

УДК 681.518:004.93'1

А.С. Довбиш, І.В. Шелехов, О.В. Коробченко

ОПТИМІЗАЦІЯ СЛОВНИКА ОЗНАК РОЗПІЗНАВАННЯ СИСТЕМИ КЕРУВАННЯ, ЩО НАВЧАЄТЬСЯ

Анотація: Запропоновано у рамках інформаційно-екстремальної інтелектуальної технології аналізу даних, математичну модель і алгоритм самонавчання системи підтримки прийняття рішень з оптимізацією словника ознак розпізнавання за схемами послідовної спадної селекції, випадкового пошуку і випадкового пошуку з адаптацією.

Ключові слова: інформаційно-екстремальна інтелектуальна технологія, оптимізація, словник ознак, система підтримки прийняття рішень, послідовна спадна селекція, випадковий пошук.

Вступ

У задачах аналізу і синтезу системи підтримки прийняття рішень (СППР), яка є складовою автоматизованої системи керування технологічним процесом (АСКТП), важливого значення набуває задача оптимізації (тут і далі в інформаційному розумінні) її вхідного математичного опису. В більшості СППР використовуються надлишкові словники ознак розпізнавання, що призводить до значних витрат і надмірної складності аналізу предметної галузі на етапі його формування. Одним із шляхів вирішення цієї проблеми є оптимізація словника ознак розпізнавання СППР. Оскільки основною функцією СППР будь-якого призначення є прийняття рішень, то оцінку її функціональної ефективності варто розглядати у рамках теоретико-інформаційного підходу [1]. Тому одним із перспективних напрямів аналізу і синтезу здатних навчатися СППР є інформаційно-екстремальна інтелектуальна технологія (ІЕІ-технологія), що ґрунтується на максимізації інформаційної спроможності системи [2–4].

У статті розглядається оптимізація словника ознак розпізнавання в рамках ІЕІ-технології за умови, що структуризація словника здійснюється за методами послідовної спадної селекції, випадкового пошуку та випадкового пошуку з адаптацією.

Постановка задачі

Розглянемо АСКТП, складовою частиною якої є здатна навчатися СППР, що функціонує в режимі кластер-аналізу вхідних даних. Нехай відомі алфавіт класів розпізнавання $\{X_m^o | m = \overline{1, M}\}$, які характеризують функціональні стани керованого технологічного процесу, і вхідна нечітка класифікована багатовимірною навчальною матрицею $\|y_{m,i}^{(j)} | i = \overline{1, N}; j = \overline{1, n}\|$ де N, n – кількість ознак розпізнавання і випробувань (спостережень) відповідно. Крім того попередньо сформовано структурований вектор параметрів навчання

$$g = \langle x_m, d_m, \delta, \Sigma \rangle, \quad (1)$$

де x_m – еталонний вектор-реалізація класу X_m^o , вершина якого визначає геометричний центр контейнера, побудованого в радіальному базисі простору ознак; d_m – радіус контейнера класу X_m^o ; δ – параметр поля контрольних допусків; Σ – структурований словник ознак розпізнавання.

Необхідно в режимі самонавчання СППР оптимізувати координати вектора параметрів (1) шляхом ітераційної процедури пошуку глобального максимуму усередненого за алфавітом класів розпізнавання інформаційного критерію функціональної ефективності (КФЕ) навчання системи в робочій (допустимій) області визначення його функції:

$$\bar{E}^* = \frac{1}{M} \sum_{m=1}^M \max_{G \cap \{k\}} E_m \quad (2)$$

де G – область допустимих значень параметрів навчання; $\{k\}$ – множина кроків навчання СППР розпізнавання реалізації класу X_m^o .

Математична модель

Математична модель СППР, що навчається в рамках ІЕТ, включає вхідний математичний опис, який подається у вигляді теоретико-множинної структури:

$$\Delta_B = \langle G, T, Z, \Omega, Y, X; \Phi_1, \Phi_2 \rangle,$$

де G – простір вхідних сигналів (факторів); T – множина моментів часу зняття інформації; Z – простір можливих функціональних станів керованого процесу; Ω – простір ознак розпізнавання; Y – вибіркова множина – багатовимірна вхідна навчальна матриця типу “об’єкт-властивість”; $\Phi_1 : G \times T \times \Omega \times Z \rightarrow Y$ – оператор формування множини Y ; $\Phi_2 : Y \rightarrow X$ – оператор перетворення матриці Y в бінарну навчальну матрицю X .

На рис. 1 показано категорійну модель у вигляді діаграми відображення множин, задіяних в процесі навчання СППР з оптимізацією словника ознак розпізнавання [2].

В діаграмі (рис. 1) оператор $\theta : X \rightarrow \mathfrak{R}^{|M|}$ буде в просторі ознак в загальному випадку нечітке розбиття $\mathfrak{R}^{|M|}$. Оператор класифікації $\psi : \mathfrak{R}^{|M|} \rightarrow l^{|l|}$ перевіряє основну статистичну гіпотезу про належність реалізацій $\{x_m^{(j)} | j = \overline{1, n}\}$ класу, X_m^o де l – кількість статистичних гіпотез. Оператор $\gamma : l^{|l|} \rightarrow \mathfrak{S}^{|q|}$ шляхом оцінки прийнятих гіпотез формує множину точнісних характеристик $\mathfrak{S}^{|q|}$, де $q = l^2$, а оператор $\varphi : \mathfrak{S}^{|q|} \rightarrow E$ обчислює множину значень інформаційного КФЕ, який є функціоналом від точнісних характеристик. Ітераційні процеси оптимізації геометричних параметрів розбиття $\mathfrak{R}^{|M|}$ і параметра поля контрольних допусків на ознаки розпізнавання

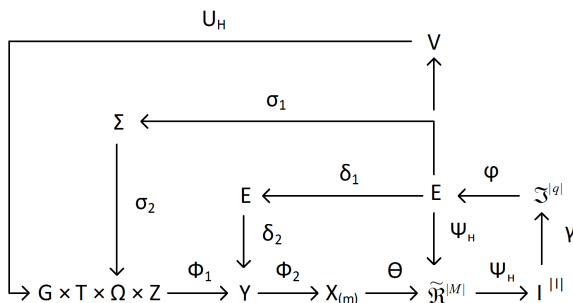


Рис. 1 – Категорійна модель оптимізації словника ознак

реалізуються оператором $r : E \rightarrow \mathfrak{R}^{|M|}$ і операторами $\delta_1 : E \rightarrow D$ і $\delta_2 : D \rightarrow Y$, де D – система контрольних допусків на ознаки розпізнавання. Крім того, в діаграму (рис. 1) введено додатковий контур, який за допомогою операторів σ_1 і σ_2 змінює структуру словника Σ . Запропонована модель передбачає перехід до використання інших типів радіально-базисних вирішальних правил у випадках, коли за результатами навчання не вдається побудувати безпомилкові за навчальною матрицею вирішальні правила. З цією метою у зовнішній контур введено множину V типів вирішальних правил. Оператор U регламентує процес навчання СППР.

Реалізація алгоритму машинного навчання

Згідно з категорійною моделлю (рис. 1) оптимізація словника ознак здійснювалася в процесі навчання СППР за трьохциклічною ітераційною процедурою пошуку глобального максимуму інформаційного КФЕ в робочій (допустимій) області визначення його функції

$$\Sigma^* = \arg \max_{G_\Sigma} \{ \max_{G_\delta} \{ \max_{G_E \cap \{k\}} \overline{E_k} \} \} \quad (3)$$

де $\overline{E_k}$ – усереднений за алфавітом класів розпізнавання інформаційний КФЕ навчання СППР, що обчислюється на кожному k -му кроці навчання; G_Σ – допустима область потужності словника ознак розпізнавання; G_δ – область допустимих значень поля контрольних допусків; G_E – робоча область допустимих значень функції КФЕ; $\{k\}$ – множина кроків навчання СППР.

Як інформаційний КФЕ навчання СППР розглядалася модифікація ентропійного (за Шенноном) критерію, яка для двох альтернативних рівноймовірних гіпотез має вигляді [2]

$$E_m^{(k)} = 1 + \frac{1}{2} \left(\frac{\alpha_m^{(k)}(d)}{\alpha_m^{(k)}(d) + D_{2,m}^{(k)}(d)} \log_2 \frac{\alpha_m^{(k)}(d)}{\alpha_m^{(k)}(d) + D_{2,m}^{(k)}(d)} + \right.$$

$$\begin{aligned}
& + \frac{\beta_m^{(k)}(d)}{D_{1,m}^{(k)}(d) + \beta_m^{(k)}(d)} \log_2 \frac{\beta_m^{(k)}(d)}{D_{1,m}^{(k)}(d) + \beta_m^{(k)}(d)} + \\
& + \frac{D_{1,m}(d)}{D_{1,m}^{(k)}(d) + \beta_m^{(k)}(d)} \log_2 \frac{D_{1,m}(d)}{D_{1,m}^{(k)}(d) + \beta_m^{(k)}(d)} + \\
& + \frac{D_{2,m}^{(k)}(d)}{\alpha_m^{(k)}(d) + D_{2,m}^{(k)}(d)} \log_2 \frac{D_{2,m}^{(k)}(d)}{\alpha_m^{(k)}(d) + D_{2,m}^{(k)}(d)} \Big), \quad (4)
\end{aligned}$$

де $\alpha_m^{(k)}(d)$ – помилка першого роду прийняття рішення на k -му кроці навчання; $\beta_m^{(k)}(d)$ – помилка другого роду; $D_{1,m}^{(k)}(d)$ – перша достовірність; $D_{2,m}^{(k)}(d)$ – друга достовірність; d – дистанційна міра, яка визначає радіуси гіперсферичних контейнерів, побудованих в радіальному базисі простору Хеммінга.

Оптимізацію словника ознак розпізнавання здійснювалася за трьома методами його редукції: послідовна спадна селекція, випадковий пошук та випадковий пошук з адаптацією.

Метод послідовно спадної селекції в процесі оптимізації словника ознак розпізнавання оперує тільки одним поточним словником. При цьому на кожній ітерації видаляються ознаки, до тих пір, поки отриманий словник не буде задовольняти певному критерію зупину. Недоліком цього методу є те, що без перевірки всіх можливих підмножин вихідного словника він не гарантує оптимального результату, оскільки ознаки, що були видалені із словника, не повертається в процесі оптимізації.

Найбільш відомими точнішими стохастичними методами з єдиним розв'язком є методи, що базуються на випадковому пошуку та його модифікаціях. За цими методами випадковим чином відбирається N_i ознак з N вхідних ознак до тих пір, поки не буде сформовано словник заданої інформативності. Модифікацією цього методу є випадковий пошук з адаптацією [5]. Це послідовно спрямована процедура, яка базується на випадковому пошуку із врахуванням результатів попередніх відборів. При цьому на початку процедури шанси всіх вхідних ознак на входження в оптимальний словник Σ^* потужністю $N_i < N$ приймаються рівними. Для випадкового відбору використовується датчик рівномірно розподілених на інтервалі $[0;1]$ випадкових (псевдовипадкових) чисел. Цей інтервал розбивається на N рівних відрізків.

Результати моделювання

Алгоритм оптимізації словника ознак розпізнавання за процедурою (3) було реалізовано в процесі навчання СППР для керування процесом вирощування великогабаритних скінтіляційних монокристалів із розплаву [6].

Навчальна матриця формувалася із структурованих векторів-реалізацій функціонального стану технологічного процесу, кожний із яких складався із 33 первинних ознак розпізнавання, які періодично зчитувалися з датчиків інформації, і вторинних, які характеризували похідні від 15 первинних ознак. Алфавіт класів розпізнавання складався із трьох класів, які характеризували три діапазони відхилення діаметру монокристалу в процесі його вирощування на заданому інтервалі спостереження

Спочатку оптимізація словника ознак розпізнавання в процесі навчання СППР здійснювалася за схемою послідовної спадної селекції, за якою на кожному кроці формувалися варіанти словників з потужністю менше на одиницю за потужність поточного словника. На кожному кроці навчання обчислювалося максимальне усереднене значення КФЕ (4) для поточного словника. На рис. 2 показано максимальні усереднені значення КФЕ, одержані при оптимізації варіантів словників, що формувалися на першому кроці навчання за схемою послідовної спадної селекції.

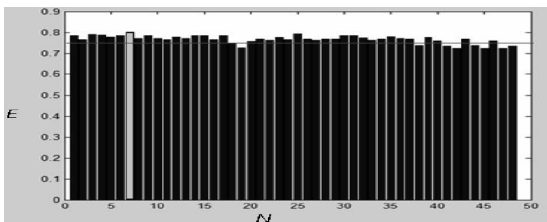


Рис. 2 – Діаграма максимальних усереднених значень КФЕ для варіантів словників, сформованих на першому кроці оптимізації

Аналіз рис. 2 показує, що на першому кроці найменш інформативною ознакою є сьома ознака, яку необхідно видалити. При цьому КФЕ досягає свого найбільшого значення, яке дорівнює $\bar{E}_1^* = 0,80$, що перевершує значення цього критерію, одержане при використанні початкового (надлишкового в інформаційному розумінні) словника ($\bar{E}_0^* = 0,73$).

Аналогічно на кожному кроці визначалася та видалялася одна найменш інформативна ознака.

На рис. 3 показано графік зміни максимального усередненого КФЕ навчання СППР для 45 варіантів словників ознак, сформованих в процесі їх оптимізації за схемою послідовної спадної селекції.

Аналіз рис. 3 показує, що в процесі оптимізації словника ознак за схемою послідовної спадної селекції значення КФЕ (4) підвищилося до 0,983 у порівнянні з початковим словником.

На рис. 4 показано графік зміни максимального усередненого КФЕ навчання СППР для 400 варіантів словників ознак, сформованих в процесі їх оптимізації за схемою випадкового пошуку.

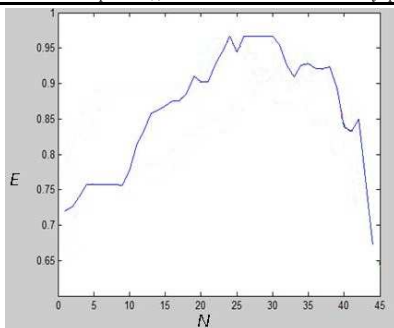


Рис. 3 – Графік зміни КФЕ при оптимізації словника ознак з використанням схеми послідовної спадної селекції

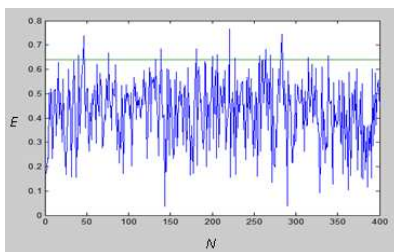


Рис. 4 – Графік зміни КФЕ при оптимізації словника ознак з використанням схеми випадкового пошуку

Аналіз рис. 4 показує, що максимальне значення КФЕ дорівнює 0,78.

На рис. 5 показано графік зміни максимального усередненого КФЕ навчання СППР для 500 варіантів словників ознак, сформованих в процесі їх оптимізації за схемою випадкового пошуку з адаптацією, яка полягала в поверненні на кожному кроці навчання випадковим чином попередньо видалених ознак.

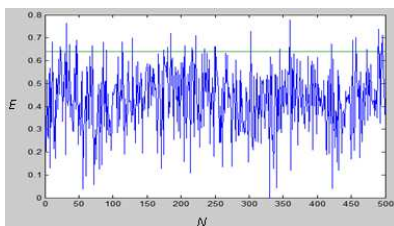


Рис. 5 – Графік зміни КФЕ при оптимізації словника ознак з використанням схеми випадкового пошуку з адаптацією

Аналіз рис. 4 показує, що максимальне значення КФЕ, одержане в процесі оптимізації словника ознак з використанням схеми випадкового пошуку з адаптацією, дорівнює 0,79.

Таким чином, аналіз рис. 3 – рис. 5 показує, що найбільшу функціональну ефективність навчання СППР одержано при оптимізації словника ознак за схемою послідовної спадної селекції.

Висновки

1. Запропоновано інформаційно-екстремальний алгоритм навчання СППР з оптимізацією словника ознак розпізнавання. При цьому показано, що найбільшу функціональну ефективність навчання СППР одержано при оптимізації словника ознак за схемою послідовної спадної селекції у порівнянні із такими схемами редуkcії, як випадковий пошук і випадковий пошук з адаптацією.

2. Для побудови безпомилкових за навчальною матрицею вирішальних правил необхідно здійснювати оптимізацію інших параметрів навчання, які впливають на функціональну ефективність машинного навчання СППР.

Список використаних джерел

1. Кузьмин И.В. Оценка эффективности и оптимизация автоматизированных систем контроля и управления / И.В. Кузьмин. – М.: Сов. Радио, 1971. – 296 с.
2. Довбиш А.С. Основи проектування інтелектуальних систем: навчальний посібник / А.С. Довбиш.– Суми: Видавництво СумДУ, 2009.– 171 с.
3. Dovbysh A.S. Information-extreme algorithm for recognizing current distribution maps in magnetocardiography / A.S. Dovbysh, S.S. Martynenko, A.S. Kovalenko, N.N. Budnyk // Journal of Automation and Information Sciences. – 2011. – V. 43.– № 2.– P. 63–70.
4. Довбиш А.С. Інтелектуальні інформаційні технології в електронному навчанні / А.С. Довбиш, А.В. Васильєв, В.О. Любчак. – Суми: Видавництво СумДУ. – 2014. – 172 с.
5. Довбиш А.С. Оцінка інформативності ознак розпізнавання за методом функціонально-статистичних випробувань / А. С. Довбиш, І.В. Шелехов // Вісник Сумського державного університету. – 2006. – №10(94). – С. 51–58.
6. Сцинтилляционные монокристаллы: автоматизированное выращивание / Суздаль В.С., Стадник П.Е., Герасимчук Л.И., Епифанов Ю.М. – Харьков: “ИСМА”, 2009.–260 с.

Отримано 19.10.2015 р.