УДК 681.3.019

М.М. БИКОВ, М.М. ФІЛАТОВА
Вінницький національний технічний університет

ВИЗНАЧЕННЯ ХАРАКТЕРИСТИК ПОТЕНЦІАЛЬНИХ КОДІВ ЗА МОДЕЛЯМИ РАНГОВИХ КОНФІГУРАЦІЙ

В статті описані геометрична, комбінаторна, алгебрична та топологічна моделі рангових конфігурацій об'єктів системи в просторі ознак, розроблені з метою визначення характеристик двійкових кодів вказаних конфігурацій, що містять інформацію як про самі об'єкти, так і про ранги відстаней між ними. Такі коди названі потенціальними за аналогією з електричним полем, в якому величини зарядів визначають напруженість поля між ними. Проведена оцінка адекватності представлених моделей.

Ключові слова: представлення і обробка інформації, інтелектуальні системи, ідентифікація станів системи, рангові конфігурації, моделі рангових конфігурацій, кодування, потенціальний код, характеристики кодів.

N. M. BYKOV, M. M. FILATOVA
Vinnitsa National Technical University

DETERMINING OF THE POTENTIAL CODES CHARACTERISTICS USING MODEL OF RANK CONFIGURATIONS

Abstract - The aim of the article is representation of models which were developed for determining of the characteristics of potential codes (DRP-codes - codes that preserves rank distances between objects encoded in the parametric space).

The paper presents a unified method for describing and processing of knowledge about the state of the control systems using rank transform of their descriptions made in different parametric spaces: deterministic, probabilistic, fuzzy and others. This problem is solved on the basis of such fundamental concepts as the degree (rank)of the distance between the classes of objects in parametric space. The definition of rank configuration is presented, the efficient low computational cost binary encoding method is described, which determines both the objects (states), and ranks about the distances between them. This code is called potential or DRP-code (code that saves the distances ranks). The method for constructing such a code was defined and its characteristics obtained from the presented models ranked configurations - geometric, algebraic, combinatorial and topological. Adequacy of the proposed models was proved.

The results can be used to unify the algorithms of information description and processing in the control systems of different purposes, and as a result to improve the efficiency of decision-making in such systems.

Keywords: representation and processing of information, intelligent systems, identification of system states, rank configuration, models ranked configurations, encoding, a potential code, codes characteristics.

Вступ

Рішення на керування в системах управління приймають на основі інформації про відстань між поточним і цільовим станами об´єкта. Для уніфікації методів і алгоритмів представлення, передачі і обробки інформації в системах прийняття рішень з різними методами опису вхідних даних – детерміністичним, імовірнісним, нечітким, наближеним, ймовірним та ін., автори запропонували в попередніх роботах описувати їх стани за допомогою рангових конфігурацій станів (елементів) системи управління. При цьому для досягнення поставленої мети виникла необхідність в представленні цих конфігурацій потенціальними кодами (DRP-codes – кодами, що зберігають ранги відстаней). На відміну від загальноприйнятих підходів, в яких кодування виконується з урахуванням тільки вимоги їх розрізнення та перешкодостійкості, в запропонованому способі двійкового представлення станів систем інформація про відстані між ними міститься в самих кодах. Такий підхід дає можливість величину рангу відстані визначати безпосередньо шляхом порівняння кодів за допомогою логічної операції. Під час реалізації такого підходу виникає низка задач, пов'язаних з необхідністю визначення таких характеристик кодів, як їх повнота, розрядність, перешкодостійкість, інформаційна ємність та конфіденційність. Для визначення вказаних характеристик були розроблені моделі рангових конфігурацій, які є предметом розгляду в даній статті.

Аналіз стану досліджень та публікацій

Зародженню ідеї узагальненого представлення інформації про стан об'єктів на основі такого фундаментального поняття, як ступінь (ранг) віддаленості між станами об'єктів у параметричному просторі послугував доведений у науково-технічній літературі факт, що ідентифікацію станів об'єкта контролю і управління і оптимізацію його роботи можна здійснити на основі інформації про відстані між його станами в параметричному просторі. Доведення цього факту для просторів з детерміністичним і імовірнісними описами можна знайти, наприклад, в [6, 11], а для просторів з нечіткими і наближеними описами – в [7–9]. Розвиваючи далі цю ідею, автори показали [1–3], що важливою інформацією для реалізації процедури ідентифікації станів і оптимізації роботи системи є не самі відстані між цими станами, а їх рангові відношення. Поняття рангової конфігурації введено для опису рангового відношення відстаней між станами системи.

Ефективність застосування рангових конфігурацій для комп'ютеризованого аналізу даних і прийняття рішень залежить від методу їх опису. Тому був розроблений ефективний щодо обчислювальних затрат двійковий метод кодування, який визначає як самі об'єкти (стани), так і інформацію про ранги відстаней між ними. Такий код названо потенціальним або DRP-кодом (кодом, що зберігає ранги відстаней). Під час розробки даного коду були використані моделі рангових конфігурацій, представлені в даній роботі.

Моделі рангових конфігурацій як засіб визначення характеристик потенціальних кодів

Визначення 1. Ранговою конфігурацією $\Re$ простору m об´єктів називають множину $(m-1)$-елементних підмножин, елементами цих підмножин є ранги відстаней, інцидентних одному і тому ж об´єкту.

Рангову конфігурацію можна отримати, якщо розмістити відстані між об'єктами у просторі параметрів в порядку зростання, а потім присвоїти їм ранговий номер.

Визначення 2. Код B, який зберігає ранги відстаней (*DRP*-код, по іншому потенціальний код), є відображення $i \to B_i$ множини $M = \{1, 2, ..m\}$ в множину $\{0,1\}^n$ двійкових послідовностей довжини n таке, що

$$\underset{i,j}{\forall}(R(d_{ij}) = r \Rightarrow R(h_{ij}) = r), \ \ r = \overline{1, m_r}, \ \ i, j \in M \ . \qquad (1)$$

У виразі (1) $R(d_{ij})$ – ранг відстаней d_{ij} між об’єктами i и j в просторі об'єктів; $R(h_{ij})$ – ранг відстані h_{ij} в просторі двійкових кодів; r – ціле число, конкретне значення рангу; m_r – максимальна величина рангу. Далі під кодованими об’єктами розумітимемо елементи систем.

Не вникаючи у властивості початкового простору об’єктів, які забезпечують можливість кодування, можна сказати, що на даному етапі мова йтиме про двійкове зображення за допомогою *DRP*-коду метричних просторів (станів систем) з заданим відношенням строгого порядку на лінійно впорядкованій множині відстаней між об’єктами, або просторів з відношенням нестрогого порядку, коли ранги рівних відстаней інцидентні одному й тому самому об’єкту.

Потрібно зразу зауважити, що потенціальний код може проектуватися для двох різних застосувань. Перше з них пов’язано з задачами ідентифікації станів систем (наприклад, задачі кластеризації), друге – з задачами розпізнавання рядків символів, перешкодостійкої передачі інформації та ін. В першому випадку виконання аксіоми ідентичності $R(d_{ii}) = 0$ при обчисленні рангів відстаней між кодами станів не потрібно, оскільки відстані між одними ж і тими ж точками не обчислюються. У даному розгляді ранг відстані $R(h_{ij}) = 0$ між двійковими словами b_i і b_i DRP-коду знаходиться за допомогою мультиплікативної операції логічного множення AND і визначається виразом:

$$R(h_{ij}) = \log_2(b_i \Lambda b_j) \, , \qquad (2)$$

де символом Λ позначена операція AND.

В другому випадку код повинен проектуватися з урахуванням адитивної операції XOR (“виключне АБО”).

Розглянемо моделі рангових конфігурацій, необхідні для визначення характеристик потенціального коду у першому випадку. Для їх представлення були розроблені геометричні, алгебричні, комбінаторні та топологічні моделі. До геометричних можна віднести багатомірні симплекси, до топологічних – граф. На рис.1 зображено рангову конфігурацію у вигляді тривимірного симплекса і графа.

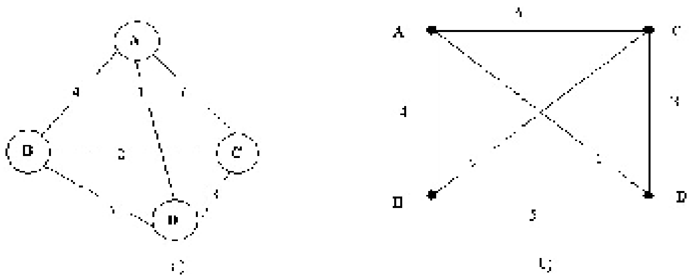


Рис. 1. Геометричні моделі рангової конфігурації: а – тривимірний симплекс; б –повний регулярний граф

Симплекс та граф дають зорову уяву про рангову конфігурацію. На рис.1 вершини **A**, **B**, **C**, **D** позначають кодовані об´єкти, а числа на ребрах симплекса і дугах графа – інцидентні їм ранги.

Алгебрична модель дає можливість компактно описати рангову конфігурацію у відповідності з **Визначенням 1** у вигляді набору підмножин рангів. Наприклад, рангову конфігурацію рис.1 можна записати в алгебричному виді як

$$K_4 = \{\{1, 4, 6\}, \{2, 4, 5\}, \{2, 3, 6\}, \{1, 3, 5\}\} \, . \qquad (3)$$

Комбінаторні моделі конфігурацій можна подати у вигляді $t-$схем і блок-схем, які широко використовуються в теорії кодування [5], а також перестановок на матриці суміжностей рангів графа. Детальний розгляд $t-$схеми і її різновиду блок-схеми показав, для моделювання рангових конфігурацій є непридатними, оскільки не виконується одна з умов, що характеризує можливість застосування такої моделі [3,4]. Тому в якості комбінаторних моделей в роботі використані матриця інцидентності графа (рис. 2а), та матриця суміжностей рангів цього графа (рис. 2б).

Ранги

Об´єкти

	6	5	4	3	2	1
A	1	0	1	0	0	1
B	0	1	1	0	1	0
C	1	0	0	1	1	0
D	0	1	0	1	0	1

а)

	A	B	C	D
A		4	6	1
B	4		2	5
C	6	2		3
D	1	5	3	

б)

Рис. 2. Комбінаторні моделі рангової конфігурації: а) – матриця інцидентностей графа рис.1, б; б) – матриця суміжностей рангів для графа рис.1,б.

Вказані комбінаторні моделі дозволяють визначити такі основні характеристики потенціальних кодів, як розрядність і повнота.

З матриці інцидентностей 2а) зразу видно, що розрядність потенціального коду n за умови визначення рангів логічною операцією "І" дорівнюватиме кількості рангів:

$$n = \frac{m(m-1)}{2}, \tag{4}$$

де m – кількість кодованих об'єктів. Наприклад, для m =4 $n = \frac{4(4-1)}{2} = \frac{12}{2} = 6$, для m =5 маємо $n = 10$ і т.д.

Матриця інцидентностей також дозволяє розробити **алгоритм 1** кодування об´єктів *DRP*-кодом, який складається з двох етапів:

1) Заповнення матриці розміром $m \times n$ нулями;

2) Перегляд матриці по стовпцях (рангах) і установка одиниці в тих стрічках, що відповідають об'єктам, інцидентних цим рангам.

На рис. 2а рядки матриці суміжності графа утворюють кодові слова шуканого для заданої рангової конфігурації DRP -коду. Як приклад визначимо ранг відстані між кодовими словами символів А і D, для чого скористаємося виразом (2):

$$R(h_{AD}) = \log_2(b_A \Lambda b_D) = \log_2(101001 \Lambda 010101) = \log_2 2^1 = 1.$$

На практиці операція логарифмування для визначення точного рангу не потрібна, оскільки відстані між кодовими словами ранжуються операцією AND.

З моделі рис.2а також видно, що за своєю формою *DRP*-код є перестановочним з постійною вагою, тому для визначення його перешкодостійкості можна скористатися відомими [4] для цих кодів залежностями. Вірогідність P_n невиявлення однієї помилки в кодовому слові (сумісна вірогідність перетворення однієї одиниці в нуль і одного нуля в одиницю) для DRP -коду рівна:

$$P_n = \binom{m-1}{1} p(1-p)^{m-2} \binom{n-m+1}{1} p(1-p)^{n-m}, \tag{5}$$

де m - кількість слів у коді. Для коду з прикладу на рис. 2а маємо:

$$P_n = \binom{3}{1} p(1-p)^2 \binom{4}{1} p(1-p)^3.$$

Прийнявши значення $p = 1 \cdot 10^{-4}$, одержимо $P_n = 12 \cdot 10^{-8}$. Оскільки для потенціального коду одиниця повинна бути невиявлена одночасно в двох кодових словах, то $P_n = 12 \cdot 10^{-16}$.

Модель рангової конфігурації у вигляді матриці суміжності рангів дозволяє дійти висновку, що множину конфігурацій можна представити групою на наборі перестановок. Ця властивість дозволяє кодувати не тільки символи, але і конфігурації, і запам'ятовувати інформацію про них в згорнутому вигляді.

Матриця суміжностей рангів також дозволяє без значних складностей визначити потужність K_m множини конфігурацій для заданої кількості об'єктів m:

Теорема 1. Потужність множини рангових конфігурацій для непоміченого (m-1)-вимірного симплекса дорівнює:

$$K_m = \frac{(m(m-1)/2)!}{m!} = \frac{n!}{m!} \tag{6}$$

Доведення. Кількість n різних рангів відстаней для (m - 1)-вимірного симплекса (рис. 1,б) дорівнює

$$n = \frac{m(m-1)}{2}.$$

З цього випливає, що загальна кількість перестановок цих рангів дорівнює $n!$, тобто $[m(m-1)/2]!$. Вони дають $n!$ сполучень з $(m-1)$ рангів по n в рядках матриці суміжностей. Серед них існує $m!$ однакових сполучень, що відповідають всім можливим перестановкам m рядків вказаної матриці.

Отже, потужність множини конфігурацій (m-1)-вимірного симплекса дорівнює:

$$K_m = \frac{(m(m-1)/2)!}{m!} = \frac{n!}{m!}.$$

Теорема доведена.

Для помічених об'єктів потужність множини рангових конфігурацій кількості перестановок рангів, що окремого доведення не потребує:

$$K_m = n!. \tag{7}$$

Зростання кількості конфігурацій за умови зростання кількості m символів відповідає закону "комбінаторного вибуху". Наприклад, при $m = 4$ кількість рангових конфігурацій K_4 непоміченого симплексу відповідно до виразу (6)

$$K_4 = \frac{6!}{4!} = \frac{1\cdot 2\cdot 3\cdot 4\cdot 5\cdot 6}{1\cdot 2\cdot 3\cdot 4} = 30,$$

для п'яти символів $K_5 = 30240$, для шести символів $K_6 \approx 1{,}8\cdot 10^9$. Для поміченого симплексу кількість конфігурацій зростає ще швидше, що видно з виразу (7). Цей факт свідчить про велику інформаційну потужність *DRP*-коду коли користувача цікавить інформація про відносини між елементами даних.

Пропускну спроможність каналу можна збільшити, передаючи не саму рангову конфігурацію, а її код, відповідний номеру рангової перестановки. Секретність коду C_R при цьому можна забезпечити вибором одного з можливих порядків перестановки рангів на матриці суміжності і визначиться вірогідністю вгадування номера цього порядку:

$$C_R = \frac{1}{n!}. \tag{9}$$

Адекватність моделей рангових конфігурацій, застосованих до визначення характеристик потенціальних точок, доведена аналітично, тому проведення експерименту з даної метою буде надлишковим.

Для доведення повноти кодів в другому випадку (для тих застосувань, які вимагають існування аксіоми ідентичності), авторами розроблена топологічна інтервальна модель рангової конфігурації, приклад якої наведений на рис.3.

Дана модель представляє собою лінійний відрізок довжиною 2^q, де q – вибрана розрядність коду, поділений на одиничні інтервали, границі інтервалів відповідають цілим числам з діапазону $[0…2^q]$. На цьому відрізку укладені інтервали, які в промасштабованому вигляді відповідають рангам конфігурації. Для моделі справедливими є такі аксіоми.

Аксіоми різницевої операції:

- аксіома симетрії;

$$\forall_{i,j\in M} \; (b_i \oplus b_j) = (b_j \oplus b_i) = \left| b_i \oplus b_j \right| = h_{ij}; \tag{10}$$

- аксіома оберненості операції.

$$b_i \oplus h_{ij} = b_j; \quad b_j \oplus h_{ij} = b_i \tag{11}$$

Аксіоми рангових інтервалів:

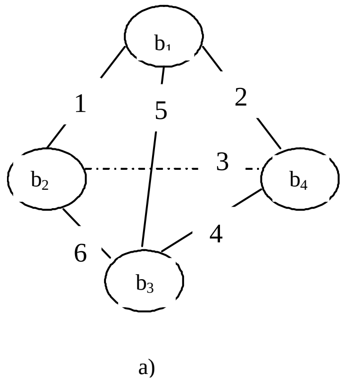
- аксіома жорсткості інтервалів в циклі

$$h_{ij}(r_k) + h_{ij}(r_{k+1}) = h_{ij}(r_k + r_{k+1}) \tag{12}$$

- аксіома трикутника для рангів в циклі

$$r_k + r_{k+1} \geq r_c \tag{13}$$

де r_k і r_{k+1} – ранги k-го і k+1-го інтервалів на моделі (ранги сусідніх інтервалів), r_c – ранг інтервалу, що замикає. Дані ранги позначені дугами на топологічному графові моделі (рис. 3б). Поняття циклу на інтервальній моделі відповідає поняттю циклу на топологічному графові.



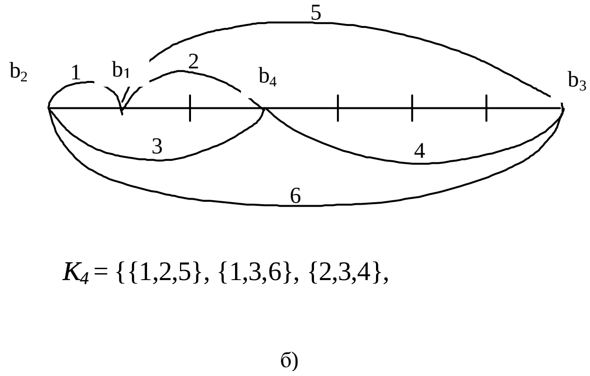


Рис. 3. Інтервальна модель рангової конфігурації

Аксіоми різницевої операції:

- аксіома симетрії;

$$\forall_{i,j \in M} \quad (b_i \oplus b_j) = (b_j \oplus b_i) = \left| b_i \oplus b_j \right| = h_{ij}; \tag{10}$$

- аксіома оберненості операції.

$$b_i \oplus h_{ij} = b_j; \quad b_j \oplus h_{ij} = b_i \tag{11}$$

Аксіоми рангових інтервалів:

- аксіома жорсткості інтервалів в циклі

$$h_{ij}(r_k) + h_{ij}(r_{k+1}) = h_{ij}(r_k + r_{k+1}) \tag{12}$$

- аксіома трикутника для рангів в циклі

$$r_k + r_{k+1} \geq r_c \tag{13}$$

де r_k і r_{k+1} – ранги *k*-го і *k*+1-го інтервалів на моделі (ранги сусідніх інтервалів), r_c – ранг інтервалу, що замикає. Дані ранги позначені дугами на топологічному графові моделі (рис. 3б). Поняття циклу на інтервальній моделі відповідає поняттю циклу на топологічному графові.

На основі інтервальної моделі сформульована і доведена наступна теорема.

Теорема 2: Допустимими є тільки ті рангові конфігурації, для яких виконуються аксіоми (12) і (13).

Як слідує з теореми 2, інтервальна модель тільки показує, що кількість рангових конфігурацій за умови використання операції XOR є обмеженою, що не дозволяє побудувати повний потенціальний код. Але вона не дає аналітичного виразу для визначення кількості можливих конфігурацій і методу побудови потенціального коду. Для оцінки адекватності цієї моделі був розроблений алгоритм, що базується на повному переборі всіх можливих сполучень з *m* по 2^q кодів цілих двійкових чисел розрядністю *q* на інтервалі [0…2^q]. Машинний експеримент підтвердив, що DRP-код у цьому випадку є неповним. Наприклад, кількість дозволених рангових конфігурацій для 3-вимірного симплекса (4 кодованих об’єкти) дорівнює 7 із 30 можливих, і коефіцієнт повноти коду відповідно дорівнює 7/30, тобто приблизно 23,3%.

Тому автори пропонують у цьому випадку використовувати потенціальний код, спроектований на базі використання модифікованої логічної операції AND “з блокуванням за порогом 1”. Вона складається з логічної операції AND над потенціальними кодами об’єктів, отриманих **алгоритмом 1**, з наступним виконанням операції XOR над результатом. Якщо сума одиниць в результаті більша 1, то результат блокується до нуля, що дає можливість забезпечити виконання аксіоми ідентичності при визначенні рангу відстані між одними і тими ж об’єктами.

Висновки

В роботі представлено уніфікований метод опису і опрацювання знань про стани системи управління за допомогою рангового перетворення їх описів, виконаних у різних параметричних просторах: детерміністичному, імовірнісному, нечіткому та ін. Цю проблему розв’язано на основі такого фундаментального поняття, як ступінь (ранг) віддаленості між станами об’єктів у параметричному просторі. Наведено визначення рангової конфігурації простору станів, описано ефективний щодо обчислювальних затрат двійковий метод кодування, який визначає як самі об’єкти (стани), так і інформацію про ранги відстаней між ними. Такий код названо потенціальним або DRP-кодом (кодом, що зберігає ранги відстаней). Визначено метод побудови такого коду і отримано його характеристики на основі представлених моделей рангових конфігурацій – геометричних, алгебричних, комбінаторних та топологічних. Доведена адекватність запропонованих моделей. Результати роботи можуть бути застосовані для підвищення ефективності процесу прийняття рішень в системах різного призначення та для підвищення перешкодостійкості процесу передачі інформації в каналах зв’язку.

Література

1. Быков Н.М. Обобщенный метод принятия решений в системах управления и контроля / Н.М. Быков, О.С. Пастушенко, О.С.Коберский // Контроль и управление в технических системах: Тез. докладов 3-й Междунар. научн.-техн. конф. Ч. 1. – Винница, 1995. – С. 223–224.
2. Биков М.М. Кластеризація даних з використанням потенціальних кодів / М.М. Биков, І.В. Кузьмін, А.І. Яковенко // Вісник Вінницького політехнічного інституту. – 2001. – № 6. – С. 61–64.
3. Биков М.М. Універсальний метод представлення інформації в інтелектуальних еволюційних системах / М.М. Биков // Відбір і обробка інформації. – 2006. – Вип. 24 (100). С. 35–42.
4. Злотник Б.М. Помехоустойчивые коды в системах связи / Злотник Б.М. – М. : Радио и связь, 1989.– 232 с.
5. Камерон П. Теория графов, теория кодирования и блок-схем / Камерон П., Линт Дж – М. : Мир, 1980. – 144 с.
6. Штейнберг С. Идентификация в системах управления / Штейнберг С.–М. : Энергоатомиздат, 1987.–196 с.
7. Kosko B. Fuzzy entropy and conditioning/ Kosko B. // Information sciences – 1966. – 2. – P.165–174.
8. Kruskal J. Multidimensional scaling by optimizing goodness-of-fit to a nonmetric hypothesis// Psychometrika – 1964. – 29. – P. 1–28, 115–129.
9. Pawlak Z., Skowron A. Rough membership functions // Advances in the Dempster-Shafer theory of evidence. Ed. Yager R., Fedrizzi M., Kacprzyk J. - New York: John Wiley & Sons, 1994. – P. 251–271.
10. Preparata F., Nievergelt J. Difference-Preserving codes // IEEE Trans. Information Theory – 1974. – 20. – P. 643–649.
11. Tou J., Gonzalez R. Pattern Recognition Principle – London-Amsterdam-Don Mill. : Addison-Wesley, 1974. – 411 p.

References

1. Bykov N.M. Obobshchennyi metod pryniatyia reshenyi v systemakh upravlenyia y kontrolia, Kontrol y upravlenye v tekhnycheskykh systemakh: Tez. dokladov 3–y Mezhdunar. nauchn.-tekhn. konf. Ch. 1, Vynnytsa, 18 – 21 sentiabria, 1995, pp. 223 – 224. [in Russian]
2. Bykov M.M. Klasteryzatsiia danykh z vykorystanniam potentsialnykh kodiv, Visnyk Vinnytskoho politekhnichnoho instytutu, 2001, No. 6, pp. 61-64. [in Ukrainian]
3. Bykov M.M. Universalnyi metod predstavlennia informatsii v intelektualnykh evoliutsiinykh systemakh, Vidbir i obrobka informatsii, 2006, No. 24(100), pp. 35-42. [in Ukrainian]
4. Zlotnyk B.M. Pomekhoustoichyvye kody v systemakh sviazy, M.: Radyo y sviaz, 1989, 232 p. [in Russian]
5. Kameron P. Teoryia hrafov, teoryia kodyrovanyia y blok-skhem, M.: Myr, 1980, 144 p. [in Russian]
6. Shteinberh S. Ydentyfykatsyia v systemakh upravlenyia, M.: Эnerhoatomyzdat, 1987, 196 p. [in Russian]
7. Kosko B. Fuzzy entropy and conditioning, Information sciences,1966, No. 2, pp.165-174.
8. Kruskal J. Multidimensional scaling by optimizing goodness-of-fit to a nonmetric hypothesis, Psychometrika, 1964. Vol. 29, pp. 1-28, 115- 129.
9. Pawlak Z., Skowron A. Rough membership functions, Advances in the Dempster-Shafer theory of evidence. Ed. Yager R., Fedrizzi M., Kacprzyk J. New York: John Wiley & Sons, 1994, pp. 251-271.
10. Preparata F., Nievergelt J. Difference-Preserving codes, IEEE Trans. Information Theory, 1974, No. 20, pp. 643-649.
11. Tou J., Gonzalez R. Pattern Recognition Principle – London-Amsterdam-Don Mill.: Addison-Wesley, 1974, 411 p.

Рецензія/Peer review : 20.7.2013 р. Надрукована/Printed :26.9.2013 р.
Рецензент: зав. кафедри захисту інформації, ВНТУ, д.т.н., проф. В.А. Лужецький