

ОЦІНКА ФУНКЦІОНАЛЬНОЇ ЕФЕКТИВНОСТІ НАВЧАННЯ АВТОМАТИЗОВАНОЇ СИСТЕМИ КЕРУВАННЯ ТЕХНОЛОГІЧНИМ ПРОЦЕСОМ

А. С. Довбиш, д-р техн. наук, професор;

О. Б. Берест, аспірант,

Сумський державний університет, м. Суми

Розглядається в рамках інформаційно-екстремальної технології використання нормованої логарифмічної статистичної інформаційної міри Кульбака для оцінки функціональної ефективності інтелектуальної системи керування вирощуванням сцинтиляційних монокристалів. Досліджено вплив кількості реалізацій на інформаційний критерій.

Ключові слова: *інформаційний критерій, інтелектуальна система керування, ознака розпізнавання, критерій функціональної ефективності, сцинтиляційний монокристал.*

ВСТУП

Основним шляхом підвищення функціональної ефективності автоматизованої системи керування технологічними процесами (АСКТП) є надання їй властивості адаптивності на основі використання ідей і методів машинного навчання та розпізнавання образів [1-2]. Але більшість відомих методів машинного навчання базується або на дистанційних, або на дистанційно-статистичних критеріях (наприклад, міра Махаланобіса [3]) близькості об'єктів, які не характеризують основної функції такої важливої складової АСКТП, як система підтримки прийняття рішень (СППР), пов'язана безпосередньо з процесом прийняття рішень. При цьому в задачах аналізу і синтезу здатної навчатися СППР центральною задачею є оцінка її функціональної ефективності, яка характеризує достовірність та оперативність прийняття рішень та інші властивості. Тому основною перевагою методів інформаційно-екстремальної, інтелектуальної технології (ІЕІ - технології) аналізу і синтезу здатних навчатися СППР є застосування для оцінки функціональної ефективності машинного навчання статистичних інформаційних мір, які є функціоналами від точнісних характеристик, що утворюють систему оцінок рішень, що приймаються [4-6]. У працях [7-9] запропоновано різні модифікації ентропійного (за Шенноном) критерію функціональної ефективності (КФЕ) та інформаційної міри Кульбака. Але не дослідженими все ще залишаються питання впливу на функціональну ефективність навчання потужності словника ознак розпізнавання та обчислювальний аспект оцінки інформаційної спроможності СППР.

У статті розглядається у рамках ІЕІ – технології розглядається використання нормованої статистичної інформаційної міри Кульбака для оцінки функціональної ефективності навчання СППР на прикладі керування технологічним процесом вирощування сцинтиляційних монокристалів.

ОСНОВНІ РЕЗУЛЬТАТИ ДОСЛІДЖЕННЯ

Для оцінки функціональної ефективності навчання СППР може бути застосовано будь-який статистичний інформаційний критерій, який відповідає вимогам інформаційних мір.

Для обчислення інформаційного КФЕ навчання СППР будемо використовувати логарифмічну статистичну міру Кульбака. Нехай у загальному випадку дано нечіткий алфавіт класів розпізнавання $\{X_m^o \mid m = \overline{1, M}\}$ і сформована у бінарному просторі ознак нечітка апріорно класифікована навчальна матриця $\|x_{m,i}^{(j)}\|$, $i = \overline{1, N}$, $j = \overline{1, n}$, де N, n – кількість ознак розпізнавання і векторів-реалізацій відповідно. Розглянемо гіперсферичний класифікатор, в якому вирішальні правила будуються на кожному k -му кроці навчання СППР шляхом відновлення в радіальному базисі простору Ω_B гіперсферичних контейнерів класів розпізнавання. Тоді диференційна інформаційна міра різноманітності класу X_m^o стосовно найближчого класу X_c^o визначається згідно з працею [4] за модифікаційним критерієм Кульбака:

$$E_{K_m}^{(k)} = [P_{t,m}^{(k)} - P_{f,m}^{(k)}] * \log_2 \frac{P_{t,m}^{(k)}}{P_{f,m}^{(k)}}, \quad (1)$$

де $P_{t,m}^{(k)}$ – повна ймовірність правильного прийняття рішень на k -му кроці навчання СППР; $P_{f,m}^{(k)}$ – повна ймовірність неправильного прийняття рішень на k -му кроці навчання.

Оскільки для двохальтернативних рішень має місце

$$\begin{aligned} P_{t,m}^{(k)} &= p(\mu_m) * D_{1,m}^{(k)}(d) + p(\mu_m) * D_{2,m}^{(k)}(d), \\ P_{f,m}^{(k)} &= p(\mu_c) * \alpha_m^{(k)}(d) + p(\mu_c) * \beta_m^{(k)}(d), \end{aligned} \quad (2)$$

де $p(\mu_m)$, $p(\mu_c)$ – безумовна ймовірність появи реалізації класу X_m^o та найближчого (сусіднього) класу X_c^o відповідно; $\alpha_m^{(k)}(d)$, $\beta_m^{(k)}(d)$ – помилка першого та другого родів відповідно; $D_{1,m}^{(k)}(d)$, $D_{2,m}^{(k)}(d)$ – перша та друга достовірності рішень, що приймаються на k -му кроці навчання відповідно; d – дистанційна міра, яка визначає радіуси гіперсферичних контейнерів, побудованих в радіальному базисі простору Хеммінга.

Після відповідної підстановки виразів (2) у формулу (1) вона при припущенні, що $p(\mu_m) = p(\mu_c) = 0,5$ (згідно з принципом Лапласа-Бернуллі), та підстановці $\alpha_m^{(k)}(d) = 1 - D_{1,m}^{(k)}(d)$, $\beta_m^{(k)}(d) = 1 - D_{2,m}^{(k)}(d)$ набуває вигляду

$$E_{K_m}^{(k)} = \left[1 - (\alpha_m^{(k)}(d) + \beta_m^{(k)}(d)) \right] * \log_2 \left(\frac{2 - (\alpha_m^{(k)}(d) + \beta_m^{(k)}(d))}{\alpha_m^{(k)}(d) + \beta_m^{(k)}(d)} \right), \quad (3)$$

де $\alpha_m^{(k)}(d) \in [0, 1]$, $\beta_m^{(k)}(d) \in [0, 1]$ є помилками першого і другого родів відповідно, що обчислюється на кожному кроці навчання і залежать від значення радіусів псевдосфер d , кількості реалізацій кожного класу, а також подій, які визначають належність і неналежність реалізацій до відповідних контейнерів.

Після підстановки у формулу (3) оцінок точнісних характеристик

$$\alpha_m^{(k)} = \frac{K_2^{(k)}}{n_{\min}}, \quad \beta_m^{(k)}(d) = \frac{K_3^{(k)}}{n_{\min}},$$

де $K_2^{(k)}$ – кількість подій, при яких приймаються помилкові рішення про те, що «своя» реалізація $\{x_m^{(j)}\}$ не належить класу X_m^o , а $K_3^{(k)}$ – кількість подій, при яких «чужа» реалізація $\{x_c^{(j)}\}$ належить класу X_m^o і n_{\min} – мінімальний обсяг репрезентативної навчальної вибірки, отримаємо робочу формулу модифікації критерію Кульбака:

$$E = \frac{1}{n} \log_2 \left\{ \frac{2n + 10^{-r} - [K_2^{(k)} + K_3^{(k)}]}{[K_2^{(k)} + K_3^{(k)}] + 10^{-r}} \right\} * [n - (K_2^{(k)} + K_3^{(k)})], \quad (4)$$

де r – число цифр у мантисі значення критерію $E_m^{(k)}$;

При цьому доданок 10^{-r} вводиться у чисельник і знаменник логарифмічного відношення правдоподібності з метою усунення проблеми ділення на нуль.

Таким чином, формула (4) на відміну від формули (3) залежить від кількості реалізацій образу, що формують багатовимірну навчальну матрицю $\|x_{m,i}^{(j)}\|$. При цьому виникає задача оцінки впливу обсягу навчальної матриці на оцінку функціональної ефективності навчання СППР, оскільки від величини глобального значення КФЕ, що обчислюється в робочій (допустимій) області визначення його функції, залежать оптимальні геометричні параметри навчання, за якими будуються вирішальні правила.

Для проведення цього дослідження будемо користуватися нормованою модифікацією критерію Кульбака, яку подамо у вигляді

$$\bar{E}_{K,m}^{(k)}(d) = \frac{E_m^{(k)}(d)}{E_{\max}^{(k)}(d)} \in [0;1], \quad (5)$$

де $E_{\max}^{(k)}(d)$ – максимальне граничне значення КФЕ, що обчислюється в робочій області визначення функції(4).

Для спрощення представимо функцію (3) у вигляді

$$z(t) = \log_2 \left(\frac{2-t}{t} \right) * (1-t), \quad (6)$$

де

$$\begin{aligned} z(t) &= E_{K,m}^{(k)}(d); \\ t &= \alpha_m^{(k)}(d) + \beta_m^{(k)}(d); t \in [0;2]. \end{aligned} \quad (7)$$

Функцію (6) будемо досліджувати за такою схемою:

- 1) область визначення функції;
- 2) точки перетину з осями координат;
- 3) похідна і критичні точки;
- 4) інтервали зростання, спадання, точки екстремумів функції та її значення в цих точках;
- 5) друга похідна і дослідження функції на опуклість та увігнутість;
- 6) контрольні точки для уточнення графіка функції;
- 7) якщо не вдалося знайти глобального максимуму функції, то застосувати наближені числові методи [10] його знаходження.

Областю визначення функції є система рівнянь, яка формується з формули (7) та за умови, що аргумент логарифма в рівнянні (6) більше нуля:

$$\begin{cases} \left(\frac{2-t}{t}\right) > 0; \\ t \in [0; 2]. \end{cases} \quad (8)$$

З першого рівняння системи (8) випливає, що $t \in (0; 2)$, оскільки при його дослідженні на кожному інтервалі визначення, а саме при $t \in (-\infty; 0)$, $t \in (0; 2)$ та $t \in (2; +\infty)$ функція $f(t) = \left(\frac{2-t}{t}\right)$ поводить себе так, як показано на рис. 1:

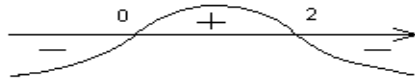


Рисунок 1 – Поведінка функції на інтервалах визначення

Отже, система рівнянь (8) набуває вигляду:

$$\begin{cases} t \in (0; 2), \\ t \in [0; 2]. \end{cases} \quad (9)$$

Таким чином, система рівнянь (9) має розв'язок $t \in (0; 2)$.

Точки перетину функції $z(t)$ з осями координат шукаються за умов: а) з координатою Oz при $t=0$; б) з координатою Ot при $z=0$. Умова а) не виконується при жодних значеннях z , а умова б) виконується лише при $t=1$, оскільки аргумент логарифму завжди більше нуля.

Інтервали зростання, спадання, точки екстремумів функції та її значення в цих точках знаходяться за першою похідною функції $z(t)$:

$$\begin{aligned} z'(t) &= [\log_2\left(\frac{2-t}{t}\right)]' * (1-t) = \log_2\left(\frac{2-t}{t}\right)' * (1-t) + (1-t)' * \log_2\left(\frac{2-t}{t}\right) = (\log_2(2-t)' - \log_2(t)') * (1-t) - \\ & - \log_2\left(\frac{2-t}{t}\right) = \left(-\frac{1}{(2-t)*\ln 2} - \frac{1}{t*\ln 2}\right) * (1-t) - \log_2\left(\frac{2-t}{t}\right) = (t-1) * \frac{2}{(2-t)*t*\ln 2} - \log_2\left(\frac{2-t}{t}\right) = \\ & = \frac{2*(t-1)}{(2*t-t^2)*\ln 2} - \log_2\left(\frac{2-t}{t}\right). \end{aligned}$$

Знаходження розв'язку рівняння $z'(t)=0$ викликає труднощі, хоча корінь в точці $t=1$ очевидний. За допомогою числових методів знайдемо решту коренів і, порівнявши значення функції при отриманих розв'язках, знайдемо глобальний максимум критерію Кульбака. Розглянемо поведінку функції в околі точки $t=1$. Оскільки $z'(1, 01) > 0$ та $z'(0,99) < 0$, то це свідчить про наявність локального або навіть глобального мінімуму функції у цій точці.

Друга похідна функції $z(t)$ має вигляд

$$\begin{aligned}
z''(t) &= (z'(t))' = \left[\frac{2 * (t - 1)}{(2 * t - t^2) * \ln 2} - \log_2 \left(\frac{2 - t}{t} \right) \right]' = \frac{1}{\ln 2} * \left(\frac{1}{2 - t} - \frac{1}{t} \right)' + \left(\frac{1}{(2 - t) * \ln 2} + \right. \\
&+ \frac{1}{t * \ln 2} \left. \right) = \frac{1}{\ln 2} * \left(\left(\frac{1}{2 - t} \right)' - \left(\frac{1}{t} \right)' \right) + \frac{2}{(2 * t - t^2) * \ln 2} = \frac{1}{\ln 2} * \left(\frac{1}{(2 - t)^2} + \frac{1}{t^2} \right) + \\
&+ \frac{2}{(2 * t - t^2) * \ln 2} = \frac{1}{\ln 2} * \left(\frac{t^2 + 4 - 4 * t + t^2}{(2 - t)^2 * t^2} \right) + \frac{1}{\ln 2} * \frac{2}{(2 - t) * t} = \\
&= \frac{1}{\ln 2} * \left(\frac{2 * t^2 - 4 * t + 4 - 2 * t^2 + 4 * t}{(2 - t)^2 * t^2} \right) = \frac{1}{\ln 2} * \left(\frac{4}{(2 - t)^2 * t^2} \right) > 0, \forall t \in (0, 2).
\end{aligned}$$

Оскільки $z''(t) > 0$ в кожній точці інтервалу $t \in (0, 2)$, то згідно з достатніми умовами опуклості і увігнутості функції на інтервалі $t \in (0, 2)$ графік функції $z(t)$ спрямовано опуклістю вниз. Для більш чіткого уявлення графіка функції в табл. 1 показано значення функції в контрольних точках із кроком 0,1.

Таблиця 1 – Табуляція функції $z(t)$ з кроком $\Delta t = 0,1$

Значення функції	3.8	2.5	1.7	1.2	0.0	0.0	0.3	0.5	2.5	3.8
Аргумент функції	0.1	0.2	0.3	0.4	1.0	1.1	1.3	1.4	1.8	1.9

Аналіз табл. 1 показує, що функція має лише один корінь на проміжку $t \in (0, 2)$, до того ж він є глобальним мінімумом.

Для знаходження максимуму функції $z(t)$ всю область визначення функції розбиваємо з кроком $\Delta t = 0,01$ та перевіряємо достатню умову знаходження кореня на кожному з інтервалів $[t_i, t_{i+1}]$. Якщо умова виконується, то за допомогою методу Ньютона знаходимо корінь із заданою точністю. Фрагмент реалізації алгоритму має вигляд

```

for(t = 0.01; t < 2.00; t+=0.01)
{
  if(F(t)*F(t+0.01)<=0){
    x0=t;
    do {
      x1=Count(x0);
      if(abs(x1-x0)<0.0001){
        boo=1;
      }
      x0=x1;
      i++;
    }while(boo);
  }
}

```

На рис. 2 показано графік функції $z(t)$.

Як видно з рис. 2, функція (6) не є взаємодозначною. При цьому максимум функції досягається на краях області визначення при $t \rightarrow 0$ та $t \rightarrow 2$.

Таким чином, згідно з формулою (7) маємо

$$\alpha_m^{(k)}(d) + \beta_m^{(k)}(d) \rightarrow 0;$$

$$\alpha_m^{(k)}(d) + \beta_m^{(k)}(d) \rightarrow 2.$$

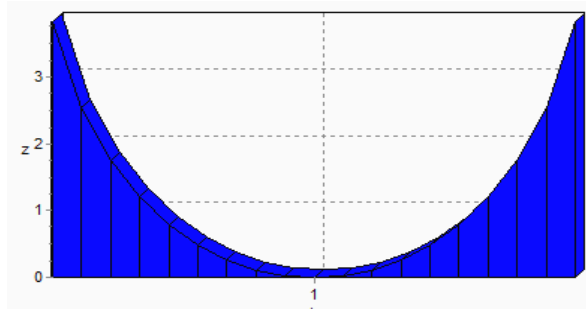


Рисунок 2 – Графік залежності функції $z(t)$ від аргументу t

Оскільки повна ймовірність правильного прийняття рішень СППР буде тим більшою, чим меншими будуть значення помилок $\alpha_m^{(k)}(d)$ і $\beta_m^{(k)}(d)$, то необхідно розглядати випадок, при якому $\alpha_m^{(k)}(d) + \beta_m^{(k)}(d) \rightarrow 0$. Для більш чіткого уявлення, як веде себе функція $z(t)$ в околі точки $t = 0$, знайдемо границю функції при $t \rightarrow 0$.

$$\begin{aligned} z(t) &= \log_2\left(\frac{2-t}{t}\right) * (1-t) = \log_2(2-t) * (1-t) - \log_2(t) * (1-t) = \\ \lim_{t \rightarrow 0} z(t) &= \log_2(2-0) * (1-0) - \lim_{t \rightarrow 0} \log_2(t) * (1-t) = 1 - \log_2(t) * (1-t) = 1 + \log_2(t) * (t-1) = \\ &= 1 + (-\infty) * (-1) = +\infty. \end{aligned}$$

Зробимо припущення, що функціональна ефективність системи буде високою, якщо одна реалізація буде неправильно класифікована в процесі навчання СППР.

Оцінку функціональної ефективності за мірою Кульбака розглянемо на прикладі технологічного процесу вирощування монокристалів. При цьому формування навчальних матриць здійснювалося за даними, одержаними при вирощуванні за методом Чохральського на установці «РОСТ». Початкова навчальна матриця кожного класу складалася із 61 вектора-реалізації, які мали по 45 первинних і вторинних ознак розпізнавання.

Розглянемо поведінку функції $z(t)$ в околі точки $t = 0$ при досить малих значеннях аргументу t (табл. 2).

Таблиця 2 – Поведінка функції $z(t)$ в околі точки $t = 0$

Значення функції	Знак першої похідної	Аргумент функції
10,95	–	0,001
8,59	–	0,005
8,09	–	0,007
7,71	–	0,009
7,41	–	0,011
6,94	–	0,015
6,74	–	0,017
6,57	–	0,019

Оскільки $\alpha^{(k)} + \beta^{(k)} = \frac{K_2^{(k)}}{n_{\min}} + \frac{K_3^{(k)}}{n_{\min}} = \frac{K_2^{(k)} + K_3^{(k)}}{n_{\min}}$, то за умови, що лише одна реалізація буде класифікована неправильно, даний вираз буде

набувати значення $\alpha^{(k)} + \beta^{(k)} = \frac{1}{61} = 0,016$, і це відповідає лише 2 %

кількості реалізацій, які не були класифіковані правильно, тоді згідно з табл. 2 максимальне значення критерію функціональної ефективності буде дорівнювати 6,55.

Визначення мінімального обсягу n_{\min} репрезентативної навчальної вибірки здійсимо за умови отримання прийнятних з практичних міркувань статистичної похибки та оперативності алгоритму його обчислення. Ці вимоги є суперечливими, що обумовлює компромісний характер розв'язання задачі. У праці [7] показано, що найбільш придатним для формування навчальних матриць є інтервал приблизно від 40 до 90 випробувань – дискретних векторів – реалізацій, які формуються шляхом опитування датчиків. У табл. 3 наведено максимальні значення КФЕ (5) залежно від кількості реалізацій за умови наявності однієї некласифікованої реалізації образу.

Таблиця 3 – Результати залежності КФЕ від кількості реалізацій

Значення функції	Кількість реалізацій	Кількість некласифікованих реалізацій, %
7,41	90	1,1
7,28	80	1,2
7,04	70	1,4
6,84	60	1,6
6,49	50	2,0
6,14	40	2,5

На рис. 3 зображено графіки залежності КФЕ за нормованим критерієм Кульбака від радіуса контейнера класу X_1^o , отриманого при різних потужності навчальної вибірки в процесі навчання з паралельної оптимізацією системи контрольних допусків на ознаки розпізнавання.

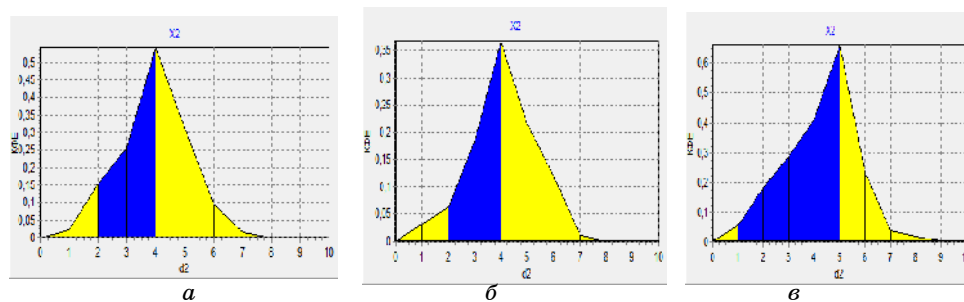


Рисунок 3 – Графіки залежності КФЕ від радіуса контейнера класу X_1^o після паралельної оптимізації для: а) 40 реалізацій; б) 50 реалізацій; в) 60 реалізацій

Аналіз рис. 3 показує, що збільшення кількості реалізацій у навчальній матриці впливає на геометричні параметри вирішальних правил, які мають тенденцію до збільшення: для випадку а) оптимальний радіус контейнера класу X_1^o дорівнює $d_1^* = 4$ (у кодових одиницях), для випадку б) – $d_2^* = 4$ і для випадку в) – $d_2^* = 5$. Побудовані вирішальні правила не є безпомилковими за навчальною матрицею, що вимагає оптимізації інших параметрів навчання [7].

ВИСНОВКИ

1. На основі проведеного дослідження функції інформаційної міри Кульбака встановлено її область визначення, інтервали зростання, спадання, точки екстремумів функції та її значення в цих точках. Показано, що міра Кульбака набуває мінімального значення в області визначення її функції за умови $\alpha_m^{(k)}(d) + \beta_m^{(k)}(d) = 1$.

2. Встановлено, що збільшення кількості векторів-реалізацій класу розпізнавання призводить до збільшення максимального значення критерію функціональної ефективності і впливає на геометричні параметри вирішальних правил.

ОЦЕНКА ФУНКЦИОНАЛЬНОЙ ЭФФЕКТИВНОСТИ ОБУЧЕНИЯ АВТОМАТИЗИРОВАННОЙ СИСТЕМЫ УПРАВЛЕНИЯ ТЕХНОЛОГИЧЕСКИМ ПРОЦЕССОМ

А. С. Довбиш, Д. Б. Берест,

Сумский государственный университет, г. Сумы

В рамках информационно-экстремальной технологии рассматривается использование нормированной логарифмической статистической информационной меры Кульбака для оценивания функциональной эффективности системы управления выращиванием скинтилляционных монокристаллов. Исследовано влияние количества реализаций на информационный критерий.

Ключевые слова: *информационный критерий Кульбака, интеллектуальная система управления, признак распознавания, критерий функциональной эффективности, скинтилляционный монокристалл.*

INFLUENCE OF VECTOR – REALIZATIONS NUMBER ON FUNCTIONAL EFFICIENCY OF RECOGNITION SYSTEM LEARNING BY KULLBACK'S MEASURE

A. S. Dobbish, O. B. Berest,

Sumy State University, Sumy

The usage of rated logarithmical statistical Kullback's measure within information - extreme technology of scintillation monocrystal growth intellectual control system is considered in this article. The influence of realization number on functional efficiency criterion is investigated.

Key words: *information measure, intellectual control system, recognition sign, functional efficiency criterion, scintillation monocrystal.*

СПИСОК ЛІТЕРАТУРИ

1. Bernhard Schölkopf. Learning with Kernels. Support Vector Machines, Regularization, Optimization, and Beyond /Bernhard Schölkopf, Alexander J. Smola. – MIT Press, Cambridge: MA, 2002.
2. H. Witten. Data Mining: Practical Machine Learning Tools and Techniques (Second Edition) / I. H. Witten, E. Frank. – Morgan Kaufmann, 2005.
3. De Maesschalck R. The Mahalanobis distance / R. De Maesschalck, D. Jouan-Rimbaud, D. L. Massart // Chemometrics and Intelligent Laboratory Systems. – 2000. – Vol. 50. – P. 1–18.
4. Луценко Е. В. Интеллектуальные информационные системы: учебное пособие / Е. В. Луценко. – Краснодар, КубГАУ, 2004. – 633 с.
5. Зайченко Ю. П. Основы проектирования интеллектуальных систем / Ю. П. Зайченко// – К.:Видавничий дім "Слово", 2004. – 352 с.
6. Евменов В. П. Интеллектуальные системы управления / В. П. Евменов. - М.: Книжный дом «Либроком», 2009. – 300 с.
7. Довбиш А. С. Основы проектирования интеллектуальных систем: навчальний посібник / А. С. Довбиш. – Суми: Вид-во СумДУ, 2009.– 171 с.
8. Краснопопосовський А. С. Інформаційний синтез інтелектуальних систем керування: Підхід, що ґрунтується на методі функціонально-статистичних випробувань / А. С. Краснопопосовський. – Суми: Видавництво СумДУ, 2004. – 261 с.
9. Шелехов І. В. Оцінка інформативності ознак розпізнавання за методом функціонально-статистичних випробувань / І. В. Шелехов // Вісник Сумського державного університету.– 2006.–№10(94). – С. 51-58.
10. Любчак В. О. Методи та алгоритми обчислень: навчальний посібник / В. О. Любчак, Л. Д. Назаренко. – Суми: Вид-во СумДУ, 2008. – 313 с.

Надійшла до редакції 19 липня 2012 р.