

ІНСТРУМЕНТАЛЬНІ ЗАСОБИ АВТОМАТИЗОВАНОГО ПРОЕКТУВАННЯ

УДК 621.382

Р.П. Базилевич, А.Р. Ждан

Національний університет “Львівська політехніка”,
кафедра програмного забезпечення автоматизованих систем

АНАЛІЗ ВХІДНИХ ДАНИХ ТА ОЦІНКА ЯКОСТІ КЛАСТЕРИЗАЦІЇ ДЛЯ АЛГОРИТМУ ОПТИМАЛЬНОГО ЗГОРТАННЯ СХЕМ

© Базилевич Р., Ждан А., 2008

Розглянуто особливості алгоритму побудови дерева оптимального згорання схеми при різних критеріях та значеннях параметрів роботи. Виконано аналіз вхідних даних, дана оцінка якості кластеризації.

The main features of algorithm of construction of tree of the optimum rolling up of chart are considered at different criteria and parameters of work. The analysis of entrance data and estimation of quality of clusterization.

Вступ

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

Для розв'язання задачі можна застосовувати метод оптимального згорання схеми, який полягає в паралельно-послідовному формуванні підсхем, що мають однакові або близькі за взаємною зв'язністю елементів характеристики [1–3]. Метод оптимального згорання схеми виявляє ієрархічну кластерну структуру схеми на основі багатокрокового процесу ідентифікації та об'єднання підсхем, які є щільнішими за інші. Це сприяє виділенню окремих сильно зв'язних груп. Одержання якісних кластерів є не менш важливим за швидкодію роботи алгоритму. Хороші результати кластеризації дають змогу спростити багато інших задач, таких, як декомпозиція, оптимізаційні задачі. Якість кластеризації можна оцінювати залежно від якості множини окремих кластерів верхніх рівнів, близьких за розмірами, які не мають входжень один в одного. Аналіз вхідних даних та оцінка якості кластеризації для алгоритму оптимального згорання схем дає змогу визначити кращі стратегії роботи алгоритму і підібрати оптимальний набір методів управління цим алгоритмом.

1. Постановка задачі

Під час побудови дерева згорання одержуємо множину кластерів з різними характеристиками. Залежно від вхідних даних, а також від вибору стратегій роботи алгоритму кластери змінюють свої характеристики. Необхідно визначити критерії оцінки якості кластеризації і дати оцінку якості отриманих результатів відносно вхідних даних і стратегій роботи алгоритму оптимального згорання схем, для одержання кращих результатів кластеризації схем. Це дає змогу спростити подальшу роботу з цими даними.

2. Аналіз вхідних даних

Аналіз вхідних даних дає можливість отримати важливу інформацію щодо вибору належних стратегій, критеріїв згорання, фільтрації довгих зв'язків, фільтрації елементів певних типів. Це забезпечує високу швидкодію роботи програмних засобів і отримання якісних результатів кластеризації схем. У табл. 2 проаналізовано кількості зв'язків різних довжин для всіх тестів.

Табл. 1 ілюструє відношення кількості елементів до кількості зв'язків схеми. Від цих співвідношень залежить кількість пар зв'язаних елементів, які необхідно утворити для швидкої побудови дерева згортання. Під парою зв'язаних елементів $\langle p_i, p_j \rangle$ розуміємо такі два елементи схеми $S=(P, E)$; $P=\{p_1, \dots, p_n\}$; $E=\{e_1, \dots, e_2\}$; що мають хоча б один інцидентний для обох них зв'язок.

Формування списку впорядкованих пар як вхідних даних для побудови дерева згортання дає змогу визначити один раз суміжні елементи і не повторяти цю процедуру на кожному кроці згортання схеми, що істотно збільшує швидкодію роботи алгоритму. Від кількості утворених пар зв'язаних елементів також залежить швидкодія і тому доцільно мінімізувати це число. Якщо зв'язок є бінарним, то він утворює одну пару. Кожен n -арний зв'язок утворює $\frac{1}{2} n(n-1)$ пар. Тобто довгі зв'язки можуть утворювати значну кількість додаткових пар. Тому доцільно вивчити можливість їхньої фільтрації при побудові пар зв'язаних елементів з оцінкою змін кластеризації та швидкодії алгоритму. Табл. 3 ілюструє зменшення кількості утворених пар при фільтрації зв'язків певної довжини. Так, неврахування зв'язків з довжиною, більшою за 10, зменшує кількість утворених пар на 24 %.

Таблиця 1

Відношення кількості елементів до кількості зв'язків

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
0,88	1,01	0,84	0,85	1,03	0,92	0,95	1,02	0,87	0,92	0,86	0,92	0,84	0,96	0,86	0,96	0,97	1,04

Таблиця 2

Характеристика тестів щодо кількості зв'язків різних довжин

№	Кількість елементів	Довжина зв'язків											кількість зв'язків
		2	3	3-5	5-10	10-15	15-20	20-30	30-40	40-50	>50		
Ibm01		12752	8341	2082	3863	1953	653	169	93	26	1	0	14411
Ibm02		19601	10692	1934	5831	3765	1164	481	45	36	1	51	19584
Ibm03		23136	17619	3084	6289	3354	981	444	84	16	1	0	27401
Ibm04		27219	19246	5160	8826	5120	385	124	159	57	15	0	31970
Ibm05		29347	18013	1073	4174	5335	2223	1879	0	0	0	0	28446
Ibm06		32331	20555	4194	9397	4351	2192	473	291	2	0	0	34826
Ibm07		45926	34500	10161	17135	9242	2895	1146	91	15	0	0	48117
Ibm08		51309	26669	8221	14158	6758	2873	802	52	0	0	0	50513
Ibm09		53395	28687	6911	15019	6698	1381	1712	545	312	74	0	60902
Ibm10		69429	42945	7742	18194	13844	4563	1435	653	21	1	0	75196
Ibm11		70558	44436	14843	26622	14257	3107	760	21	0	0	0	81454
Ibm12		71076	44312	8064	15741	13406	6617	1080	877	0	0	0	77240
Ibm13		84199	55301	17679	30699	12716	5100	1778	144	0	0	0	99666
Ibm14		147605	85328	23041	46229	23661	6675	2199	310	7	0	0	152772
Ibm15		161570	101899	28516	55307	28123	8605	4877	1571	23	0	0	186608
Ibm16		183484	103586	22996	47522	38010	12569	4022	1553	19	2	0	190048
Ibm17		185495	102749	18273	36154	39524	19044	5039	2070	16	0	0	189581
Ibm18		210613	97538	42159	73802	26343	11772	7156	739	470	73	24	201920

Таблиця 3

Застосування фільтрації зв'язків певної довжини

Тест	Фільтрація зв'язків певної довжини	Кількість сформованих пар
Ibm01	< 10	82807
	< 20	99207
	< 30	102944
	< 40	108253

3. Методи та критерії оцінки якості алгоритмів кластеризації схем

Під час оцінювання якості кластеризації схеми найчастіше враховуються такі найважливіші показники, як кількість зв'язків між кластерами та кількість внутрішніх зв'язків. Для дослідження роботи алгоритму оптимального згортання схем і для кращої оцінки якості кластеризації проаналізовано згортання на всіх ітераціях. Визначається кількість згорнутих пар елементів на певному рівні, середня, максимальна і мінімальна кількість внутрішніх і зовнішніх зв'язків. Аналіз цієї інформації створює умови для вироблення ефективних методів управління процесом кластеризації. Факторами, що дають змогу керувати кластеризацією схем, є критерії згортання, фільтрація довгих зв'язків, вилучення плаваючих елементів, тип згортання – вимушене чи вільне, а також їхні комбінації. Кластеризація схем є допоміжною задачею для багатьох задач проектування електронних складних схем. Тому оцінка якості кластеризації схем є важливою. Досліджено вплив таких показників на якість кластеризації схем:

- 1) $K_1 = m_c^{Ext}$;
- 2) $K_2 = m(c_1)^{Ext} + m(c_2)^{Ext}$;
- 3) $K_3 = m(c_1)^{Ext} + m(c_2)^{Ext} - m^{ExtC1C2}$;
- 4) $K_4 = m_c^{Ext} / m_c^{Int}$;
- 5) $K_5 = m^{ExtC1C2}$;
- 6) $K_6 = n_{elem} / (n(c_1) + n(c_2))$;

де m_c^{Ext} – середня кількість зовнішніх зв'язків на всіх рівнях згортання; m_c^{Int} – середня кількість внутрішніх зв'язків на всіх рівнях згортання; $m(c_1)^{Ext}$ і $m(c_2)^{Ext}$ – кількість зовнішніх зв'язків двох найбільших кластерів дерева T^R , що не мають входження один в одного; $m^{ExtC1C2}$ – кількість зовнішніх зв'язків між двома найбільшими кластерами дерева T^R , що не мають входження один в одного; n_{elem} – загальна кількість елементів у схемі; $n(c_1)$ і $n(c_2)$ – кількість елементів у першому і другому найбільшому кластері, що не мають входження один в одного.

4. Експериментальні дослідження алгоритму оптимального згортання схеми

Експериментальні дослідження виконані для тесту Ibm01 [4]. Вивчено вплив критеріїв згортання, процента згорнутих пар на кожному кроці, фільтрації довгих зв'язків на якісні характеристики методу оптимального згортання – такі, як час згортання, кількість ітерацій, кількість внутрішніх та зовнішніх зв'язків, параметри двох найбільших кластерів. Необхідно було визначити, за яких значень цих чинників забезпечується утворення кращих кластерів, як вони впливають на час побудови дерева згортання, що є особливо важливим для задач великої розмірності.

У табл. 4–9 відображено такі характеристики дерева оптимального згортання схем:

- кількість ітерацій у кластеризації схеми;
- максимальна кількість зовнішніх зв'язків дерева згортання;
- максимальна кількість внутрішніх зв'язків дерева згортання;
- середня кількість зовнішніх зв'язків на всіх ітераціях згортання;
- середня кількість внутрішніх зв'язків на всіх ітераціях згортання;
- ідентифікація двох найбільших кластерів, які не входять один в одного;
- розмірність двох найбільших кластерів (кількість об'єднаних елементів);
- кількість зовнішніх зв'язків двох найбільших кластерів;
- кількість внутрішніх зв'язків двох найбільших кластерів;
- кількість зв'язків між двома найбільшими кластер.

Також у табл. 7–9 відображена інформація про середню кількість зовнішніх, внутрішніх зв'язків, про кількість згорнутих пар на цьому рівні. Аналіз виконано для 10 %, 20 %, 30 %, ..., 100 % ітерацій від всієї їхньої кількості, а також для (10–20) %, (20–30) %, ..., (90–100) %.

Ця інформація може бути використана для оцінки якості кластеризації та вивчення кластеризації і вироблення ефективних підходів для керуванням цим процесом.

Табл. 4 відображає характеристики кластеризації при різних критеріях згортання:

- 1) $h_1 = E^{Ext}(P_T) - E^{Int}(P_T)$;
- 2) $h_2 = E^{Ext}(P_T) - E^{Int}(P_T) - E^{Com}(P_T)$;
- 3) $h_3 = E^{Ext}(P_T)$;
- 4) комбінований критерій з η_1 і η_2 .

Таблиця 4

**Характеристики кластеризації для тесту Іbm01
за різних критеріїв згортання**

Іbm01		1	2	3	4
Кількість ітерацій		329	308	254	337
Мах. зовн. зв.		603	573	702	475
Мах. внутр. зв.		263	358	229	309
Середня кількість зовн. зв.		12,3	12	12,6	12,2
Середня кількість внутр. зв.		2,86	2,86	2,86	2,86
Кластер 1	Назва	12745	12743	12746	12745
	Кіл. елем	5990	5929	6059	6078
	Кіл. зовн. зв.	452	390	551	415
	Кіл. внутр. зв.	80	90	178	112
Кластер 2	Назва	12748	12748	12750	12746
	Кіл. елем	4022	5667	3666	4356
	Кіл. зовн. зв.	603	573	702	416
	Кіл. внутр. зв.	239	167	229	92
Кількість зв. між Кластер1 і Кластер2		263	240	246	152
Кількість пар		109183			
Кількість елементів		12752			
Кількість зв'язків		14411			

Таблиця 5

**Характеристики кластеризації для тесту Іbm01
при фільтрації довгих зв'язків і плаваючих елементів**

Іbm01	6ф	5	7	10	15	20	фе	
Кількість відфільтрованих елементів	0	1924	718	237	113	60	781	
Кількість ітерацій	329	351	350	351	345	348	319	
Мах. зовн. зв.	603	2662	1102	617	652	612	725	
Мах. внутр. зв.	263	317	296	334	315	339	224	
Середня кількість зовн. зв.	12,3	15,0	13,3	12,6	12,5	12,4	13,4	
Середня кількість внутр. зв.	2,86	2,9	2,9	2,8	2,8	2,8	3	
Кластер1	Назва	12745	10825	12030	12509	12635	12689	11966
	Кіл. елем	5990	4914	4912	5608	6167	6343	5504
	Кіл. зовн. зв.	452	1383	854	610	652	548	620
	Кіл. внутр. зв.	80	222	273	188	278	251	184
Кластер2	Назва	12748	10824	12031	12512	12627	12680	11962
	Кіл. елем	4022	3809	4382	4711	4209	5609	4038
	Кіл. зовн. зв.	603	1215	784	609	397	421	619
	Кіл. внутр. зв.	239	178	175	287	118	136	149
Кількість зв. між Кластер1 і Кластер2		263	144	219	222	138	237	34
Кількість елементів		12752						
Кількість зв'язків		14411						

Таблиця 6

**Характеристики кластеризації для тесту Івм01
за різних параметрів згортання**

Івм01		найкращі	20%
Кількість ітерацій		329	72
Мах. зовн. зв.		603	574
Мах. внутр. зв.		263	289
Середня кількість зовн. зв.		12,3	12,2
Середня кількість внутр. зв.		2,86	2,8
Кластер 1	Назва	12745	12742
	Кіл. елем.	5990	5706
	Кіл. зовн. зв.	452	378
	Кіл. внутр. зв.	80	84
Кластер 2	Назва	12748	12749
	Кіл. елем.	4022	4688
	Кіл. зовн. зв.	603	540
	Кіл. внутр. зв.	239	252
Кількість зв. між Кластер1 і Кластер2		263	265
Кількість елементів		12752	
Кількість зв'язків		14411	

Таблиця 7

**Характеристики згортання для тесту Івм01
за різних критеріїв згортання на різних рівнях згортання**

Відсоток згорнутих елементів	Середня кількість зовнішніх зв'язків				Середня кількість внутрішніх зв'язків				Кількість згорнутих пар			
	1	2	3	4	1	2	3	4	1	2	3	4
10	4,8	5	5,2	5,4	1,6	1,4	1,5	1,6	6365	7122	7530	7776
20	6,3	6,6	7,3	6,5	1,9	1,8	1,9	1,9	9551	10223	11024	9746
30	7,5	7,6	8,4	7,6	2,0	2,0	2,1	2,0	11239	11406	11939	11291
40	8,3	8,7	9,2	8,2	2,2	2,1	2,2	2,1	11839	12143	12299	11780
50	9,2	9,2	9,9	9,0	2,3	2,3	2,3	2,2	12314	12378	12509	12246
60	9,7	9,8	10,4	9,8	2,4	2,4	2,4	2,3	12459	12529	12602	12483
70	10,4	10,3	10,8	10,4	2,4	2,4	2,5	2,4	12607	12632	12655	12600
80	10,9	10,8	11,2	10,9	2,5	2,5	2,5	2,5	12676	12680	12692	12669
90	11,5	11,2	11,7	11,4	2,6	2,6	2,6	2,6	12716	12718	12723	12714
0-10	4,8	5	5,2	5,4	1,6	1,4	1,5	1,6	6365	7122	7530	7776
10-20	9,3	10,4	11,8	10,8	2,4	2,8	2,8	2,7	3186	3101	3494	1970
20-30	14,6	16,1	21,9	14,9	2,8	3,8	4,3	2,9	1688	1183	915	1545
30-40	22,2	24,9	34,6	22,3	4,3	5,1	6,1	4,5	600	737	360	489
40-50	32,8	39,0	51,9	30,8	5,6	7,6	9,2	5,6	475	235	210	466
50-60	49,0	53,2	76,1	46,8	8,9	9,7	11,0	8,0	145	151	93	237
60-70	68,3	79,5	107	70,5	9,5	12,2	14,5	12	148	103	53	117
70-80	112,6	119	148	104	15,8	18,6	20,6	15	69	48	37	69
80-90	182	182	228	164	28,1	34,0	34,7	29	40	38	31	45
90-100	323	312	417	284	97,9	103	130	91	34	32	27	36

**Характеристики згорання для тесту Івм01
за різних параметрів фільтрації на різних рівнях згорання**

Відсоток згорнутих елементів	Середня кількість зовнішніх зв'язків						Середня кількість внутрішніх зв'язків						Кількість згорнутих пар					
	бф	7	10	15	20	фе	бф	7	10	15	20	фе	бф	7	10	15	20	фе
10	4,8	5,4	4,8	4,8	4,8	5,6	1,6	1,7	1,5	1,6	1,6	1,8	6365	6769	5965	6224	6319	7295
20	6,3	6,7	6,4	6,3	6,3	6,8	1,9	1,9	1,8	1,9	1,9	2	9551	9035	9162	9328	9430	9526
30	7,5	7,6	7,5	7,5	7,4	7,9	2,0	2,0	2,0	2,0	2	2,1	11239	10279	10799	10946	11024	10626
40	8,3	8,2	8,1	8,1	8	8,6	2,2	2,1	2,1	2,1	2,1	2,2	11839	10743	11342	11550	11539	11110
50	9,2	9,3	9,1	8,9	8,9	9,7	2,3	2,2	2,2	2,2	2,2	2,3	12314	11402	11930	12021	12108	11545
60	9,7	10	9,8	9,7	9,6	10,3	2,4	2,4	2,3	2,3	2,3	2,5	12459	11688	12201	12304	12360	11707
70	10,4	10,9	10,6	10,4	10,3	11	2,4	2,5	2,4	2,4	2,4	2,5	12607	11869	12363	12469	12530	11833
80	10,9	11,5	11,1	11	10,9	11,5	2,5	2,5	2,5	2,5	2,5	2,6	12676	11947	12432	12556	12607	11891
90	11,5	12,1	11,7	11,5	11,5	12,2	2,6	2,6	2,6	2,6	2,6	2,7	12716	11996	12477	12600	12652	11936
0-10	4,8	5,4	4,8	4,8	4,8	5,6	1,6	1,7	1,5	1,6	1,6	1,8	6365	6769	5965	6224	6319	7295
10-20	9,3	10,6	9,4	9,4	9,4	11,1	2,4	2,6	2,4	2,5	2,5	2,6	3186	2266	3197	3104	3111	2231
20-30	14,6	14,6	14,1	14,1	14,1	17,3	2,8	2,8	2,8	2,8	2,8	3,4	1688	1244	1637	1618	1594	1100
30-40	22,2	19,8	20	20,6	20,3	24,4	4,3	4,1	4,0	4,2	4,3	4,5	600	464	543	604	515	484
40-50	32,8	26,4	27,6	27,9	27,5	35,5	5,6	4,6	4,9	4,9	4,8	5,8	475	659	588	471	569	435
50-60	49,0	41,5	42,7	41,4	42,1	53	8,9	7,4	7,4	7,2	7,0	8,5	145	286	271	283	252	162
60-70	68,3	63,4	65,9	62,3	63,6	78,5	9,5	9,5	10,5	9,0	9,5	10,1	148	181	162	165	170	126
70-80	112	101	108	102,8	104	124	15,8	12,1	14,8	15,8	15,1	16	69	78	69	87	77	58
80-90	182	171	171	167	167	199	28,1	27,1	25	26,3	25,2	26,6	40	49	45	44	45	45
90-100	323	426	372	321	329	421	97,9	83,8	92	89,5	90,4	102	34	36	36	37	38	33

Таблиця 9

**Характеристики згорання для тесту Івм01
за різних параметрів згорання на різних рівнях згорання**

Відсоток згорнутих елементів	Середня кількість зовнішніх зв'язків		Середня кількість внутрішніх зв'язків		Кількість згорнутих пар	
	найкращі	20%	найкращі	20%	найкращі	20%
10	4,8	4,7	1,6	1,6	6365	5977
20	6,3	6,7	1,9	1,9	9551	9513
30	7,5	7,5	2,0	2	11239	11206
40	8,3	8,6	2,2	2,2	11839	12016
50	9,2	9,6	2,3	2,3	12314	12438
60	9,7	10,3	2,4	2,4	12459	12605
70	10,4	11	2,4	2,5	12607	12684
80	10,9	11,5	2,5	2,6	12676	12722
90	11,5	11,9	2,6	2,7	12716	12741
0-10	4,8	4,7	1,6	1,6	6365	5977
10-20	9,3	8,9	2,4	2,4	3186	3536
20-30	14,6	14,4	2,8	2,9	1688	1693
30-40	22,2	23,3	4,3	4,4	600	810
40-50	32,8	38,8	5,6	6,6	475	422
50-60	49,0	65,7	8,9	10	145	167
60-70	68,3	113	9,5	18,2	148	79
70-80	112,6	192	15,8	32,4	69	38
80-90	182	288	28,1	75	40	19
90-100	323	357	97,9	157	34	9

Значення критеріїв оцінки якості кластеризації за різних критеріїв згортання $K_1=12$ (4-й критерій), $K_2=831$ (4-й критерій), $K_3=679$ (4-й критерій), $K_4=4.20$ (2-й критерій), $K_5=152$ (4-й критерій), $K_6=1,1$ (2-й критерій).

Табл. 5 відображає характеристики кластеризації за різних параметрів фільтрації довгих зв'язків і плаваючих елементів:

бф – без фільтрації;

фільтрація зв'язків, довжина яких більша за 5, 7, 10, 15, 20;

фе – фільтрація плаваючих елементів.

Значення критеріїв оцінки якості кластеризації, при різних параметрах фільтрації довгих зв'язків і плаваючих елементів $K_1=12,3$ (без фільтрації), $K_2=969$ (фільтрація зв'язків, довжина яких більша за 20), $K_3=732$ (фільтрація зв'язків, довжина яких більша за 20), $K_4=4.30$; (без фільтрації), $K_5=34$ (фільтрація плаваючих елементів), $K_6=1.07$ (фільтрація зв'язків, довжина яких більша за 20).

Значення критеріїв оцінки якості кластеризації, за різних параметрів згортання $K_1=12,2$ (20 % – вимушене згортання), $K_2=1055$ (вільне згортання), $K_3=792$ (вільне згортання), $K_4=4.30$ (вільне згортання), $K_5=144$ (20 % – вимушене згортання), $K_6=1,27$ (вільне згортання).

Висновок

Для дослідження якості кластеризації були підібрані схеми із набору тестових файлів Ibm01 – Ibm18 з різними параметрами вхідних даних. Із результатів досліджень випливають такі висновки:

- Застосування фільтрації дуже довгих зв'язків, а також фільтрації плаваючих елементів приводить до підвищення швидкодії алгоритму, не погіршуючи якості кластеризації, а інколи приводить до підвищення швидкодії і якості кластеризації.

- Застосування критеріїв згортання η_1 , η_2 і комбінованого критерію з η_1 і η_2 дають хороші результати швидкодії та якості кластеризації для всіх тестів. Критерії згортання η_1 і комбінований критерій дають кращі результати якості кластеризації для тестів з більшою кількістю довгих зв'язків.

На основі виконаних тестів можна зробити висновок про ефективність розробленого алгоритму та програмного забезпечення і доцільність використання розвинутого алгоритму для розв'язування задач великої розмірності.

1. *Базилевич Р.П. Декомпозиционные и топологические методы автоматизированного конструирования электронных устройств. – Львов: Вища школа. Изд-во при Львов. гос. ун-те, Львів. 1981. – 168 с.* 2. *Базилевич Р.П., Подольський І.В. Особливості організації пакету програм для ієрархічної кластеризації схем // Вісник НУЛП “Радіоелектроніка та телекомунікації”, № 440, 2002. – Львів. – 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, НАН України, № 3, 2002. – С. 474–483.* 4. *Charles J. Alpert, The ISPD98 Circuit Benchmark Suite. ISPD98 Monterey CA USA, 1998.*