

ЕКОНОМІКО-МАТЕМАТИЧНЕ МОДЕЛЮВАННЯ

Ярослав НИКОЛАЙЧУК, Володимир ШАРЯК

РЕАЛІЗАЦІЯ БАГАТОРІВНЕВОЇ РЕКУРЕНТНОЇ БАЗИ ДАНИХ В БАЗИСІ ГАЛУА МЕРЕЖЕВОГО МАРКЕТИНГУ КОРПОРАЦІЇ «ТЯНЬ ШІ»

Викладено принципи побудови рекурентної бази даних в базисі Галуа, проведено аналіз ефективності кодування ідентифікаційних даних на прикладі ієрархічно-реляційної бази даних китайської корпорації «ТЯНЬ ШІ».

Потужний розвиток сучасних інформаційних систем потребує відповідного розвитку теоретичних засад реалізації бази даних (БД) та бази знань (БЗ). Особливо великі масиви даних формуються, підлягають зберіганню та оперативно циркулюють в каналах зв'язку міжнародних корпорацій, до яких належить міжнародна китайська корпорація «ТЯНЬ ШІ».

Аналіз організаційної структури БД, якими користуються великі міжнародні корпорації показує, що найбільш широким застосуванням для кодування ідентифікаційних даних є багаторівнева ієрархічно-реляційна структура [1, 2, 3, 4]. При цьому очевидно є велика надлишковість кодування та їх низька захищеність, що аналізувалося в роботах [5, 6], де на основі ієрархічних структур обґрунтовані принципи та показана ефективність кодування ІД на основі багатомірної рекурентної організації БД в базисі Галуа. Наприклад, схема діяльності інфраструктури корпорації «Тянь Ші» зображена на рис. 1.

Аналіз характеристики кодування ІД в ієрархічно-реляційній БД структурі багаторівневого маркетингу корпорації «ТЯНЬ ШІ» свідчить, що функції корпорації «ТЯНЬ ШІ» реалізуються п'ятьма ієрархічними рівнями дистриб'юторів на основі тетраедного дерева, що забезпечує структуру дистриб'юторської організації у складі 1024 осіб. Дистриб'ютор 4* - зірок І - рівня (наставник) традиційно в 1-рівень підписує нових 4 заявників (нових дистриб'юторів 1*), на 2 - рівні - 16 дистриб'юторів, 3 - 64 дистриб'ютори, 4 - 256 дистриб'юторів, 5 - 1024 дистриб'юторів, при цьому число рівнів не перевищує п'яти показана в табл. 1. Маркетинг-план мереженого маркетингу корпорації «ТЯНЬ ШІ» на конференції 6-річниці «ТЯНЬ ШІ» 03.08.2001 року в м. Санкт-Петербург Російської федерації, світовими економістами признаний як найкраща модель багаторівневого маркетингу в світі, тому авторами розроблено і в перше була запропонована в західному регіоні «ТЯНЬ ШІ-Україна» багаторівнева ієрархічно-реляційна структура БД як найбільш ефективна в роботі мереженого маркетингу корпорації, що показує застосування ієрархічно-рекурентної організації БД в базисі

Галуз є найбільш ефективними та може забезпечити зменшення надлишковості кодування даних, які формуються, зберігаються і передаються по міжнародних каналах зв'язку більше на один два порядки.

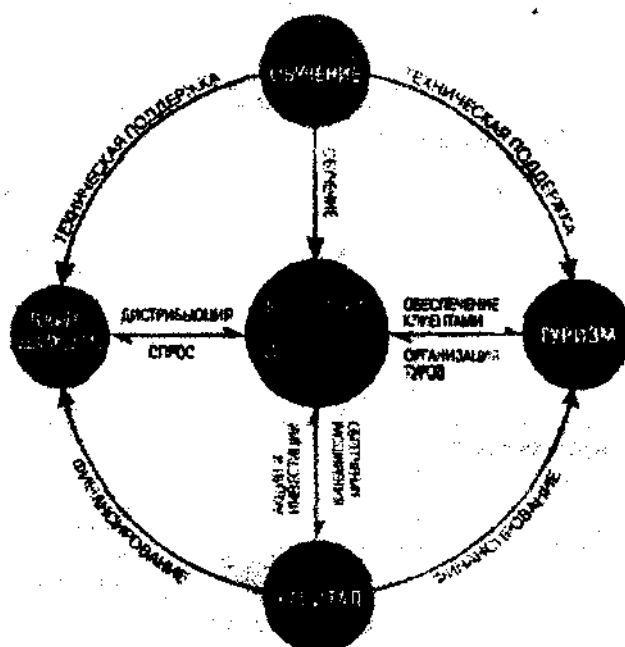


Рис. 1. Циклічний граф шести мереж міжнародної корпорації „Тянь Ші”.

Таблиця 1

**Багаторівнева ієрархічно-реляційна структура БД
корпорації „Тянь Ші”-Україна**

Ступінь Зірка*	Звання дистриб'юторів	Кількість дистриб'юторів із званням	Кількість дистриб'юторів по рівнях	Частотність, %
8* 5	Лідер дистриб'юторської організації	1	1	0,98
7* 4	Лідер дистриб'юторської структури	4	16	1,56
6* 3	Керівник офісу	16	48	4,68
5* 2	Верхній наставник	64	129	12,98
4* 1	Наставник	156	345	33,69
3* 0	Дистриб'ютор	420	285	27,83
2* 0	Споживач	280	105	10,25
1* 0	Заявник	83	95	9,27
Всього:		1024	1024	100

Характеристика частотності $F(r)$ розраховується у відсоткових значеннях реального числа дистриб'юторів на кожному рівні R^* від їх загального числа, що наближається до логарифмічного Гаусівського розподілу згідно рис. 2.

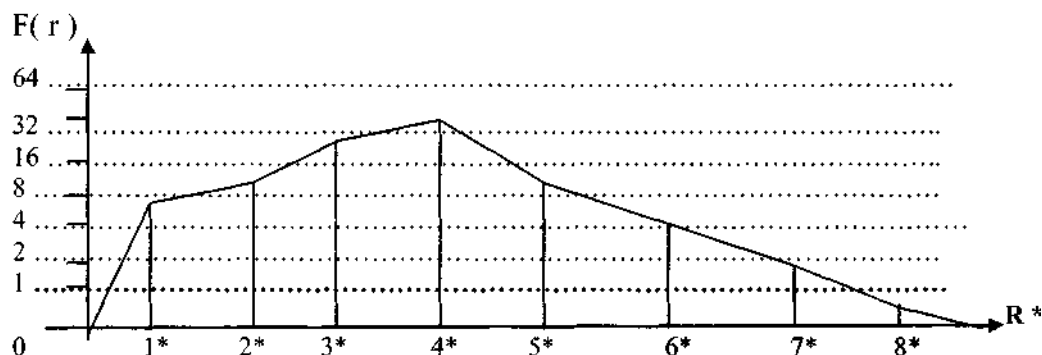


Рис. 2. Характеристика розподілу числа дистриб'юторів $F(r)$ на різних рівнях * структурного дерева корпорації «Тянь Ші»

На рис. 3 – відображена ієрархічна структура корпорації «Тянь Ші» системи ідентифікації дистриб'юторської організації багаторівневого маркетингу – (БРМ), яка належить до класу дерев зі зв'язаними списками [2, 3], де:

n -корінь – президент міжнародної китайської транснаціональної корпорації «Тянь Ші», $n_1, \dots, n_k$ підкорені – віце-президенти корпорації, почесні члени правління корпорації, ЗЛ6* – «Золоті льви шісти зірок», $T_1, \dots, T_n$ дерева – ЗЛ5* – виконавчі директори, які є менеджерами вищого рівня.

ЗЛ4* – ЗЛ – піддерева – «Золоті льви чотирьох зірок», СрЛ – «Срібні льви», міжнародні лектори, БрЛ – «Бронзові льви» – керівники середніх підрозділів, національні лектори,

які є менеджерами середнього рівня.

$k_1, \dots, k_n$ – вузли- дистриб'ютори 8*- зірок, регіональні лектори, 7*-лідери дистриб'юторських організацій, 6*-майстри, утримувачі офісів, завідувачі майстерень, відділів, 5*-бригадири, керівники –координатори груп, 4*- ланкові, керівники груп, які є менеджерами нижчого рівня.

3*- наставники, які працюють по спонсорській програмі, 2*- дистриб'ютори, які працюють по торговій програмі, - дистриб'ютори 1* зірки, заявники, учні, які працюють по споживчій програмі.

$l_1, \dots, l_k$ - листя – кандидати в дистриб'ютори, споживачі.

Структури представлені в табл.1 є типовими на всіх рівнях корпорації, які охоплюють міста та регіони більше 200 країн світу.

Наприклад, в Україні є п'ять регіональних представництв корпорації «ТЯНЬ ШІ» (східний, північний, центральний, південний та західний), які охоплюють усі 25 областей України. Університет «ТЯНЬ ШІ» має наміри співпрацювати з THEU.

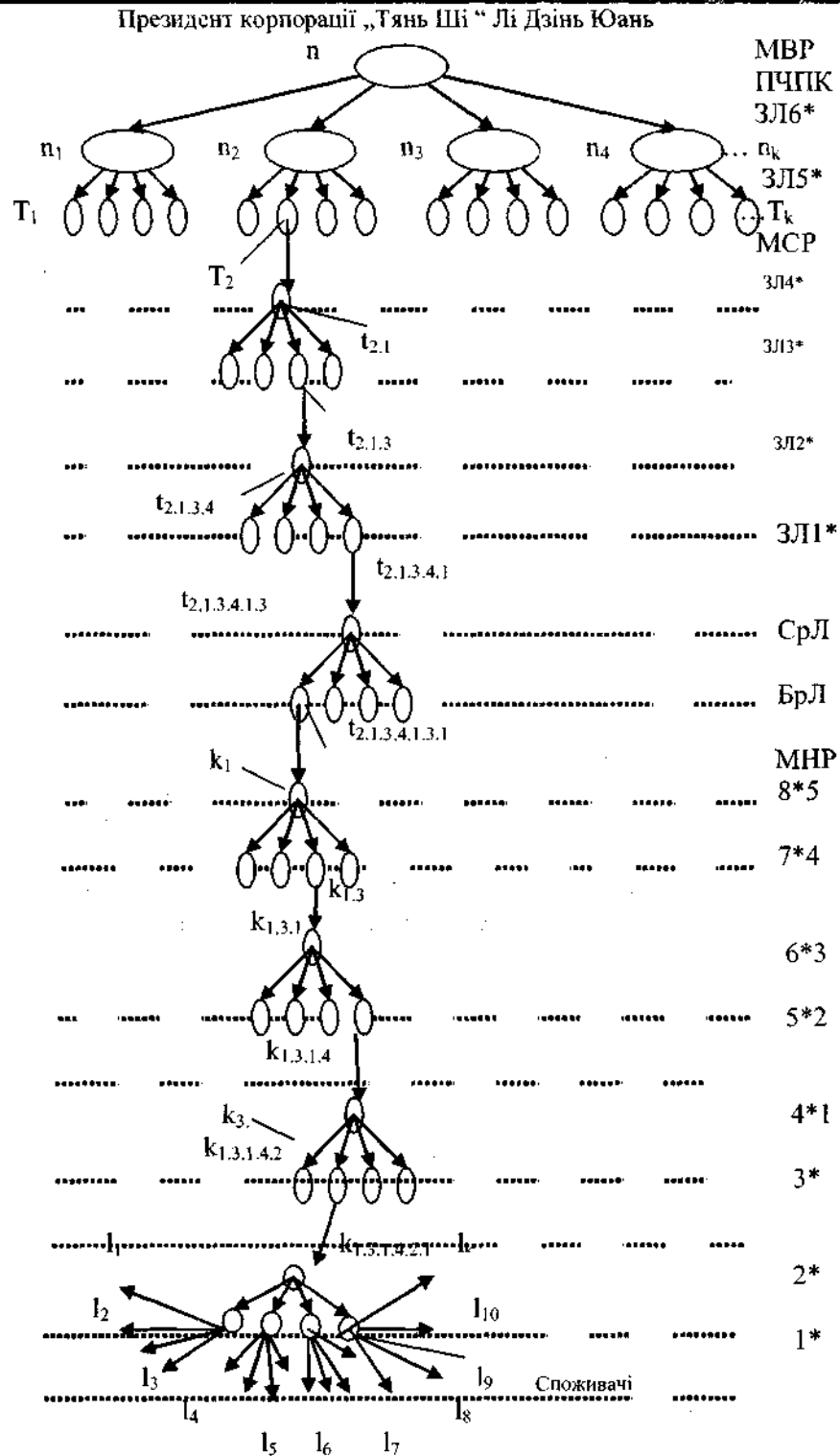


Рис. 3. Ациклічний граф. Велике дерево Т- організаційної структури корпорації «Тянь Ші».

В табл. 2. подана структура ієрархічно-реляційної організації ІД корпорації, яка включає наступні дані: НП-номер підприємства, НКД – ідентифікаційний номер дистриб'ютора, ФІОД – ПІБ, дистриб'ютора, НСД – ідентифікаційний номер наставника, ФІОСД – ПІБ, наставника, СТЕ – ступінь, ПОН – відстань до верхнього наставника, ЛЗ – особисті закупки, ОМЗ – товарообіг місячних закупок, СО – структурний об'єм, ЛНО – особистий нагромаджувальний об'єм, ОНОС – основний нагромаджувальний об'єм структури.

Таблиця 2

**Середньостатистичні параметри даних ієрархічно-реляційної БД
п'ятирівневої системи БРМ**

НП	НКД	ФІОД	НСД	ФІОСД	СТЕ	ПОН	ЛЗ	ОМЗ	СО	ЛНО	ОНОС
6·X	8·X	24·X	8·X	24·X	1·X	2·X	3·X	5·X	5·X	9·X	12·X
...	...	...	...	...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...	...	...	...	...
6·X	8·X	24·X	8·X	24·X	1·X	2·X	3·X	5·X	5·X	9·X	12·X

де: X – байт,

ІД

ОД

ІД – ідентифікаційні дані дистриб'ютора, ОД – організаційні дані дистриб'ютора

В табл. 2. символ n·X – означає число алфавітно-цифрових даних в кожній позиції двовірного файлу (n = 1, 2, ...). Таким чином при заданій кількості користувачів згідно табл. 2 можна розрахувати об'єм алфавітно-цифрових даних одного кортежа БД корпорації «ТЯНЬ ШІ».

Для зберігання даних кортежа реляційної БД, виходячи з умови байт-орієнтованого кодування, загальний об'єм даних може бути обчислений згідно формули:

$$I_j = \sum_{i=1}^m \sum_{k=1}^k d_{ij}, \quad (1)$$

де d_{ij} – об'єм даних ij – го атрибута кортежа ($i=1,2,\dots,k$) – число атрибутів, ($j = 1, 2, \dots, m$) – число кортежів.

Аналіз табл. 2 показує, що об'єми оперативних даних (ОД) діяльності дистриб'юторів визначаються доменами: ЛЗ, ОМЗ, СО, ЛНО та ОНОС, а ідентифікаційні дані описуються доменами: НП, НКД, ФІОД, НСД, ФІОСД, СТЕ та ПОН. Таким чином ефективність кодування БД ієрархічно-реляційної архітектури, можна розрахувати за формулами відносної та приведеної оцінки:

$$K_{e_1} = \frac{I_{ID}}{I_{OD}}; K_{e_2} = \frac{I_{ID} - I_{OD}}{I_{OD} + I_{ID}} \cdot 100\%. \quad (2)$$

Ефективність кодування БД ієрархічно-реляційної архітектури, яка відповідає корпорації «ТЯНЬ ШІ», можна розрахувати за формулою (2):

$$I_{OD} = 34 \cdot X; I_{ID} = 73 \cdot X. \quad (3)$$

Підставляючи дані у вирази формули (2) отримуємо значення коефіцієнта ефективності у вигляді .

$$K_{e_1} = 2,147; K_{e_2} = 36,4 \%$$

що відповідає більш ніж двохкратній надлишковості кодування (ІД) по відношенню до (ОД) або 36,4 % відносній надлишковості кодування (ІД).

Практично у корпорації «Тянь Ші» використовується універсальна система кодування ІД у вигляді двох 8 – розрядних десяткових кодів, які ідентифікують код наставника і дистриб'ютора в будь-якій структурі і підрозділі в будь-якій країні світу на ринку яких працює корпорація «Тянь Ші». Це складає 64 біт двійково-десятькового коду, або символно-байтового стандартного коду. При цьому об'єм бази ідентифікаційних даних менеджерів нижчого рівня корпорації «Тянь Ші» складає $11200 \times 128 = 1433600$, а загальна кількість можливих ідентифікованих членів складає 2^{64} осіб дистриб'юторів і споживачів населення планети Земля.

На рис.4 відображена ієрархічна структура менеджерів нижчого рівня корпорації «Тянь Ші». Для визначення певної ланки (наприклад дистриб'юторської організації №1) структури від лідера до дистриб'ютора і заявника, згідно рис. 3, відслідковуємо шлях: 0 – 1 – 2 – 3 – 1 – 2 – ... тобто лідер – дистриб'ютор підписує в свою групу 1 – 4 – ... нових заявників, які дублюють своїх наставників. Таким чином, наприклад, відслідковуючи шлях присвоюємо номер дистриб'ютору 1 – рівня – код 01, 2 – рівня – код 012, 3 – рівня – код 0123, 4 – рівня – код 01231, 5 – рівня – код 012312, новому заявнику – код 0123124. Відповідно з десяткової системи числення переходимо на кодування даних в двійковій системі числення в базисі Галуа, що показано в табл.4.

Якщо взяти окрему структуру дистриб'юторської організації (вузла k_1) і починаючи з лідера і 1×0 -рівня та пройти шлях до 8×5 - рівня зафіксувавши шлях та згідно порядку підписання контракту та розташуванні на певному рівні у групах буде присвоєний відповідний номер, який можна закодувати в базисі Галуа (рис. 3, 4.) що відображено в табл. 4. Відкриваючи код ієрархічної структури за його номером, відкривається реляційна БД всієї структури кожного дистриб'ютора світу корпорації «ТЯНЬ ШІ».

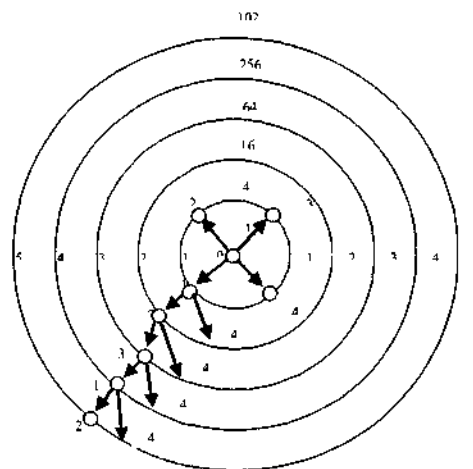


Рис. 4. Ациклічний граф. Вузли $k_1, \dots, k_n$. Дерева T - Ієрархічна структура менеджерів нижчого рівня дистриб'юторів $8 \times 5 - 4 \times 1, 3^*, 2^*$, та $1^*, \dots, 1^*_{k_n}$ листів -заявників 1^* зірок, споживачів $1-1024 \dots$

Проведений аналіз дозволяє обґрунтувати принципи організації багаторівневої рекурентної БД в базисі Галуа.

Теоретичні положення кодування ІД в багаторівневій рекурентній організації БД викладені в [5, 6, 7, 8]. Рекурентно-реляційна організація БД, яка еквівалентна ієрархічно-реляційної організації, що використовується в корпорації «ТЯНЬ ШІ», представлена в табл. 3.

Таблиця 3

Кодування символів дистриб'юторів БРМ та їх структур

№ 1	Ідент.№	П.І.Б.	Ідент.№	П.І.Б.	Ключі -	Символи-	Коди
Д*	Дистри-	Дистри-	Наставника	Наставника	коди	коди	Галуа
б'ютора	б'ютора						
8*5	8·X	24·X	8·X	24·X	0	0	00
7*4	8·X	24·X	8·X	24·X	01	.1	01
6*3	...	...	...	...	012	..2	10
5*2	...	...	...	...	0123	...3	11
4*1	...	...	...	...	01231	1	01
3*	8·X	24·X	8·X	24·X	012312	2	10
2*	8·X	24·X	8·X	24·X	0123124	4	00
1*	8·X	24·X	8·X	24·X			...

Числовий аналіз кодування ідентифікаторів НП – 11200·X, показує, що число символів, які потрібно зберігати в БД корпорації «ТЯНЬ ШІ» про код країни та ідентифікаційний номер дистриб'ютора, дорівнює – 11200 · X. При кодуванні аналогічних даних на основі рекурентної БД в базисі Галуа маємо 208·X · 300.

Звідки коефіцієнт надлишковості кодування ІД в БД корпорації «ТЯНЬ ШІ» згідно

$$\text{виразу (2) складає: } K_{\epsilon_1} = \frac{11200}{208} = 53,85.$$

Відповідно коефіцієнт надлишковості кодування ідентифікаційних кодів дистриб'юторів згідно рис.1. в першому випадку складає: 24·X, а в другому випадку: 11·X.

Результати розрахунків кодування ідентифікаційних даних по рівнях дистриб'юторської організації на основі рекурентної БД у базисі Галуа відображено в табл.4, де G_i - послідовність бітів коду Галуа згідно певного ключа, а $\bar{G}_i$ - інвертовані біти Галуа, які вказують початок ієрархічного підрівня БД.

Таблиця 4

**Кодування ідентифікаційних даних на основі рекурентних
кодів поля Галуа**

XXX	0	1	2	3	4	5	Код Галуа
... 0	8·X						GGGGGG...G $\overline{G}$ GGGGG
... 1	8·X	8·X					GGGGG...GG $\overline{G}$ GGGGG
... 2		8·X	8·X				GGGG...GGG $\overline{G}$ GGGGG
... 3			8·X	8·X			GGG...GGGG $\overline{G}$ GGGGG
... 4				8·X	8·X		GG...GGGGG $\overline{G}$ GGGGG
... 5					8·X	8·X	G...GGGGGG $\overline{G}$ GGGGG
XXX						8·X	...

У ХХІ столітті корпорацією планується реалізація 80% роздрібною торгівлі в мережевому маркетингу і 20% в класичній роздрібній торгівлі. Завдяки мережевому п'ятирівневому маркетингу тетраїдних груп дистриб'юторських організацій компанії вперше пропонувано ієрархічно-рекурентну організацію БД в базисі Галуа.

Проведений аналіз оцінки ефективності кодування ІД та ОД на прикладі корпорації «ТЯНЬ ШІ» показує, що застосування ієрархічно-рекурентної організації БД в базисі Галуа є найбільш ефективними та може забезпечити зменшення надлишковості кодування даних, які формуються, зберігаються і передаються по міжнародних каналах зв'язку більше на один два порядки. При цьому, враховуючи відносно низьку вартість засобів зберігання даних на фізичних носіях, які в діапазоні 10-10 Гбайт, ефекти зниження собівартості руху даних в транснаціональних міжнародних корпораціях можуть досягти в каналах зв'язку локальні та глобальні комп'ютерні мережі та Інтернет, що при інтенсивному трафіку обміну даними 0–1 Гбайт/с може скласти сотні тисяч у.о. на рік.

Література

1. Дж. Мартин. Организация баз данных в вычислительных системах. – М.: «Мир». 1980. – С. 662.
2. Ахо и др. Структуры данных и алгоритмы. Пер. с англ. – М.: дом Вильямс, 2001. – С. 384.
3. Гарсия-Молина и другие. Системы баз данных. – М.: дом «Вильямс», 2004. – С. 1088.
4. Пасічник В. В., Резніченко В. А. Організація баз даних та знань. – К.: Видавнича група BHV, 2006. – С. 384.

5. В. В. Шаряк, Я. М. Николайчук. Теоретико-числові базиси та їх застосування при організації бази даних // Оптико-електронні інформаційно-енергетичні технології. Вінниця. ВНТУ. № 2(12), 2006. – С. 59–66.
6. В. В. Шаряк, Я. М. Николайчук. Системні характеристики баз даних та перспективні напрямки їх розвитку // Моделювання та інформаційні технології. Київ. ІЕД НАН України. №1(1), 2007. – С. 38–45.
7. В. В. Шаряк. Архітектура і кодування баз даних на основі теоретико-числових базисів // Вісник. Тернопіль. ТДТУ ім. І.Пулля. Том.12. № 1, 2007. – С. 171–179.
8. В. В. Шаряк. Методи дослідження системних характеристик моделей бази даних. // Вісник. Тернопіль. ТДТУ ім. І.Пулля. Том 13. № 2, 2008. – С. 116–121.

Матеріал подається для публікації вперше і не був раніше опублікований.

Редакція отримала матеріал 10 листопада 2008 р.