

ІНТЕГРОВАНА ІНСТРУМЕНТАЛЬНА СИСТЕМА ДЛЯ РІШЕННЯ ЗАДАЧ ДІАГНОСТИЧНОГО КОНТРОЛЮ

Канд. техн. наук В.Д. Павленко, О.О. Фомін, В.О. Шубаєв

Одеський державний політехнічний університет

65044 пр. Шевченка 1, Одеса, Україна

тел. (0482) 288-170, e-mail: pavlenko_vitalij@mail.ru

Розглядається комплексний підхід до проблеми діагностичного контролю, що включає спільне використання програмних засобів і алгоритмів статистичної класифікації. Аналізується склад алгоритмів для рішення широкого кола задач діагностичного контролю: навчання, класифікації, дослідження інформативності діагностичних ознак об'єкта.

Пропонується структура організації алгоритмічної і інтерфейсної частин, інтегрованих в інструментальну систему для рішення задач діагностичного контролю.

ВСТУП

Постійне ускладнення технічних систем ставить нові проблеми забезпечення їх якості і надійності. У системах управління якістю при виробництві складних систем і їх сертифікації інформаційні етапи процесу управління об'єднуються операцією діагностичного контролю, задача якої - отримання інформації про фактичний стан об'єкта контролю і прийняття на цій підставі рішення про віднесення його до тієї або іншої із заданих категорій якості. Впровадження прогресивних інформаційних технологій при автоматизації операцій контролю в системах управлінні якістю і сертифікація складних систем дозволяє використати для рішення задач

діагностичного контролю ефективні процедури інтелектуальної обробки інформації, засновані на методах статистичної класифікації (розпізнавання образів) [1].

У роботі пропонується комплексний підхід до проблеми підвищення ефективності систем управління якістю і сертифікації, заснований на розвитку як алгоритмічних, так і програмних засобів. Тут розглядається інтегрована інструментальна система (ІС) для рішення задач діагностичного контролю, що поєднує в собі нові можливості організації інтерфейсу користувача на підставі прогресивних інформаційних технологій, а також ефективні статистичні алгоритми навчання, класифікації і дослідження діагностичної цінності ознак при багатокласовому розпізнаванні на підставі даних, поданих як звичайними вибірками, так і вибірками з нечітким описом класів.

1. АЛГОРИТМІЧНІ ЗАСОБИ ІС

Розроблена ІС містить декілька груп алгоритмів, для багатостороннього дослідження статистичних даних в задачах управління якістю і сертифікації. У програмі реалізовані алгоритми попередньої обробки вибірки даних: заповнення прогалів в даних статистичними методами [2]. Ця проблема набуває особливе значення при роботі з реальними даними, коли внаслідок цілого ряду

об'єктивних і суб'єктивних причин в описі деяких об'єктів статистичної вибірки відсутні окремі компоненти (діагностичні ознаки), а збір нової інформації про ці не повністю описані об'єкти часто зв'язаний з додатковими витратами ресурсів (часових, матеріальних і т.д.) або ж взагалі не є можливим.

Друга група алгоритмів призначена для побудови вирішальних правил статистичної класифікації. В якості вирішальних правил статистичної класифікації використовуються поліноми першого, другого і третього ступеня. Для побудови вирішальних правил статистичної класифікації використовуються параметричні і непараметричні методи навчання розпізнаванню образів: метод статистичних рішень і метод стохастичної апроксимації [1, 2]. Розвиток цих методів дозволив отримати більш ефективні процедури побудови вирішального правила на підставі методу статистичних рішень і оптимізації вибору порога розпізнавання:

$$d(x) = -\frac{1}{2}(x - m_1)^T S_1^{-1}(x - m_1) + \frac{1}{2}(x - m_2)^T S_2^{-1}(x - m_2) + \frac{1}{2} \ln \frac{|S_2|}{|S_1|} + \lambda \quad (1)$$

де x - вектор параметрів стану об'єкта розмірністю n ; m_i - вектор середніх значень вимірювань параметрів стану об'єктів для i -го класу ($i=1,2$); S_i - коваріаційна матриця параметрів стану, що вимірюються для i -го класу; $|S_i|$ - визначник матриці S_i ; S_i^{-1} - матриця, зворотна матриці S_i , λ - поріг розпізнавання.

Другий метод для ефективного пошуку вирішального правила полягає у використанні в якості початкових значень коефіцієнтів квадратичного вирішального правила в алгоритмі методу стохастичної апроксимації значень, попередньо отриманих на етапі навчання за допомогою методу статистичних рішень, що дозволяє значно скоротити кількість ітерацій в рекурентній процедурі методу стохастичної

апроксимації і дає можливість отримати вирішальне правило з більш високим процентом правильного розпізнавання. При цьому перехід між розглянутими методами здійснюється по формулах:

$$A = \frac{1}{2}(S_2^{-1} - S_1^{-1})$$

$$b = S_1^{-1}m_1 - S_2^{-1}m_2 \quad (2)$$

$$c = -\frac{1}{2}(m_1^T S_1^{-1} m_1 - m_2^T S_2^{-1} m_2) + \frac{1}{2} \ln \frac{|S_2|}{|S_1|}$$

де A - матриця коефіцієнтів при квадратичних членах вирішального полінома; b - вектор коефіцієнтів при лінійних членах вирішального полінома; c - вільний коефіцієнт.

Важливим етапом всякого процесу класифікації є вибір набору діагностичних параметрів. При цьому звичайно прагнуть обмежитися як можна меншою їх кількістю. Скорочення опису об'єктів спрощує збір додаткових даних, робить матеріал більш оглядним, а пристрій, що розпізнає - більш простим і швидкодіючим. Зменшення параметрів опису об'єктів, як правило, поліпшує якість прогнозу [3]. Крім того, скорочення числа діагностичних ознак дозволяє знизити мінімально необхідну кількість об'єктів статистичної вибірки, зберігаючи її репрезентативність. Тому після збору статистичних даних необхідно розглянути можливість виключення частини ознак так, щоб помилки розпізнавання були менш допустимих.

Очевидно, що кількість ознак, необхідна для успішного рішення деякої задачі розпізнавання, залежить від розділяючих якостей вибраних ознак. Задача вибору ознак ускладнюється звичайно тією обставиною, що найбільш важливі ознаки не завжди легко виміряти або, як виявляється в багатьох випадках, відповідні можливості вимірювання стримуються економічними факторами. У ІС включена група алгоритмів, спрямованих на дослідження інформативності параметрів статистичної вибірки і відбір найбільш

інформативних наборів з них. Ці алгоритми представляють різні методи перебору параметрів статистичної вибірки: повний, скорочений і випадковий перебір (метод Монте-Карло) [3].

При роботі з багатокласовими статистичними вибірками для пошуку найкращого рішення в задачах дослідження інформативності пропонується використати критерії $F(P)$, що враховують різницю в якості розпізнавання різних класів на кожному наборі діагностичних параметрів:

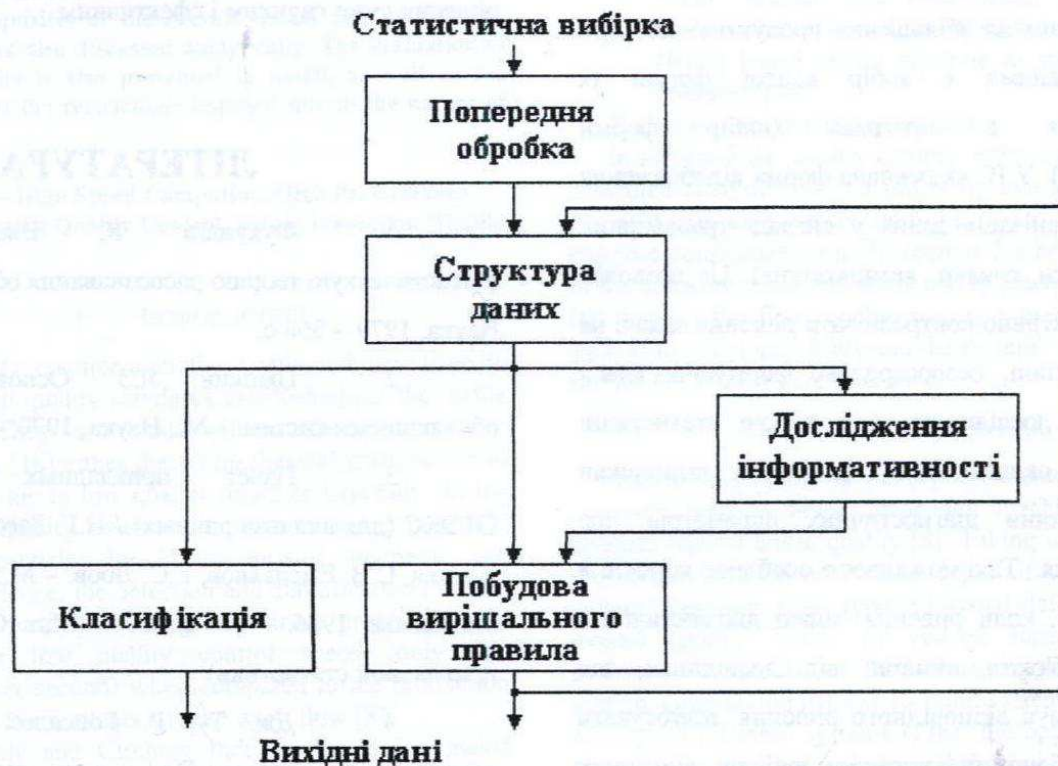
$$F(P) = \frac{\sum_{i=1}^m P_i}{\alpha \cdot (P_{max} - P_{min}) + 1} \quad (3)$$

$$F(P) = \frac{\sum_{i=1}^m P_i}{e^{\beta \cdot (P_{max} - P_{min})}} \quad (4)$$

де P_i - відсоток правильного розпізнавання i -го класу, P_{max} і P_{min} (максимальне і мінімальне значення відсотка правильного розпізнавання серед заданих

класів на наборі діагностичних параметрів, що розглядається; m - кількість класів в задачі; α, β - вагові коефіцієнти (величина зворотна сумі $\sum_{i=1}^m P_i$).

Використання множини алгоритмів обробки статистичної вибірки викликає необхідність розв'язання проблеми організації обчислювального процесу в ІС. Для забезпечення ефективної взаємодії методів обробки даних в ІС, що пропонується, реалізований оригінальний засіб інтеграції перерахованих алгоритмів в єдиний обчислювальний процес. На мал. 1 представлена структурна схема організації програмного інтерфейсу, що реалізує обчислювальні алгоритми статистичного розпізнавання і дослідження інформативності діагностичних об'єктів.



Мал. 1. Структурна схема організації програмного інтерфейсу ІС

2. ІНТЕРФЕЙСНІ ЗАСОБИ ІС

З метою скорочення втрат часу на непродуктивні операції підготовки і узгодження даних між різними алгоритмами і забезпечення ефективної взаємодії методів обробки даних, а також зручності роботи зі статистичними вибірками розроблений сучасний графічний інтерфейс користувача ІС, відповідний стандартам операційної системи Windows 95/98, що дозволяє гранично спростити, стандартизувати і візуалізувати процеси маніпулювання з даними, алгоритмами дослідження початкових даних, дозволяючи створювати, змінювати статистичні вибірки, переглядати результати досліджень і здійснювати вибір переважних рішень.

Основні функції інтерфейсу користувача полягають в наданні простого і інтуїтивно зрозумілого засобу відображення початкових даних і формування повчальних статистичних вибірок. Першим кроком до збільшення продуктивності при роботі з даними є вибір вдалої форми їх представлення в програмі (вибір форми відображення). У ІС як основна форма відображення прийнята організація даних у вигляді трьохмірних таблиць (класи, ознаки, вимірювання). Це дозволяє швидко і ефективно контролювати рішення задачі на будь-якому етапі, безпосередньо формуючи класи об'єктів, що досліджуються, і робочі статистичні вибірки, що складаються з довільних підмножин повної множини діагностичних параметрів, що досліджуються. Такі можливості особливо корисні в тих випадках, коли рішення задачі діагностичного контролю об'єктів вимагає від дослідника, що здійснює пошук відповідного рішення, підготувати цілий ряд робочих статистичних вибірок, відмінних одна від одної лише набором включених до їх складу діагностичних параметрів об'єктів, що досліджуються, або кількістю і складом класів

об'єктів, що розглядаються, тоді як непродуктивний час, затрачений на безпосереднє формування робочих вибірок перевищує у багато разів час, що затрачується на рішення задачі.

ВИСНОВКИ

Використання розробленої ІС для рішення практичних задач діагностичного контролю в системі управління якістю і сертифікація, розглянутих в роботі, демонструє істотне скорочення часу, що відводиться на пошук задовільних рішень задачі, скорочення кількості діагностичних параметрів в описі об'єкта досліджень без зниження якості класифікації його можливих станів, зниження об'єму статистичної вибірки і отримання вирішальних правил, що забезпечують надійне розпізнавання.

Реалізація різноманітних методів в ІС дозволяє зробити процес пошуку задовільного рішення дуже гнучким і ефективним.

ЛІТЕРАТУРА

1. Фукунага К. Введение в статистическую теорию распознавания образов. - М.: Наука, 1970. - 368 с.
2. Цыпкин Я.З. Основы теории обучающихся систем. - М.: Наука, 1970. - 399 с.
3. Пакет прикладных программ ОТЭКС (для анализа данных) / Н.Г. Загоруйко, В.Н. Ёлкина, С.В. Емельянов, Г.С. Лбов. - М.: Финансы и статистика, 1986. - 160 с.: ил. - (Мат. Обеспечение прикладной статистики).
4. Дж. Ту, Р. Гонсалес. Принципы распознавания образов: Пер. с англ. под ред. Ю.И. Журавлёва. - М.: Мир, 1978 - 411 с.