, \xi^n = x^n) &= \gamma_4^1(B_4 \mid x_3^n, x_4^n, x_5^n) \\ P(\alpha_5^{n+1} \in B_5 \mid \xi^0 = x^0, \alpha^0 = a^0, \xi^1 = x^1, \dots, \xi^n = x^n) &= \gamma_5^1(B_5 \mid x_2^n, x_4^n, x_5^n) \end{aligned} \quad (2.1.1.5)$$

З (1.4.5) визначимо ймовірності вибору дій для $t = 0$

$$\begin{aligned} P(\alpha^{n+1} \in B_1 \mid \xi^0 = x^0, \alpha^0 = a^0, \xi^1 = x^1, \dots, \xi^n = x^n) &= \\ = \prod_{i \in V} P(\alpha^{n+1} \in B_1 \mid \xi^0 = x^0, \alpha^0 = a^0, \xi^1 = x^1, \dots, \xi^n = x^n) &= \\ = \prod_{i \in V_1} \pi_i^n(B_{i=1} \mid x_1^n, x_2^n, x_3^n) \times \prod_{j \in V_2} \gamma_j^n(B_2 \mid x_1^n, x_2^n, x_5^n) &= \\ \pi_1^n(B_1 \mid x_1^n, x_2^n, x_3^n) \times \pi_2^n(B_2 \mid x_1^n, x_2^n, x_5^n) \times \\ \times \gamma_3^n(B_{j=3} \mid x_1^n, x_3^n, x_4^n) \times \gamma_4^n(B_4 \mid x_3^n, x_4^n, x_5^n) \times \gamma_5^n(B_5 \mid x_2^n, x_4^n, x_5^n) \end{aligned} \quad (2.1.1.6)$$

Знайдемо функцію виграшу для обох коаліцій. З (1.3.2) отримаємо вигляд функції виграшу для коаліції K_1 :

$$\begin{aligned} r_1(x, a) &= \sum_{\{1,2\}} \sum_{j \in N(\{1,2\})} r_{ij}(x_i, x_j, a_i, a_j) = \\ &= \sum_{j \in N(1)} r_{1j}(x_1, x_j, a_1, a_j) + \sum_{j \in N(2)} r_{2j}(x_2, x_j, a_1, a_j) = \\ &= r_{12}(x_1, x_2, a_1, a_2) + r_{13}(x_1, x_3, a_1, a_3) + r_{21}(x_2, x_1, a_2, a_1) + r_{25}(x_2, x_5, a_2, a_5) \end{aligned} \quad (2.1.1.7)$$

Аналогічно і для коаліції K_2 :

$$\begin{aligned}
r_1(x, a) &= \sum_{\{3,4,5\}} \sum_{i \in N(\{3,4,5\})} r_{ji}(x_j, x_i, a_j, a_i) = \\
&= \sum_{j \in N(3)} r_{3i}(x_3, x_i, a_3, a_i) + \sum_{j \in N(4)} r_{4i}(x_4, x_i, a_4, a_i) + \sum_{j \in N(5)} r_{5i}(x_5, x_i, a_5, a_i) = \\
&= r_{31}(x_3, x_1, a_3, a_1) + r_{34}(x_3, x_4, a_3, a_4) + \\
&+ r_{43}(x_4, x_3, a_4, a_3) + r_{45}(x_4, x_5, a_4, a_5) + \\
&+ r_{54}(x_5, x_4, a_5, a_4) + r_{52}(x_5, x_2, a_5, a_2)
\end{aligned}
\tag{2.1.1.8}$$

Так, як розглядаємо гру з нульовою сумою, то підставимо в рівняння (1.3.4):

$$\left(\begin{array}{l} r_{12}(x_1, x_2, a_1, a_2) + r_{13}(x_1, x_3, a_1, a_3) + \\ + r_{21}(x_2, x_1, a_2, a_1) + r_{25}(x_2, x_5, a_2, a_5) \end{array} \right) = - \left(\begin{array}{l} r_{31}(x_3, x_1, a_3, a_1) + r_{34}(x_3, x_4, a_3, a_4) + \\ + r_{43}(x_4, x_3, a_4, a_3) + r_{45}(x_4, x_5, a_4, a_5) + \\ + r_{54}(x_5, x_4, a_5, a_4) + r_{52}(x_5, x_2, a_5, a_2) \end{array} \right)
\tag{2.1.1.9}$$

І після скорочення ми отримаємо,

$$\left(r_{13}(x_1, x_3, a_1, a_3) + r_{25}(x_2, x_5, a_2, a_5) \right) = - \left(r_{31}(x_3, x_1, a_3, a_1) + r_{52}(x_5, x_2, a_5, a_2) \right)
\tag{2.1.1.10}$$

2.2 Алгоритм пошуку оптимальної стратегії

2.2.1 Теоретичний матеріал для пошуку оптимальної стратегії

Згідно з теоремою 3.3 [1, с.293] позначимо ціну гри як c , та визначимо функцію $v(x)$ на Банахову просторі таким чином, щоб наступна рівність мала місце:

$$c + v(x) = \sup_{\zeta} \inf_{\lambda} K_v(x, \zeta, \lambda), \quad \zeta \in P_{A_1}, \lambda \in P_{A_1},$$

де

$$K_u(x, \zeta, \lambda) = r(x, \zeta, \lambda) + \int u(y)Q(dy / x, \zeta, \lambda).
\tag{2.2.2.1}$$

Застосуємо теоретичний результат з роботи [1] і знайдемо оптимальні стратегії для гри.

Згідно з формулою доведення (18) теореми 3.3 [1, с. 296] $u(x)$ можна знайти, як

$$u(x) = \sup_{\zeta} \inf_{\lambda} K_u(x, \zeta, \lambda) - \int u(y)Q(dy / x, \zeta, \lambda), \quad \zeta \in P_{A_1}, \lambda \in P_{A_2},
\tag{2.2.2.2}$$

якщо

$$\mu(y) = Q(dy / x, \zeta, \lambda).$$

Тоді з (2.2.2.1) та (2.2.2.2) можемо отримати формулу для обчислення сідлової точки:

$$c = \int u^*(y) dQ(y) \quad (2.2.2.3)$$

2.2.2 Програмний застосунок

Для чисельного прикладу, описаного в підрозділі 2.1, спроектовано алгоритм і створено застосунок для пошуку оптимальної стратегії, сідлової точки та ціни гри.

Проект складається з чотирьох частин (рис 2.2.2.1):

- 1) Папка *gamers* – клас, що містить інформацію про гравців;
- 2) Папка *Q* – класи для обчислення та зберігання інформації повного ядра гри та для кожного стану кожного гравця;
- 3) Папка *U* містить класи, що описують функцію $u(x)$;
- 4) *functions.php* – скрипт, в якому зберігаються усі функції обчислення;
- 5) *run.php* – індексний файл, що запускає алгоритм пошуку оптимальної стратегії, сідлової точки та обчислення ціни гри.



рис 2.2.2.1 Загальний вигляд проекту

Директорія *gamers* складається з одного файлу *Gamer.php* (рис 2.2.2.2).



рис. 2.2.2.2 директорія gamers

Gamer.php описує гравця гри. Кожен об'єкт гравця зберігає чотири поля:

- 1) *\$gamerNumber* – порядковий номер гравця, тип *Integer*;
- 2) *\$x_states* – стани гравця згідно з (2.1.4), тип *Array*;
- 3) *\$actions* – можливі дії, згідно з (2.1.3), тип *Array*;
- 4) *\$connections* – зв'язки з іншими гравцями, або масив сусідів, тип *Array*.

Та функцію `__construct()`, що приймає значення цих параметрів та задає (рис. 2.2.2.3).

```
public function __construct($gamerNumber,$x_states,$actions,$connections){
    $this->gamerNumber = $gamerNumber;
    $this->x_states = $x_states;
    $this->actions = $actions;
    $this->connections = $connections;
}
```

рис. 2.2.2.3

Директорія *Q* складається з двох класів-об'єктів (рис. 2.2.2.4):

- 1) *Qfull.php* – ядро переходу;
- 2) *Qmini.php* – об'єкт, що описує ймовірність переходу у кожен стан кожного гравця



рис. 2.2.2.4

Кожен об'єкт *Qmini.php* містить 4 параметри:

- 1) *\$mainX* – стан, у який перейде гравець, тип *String*;
- 2) *\$x_states* – стани сусідів, які впливають на ймовірність переходу згідно з постановкою задачі, тип *Array*;
- 3) *\$action* – дія, яку обирає гравець, тип *String*;
- 4) *\$probability* – ймовірність переходу, тип *float*;

Та функцію `__construct()`, що приймає значення цих параметрів та створює об'єкт *Qmini* (рис 2.2.2.5).

```
public function __construct($mainX,$x_states,$action, $probability){
    $this->mainX = $mainX;
    $this->x_states = $x_states;
    $this->action = $action;
    $this->probability = $probability;
}
```

рис. 2.2.2.5

Директорія *U* складається з трьох класів-об'єктів (2.2.2.6):

- 1) *Umini.php* – функція $u(x)$ для заданого стану системи
- 2) *Utime.php* – об'єкт усіх функцій $u(x)$ для заданого часу t ;
- 3) *Ufull.php* – об'єкт усіх $u(x)$ гри, складається з об'єктів *Utime*.



рис. 2.2.2.6

Об'єкт *Umini.php* містить параметри:

- 1) $\$x_states$ – задані стани системи для функції $u(x)$, тип *Array*;
- 2) $\$actions$ – дії, тип *Array*;
- 3) $\$time$ – час, тип *Integer*;
- 4) $\$value$ – отримане значення $u(x)$

Та функцію `__construct()`, що приймає значення цих параметрів та створює об'єкт *Umini* (рис. 2.2.2.7).

```
public function __construct($x_states,$actions,$time,$value){
    $this->x_states = $x_states;
    $this->actions = $actions;
    $this->time = $time;
    $this->value = $value;
}
```

рис. 2.2.2.7

Об'єкт *Utime.php*, параметри:

- 1) $\$U_minis$ – набір $u(x)$ ($Uminis$ об'єктів) для кожного стану системи, тип *Array*;
- 2) $\$time$ – час t , тип *Integer*.

Та методи задання параметрів $setUminis()$, $setTime()$, (рис 2.2.2.8).

```
public function setUminis($U_minis){
    $this->U_minis = $U_minis;
}
public function setTime($time){
    $this->time = $time;
}
```

рис. 2.2.2.8

Об'єкт *Ufull.php*, параметри:

- 1) $\$U_times$ – набір $u(x)$ ($Utime$ об'єктів), тип *Array*;

Методи (рис. 2.2.2.9):

- 1) $setUTimes()$ – задає масив об'єктів *Utime*;
- 2) $addUtime()$ – додає в масив $\$U_times$ об'єкт *Utime*.

```
public function setUTimes($U_times){
    $this->U_times = $U_times;
}
public function addUtime($Utime){
    array_push( &array: $this->U_times, $Utime);
}
```

рис. 2.2.2.9

В файлі *functions.php* наведені усі необхідні методи:

- 1) для обчислення ймовірностей кожного стану гравця;
- 2) для обчислення ймовірностей кожного стану системи;
- 3) обчислення загального числа станів системи;
- 4) функцій (2.2.2.1), (2.2.2.2) та (2.2.2.3);
- 5) перевірка на існування сідлової точки
- 6) отримання ціни гри.

2.2.3 Алгоритм пошуку оптимальної стратегії

Файл *run.php* – запускає алгоритм пошуку оптимальної стратегії, сідлової точки та ціни гри.

На початку відбувається завантаження усіх необхідних компонентів для роботи (2.2.3.1).

```
<?php
// import all elements
include("./gamers/Gamer.php");
include "./Q/Qmini.php";
include "./Q/Qfull.php";
include "./U/Umini.php";
include "./U/Utime.php";
include "./U/Ufull.php";
include "functions.php";
ini_set( option: 'precision', value: 100);
set_time_limit( seconds: 0);
```

рис. 2.2.3.1

Далі задається необхідну кількість гравців у грі, кількість можливих станів для гравців, кількість можливих дій гравців. Визначається для кожного гравця

```
// set gamers
$Number_of_Gamers = 5;
$Number_of_States_for_each_gamer = 3;
$Number_of_Action_for_each_gamer = 2;
//set Graph connections
$all_connections = array(
    1 => array(2,3),
    2 => array(1,5),
    3 => array(1,4),
    4 => array(3,5),
    5 => array(2,4)
);
```

сусідів (рис. 2.2.3.2).

рис. 2.2.3.2

Створюється масив гравців, об'єктів Gamer (рис. 2.2.3.4).

```
//create gamers
$gamers = createGamersArray($Number_of_Gamers,$Number_of_States_for_each_gamer,$Number_of_Action_for_each_gamer, $all_connections);
```

рис. 2.2.3.4

Задаються ймовірності переходів для кожного стану гравців (рис. 2.2.3.5).

```
//calc Qminis for each gamer-states-actions
$Qminis = setQminis($gamers,$Number_of_States_for_each_gamer,$Number_of_Action_for_each_gamer);
```

рис. 2.2.3.5

Обчислюються усі можливі стани та дії гравців, і ядро гри (2.2.3.6).

```
$Q_states = getQ_states($gamers);
$Q_actions = getQ_actions($gamers);

//set Qfull
$QFulls = setQFull($Qmins,$Q_states,$Q_actions);
```

рис. 2.2.3.6

Обчислюються $u(x)$ для кожного етапу гри. На кожному етапі

```
for ($TIME = 1; $TIME <= $N; $TIME++) {
    // calc U mini for each state
    $UMinis = getUminis($Q_states,$UFULL,$QFulls,$reward,$TIME);

    // set Utime from Uminis
    $UTime = new Utime();
    $UTime->setTime($TIME);
    $UTime->setUMinis($UMinis);

    // set Ufull from Utimes
    $UFULL->addUtime($UTime);

    /* compare current Utime and previous UTime
    IF current Utime = previous Utime
    THEN stop, and calc game price C
    ELSE continue
    */

    if($TIME > 2){
        $GameValueExists = checkGameValueExists($UFULL->U_times[$TIME], $UFULL->U_times[$TIME - 1]);
        if($GameValueExists){
            break;
        }
    }
}
```

перевіряються умови існування сідлової точки (рис. 2.2.3.7).

рис. 2.2.3.7

Якщо сідлова точка існує, обраховується ціна гри, в іншому випадку повідомляється про відсутність (рис. 2.2.3.8).

```
$gameValue = 0;
if($GameValueExists){
    $gameValue = getGameValue($UFULL, $TIME);
}else{
    echo "Game value status: Not Exists for current time " . $TIME;
}

$time = microtime( as_float: true) - $start;
echo "Total time:" . $time;
```

рис. 2.2.3.8

2.3 Висновок

В розділі розглянуто чисельний приклад гри за допомогою графу, з визначеними станами та діями для кожного гравця. Спроектвано алгоритм пошуку оптимальної стратегії, перевірки існування сідлової точки та обчислення ціни гри. Створено програмний застосунок.

3 Аналіз функції виграшу та задачі максимізації

3.1 Вступ

У даному розділі розглянуто умови існування сідлової точки. Досліджена однокрокова гра, середній виграш з дисконтованим критерієм, за яких умов гравцям краще обирати виграш на першому етапі, замість продовження. Додатково була проаналізована умова максимізації середнього виграшу, коли гравці коаліції зацікавлені лише в отриманні максимальної винагороди.

3.2 Дослідження функції середнього виграшу для скінченої гри

Розглянемо варіант існування сідлової точки для однокрокової гри. Згідно з формулою (1.5.1), при $t = 1$

$$\phi(x, \pi, \gamma) = \frac{1}{t+1} E_{\pi, \gamma}^x \sum_{k=0}^t r(\xi^k, \alpha^k) = \frac{1}{2} E_{\pi, \gamma}^x (r(\xi^0, \alpha^0) + r(\xi^1, \alpha^1)).$$

$$(3.2.1)$$

Для кожного наступного t визначимо критерій дисконтування β . Гравців цікавить виграш на t етапі гри, а не у майбутніх етапах гри. Так, як кожна коаліція K_1 та K_2 не може накопичувати ці виграші до нескінченості, то функція середньої винагороди буд обмеженою.

Тоді формула (1.5.2) прийме вигляд:

$$\mathcal{G}(x, \pi, \gamma) = \frac{1}{t+1} E_{\pi, \gamma}^x \sum_{k=0}^t r(\xi^k, \alpha^k) \beta^k = \frac{1}{2} E_{\pi, \gamma}^x (r(\xi^0, \alpha^0) + r(\xi^1, \alpha^1) \beta)$$

$$(3.2.2)$$

Нехай виграш у початковому стані буде рівний 0, тоді з (4.2.2) отримаємо наступний вид:

$$\mathcal{G}(x, \pi, \gamma) = \frac{1}{2} E_{\pi, \gamma}^x (r(\xi^1, \alpha^1) \beta). \quad (3.2.3)$$

Розглянемо інший варіант події, коли гра триває у два етапи. Тоді використовуючи формулу (4.2.1) отримаємо середню винагороду:

$$\phi(x, \pi, \gamma) = \frac{1}{t+1} E_{\pi, \gamma}^x \sum_{k=0}^t r(\xi^k, \alpha^k) = \frac{1}{3} E_{\pi, \gamma}^x (r(\xi^0, \alpha^0) + r(\xi^1, \alpha^1) + r(\xi^2, \alpha^2))$$

$$(3.2.4)$$

Використовуючи критерій дисконтування отримаємо:

$$\mathcal{G}(x, \pi, \gamma) = \frac{1}{t+1} E_{\pi, \gamma}^x \sum_{k=0}^t r(\xi^k, \alpha^k) \beta^k = \frac{1}{3} E_{\pi, \gamma}^x \left(r(\xi^0, \alpha^0) + r(\xi^1, \alpha^1) \beta + r(\xi^2, \alpha^2) \beta^2 \right). \quad (3.2.5)$$

Перевіримо за яких умов, гравцям вигідніше грати два етапи, замість одного. Задача зводиться до пошуку цілі максимізації середнього виграшу. З формул (4.2.3) (4.2.5) отримаємо нерівність:

$$\begin{aligned} \frac{1}{2} \left(r(\xi^1, \alpha^1) \beta \right) &\leq \frac{1}{3} \left(r(\xi^1, \alpha^1) \beta + r(\xi^2, \alpha^2) \beta^2 \right) \\ \frac{1}{2} \left(r(\xi^1, \alpha^1) \beta \right) &\leq \frac{1}{3} \beta \left(r(\xi^1, \alpha^1) + r(\xi^2, \alpha^2) \beta \right) \\ 3 \times r(\xi^1, \alpha^1) &\leq 2 \times \left(r(\xi^1, \alpha^1) + r(\xi^2, \alpha^2) \beta \right) \quad , \\ r(\xi^1, \alpha^1) &\leq 2 \times r(\xi^2, \alpha^2) \beta \\ r(\xi^1, \alpha^1) (2 \times \beta)^{-1} &\leq r(\xi^2, \alpha^2) \end{aligned} \quad (3.2.6)$$

Отже, якщо виконані умови (4.2.6), то гравцям коаліції слід грати в два етапи.

Перевіримо, коли гра буде тривати в n етапів. Як наслідок отримаємо твердження.

Твердження 3.2.1: Якщо наступна нерівність

$$\left((k-2) \times r(\xi^1, \alpha^1) - k \sum_{l=2}^{t-1} r(\xi^l, \alpha^l) \beta^l \right) \times \beta^{-t} \leq r(\xi^t, \alpha^t)$$

має місце, то гравцям коаліції рекомендується продовжувати до n -ого етапу й отримати максимальний підсумковий виграш.

3.3 Задача максимізації виграшу в умовах неповної інформації

Додамо умову для гри.

Нехай коаліція має набір власних дій, але жодному з гравців коаліції не будуть відомі можливі дії гравців іншої коаліції. Гра розігрується в $\forall t \in \mathbb{N}$ кількість етапів. Після кожного раунду, гравці коаліції можуть або отримати виграш, що обраховується за допомогою формули (1.5.3), або продовжити грати далі.

Задача коаліції – знайти правило зупинки, щоб отримати максимальний виграш серед усіх t етапів гри з нульовою сумою.

Нехай існує послідовність вигравів на кожному етапі гри:

$$r_i = \{r(\xi^0, \alpha^0)\beta^0, r(\xi^1, \alpha^1)\beta^1, r(\xi^2, \alpha^2)\beta^2, \dots, r(\xi^n, \alpha^n)\beta^n\}, i = 0, \dots, n$$

Позначимо $F(x)$ – функцію корисності від отримання виграву коаліції на i -тому кроці

$$F(x(\xi, \alpha)) = r_i$$

Нехай гравці коаліції можуть впорядкувати за зростанням середню винагорода. Таким чином, нехай кожна середня винагорода після етапу гри має ранг.

Отже, задача коаліції – пошук правила зупинки, що максимізує (1.5.3).

Позначимо $F^*(a, k)$ – середню корисність для коаліції від оптимального продовження, коли буде проведено k етапів гри, і в загальному рейтингу середній отриманий виграв упорядковується на a -тому місці серед вже переглянутих середніх вигравів. Позначимо за $F_0(a, k)$ – середню корисність від завершення гри на k -тому етапі гри і отримання середнього виграву.

Тоді, якщо система знаходиться на останньому заданому n етапі, то має місце рівність:

$$F^*(a, n) = F_0(a, n) = F(a), a = \{1, \dots, n\} \quad (3.3.1)$$

Так, як середня асимптотична винагорода отримана на $k+1$ етапі гри має однакову ймовірність зайняти відповідний ранг серед n середніх вигравів, то ранг з рівною ймовірністю прийме значення від 1 до $k+1$.

Нехай система перейшла на етап $k+1$. Для коаліції загальний виграв позначимо за b , тоді середня корисність від проведення ще одного етапу буде:

$$\frac{1}{k+1} \sum_{b=1}^{k+1} F^*(b, k+1) \quad (3.3.2)$$

Згідно з твердженням 5 [3, с.329] та з означення F^* і F_0 отримаємо:

$$F^*(a, k) = \max \left\{ F_0(a, k), \frac{1}{k+1} \sum_{b=1}^{k+1} F^*(b, k+1) \right\} \quad (3.3.3)$$

Отже, згідно з оптимальною процедурою, гравцям коаліції вигідніше продовжити гру, якщо

$$F^*(a, k) > F_0(a, k),$$

і закінчити за умовою:

$$F^*(a, k) = F_0(a, k).$$

Позначимо за $P(a, k)$ ймовірність вибору асимптотичної середньої винагороди для будь-якого етапу гри.

Додамо додаткову умову для функції корисності. Нехай гравці коаліції зацікавлені лише в отриманні максимального виграшу, а будь-який інший не є привабливим. Відповідно отримаємо рівність:

$$\begin{cases} F(1) = 1 \\ F(b) = 0, b = \{2, \dots, n\} \end{cases} \quad (3.3.4)$$

Ціль гравців коаліції оновиться до максимізації ймовірності отримання середньої винагороди на етапі k , серед усіх етапів n .

З умови отримаємо:

$$\begin{cases} F_0(1, k) = \frac{k}{n} \\ F_0(a, k) = 0, k = \{2, \dots, n\} \end{cases} \quad (3.3.5)$$

Тоді оптимальну процедуру опишемо наступним чином. Нехай k^* – найменше число, для якого виконується рівність

$$F^*(1, k) = \frac{k}{n}. \quad (3.3.6)$$

Тоді, якщо загальний середній виграш на k^* етапі гри має ранг 1, то процес гри закінчується і гравці коаліції отримують виграш.

Якщо ранг середньої винагороди на k^* етапі рівний 1, то ймовірність, з якою цей виграш буде мати ранг 1 серед усіх n можливих виграшів буде:

$$P(1, k) = \frac{k^*}{n}. \quad (3.3.7)$$

З формули 13 [3, с.332] отримаємо наступне твердження. Згідно з оптимальною процедурою гравцям коаліції рекомендується грати приблизно до

$$\frac{n}{e}$$

етапу, а далі зупинитись та отримати середній виграш, за умови, якщо його ранг буде дорівнювати 1 серед вже переграних етапів. Відповідна ймовірність гравців коаліції отримати максимальний середній виграш серед n етапів гри буде:

$$\frac{1}{e}.$$

3.4 Висновок

У даному розділі була досліджена однокрокова гра, середній виграш з дисконтованим критерієм, сформовані умови продовження гри задля отримання більшого виграшу. Отримане наближене значення етапу, на якому гравці коаліції заробляють максимальну винагороду, якщо стратегії не відомі один одному.

Висновок

В роботі було розглянуто стохастичну коаліційну гру на графі зі скінченою кількістю гравців у кожній коаліції. На теоретичному матеріалі побудовано алгоритм пошуку оптимальної стратегії, наведено чисельний приклад і створено програмний застосунок, який дозволяє знайти сідлову точку та обчислити ціну гри.

Література

1. Chorney R.K., Hans Daduna, Knopov P. S., Stochastic games for distributed players on graphs // *Mathematical methods of Operations Research*, (2004), (60): P. 279–298, DOI 10.1007/s001860400374;
2. Moris H. DeGroot, Optimal statistical decisions, *McGraw-Hill Series in probability and statistics*, 1970, C. 328-333;
3. Michael Kearns, Michael L. Littman, Satinder Singh, Graphical Models for Game Theory, 2005;
4. Game Theory First // *Stanford Encyclopedia of Philosophy*, 1997, <https://plato.stanford.edu/entries/game-theory/>
5. Friesner D., Axelsen D., Using Game Theory to teach principles of Microeconomics // *Journal For Economic Educators*, V. 6, N. 1, 2006
6. Dixit A. K., Nalebuff B. J., The Art of Strategy A Game Theorist's Guide to Success in Business and Life // *Library of Congress Cataloging-in-Publication Data*, 2008, ISBN: 978-0-393-33717-4, C. 1-512;
7. Katz A., The Strategic Structure of Offer and Acceptance: Game Theory and the Law of Contract Formation, *University of Michigan Law*, 1990, C.3-15
8. Rosenthal Ed. C., The Complete Idiot's Guide to Game Theory // *Penguin Random House LLC*, 2011, ISBN: 9781615640553
9. Luce R. D., Raiffa H., Games and Decisions: Introduction and Critical Survey // *Dover Publications inc. New York*, 2012, ISBN: 0486134830;
10. Zhang W., Game Theory and Society // *China Perspectives*, 2017, ISBN: 9781351330985;