

## ІЄРАРХІЧНА КЛАСТЕРИЗАЦІЯ СКЛАДНИХ СХЕМ

© Базилевич Р., Ждан А., 2008

**Розглянуто особливості алгоритмічної та програмної реалізації побудови дерева оптимального згортання схеми. Розкрито основні підходи до формування пар елементів для утворення кластерів. Проаналізовано експериментальні результати.**

**The features of algorithmic and programmatic realization of tree construction by the optimum reduction are considered. Basic approaches of forming of elements pair are exposed for clusters. The analysis of experimental results are provided for the few tests.**

### Вступ

Розбиття складних схем на частини є однією з важливих часткових задач при проектуванні сучасних мікропроцесорів, складних електронних схем, НВІС, багатопроцесорних систем. Важкість розв'язування задачі обумовлена неполіноміальною обчислювальною складністю та великою розмірністю схем (сотні тисяч базових елементів).

Для розв'язання задачі можна застосовувати метод оптимального згортання схеми, який полягає в паралельно-последовному формуванні підсхем, що мають однакові або близькі за взаємною зв'язністю елементів характеристики [1]. Метод виявляє ієрархічну кластерну структуру схеми на основі багатокрокового процесу ідентифікації та об'єднання підсхем, які є щільніші (густіші), ніж інші. Бажано виявити реальні згустки схеми, їхнє взаємне входження – менших у більші – та зв'язки між ними. Це сприятиме виділенню окремих сильно зв'язних груп, а також вказуватиме на їхнє взаємне входження. Наявність такої інформації істотно полегшить розв'язування наступної задачі – розбиття схеми на частини із забезпеченням бажаних характеристик. Для виявлення природних згустків (кластерів) схеми можуть бути використані певні формальні критерії, які враховують такі показники, як зв'язність між елементами чи їх групами, а також якісні, що враховують функційні властивості груп елементів.

### 1. Формулювання задачі

Більшість реальних складних схем можна представити як пару:  $S=(P, E)$ ; (1), де  $P=\{p_1, \dots, p_n\}$  – множина елементів,  $E=\{e_1, \dots, e_m\}$  – множина зв'язків між елементами.

Необхідно виявити сильно зв'язні згустки системи та побудувати дерево оптимального згортання  $T^R$ , яке відображає ієрархічне входження малих згустків (кластерів) у більші та корінь якого відповідає всій схемі. У випадках, коли система містить кілька незв'язаних підсистем, будується ліс дерев  $F^R=\{T^R_1, \dots, T^R_n\}$ .

### 2. Алгоритмізація задачі формування кластерів

Метод оптимального згортання схеми ґрунтується на виявленні сильно зв'язних згустків схеми – кластерів. На кожному кроці два елементи (кластери) об'єднуються в один, кластери формують дерево оптимального згортання  $T^R$ , яке відображає ієрархічне входження малих згустків у більші та корінь якого відповідає всій схемі. Базові кроки алгоритму побудови дерева оптимального згортання схеми  $T^R$  є такі:

#### 2.1. Формування списку пар елементів/кластерів, зв'язаних між собою

Для побудови пар елементів  $P_T$  використовується інформація зі списків *NetList* і *ElementList*. Для кожного зв'язка із множини зв'язків існує множина елементів, для яких цей зв'язок є спільним.  $Net_1=\{p_1, p_2, p_4, p_8\}$  зв'язок  $Net_1$  формує пари з елементів  $P_{T1}=p_1p_2$ ;  $P_{T2}=p_1p_4$ ;  $P_{T3}=p_1p_8$ ;  $P_{T4}=p_2p_4$ ;  $P_{T5}=p_2p_8$ ;  $P_{T6}=p_4p_8$ .

Зв'язок, який об'єднує  $n$  елементів, формує число  $M = (n-1)*n/2$  пар. Для кожної пари елементів визначаються множини зовнішніх, внутрішніх та довгих зв'язків і критеріїв згортання. Множина внутрішніх зв'язків пари є результатом перетину множин зв'язків елементів, які утворюють пару.

Множина довгих зв'язків утворюється із зв'язків, які об'єднують більше двох елементів. Множина зовнішніх зв'язків – це симетрична різниця множин зв'язків елементів, які утворюють пару.

## 2.2. Визначення критерію об'єднання для виділених пар

Вибір критерію згортання є одним з важливих для побудови дерева оптимального згортання. Досліджуються три критерії згортання:

- 1)  $\eta_1 = E^{Ext}(P_T) - E^{Int}(P_T)$ ;
- 2)  $\eta_2 = E^{Ext}(P_T) - E^{Int}(P_T) - E^{com}(P_T)$ ;
- 3)  $\eta_3 = E^{Ext}(P_T)$

де:

$E^{Ext}(PT)$  – кількість зовнішніх зв'язків пари елементів;

$E^{Int}(PT)$  – кількість внутрішніх зв'язків пари елементів;

$E^{com}(PT)$  – кількість довгих зв'язків пари елементів.

## 2.3. Упорядкування пар за значенням критерію

Для зручності і швидшого опрацювання даних множина пар кандидатів на згортання в кластери впорядковується за значенням критерію згортання. Впорядкування здійснюється один раз на етапі формування множини пар методом вставки пар у потрібні позиції. В процесі згортання модифіковані пари заносяться у потрібні позиції у списку пар, не порушуючи порядку слідування, заданого сортуванням на етапі формування множини пар. Такий підхід зменшує затрати часу на впорядкування.

## 2.4. Вибір пар елементів/кластерів для об'єднання

Розглядаються два підходи до вибору пар елементів/кластерів для об'єднання:

а) вільне згортання, при якому для об'єднання беруться усі пари з початку впорядкованого списку з однаковим найкращим значенням критерію;

б) вимушене згортання, при якому для об'єднання береться певний відсоток пар з початку впорядкованого за значенням критерію списку.

Відібрані пари формують множину пар елементів кандидатів для об'єднання. З цієї множини для об'єднання (утворення нових кластерів) відбираються пари, які не конфліктують між собою. Конфліктуючі пари – це пари, в яких є спільний один з елементів:  $P_{T1} = p_1p_2$ ,  $P_{T2} = p_1p_3$ ,  $p_1$  – спільний елемент.

## 2.5. Вилучення пар елементів із списку впорядкованих пар

Всі відібрані пари, які згорнулися у кластери, усуваються із списку впорядкованих пар. Пари, які конфліктували, заносяться на початок списку. У результаті згортання список пар постійно скорочується.

## 2.6. Модифікація впорядкованого списку пар

Список впорядкованих пар, що залишилися після вилучення згорнутих в кластери елементів, модифікується. Модифікація проводиться над тими парами, які мають спільні елементи із парами, що згорнулися, спільний елемент замінюється кластером, що утворився. Процес модифікації аналогічний до процесу створення нової пари. У результаті модифікації можуть утворюватися дублюючі пари, які потрібно вилучати зі списку пар. На одному кроці список скорочується за рахунок згортання і за рахунок вилучення дублюючих пар. Якщо для кластера не існує пар, що модифікуються, це означає, що кластер не зв'язаний з іншими елементами схеми. Ці кластери виділяються окремо для полегшення подальшої роботи з ними. Вони утворюють окремі групи, не зв'язані з іншими кластерами, тобто окремими елементами лісу.

У результаті модифікації може змінюватися значення критерію згортання окремих пар; такі пари потрібно перевпорядкувати, що здійснюється вставлянням у відповідні місця списку.

### 3. Опис структур даних

Структури даних програми використовують однонапрямлені і двонапрямлені списки та списки з розгалуженням. Також використовуються динамічні масиви різних структур і простих типів даних. Зв'язки кожної пари подають у вигляді списків, які характеризують її внутрішні, зовнішні та довгі зв'язки. Робота зі списками зв'язків побудована так, щоб звести до мінімуму затрати пам'яті і повторні обчислення.

Пари елементів будуються зі списків *ElementList* і *NetList*. *ElementList* – це спискова структура, кожен зв'язок якої підпорядкований одному елементу схеми та містить список зв'язків, інцидентних до цього елемента (рис. 1), *NetList* – спискова структура, кожен елемент якої відповідає одному зв'язку схеми та містить список елементів, інцидентних до цього (рис. 2). Пара елементів – це структура, яка містить: назву першого елемента; назву другого зв'язку; посилання на список зовнішніх зв'язків; посилання на список внутрішніх зв'язків, посилання на список довгих зв'язків, інформацію про кількість внутрішніх і зовнішніх зв'язків; значення критерію згортання; два прапорці, які вказують, з чого утворена пара (елементів, кластерів, елемента і кластера); посилання на список зв'язків для першого елемента (кластера); посилання на список зв'язків для другого елемента (кластера).

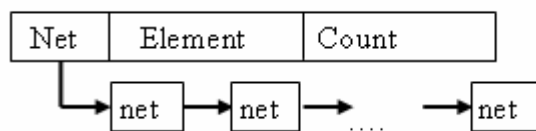


Рис. 1. Представлення структури NetList

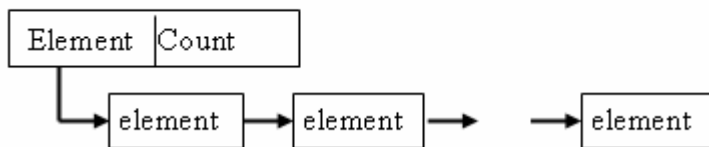


Рис. 2. Представлення структури ElementList

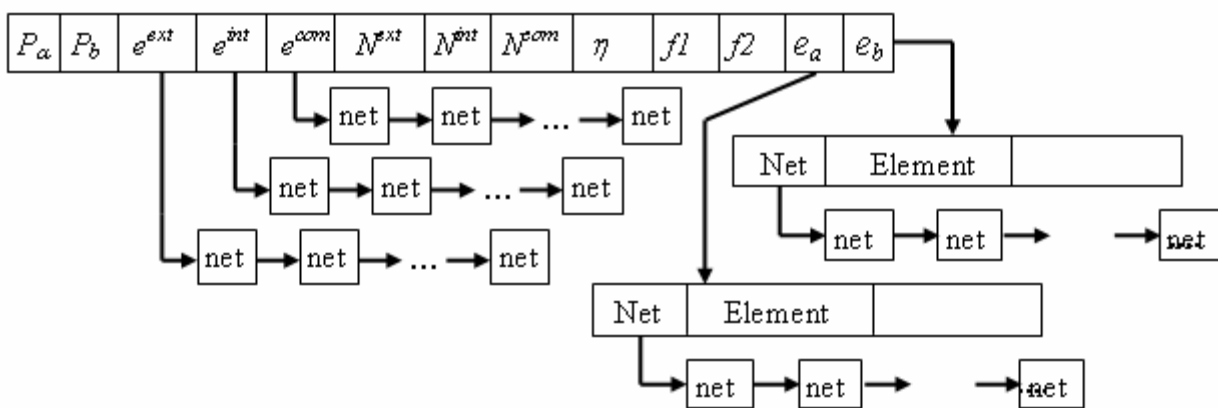


Рис. 3. Представлення структури даних для пари елементів/кластерів

На рис. 3 подано інформацію, що зберігається в пам'яті про пару зв'язаних елементів. Введено такі позначення:  $P_a$  і  $P_b$  – елементи, або кластери, які утворюють пару;  $e^{ext}$ ,  $e^{int}$ ,  $e^{com}$  – список внутрішніх, зовнішніх довгих зв'язків між  $P_a$  і  $P_b$ ;  $\eta$  – значення критерію згортання;  $N^{ext}$ ,  $N^{int}$ ,  $N^{long}$  – кількість

внутрішніх, зовнішніх і довгих зв'язків пари  $P_a$  і  $P_b$ ;  $f1, f2$  – прапорці, які вказують, чи  $P_a$  і  $P_b$  є елементами, чи кластерами;  $e_a, e_b$  – посилання на списки всіх зв'язків для елементів (кластерів). Пари, які згорнулись, записуються в масив. Елементи цього масиву є структурами і містять таку інформацію:  $P_a$  і  $P_b$  – назва елементів чи кластерів, які утворили цей кластер;  $C$  – назва утвореного кластера;  $N_c$  – кількість елементів, які об'єднує кластер  $C$ ;  $Itr$  – кількість ітерацій, потрібних для згортання кластера  $C$ . На рис. 4. представлена інформація про модель дерева згортання.

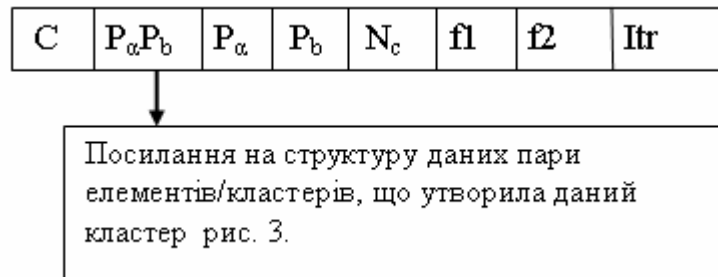


Рис. 4. Представлення структури даних дерева згортання

#### 4. Особливості програмної реалізації

Для розроблення програмної системи використано об'єктно-орієнтований підхід. Програмна система може використовуватись як допоміжний засіб у САПР, а також для задач пакування, та інших.

Для побудови пари елементів визначається значення її критерію згортання, і пара додається до списку пар. Для ефективнішого опрацювання списку пар посилання на ці пари групуються за значенням критерію (рис. 5). Формуються блоки посилання (К) на пари (D) з однаковим значенням критерію.

Доступ до блоків здійснюється за допомогою масиву. Наприклад, для того, щоб здійснити доступ до блоку зі значенням критерію пар, який дорівнює 3, потрібно звернутись до 3-го елемента масиву, який є вказівником на початок відповідного блоку. У процесі згортання це дає можливість вибирати пари, які згортатимуться з окремого блоку або з декількох блоків залежно від вибору параметрів згортання.

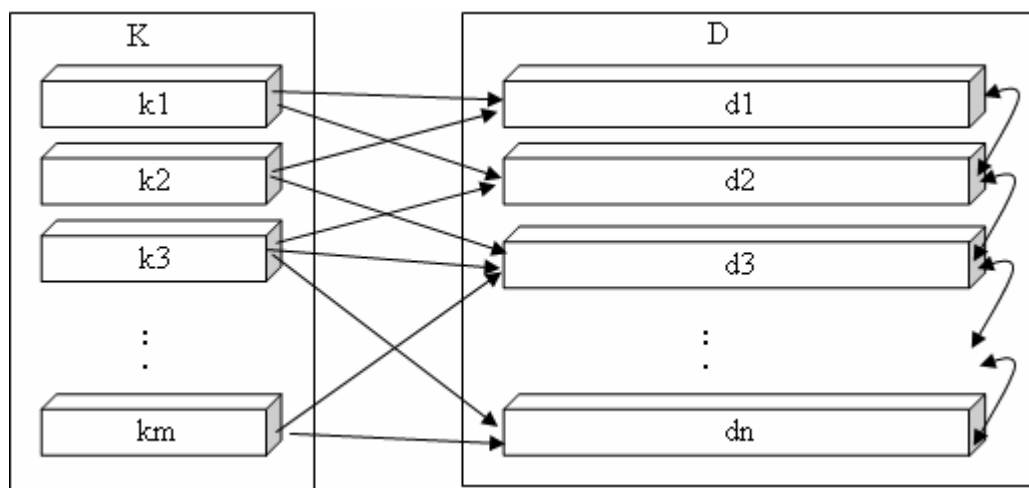


Рис. 5. Групування списку пар за значенням критерію

Для підвищення ефективності визначення пар для модифікації застосовуємо підхід, при якому посилання на пари групуються за значенням назв елементів. Доступ до пар здійснюється за допомогою масиву, аналогічного до масиву блоків значень критеріїв. У цьому випадку блоки об'єднують у пари, в яких присутні однакові елементи чи кластери. Схема групування списку пар відносно елементів, які утворюють пару, аналогічна до схеми на рис. 5.

Для модифікації списку пар використовуємо масиви, індексами в яких є елементи (для масиву, який індексує елементи) і кластери для відповідного масиву, який індексує кластери.

Значення цих масивів є списками вказівників на пари зі списку пар. Наприклад, 5-й елемент масиву (індексує елемент) вказує на список (блок), в якому містяться вказівники на всі пари, в яких міститься 5-й елемент, відповідно 5-й елемент масиву (індексує кластер) вказує на всі пари, в яких міститься 5-й кластер. Це дає можливість при модифікації пар не затрачати додатково час на пошук пар у всьому списку, по чергово переглядаючи його, а за відповідними індексами отримувати список необхідних пар.

Усі модифіковані пари записуються до спискової структури, яка зберігає список модифікованих пар. У цьому списку пари дублюються, для вилучення їх із цього списку і загального списку пар використовується такий самий підхід, як при модифікації. У цьому випадку звертаємось за цим індексом до масиву і одержуємо вказівник на список пар. Якщо трапляється ситуація, що в списку є вказівники, які вказують на декілька пар, то дубльовані пари усуваються. Після вилучення таких пар вставляють пару до списку пар за відповідним значенням критерію згортання. Пари вставляються за аналогічним принципом, що і під час побудови списку пар з використанням блоків (рис. 5). Залежно від значення критерію згортання пари вставляють у відповідні блоки.

### 5. Експериментальні дослідження процесу згортання схеми

Дослідження проводились на основі реальних схем фірми *ibm* [4] на пакеті з 18 тестів *ibm01-ibm18*. Розмірність тестів знаходиться в діапазоні від 12752 до 210613 логічних елементів.

У таблиці наведено характеристики тестів, затрати часу на побудову дерева згортання, загальний час (побудова списку елементів, зв'язків, пар елементів, дерева згортання). За критерій згортання вибрано різницю внутрішніх і зовнішніх зв'язків.

$$\eta_l = E^{Ext}(P_T) - E^{Int}(P_T).$$

#### Згортання найкращих пар елементів згідно з критерієм

№ тесту	Кількість елементів	Кількість зв'язків	Час побудови списків (елементів, зв'язків)	Кількість пар елементів	Час побудови пар елементів, с	Час побудови дерева згортання, с	Загальний Час роботи, с
1	12752	14411	1	109183	4	<b>5</b>	10
2	19601	19584	2	343409	22	<b>20</b>	44
3	23136	27401	3	206069	15	<b>18</b>	36
4	27219	31970	4	229423	22	<b>19</b>	45
5	29347	28446	5	349676	30	<b>29</b>	64
6	32331	34826	6	321308	37	<b>26</b>	69
7	45926	48117	11	373328	79	<b>37</b>	127
8	51309	50513	13	732550	113	<b>72</b>	198
9	53395	60902	17	478777	185	<b>49</b>	251
10	69429	75196	30	707969	239	<b>65</b>	334
11	70558	81454	31	508442	279	<b>56</b>	303
12	71076	77240	34	806965	258	<b>104</b>	396
13	84199	99666	58	744500	494	<b>94</b>	446
14	147605	152772	129	1125147	1082	<b>208</b>	1419
15	161570	186608	314	1751474	1020	<b>400</b>	2134
16	183484	190048	390	1923995	1600	<b>476</b>	2466
17	185495	189581	465	2235716	1850	<b>534</b>	2849
18	210613	201920	478	2221860	1720	<b>510</b>	2708

Рисунки 6–8 ілюструють графіки залежності часу побудови дерева згортання від кількості елементів, зв'язків та утворених пар схеми. Як бачимо, залежність хоча і є показниковою, маємо невеликий степінь (1,55), що робить алгоритм придатним для розв'язування задач великої розмірності.

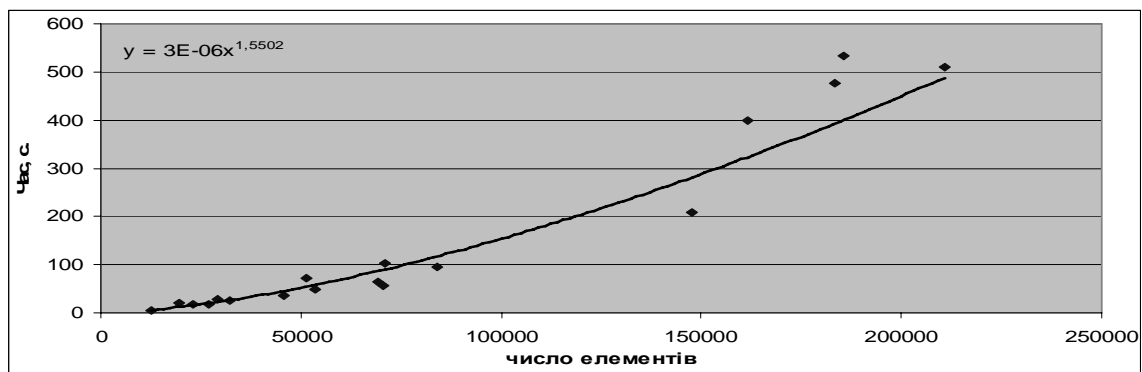


Рис. 6. Залежність часу побудови дерева згорання від кількості елементів схеми

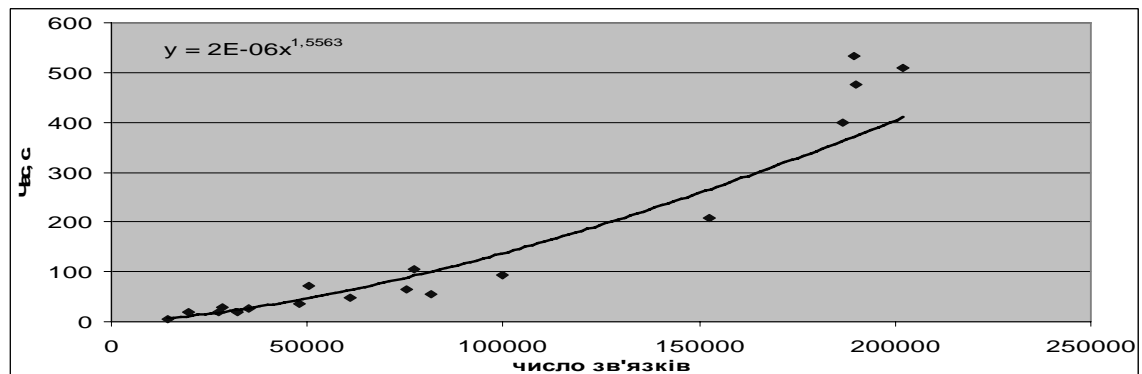


Рис. 7. Залежність часу побудови дерева згорання від кількості зв'язків схеми

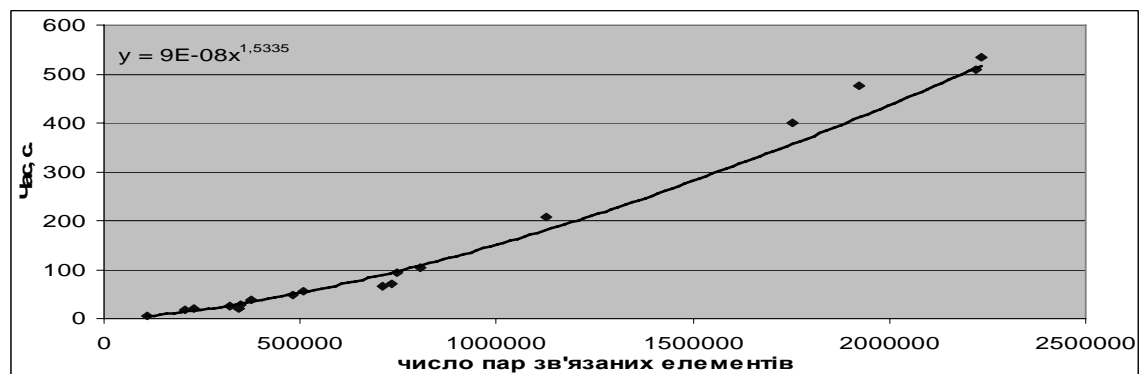


Рис. 8. Залежність часу побудови дерева згорання від кількості пар зв'язаних елементів схеми

### Висновок

Експериментальні дослідження розробленого алгоритму на тестах *ibm01–ibm18* дають добрі результати. Розглянуто декілька підходів програмної реалізації алгоритму згорання схем. Проведено експериментальні дослідження з різними значеннями параметрів. На основі проведених експериментів можна зробити висновок про доцільність використання розвинутого алгоритму для розв'язування задач великої розмірності.

1. Базилевич Р.П. Декомпозиційні та топологічні методи автоматизованого конструювання електронних пристроїв. – Львів: Вища школа. Вид-во при Львів. гос. ун-те, 1981. – 168 с. 2. Базилевич Р.П., Подольський І.В. Особливості організації пакета програм для ієрархічної кластеризації схем // Вісник Нац. ун-ту "Львівська політехніка" "Радіоелектроніка та телекомунікації", 2002. – № 440. – С. 139–144. 3. Базилевич Р.П., Подольський І.В., Ієрархічна кластеризація – ефективний засіб розв'язування не поліноміальних комбінаторних задач схемного типу високої розмірності // "Штучний інтелект". – 2002. – №3. – С. 474–483. 4. Charles J. Alpert, The ISPD98 Circuit Benchmark Suite. ISPD98 Monterey CA USA, 1998.5. R. P. Bazylevych, R. A. Melnyk and O. G. Rybak. "Circuit Partitioning for FPGAs by the Optimal Circuit Reduction Method", VLSI DESIGN 2000, Vol. 11, No. 3, pp. 237–248.