

ДИВАК МИКОЛА

Західноукраїнський національний університет

<https://orcid.org/0000-0002-9049-4993>e-mail: mdy@wunu.edu.ua

КІНДЗЕРСЬКИЙ ОЛЕКСАНДР

Західноукраїнський національний університет

e-mail: o.kindzersky@gmail.com

ДОСЛІДЖЕННЯ ЕФЕКТИВНОСТІ ПАРАЛЕЛЬНОЇ ОБЧИСЛЮВАЛЬНОЇ СХЕМИ ІДЕНТИФІКАЦІЇ ІНТЕРВАЛЬНИХ ДИСКРЕТНИХ МОДЕЛЕЙ НА ОСНОВІ РОЙОВОГО ІНТЕЛЕКТУ

Проблематика досліджень, наведених у праці стосується організації обчислень для розв'язування NP-складних задач. Розглянуто задачу параметричної ідентифікації дискретних моделей об'єктів з розподіленими параметрами на основі аналізу інтервальних даних, яка відноситься до вище згаданого класу. Запропоновано для її розв'язування застосувати обчислювальні алгоритми, які побудовано на основі поведінкових моделей колонії медоносних бджіл. Застосування запропонованого обчислювального алгоритму, підвищує ефективність розв'язування задач параметричної ідентифікації інтервальних дискретних моделей об'єктів з розподіленими параметрами, проте таке застосування відзначається високою часовою складністю реалізації обчислювальних процедур. Тому, у дослідженнях, наведених у праці запропоновано використовувати алгоритми паралельних обчислень з метою зниження часової складності. Для оцінки ефективності застосування паралельних обчислень у випадку ідентифікації інтервальних дискретних моделей об'єктів з розподіленими параметрами, запропоновано використовувати обчислювальні експерименти, побудовані на ряді прикладів моделювання процесів розподілу та забруднення повітря шкідливими викидами автотранспорту.

В результаті обчислювальних експериментів для різних параметрів налаштувань обчислювальної схеми, побудованої на основі поведінкових моделей колонії медоносних бджіл, на 4 ядерному процесорі встановлено, що у всіх експериментах досягається підвищення ефективності, причому чим вища складність задачі тим більша ефективність розпаралелення. Разом з тим, середня кількість поколінь для знаходження розв'язку паралельного алгоритму у деяких експериментах значно вища ніж для послідовного алгоритму. Цей факт говорить про те, що алгоритм дуже чутливий до початково згенерованих точок в просторі пошуку розв'язків. В цілому, встановлено, що існує доцільність розпаралелювання обчислювальної схеми розв'язування задачі параметричної ідентифікації на інших більш перспективних паралельних архітектурах, наприклад на графічних конвеєрах.

Ключові слова: алгоритм Штучної Бджолиної Колонії, параметрична ідентифікація, екологічне моделювання, паралельні обчислювальні схеми, аналіз інтервальних даних.

DYVAK MYKOLA

West Ukrainian National University

KINDZERSKYI OLEKSANDR

West Ukrainian National University

INVESTIGATION OF THE EFFICIENCY OF PARALLEL COMPUTATIONAL SCHEME FOR IDENTIFICATION OF INTERVAL DISCRETE MODELS BASED ON SWARM INTELLIGENCE

The research addressed in the paper focuses on organizing computations to solve NP-complex problems. Specifically, it examines the task of parametric identification of discrete models with distributed parameters based on the analysis of interval data. Computational algorithms inspired by the behavioral models of honeybee colonies are proposed to address this task. The application of the proposed algorithm enhances the efficiency of solving parametric identification problems for interval discrete models with distributed parameters, albeit with high computational time complexity. Therefore, the study suggests using parallel computing algorithms to reduce the time complexity. To assess the effectiveness of parallel computing in identifying interval discrete models with distributed parameters, computational experiments are proposed using examples of modeling air distribution and pollution by automotive exhaust emissions.

The computational experiments, based on behavioral models of honeybee colonies on a four-core processor, demonstrate increased efficiency in all experiments, with higher task complexity leading to greater parallelization efficiency. However, it is noted that the average number of generations required for the parallel algorithm to find a solution is significantly higher in some experiments compared to the sequential algorithm. This sensitivity indicates that the algorithm is highly influenced by the initially generated points in the solution search space. Overall, the study establishes the feasibility of parallelizing the computational scheme for solving parametric identification problems on other promising parallel architectures, such as graphical processors.

Keywords: Artificial Bee Colony algorithm, parametric identification, environmental modeling, parallel computing schemes, interval approach.

Постановка проблеми

Проблематика досліджень, наведених у праці стосується організації обчислень для розв'язування NP-складних задач. Однією із таких задач є розв'язування задачі параметричної ідентифікації дискретних моделей об'єктів з розподіленими параметрами на основі аналізу інтервальних даних. Варто зазначити, що такі моделі у даній праці застосовані для моделювання процесів розподілу та забруднення повітря шкідливими викидами автотранспорту. Такі математичні моделі, хоча і характеризуються «гарантованими» прогностичними властивостями, проте вимагають застосування для їх ідентифікації складних оптимізаційних процедур з багатоекстремальними функціями мети та часто ці функції представлені в дискретному вигляді. Тому для розв'язування такого класу задач застосовують метаевристичні алгоритми оптимізації, які, своєю чергою, використовують певні загальні стратегії пошуку для знаходження прийнятного рішення проблеми

оптимізації в умовах складності та невизначеності. Зазвичай, ці алгоритми симулюють еволюційні процеси, щоб отримати оптимальне чи квазі-оптимальний розв'язок в просторі можливих розв'язків. Серед метаевристичних алгоритмів виділяють клас алгоритмів, які симулюють конкретні аспекти поведінки природних систем чи інших організованих структур, зокрема ройові алгоритми штучної бджолиної колонії. Незважаючи на те, що застосування обчислювальних алгоритмів, побудованих на поведінкових моделях колонії медоносних бджіл, підвищило ефективність розв'язування задач параметричної ідентифікації інтервальних дискретних моделей об'єктів з розподіленими параметрами, проте таке застосування відзначається високою часовою складністю реалізації обчислювальних процедур. У дослідженнях, наведених у праці запропоновано використовувати алгоритми паралельних обчислень з метою зниження часової складності. На сьогоднішній день, в рамках розв'язування вище зазначеного класу задач не існує єдиного загального підходу до оцінки ефективності застосування паралельних обчислень. Тому, для оцінки ефективності застосування паралельних обчислень у випадку ідентифікації інтервальних дискретних моделей об'єктів з розподіленими параметрами, запропоновано використовувати обчислювальні експерименти, побудовані на ряді прикладів моделювання процесів розподілу та забруднення повітря шкідливими викидами автотранспорту.

Аналіз публікацій

Для розв'язування широкого кола задач, пов'язаних із підготовкою та прийняттям рішень, як інструментарій, набули застосування математичні моделі, реалізовані за допомогою комп'ютерних програм [1-3]. Зокрема, до такого класу належать математичні моделі, які представляють у дискретному вигляді [4]. Особливістю процесів побудови цих моделей є необхідність отримання експериментальних даних за результатами спостереження за об'єктом чи явищем з метою подальшої ідентифікації математичної моделі на основі результатів спостережень [4-7]. Варто зазначити, що результати спостережень практично завжди містять похибки спостережень та неточності в даних. Крім того, сама модель є деяким спрощеним представленням реального об'єкту. Тому це спрощення також призводить до похибок [3,8-13] та неточностей в процесі моделювання. Однією із форм представлення цих похибок є використання для представлення результатів експерименту інтервального способу опису неточності [11-15]. У цьому випадку математична модель будується у вигляді рівняння, яке передбачає розрахунок оцінок характеристик об'єкта у вигляді числових інтервалів [16-18]. Такий підхід, на відміну від стохастичного, забезпечує «гарантовані» прогностичні властивості математичної моделі, не вимагає великої кількості спостережень, але одночасно знижує певним чином точність результуючої моделі [19-26]. Застосування інтервального підходу до побудови моделі також вимагає розв'язування задач структурної та параметричної ідентифікації, які в цьому випадку є NP-складними. Тому для їх розв'язування застосовують метаевристичні алгоритми - клас алгоритмів оптимізації, які використовують певні загальні стратегії пошуку для знаходження прийняттого рішення проблеми оптимізації в умовах складності або невизначеності [16-18]. Ці алгоритми симулюють еволюційні процеси, такі як мутація, селекція, та інші стратегії, щоб отримати оптимальне чи квазі-оптимальне рішення в просторі можливих варіантів [16,18, 27-30]. Серед метаевристичних алгоритмів виділяють клас алгоритмів, які симулюють конкретні аспекти поведінки природних систем чи інших організованих структур. Ось кілька прикладів:

1. алгоритм Штучної Бджолиної Колонії (Artificial Bee Colony Algorithm), - імітує поведінку рою медоносних бджіл, де бджоли обмінюються інформацією про якість рішень та спрямовують членів рою до найперспективніших ділянок щодо якості та кількості нектару [16,18, 27-30];
2. алгоритм «Джерела Світла» (Firefly Algorithm) - інспірований способом, яким світляки приваблюють один одного, отже «кращі» рішення привертають більше уваги за якістю рішення [30];
3. алгоритм Вовчої Зграї (Wolf Algorithm) - моделює взаємодію та ієрархію вовчої зграї, де вовки діють разом для досягнення спільної мети [27].

Попередній аналіз метаевристичних алгоритмів, наведений в працях [16,18, 27-30] показав, що для розв'язування задачі параметричної ідентифікації дискретної моделі, найбільш ефективним є алгоритм на підставі ройового інтелекту бджолиної колонії [16,18, 27-29]. Незважаючи на високу ефективність цього алгоритму, він залишається достатньо складним з обчислювальної точки зору. Тому у роботі запропоновано паралельну схему його реалізації [31-35]. Метою праці є дослідження ефективності паралельної обчислювальної схеми реалізації цього алгоритму на прикладних задачах моделювання для екологічного моніторингу.

Постановка задачі

У цій праці розглядаємо моделі для об'єктів з розподіленими параметрами. Як вже зазначалось, задача параметричної ідентифікації інтервальних дискретних моделей таких об'єктів ґрунтується на інтервальних даних у формі

$$[z_{i,j,h,k}^-, z_{i,j,h,k}^+], i = 0, \dots, I, j = 0, \dots, J, h = 0, \dots, H, k = 0, \dots, K, \quad (1)$$

де $[z_{i,j,h,k}^-, z_{i,j,h,k}^+]$ - числовий інтервал характеристики об'єкта, отриманої за результатами спостережень за цією характеристикою в точці з дискретними координатами $i = 0, \dots, I, j = 0, \dots, J, h = 0, \dots, H$ та на часовій дискреті, $k = 0, \dots, K$.

Враховуючи множинний спосіб представлення результатів експерименту, тобто у вигляді інтервалів (1), задача параметричної ідентифікації математичної моделі об'єкта переважно має множину рівнозначних розв'язків [18]. Математичну модель об'єкта розглядаємо, як дискретне (різницеве) рівняння у такому вигляді:

$$v_{i,j,h,k}(\vec{V}) = f_1(\vec{V}) * g_1 + f_2(\vec{V}) * g_2 + \dots + f_m(\vec{V}) * g_m$$

$$i = d, \dots, I, j = d, \dots, J, h = d, \dots, H, k = d, \dots, K, \quad (2)$$

де $v_{i,j,h,k}$ - означає модельоване значення характеристики об'єкта; d - порядок різницевої схеми (2); $\vec{g}$ - вектор параметрів моделі, значення яких необхідно оцінити на основі інтервальних даних; $f_1(\vec{V}), f_2(\vec{V}), \dots, f_m(\vec{V})$ - набір базисних функцій.

Тобто для кожної базисної функції матимемо у виразі (2), такий вектор:

$$\vec{V} = (v_{i-d,j-d,h-d,k-d}, v_{i-d+1,j-d,h-d,k-d}, \dots, v_{i-d+1,d,d,d}, \dots, v_{i,j,h-1,k})^T \quad (3)$$

Позначимо оцінки вектора параметрів за $\vec{g}$. Тоді, математична модель характеристики об'єкта матиме вигляд інтервального різничевого рівняння:

$$[v_{i,j,h,k}(\vec{V})] = [f_1(\vec{V})] * \hat{g}_1 + [f_2(\vec{V})] * \hat{g}_2 + \dots + [f_m(\vec{V})] * \hat{g}_m$$

$$i = d, \dots, I, j = d, \dots, J, h = d, \dots, H, k = d, \dots, K, \quad (4)$$

де $\vec{V} = [\hat{v}_{i-d,j-d,h-d,k-d}, \dots, \hat{v}_{i,j,h-1,k}]$ - інтервальний вектор з компонентами, які означають обчислені інтервальні оцінки $[\hat{v}_{i,j,h,k}^-; \hat{v}_{i,j,h,k}^+]$ характеристики.

Спираючись на гіпотезу, що обчислені інтервальні оцінки $[v_{i,j,h,k}^-; v_{i,j,h,k}^+]$ модельованої характеристики мають належати отриманим експериментально числовим інтервалам $[z_{i,j,h,k}^-; z_{i,j,h,k}^+]$ цієї ж характеристики об'єкта в точці з дискретними координатами $i = 0, \dots, I, j = 0, \dots, J, h = 0, \dots, H$ та у дискретні моменти часу $k = 0, \dots, K$, отримуємо математичну задачу для обчислення оцінки $\vec{g}$ вектора параметрів моделі [16]:

$$\begin{cases} [\hat{v}_{0,0,0,0}^-; \hat{v}_{0,0,0,0}^+] \subseteq [z_{0,0,0,0}^-; z_{0,0,0,0}^+], \dots, \\ [\hat{v}_{d-1,d-1,d-1,d-1}^-; \hat{v}_{d-1,d-1,d-1,d-1}^+] \subseteq [z_{d-1,d-1,d-1,d-1}^-; z_{d-1,d-1,d-1,d-1}^+] \\ z_{i,j,h,k}^- \leq [f_1(\vec{V})] * \hat{g}_1 + [f_2(\vec{V})] * \hat{g}_2 + \dots + [f_m(\vec{V})] * \hat{g}_m \leq z_{i,j,h,k}^+; \\ i = d, \dots, I, j = d, \dots, J, h = d, \dots, H, k = d, \dots, K \end{cases} \quad (5)$$

Отримана система є інтервальною системою нелінійних алгебричних рівнянь (ISNAP) відносно невідомих компонент вектора $\vec{g}$ оцінок параметрів моделі. Причому, у цій ISNAP вирази $[\hat{v}_{0,0,0,0}^-; \hat{v}_{0,0,0,0}^+] \subseteq [z_{0,0,0,0}^-; z_{0,0,0,0}^+], \dots, [\hat{v}_{d-1,d-1,d-1,d-1}^-; \hat{v}_{d-1,d-1,d-1,d-1}^+] \subseteq [z_{d-1,d-1,d-1,d-1}^-; z_{d-1,d-1,d-1,d-1}^+]$ - це початкові умови для різницевої схеми (4).

Математичний вираз (4) називаємо інтервальною дискретною моделлю об'єкта.

Враховуючи високу обчислювальну складність (комбінаторну) розв'язування цієї ISNAP, на практиці, обчислюють тільки точкові оцінки параметрів $\vec{g}$. При цьому процес оцінювання точкових значень параметрів моделі трансформують до розв'язування такої оптимізаційної задачі [16,18]:

$$\delta(\hat{g}_l) \xrightarrow{\hat{g}_l} \min, \hat{g}_l \in [g_{jl}^{low}; g_{jl}^{up}], j = 1, \dots, m, l = 1, \dots, S \quad (6)$$

де $g_{jl}^{low}, g_{jl}^{up}$ - нижнє та верхнє можливе значення кожного параметра моделі.

Зазвичай, для розв'язування нелінійної оптимізаційної задачі (6) використовують ітераційні процедури, де функції $\delta(\hat{g}_l)$ має такий вигляд [16,18]:

$$\delta(\hat{g}_l) = \max_{i=1, \dots, N} \{ |mid([\hat{v}_k]) - mid([z_k^-; z_k^+])| \},$$

$$\text{якщо } [\hat{v}_k] \cap [z_k^-; z_k^+] = \emptyset, \exists k = 0, \dots, K \quad (7)$$

$$\delta(\hat{g}_l) = \max_{i=1, \dots, N} \{ wid([\hat{v}_k]) - wid([\hat{v}_k] \cap [z_k^-; z_k^+]) \},$$

$$\text{якщо } [\hat{v}_k] \cap [z_k^-; z_k^+] \neq \emptyset, \forall k = 0, \dots, K \quad (8)$$

Якість оцінки параметрів, задаємо величиною $\delta(\hat{g}_l)$ у вигляді різниці центрів найбільш віддалених між собою модельованого та експериментального інтервалів для кожної часової дискрети, якщо ці інтервали не перетинаються. У випадку перетину цих інтервалів, функцію $\delta(\hat{g}_l)$ визначатимемо найменшою шириною їх перетину.

Очевидно, що для забезпечення збіжності ітераційної процедури при розв'язуванні задачі (6), необхідно забезпечити таку послідовність оцінювання значень функції $\delta(\hat{g}_1), \dots, \delta(\hat{g}_l)$, яка забезпечує такі умови [16]:

$$\delta(\hat{g}_1) > \delta(\hat{g}_2) > \dots, \delta(\hat{g}_{l=L}) < \Omega, \quad (9)$$

де Ω позначає область розв'язків ISNAP.

Паралельна обчислювальна схема методу

Для розв'язування задачі параметричної ідентифікації інтервальної дискретної моделі (4), обрано метаевристичний алгоритм штучної бджолоїної колонії, який ґрунтується на принципах ройового інтелекту

[16,18, 27-30]. Основна ідея АКБ полягає у реалізації поведінки колонії медоносних бджіл у процесі пошуку нектару.

У контексті поведінки бджолоїної колонії, збір нектару поділяється на фази [16]: фаза бджіл розвідників, фаза робочих бджіл і фаза бджіл дослідників. Бджоли-розвідники шукають та оцінюють джерела нектару у випадковому напрямі на початковій фазі. Якість джерела нектару, - це певна кількісна характеристика, яка залежить від кількості нектару та відстані до вулика. Після повернення до вулика бджоли-розвідники обмінюються інформацією з робочими бджолами. Останні, - обирають джерело нектару до якого полетять. Чим більша цінність джерела нектару, тим більше бджіл до нього полетить. Якщо джерело вичерпалося, знову настає фаза бджіл-розвідників.

В контексті алгоритму розв'язування оптимізаційної задачі (6), вище описані принципи функціонування колонії (рою) медоносних бджіл матимуть нижче описану реалізацію.

Перед початком роботи алгоритму задаємо вхідні дані та параметри алгоритму: інтервальні дані (1); структуру моделі (4); S – чисельність усієї популяції бджіл; MCN – загальну кількість ітерацій, $mcp=1$ – номер поточної ітерації, $LIMIT$ – числове значення, яке визначає вичерпність джерела.

Фаза ініціалізації(бджіл розвідників). Вектори, які визначають можливі точки мінімуму функції мети із (6) – вектори оцінок параметрів позначаємо за $\hat{g}_l$. У контексті поведінкової моделі бджолоїної колонії це означає, що кожен вектор координат джерела нектару відповідає одній l -тій бджолі. На цій фазі ініціалізуємо вектори $\hat{g}_l$ випадковим чином, використовуючи таку формулу:

$$[\hat{g}_{jl}^-, \hat{g}_{jl}^+] = [g_{jl}^{low}] + rand(0,1) * (g_{jl}^{up} - g_{jl}^{low}), j = 1, \dots, m, l = 1, \dots, S \quad (10)$$

Далі, обчислюємо значення функції мети за формулою (7) чи (8). Як бачимо, обчислення значень функції мети незалежні між собою для кожного із згенерованих за формулою (10) векторів $\hat{g}_l$. Тому ці обчислення доцільно проводити в паралельних потоках. Якщо кількість обчислювальних вузлів N , то усю множину векторів $\hat{g}_l, l = 1, \dots, S$ розділимо на N частин.

Фаза робочих бджіл. В контексті оптимізаційної задачі ця фаза означає пошук нових оцінок розв'язків, які забезпечують зменшення значення функції мети. Такі оцінки вектора параметрів обчислюємо за формулою:

$$\hat{g}_{jl}^{mcn} = \hat{g}_{jl} + \Phi_{jl} * (\hat{g}_{jl} - \hat{g}_{jp}), j = 1, \dots, m, p \neq l = 1, \dots, S \quad (11)$$

де Φ_{jl} – випадкове число з діапазону $[-1;1]$, $j = 1, \dots, m$ – випадково обраний індекс параметра, $\hat{g}_p$ – випадково обраний вектор координат нектару із $p \neq l = 1, \dots, S$.

Якщо отримане значення параметру не задовольняє умові $\hat{g}_l \in [g_{jl}^{low}; g_{jl}^{up}]$, то замість формули (11) використовуємо таку формулу:

$$\hat{g}_{jl}^{mcn} = \hat{g}_{jl} - \Phi_{jl} * (\hat{g}_{jl} - \hat{g}_{jp}), j = 1, \dots, m, p \neq l = 1, \dots, S \quad (12)$$

Далі, попарно порівнюємо значення функції мети для поточних та обчислених оцінок параметрів:

$$\hat{g}_l = \{\hat{g}_l, \text{якщо } \delta(\hat{g}_l) \leq \delta(\hat{g}_l^{mcn})\} \text{ або } \hat{g}_l = \{\hat{g}_l^{mcn}, \text{якщо } \delta(\hat{g}_l) > \delta(\hat{g}_l^{mcn})\} \quad (13)$$

Після цього збільшуємо лічильник спроб (вичерпаності джерела нектару) $limit < LIMIT$ для отриманого значення $\hat{g}_l$, щоб мати можливість вийти з локального мінімуму в майбутньому.

Як бачимо, обчислення значень функції мети незалежні між собою для кожного із згенерованих за формулами (11), (12) векторів $\hat{g}_l$. Тому ці обчислення також доцільно проводити в паралельних потоках.

Фаза бджіл дослідників. В контексті оптимізаційної задачі, на цій стадії визначаємо найбільш ймовірні точки (вектори значень параметрів), в околі яких необхідно проводити детальне дослідження функції мети. Для цього для кожної точки, встановленої на попередній фазі, обчислюємо ймовірність P_l за наступною формулою:

$$P_l = \frac{1 - \delta(\hat{g}_l)}{\sum_{l=1}^S (1 - \delta(\hat{g}_l))} \quad (14)$$

Варто зазначити, що попередньо значення $\delta(\hat{g}_l)$ нормуємо до 1.

На підставі розрахованих ймовірностей за формулою (14) визначаємо кількість точок для дослідження околу можливих локальних мінімумів функції мети – $m_l = \text{int}(P_l * S)$ точок. Відповідно у випадку $m_l > 0$, генеруємо m_l сусідніх точок за формулами (11) чи (12). Після обчислення координат кожної точки околу, організуємо N потоків обчислень, подібно, як на попередніх фазах. В кожному потоці для точок, в околі яких $m_l > 0$, обчислюємо значення функції мети і попарно порівнюємо ці значення для поточних та обчислених оцінок параметрів за формулою (13).

Наступний кроком є пошук точки з найменшим значенням функції мети на даній ітерації алгоритму. Якщо значення функції мети $\delta(\hat{g}_l) = 0$, то завершення процедури оптимізації. В протилежному випадку - покладемо $limit = 0$ і переходимо на початок цієї фази.

Фаза бджіл розвідників. Ця фаза використовується для уникнення зациклення на локальних мінімумах функції мети. В контексті поведінкової моделі бджолоїної колонії, на цю фазу переходимо, коли джерело нектару є вичерпаним. Для цього у кожній точці перевіряємо чи не перевищує її лічильник спроб $limit$ граничне значення $LIMIT$. Якщо- так, то відповідний параметр моделі, замінюємо оцінкою, яку обчислюємо за формулою (10). Після цього переходимо до фази робочих бджіл.

Експериментальні дослідження методу

Дослідження ефективності запропонованої обчислювальної схеми проведемо на прикладах побудови інтервальних моделей, які описують процеси забруднення повітря шкідливими викидами автотранспорту.

Приклад 1.

Розглянемо параметричну ідентифікацію математичної моделі процесу поширення концентрації окису вуглецю на прямолінійній ділянці вулиці внаслідок рівномірного руху транспортного потоку з постійною потужністю викидів. Вказаний процес описано різницеvim рівнянням у такому загальному вигляді:

$$v_k = g_1 + g_2 * v_{k-1} + g_3 * v_{k-2}, k = 2, \dots, K$$

Для ідентифікації для ідентифікації параметрів цієї моделі отримано концентрації окису вуглецю з відносною похибкою 5%. Вимірювання проводилися перпендикулярно до дороги на відстані від 0 до 100 метрів кожні 10 метрів [16].

Таблиця 1

Інтервальні дані вимірювання концентрації окису вуглецю

Дискрета, k	Виміряна концентрація CO, z_k	Нижня межа виміряної концентрації CO, z_k^-	Верхня межа виміряної концентрації CO, z_k^+
k0	55	52,25	57,75
k1	47	44,65	49,35
k2	43	40,85	45,15
k3	37	35,15	38,85
k4	32	30,4	33,6
k5	30	28,5	31,5
k6	26	24,7	27,3
k7	23	21,85	24,15
k8	20	19	21
k9	18	17,1	18,9
k10	16	15,2	16,8

За початкові умови покладемо інтервальні оцінки з відхиленнями у межах $\pm 2.2\%$ виміряної концентрації окису вуглецю. В результаті застосування вищеописаного методу параметричної ідентифікації моделі, отримали:

$$\hat{v}_k = 0,48711988 + 0,5943693 * \hat{v}_{k-1} + 0,24638686 * \hat{v}_{k-2}$$

Вимірювання часу виконання програми алгоритму проводилися на персональному комп'ютері з такими характеристиками: 4 ядерний CPU: Intel(R) Core(TM) i5-6500 CPU @ 3.20GHz; RAM: 16Gb Kingston DDR3 SDRAM.

Загалом, проведено по 100 експериментів для таких початкових параметрів алгоритму: $S = 8, 16, 32, 64$, MCN – до знаходження розв'язку, $\text{LIMIT} = S/2$.

Результати обчислення часової складності виконання алгоритму наведено в таблиці 2.

Таблиця 2

Порівняльна характеристика послідовної і паралельної реалізації алгоритму бджолоїної колонії

S, кількість бджіл	Час послідовного алгоритму, мс	Середня кількість поколінь для знаходження розв'язку послідовного алгоритму, MCN	Час паралельного алгоритму, мс	Середня кількість поколінь для знаходження розв'язку паралельного алгоритму, MCN	Коефіцієнт	Коефіцієнт з урахуванням кількості ітерацій
8	10.11	515.24	9.72	559.58	1.04	1.13
16	5.35	158.51	4.14	162.49	1.29	1.32
32	5.28	77.43	3.34	80.2	1.58	1.64
64	6.04	43.6	4.21	52.93	1.43	1.74

Як бачимо, для усіх проведених обчислювальних експериментів, ефективність паралельного алгоритму незначно вища ніж послідовного.

Приклад 2.

Моделювання динаміки розподілу концентрацій діоксиду азоту внаслідок забруднення повітря викидами автотранспорту.

Вказаний процес описано різницеvim рівнянням у такому загальному вигляді:

$$v_{k,i,j} = g_1 + g_2 * v_{k-1,i,j} + g_3 * v_{k,i,j-1}, k = 1, \dots, K, i = 0, \dots, I, j = 1, \dots, J$$

Для ідентифікації параметрів моделі, за допомогою рою безпілотних літальних апаратів ІДМДО проведено вимірювання концентрації забруднення з відносною похибкою 15% на площині, з визначеними дискретними координатами i, j та в одночасно у заданий дискретний момент k .

За початкові умови покладемо інтервальні оцінки з відхиленнями у межах $\pm 3\%$ вимірної концентрації діоксиду азоту, дискрету $k = 0$ і перший стовпець для кожної дискрети $k=1..9$.

Таблиця 3

Дані вимірювання концентрації діоксиду азоту

k	i/j					k	i/j				
0	14.87	15.23	15.92	16.51	17.09	5	26.51	26.83	27.58	28.54	29.14
	19.32	20.76	20.43	21.54	21.97		31.65	32.54	32.75	32.97	33.61
	23.36	24.45	25.11	25.69	26.21		38.23	38.62	39.51	39.94	40.24
	29.23	30.87	32.04	31.18	31.78		39.82	40.35	40.62	41.42	41.93
	39.24	38.97	40.54	40.87	41.48		47.32	48.54	48.98	49.35	49.76
1	14.69	15.32	15.73	16.45	16.91	6	30.56	31.14	31.87	32.45	32.87
	17.23	18.82	19.34	19.45	19.82		35.76	36.22	36.54	37.74	37.41
	26.57	27.76	28.32	29.37	29.95		39.43	39.91	40.64	40.32	40.87
	35.36	35.92	36.51	37.35	37.73		47.56	48.12	48.73	49.26	49.92
	31.54	32.45	32.87	33.24	33.86		54.23	55.36	55.89	56.43	57.41
2	22.14	22.65	22.98	23.67	24.26	7	31.65	32.54	32.87	33.35	33.91
	27.36	27.92	28.83	29.16	29.78		38.35	38.77	39.42	40.65	41.76
	31.54	32.41	32.65	33.13	33.74		46.65	47.23	48.76	49.12	49.87
	38.27	38.82	39.53	40.46	40.82		55.41	55.72	56.23	56.84	57.48
	39.45	40.12	40.76	41.54	41.95		63.72	64.23	64.76	65.43	65.82
3	27.31	27.78	28.34	28.87	29.54	8	26.23	27.86	28.54	29.41	29.81
	31.23	32.35	32.76	32.85	33.25		34.65	34.84	35.26	36.58	37.11
	39.78	40.43	40.78	41.43	41.91		39.85	40.63	41.34	41.57	41.83
	46.23	46.54	47.56	47.98	48.45		55.76	56.23	56.81	57.15	57.78
	31.34	31.76	32.23	32.54	33.13		47.56	48.43	48.87	49.54	49.93
4	22.47	22.81	23.63	23.68	24.65	9	22.34	22.95	23.43	23.87	24.65
	27.54	28.46	28.81	28.97	29.46		30.25	30.74	31.15	31.78	32.21
	34.25	34.73	35.42	35.81	36.53		34.76	35.65	35.98	36.76	37.23
	39.34	40.45	40.76	41.54	41.82		38.94	39.65	39.92	40.45	41.24
	29.73	30.35	30.85	31.53	31.95		34.21	34.76	35.34	36.78	37.12

В результаті застосування методу параметричної ідентифікації моделі, отримали:

$$\hat{v}_{k,i,j} = 0,8147161 - 0,040205844 * \hat{v}_{k-1,i,j} + 1,0450916 * \hat{v}_{k,i,j-1}$$

Як і у попередньому випадку, проведено по 100 обчислювальних експериментів для заданих параметрів алгоритму: $S = 4, 8, 16, 32$, MCN – до знаходження розв'язку, $LIMIT = S/2$.

Результати обчислення часової складності виконання алгоритму з розпаралеленням обчислювальних потоків та без розпаралелення наведено в таблиці 4.

Як бачимо, для цієї задачі, ефективність паралельного алгоритму майже для всіх обчислювальних експериментів в понад два рази вища ніж послідовного.

Таблиця 4

Порівняльна характеристика послідовної і паралельної реалізації алгоритму бджолоїної колонії

S, кількість бджіл	Час послідовного алгоритму, мс	Середня кількість поколінь для знаходження розв'язку послідовного алгоритму, MCN	Час паралельного алгоритму, мс	Середня кількість поколінь для знаходження розв'язку паралельного алгоритму, MCN	Коефіцієнт	Коефіцієнт з урахуванням кількості ітерацій
4	20.23	109.35	14.07	137.72	1.44	1.81
8	16.23	46.16	7.83	49.37	2.07	2.22
16	20.69	27.15	9.54	27.5	2.17	2.2
32	28.26	18.08	12.1	19.21	2.34	2.48

Приклад 3.

Розглянемо приклад застосування запропонованого методу до моделювання розподілу концентрації діоксиду азоту, який спричинено шкідливими викидами у вихлопних газах автотранспорту, без урахування його поширення. Для отримання експериментальних даних вимірювання проводили на висоті $h=1,5$ метра від дорожнього полотна на ділянці розміром 32.4328 на 32.4328 метра із дискретизацією заданої ділянки по 20 точок для кожної координати, враховуючи відносну похибку вимірювання 20%.

Таблиця 5

Вимірювання концентрацій діоксиду азоту NO2 на заданих дискретах

i/j	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
0	45,134	45,604	45,134	44,663	44,663	44,193	44,193	44,663	45,134	45,604	45,604	45,134	47,014	47,014	47,01	47,484	47,484	48,895	48,425	47,484
1	48,425	48,425	48,895	49,835	50,305	49,835	49,365	49,365	49,365	49,365	49,835	49,835	50,305	49,835	49,84	50,775	49,835	48,895	48,895	48,425
2	49,835	49,365	49,365	50,775	49,835	49,835	49,835	49,365	48,895	49,835	50,305	50,775	51,245	51,245	51,72	52,186	51,245	50,775	51,245	50,775
3	48,895	47,484	47,014	47,014	47,014	46,544	47,954	47,014	46,074	47,484	47,954	47,954	48,425	47,954	47,95	48,895	48,425	47,484	47,484	46,544
4	46,074	46,074	45,604	44,663	45,134	44,663	45,134	44,663	44,193	45,134	44,193	44,193	45,134	44,663	44,19	45,134	45,134	44,663	45,604	45,604
5	46,544	46,544	47,014	47,484	48,895	48,895	48,425	48,895	48,425	49,365	49,365	49,835	51,716	51,716	51,72	52,656	52,656	53,126	53,126	55,007
6	56,887	58,298	58,768	59,708	60,648	61,118	60,648	61,589	61,589	61,589	62,059	61,589	61,118	62,059	61,59	61,589	62,529	62,059	61,589	62,529
7	62,529	63,469	63,469	63,939	65,35	65,35	65,35	65,82	65,82	64,88	64,409	63,469	62,999	62,999	62,06	61,589	62,059	60,648	60,178	61,118
8	60,648	60,178	60,648	60,178	60,648	61,118	60,648	61,589	61,589	61,118	62,059	62,529	61,589	62,059	63	63,469	64,88	65,35	67,23	67,7
9	67,23	68,171	69,581	70,051	70,051	70,991	71,462	72,402	73,342	73,812	75,223	76,163	77,103	78,514	79,92	80,394	81,335	82,275	83,215	84,626
10	85,566	85,566	87,917	88,387	88,857	89,797	92,148	93,088	94,028	96,379	97,319	99,2	99,67	100,14	101,1	102,02	101,08	102,02	101,55	101,55
11	101,08	100,61	100,14	100,14	100,14	99,67	100,61	99,67	98,26	97,789	96,379	95,909	94,969	94,028	94,5	94,028	93,088	92,148	90,737	88,387
12	88,387	86,506	86,036	86,036	85,096	84,155	84,626	83,685	83,215	84,155	82,745	81,335	80,864	79,924	80,39	78,984	78,044	78,514	77,573	76,633
13	77,103	76,633	76,163	76,163	75,693	75,223	75,693	74,753	74,753	74,753	74,282	73,812	74,282	73,342	72,4	71,932	71,462	71,462	70,521	69,581
14	69,581	70,051	69,111	69,111	68,171	67,23	67,23	66,29	66,29	65,35	64,88	65,35	66,76	66,29	67,23	68,171	67,7	67,23	67,23	66,76
15	66,76	66,76	65,82	66,76	65,35	64,88	64,88	64,409	62,999	63,469	63,469	62,529	62,999	62,529	62,53	62,999	62,529	62,059	62,999	62,059
16	61,118	60,648	59,708	60,178	59,708	59,238	60,178	59,708	59,238	59,708	59,238	58,768	58,298	57,827	57,83	57,827	57,357	56,887	57,827	56,887
17	57,357	58,298	57,357	57,357	58,298	58,298	61,118	61,589	62,059	63,939	63,939	64,409	65,35	65,82	65,82	65,82	65,82	65,82	66,76	66,76
18	66,76	67,7	66,76	66,76	66,29	65,82	64,409	64,409	63,469	63,939	63,469	62,059	62,529	61,118	61,12	60,648	61,118	58,768	59,238	58,768
19	58,768	58,298	57,357	56,887	57,357	56,887	56,417	57,827	56,417	55,477	55,477	55,477	55,947	55,477	55,01	54,066	54,066	53,126	52,186	52,186

Вказаний процес описано різницею рівнянням у такому загальному вигляді:

$$v_{i,j} = g_1 + g_2 * v_{i,j-1} + g_3 * v_{i-1,j} + g_4 * v_{i-1,j-1} + g_5 * v_{i,j-2} + g_6 * v_{i,j-3},$$

$$i = 1, \dots, I, j = 3, \dots, J$$

За початкові умови покладемо інтервальні оцінки з відхиленнями у межах $\pm 5\%$ виміряної концентрації діоксиду азоту, перший стовпець і 3 перші рядки. В результаті методу параметричної ідентифікації ІДМ, отримали таку модель:

$$\hat{v}_{i,j} = 0,38556296 + 0,9205657 * \hat{v}_{i,j-1} - 0,7482962 * \hat{v}_{i-1,j} + 0,7141639 * \hat{v}_{i-1,j-1} - 0,12739463 * \hat{v}_{i,j-2} + 0,23220368 * \hat{v}_{i,j-3}$$

Для цього випадку проведено 10 експериментів для заданих параметрів алгоритму: $S = 32$, MCN – до знаходження розв'язку, LIMIT = S/2.

Результати обчислення часової складності виконання алгоритму для цієї задачі з розпаралеленням обчислювальних потоків та без розпаралелення, наведено в таблиці 6.

Таблиця 6

Порівняльна характеристика послідовної і паралельної реалізації алгоритму бджолоїної колонії

Номер запуску	Час послідовного алгоритму, мс	Середня кількість поколінь для знаходження розв'язку послідовного алгоритму, MCN	Час паралельного алгоритму, мс	Середня кількість поколінь для знаходження розв'язку паралельного алгоритму, MCN	Коефіцієнт	Коефіцієнт з урахуванням кількості ітерацій
1	28439,1	6896,2	3689,8	2158,5	7,70	2,41
2	41920,5	4602	5219,2	1442,5	8,03	2,52
3	10926,5	1766,4	11294,1	4573,1	0,97	2,5

Як бачимо, для цієї задачі, ефективність (за кількістю ітерацій) паралельного алгоритму для всіх обчислювальних експериментів в понад два рази вища ніж послідовного. Разом з тим, середня кількість поколінь для знаходження розв'язку паралельного алгоритму у третьому експерименті значно вища ніж для послідовного алгоритму. Цей факт говорить про те, що алгоритм дуже чутливий до початково згенерованої точки за формулою (10).

Висновки

Задача моделювання процесів забруднення приземистого шару атмосфери вихлопними газами автотранспорту є однією із найважливіших при моніторингу довкілля міст. Розглянуто приклад застосування параметричної ідентифікації інтервальної дискретної моделі складних об'єктів на підставі АБК до моделювання розподілу концентрації шкідливих викидів.

В результаті аналізу обчислювальної схеми, встановлено, що базова схема обчислень в основі якої є моделювання процесу поведінки рою медоносних бджіл, має достатню природу для розпаралелювання. Запропоновано нову схему паралельних обчислень розв'язування задачі параметричної ідентифікації інтервальної дискретної моделі складних об'єктів.

В результаті обчислювальних експериментів для різних параметрів налаштувань обчислювальної схеми АБК на 4 ядерному процесорі встановлено, що у всіх експериментах досягається підвищення ефективності, причому чим вища складність задачі тим більша ефективність розпаралелення. Разом з тим, середня кількість поколінь для знаходження розв'язку паралельного алгоритму у деяких експериментах значно вища ніж для послідовного алгоритму. Цей факт говорить про те, що алгоритм дуже чутливий до початково згенерованої точки за формулою (10). Проте, в цілому, існує можливість та доцільність розпаралелювання обчислювальної схеми розв'язування задачі параметричної ідентифікації на інших більш перспективних паралельних архітектурах, наприклад, на графічних конвеєрах, що буде предметом подальших досліджень.

References

1. Goddard, W. & Melville, S. "Research Methodology: An Introduction" 2nd edition, Blackwell Publishing, 2004
2. Box, G. E. P., Hunter, W. G., & Hunter, J. S. Statistics for Experimenters: Design, Innovation, and Discovery. Wiley, 2005, 654 p.
3. Shadish, W. R., Cook, T. D., & Campbell, D. T. (). Experimental and Quasi-Experimental Designs for Generalized Causal Inference, Houghton Mifflin, 2002, 623 p.
4. Palais, R. S., Osserman, R., & Brin, M. Differential Equations, Mechanics, and Computation. Yale University Press, 2021, 392 p.
5. Braun, M. Differential Equations and Their Applications: An Introduction to Applied Mathematics, 4th Edition. Springer, 2020, 559 p.
6. Ahmad, B., & Alsaedi, A. Fractional Order Difference Equations. Springer, 2021, 318 p.
7. Baleanu, D., Wang, Z., & Raja, M. A. Z. Recent Advances in Fractional Difference Calculus and Its Applications. Hindawi, 2021, 308 p.
8. M. Dyvak, I. Voytyuk, N. Porplytsya and A. Pukas, "Modeling the process of air pollution by harmful emissions from vehicles," 2018 14th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET), Slavske, 2018, pp. 1272-1276, doi: 10.1109/TCSET.2018.8336426.
9. M. Dyvak, N. Porplytsya, "Formation and Identification of a Model for Recurrent Laryngeal Nerve Localization During the Surgery on Neck Organs", Advances in Intelligent Systems and Computing III. CSIT 2018, Cham: Springer, vol.871, pp. 391-404, 2019.

10. N. Ocheretnyuk, M. Dyvak, T. Dyvak, I. Voytyuk "Structure identification of interval difference operator for control the production process of drywall," in Proc. 12th Int. Conf. on Experience of Designing and Application of CAD Systems in Microelectronics (CADSM), 2013, pp. 262-264.
11. Dyvak, M., Papa, O., Melnyk, A., Pukas, A., Porplytsya, N., Rot, A. Interval model of the efficiency of the functioning of information web resources for services on ecological expertise (2020) *Mathematics*, 8 (12), art. no. 2116, pp. 1-12.
12. M. Dyvak, N. Kasatkina, A. Pukas, N. Padletska. "Spectral analysis the information signal in the task of identification the recurrent laryngeal nerve in thyroid surgery". *Przegląd Elektrotechniczny*, vol. 89, no. 6, pp. 275-277, 2013.
13. Porplytsya, N., Dyvak, M.: Interval difference operator for the task of identification recurrent laryngeal nerve. In: *Proceedings of the 16th International Conference on Computational Problems of Electrical Engineering (CPEE'2015)*, pp. 156-158 (2015).
14. Dyvak, M., Pukas, A., Oliynyk, I., Melnyk, A. Selection the 'Saturated' Block from Interval System of Linear Algebraic Equations for Recurrent Laryngeal Nerve Identification *Proceedings of the 2018 IEEE 2nd International Conference on Data Stream Mining and Processing, DSMP 2018*, 2018, pp. 444–448, 8478528
15. M. Dyvak, V. Manzhula, O. Kozak. New method tolerance estimation of the parameters set of interval model based on saturated block of ISLAE, *Proceeding of the IX-th International Conference CADSM'2007, Lviv–Polyana*, 2007, p. 376-379.M.
16. Dyvak Mykola. (2020). Parameters Identification Method of Interval Discrete Dynamic Models of Air Pollution Based on Artificial Bee Colony Algorithm. 130-135. 10.1109/ACIT49673.2020.9208972.
17. Ocheretnyuk, N.; Voytyuk, I.; Dyvak, M.; Martsenyuk, Ye. Features of structure identification the macromodels for nonstationary fields of air pollutions from vehicles. In *Proceedings of the Modern Problems of Radio Engineering, Telecommunications and Computer Science—Proceedings of the 11th International Conference, Lviv, Ukraine, 17–19 May, 2012*; p. 444.
18. M. Dyvak; N. Porplytsya; Y. Maslyiak and N. Kasatkina. Modified artificial bee colony algorithm for structure identification of models of objects with distributed parameters and control. *2017 14th International Conference The Experience of Designing and Application of CAD Systems in Microelectronics (CADSM), Lviv, 2017*, pp. 50-54.
19. Baleanu, D., Çakır, A. F., & Abdeljawad, T. *Fractional Order Nonlinear Systems: Modeling, Analysis, and Simulation*. Elsevier, 2020, 418 p.
20. Gallant, A. R. *Nonlinear statistical models*. John Wiley & Sons, 1987, 353 p.
21. Chen, R., & Li, X. *Parametric statistical change point analysis*. Birkhäuser, 2009, 290 p.
22. Pakes, A. *Identification, Information, and Learning: The 21st Century Toolbox for Economics*. Harvard University Press, 2021, 368 p.
23. Kostin, A., & Kopnov, V. Parametric identification of dynamic systems: Modern approaches and algorithms. *ISA Transactions*, 117, 2021, pp. 49-64
24. Balakrishnan, V., & Balachandran, B. *Model-Based Parameter Estimation: Theory and Applications*. Springer, 2021, 409 p.
25. Piga, D., & Stefanoiu, D. *Identification of Parametric Models from Experimental Data*. Springer, 2020, 319 p.
26. Hassani, V., & Schoukens, M. *Estimation of Parametric Models from Experimental Data*. Springer, 2019, 225 p.
27. Bhattacharya, A., & Chatterjee, A. Multi-Objective Grey Wolf Optimizer and Artificial Bee Colony Algorithm for Electrical Distribution Network Reconfiguration. *Applied Soft Computing*, 121, 107086, 2022, 17 p.
28. Basnet, B. B., Duan, J., Pan, Q. K., & Bhattarai, M. Design optimization of ultra-wideband antenna using the hybrid artificial bee colony algorithm. *Wireless Personal Communications*, 2022, pp. 1-19.
29. Zafar, M. A., Khan, A., Tariq, M., & Naeem, M. A. A Comprehensive Review on Artificial Bee Colony Algorithm and Its Variants. *Arabian Journal for Science and Engineering*, 47(1), 2022, pp. 113-133.
30. Yang, X. S. "Firefly algorithm, Levy flights and global optimization." In *Research and Development in Intelligent Systems XXVI*, Springer, 2010, pp. 209-218
31. Parashar, M. & Sun, X-H. *Parallel Computing: Algorithms, Architectures, and Applications*. CRC Press, 2021, 566 p.
32. Ivanov, P., & Petrov, V. (Eds.). *Computational Intelligence in Parallel and Distributed Systems: Concepts, Techniques, and Applications*. IGI Global, 2021, 420 p.
33. Chen, Y., & Zhang, L. *Parallel Metaheuristics: A New Class of Algorithms*. Springer, 2018, 300 p.
34. Wang, J., & Wu, Q. *Parallel and Distributed Computing: Concepts and Technologies*. Taylor & Francis Group, 2019, 450 p.
35. Mattson, T.G., Sanders, B.A., & Massingill, B.L. *Patterns for parallel programming*. Addison-Wesley Professional, 2004, 368 p.