

С. С. ТОВКАЧ

Національний авіаційний університет «НАУ», Київ, Україна

## ПІДХІД НЕЧІТКОЇ КЛАСТЕРИЗАЦІЇ В РОЗПОДІЛЕНИХ ІНФОРМАЦІЙНИХ СИСТЕМАХ АВІАЦІЙНИХ ДВИГУНІВ

Стаття присвячена розробці системних підходів до побудови розподіленої інформаційної системи (РІС) авіаційних газотурбінних двигунів (ГТД). Визначено, що суттєвим для інтеграції авіадвигунобудівної галузі у світове суспільство розробників і виробників є використання CALS-технологій (Continuations Acquisition and Life-Cycle Support), які повинні забезпечити конкурентоспроможність продукції на світовому ринку. Актуальність використання CALS-технологій пов'язано з тим, що на сьогоднішній день у відповідності з вимогами ринку провідними фірмами світу встановлені терміни створення нової конструкції двигуна цивільної авіації п'ятого і шостого поколінь. Розглянуто блочно-модульний принцип побудови двигуна – математичні моделі і програмне забезпечення – із задоволенням критеріїв дивергенції, трансформації та конвергенції. Для спрощення пошуку оптимальної технології побудови розподіленої інформаційної системи авіаційного двигуна запропоновано використання підходу нечіткої кластеризації – спосіб проектування із знаходженням нового знання про ГТД із високоефективними показниками експлуатації. За рахунок виявлення методів аналізу знань та основних методів кластеризації – К-середніх, графові алгоритми кластеризації, алгоритми сімейства FOREL, ієрархічна кластеризація, нейронна мережа Кохонена, алгоритми сімейства KRAB, алгоритми нечітких середніх, субтрактивної кластеризації – визначено застосування нечіткої кластеризації в розподілених інформаційних системах авіаційних двигунів. Для зручної реалізації визначеного методу розглянуто множину даних об'єктів інформаційної системи ГТД, які містяться в експериментальних файлах. За результатами виконання процедури нечіткої кластеризації фіксуються координати центри класів, належність кожної сукупності даних до класів, значення цільової функції, які мають наближений характер і використовуються для попередньої структуризації даних. Після проведених досліджень визначено, що об'єднання в собі алгоритмів кластеризації повинно сприяти побудові більш точної моделі ГТД і підвищення її швидкодії.

**Ключові слова:** авіаційний двигун; CALS-технології; розподілена інформаційна система; системний підхід; нечітка кластеризація; методи оптимізації; візуалізація.

### Вступ

Створення сучасних газотурбінних двигунів (ГТД) передбачає ще на стадії проектування забезпечення суперечливих вимог економічності, ресурсу та надійності.

Процес створення і розвитку систем автоматизованого проектування (САПР) пройшов кілька етапів [1]:

1) розробку засобів, методів, моделей і алгоритмів, яка дозволила, з одного боку, знизити обсяг ручної праці інженера (підготовку даних і креслень, обробку результатів розрахунків і випробувань), а з іншого - здійснювати на електронно-обчислювальних машинах (ЕОМ) застосовувати в процесі створення ГТД інженерні методи і моделі розрахунку;

2) глибоку інтеграцію обчислювальної техніки в процес проектування, впровадження дво- і тривимірних моделей деталей і вузлів для оцінки різних процесів в ГТД і напружено-деформованого стану (НДС) окремих елементів;

3) впровадження засобів і методів багато дисциплінарного математичного моделювання, оптимізації конструкції і технологічних процесів, а також паралельного і розподіленого проектування.

При цьому намітилась тенденція (для досягнення конкурентоспроможності двигунів на ринку за рахунок застосування нових методів проектування і технологій) скоротити в найближчі роки не менше ніж на 50 % час на проектування і на 15...60 % виробничі витрати. Досягнення таких цілей можливо за рахунок революційних перетворень в інформаційних стратегіях при вдосконаленні діяльності в області різномірних процесів на всіх стадіях життєвого циклу двигуна (ЖЦД). При цьому повинні здійснюватися неперервне керування його параметрами (від ескізного проекту до виведення двигуна з експлуатації) та електронна інтеграція підприємств, які забезпечують процес керування. Одним з найважливіших результатів впровадження CALS-технологій повинно стати різке скорочення обсягів інформації.

## 1. Постановка задачі

Проблема роботи з великими потоками різно-рідної інформації є в даний час визначальною в розвитку систем, які супроводжують ЖЦД. В інформаційне поле необхідно включати не тільки інформацію, що знаходиться в базах даних, а й програмне забезпечення із обробки.

Програмне забезпечення в системах проектування засноване на великій кількості предметних моделей, які описують різні процеси, що відбуваються в двигуні, або з його деталями і вузлами. Всі програми і математичні моделі, які супроводжують ЖЦД, оперують різномірною інформацією із загального інформаційного поля (рис. 1). Дії зі створення і обробки цієї інформації повинні бути уніфіковані. Як показав досвід, сформульований в середині 70-х років І. А. Биргером загальні принципи автоматизованого проектування (модульності, ієрархії, адаптації, інформаційної єдності, ітерації), не тільки не втратили свого значення [1], але і з переходом на CALS-технології стали більш актуальними. Недооцінка значущості цих принципів призводить до зниження ефективності використання САПР.

Процес створення нової конструкції є багатостадійним і ітераційним. Багатостадійність полягає в тому, що на кожному кроці можливе повернення до попередньої стадії для визначення та уточнення параметрів проекту. Набагато це залежить від того, що на кожній стадії автоматизованого проектування

відбувається уточнення і ускладнення математичних моделей, які описують зміну газодинамічних параметрів двигуна, теплового стану, напруг і деформацій в деталях і вузлах його конструкції, аналізується можливість виготовлення конструктивних елементів на основі сучасних досягнень в області технологій виробництва. Методи і моделі випробувань і діагностики стану двигуна в експлуатації повинні закладатися на ранній стадії проектування двигуна для того, щоб забезпечити неперервний процес його супроводу протягом всього ЖЦД.

Вимоги забезпечення узгодженості параметрів проекту, одержуваних з предметних моделей для різних вузлів двигуна, призводять до ряду протиріч, вирішення яких можливе тільки при поверненні до попередніх етапів проектування для корекції параметрів, які є вихідними даними на наступних етапах роботи. Це визначає ітераційний принцип проектування.

Схема автоматизованого створення конструкції передбачає блочно-модульний принцип побудови. Причому, можливість використання модулів, що спираються на математичні моделі різного рівня складності і ієрархії [1, 2]. На основі загальних уявлень про процес проектування відбувається циклічне проходження стадії аналізу, синтезу та оцінки [2, 3]. Так, як ці стадії повторюються багаторазово, математичні моделі повинні задовольняти критеріям дивергенції, трансформації та конвергенції.

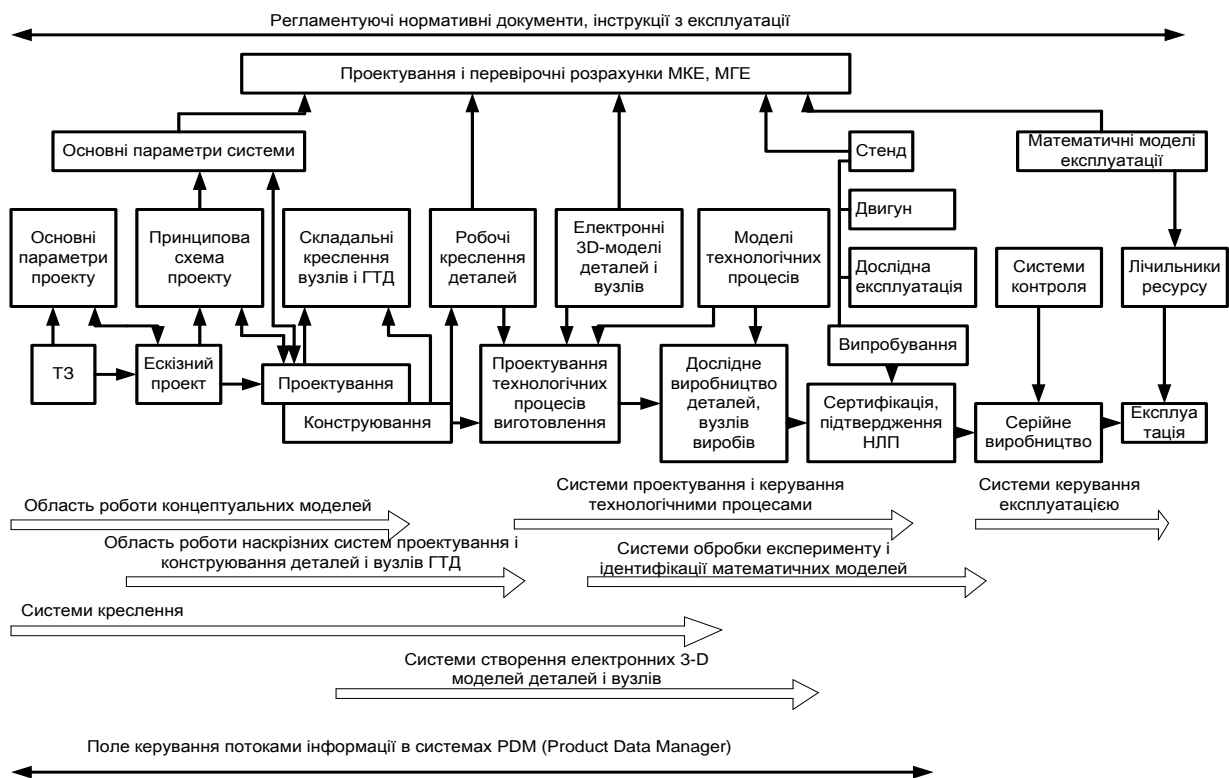


Рис. 1. Математичні моделі і програмне забезпечення, які супроводжують життєвий цикл ГТД

У наукових дослідженнях і практичних розробках [2 - 4] з метою поєднання кращих властивостей просторових методів блочно-модульної побудови для розрахунку розподіленої інформаційної двигуна використовуються CALS-технології та різні методи і моделі керування процесом проектування із застосуванням алгоритмів оптимізації для визначених даних інформаційного середовища ГТД.

У зв'язку з цим актуальною є задача формування оптимальної технології побудови розподіленої інформаційної системи авіаційного двигуна, вузли якої зможуть одночасно функціонувати в межах робочого простору, – так званого, підходу нечіткої кластеризації щодо способу проектування із знаходженням нового знання про ГТД із високоефективними показниками експлуатації.

## 2. Підхід нечіткої кластеризації в інформаційних системах ГТД

Розглянемо основні види моделей, що використовуються для знаходження нового знання на основі даних інформаційного сховища ГТД та задовольняють критеріям дивергенції, трансформації та конвергенції. Метою інтелектуальних технологій є знаходження нового знання, яке користувач може надалі застосувати для поліпшення результатів своєї діяльності. Результат моделювання - виявлені відношення в даних.

Можна виділити, принаймні, шість методів виявлення й аналізу знань:

- 1) класифікація,
- 2) регресія,
- 3) прогнозування часових послідовностей (рядів),
- 4) кластеризація,
- 5) асоціація,
- 6) послідовність.

Перші три використовуються головним чином для передбачення, у той час як останні зручні для опису існуючих закономірностей в даних [3].

**Кластеризація** - розділення елементів на групи за певними ознаками. Відрізняється від класифікації тим, що самі групи заздалегідь не задані. За допомогою моделі кластеризації засоби інтелектуальних обчислень самостійно виділяють різні однорідні групи даних.

При **чіткій** кластеризації об'єкт може належати лише одному кластеру. При **нечіткій** кластеризації об'єкт може належати до декількох кластерів, при цьому зазначається ступінь його приналежності для кожного з них, який характеризує ступінь зв'язку об'єкта з певним кластером.

Ієрархічні методи кластеризації будують ієрархічну (деревовидну) структуру. Неієрархічні просто розділяють елементи на групи.

Існують такі методи кластеризації:

– К-середніх (K-means). Є методом чіткої кластеризації. Будується деяка цільова функція  $\Phi(\circ)$ , що виражає якість поточного розбиття зображення на  $k$  кластерів із центрами у точках  $C_i$ ,  $i = 1, \dots, k$ ;  $k$  – задано. Вибравши в початковий момент центри кластерів довільним чином, далі для кожного елемента ітеративно визначаємо його приналежність до одного із  $k$  кластерів і обчислюємо нові значення для центрів кластерів, намагаючись при цьому мінімізувати функцію  $\Phi(\circ)$ ;

- графові алгоритми кластеризації;
- статистичні алгоритми кластеризації;
- алгоритми сімейства FOREL;
- ієрархічна кластеризація або таксономія;
- нейронна мережа Кохонена;
- ансамбль кластеризаторів;
- алгоритми сімейства KRAB;
- EM-алгоритм;
- алгоритм, заснований на методі просівання.

В Matlab є такі основні методи нечіткої кластеризації:

- FCM (Fuzzy C-Means). Алгоритм нечітких середніх;
- субтрактивна кластеризація.

## 3. Метод нечіткої кластеризації для побудови інформаційних систем ГТД

Для більш зручної реалізації нечіткої кластеризації розглянемо множину даних розмірності  $140 \times 2$ , які містяться у файлі *fcmdata.dat* або *gteinformation.dat*. Матриця  $140 \times 2$  відповідає 140 об'єктам інформаційної системи ГТД, для кожного з яких виконане вимірювання за двома ознаками.

Процедуру нечіткої кластеризації (рис. 2) можна записати у вигляді командного M-файла або шляхом використання графічного інтерфейсу, який викликається за допомогою команди *findcluster*: Методи: алгоритм нечітких середніх та субтрактивна кластеризація.

За результатами виконання нечіткої кластеризації можна визначити координати центри класів (*center*), належність кожної сукупності даних до класів (*matrix U*), значення цільової функції (*obj\_fcn*) Отримана інформація має наближений характер і використовується для попередньої структуризації системи ГТД.

## Висновок

Авіаційний ГТД являється добре структурованою конструкцією, тому його розробка на етапах трансформації і конвергенції передбачає паралельну

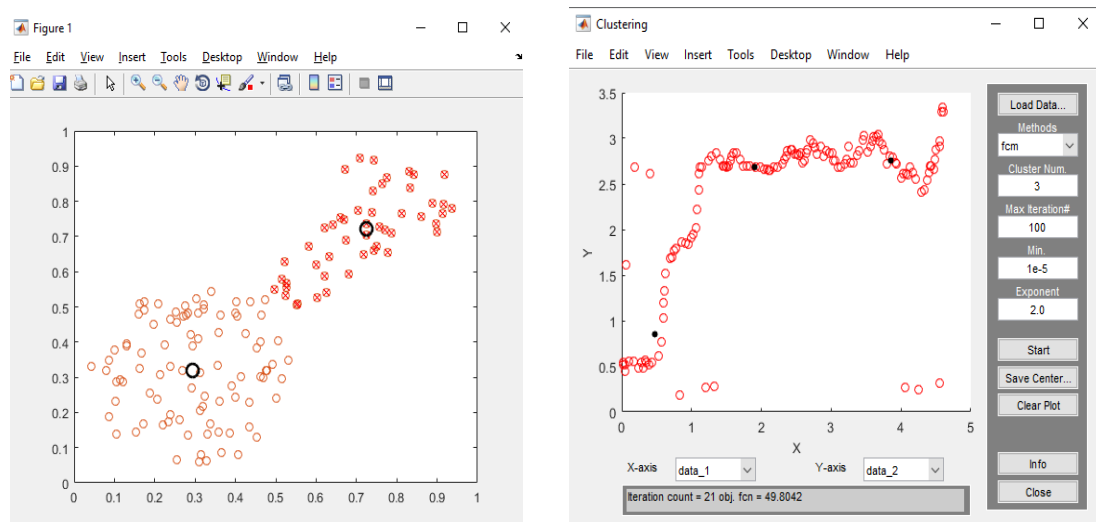


Рис. 2. Застосування нечіткої кластеризації для об'єктів інформаційної системи ГТД: візуалізація за допомогою командного рядка (зліва) та графічний інтерфейс візуалізації (справа)

розробку вузлів та деталей. Основний об'єм інформації, який є базою для інформаційного поля, створюється системами геометричного моделювання, обчислювальними комплексами математичного моделювання і технологічними системами, які формуються у підході до побудови систем ГТД.

В процесі дослідження підходу нечіткої кластеризації в інформаційних системах ГТД визначено напрямок знаходження знань (Data Mining), який є невід'ємною частиною концепції електронних сховищ даних та організації інтелектуальних обчислень.

Об'єднання в собі алгоритмів нейронних мереж і технології дерев рішень повинно сприяти побудові більш точної моделі і підвищенню її швидкодії. Програми візуалізації даних у певному сенсі не є засобом аналізу інформації, оскільки вони тільки представляють її користувачу. Проте, візуальне представлення, скажемо, відразу чотирьох змінних досить виразно узагальнює надвеликі обсяги даних. Отже, виробники розуміють, що для рішення кожної проблеми варто застосовувати оптимальний підхід.

## Література

1. ЦИАМ – Центральный институт авиационного моторостроения имени И. П. Баранова [Электронный доступ]. – Режим доступа: <http://www.ciam.ru/> – 10.06.2020.
2. Malhotra, V. K. An Analysis of Fuzzy Clustering [Text] / V. K. Malhotra, K. Harleen, M. Afshar Alam

// International Journal of Computer Application. – 2014. – No. 19, iss. 94. – P. 9-12. DOI:10.5120/16497-6578.

3. Lorenzon, F. Parallel Computing Hits the Power Wall: principles, challenges, and survey of solutions [Text] / F. Lorenzon, B. Filho, A. Carlos. – Springer, Cham, 2019. – 88 p. ISBN: 978-3-030-28719-1.

4. Товкач, С. С. Гібридний генетичний підхід до побудови розподіленої системи автоматичного керування авіаційних двигунів [Текст] / С. С. Товкач // Вісник двигунобудування. – 2019. – № 2. – С. 57-62.

## References

1. CIAM – Central'nyj institut aviacionnogo motorostroeniya imeni I. P. Baranova [Central Institute of Aviation Motors named after I. P. Baranov]. Available at: <http://www.ciam.ru/> (accessed 10.06.2020).
2. Malhotra, V. K., Kaur, H., Alam, M. A. An Analysis of Fuzzy Clustering. *International Journal of Computer Applications*, 2014, no. 19, iss. 94, pp. 9-12. DOI:10.5120/16497-6578.
3. Lorenzon, F., Filho, B., Carlos, A. *Parallel Computing Hits the Power Wall: principles, challenges, and survey of solutions*, Springer, Cham, 2019. 88 p. ISBN: 978-3-030-28719-1.
4. Tovkach, S. S. Gibrydnyj genetychnyj pidhid do pobudovy rozpodilenoj systemy avtomatychnoho keruvannya aviacijnyx dyguniv [Hybrid genetic approach for building the distributed automatic control system of aviation engines]. *Visnyk dvygunobuduvannya – Herald of Aeroenginebuilding*, 2019, no. 2, pp. 57-62.

Надійшла до редакції 10.06.2020, розглянута на редколегії 15.08.2020

## ПОДХОД НЕЧЕТКОЙ КЛАСТЕРИЗАЦИИ В РАСПРЕДЕЛЕННЫХ ИНФОРМАЦИОННЫХ СИСТЕМАХ АВИАЦИОННЫХ ДВИГАТЕЛЕЙ

*С. С. Товкач*

Статья посвящена разработке системных подходов к построению распределенной информационной системы (РИС) авиационных газотурбинных двигателей (ГТД). Определено, что существенным для интеграции авиадвигателестроительной отрасли в мировое сообщество разработчиков и производителей является использование CALS-технологий (Continuons Acquisition and Life-Cycle Support), которые должны обеспечить конкурентоспособность продукции на мировом рынке. Актуальность использования CALS-технологий связано с тем, что на сегодняшний день в соответствии с требованиями рынка ведущими фирмами мира установлены сроки создания новой конструкции двигателя гражданской авиации пятого и шестого поколений. Рассмотрены блочно-модульный принцип построения двигателя - математические модели и программное обеспечение - с удовлетворением критериев дивергенции, трансформации и конвергенции. Для упрощения поиска оптимальной технологии построения распределенной информационной системы авиационного двигателя предложено использование подхода нечеткой кластеризации - способ проектирования с нахождением нового знания о ГТД с высокоэффективными показателями эксплуатации. За счет выявления методов анализа знаний и основных методов кластеризации – K-средних, графовые алгоритмы кластеризации, алгоритмы семейства FOREL, иерархическая кластеризация, нейронная сеть Кохонена, алгоритмы семейства KRAB, алгоритмы нечетких средних, субтрактивная кластеризация – определены применения нечеткой кластеризации в распределенных информационных системах авиационных двигателей. Для удобной реализации определенного метода рассмотрено множество данных объектов информационной системы ГТД, содержащиеся в экспериментальных файлах. По результатам выполнения процедуры нечеткой кластеризации фиксируются координаты центры классов, принадлежность каждой совокупности данных в классы, значение целевой функции, которые имеют приближенный характер и используются для предварительной структуризации данных. После проведенных исследований определено, что объединение в себе алгоритмов кластеризации должно способствовать построению более точной модели ГТД и повышению ее быстродействия.

**Ключевые слова:** авиационный двигатель; CALS-технологии; распределенная информационная система; системный подход; нечеткая кластеризация; методы оптимизации; визуализация.

## APPROACH OF FUZZY CLUSTERING IN DISTRIBUTED INFORMATION SYSTEMS OF AVIATION ENGINES

*S. Tovkach*

The article is devoted to the development of systematic approaches for the construction of a distributed information system (DIS) of aviation gas turbine engines (GTE). It is determined that the use of CALS-technologies (Continuous Acquisition and Life-Cycle Support), which should ensure the competitiveness of products on the world market, is essential for the integration of the aircraft engine industry into the world community of developers and manufacturers. The relevance of the use of CALS-technologies is due to the fact that today, in accordance with market requirements, the world's leading companies have set deadlines for the creation of a new design of the civil aviation engine of the fifth and sixth generations. The block-modular principle of engine construction - mathematical models and software - with the satisfaction of the criteria of divergence, transformation, and convergence has been considered. For a simplified search of the optimal technology for building a distributed information system of an aviation engine, the use of a fuzzy clustering approach is proposed, which is a design method with finding new knowledge about the gas turbine engine with highly efficient performance. By identifying methods of knowledge analysis and basic methods of clustering, that are K-means, graph clustering algorithms, algorithms of the FOREL family, hierarchical clustering, Kohonen neural network, algorithms of the KRAB family, fuzzy mean algorithms, subtractive, the application clustering in distributed information systems of aviation engines have been determined. For the convenient implementation of the defined method, a set of data objects of the GTE information system, which are contained in the experimental files, are considered. According to the results of the fuzzy clustering procedure, the coordinates of the class centers, the belonging of each data set to the classes, the values of the objective function, which have an approximate character, and are used for preliminary structuring of the data, are fixed. After research, it was determined that the integration of clustering algorithms should help build a more accurate model of the gas turbine engine and increase its speed.

**Keywords:** aviation engine; CALS-technologies; distributed information system; system approach; fuzzy clustering; optimization methods; visualization.

**Товкач Сергій Сергійович** – канд. техн. наук, доцент, доцент кафедри автоматизації та енергоменеджменту, Аерокосмічний факультет Національного авіаційного університету «НАУ», Київ, Україна.

**Serhii Tovkach** – Candidate of Technical Science, Associate Professor, Associate Professor of Automation & Power Management Department, Aerospace faculty, National Aviation University “NAU”, Kyiv, Ukraine, e-mail: serhii.tovkach@nau.edu.ua, serhii.tovkach@npp.nau.edu.ua, ORCID Author ID: 0000-0002-8740-298X, Scopus Author ID: 57206192351, ResearcherID: Q-2695-2019, ResearchgateID: Serhii\_Tovkach, <https://scholar.google.com.ua/citations?user=gJqZwiUAAAJ&hl=ru>.