

УДК 629.735.083.03.004.58:004.801.3(045)

О.С. Якушенко, канд. техн. наук, старш. наук. співроб.

ДІАГНОСТУВАННЯ ТЕХНІЧНОГО СТАНУ ГАЗОТУРБІННИХ ДВИГУНІВ МЕТОДОМ ПОБУДОВИ БІНАРНОГО ДЕРЕВА

Розглянуто проблему діагностування технічного стану газотурбінних двигунів з використанням методу побудови бінарного дерева. Наведено вихідні дані та результати діагностування при використанні зазначеного методу.

In the paper the problem of binary trees using for identification of aircraft gas turbine engines technical condition is considered. The results of the method using for aircraft gas turbine engines technical condition identification are given.

Постановка проблеми

Натепер розроблено і впроваджено в експлуатаційні підрозділи цивільної авіації ряд систем діагностування технічного стану (ТС) авіаційної техніки. Ці системи різняться ступенем автоматизації, повнотою діагностування і використовуваними у процесі їх реалізації технічними засобами.

Результатом діагностування газотурбінних двигунів (ГТД) є встановлення первинного діагнозу щодо їх ТС. Щоб підвищити якість й оперативність розпізнавання ТС ГТД, треба виконувати автоматизований аналіз польотної інформації, результатом якого є класифікація ТС ГТД – віднесення двигуна до одного з можливих технічних станів.

Аналіз досліджень і публікацій

Останнім часом найперспективнішими напрямами створення комп'ютерних систем автоматизованого розпізнавання ТС ГТД є такі, що базуються на використанні методів математичної статистики, штучного інтелекту та методів розпізнавання образів. Ця робота є продовженням серії робіт, присвячених дослідженню цих методів. У роботах [1; 2] було наведено результати використання методів штучного інтелекту під час діагностування ТС ГТД. У цій роботі розглядається питання діагностування ГТД з використанням методів розпізнавання образів, а саме: діагностування методом побудови бінарного дерева.

Мета

Основні проблеми, що виникають під час використання методу бінарного дерева, такі:

- формування базового набору даних, який має досить повно описувати принаймні основні класи ТС об'єкта та дозволяти коректно оцінювати відстані між ними;
- вибір відстані між об'єктами, від правильності вибору якого значною мірою залежить якість розпізнавання.

Математичне формулювання проблеми

Розпізнавання образів у широкому значенні полягає в автоматичному розпізнаванні різних систематичних повторювань. Досліджуваний методом розпізнавання образів об'єкт має бути віднесений до визначеного класу. При цьому під класом слід розуміти групу об'єктів, що характеризуються набором загальних властивостей. Прикладами технічних класів можуть бути, наприклад, класи «Справний ГТД» – «Несправний ГТД» або «ГТД з деградованою проточною частиною», «ГТД з руйнуванням підшипника» тощо.

Класифікація включає всі процеси, що закінчуються вказівкою на приналежність об'єкта до конкретного класу, тобто включає прості алгоритми, в основу яких покладено використання результатів деяких вимірюваного і граничного значень, наприклад визначення того, чи знаходиться значення вібрації у припустимих межах.

При цьому слід враховувати вимоги сучасної концепції «сплавлення» даних, яка передбачає одночасне використання під час діагностування різних джерел інформації. Такий підхід до діагностування вимагає розширеної номенклатури діагностичних параметрів.

Існуючі методи класифікації можна розділити на методи розпізнавання «з учителем» і «без учителя» [3; 4].

Для розпізнавання «з учителем» треба мати класифіковані вибірки, кожна з яких складається з параметрів функціонування різних ГТД, що належать до наперед відомого класу ТС. Найпростішим випадком є формування двох вибірок. У першу з них входять параметри функціонування тільки справних ГТД, у другу – тільки несправних. Отже, для кожного ГТД у вибірці має бути визначено «вчителем» (дослідником), до якого класу цей ГТД відноситься.

Потреба у використанні методу «без учителя» виникає, якщо для навчальних даних заздалегідь невідомо, до якого класу вони належать (навчання «без учителя»).

Іноді навіть невідомо, скільки та які класи можуть бути сформовані. Тоді необхідно проводити самонавчання системи. Саме до таких методів належить метод класифікації побудовою бінарного дерева.

Цей метод є способом поділяти набір об'єктів у групи чи кластери таким чином, щоб властивості об'єктів в одному кластері були дуже схожі, а в різних кластерах дуже відмінними.

Основна процедура такої класифікації полягає у такому [5; 6]:

- знайти подобу або відмінність між кожною парою об'єктів у наборі даних, обчислюючи відстані між всіма об'єктами у навчальній вибірці;
- згрупувати об'єкти в бінарне ієрархічне кластерне дерево, зв'язуючи разом пари об'єктів, що знаходяться в найбільшій близькості (об'єкти з'єднуються в подвійні кластери, сформовані кластери групуються у великі кластери, доки не буде сформоване єдине ієрархічне дерево, яке включає всі об'єкти);
- визначити, де розділити ієрархічне дерево на кластери.

Приклад, що ілюструє викладений підхід показано на рис. 1. Лінії на рис.1 зв'язують найближчі об'єкти між собою. На рис.1, б показано отриману дендрограму (бінарне дерево). На ній по горизонтальній осі проставлено номери точок у першому наборі даних. Зв'язки ж між об'єктами (рис.1, б) подано як інвертовані вниз U -подібні лінії. Висота U вказує евклідову відстань між об'єктами.

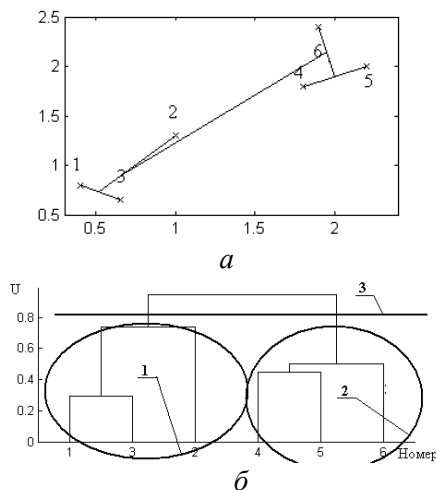


Рис. 1. Кластеризація двовимірного набору даних методом побудови бінарного дерева:

a – вихідний набір даних та найкоротші відстані між об'єктами;

1–6 номери точок;

б – дендрограма, розділена на два кластери:

1, 2 – номери кластерів;

3 – відокремлювальна лінія

Після того, як ієрархічне дерево подвійних кластерів створено, потрібно розділити ієрархію на великі кластери. При цьому слід визначити кількість класів або максимальну відстань між об'єктами в одному кластері. Якщо в розглянутому прикладі як максимальну відстань між об'єктами в одному кластері вибрати відстань 0,8, то отримуємо два кластери, як показано на рис.1, б. На заключному етапі дослідження треба перевірити, що отримані кластери поєднують об'єкти, які мають істотну подібність. Для проведення такого аналізу для ГТД слід отримати додаткові дані про реальний технічний стан всіх об'єктів, поданих у навчальній вибірці.

Під час обробки нових параметрів функціонування ГТД проводять повторну побудову бінарного дерева з урахуванням нових даних. Діагноз установлюють залежно від того, до якого кластера найбільш близькі дані діагностованого ГТД.

Одним з основних питань використання цього методу є вибір відстані між об'єктами. Найчастіше використовують:

– евклідову відстань:

$$\Delta_{rs}^2 = (x_r - x_s)(x_r - x_s)',$$

де x_r, x_s – вектори координат об'єктів r та s ;

– нормовану евклідову відстань:

$$\Delta_{rs}^2 = (x_r - x_s)D^{-1}(x_r - x_s)',$$

де D – діагональна матриця, складена з дисперсій відповідних компонентів координат об'єктів, обчислених за всіма об'єктами навчальної вибірки;

– відстань Махаланобіса:

$$\Delta_{rs}^2 = (x_r - x_s)V^{-1}(x_r - x_s)',$$

де V – проста коваріаційна матриця;

– відстань сумарного абсолютного відхилення:

$$\Delta_{rs} = \sum_{j=1}^n |x_{rj} - x_{sj}|;$$

– відстань Мінковські (Minkowski):

$$\Delta_{rs} = \left\{ \sum_{j=1}^n |x_{rj} - x_{sj}|^p \right\}^{1/p},$$

де p – параметр у діапазоні від 1 до 2.

Вибір відстані залежить від особливостей набору даних. Найдоцільніше обирати відстань, яка забезпечує найкраще розділення класів. Як критерій вибору може бути прийнятий відсоток помилково кластеризованих даних.

Іншою проблемою створення розглянутого методу є отримання навчальної вибірки. Для досить повного опису одного класу ТС на одному режимі роботи ГТД необхідно мати 20–200 і більше різних прикладів цього ТС. Кожен такий приклад має включати діагностичну інформацію

для однієї з можливих комбінацій функціональних характеристик вузлів ГТД. Отримання такої інформації в експлуатації, враховуючи низьку частоту виникнення дефектів та відмов і їх комбінацій – процес досить тривалий. Використання результатів стендових випробувань доцільно, враховуючи високу собівартість, тільки на заключному етапі створення системи класифікації для перевірки її працездатності. Єдиним реальним джерелом основного обсягу інформації можуть бути тільки результати проведення числового експерименту з використанням математичної моделі робочого процесу ГТД.

Схему діагностування з використанням розглянутого методу визначення ТС показано на рис. 2.



Рис. 2. Схеми діагностування

Розглядаючи цю схему слід зупинитися на переході від параметрів функціонування до ознак стану (параметрів, значення яких залежить, насамперед, від ТС об'єкта). Реєстровані параметри функціонування залежать не тільки, а часто, і не стільки, від ТС, а передусім від режиму та зовнішніх умов експлуатації ГТД. Тому до алгоритму діагностування потрібно включити відомі методи зменшення мінливості прояву несправності, пов'язані, наприклад, зі зведенням параметрів робочого процесу до заданих атмосферних умов, обчисленням відхилень реальних параметрів від еталонних.

Але аналіз методу діагностування будовою бінарного дерева вказує, крім загальновідомих особливостей проведення розрахунку ознак ТС, на деякі додаткові особливості процесу переходу від параметрів функціонування до ознак стану.

Під час розробки методу розпізнавання ТС ГТД треба врахувати той факт, що кластеризація виконується за обраною відстанню між окремими точками n -вимірному простору. При цьому значення кожної ознаки стану в n -вимірному просторі ознак змінюється у своїх межах і залежить від одиниць виміру параметрів та методу розрахунку. При цьому діапазони зміни ознак стану істотно різняться. На формування відстані між точками буде найсуттєвіше впливати ознака стану, що має найбільше значення різниці між максимальним Δ_{\max} та мінімальним Δ_{\min} значенням:

$$(\Delta_{\max} - \Delta_{\min})_{\max}$$

Ознака, що має мінімальне значення цього параметра, майже не буде брати участі у процесі

формування кластерів. Для забезпечення однакової участі всіх ознак у процесі кластеризації їх значення мають бути помножені на індивідуальні коефіцієнти:

$$k_i = \frac{(\Delta_{\max} - \Delta_{\min})_{\max}}{\Delta_{\max i} - \Delta_{\min i}}, \quad (1)$$

де i – номер ознаки стану.

Після такого нормування значення промасштабованих ознак стану може бути істотно менше або більше одиниці. Для запобігання втрати значимих чисел або отримання надвеликих/надмалих проміжних значень під час розрахунку може виникнути потреба в уведенні загального для всіх ознак коефіцієнта K . Тоді значення i -ї ознаки стану у j -му прикладі, яке буде використане при кластеризації, можна розрахувати як

$$\Delta'_{i,j} = K k_i \Delta_{i,j}, \quad (2)$$

де $\Delta_{i,j}$ – вхідне значення ознаки.

Числовий експеримент

Під час проведення числового експерименту було використано дані, в яких подано шість різних технічних станів авіаційного ГТД ПС-90А. Докладнішу характеристику методу отримання цих даних викладено в роботі [2]. Значення Δ_{\max} , Δ_{\min} та $\Delta_{\max} - \Delta_{\min}$ для кожного з параметрів, які входять до навчальної вибірки, наведено у таблиці. Значення параметрів, які входять до навчальної вибірки, було промасштабовано за залежностями (1), (2). Для параметра K було взято значення $K = 1000$.

Діапазони зміни вихідних ознак ТС у навчальному наборі даних

Ознака Δ	Δ_{\min}	Δ_{\max}	$\Delta_{\max} - \Delta_{\min}$
За частотою обертання ротора високого тиску n_{BT}	-0,0050	0,0207	0,0258
За повним тиском у зовнішньому контурі P_B^*	-0,0098	0,0360	0,0458
За повним тиском за компресором P_K^*	-0,0528	0,0929	0,1457
За повною температурою за компресором T_K^*	-0,0401	0,0187	0,0588
За повною температурою за турбіною T_T^*	-0,1441	-0,0125	0,1316
За відносним тиском за турбіною P_T^*	-0,0123	0,0363	0,0486
За витратою палива G_{II}	0,1894	0,0118	0,2012
$(\Delta_{\max} - \Delta_{\min})_{\max}$	–	–	0,2012

Отримані у такий спосіб значення $\Delta'_{i,j}$ показано на рис. 3.

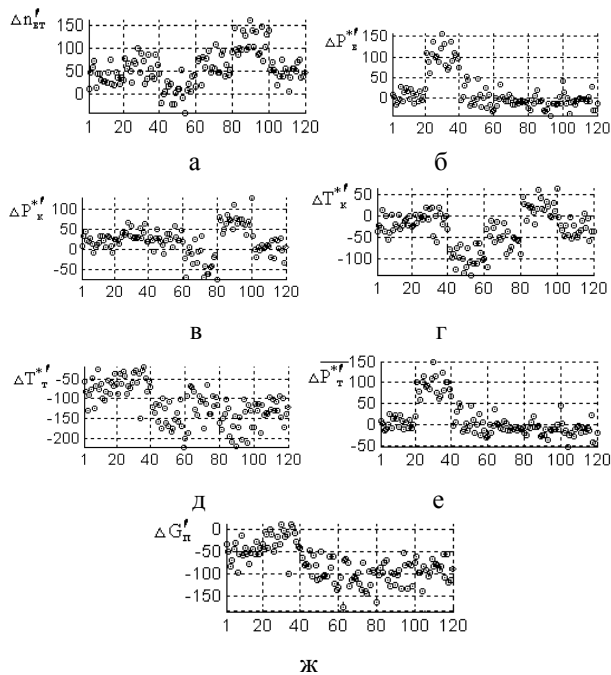


Рис. 3. Промасштабовані ознаки технічного стану:
 а – за частотою обертання ротора високого тиску ГТД;
 б – за повним тиском за вентилятором;
 в, г – за повним тиском та температурою за компресором;
 д – за повною температурою за турбіною;
 е – за витратою палива;
 ж – за відношенням повного тиску за турбіною до тиску на вході до ГТД

Після масштабування всі ознаки стану змінюються у межах 150 одиниць.

Було проведено спробу розпізнавання отриманого навчального набору даних з використанням методу бінарного дерева. Як одиницю виміру відстані між об'єктами було використано евклідову відстань. Отримане в результаті розрахунків бінарне дерево зображено на рис. 4. Розділення цього дерева на шість класів показало близьку до нульової ефективність розглянутого методу. При цьому майже всі точки було віднесено до одного класу і тільки вісім точок до п'ятих інших класів.

Причиною цього стало те, що у вхідній вибірці класи ТС ГТД не строго розділені, а, навпаки, мають ділянки перетинання, коли класи, які межують, накладаються один на одного.

Спроба зміни відстані не дала суттєвого поліпшення результату класифікації.

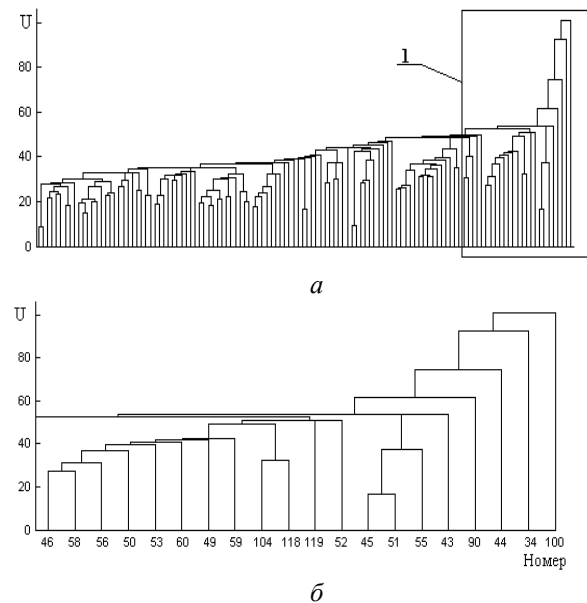


Рис. 4. Дендрограма, отримана в результаті кластеризації:
 а – загальна дендрограма;
 б – збільшений вигляд ділянки 1 загальної дендрограми

Висновки

Головним недоліком описаного методу кластеризації є те, що він не працює у разі перекриття класів хоча б в 1–2 точках і може використовуватися тільки за умови чіткого розподілу класів. Тому для отримання більш стійких та надійних методів класифікації ТС ГТД потрібно перейти до методів, більш стійких щодо наявності перекриття класів та впливу стохастичного характеру даних, які складають навчальну множину. Це, передусім, створення різних модифікацій методу кластеризації за найближчим сусідом, методи дискримінантного аналізу та мінімізації ризику.

Література

1. Якушенко О.С. Нейронні мережі для діагностування ГТД // Вісн. НАУ. – 2004. – № 2 (20). – С. 67–71.
2. Кучер О.Г., Якушенко О.С., Сухоруков В.Ю. Розпізнавання технічного стану авіаційних ГТД з використанням нейронних мереж // Вісн. двигунобудування. – 2002. – № 1. – С. 101–106.
3. Патрик Э. Основы теории распознавания образов: Пер. с англ. / Под ред. Б. Левина. – М.: Сов. радио, 1980. – 408 с.
4. Медведев В.С., Потемкин В.Г. Нейронные сети. Matlab 6 / Под общ. ред. В.Г. Потемкина. – М.: Диалог-МИФИ, 2002. – 496 с.
5. Потемкин В.Г. Система MatLab 5 для студентов. – М.: Диалог-МИФИ, 1998. – 314 с.
6. Потемкин В.Г. Система MatLab: Справочное пособие. – М.: Диалог-МИФИ, 1998. – 350 с.