

Міністерство освіти і науки України
Національний університет «Києво-Могилянська академія»
Факультет інформатики
Кафедра математики

Магістерська робота

освітній ступінь – магістр

на тему: **«ДОСЛІДЖЕННЯ СТОХАСТИЧНОЇ
ПОВЕДІНКИ КЛІТИННИХ АВТОМАТІВ»**

Виконав: студент 2-го року
навчання
освітньо-наукової програми
«Системний аналіз»,
спеціальності 124 Системний аналіз

Глушенков Сергій Михайлович

Керівник: Чорней Р.К.,
кандидат фіз.-мат. наук, доцент

Рецензент _____
(прізвище та ініціали)

Кваліфікаційна робота захищена
з оцінкою _____

Секретар ЕК _____

« ____ » _____ 2022 р.

Міністерство освіти і науки України
Національний університет «Києво-Могилянська академія»
Факультет інформатики
Кафедра математики

ЗАТВЕРДЖУЮ

Зав. кафедри математики,
проф., док. фіз.-мат. наук,

_____ Б.В. Олійник
(підпис)

« ____ » _____ 2022 р.

ІНДИВІДУАЛЬНЕ ЗАВДАННЯ
на магістерську роботу

студенту 2 року навчання МП «Системний аналіз»

Глушенкову Сергію Михайловичу

Тема: **Дослідження стохастичної поведінки клітинних автоматів**

Вихідні дані:

- Доведення існування оптимальних стратегій
- Програмний застосунок для демонстраційних цілей

Зміст ТЧ до дипломної роботи:

Реферат

Вступ

Розділ 1: Взаємодія стохастичних клітинних автоматів та оптимальні стратегії керування

Розділ 2: Задача поширення лісових пожеж

Висновок

Список використаних джерел

Дата видачі « ____ » _____ 2022 р.

Керівник: канд. фіз.-мат. наук, доцент,
Чорней Р.К.

_____ (підпис)

Завдання отримав

Глушенков С.М.

_____ (підпис)

Тема: Дослідження стохастичної поведінки клітинних автоматів

Календарний план виконання роботи:

№ п/п	Назва етапу дипломного проекту (роботи)	Термін виконання етапу	Примітка
1.	Вибір теми дипломної роботи.	20.09.2021	
2.	Огляд необхідної літератури за темою роботи.	01.10.2021	
3.	Визначення основних завдань дипломної роботи.	05.10.2021	
4.	Доведення існування оптималь- ної стратегії і написання теоре- тичної частини.	13.02.2022	
5.	Аналіз алгоритму покращення стратегії.	06.05.2022	
6.	Пошук прикладних задач за- стосування стохастичних клі- тинних автоматів.	16.05.2022	
7.	Розробка застосунку для роз- глянутої прикладної задачі.	10.06.2022	
8.	Аналіз отриманих результатів та попередній захист роботи.	17.06.2022	
9.	Корегування роботи за резуль- татами попереднього захисту.	29.06.2022	
10.	Захист магістерської роботи.	08.07.2022	

Студент **Глушков С.М.** _____
(підпис)

Керівник: канд. фіз.-мат. наук, доцент,
Чорней Р.К. _____
(підпис)

«___» _____ 2022 р.

Зміст

Анотація	5
Вступ	6
1 Взаємодія стохастичних клітинних автоматів та оптимальні стратегії керування	8
1.1 Опис стохастичних клітинних автоматів	8
1.2 Структура керування і стратегії	9
1.3 Існування оптимальної стратегії керування	13
1.4 Знаходження оптимальної стратегії	20
2 Задача поширення лісових пожеж	23
2.1 Постановка задачі	23
2.2 Програмна реалізація	26
2.3 Порівняння стратегій	30
Висновок	34
Список використаних джерел	35

Анотація

Обсяг роботи: 35 сторінок, 8 рисунків, 5 джерел посилань.

СТОХАСТИЧНІ КЛІТИННІ АВТОМАТИ, ОПТИМАЛЬНЕ КЕРУВАННЯ, ОПТИМАЛЬНІ СТРАТЕГІЇ, ЛІСОВІ ПОЖЕЖІ

Об'єктом роботи є стохастичні локально взаємодіючі клітинні автомати. Предметом роботи є вплив оптимального керування на стохастичну поведінку клітинних автоматів.

Метою роботи є доведення існування оптимальної стратегії керування стохастичними клітинними автоматами, та перевірка оптимальності розглянутих стратегій на прикладній задачі.

Методи розробки: Побудова моделей, заснованих на марковських процесах, на основі синхронних локально взаємодіючих стохастичних клітинних автоматів, знаходження оптимального управління методами динамічного програмування, програмування на мові C++ у середовищі Xcode (Version 13.1), використання графічної бібліотеки SFML для візуального відтворення процесів.

Результат роботи: Наведений формальний опис локально взаємодіючих стохастичних клітинних автоматів, показано існування оптимальної стратегії керування в множині детермінованих марковських стаціонарних стратегій, розглянута прикладна задача поширення лісових пожеж з доданим керуванням. Розроблений програмний застосунок, який дозволяє візуально симулювати процес лісової пожежі та керувати процесом гасіння. За допомогою застосунку порівняно різні типи стратегій, їх вплив на стохастичну поведінку клітинних автоматів, та порівняна ефективність мінімізації функції витрат.

Вступ

Оцінка сучасного стану об'єкта розробки. Клітинні автомати дозволяють моделювати широкий спектр складних систем з локальною взаємодією. Не дивлячись на те, що в загальному поведінка окремо взятих клітинних автоматів може бути дуже простою, вдала їх комбінація або задання нестандартних правил взаємодії може значно ускладнити поведінку системи в загальному і може призводити до доволі неоднозначних та непередбачуваних результатів. Одним із прикладів застосування стохастичних клітинних автоматів є виникнення і поширення лісових пожеж, і для цього прикладу клітинні автомати є досить потужним інструментом. Достатньо велика кількість наукових досліджень в даній галузі використовує саме стохастичні клітинні автомати у якості моделі для симуляції. Однак, майже всі ці роботи присвячені лише передбаченню поширення та динаміці розвитку пожежі, в той час як проблемі гасіння такої пожежі приділяється несправдливо мало уваги.

Актуальність роботи та підстави її виконання. Оскільки щороку абсолютно в різних кутках нашої планети людство постійно вимушено мати справу з гасінням лісових пожеж, питання оптимальності гасіння та мінімізації збитків є досить нагальним. З одного боку, може здатися, що оптимальна стратегія є досить очевидною – просто якомога швидше гасити усе, що палає. Однак, враховуючи можливі незручності та фактори, які перешкоджатимуть гасінню, наприклад, відсутність природних водойм для швидкого поповнення резервуарів пожежної техніки, або важкодоступна місцевість, яка змушує використовувати ресурси більш заощадливо, не кажучи вже про велику кількість параметрів які роблять поширення пожежі випадковим і непередбачуваним. Тому, в таких випадках і постає питання прорахунку оптимальної стратегії. Використання саме стохастичних клітинних автоматів дозволяє краще наблизити модель до реалістичної поведінки, та допомагає визначити оптимальне рішення, яке буде більш стійким до непередбачуваних та випадкових подій, що в результаті може демонструвати суттєве зменшення витрат та шкоди, яку приносять лісові пожежі на великих часових проміжках.

Об'єкт та предмет дослідження. Об'єктом даної роботи є стохастичні локально взаємодіючі клітинні автомати. Предметом роботи є вплив оптимального керування на стохастичну поведінку клітинних автоматів.

Мета й завдання роботи. Метою даної роботи є доведення існування оптимальної стратегії керування стохастичними клітинними автоматами, та перевірка оптимальності розглянутих стратегій на прикладній задачі. Для цього потрібно виконати основні чотири задачі:

- Дати формальний опис синхронних локально взаємодіючих стохастичних клітинних автоматів.
- Для критеріїв очікуваних витрат в одиницю часу і дисконтованого критерію показати існування оптимальних стратегій.
- Розглянути застосування стохастичних клітинних автоматів на прикладній задачі поширення лісових пожеж.
- Розробити програмний застосунок для знаходження оптимальної стратегії керування в розглянутій задачі.

1 Взаємодія стохастичних клітинних автоматів та оптимальні стратегії керування

1.1 Опис стохастичних клітинних автоматів

Маємо неорієнтований граф локально взаємодіючих стохастичних клітинних автоматів $G(V, N)$:

V – множина вершин (клітинних автоматів) графа, $V \in \{1, 2, \dots, S\}$;

$N(i)$ – множина суміжних автоматів для автомата i :

$N(i) = \{j \in \{1, 2, \dots, S\} / \{i\} \mid d(i, j) = 1, i \in \{1, 2, \dots, S\}\}$, де $d(i, j)$ – відстань між вершинами графа G .

В кожен момент часу $t \in \mathbb{N}$:

- кожен автомат $i \in V$ знаходиться в одному з можливих станів $\xi_i^t = x_i \in X_i, |X_i| < \infty$, X_i – множина можливих станів автомата i ;
- на вхід кожного автомата $i \in V$ подається сигнал α_i^t з множини входів цього автомата: $\alpha_i^t = y_i \in I_i, |I_i| < \infty$. При цьому сигнали α_i^t не залежать від сигналів інших автоматів та від історії системи.

На перехід автомата i з одного стану в інший впливають лише вхідні сигнали α_i^t , а також стани “сусідів” автомата з множини $N(i)$, що забезпечує локальну взаємодію між автоматами [1, с. 89].

В даній роботі будуть розглядатися системи синхронних клітинних автоматів, тобто перехід в наступний момент часу для всіх клітин відбувається одночасно.

При переході системи в наступний момент часу $t + 1$, кожен автомат $i, i \in V$ подає на вихід сигнал γ_i^{t+1} з множини виходів цього автомата: $\gamma_i^{t+1} = o_i \in O_i, |O_i| < \infty$.

Таким чином, якщо в момент часу t автомат i знаходиться в стані $\xi_i^t = x_i \in X_i$, з околom його сусідів $N(i)$, і на вхід цього автомата подається сигнал $\alpha_i^t = y_i \in I_i$, то при переході в наступний момент часу $t + 1$ автомат i перейде в наступний стан $\xi_i^{t+1} = \tilde{x}_i \in X_i$, і на вихід автомат надішле сигнал

$\gamma_i^{t+1} = o_i \in O_i$ з ймовірністю $[1, \text{с. } 90]$:

$$Q(x_i, x_j : j \in N(i), y_i; \tilde{x}_i, o_i) = \Pr(\xi_i^{t+1} = \tilde{x}_i, \gamma_i^{t+1} = o_i | \xi_i^t = x_i, \xi_j^t = x_j, j \in N(i), \alpha_i^t = y_i), \quad (1)$$

$$0 \leq Q(x_i, x_j : j \in N(i), y_i; \tilde{x}_i, o_i) \leq 1$$

Кожному вихідному сигналу $\gamma_i^t = o_i$ автомата i ставиться у відповідність певне числове значення за допомогою функції перетворення: $T^o : O_i \rightarrow \mathbb{R}, \forall i \in V$. Аналогічно і з вхідним сигналом $\alpha_i^t = y_i$ і відповідною функцією $T^i : I_i \rightarrow \mathbb{R}, \forall i \in V$. Дане перетворення виконується для використання вхідних та вихідних сигналів в числовій функції оцінки керування системою клітинних автоматів. Тобто, маємо числову функцію оцінювання $f : \mathbb{R} \rightarrow \mathbb{R}$, яка залежить від вихідного сигналу автомата при переході в момент часу $t + 1$, а також функція $g : \mathbb{R} \rightarrow \mathbb{R}$, що залежить від вхідного сигналу в момент часу t , якщо автомат знаходився в стані ξ_i^t . Сукупну функцію витрат при переході клітинного автомата в момент часу $t+1$ позначимо через $r : \mathbb{R} \rightarrow \mathbb{R}$. Таким чином, функція витрат за момент часу t рахується наступним чином:

$$F^t = \sum_{i=1}^S g(\xi_i^t, \alpha_i^t) + f(\gamma_i^{t+1}) = \sum_{i=1}^S r(\xi_i^t, \alpha_i^t, \gamma_i^{t+1}). \quad (2)$$

Основною задачею оптимального керування системою стохастичних клітинних автоматів є оптимізація подання вхідних сигналів α_i^t таким чином, щоб глобальна функція витрат F прямувала до бажаного значення (в залежності від задачі, це може бути максимізація, мінімізація, або певне конкретне значення): $\sum_{t=0}^{n-1} F^t = F \rightarrow \tilde{F}, n \in \mathbb{N}$.

1.2 Структура керування і стратегії

Визначимо поняття стратегії керування. Стратегією будемо називати послідовність вхідних сигналів, які подаються на вхід клітинного автомату/системи клітинних автоматів в кожен момент часу $t, t \in \mathbb{N}$.

В загальному випадку, керування вхідними сигналами клітинного автомату здійснюється за допомогою застосування певних політик вибору рішення, які базуються на історії станів автомату і його околу, та попередньо прийнятих рішень. Тобто, якщо в момент часу t для автомату i відома історія станів $h_i^t = (x_i^0, x_i^1, \dots, x_i^{t-1}, x_i^t)$, $\forall x_i^t \in X_i$ разом з історією станів його околу $h_{N(i)}^t = \{(x_j^0 : j \in N(i)), (x_j^1 : j \in N(i)), \dots, (x_j^t : j \in N(i))\}$, $\forall \xi_j^t \in X_j$, і до цього були прийняті рішення $\omega_i^{t-1} = (\alpha_i^0, \alpha_i^1, \dots, \alpha_i^{t-1})$, $\forall \alpha_i^k \in I_i$, $k = 0, \dots, t-1$, то рішення про надсилання вхідного сигналу $\alpha_i^t = y_i \in I_i$ приймається з ймовірністю:

$$\Pr(\alpha_i^t = y_i \mid \{\xi_i^0 = x_i^0, \xi_j^0 = x_j^0 : j \in N(i), \alpha_i^0 = y_i^0\}, \{\xi_i^1 = x_i^1, \xi_j^1 = x_j^1 : j \in N(i), \alpha_i^1 = y_i^1\}, \dots, \{\xi_i^{t-1} = x_i^{t-1}, \xi_j^{t-1} = x_j^{t-1} : j \in N(i), \alpha_i^{t-1} = y_i^{t-1}\}, \xi_i^t = x_i^t, \xi_j^t = x_j^t : j \in N(i)) = \Pr(\alpha_i^t = y_i \mid h_i^t, h_{N(i)}^t, \omega_i^{t-1}), \quad (3)$$

і при цьому $\sum_{y_i \in I_i} \Pr(\alpha_i^t = y_i \mid h_i^t, h_{N(i)}^t, \omega_i^{t-1}) = 1$.

Дані стратегії прийняття рішень називають рандомізованими допустимими стратегіями [1, с. 76].

Стратегія буде мати марковську властивість, якщо ймовірність вибору того чи іншого вхідного сигналу для автомату не залежить від минулих станів, околів, та прийнятих рішень, а лише від поточного стану з його околів:

$$\Pr(\alpha_i^t = y_i \mid h_i^t, h_{N(i)}^t, \omega_i^{t-1}) = \Pr(\alpha_i^t = y_i \mid \xi_i^t = x_i, \xi_j^t = x_j : j \in N(i)). \quad (4)$$

Якщо ж ймовірність вибору стратегії для одного й того ж стану клітинного автомату та його околу не залежить від моменту часу t , тобто:

$$\Pr(\cdot \mid \xi_i^{t'} = x_i, \xi_j^{t'} = x_j : j \in N(i)) = \Pr(\cdot \mid \xi_i^{t''} = x_i, \xi_j^{t''} = x_j : j \in N(i)), \quad (5)$$

і при цьому $t' \neq t''$, то такі стратегії називаються стаціонарними (марковськими) допустимими стратегіями.

Допустимі стаціонарні (марковські) стратегії називаються детермінованими, якщо для всіх станів автомату та його околу, вхідний сигнал виби-

рається однозначно, тобто: $\Pr(\cdot \mid \xi_i^t = x_i, \xi_j^t = x_j : j \in N(i)), \forall x_i \in X_i$ – одноточкова міра [1, с. 77].

Окрім розглянутих типів стратегій, потрібно розглянути ще два різновиди, які залежать від масштабності керованого об'єкту: а саме локальні та глобальні стратегії (надалі під стратегіями будуть вважатись саме допустимі стаціонарні марковські детерміновані стратегії).

Локальною будемо називати стратегію, яка застосовується лише для одного конкретного клітинного автомату. Нехай клітинний автомат i може приймати $|X_i|$ різних станів, і кожен автомат з його околу $N(i)$ може приймати $|X_j|$ станів, $\forall j \in N(i)$. Тоді загальна кількість станів автомату i з його окомом $N(i)$ буде дорівнювати $c_i = |X_i| \prod_{j \in N(i)} |X_j|$. Позначимо ці стани $\psi_{i,k} = \{\xi_i = x_i, \xi_j = x_j : \forall j \in N(i)\}$, $k = 1, \dots, c_i$. Тоді локальною стратегією керування для цього автомата буде:

$$\omega_i = \{(y_i^1 \mid \psi_i = \psi_{i,1}), (y_i^2 \mid \psi_i = \psi_{i,2}), \dots, (y_i^{c_i} \mid \psi_i = \psi_{i,c_i})\},$$

$$y_i^j \in I_i, j = 1, \dots, c_i. \quad (6)$$

Множину усіх можливих локальних стратегій автомата i позначимо Ω_i і маємо: $\omega_i \in \Omega_i$, $|\Omega_i| = \prod_{j=1}^{c_i} |I_i|$.

Глобальна стратегія передбачає керування усією системою клітинних автоматів. Нехай маємо систему із S клітинних автоматів, кожен із яких може знаходитись в $|X_i|$ можливих станів, та має множину вхідних сигналів I_i , $i = 1, \dots, S$. Тоді кількість усіх можливих станів такої системи дорівнює: $c = \prod_{i=1}^S |X_i|$. Стани системи позначимо як $\psi_k = \{\xi_1 = x_1, \xi_2 = x_2, \dots, \xi_S = x_S\}$, $k = 1, \dots, c$; $\forall i \in V : x_i \in X_i$. Тоді глобальною стратегією керування для цієї системи буде:

$$\omega = \{(\omega_1 \mid \psi = \psi_1), (\omega_2 \mid \psi = \psi_2), \dots, (\omega_c \mid \psi = \psi_c)\},$$

$$\omega_k = (y_1, y_2, \dots, y_S)^T, \forall k = 1, \dots, c, y_i \in I_i, \forall i \in V. \quad (7)$$

Множину усіх можливих глобальних стратегій системи позначимо Ω , і тоді

$$\omega \in \Omega, |\Omega| = \left(\prod_{i=1}^S |I_i|\right)^c.$$

Також позначимо через γ^{t+1} вектор, що містить усі вихідні сигнали системи при переході системи в момент часу $t+1$: $\gamma^{t+1} = (\gamma_1^{t+1}, \gamma_2^{t+1}, \dots, \gamma_S^{t+1})$, $\gamma_i^{t+1} = o_i \in O_i$.

Тепер можемо переписати глобальну числову функцію витрат з урахуванням введених позначень стратегії наступним чином – якщо особа що приймає рішення обирає глобальну стратегію ω , то значення чисельної функції оцінки керування F_ω буде дорівнювати:

$$F_\omega = \sum_{t=0}^{n-1} F_\omega^t = \sum_{t=0}^{n-1} r(\psi^t, \omega, \gamma^{t+1}) = \sum_{t=0}^{n-1} \sum_{i=1}^S r(\xi_i^t, \alpha_i^t, \gamma_i^{t+1}), \quad (8)$$

де $\alpha_i^t \in \omega$, $\forall i \in V$, F_ω^t – функція оцінки керування системою в момент часу t для стратегії ω .

Для оптимізації стратегії керування необхідно мати критерій попарного порівняння стратегій для обрання найкращої. Розглянемо три випадки порівняння для різних бажаних значень глобальної числової функції витрат:

1. Максимізація функції витрат $\tilde{F} = \max_{\omega} F_\omega$. Будемо вважати, що стратегія $\hat{\omega}$ краща за стратегію $\hat{\omega}$, якщо: $F_{\hat{\omega}} > F_{\hat{\omega}}$, $\hat{\omega}, \hat{\omega} \in \Omega$.
2. Мінімізація функції витрат $\tilde{F} = \min_{\omega} F_\omega$. Будемо вважати, що стратегія $\hat{\omega}$ краща за стратегію $\hat{\omega}$, якщо: $F_{\hat{\omega}} < F_{\hat{\omega}}$, $\hat{\omega}, \hat{\omega} \in \Omega$.
3. Досягнення бажаного значення $\sum_{t=0}^{n-1} F_\omega^t \rightarrow \tilde{F}$. Якщо для стратегій $\hat{\omega}$ та $\hat{\omega}$ виконується наступна нерівність: $|\tilde{F} - F_{\hat{\omega}}| < |\tilde{F} - F_{\hat{\omega}}|$, $\hat{\omega}, \hat{\omega} \in \Omega$, то будемо вважати, що стратегія $\hat{\omega}$ краще за стратегію $\hat{\omega}$.

Наступним кроком потрібно довести існування оптимальних детермінованих стаціонарних марковських стратегій.

1.3 Існування оптимальної стратегії керування

Процес еволюції системи з глобальною стратегією керування будемо описувати парою (ψ, ω) , де $\psi = (x_1, x_2, \dots, x_s) \in X_1 \times X_2 \times \dots \times X_S$ – стан системи, $\omega \in \Omega$ – обрана стратегія.

Теорема 1.1. *Розглянемо систему синхронних локально взаємодіючих стохастичних клітинних автоматів, задану на графі $G(V, N)$ зі скінченним простором станів X_i та скінченним простором керувань I_i , $\forall i \in V$. Нехай множина допустимих керувань системою не залежить від моменту часу t . Тоді серед множини усіх допустимих стаціонарних марковських детермінованих стратегій Ω існує оптимальна стратегія $\omega^* \in \Omega$.*

Доведення: (див. [1, с. 103]) В даному доведенні будемо розглядати мінімізацію глобальної функції витрат (доведення для максимізації та наближення до оптимального значення будуть аналогічними).

(I) Якщо система в момент часу $t = 0$ знаходилась в початковому стані $\psi^0 = (\xi_1^0, \xi_2^0, \dots, \xi_s^0)$, і для керування обрана стратегія ω , тоді:

$$E_{\psi^0}^{\omega} = \sum_{t=0}^{n-1} \sum_{i=1}^S \sum_{\xi_i^{t+1} \in X_i} Q(\xi_i^t, \xi_j^t : j \in N(i), \alpha_i^t; \xi_i^{t+1}, \gamma_i^{t+1}) r(\xi_i^t, \alpha_i^t, \gamma_i^{t+1}), \alpha_i^t \in \omega, \quad (9)$$

де $E_{\psi^0}^{\omega}$ – математичне сподівання, що відповідає керованому процесу (ψ, ω) , якщо початковий стан системи ψ^0 .

Для кожної можливої стратегії $\omega \in \Omega$ визначимо очікувані дисконтовані витрати для деякого дисконтованого множника $\beta \in (0; 1)$, якщо процес почався зі стану ψ :

$$\mathcal{F}_{\psi}^{\omega}(\beta) = E_{\psi}^{\omega} \sum_{t=0}^{n-1} \beta^t r(\psi^t, \omega, \gamma^{t+1}). \quad (10)$$

І тоді $\mathcal{F}_{\psi}(\beta) = \inf_{\omega \in \Omega} \mathcal{F}_{\psi}^{\omega}(\beta)$ – інфімум по всім допустимим стратегіям.

(II) Оскільки множини вхідних сигналів $I_i, \forall i \in V$ скінченні, і при цьому вони не залежать від історії системи та інших автоматів, можемо показати що простір допустимих стратегій Ω є компактним. В кожен фіксований момент часу t приймається рішення $\omega^t = (\omega_1^t, \omega_2^t, \dots, \omega_s^t)^T$, $\omega_i^t \in \Omega_i = I_1 \times I_2 \times \dots \times I_S$. При цьому, $\forall i \in V$, I_i – є компактним, як простір зі скін-

ченною кількістю точок. А так як Ω_i дорівнює добутку всіх множин I_i , то за теоремою Тихонова, Ω_i також компактний. А отже і $\prod_i^S \Omega_i = \Omega$ – компактний простір.

Функціонал $\mathcal{F}_\psi^\omega(\beta)$ як функція від ω є неперервним для всіх ψ і $0 \leq \beta \leq 1$, оскільки кожен його член із (10) є неперервною функцією [2, Лема 1, с. 38].

Використовуючи компактність простору допустимих стратегій та неперервність функціоналу $\mathcal{F}_\psi^\omega(\beta)$, можна показати, що для будь-якого $\beta \in (0; 1)$ очікувані дисконтовані витрати $\mathcal{F}_\psi^\omega(\beta)$ досягають свого мінімального значення на множині допустимих стратегій Ω . Тобто, існує така стратегія $\tilde{\omega} = \tilde{\omega}(\beta) = (\tilde{\omega}^0(\beta), \tilde{\omega}^1(\beta), \dots, \tilde{\omega}^t(\beta), \dots) \in \Omega$:

$$\begin{aligned} \tilde{\omega}(\beta) &= \{\tilde{\omega}_{\psi_k}(\beta), k = 1, \dots, c\}, \\ \tilde{\omega}_{\psi_k}(\beta) &= \{\tilde{\omega}_{\psi_k}^0(\beta), \tilde{\omega}_{\psi_k}^1(\beta), \dots, \tilde{\omega}_{\psi_k}^t(\beta), \dots\}, \\ \text{де } \tilde{\omega}_{\psi_k}^t(\beta) &= (\alpha_1^t, \alpha_2^t, \dots, \alpha_S^t) \in I_1 \times I_2 \times \dots \times I_S \text{ таке, що для всіх } \psi \in \\ X_1 \times X_2 \times \dots \times X_S: \end{aligned}$$

$$\mathcal{F}_\psi^{\tilde{\omega}}(\beta) = \mathcal{F}_\psi(\beta) = \inf_{\omega \in \Omega} \mathcal{F}_\psi^\omega(\beta). \quad (11)$$

Наступним кроком в доведенні [1, с. 104] потрібно показати, що якщо при зафіксованому β для критерію повних дисконтованих витрат існує оптимальна стратегія $\tilde{\omega}$, то серед допустимих стаціонарних марковських стратегій також існує деяка стратегія $\omega^* = \omega^*(\beta)$, для якої: $\mathcal{F}_\psi^{\tilde{\omega}}(\beta) = \mathcal{F}_\psi^{\omega^*}(\beta)$.

(III) Використовуючи (11), можна сказати що стратегія $\tilde{\omega}$ має наступну властивість:

$$\mathcal{F}_\psi^{\tilde{\omega}}(\beta) \leq \mathcal{F}_\psi^\omega(\beta), \omega \in \Omega. \quad (12)$$

Використовуючи умовне математичне сподівання, для будь-якого $p \geq 0$ маємо:

$$\begin{aligned} \mathcal{F}_{\psi_0}^\omega(\beta) &= E_{\psi_0}^\omega \left\{ \sum_{t=0}^p \beta^t r(\psi^t, \omega^t, \gamma^{t+1}) + \right. \\ &\quad \left. E^\omega \left(\sum_{t=p+1}^{\infty} \beta^t r(\psi^t, \omega^t, \gamma^{t+1}) \mid \psi^0 = \psi_0, \dots, \psi^p = \psi_p \right) \right\}. \end{aligned}$$

Використавши принцип Беллмана, покажемо: якщо $\psi_0, \psi_1, \dots, \psi_p$ – пер-

ші $p + 1$ станів процесу $(\psi, \tilde{\omega}(\beta))$, то для будь-якої стратегії $\omega \in \Omega$, згідно з (12), виконується:

$$\begin{aligned} E^{\tilde{\omega}}\left(\sum_{t=p+1}^{\infty} \beta^t r(\psi^t, \tilde{\omega}^t, \gamma^{t+1}) \mid \psi^0 = \psi_0, \dots, \psi^p = \psi_p\right) &\leq \\ &\leq E^{\omega}\left(\sum_{t=p+1}^{\infty} \beta^t r(\psi^t, \omega^t, \gamma^{t+1}) \mid \psi^0 = \psi_0, \dots, \psi^p = \psi_p\right) \end{aligned} \quad (13)$$

майже всюди згідно з мірою процесу $(\psi, \tilde{\omega}(\beta))$ [1, с. 105].

Нехай \mathcal{X}_p – сигма-алгебра, породжена множиною $(\psi^0, \psi^1, \dots, \psi^p)$. Припустимо, що існує стратегія $\hat{\omega} \in \Omega$ і існує певна множина $\mathcal{M}_p = \{\psi^0 = \psi_0, \dots, \psi^p = \psi_p\} \in \mathcal{X}_p$ з додатною ймовірнісною мірою процесу $(\psi, \tilde{\omega}(\beta))$, що має місце:

$$\begin{aligned} E^{\tilde{\omega}}\left(\sum_{t=p+1}^{\infty} \beta^t r(\psi^t, \tilde{\omega}^t, \gamma^{t+1}) \mid \psi^0 = \psi_0, \dots, \psi^p = \psi_p\right) &> \\ &> E^{\hat{\omega}}\left(\sum_{t=p+1}^{\infty} \beta^t r(\psi^t, \hat{\omega}^t, \gamma^{t+1}) \mid \psi^0 = \psi_0, \dots, \psi^p = \psi_p\right), \end{aligned}$$

і при цьому (13) виконується на доповненні множині $\overline{\mathcal{M}_p}$. Таким чином, можемо створити комбіновану стратегію $\hat{\omega}(\beta) = \{\hat{\omega}_{\psi_k}(\beta), k = 1, \dots, c\}$, $\hat{\omega}_{\psi_k}(\beta) = \{\hat{\omega}_{\psi_k}^t(\beta), t \geq 0\}$, для якої $\forall j : j \leq p$, політика співпадає з політикою стратегії $\tilde{\omega}(\beta)$:

$$\hat{\omega}_{\psi_k}^t(\psi^0, \dots, \psi^j) = \tilde{\omega}_{\psi_k}^t(\psi^0, \dots, \psi^j),$$

а для $j \geq 1$:

$$\begin{aligned} \hat{\omega}_{\psi_k}^{p+j}(\psi^0, \dots, \psi^p, \dots, \psi^{p+j}) &= \\ &= \begin{cases} \hat{\omega}_{\psi_k}^{p+j}(\psi^0, \dots, \psi^p, \dots, \psi^{p+j}), & (\psi^0, \dots, \psi^p) \in \mathcal{M}_p; \\ \tilde{\omega}_{\psi_k}^{p+j}(\psi^0, \dots, \psi^p, \dots, \psi^{p+j}), & (\psi^0, \dots, \psi^p) \in \overline{\mathcal{M}_p}. \end{cases} \end{aligned}$$

Отже, для стратегії $\widehat{\omega}(\beta) \in \Omega$, згідно з [1, с. 106] маємо:

$$\mathcal{F}_{\psi^0}^{\widehat{\omega}}(\beta) = E_{\psi^0}^{\widehat{\omega}} \left\{ \sum_{t=0}^p \beta^t r(\psi^t, \widehat{\omega}^t, \gamma^{t+1}) + \right. \\ \left. E^{\widehat{\omega}} \left(\sum_{t=p+1}^{\infty} \beta^t r(\psi^t, \widehat{\omega}^t, \gamma^{t+1}) \mid \psi^0 = \psi_0, \dots, \psi^p = \psi_p \right) \right\} < \mathcal{F}_{\psi^0}^{\widetilde{\omega}}(\beta),$$

що суперечить нерівності (12). Тому надалі будемо вважати, що (13) виконується для оптимальної стратегії $\widetilde{\omega}(\beta)$ для будь-якої послідовності (ψ^0, \dots, ψ^p) з ймовірністю $\Pr_{\psi^0}^{\widetilde{\omega}} > 0$. Позначимо:

$$\Phi^{\widetilde{\omega}}(\psi^0, \dots, \psi^p) = E^{\widetilde{\omega}} \left(\sum_{t=p+1}^{\infty} \beta^t r(\psi^t, \widetilde{\omega}^t, \gamma^{t+1}) \mid \psi^0 = \psi_0, \dots, \psi^p = \psi_p \right), \quad (14)$$

і аналогічно до [1, с. 107], покажемо, що $\Phi^{\widetilde{\omega}}(\psi^0, \dots, \psi^p)$ має марковську властивість, тобто залежить від історії попередніх станів лише через ψ^p : $\Phi^{\widetilde{\omega}}(\psi^0 = \psi_q, \dots, \psi^{p-1} = \psi_w, \psi^p = \psi_k) = \Phi^{\widetilde{\omega}}(\psi^0 = \psi_l, \dots, \psi^{p-1} = \psi_m, \psi^p = \psi_k)$, $\forall q, w, l, m = 1, \dots, c$.

Зафіксуємо момент часу k , і розглянемо всі можливі послідовності станів $(\psi^0, \dots, \psi^{k-1}, \psi^k)$ процесу $(\psi, \widetilde{\omega})$, такі що в момент часу k система знаходиться в стані ψ^k , і $\Phi^{\widetilde{\omega}}(\psi^0, \dots, \psi^k)$ відповідає (14). Оскільки множина всіх можливих станів скінченна $c = \prod_{i=1}^S |X_i| < \infty$, то множина всіх послідовностей $(\psi^0, \dots, \psi^{k-1}, \psi^k)$ також скінченна, і тому можемо стверджувати, що $\Phi^{\widetilde{\omega}}(\psi^0, \dots, \psi^k)$ досягає свого мінімуму на деякій послідовності $(\hat{\psi}^0, \dots, \hat{\psi}^{k-1}, \psi^k)$, і ця послідовність реалізується з додатною ймовірнісною мірою $\Pr_{\psi^0}^{\widetilde{\omega}}$.

Розглянемо будь-яку іншу послідовність, для якої в момент часу k система також перебувала в стані ψ^k : $(\check{\psi}^0, \dots, \check{\psi}^{k-1}, \psi^k)$, і оскільки мінімум $\Phi^{\widetilde{\omega}}$ вже досягається на попередній ймовірності, маємо:

$$\Phi^{\widetilde{\omega}}(\hat{\psi}^0, \dots, \hat{\psi}^{k-1}, \psi^k) \leq \Phi^{\widetilde{\omega}}(\check{\psi}^0, \dots, \check{\psi}^{k-1}, \psi^k). \quad (15)$$

Покажемо, що в (15) можлива рівність. Підемо від супротивного, і при-

пустимо, що для деякої послідовності $(\ddot{\psi}^0, \dots, \ddot{\psi}^{k-1}, \psi^k)$ виконується лише строга нерівність:

$$\Phi^{\tilde{\omega}}(\hat{\psi}^0, \dots, \hat{\psi}^{k-1}, \psi^k) < \Phi^{\tilde{\omega}}(\ddot{\psi}^0, \dots, \ddot{\psi}^{k-1}, \psi^k).$$

Тоді розглянемо нову стратегію $\ddot{\omega}$, політика якої до моменту часу k збігається з політикою стратегії $\tilde{\omega}$, а після ітерації k , для будь-якої послідовності $(\psi^0, \dots, \psi^{k-1}, \psi^k, \psi^{k+1}, \dots, \psi^{k+m})$, політика буде ідентичною до політики $\tilde{\omega}$, але за умови, що до моменту часу k для неї спостерігалась послідовність $(\hat{\psi}^0, \dots, \hat{\psi}^{k-1}, \psi^k)$:

$$\begin{aligned} \ddot{\omega}^{k+m}(\psi^0, \dots, \psi^{k-1}, \psi^k, \psi^{k+1}, \dots, \psi^{k+m}) = \\ = \tilde{\omega}^{k+m}(\hat{\psi}^0, \dots, \hat{\psi}^{k-1}, \psi^k, \psi^{k+1}, \dots, \psi^{k+m}). \end{aligned}$$

Таким чином, для нової стратегії $\ddot{\omega}$ маємо:

$$\begin{aligned} E^{\ddot{\omega}}\left(\sum_{t=p+1}^{\infty} \beta^t r(\psi^t, \ddot{\omega}^t, \gamma^{t+1}) \mid \psi^0 = \psi_0, \dots, \psi^p = \psi_p\right) \Big\} < \\ < E^{\tilde{\omega}}\left(\sum_{t=p+1}^{\infty} \beta^t r(\psi^t, \tilde{\omega}^t, \gamma^{t+1}) \mid \psi^0 = \psi_0, \dots, \psi^p = \psi_p\right) \Big\}, \end{aligned}$$

що суперечить (13), і тому дійсно $\Phi^{\tilde{\omega}}(\psi^0, \dots, \psi^k)$ не залежить від повної історії станів, а лише від ψ^k .

Для довільної послідовності станів $(\psi^0, \dots, \psi^{k-1}, \psi^k, \psi^{k+1}, \dots, \psi^{k+m})$ керованого процесу $(\psi, \tilde{\omega})$, в якому зафіксований лише стан ψ^k в момент часу k , замінимо політику прийняття рішень з моменту часу k [1, с. 108]:

$$\begin{aligned} \tilde{\omega}^{k+m}(\psi^0, \dots, \psi^{k-1}, \psi^k, \psi^{k+1}, \dots, \psi^{k+m}) \rightarrow \\ \rightarrow \tilde{\omega}^{k+m}(\hat{\psi}^0, \dots, \hat{\psi}^{k-1}, \psi^k, \psi^{k+1}, \dots, \psi^{k+m}), \quad m \geq 1. \end{aligned}$$

Так як це не впливає на значення $\Phi^{\tilde{\omega}}(\psi^0, \dots, \psi^k)$, то можемо визначити нову стратегію $\check{\omega}$, політика якої співпадає з $\tilde{\omega}$ до моменту часу k , а потім

визначається наступним чином:

$$\begin{aligned}\tilde{\omega}^{k+m}(\psi^0, \dots, \psi^{k-1}, \psi^k, \psi^{k+1}, \dots, \psi^{k+m}) = \\ \tilde{\omega}^{k+m}(\hat{\psi}^0, \dots, \hat{\psi}^{k-1}, \psi^k, \psi^{k+1}, \dots, \psi^{k+m}), \quad m \geq 0,\end{aligned}$$

причому умовне математичне сподівання при даній стратегії співпадає з матсподіванням при оптимальному керуванні $\tilde{\omega}$.

Використовуючи незалежність $\Phi^{\tilde{\omega}}(\psi^0, \dots, \psi^{k-1}, \psi^k)$ від історії попередніх станів $(\psi^0, \dots, \psi^{k-1})$, для довільного k , можна побудувати послідовність рішень

$$\tilde{\omega}_{(k)} = (\omega_{(k)}^j : j \in \mathbb{N}), \quad k \in \mathbb{N},$$

яка буде зберігати незалежність $\Phi^{\tilde{\omega}}$ для деякого фіксованого k .

Для оптимальної стратегії $\tilde{\omega}$ маємо послідовність керувань:

$$\tilde{\omega}(\psi^0, \psi^1, \dots, \psi^k, \dots) = (\tilde{\omega}^0(\psi^0), \tilde{\omega}^1(\psi^0, \psi^1), \dots, \tilde{\omega}^k(\psi^0, \psi^1, \dots, \psi^k), \dots).$$

Аналогічно до формування стратегії $\tilde{\omega}$, розглянемо для $k = 1$ керуючий вплив:

$$\begin{aligned}\tilde{\omega}_{(1)}(\psi^0, \psi^1, \dots, \psi^k, \dots) = \\ (\omega_{(1)}^0(\psi^0), \omega_{(1)}^1(\psi^1), \omega_{(1)}^2(\psi^1, \psi^2), \dots, \omega_{(1)}^k(\psi^1, \psi^2, \dots, \psi^k), \dots),\end{aligned}$$

який буде давати таке саме умовне матсподівання, як і оптимальна стратегія $\tilde{\omega}$: $\Phi^{\tilde{\omega}_{(1)}}(\psi^0) = \Phi^{\tilde{\omega}}(\psi^0)$. Тому маємо наступне: $\omega_{(1)}^0(\psi^0) = \tilde{\omega}^0(\psi^0)$.

Продовжуючи ті самі міркування для $k = 2$ і вже для стратегії $\tilde{\omega}_{(1)}$, отримаємо наступну стратегію:

$$\begin{aligned}\tilde{\omega}_{(2)}(\psi^0, \psi^1, \dots, \psi^k, \dots) = \\ (\omega_{(2)}^0(\psi^0), \omega_{(2)}^1(\psi^1), \omega_{(2)}^2(\psi^2), \omega_{(2)}^3(\psi^2, \psi^3), \dots, \omega_{(1)}^k(\psi^2, \psi^3, \dots, \psi^k), \dots),\end{aligned}$$

і відповідно: $\omega_{(2)}^0(\psi^0) = \omega_{(1)}^0(\psi^0)$, $\omega_{(2)}^1(\psi^1) = \omega_{(1)}^1(\psi^1)$.

Продовжуючи застосування даної процедури, врешті-решт отримаємо послідовність $\tilde{\omega}_{(n)}$, $n \in \mathbb{N}$. Згідно з [1, с. 110], за допомогою методу діа-

гоналізації Вейєрштрасса, можна отримати марковську стратегію $\dot{\omega}(\beta) = \{\dot{\omega}_{\psi_k}, k = 1, \dots, c\}$, для якої:

$$\dot{\omega}_{(\cdot)}(\psi^0, \psi^1, \dots, \psi^k, \dots) = (\omega_{(\cdot)}^0(\psi^0), \omega_{(\cdot)}^1(\psi^1), \omega_{(\cdot)}^2(\psi^2), \dots, \omega_{(\cdot)}^k(\psi^k), \dots),$$

і при цьому $\omega_{(\cdot)}^j(\psi^j) = \omega_{(j)}^j(\psi^j)$, $\forall j \in \mathbb{N}$.

Таким чином, загальноочікувані дисконтовані витрати для всіх допустимих стратегій $\tilde{\omega}_{(n)}$, $n \in \mathbb{N}$ є однаковими, і в точності такими як для оптимальної стратегії: $\mathcal{F}_{\psi^0}^{\tilde{\omega}_{(n)}}(\beta) = \mathcal{F}_{\psi^0}^{\tilde{\omega}}(\beta)$, $\psi^0 \in X_1 \times X_2 \times \dots \times X_S$. Більше того, маємо:

$$\lim_{n \rightarrow \infty} \mathcal{F}_{\psi^0}^{\tilde{\omega}_{(n)}}(\beta) = \mathcal{F}_{\psi^0}^{\dot{\omega}}(\beta) \implies \mathcal{F}_{\psi^0}^{\dot{\omega}}(\beta) = \mathcal{F}_{\psi^0}^{\tilde{\omega}}(\beta). \quad (16)$$

(IV) Згідно з [1, Стаціонарність, с. 104], з використанням рівняння Беллмана, можна показати, що якщо для марковської стратегії виконується (16), то дана стратегія є стаціонарною марковською. Таким чином, існує стаціонарна марковська стратегія $\omega^*(\beta)$, для якої:

$$\mathcal{F}_{\psi^0}^{\omega^*}(\beta) = \min_{\omega \in \Omega} \mathcal{F}_{\psi^0}^{\omega}(\beta).$$

Можна знайти нескінченну послідовність дисконтованих множників $\beta_\alpha \rightarrow 1$ при $\alpha \rightarrow \infty$, що оптимальні стратегії $\omega^*(\beta_\alpha)$ будуть збігатись: $\omega^*(\beta_\alpha) = \tilde{\omega}$, $\tilde{\omega} = \{\tilde{\omega}_{\psi_k} = \tilde{\omega}_{\psi_k}(\psi), k = 1, \dots, c\}$.

Візьмемо довільну послідовність $\beta_{\alpha'} \rightarrow 1, \alpha' \rightarrow \infty$ з відповідними стратегіями $\omega^*(\beta_{\alpha'}) = \{\omega_{\psi_k}^*(\beta_{\alpha'}), k = 1, \dots, c\}$, які визначаються через функції $\omega^*(\beta_{\alpha'}) = (\omega_{\psi_k}^{*t}(\beta_{\alpha'}) : t \in \mathbb{N})$, де $(\omega_{\psi_k}^{*t}(\beta_{\alpha'}))$ визначені на скінченній множині, таким чином, маючи скінченну кількість значень.

Отож, маємо нескінченну підпослідовність $\beta_{\alpha'} \rightarrow 1$ послідовності $\beta_\alpha \rightarrow 1$, якій відповідає та сама стратегія $\tilde{\omega}$.

Процес $(\psi, \tilde{\omega})$ є однорідним марковським ланцюгом з ймовірностями переходу:

$$\Pr\{\psi_i | \psi_j, \tilde{\omega}_{\psi_j}, j = 1, \dots, c\} = Q(\psi_i | \psi_j).$$

Згідно з доведенням [1, с. 112], для фіксованої стратегії $\tilde{\omega}$, використову-

ючи збіжність за Чезаро існує границя

$$\lim_{T \rightarrow \infty} \sup \frac{1}{T+1} \sum_{t=0}^T E_{\psi}^{\tilde{\omega}} r(\psi^t, \tilde{\omega}, \gamma^{t+1}) = \rho(\psi, \tilde{\omega}).$$

Застосувавши абелеву теорему до цієї границі – отримаємо дисконтовані витрати для $\beta_\alpha \rightarrow 1$. Отже, теорема доведена. \square

1.4 Знаходження оптимальної стратегії

Теорема 1.2. *Процес еволюції системи синхронних локально взаємодіючих стохастичних клітинних автоматів ергодичний і його єдиний граничний і стаціонарний розподіл $\pi^\omega = (\pi^\omega(\psi_k) : k = 1, \dots, c)$ такий, що:*

$$\pi^\omega(\psi_k) = \pi^{\omega_k}(x_1, x_2, \dots, x_S) = \prod_{i=1}^S \left(\frac{(1 - p(x_i, y_i))}{p(x_i, y_i)} \right) G(\psi_k)^{-1},$$

де $p(x_i, y_i)$ – ймовірність i -го клітинного автомату залишитись в стані x_i , якщо на вхід був поданий сигнал y_i , $G(\psi_k)$ – нормуюча константа.

Доведення даної теореми аналогічне до доведення в [3, Theorem 5.1].

Через $Q(\psi^{t+1} = \psi_p \mid \psi^t = \psi_k, \omega_k)$ позначимо ймовірність переходу системи зі стану ψ_k в стан ψ_p за умови використання керування ω_k , $k, p = 1, \dots, c$.

[1, с. 116] Процедура покращення стратегії керування полягає у наступному – вибирається довільна початкова стратегія ω_0 , і проводиться наступний алгоритм оптимізації:

1. Для обраної стратегії ω і для певної (невідомої) функції $v = (v(\psi_k) : k = 1, \dots, c)$ розв'язується система рівнянь:

$$\begin{cases} R_{\psi_k}^\omega + v(\psi_k) = r(\psi_k, \omega_k) + \sum_{p=1}^c Q(\psi_p \mid \psi_k, \omega_k) v(\psi_p), & k = 1, \dots, c \\ \sum_{k=1}^c \pi^{\omega_k}(\psi_k) v(\psi_k) = 0 \end{cases},$$

2. Для кожного $k = 1, \dots, c$ визначимо A^k як множину рішень ω_k^* , що

задовольняє

$$\sum_{i=1}^c Q(\psi_i | \psi_k, \omega_k^*) R_{\psi_i}^\omega = R_{\psi_k}^\omega,$$

і при цьому:

$$\begin{aligned} r(\psi_k, \omega_k^*) + \sum_{p=1}^c Q(\psi_p | \psi_k, \omega_k^*) v(\psi_p) \\ < r(\psi_k, \omega_k) + \sum_{p=1}^c Q(\psi_p | \psi_k, \omega_k) v(\psi_p) = R^\omega + v^\omega(\psi_k). \end{aligned}$$

3. Для всіх можливих станів системи ψ_k , для яких множина A^k непорожня, змінюємо керування ω_k на $\omega_k^* \in A^k$, таким чином формуючи нову стратегію ω^* – і повторюємо процедуру покращення. Якщо ж $\forall k = 1, \dots, c : A^k = \emptyset$, то стратегія ω є оптимальною.

Теорема 1.3. Використовуючи процедуру вдосконалення стратегій, оптимальна стратегія досягається за обмежену кількість ітерацій.

Доведення: (див. [4, Chapter 6, Theorem 1]) Нехай $\omega^1, \omega^2, \dots$ - стратегії, отримані шляхом використання процедури покращення, починаючи з довільної стратегії ω^1 . Оскільки множина допустимих стратегій Ω скінченна, і процедура призводить до постійного зменшення очікуваних витрат R^ω , існує такий номер ітерації покращення n – для якого $\omega^n = \omega^{n+1}$, тобто наступні ітерації покращення не змінюють стратегію. Тоді маємо:

$$R_{\psi_k}^{\omega^n} + v(\psi_k) = \min_{\omega \in \omega_{\psi_k}} \left\{ r(\psi_k, \omega) + \sum_{p=1}^c Q(\psi_p | \psi_k, \omega) v(\psi_p) \right\}, \quad (17)$$

і

$$R_{\psi_k}^{\omega^n} = \min_{\omega} \sum_{p=1}^c Q(\psi_p | \psi_k, \omega) R_{\psi_p}^{\omega^n}, \quad (18)$$

де ω_{ψ_k} з (17) – множина рішень ω , для яких виконується рівність (18). Тепер покажемо, що якщо для стратегії ω^n одночасно виконується (17) та (18), то ω^n є оптимальною, причому ω^n призводить не лише до локального, а й до абсолютного мінімуму.

Нехай $\tilde{\omega}$ – випадкова стратегія із множини допустимих стратегій. На

підставі виконання (17) та (18) для стратегії ω^n , використовуючи критерій дисконтованих витрат, маємо:

$$\mathcal{F}_{\psi_k}^{\omega^n}(\beta) \leq r(\psi_k, \tilde{\omega}, \gamma \mid \psi_k, \tilde{\omega}) + \beta \sum_{j=1}^c Q(\psi_j \mid \psi_k, \tilde{\omega}) \mathcal{F}_{\psi_j}^{\omega^n}(\beta), \quad \forall \beta \rightarrow 1, \quad \forall k = 1, \dots, c.$$

Використовуючи доведення з [4, Chapter 4, Theorem 2], маємо

$$\mathcal{F}_{\psi_k}^{\omega^n}(\beta) \leq \mathcal{F}_{\psi_k}^{\tilde{\omega}}(\beta), \quad \forall \beta \rightarrow 1, \quad \forall k = 1, \dots, c.$$

Згідно з [4, с. 69]:

$$R_{\psi_k}^{\omega^n} \leq R_{\psi_k}^{\tilde{\omega}}, \quad \forall k = 1, \dots, c$$

Оскільки $\tilde{\omega}$ – випадкова стратегія, то теорема доведена. \square

2 Задача поширення лісових пожеж

2.1 Постановка задачі

В даній роботі буде розглянута розширена модель поширення лісових пожеж [5].

Територія поширення лісової пожежі розглядається як двовимірна сітка розміром $L_x \times L_y$. Кожна клітина сітки моделює частину лісу, і є клітинним автоматом, який визначається:

- позицією $(i, j) \in \mathbb{Z}^2$ у сітці, де $i = 1, \dots, L_x$ – відповідний рядок, а $j = 1, \dots, L_y$ – відповідний стовпчик;
- скінченною множиною станів автомата $X_{i,j}^t \in \{E, A, B, D\}$, яка описує поточний стан ділянки лісу в кожен момент часу. Стан E (Empty) позначає ділянку лісу, яка не містить достатньої кількості рослинності щоб загорітись (імітація неоднорідності лісу). Стан A (Alive) позначає ділянку, яка знаходиться в потенціальній небезпеці. Стан B (Burning) позначає область лісу, яка горить у даний момент часу. Стан D (Dead) визначає клітину, яка вже повністю згоріла;
- множиною сусідніх клітин $N(i, j)$. В даній моделі “сусідами” будуть вважатися клітини з околу фон Неймана 1-го порядку: 4 клітини навколо обраної, що мають з нею спільне ребро [5, с.2].

Однак, розглянута модель не містить в собі керування вхідними сигналами, тому в якості доповнення буде розглянуте керування, яке імітує гасіння лісової пожежі в певному регіоні, наприклад, за допомогою авіації (пожежних літаків та гелікоптерів). Для запобігання надлишкового ускладнення моделі, будемо вважати, що для здійснення керування, особа що приймає рішення має необмежений ресурс для керування – тобто, за один крок еволюції системи немає жодних обмежень щодо масштабів витрат на гасіння пожежі.

Отож, для керування введемо скінченну множину вхідних символів $I_{i,j} \in \{Z, F\}$, де $\alpha_{i,j}^t \in I_{i,j}$ – об’єм води або іншої вогнегасної речовини, яка виливається на ділянку з позицією (i, j) в момент часу t . Тут Z (Zero) – це бездіяльність в обраній позиції, F (Full) – літак виливає повний бак для

гасіння пожежі в обраній ділянці.

Також для кожного клітинного автомата потрібно ввести множину вихідних сигналів, які ми будемо використовувати для підрахунку функції витрат. Нехай для усіх автоматів дана множина буде однаковою, і позначатиме зміну стану стану автомата $O_{i,j} = \{A \rightarrow A, A \rightarrow B, B \rightarrow B, B \rightarrow A, B \rightarrow D, D \rightarrow D, D \rightarrow A\}$

У якості чисельної функції оцінювання керування системою введемо однокрокову чисельну функцію витрат $g(\xi_{i,j}^t, \alpha_{i,j}^t)$ яка залежить від обраної стратегії гасіння пожежі в момент часу t , якщо автомат знаходився в стані $\xi_{i,j}^t$, та функцію оцінювання збитків від пожежі $f(\gamma_{i,j}^{t+1})$, яка залежить від вихідних сигналів системи при переході в момент часу $t + 1$. Під однокроковими витратами на керування системою матимемо на увазі витрати на паливо та вогнегасну речовину для авіаційної техніки на гасіння пожежі в кожній клітині. Вихідні сигнали системи будемо інтерпретувати як збитки, завдані пожежею в конкретній області в кожен момент часу t , тобто, кожна клітина, яка при переході з моменту часу t в $t + 1$ змінює свій стан (або залишається в ньому) – на вихід подає сигнал з множини виходів $\gamma_{i,j}^{t+1} = o_{i,j} \in O_{i,j}$.

Глобальна функція витрат в кожен момент часу t рахується як сума витрат по усім ділянкам лісу: $F^t = \sum_{i=1}^S g(\xi_i^t, \alpha_i^t) + f(\gamma_i^{t+1})$. Задача оптимального керування в даній моделі полягає у мінімізації глобальної функції витрат F за допомогою вибору оптимальної стратегії ω : $\tilde{F} = \min_{\omega} F_{\omega}$.

Еволюція системи відбувається наступним чином:

- *Генерація лісу.* В момент часу t_0 для кожної клітини системи генерується її початковий стан $\xi_{i,j}^0 \in \{E, A\}$. З ймовірністю q_e клітина набуває стан E , тобто не зможе загорітися взагалі, і відповідно з ймовірністю $(1 - q_e)$ – стан A .
- *Початок пожежі.* Після початкової генерації лісу, випадковим чином вибирається довільна клітина (або декілька клітин), яка має стан A , і стан обраних клітинних автоматів змінюється на B – таким чином розпочинається процес лісової пожежі.

Також, для кожної клітини у стані A існує мала, але ненульова ймовірність загорітись p_b , незалежно від того чи є поряд палаючі

клітини.

- *Поширення пожежі.* [5, с.3] Якщо в момент часу t клітина знаходиться в стані $\xi_{i,j}^t = A$, на вхід подається сигнал $\alpha_{i,j}^t = y_{i,j} \in I_{i,j}$, і в околі її сусідів $N(i, j)$ є хоча б одна клітина яка має стан B , то в наступний момент часу її стан може змінитися на $\xi_{i,j}^{t+1} = B$, а на вихід буде подано сигнал $(A \rightarrow B) \in O_{i,j}$ з ймовірністю:

$$Q(A, n_{i,j}^t, y_{i,j}; B, (A \rightarrow B)) = Pr(\xi_{i,j}^{t+1} = B, \gamma_{i,j}^{t+1} = (A \rightarrow B) | \xi_{i,j}^t = A, 1 \leq n_{i,j}^t \leq |N(i, j)|, \alpha_{i,j}^t = y_{i,j}),$$

де $n_{i,j}^t$ – кількість сусідніх клітин для клітини на позиції (i, j) , які мають стан B в момент часу t .

- *Вигорання лісу.* В кожен момент часу t ділянка лісу яка горить, може згоріти повністю, тобто, якщо на вхід подається сигнал $\alpha_{i,j}^t = y_{i,j} \in I_{i,j}$, клітина з околom $N(i, j)$ зі стану $\xi_{i,j}^t = B$ може перейти в стан $\xi_{i,j}^{t+1} = D$, а на вихід буде подано сигнал $(B \rightarrow D) \in O_{i,j}$ з ймовірністю

$$Q(B, N(i, j), y_{i,j}; D, (B \rightarrow D)) = Pr(\xi_{i,j}^{t+1} = D, \gamma_{i,j}^{t+1} = (B \rightarrow D) | \xi_{i,j}^t = B, \alpha_{i,j}^t = y_{i,j}).$$

- *Гасіння пожежі.* У випадку спроби гасіння пожежі в клітині (i, j) , якщо на вхід був поданий сигнал $\alpha_{i,j}^t = F$, а в її околі рівно $n_{i,j}^t$ палаючих клітин, то її стан в момент часу $t + 1$ може змінитися з $\xi_{i,j}^t = B$ на $\xi_{i,j}^{t+1} = A$, і на вихід буде поданий сигнал $(B \rightarrow A) \in O_{i,j}$ з ймовірністю:

$$Q(B, n_{i,j}^t, F; A, (B \rightarrow A)) = Pr(\xi_{i,j}^{t+1} = A, \gamma_{i,j}^{t+1} = (B \rightarrow A) | \xi_{i,j}^t = B, 0 \leq n_{i,j}^t \leq |N(i, j)|, \alpha_{i,j}^t = F),$$

- *Відновлення лісу.* Повністю згорівший ліс з часом має властивість відновлюватися. Нехай в момент часу t в околі клітини рівно $n_{i,j}^t$ сусідів в стані A , і при цьому жодної клітини в стані B , тоді стан

цього автомату в момент часу $t + 1$ може змінитися з $\xi_{i,j}^t = D$ на $\xi_{i,j}^{t+1} = A$, і на вихід буде поданий сигнал $(D \rightarrow A) \in O_{i,j}$ з ймовірністю:

$$Q(D, n_{i,j}^t; A, (D \rightarrow A)) = Pr(\xi_{i,j}^{t+1} = A, \gamma_{i,j}^{t+1} = (D \rightarrow A) | \xi_{i,j}^t = D, 0 \leq n_{i,j}^t \leq |N(i, j)|),$$

при цьому, якщо хоча б одна клітина з околу буде знаходитись в стані **B**, то дана ймовірність буде дорівнювати 0.

Умовно позначимо ймовірність живої клітини загорітись як p_B , палаючої клітини згоріти: p_D , мертвої клітини відновитись: p_A , і ймовірність загасити палаючу клітину як: p_E (E– extinguish). Якщо під час генерації лісу клітина не перейшла в статус **E**–Empty, то використовуючи дані позначення, її подальшу поведінку можна описати ланцюгом Маркова:

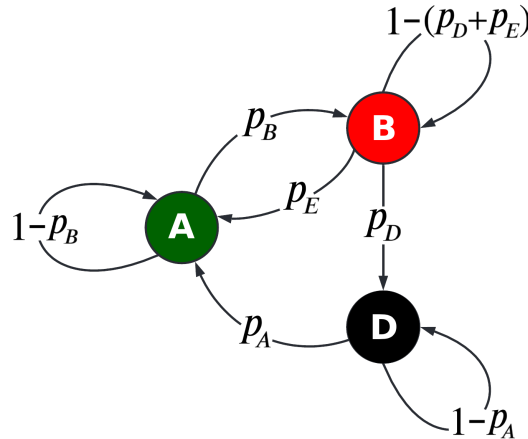


Рис. 1: Ланцюг Маркова для "живої" ділянки лісу

2.2 Програмна реалізація

Для більшого розуміння процесів які відбуваються, та для легшого сприйняття інформації – розроблений програмний застосунок, який дозволяє візуально спостерігати за процесом та керувати гасінням пожежі.

Для візуального розділення станів на сітці, стан **E** позначений коричневим кольором, стан **A** – зеленим, стан **B** – червоним, стан **D** – чорним. Для зручності розмір сітки можна регулювати в залежності від потреб.

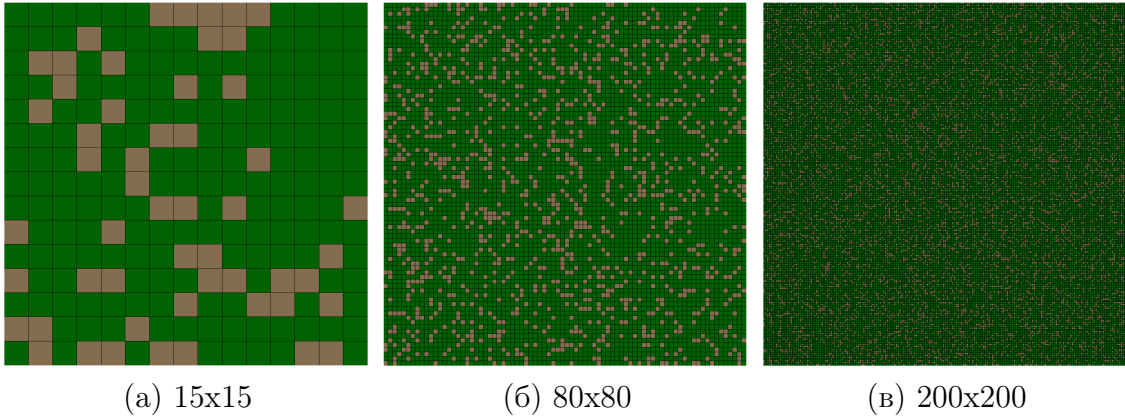


Рис. 2: Сгенерований початковий ліс з 20% "пустих" клітин

Для того щоб симулювати поширення пожежі, потрібно задати ймовірності перекидання вогню на сусідні клітини, вигорання клітин, їх гасіння та відновлення. Для спрощення моделі, дані ймовірності переходів між станами для кожного клітинного автомату є однаковими і розраховуються наступним чином:

- Коли в момент часу t клітина знаходиться в стані A , і якщо на вхід подається вхідний сигнал $\alpha_{i,j}^t \in \{Z, F\}$, то при переході системи в момент часу $t + 1$, клітина може перейти в стан B з ймовірністю:

$$\begin{cases} p_{i,j}^b = 0.35 \cdot \text{sgn}(x) + 0.2 \cdot (x - 1) - 0.3 \cdot T_{i,j}^i(\alpha_{i,j}^t), & 1 \leq x \leq 4 \\ p_{i,j}^b = 1 \times 10^{-6}, & x = 0 \end{cases},$$

де x – кількість палаючих сусідів навколо клітини на позиції (i, j) , $\text{sgn}(x)$ – знак числа x ($\text{sgn}(x) \in \{0, 1\}$), $T_{i,j}^i(\alpha_{i,j}^t) \in \{0, 1\}$ – функція перетворення вхідного сигналу $\alpha_{i,j}^t$ у числове значення.

- Для клітини в стані B ймовірність зміни стану на D при переході в момент часу $t + 1$ вираховується за формулою:

$$p_{i,j}^d = 0.2 - 0.04 \cdot x,$$

де x – кількість палаючих або живих сусідів навколо клітини на позиції (i, j) .

- Для клітини в стані B при переході в момент часу $t + 1$, ймовірність зміни стану на A (тобто гасіння даної клітини), якщо на вхід пода-

ється вхідний сигнал $\alpha_{i,j}^t \in \{Z, F\}$, дорівнює:

$$\begin{cases} p_{i,j}^e = 0.8 - 0.2 \cdot x, \alpha_{i,j}^t = F \\ p_{i,j}^b = 0, \alpha_{i,j}^t = Z \end{cases},$$

де x – кількість палаючих сусідів навколо клітини на позиції (i, j) .

- Для клітини в стані D ймовірність зміни стану на A при переході в момент часу $t + 1$ вираховується за формулою:

$$\begin{cases} p_{i,j}^r = 0.0002 + 0.001 \cdot x_a, 0 \leq x_a \leq 4 \\ p_{i,j}^b = 0, x_b > 0 \end{cases},$$

де x_a – кількість живих клітин навколо клітини на позиції (i, j) , x_b – кількість палаючих клітин в околі $N(i, j)$.

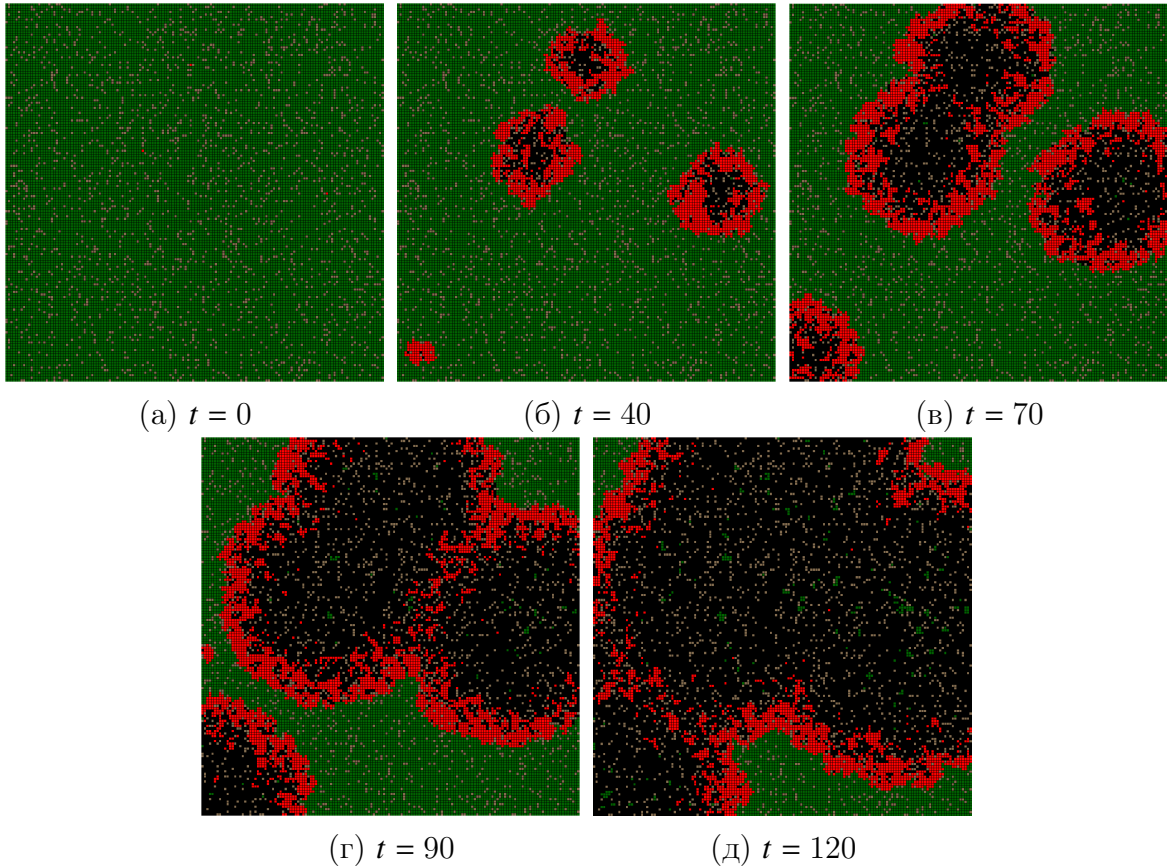


Рис. 3: Поширення пожежі з трьох початкових точок

Коефіцієнти у формулах за потреби можна корегувати, однак дані значення були вибрані методом спроб і помилок і є оптимальними для рівно-

мірного поширення пожежі та попередження занадто швидкого (або надто довгого) вигорання лісу.

В рамках розробленого програмного застосунку передбачено керування системою як і користувачем, так і автоматичне керування за прорахованою оптимальною стратегією.

Керування системою відбувається наступним чином: за замовчуванням на вхід усіх автоматів подається вхідний сигнал Z – Zero, тобто бездіяльність у кожній з областей. Якщо користувач хоче загасити пожежу в деякій клітині (або попередити поширення вогню), то він має змогу натиснути на неї, і тоді для цієї клітини вхідний сигнал зміниться на F – Full, що зміниться і на візуальному відображенні клітини у вигляді краплі води. Якщо користувач хоче змінити стратегію (або ж натискання було помилковим) – повторне натискання на клітину повертає вхідний сигнал Zero.

Після остаточного прийняття рішення та формування бажаної стратегії, після натискання відповідної клавіші користувачем, відбувається один крок еволюції системи з переходом у момент часу $t + 1$ і підрахунком функції витрат.

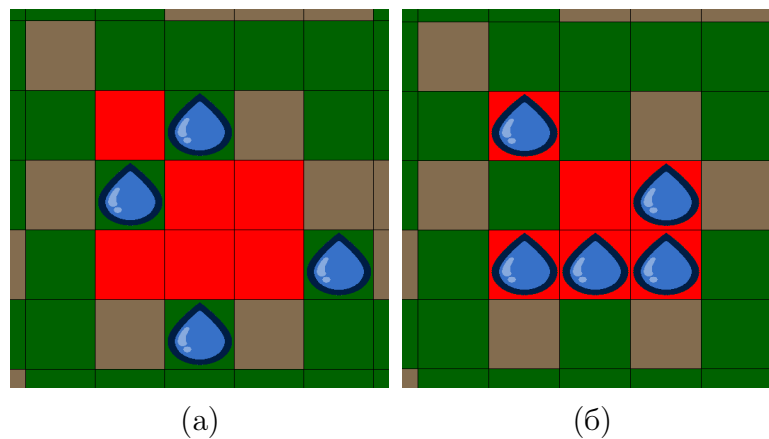


Рис. 4: Приклад керування вхідними сигналами автоматів

Введемо наступну функцію витрат:

- Нехай однокрокові витрати на керування клітинним автоматом при вхідному сигналі F будуть складати $g(\xi_{i,j}^t, F) = 15$ у.о. ($\xi_{i,j}^t \in \{A, B\}$).
- При переході автомату зі стану A в стан B , на вихід буде подаватися сигнал, відповідне чисельне значення якого буде коштувати

$f(\gamma_{i,j}^{t+1}) = 10 \text{ у.о.}, \gamma_{i,j}^{t+1} = (A \rightarrow B) \in O_{i,j}.$

- При переході автомату зі стану B в стан D , чисельні збитки будуть становити $f(\gamma_{i,j}^{t+1}) = 40 \text{ у.о.}, \gamma_{i,j}^{t+1} = (B \rightarrow D) \in O_{i,j}.$
- Якщо при переході в момент часу $t + 1$ автомат залишився в стані B , то незалежно від керування, функція витрат зросте на половину завданих збитків від переходу в цей стан, тобто $f(\gamma_{i,j}^{t+1}) = 5 \text{ у.о.}, \gamma_{i,j}^{t+1} = (B \rightarrow B) \in O_{i,j}.$
- Аналогічно до попереднього пункту, розраховуються витрати і для автомату, що при переході в момент часу $t + 1$ залишився в стані D , тобто $f(\gamma_{i,j}^{t+1}) = 20 \text{ у.о.}, \gamma_{i,j}^{t+1} = (D \rightarrow D) \in O_{i,j}.$
- Усі інші переходи не призводять до додаткових витрат.

Усі наведені значення також можна корегувати відповідно до потреб користувача чи необхідності внесення змін у поведінку моделі.

2.3 Порівняння стратегій

Для введення автоматичного керування системою потрібно додати процедуру розрахунку оптимальної стратегії. Як було зазначено раніше, можна розглядати два типи стратегій: глобальні – для всієї системи, та локальні – для кожного клітинного автомату та його околу. При виборі оптимальної стратегії, постають ключовими дві проблеми:

1. При використанні глобальної стратегії кількість усіх можливих станів системи зростає експоненційно. При 3 активних станах кожного клітинного автомата (не враховуючи стан Empty), кількість станів системи залежно від розміру сітки буде дорівнювати: при 2×2 – 81 стан, при 3×3 – **19 683** стани, при 4×4 – **43 046 721** стан, і.т.д. Враховуючи той факт, що у процедурі покращення стратегії необхідно розв'язувати систему лінійних рівнянь із кількістю невідомих, що дорівнюють кількості станів – при оптимізації стратегії вже для сітки 3×3 процес покращення буде потребувати багато ресурсів і часу, що робить використання глобальної стратегії нераціональним.
2. Для великих розмірностей сітки використання локальних стратегій значно спрощує обрахунки та зменшує час оптимізації. Однак, не

можна стверджувати що використання локальних стратегій дасть такий самий результат що і при використанні глобальної стратегії. Тобто, мінімізація витрат на локальному рівні не обов'язково призведе до глобальної мінімізації.

Використання локальних стратегій легко вирішує першу проблему, однак неминуче тягне за собою другу. Для того щоб використовувати локальні стратегії замість глобальної, потрібно перевірити наскільки сильно відрізняється функція витрат для локальних і для глобальних стратегій, якщо використовувати їх на великих проміжках часу. Якщо ж відхилення буде незначним, то можна припустити, що при великій кількості ітерацій процесів, використання локальних стратегій призводить до мінімізації глобальної функції витрат.

Дане припущення будемо перевіряти на гранично допустимій для обчислень сітці 3x3 та 15% пустих клітинок (для більших розмірностей на розрахунок оптимальної глобальної стратегії буде витрачено дуже багато часу). Для цього згенеруємо дві однакові сітки та запустимо паралельно два процеси з використанням глобальних та локальних стратегій.

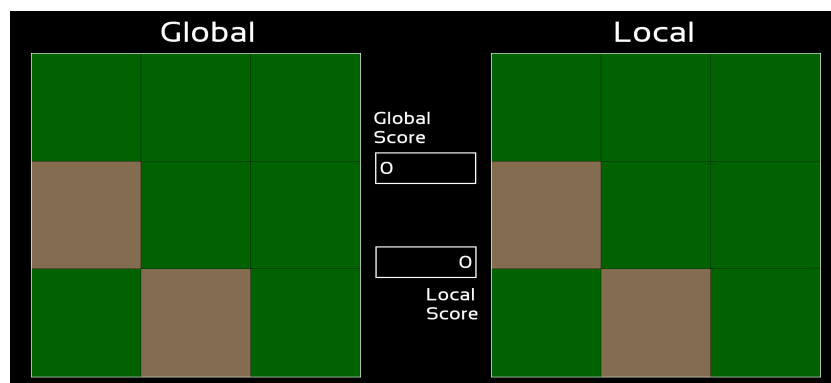


Рис. 5: Стартова конфігурація для порівняння стратегій

Під час симуляції процесів, кожен мільйон ітерацій спостерігалися проміжні значення функції витрат для порівняння динаміки зростання витрат залежно від типу стратегії, і отриманий наступний результат:

Отож, після 10 мільйонів ітерацій спостерігаємо доволі близьке розташування гілок графіка, і отримані наступні фінальні значення функцій витрат: для глобальних стратегій – 141 590 у.о., для локальних – 151 795. Різниця складає 7,2 % від значення функції витрат для глобальних стратегій. Дане

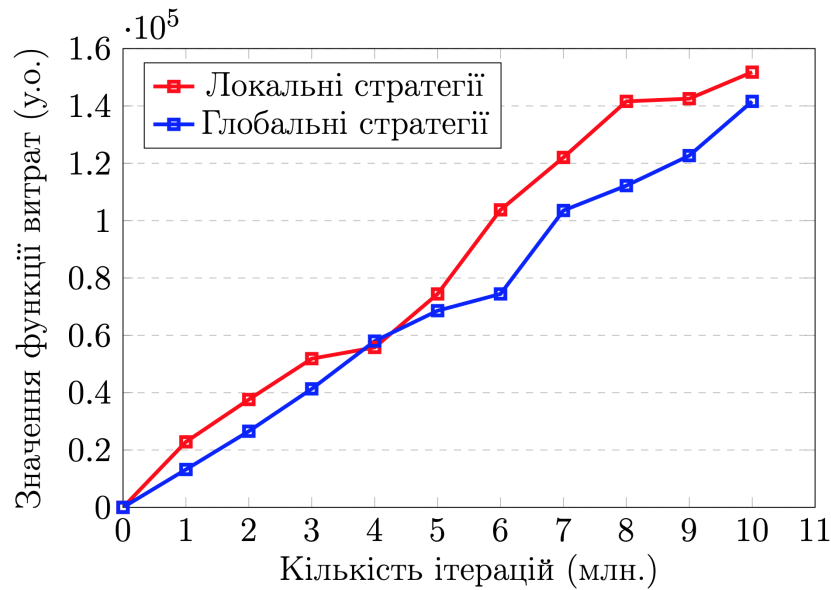


Рис. 6: Динаміка функції витрат для глобальних і локальних стратегій

відхилення не можна назвати несуттєвим, що означає що використання локальних стратегій не призводить до глобальної мінімізації функції витрат, однак, даний показник є достатньо невеликим, щоб дозволити використовувати локальні стратегії для сіток лісу більших розмірностей.

Проте, для даної конкретної задачі про лісову пожежу може виникнути природне питання: чому просто не гасити всі палаючі ділянки лісу? Тому для оцінки оптимальності застосування локальних стратегій розглянемо ще дві стратегії на більших розмірностях сітки лісу.

У якості першої стратегії для порівняння візьмемо таке керування, яке гасить усі палаючі клітини в будь-який момент часу. Для другої стратегії задамо керування, метою якого буде не допустити поширення пожежі і будемо надсилати вхідний сигнал F на всі живі суміжні клітини поряд з палаючою.

Даний експеримент розглянемо вже на сітці 20×20 з 15% пустих клітин протягом 200 тисяч ітерацій.

Як можна побачити на рис. 7 – розраховані локальні стратегії значно краще мінімізують функцію витрат, що свідчить про їх оптимальність, порівняно з двома запропонованими стратегіями.

Нарешті, можемо порівняти динаміку поширення пожежі на великій розмірності сітки за відсутності гасіння, та за керуванням за оптимальною стратегією. Для фінального порівняння візьмемо розмір сітки 60×60 та 10%

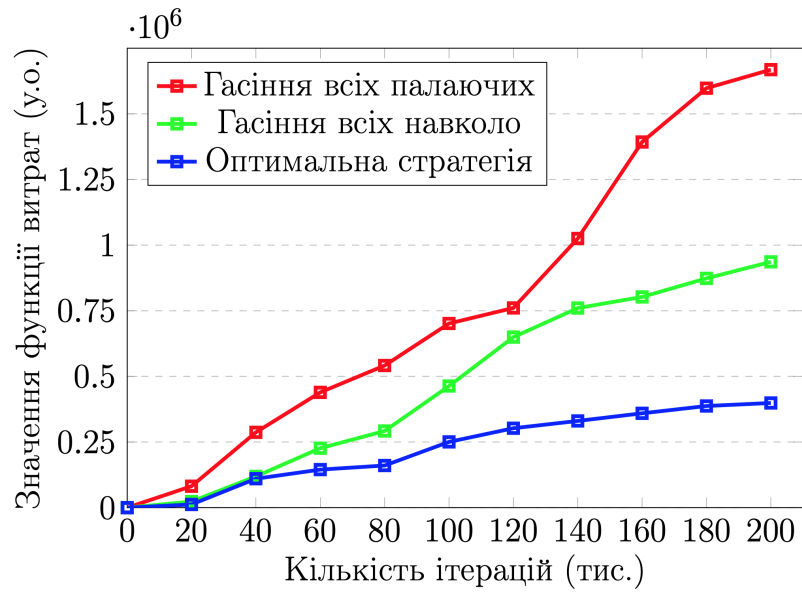


Рис. 7: Порівняння функції витрат за різних стратегій

"порожніх" клітин, і подивимось на результат після 10,20,30 та 40 ітерацій (рис.8).

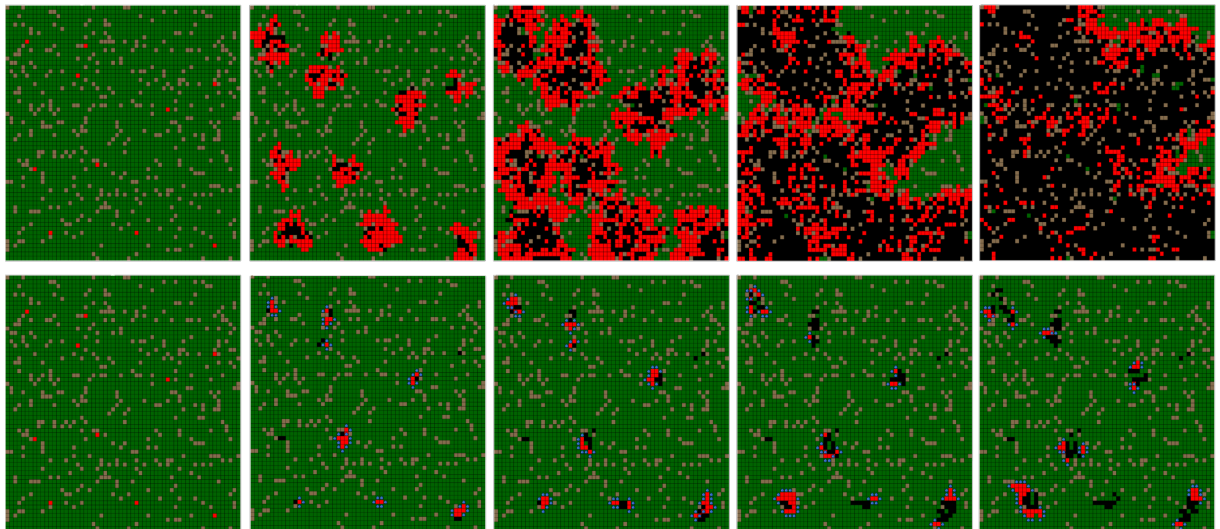


Рис. 8: Динаміка поширення пожежі без керування та за оптимальною стратегією

Висновок

У даній роботі були розглянуті синхронні локально взаємодіючі стохастичні клітинні автомати, для яких на основі критеріїв дисконтованих витрат та витрат в одиницю часу показано існування оптимальних стратегій керування в множині детермінованих стаціонарних марковських стратегій.

Змодельована прикладна задача поширення лісових пожеж, з додатковим ускладненням у вигляді керування. Розглянута процедура покращення стратегії, під час якої було виявлено, що для громіздких систем клітинних автоматів використання глобальних стратегій є доволі ресурсовитратним, на відміну від використання локальних стратегій, які дозволяють працювати з великою кількістю автоматів. Але з іншого боку, для локальних стратегій мінімізація по кожному автомату не призводить до мінімізації глобальної функції витрат, хоч і відхилення від глобальних стратегій було незначним, що вимагає від використання локальних стратегій врахування похибки відхилення від оптимального значення. Таким чином, потрібно обережно вибирати тип стратегії в залежності від кількості можливих станів кожного автомата, їх загальної кількості в системі, та від потужності множини входних сигналів.

Розроблений програмний застосунок може використовуватись і надалі для досліджень процесів поширення лісових пожеж та різних стратегій керування. Так як усі значення коефіцієнтів у ймовірностях переходів були підібрані методом спроб та помилок, отримані результати можуть бути вдосконалені шляхом підбору реалістичніших ймовірностей локальної взаємодії між автоматами, наприклад, з використанням реальних даних та статистики справжніх лісових пожеж.

Список використаних джерел

- [1] Чорней Р.К. Локальне керування в мережах. Національний університет "Києво-Могилянська академія 2021 – 311 с.
- [2] Viskov O. V., Shiryayev A. N. On controls leading to optimal stationary states // Trudy Mat. Inst. Steklov. — 1964. — Vol. 71. — P. 35–45. — [in Russian].
- [3] Agapie, Alexandru & Andreica, Anca & Giuclea, Marius. (2014). Probabilistic Cellular Automata. Journal of computational biology : a journal of computational molecular cell biology. 21. 10.1089/cmb.2014.0074.
- [4] Derman C. Finite State Markovian Decision Processes. – New York, London : Academic Press, 1970.
- [5] Rodolfo Maduro Almeida and Elbert E N Macau 2011 J. Phys.: Conf. Ser. 285 012038