

บทที่ 5

ขั้นตอนวิธีสำหรับปัญหาการคัดลอกข้อมูลที่น้อยที่สุดที่กำหนดความต้องการของผู้ใช้

และผลการทดลอง

ในบทที่ 4 ได้พิสูจน์ไปแล้วว่าปัญหาการคัดลอกข้อมูลที่น้อยที่สุดที่กำหนดความต้องการของผู้ใช้ (Minimum Replication Problem with User Preferences) หรือ MRPUP มีความยากอย่างน้อยเท่ากับปัญหาเซตคัฟเวอร์ (Set Cover) นั่นก็หมายความว่าไม่มีขั้นตอนวิธีที่ให้คำตอบที่แน่นอนได้ภายในเวลาที่สมเหตุสมผล ดังนั้นในบทนี้จึงมุ่งความสนใจไปยังขั้นตอนวิธีที่เหมาะสม หรือขั้นตอนวิธีที่หยุดทำงานภายในเวลาแบบพหุนาม แต่ให้เซตคำตอบที่ใกล้เคียงกับคำตอบที่ดีที่สุดแทน

5.1 ขั้นตอนวิธี MRPUP

ในขั้นตอนของการออกแบบขั้นตอนวิธีมีสมมติฐานว่า $u^* = m$ เมื่อ m คือจำนวนผู้ใช้ทั้งหมดในระบบ และพ่อนปรนเงื่อนไข $e(U, T, V(B)) \leq p^*$ ซึ่งสมมติฐานนี้ไม่ได้ขัดแย้งกับนิยามของปัญหา ซึ่งหากขั้นตอนวิธีสามารถทำงานได้กับกรณีที่ $u^* = m$ ขั้นตอนวิธีนี้ก็สามารถทำงานได้กับกรณีที่ $u^* < m$ เช่นกัน ขั้นตอนวิธีนี้เรียกว่า ขั้นตอนวิธี MRPUP (MRPUP algorithm) ซึ่งกำหนดดังต่อไปนี้

MRPUP ALGORITHM (M, T)

- 1 {Input: model M and set T of preference lists}
- 2 {Output: set B of computers}
- 3 begin
- 4 repeat
- 5 pick $u_i \in U$ for some i
- 6 find $c_j \in C$ that minimizes $\rho(u_i, t_i, \Phi(c_j))$ over C
- 7 $B = B \cup \{c_j\}$
- 8 $U = U - \Phi(c_j)$
- 9 until $U = \emptyset$
- 10 end

ขั้นตอนวิธี MRPUP ออกแบบโดยใช้หลักการของขั้นตอนวิธีเชิงละโมบ (Greedy Algorithm) ซึ่งมีความซับซ้อนของเวลาในการทำงานเป็นแบบพหุนาม ต่อไปจะแสดงว่าขั้นตอน

วิธีนี้ทำงานได้ถูกต้องและให้คำตอบที่เหมาะสมสำหรับปัญหาการคัดลอกข้อมูลที่น้อยที่สุดที่กำหนดความต้องการของผู้ใช้ กำหนดให้ $\Phi(B) = \bigcup_{c \in B} \Phi(c)$

ทฤษฎีบท 5.1 ขั้นตอนวิธี MRPUP ให้เซตคำตอบ B ของคอมพิวเตอร์ซึ่ง $|\Phi(B)| = m$

พิสูจน์ โดยใช้อุปนัยทางคณิตศาสตร์กับตัวแปร m

Basis step

เมื่อ $m = 1$ นั่นคือ $U = \{u_1\}$ ขั้นตอนวิธีในบรรทัดที่ 4 ถึง 9 จะหา c_j ซึ่ง $\Phi(c_j) = \{u_1\}$ และเพิ่ม c_j เข้าไปในเซตคำตอบ B ดังนั้นขั้นตอนวิธี MRPUP ทำงานได้ถูกต้องสำหรับ $m = 1$

Inductive step

กำหนดให้ a เป็นค่าคงที่บวกใดๆ สมมติให้ขั้นตอนวิธีให้เซต B โดยที่ $|\Phi(B)| = m - a$ เมื่อขั้นตอนวิธีในบรรทัดที่ 9 ทำงาน ได้ว่าขั้นตอนวิธีนั้นจะทำงานในบรรทัดที่ 5 ถึง 8 เพื่อหา c_j ซึ่ง $|\Phi(c_j)| = a$ แล้วเพิ่ม c_j เข้าไปในเซต B และทำซ้ำจนกว่าเซต U จะเป็นเซตว่าง ขั้นตอนวิธีจะหยุดทำงานและให้เซตคำตอบ B โดยที่ $|\Phi(B)| = m$

จากหลักการอุปนัยทางคณิตศาสตร์ ได้ว่า ขั้นตอนวิธี MRPUP ให้เซตคำตอบ B ของคอมพิวเตอร์ซึ่ง $|\Phi(B)| = m$ สำหรับทุก $m \in \mathbb{N}$ ■

ขั้นตอนวิธี MRPUP ให้เซตคำตอบที่เหมาะสมโดยที่ผู้ใช้แต่ละคนสามารถเข้าถึงข้อมูลบนเครื่องคอมพิวเตอร์อย่างน้อย 1 เครื่องในเซต B แต่ค่าจากฟังก์ชันความต้องการของผู้ใช้ทั้งหมดก็ยังไม่ใช่ค่าที่น้อยที่สุด เนื่องจากยังไม่มีขั้นตอนวิธีเชิงประมาณที่ดีสำหรับปัญหานี้ ขั้นตอนวิธีที่นำเสนอข้างบนเป็นขั้นตอนวิธีที่ดีที่สุดสำหรับเวลานี้ โดยใช้ผลจากการทดลองสำหรับยืนยันความเหมาะสมของขั้นตอนวิธีที่นำเสนอ

5.2 ผลการทดลอง

ในการทดลองเซตคอมพิวเตอร์ C เซตผู้ใช้ U เซตรายการความต้องการของผู้ใช้ T และฟังก์ชันการเข้าถึง Φ ถูกสร้างขึ้นแบบสุ่ม และขนาดของเซต C กำหนดให้ไม่เกิน 25 แต่ละกรณีตัวอย่างที่ใช้ในการทดลองแบ่งเป็นกลุ่มตามจำนวนของผู้ใช้คือ 20 30 40 50 และ 60 คน และแต่ละกลุ่มมีกรณีตัวอย่าง 20 กรณีสำหรับทำการทดลอง ในการทดลองนี้สร้างฟังก์ชันการเข้าถึง Φ โดยเลือกผู้ใช้ จากนั้นเลือกจำนวนและเครื่องคอมพิวเตอร์ที่เขาสามารถใช้งานได้อย่างสุ่ม และทำอย่างนี้สำหรับทุกผู้ใช้ $u \in U$ เซตรายการความต้องการของผู้ใช้สร้างขึ้นโดยใช้การเรียงสับเปลี่ยนแบบ

ส่วนและสุดท้ายกำหนดเขต P S O และ A เช่นเดียวกับทฤษฎีบท 3.3 จากนั้นทำการทดลองเพื่อหาค่าเฉลี่ยของค่าที่ได้จากฟังก์ชัน $e(U, T, H)$ ที่ได้จากเขตคำตอบของขั้นตอนวิธี MRPU P เปรียบเทียบกับค่าที่ได้จากเขตคำตอบของขั้นตอนวิธีแบบบรูทฟอส (brute - force algorithm) ซึ่งได้ผลการทดลองดังตารางต่อไปนี้

ครั้งที่ ที่	$e(U, T, H)$		$\frac{e(U, T, H)_{GR}}{e(U, T, H)_{BF}}$	ครั้งที่ ที่	$e(U, T, H)$		$\frac{e(U, T, H)_{GR}}{e(U, T, H)_{BF}}$
	Greedy	Brute Force			Greedy	Brute Force	
1	2826	2509	1.1263452	11	3305	3122	1.0586163
2	2231	1970	1.1324873	12	2192	2192	1.0000000
3	3664	3326	1.1016236	13	3130	2743	1.1410864
4	2404	2404	1.0000000	14	2257	1934	1.1670114
5	3258	2847	1.1443625	15	3620	2493	1.4520658
6	3196	2916	1.0960219	16	5109	3221	1.5861534
7	3231	2829	1.1420997	17	2853	2287	1.2474858
8	3002	3002	1.0000000	18	3146	2122	1.4825636
9	2971	2162	1.3741906	19	3445	2202	1.5644868
10	3062	2463	1.2431994	20	4266	3089	1.3810295
เฉลี่ย							1.2220414

ตาราง 5.1: ผลการทดลองในกรณีตัวอย่างที่มีผู้ใช้ 20 คน

ครั้งที่ ที่	$e(U, T, H)$		$\frac{e(U, T, H)_{GR}}{e(U, T, H)_{BF}}$	ครั้งที่ ที่	$e(U, T, H)$		$\frac{e(U, T, H)_{GR}}{e(U, T, H)_{BF}}$
	Greedy	Brute Force			Greedy	Brute Force	
1	13946	12643	1.1030610	11	12155	7983	1.5226105
2	8787	7045	1.2472676	12	13341	9955	1.3401306
3	9519	8187	1.1626970	13	11267	8270	1.3623942
4	14293	9735	1.4682075	14	14454	9878	1.4632517
5	18171	11396	1.5945068	15	10046	8510	1.1804935
6	11569	7680	1.5063802	16	11699	9996	1.1703681
7	10403	7111	1.4629447	17	13125	9961	1.3176388
8	13049	10057	1.2975042	18	11551	8771	1.3169536
9	9852	8205	1.2007313	19	9388	8237	1.1397353
10	12955	9257	1.3994815	20	8850	7824	1.1311350
เฉลี่ย							1.3193747

ตาราง 5.2: ผลการทดลองในกรณีตัวอย่างที่มีผู้ใช้ 30 คน

ครั้งที่ ที่	e(U, T, H)		$\frac{e(U, T, H)_{GR}}{e(U, T, H)_{BF}}$	ครั้งที่ ที่	e(U, T, H)		$\frac{e(U, T, H)_{GR}}{e(U, T, H)_{BF}}$
	Greedy	Brute Force			Greedy	Brute Force	
1	40412	26994	1.4970734	11	41826	27100	1.5433948
2	32177	23657	1.3601471	12	46152	30376	1.5193574
3	33535	21635	1.5500347	13	31825	27183	1.1707685
4	31755	21664	1.4657958	14	23376	16533	1.4138995
5	27335	22524	1.2135944	15	43394	26751	1.6221450
6	36955	26734	1.3823221	16	27256	21822	1.2490148
7	29821	20668	1.4428585	17	42953	34458	1.2465320
8	25697	21028	1.2220373	18	42235	32070	1.3169629
9	34958	23437	1.4915732	19	32158	25385	1.2668111
10	22490	20762	1.0832290	20	30786	20926	1.4711842
เฉลี่ย							1.3764368

ตาราง 5.3: ผลการทดลองในกรณีตัวอย่างที่มีผู้ใช้ 40 คน

ครั้งที่ ที่	e(U, T, H)		$\frac{e(U, T, H)_{GR}}{e(U, T, H)_{BF}}$	ครั้งที่ ที่	e(U, T, H)		$\frac{e(U, T, H)_{GR}}{e(U, T, H)_{BF}}$
	Greedy	Brute Force			Greedy	Brute Force	
1	56859	47234	1.2037727	11	81972	55810	1.4687690
2	82002	49433	1.6588514	12	72951	56526	1.2905742
3	62968	41896	1.5029597	13	98674	64104	1.5392799
4	88885	65947	1.3478248	14	77831	47959	1.6228654
5	63072	53072	1.1884233	15	62879	48894	1.2860269
6	86301	62662	1.3772462	16	70833	56456	1.2546585
7	79305	49386	1.6058195	17	61078	51213	1.1926269
8	65444	44238	1.4793616	18	78058	56635	1.3782643
9	65726	47259	1.3907615	19	68808	56112	1.2262618
10	78904	62086	1.2708823	20	80225	70802	1.1330895
เฉลี่ย							1.3709160

ตาราง 5.4: ผลการทดลองในกรณีตัวอย่างที่มีผู้ใช้ 50 คน

ครั้งที่	e(U, T, H)		$\frac{e(U, T, H)_{GR}}{e(U, T, H)_{BF}}$	ครั้งที่	e(U, T, H)		$\frac{e(U, T, H)_{GR}}{e(U, T, H)_{BF}}$
	Greedy	Brute Force			Greedy	Brute Force	
1	119043	91097	1.3067719	11	164049	131008	1.2522060
2	210180	139593	1.5056629	12	159302	133487	1.1933896
3	139835	95483	1.4645015	13	129848	102598	1.2655997
4	174494	118561	1.4717656	14	153575	118420	1.2968671
5	143326	121588	1.1787841	15	132363	103197	1.2826245
6	116486	89378	1.3032961	16	149304	107303	1.3914243
7	118811	93033	1.2770845	17	119735	100413	1.1924253
8	157538	101956	1.5451567	18	145079	105814	1.3710757
9	113207	85303	1.3271163	19	134611	93978	1.4323671
10	166720	113327	1.4711410	20	113742	87972	1.2929341
เฉลี่ย							1.3411097

ตาราง 5.5: ผลการทดลองในกรณีตัวอย่างที่มีผู้ใช้ 60 คน

ข้อสังเกตค่าที่ได้จากการเปรียบเทียบนั้นถ้ามีค่าใกล้กับ 1 มากเท่าไร ก็หมายความว่าขั้นตอนวิธีนั้นให้คำตอบที่ใกล้กับคำตอบที่ดีที่สุดมากเท่านั้น และขั้นตอนวิธีที่นำเสนอให้ผลลัพธ์โดยเฉลี่ยอยู่ที่ 1.38 เท่าของคำตอบที่ดีที่สุด