

ІНФОРМАЦІЙНО-ЕКСТРЕМАЛЬНИЙ АЛГОРИТМ ПОБУДОВИ УНІМОДАЛЬНОГО КЛАСИФІКАТОРА

Розглядається в рамках інформаційно-екстремальної інтелектуальної технології алгоритм навчання системи підтримки прийняття рішень з унімодальним класифікатором, який доцільно використовувати для впорядкованих класів розпізнавання.

Вступ

Підвищення функціональної ефективності автоматизованих систем керування (АСК) складними технологічними процесами, що функціонують за умови апріорної невизначеності, на практиці досягається шляхом надання їм властивості адаптивності на основі машинного навчання та розпізнавання образів [1 – 3]. Одним із перспективних підходів щодо аналізу і синтезу інтелектуальної системи підтримки прийняття рішень (СППР), яка входить до складу сучасної адаптивної АСК, є використання ідей і методів інформаційно-екстремальної інтелектуальної технології (ІЕІ-технологія), що ґрунтується на максимізації інформаційної спроможності СППР шляхом введення у процесі навчання додаткових інформаційних обмежень [4,5].

У наукових працях [6 – 8] розглядалося питання оптимізації у рамках ІЕІ-технології просторово-часових параметрів функціонування, які впливають на функціональну ефективність СППР для мультимодального класифікатора, що характеризується наявністю декількох центрів розсіювання реалізацій класів із заданого алфавіту. Але на практиці існують задачі, які потребують застосування унімодального класифікатора з одним загальним центром розсіювання для реалізацій усіх класів розпізнавання. До цих задач відносяться, наприклад, задачі керування із триальтернативною системою оцінок типу «Норма», «Менше норми» і «Більше норми».

У статті запропоновано алгоритм навчання СППР з оптимізацією контрольних допусків на ознаки розпізнавання для унімодального класифікатора.

1. Постановка задачі

Нехай для алфавіту нечітких класів розпізнавання $\{X_m^0 \mid m = \overline{1, M}\}$, які характеризують M допустимих функціональних станів складного організаційно-технічного комплексу, сформовано апріорно класифіковану нечітку навчальну матрицю типу «об'єкт-властивість» $\|y_{m,i}^{(j)}\|$, $i = \overline{1, N}$, $j = \overline{1, n}$, де N , n – кількість ознак розпізнавання і випробувань відповідно. Крім того дано вектор параметрів функціонування СППР $g = \langle x_1, d_{m,1}, d_{m,2}, \delta \rangle$, де x_1 δ – еталонний вектор-реалізація базового класу X_1^0 , який визначає геометричний центр унімодального класифікатора – центр розсіювання векторів-реалізацій усіх класів розпізнавання; $d_{m,1}$, $d_{m,2}$ – менший і більший радіуси контейнера класу X_m^0 , що відновлюється в радіальному базисі простору ознак розпізнавання, відповідно; δ – параметр поля контрольних допусків. При цьому задано такі обмеження: $x_1 \in \Omega_B$, де Ω_B – бінарний простір ознак розпізнавання (Хеммінга); $d_{m,1} = d_{m-1,2} < [0; d(x_m \oplus x_c) - 1]$, де $d(x_m \oplus x_c)$ – кодова відстань центру класу X_m^0 від центру найближчого (сусіднього) до нього класу X_c^0 і параметр $\delta \in [0; \delta_H / 2]$, де δ_H – нормоване (експлуатаційне) поле допусків для відносної шкали вимірювання ознак, яке є областю значень для параметра контрольного поля допусків δ . При цьому нижній допуск на i -ту ознаку дорівнює $A_{H,i} = y_{1,i} - \delta$, а верхній допуск – $A_B = y_{1,i} + \delta$. Тут $y_{1,i}$ – i -та ознака еталонного вектора-реалізації y_1 базового класу X_1^0 . На етапі навчання необхідно оптимізувати параметри навчання за усередненим інформаційним критерієм функціональної ефективності (КФЕ)

де E_m^* – глобальний максимум інформаційного КФЕ навчання СППР розпізнавати реалізації класу X_m^0 , що знаходиться в робочій області визначення його функції, і на етапі екзамену прийняти високостовірне рішення про належність вектора-реалізації образу, що розпізнається, до деякого класу із заданого алфавіту $\{X_m^0\}$.

2. Категорійна модель навчання унімодальної СППР

Категорійну модель побудови унімодального класифікатора для побудови оптимального (тут і далі в інформаційному розумінні) розбиття простору ознак на класи розпізнавання, яке утворює сукупність вирішальних правил, або класифікатор, подамо у вигляді діаграми відображень множин. Як і у випадку мультимодальності розподілу реалізацій образу категорійна модель унімодальної СППР містить оператор формування вхідного математичного опису (багатовимірної навчальної матриці типу «об'єкт-властивість»)

$$F: G \times T \times \Omega \times Z \rightarrow Y, \quad (2)$$

де G – простір входних сигналів (факторів), які діють на СППР; T – множина моментів часу зняття інформації; Ω – простір ознак розпізнавання; Z – простір можливих функціональних і технічних станів СППР; Y – вибіркова множина значень рецепторів (вхідна навчальна матриця).

Крім того, для інформаційно-екстремального класифікатора обов'язковими є оператори:

- побудови у загальному випадку нечіткого розбиття $\tilde{\mathfrak{R}}^{[M]}$ простору ознак на класи розпізнавання

$$\theta: Y \rightarrow \tilde{\mathfrak{R}}^{[M]}; \quad (3)$$

– перевірки основної статистичної гіпотези про належність реалізації $x_m^{(j)}$ класу X_m^O

$$\Psi: \widetilde{\mathfrak{R}}^{[M]} \rightarrow \mathbf{I}^{[I]}, \quad (4)$$

де $I^{[1]} = \{\gamma_1, \dots, \gamma_l\}$ – множина статистичних гіпотез;

- формування за результатами перевірки основної статистичної гіпотези множини точнісних характеристик

$$\gamma:\mathbf{I}^{|\mathbf{l}|}\rightarrow\mathfrak{I}^{|\mathbf{q}|}, \quad (5)$$

де $q = l^2$ – кількість точнісних характеристик;

– обчислення терм-множини E , елементами якої є значення КФЕ навчання СППР як функціонала від точнісних характеристик

$$\varphi: \mathfrak{I}^{|\mathbf{q}|} \rightarrow \mathbb{E}. \quad (6)$$

Аналогічно діаграма відображення множин, що формуються при навчанні унімодальної СППР, у зворотному зв'язку містить побудовані за принципом повної композиції контури оптимізації відповідних параметрів навчання. Основною відмінністю моделі унімодальної СППР є наявність додаткового контуру вибору відповідного вирішального правила для реалізації оператора перевірки основної статистичної гіпотези. На рис. 1 з урахуванням виразів (2)–(6) показано діаграму відображення множин, що формуються в процесі навчання унімодальної СППР з оптимізацією системи контрольних допусків (СКД) на ознаки розпізнавання.

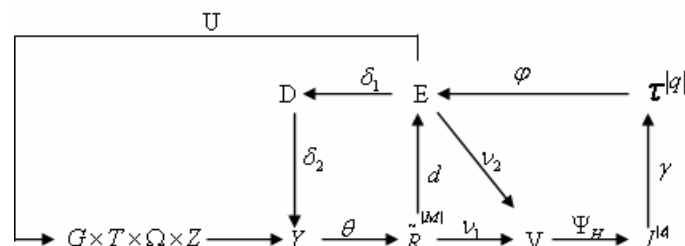


Рис. 1. Категорійна модель навчання унімодальної СППР

У діаграмі (рис.1) терм-множина D складається із допустимих значень СКД на ознаки розпізнавання, а множина V є впорядкованим вектором вирішальних правил. Контур

$$\boxed{\rightarrow v_1 \rightarrow \Psi_H \rightarrow \gamma \rightarrow \varphi \rightarrow d} \quad (7)$$

оптимізує геометричні параметри побудованих у радіальному базисі простору ознак розпізнавання контейнерів класів розпізнавання, а послідовність операторів контуру

$$\boxed{\rightarrow \theta \rightarrow v_1 \rightarrow \Psi_H \rightarrow \gamma \rightarrow \varphi \rightarrow \delta_1 \rightarrow \delta_2} \quad (8)$$

оптимізує контрольні допуски на ознаки розпізнавання. У контурі (7) оператор v_1 відображує розбиття $\tilde{\mathcal{R}}^{[M]}$ на відповідні вирішальні правила, оператор Ψ_H перевіряє основну статистичну гіпотезу за відповідним вирішальним правилом і оператор d змінює геометрію контейнерів відповідних класів розпізнавання. Таким чином, для унімодального класифікатора оператор (4) розглядається як композиція операторів $\Psi = v_1 \circ \Psi_H$. У контурі (8) оператор δ_1 дає дозвіл на формування наступних кроків навчання СППР у процесі пошуку глобального максимуму КФЕ в робочій області його визначення, а оператор δ_2 формує відповідні контрольні допуски на ознаки розпізнавання.

Специфічним контуром для діаграми відображення множин при побудові унімодального класифікатора є контур

$$\boxed{\rightarrow \Psi_H \rightarrow \gamma \rightarrow \varphi \rightarrow v_2} \quad (9)$$

де v_2 – оператор переходу до наступного вирішального правила після оптимізації контейнера попереднього класу розпізнавання.

Оператор $U: E \rightarrow G \times T \times \Omega \times Z$ регламентує процес навчання СППР.

3.Алгоритм побудови унімодального класифікатора

Алгоритм навчання унімодальної СППР з оптимізацією СКД на ознаки розпізнавання, як і для мультимодальної, полягає у реалізації структурованої двоцикличованої ітераційної процедури пошуку глобального максимуму інформаційного КФЕ (1) в робочій області визначення його функції. Для інформаційно-екстремального алгоритму паралельної оптимізації СКД, при якому параметр поля контрольних допусків δ_K змінюється одночасно для всіх ознак розпізнавання, така процедура має вигляд

$$\delta_K^* = \arg \max_{G_\delta} \{ \max_{G_E} \bar{E} \}, \quad (10)$$

де G_δ – область допустимих значень відповідних контрольних допусків на ознаки розпізнавання; G_E – область допустимих значень інформаційного КФЕ (1) навчання СППР.

У процедурі (10) внутрішній цикл реалізує базовий алгоритм навчання [1], основними задачами якого є:

- обчислення інформаційного КФЕ навчання СППР;
- пошук глобального максимуму КФЕ в робочій (допустимій) області визначення його функції;
- оптимізація геометричних параметрів контейнерів класів розпізнавання.

При цьому специфіка базового алгоритму навчання унімодальної СППР полягає у відсутності процедури визначення для кожного класу найближчого сусіда, оскільки класи розпізнавання априорно є впорядкованими, що суттєво підвищує оперативність навчання.

Розглянемо у рамках ІЕІ-технології алгоритм навчання унімодальної СППР з паралельною оптимізацією контрольних допусків на ознаки розпізнавання (11). Вхідні дані: масив реалізацій класів розпізнавання $\{y_{m,i}^{(j)} \mid m = \overline{1, M}; i = \overline{1, N}; j = \overline{1, n}\}$; система нормованих допусків $\{\delta_{H,i} \mid i = \overline{1, N}\}$, що визначає область значень відповідних контрольних допусків. Попередньо для кожної ознаки визначається ціна градації її шкали виміру, що дозволяє обчислювати на кожному кроці навчання нижній і верхній контрольні допуски відповідно:

$$A_{KH,i} = y_{1,i} - \delta; A_{KB,i} = y_{1,i} + \delta, \quad (11)$$

де $y_{1,i}$ – вибіркове середнє значення i -ї ознаки у векторах-реалізаціях базового класу X_1^0 .

Реалізація алгоритму навчання унімодальної СППР з паралельною оптимізацією контрольних допусків на ознаки розпізнавання здійснюється за такою схемою:

- 1) обнуляється лічильник кроків зміни параметра δ (кроків навчання): $l:=0$;
 - 2) $\delta:l:=l+1$;
 - 3) на кожному кроці навчання за формулами (10) обчислюються нижній $A_{HK,i}[l]$ і верхній $A_{BK,i}[l]$ контрольні допуски для всіх ознак розпізнавання;
 - 4) реалізується базовий алгоритм навчання і визначається поточний глобальний максимум усередненого за алфавітом класів розпізнавання інформаційного критерію $\bar{E}[l]$ в робочій області визначення його функції.
 - 5) якщо в робочій області визначення функції інформаційного критерію має місце $\bar{E}[l] \leq \text{extr max } \bar{E}$, де $\text{extr max } \bar{E}$ – граничний максимум усередненого за алфавітом класів розпізнавання інформаційного критерію (1), то виконується пункт 6, інакше – пункт 8 ($\bar{E}[0] = 0$);
 - 6) якщо $\delta \leq \delta_H / 2$, то виконується пункт 2, інакше – пункт 7;
 - 7) $\text{extr max } \bar{E} := \max_{\{\delta\}} \bar{E}^*[l]$;
 - 8) $\delta^* := \arg \text{extr max } \bar{E}$;
 - 9) для параметра δ^* обчислюються оптимальні нижні $\{A_{HK,i}^*\}$ і верхні $\{A_{BK,i}^*\}$ контрольні допуски на ознаки розпізнавання;
 - 1) виведення оптимальних двійкових еталонних векторів-реалізацій $\{x_m^*\}$, вершини яких визначають геометричні центри контейнерів класів розпізнавання;
 - 2) ЗУПИН.
- Як КФЕ навчання може використовуватися одна із статистичних інформаційних мір. Наприклад, для двоальтернативних рішень модифікований критерій Кульбака, що використовується для оцінки функціональної ефективності навчання СППР розпізнавати реалізації класу X_m^0 , має вигляд [5]

$$E_m^{(k)} = \log_2 \left(\frac{2 - (\alpha_m^{(k)}(d) + \beta_m^{(k)}(d))}{\alpha_m^{(k)}(d) + \beta_m^{(k)}(d)} \right) * \left[1 - (\alpha_m^{(k)}(d) + \beta_m^{(k)}(d)) \right], \quad (12)$$

де $\alpha_m^{(k)}(d)$, $\beta_m^{(k)}(d)$ – точнісні характеристики: похибки першого та другого роду відповідно, що обчислюються на k -му кроці навчання; d – дистанційна міра, що визначає радіуси контейнерів, побудованих у радіальному базисі простору ознак розпізнавання.

Оцінками похибок першого та другого роду при прийнятті рішень у процесі навчання є емпіричні частоти

$$\tilde{\alpha}_m^{(k)}(d) = \frac{K_{1,m}^{(k)}}{n_{\min}}; \tilde{\beta}_m^{(k)}(d) = \frac{K_{2,m}^{(k)}}{n_{\min}}, \quad (13)$$

де $K_{1,m}^{(k)}$ – кількість подій, що визначають неналежність реалізацій образу контейнеру класу X_m^0 , якщо насправді вони є реалізаціями цього класу; $K_{2,m}^{(k)}$ – кількість подій, що визначають належність реалізацій образу контейнеру класу X_m^0 , якщо насправді вони належать іншому класу; $n_{\min}$ – мінімальний обсяг репрезентативної навчальної вибірки [1].

Суми $K_{1,m}^{(k)}$ і $K_{2,m}^{(k)}$ обчислюються на k -му кроці навчання унімодальної СППР відповідно ($K_{1,m}^{(k)}[0] = 0$, $K_{2,m}^{(k)}[0] = 0$):

if $x_m^{(j)} \notin X_m^0$ then $K_{1,m}^{(k)}[j] := K_{1,m}^{(k)}[j-1] + 1$;

if $x_c^{(j)} \in X_m^0$ then $K_{2,m}^{(k)}[j] := K_{2,m}^{(k)}[j-1] + 1$,

де $x_c^{(j)}$ – j -та реалізація найближчого (сусіднього) класу X_c^0 .

Належність деякої реалізації $x^{(j)}$, наприклад, класу X_m^0 для унімодального класифікатора здійснюється за правилом

if $d_{m-1}^* < d[x_1 \oplus x^{(j)}] < d_m$ then $x^{(j)} \in X_m^0$ else $x^{(j)} \notin X_m^0$,

де d_{m-1}^* – визначений у процесі навчання оптимальний радіус контейнера лівого сусіднього класу; $d[x_1 \oplus x^{(j)}]$ – кодова відстань вектора $x^{(j)}$ від центра розсіювання реалізацій, що визначається вершиною усередненого вектора x_1 ; $\oplus$ – символ операції складання за модулем два; d_m – поточний радіус контейнера класу X_m^0 , що відновлюється в бінарному просторі ознак розпізнавання.

Аналіз вищенаведеного алгоритму показує, що він дозволяє навчання системи без досягнення граничного максимум інформаційного критерію (1), що відповідає принципу відкладених рішень і надає можливість здійснювати оптимізацію інших параметрів навчання з метою побудови безпомилкового за навчальною матрицею класифікатора.

4. Приклад реалізації алгоритму навчання унімодальної СППР

Розглянемо результати реалізації запропонованого алгоритму на прикладі паралельної оптимізації контрольних допусків на ознаки розпізнавання при навчанні СППР, що використовується в АСК технологічним процесом виробництва фосфорної кислоти у ВАТ «Сумихімпром». Вхідну апріорно класифіковану навчальну матрицю було сформовано за експериментальними даними, отриманими для трьох класів, що характеризували зміст P_2O_5 на виході п'ятого ступіню екстрактора за такою триальтернативною системою оцінок: «Норма», «Менше норми» і «Більше норми». При цьому кількість ознак розпізнавання дорівнювала 65, із них 15 поступали за результатами хімічного аналізу, а інші значення різних за природою ознак розпізнавання поступали в СППР безпосередньо з датчиків інформації. Як базовий було обрано клас X_1^0 , що характеризував функціональний стан технологічного процесу «Більше норми». Клас X_2^0 характеризував заданий режим технологічного процесу «Норма» і клас X_3^0 – «Менше норми».

На рис. 1 показано графік залежності критерію Кульбака (12) від параметра поля контрольних допусків, одержаний у процесі паралельної оптимізації контрольних допусків на ознаки розпізнавання. На графіку темні ділянки позначають робочі області визначення функції інформаційного критерію, в яких перша і друга достовірності перевершують відповідно похибки першого та другого роду і одночасно виконується умова $d_m > d_{m-1}$. При цьому у загальному випадку з кількістю класів більше двох можуть існувати декілька робочих областей.

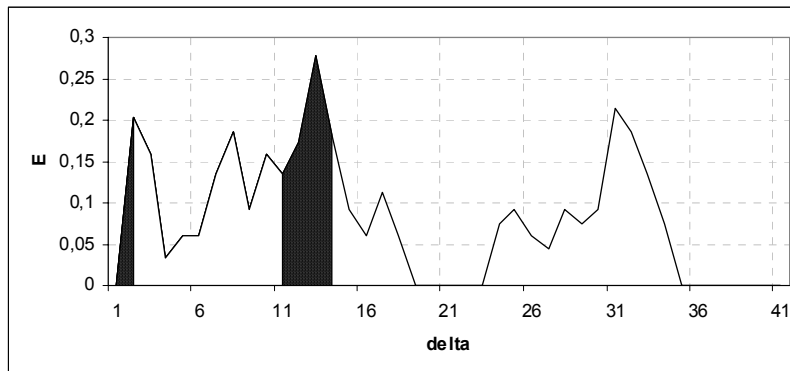


Рис. 2. Графік залежності критерію Кульбака від параметра поля контрольних допусків

Аналіз рис.2 показує, що оптимальний параметр поля контрольних допусків дорівнює $\delta \pm 13$ (у відносних одиницях) при значенні максимуму усередненого критерію $\bar{E}^* = 0,28$.

При цьому побудова мультимодального класифікатора з паралельною оптимізацією контрольних допусків на ознаки розпізнавання для цих класів і тієї самої вхідної навчальної матриці забезпечує вдвічі менше значення максимуму критерію $\bar{E}^*$.

Оскільки метою навчання СППР є відновлення в просторі ознак оптимальних контейнерів класів розпізнавання, то на рис.3 і рис.4 наведено графіки залежності КФЕ (12) від радіусів контейнерів класів X_1^0 і X_2^0 .

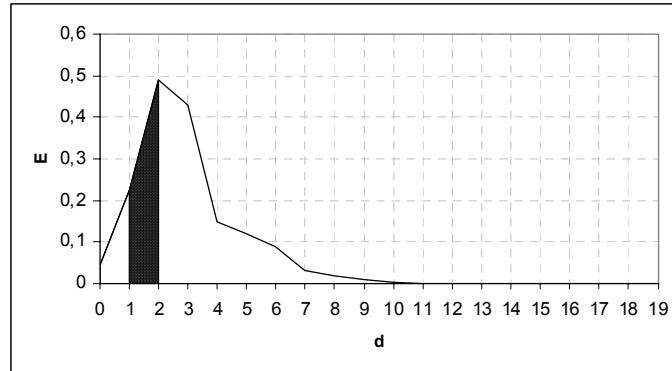


Рис. 3. Графік залежності критерію Кульбака від радіусу контейнера класу X_1^0

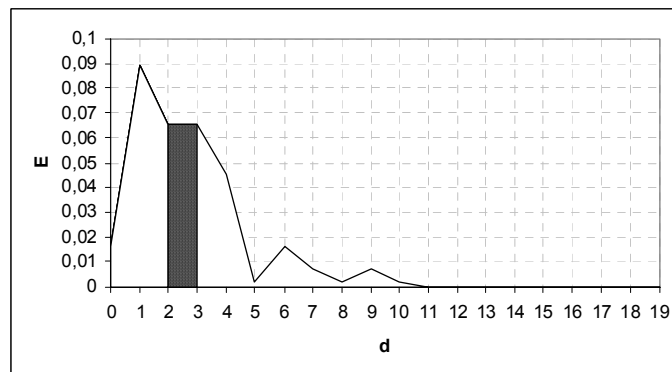


Рис. 4. Графік залежності критерію Кульбака від радіусу контейнера класу X_2^0

Аналіз рис. 3 і 4 показує, що оптимальні радіуси контейнерів класів X_1^0 і X_2^0 дорівнюють двом кодовим одиницям. Оскільки центр ваги реалізацій класу X_2^0 зміщений від центра розсіювання реалізацій всіх класів на відстань $d(x_1 \oplus x_2) = 3$, де x_2 — усереднений вектор-реалізація для класу X_2^0 , оптимальний радіус для класу X_1^0 , дорівнює $d_1^* = 2$ і для класу X_2^0 — $d_2^* = 5$.

Отже, аналіз значень КФЕ показує, що одержані в прикладі оптимальні параметри навчання не забезпечують побудову безпомилкового за навчальною матрицею класифікатора, тобто вони можуть розглядатися як квазіоптимальні. Для побудови безпомилкового за навчальною матрицею класифікатора згідно з принципом відкладених рішень необхідна оптимізація інших параметрів навчання [4,5].

Висновки

1. Для впорядкованих класів розпізнавання, які мають єдиний центр розсіювання векторів-реалізацій всіх класів, використання для навчання системи підтримки прийняття рішень унімодального класифікатора дозволяє підвищити достовірність і оперативність навчання у порівнянні з аналогічним мультимодальним класифікатором.

2. Основна відмінність унімодального класифікатора від мультимодального полягає у відновленні в процесі навчання контейнерів класів розпізнавання з єдиним геометричним центром, що більше відповідає реальному розподілу векторів реалізацій впорядкованих класів розпізнавання у бінарному просторі ознак.

Список літератури: 1. *Васильев В.И.* Распознающие системы. Справочник. 2-е изд., перераб. и доп.- Киев: Наукова думка, 1983. 422 с. 2. *Vapnik V.* Statistical Learning Theory, John Wiley&Sons. New York, 1998. 732 p. 3. *Advances in Learning Theory: Methods, Models and Application* / J.A.K. Suykens, G. Horvath, S. Basu, C. Micchelli, J. Vandewalle // IOS Press NATO-ASI Series in Computer and Systems Sciences, Amsterdam, The Nether-Lands, 2003. 432 p. 4. Інформаційний синтез інтелектуальних систем керування, що навчаються: Підхід, що ґрунтується на методі функціонально-статистичних випробувань / Краснополюсовський А.С. Суми: Видавництво СумДУ, 2003. 264 с. 5. *Основи проектування інтелектуальних систем: Навчальний посібник.* / Довбиш А.С. Суми: Видавництво Сум ДУ, 2009. 171 с. 6. *Довбиш А.С., Козинець М.В., Котенко С.М.* Оптимізація контрольних допусків на ознаки розпізнавання в інформаційно-екстремальних методах автоматичної класифікації // Вісник Сумського державного університету. Серія «Техніка», №1, 2007. С. 169-178. 7. *Довбиш А.С., Мартиненко С.С.* Інформаційно-екстремальний метод розпізнавання електронограм // Вісник СумДУ. Технічні науки. 2009. №2. С. 85-92. 8. *Краснополюсовський А.С.* Оптимізація контрольних допусків на ознаки розпізнавання за методом функціонально-статистичних випробувань // Штучний інтелект. 2003. №1. С. 53 – 61.

Надійшла до редколегії 10.09.2010

Довбиш Анатолій Степанович, д-р техн. наук, професор, завідувач кафедри інформатики Сумського державного університету. Наукові інтереси: інформаційний аналіз і синтез інтелектуальних систем, що навчаються (самонавчаються). Адреса: Україна, 40035, Суми, вул. Заливна, 7, кв. 40, тел. 050-307-83-74, e-mail: kras@id.sumdu.edu.ua.

Ліщинський Олександр Володимирович, аспірант кафедри інформатики Сумського державного університету. Наукові інтереси: інформаційний аналіз і синтез інтелектуальних систем, що навчаються (самонавчаються). Захоплення та хобі: туризм, музика, плавання. Адреса: Україна, 40007, Суми, вул. Римського-Корсакова, 2, тел. 0668819256, e-mail: delaver@i.ua.

Востоцький Віталій Олексійович, аспірант кафедри інформатики, молодший науковий співробітник науково-дослідної лабораторії інтелектуальних систем Сумського державного університету. Наукові інтереси: інформаційний аналіз і синтез інтелектуальних систем, що навчаються (самонавчаються), класифікаційне керування, інтелектуальний аналіз даних. Захоплення та хобі: читання, плавання, музика, програмування, фотографування. Адреса: Україна, 40007, Суми, вул. Римського-Корсакова, 2, тел. 0956303310, e-mail: v.vostotskiy@id.sumdu.edu.ua.