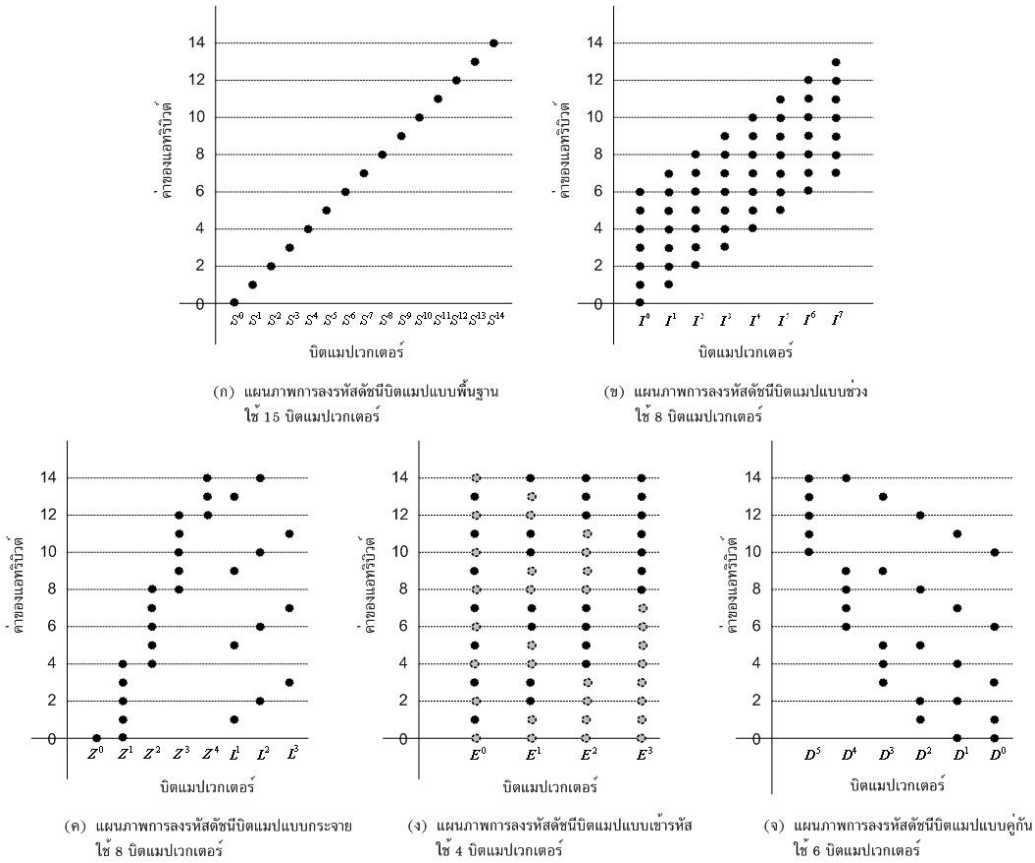
บทที่ 5

การวิเคราะห์และผลการทดลอง

สำหรับบทนี้จะกล่าวถึงการวิเคราะห์เปรียบเทียบค่าใช้จ่ายเกี่ยวกับพื้นที่ใช้ในการจัดเก็บดัชนี ค่าใช้จ่ายเกี่ยวกับเวลาที่ใช้ในการสอบถามแบบค่าเท่ากัน และประสิทธิภาพในแง่ Space-Time Trade-off (การแลกเปลี่ยนระหว่างประสิทธิภาพของพื้นที่กับเวลา) ของดัชนีบิตแมปแบบคู่กันคิดค้นขึ้นและดัชนีบิตแมปที่เคยมีมาทั้ง 4 ชนิด ได้แก่ ดัชนีบิตแมปแบบพื้นฐาน ดัชนีบิตแมปแบบช่วง ดัชนีบิตแมปแบบกระจาย และดัชนีบิตแมปแบบเข้ารหัส

5.1 ค่าใช้จ่ายเกี่ยวกับพื้นที่ที่ใช้ในการจัดเก็บดัชนี

จากที่ได้กล่าวมา ในบทที่ 3 และ 4 จะเห็นว่าดัชนีบิตแมปแต่ละชนิดมีการลงรหัสและสอบถามที่แตกต่างกัน พิจารณารูปแบบการลงรหัสที่แตกต่างกันของดัชนีบิตแมปทั้ง 5 ชนิดดังภาพประกอบ 5-1



ภาพประกอบ 5-1 แสดงแผนภาพการลงรหัสดัชนีบิตแมปแบบทั้ง 5 ชนิด เมื่อ คาร์ดินอลิตี้ (C) เท่ากับ 15

จากภาพประกอบ 5-1 การทำดัชนีบิตแมปแบบพื้นฐานใช้พื้นที่มากที่สุด คือ ในที่นี้คาร์ดินอลิตี้เท่ากับ 15 ดัชนีบิตแมปแบบพื้นฐานใช้ 15 บิตแมปเวกเตอร์ดังภาพประกอบ 5-1 (ก) ส่วนดัชนีบิตแมปแบบเข้ารหัสใช้พื้นที่ที่สุด คือ 4 บิตแมปเวกเตอร์ดังภาพประกอบ 5-1 (ง) ดัชนีบิตแมปแบบช่วงใช้พื้นที่เป็นครึ่งหนึ่งของดัชนีบิตแมปแบบพื้นฐาน คือ 8 บิตแมปเวกเตอร์ดังภาพประกอบ 5-1 (ข) แต่ยังใช้พื้นที่มากกว่าดัชนีบิตแมปแบบเข้ารหัส จากภาพประกอบ 5-1 (ข) และ 5-1 (ค) จะเห็นว่า ดัชนีบิตแมปแบบกระจายและดัชนีบิตแมปแบบช่วงใช้ 8 บิตแมปเวกเตอร์เท่ากัน ซึ่งในกรณีที่คาร์ดินอลิตี้เท่ากับ 15 นี้จะยังใช้พื้นที่เท่ากันอยู่ แต่เมื่อคาร์ดินอลิตี้มีค่ามากขึ้น ดัชนีบิตแมปแบบกระจายจะใช้พื้นที่น้อยกว่าดัชนีบิตแมปแบบช่วง ($\lceil 2\sqrt{C} \rceil < \lceil C/2 \rceil$ เมื่อ $C \geq 19$) แต่ยังคงใช้พื้นที่มากกว่าดัชนีบิตแมปแบบเข้ารหัส สำหรับดัชนีบิตแมปแบบคู่กันที่คิดค้นขึ้น ใช้พื้นที่น้อยกว่าดัชนีบิตแมปแบบกระจาย คือ 6 บิตแมปเวกเตอร์ดังภาพประกอบ 5-1 (จ) แต่ใช้พื้นที่มากกว่าดัชนีบิตแมปแบบเข้ารหัส

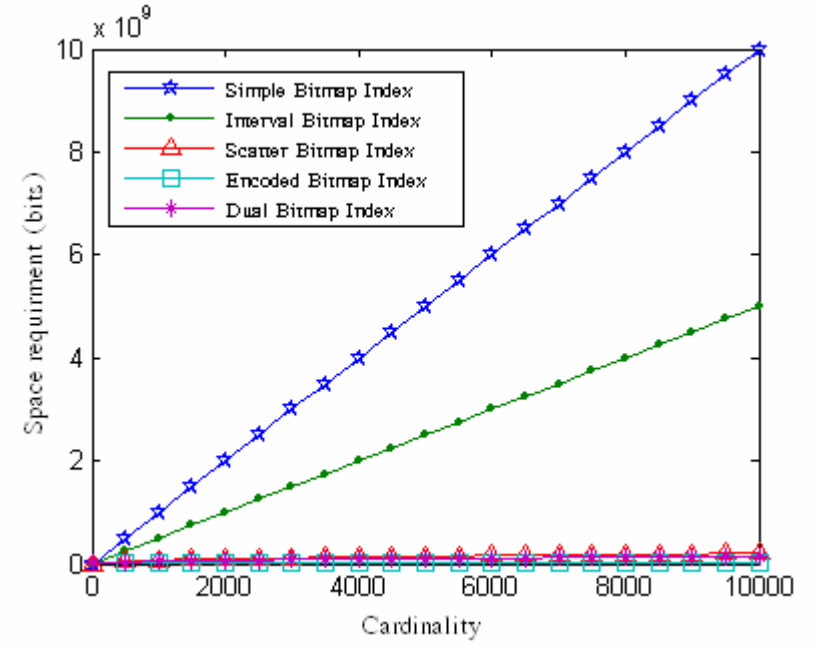
ต่อไปเราจะพิจารณาเปรียบเทียบพื้นที่ที่ใช้ในการจัดเก็บดัชนีของดัชนีบิตแมปทั้ง 5 ชนิดในรูปทั่วไป ดังตาราง 5-1

ตาราง 5-1 แสดงพื้นที่ที่ใช้ในการจัดเก็บดัชนีบิตแมปทั้ง 5 ชนิด เมื่อ C คือ คาร์ดินอลิตี้ของแอทริบิวต์ และ N คือ จำนวนเรคอร์ดของแอทริบิวต์

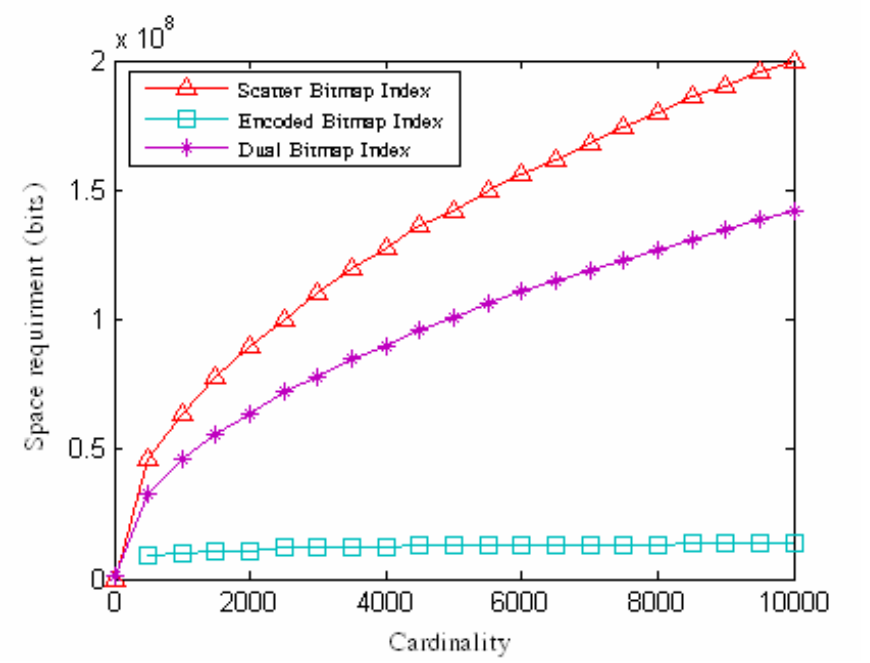
ชนิดของดัชนีบิตแมป	พื้นที่ที่ใช้ในการจัดเก็บดัชนี (บิต)
ดัชนีบิตแมปแบบพื้นฐาน	CN
ดัชนีบิตแมปแบบช่วง	$\lceil C/2 \rceil N$
ดัชนีบิตแมปแบบกระจาย	$\lceil 2\sqrt{C} \rceil N$
ดัชนีบิตแมปแบบเข้ารหัส	$\lceil \log_2 C \rceil N$
ดัชนีบิตแมปแบบคู่กัน	$\lceil \sqrt{2C + 0.25} + 0.5 \rceil N$

จากตาราง 5-1 จะเห็นได้ว่า สำหรับการทำดัชนีบิตแมปบนแอทริบิวต์ที่มี N เรคอร์ด พื้นที่ที่ใช้ในการจัดเก็บดัชนีบิตแมปแต่ละชนิดจะแปรผันตรงกับคาร์ดินอลิตี้ (C) ของแอทริบิวต์ที่นำมาทำดัชนี

ต่อไปเราจะเปรียบเทียบให้เห็นการใช้พื้นที่ในการจัดเก็บดัชนีของดัชนีบิตแมปแต่ละชนิด โดยพิจารณาจากกราฟดังภาพประกอบ 5-2



ภาพประกอบ 5-2 กราฟแสดงการเปรียบเทียบพื้นที่ที่ใช้ในการสร้างดัชนีบิตแมปทั้ง 5 ชนิด เมื่อแอทริบิวต์ที่นำมาทำดัชนีมี 1,000,000 เรคอร์ด ($N = 1{,}000{,}000$)



ภาพประกอบ 5-3 กราฟแสดงการเปรียบเทียบพื้นที่ที่ใช้ในการสร้างดัชนีบิตแมปทั้ง 3 ชนิด เมื่อแอทริบิวต์ที่นำมาทำดัชนีมี 1,000,000 เรคอร์ด ($N = 1{,}000{,}000$)

พิจารณาภาพประกอบ 5-2 แสดงความสัมพันธ์ระหว่างพื้นที่ที่ใช้ในการสร้างดัชนีกับคาร์ดินอลิตี้ของแอทริบิวต์ที่เลือกมาทำดัชนีบิตแมปทั้ง 5 ชนิด ซึ่งคาร์ดินอลิตี้มีค่าตั้งแต่ 0 ถึง 10,000 จะเห็นได้ว่าเราสามารถแบ่งกลุ่มดัชนีบิตแมปตามการใช้พื้นที่จัดเก็บดัชนีได้เป็น 2 กลุ่ม คือ

- **กลุ่มที่ใช้พื้นที่มาก** ได้แก่ ดัชนีบิตแมปแบบพื้นฐานและดัชนีบิตแมปแบบช่วง โดยดัชนีบิตแมปแบบช่วงใช้พื้นที่น้อยกว่าดัชนีบิตแมปแบบพื้นฐาน
- **กลุ่มที่ใช้พื้นที่น้อย** ได้แก่ ดัชนีบิตแมปแบบเข้ารหัส, ดัชนีบิตแมปแบบกระจาย และดัชนีบิตแมปแบบคู่กัน เมื่อพิจารณาภาพประกอบ 5-3 ซึ่งแสดงการเปรียบเทียบการใช้พื้นที่ในการสร้างดัชนีบิตแมปในกลุ่มที่ใช้พื้นที่น้อย จะเห็นว่าดัชนีบิตแมปแบบเข้ารหัสใช้พื้นที่น้อยที่สุด รองลงมาคือดัชนีบิตแมปแบบคู่กันและดัชนีบิตแมปแบบกระจายตามลำดับ

โดยสรุปแล้ว หากต้องการทำดัชนีบิตแมปมีประสิทธิภาพในการใช้พื้นที่มากที่สุด ควรเลือกใช้ดัชนีบิตแมปแบบเข้ารหัส เพราะใช้พื้นที่ในการจัดเก็บดัชนีน้อยที่สุด

5.2 ค่าใช้จ่ายเกี่ยวกับเวลาที่ใช้ในการสอบถามแบบค่าเท่ากัน

ดัชนีบิตแมปแต่ละชนิดมีรูปแบบการลงรหัสที่แตกต่างกัน นอกจากจะทำให้ใช้พื้นที่ในการจัดเก็บดัชนีต่างกันแล้ว ยังทำให้ขั้นตอนการสอบถามแบบค่าเท่ากันแตกต่างกันด้วย ทั้งจำนวนบิตแมปเวกเตอร์ที่อ่านและจำนวนครั้งในการดำเนินการตรรกะระหว่างบิตแมปเวกเตอร์เป็นผลให้เวลาที่ใช้ในการสอบถามแบบค่าเท่ากันต่างกันด้วย เวลาในการสอบถามแบบค่าเท่ากันประกอบด้วย เวลาที่ใช้ในการอ่านบิตแมปเวกเตอร์ ถ้าต้องอ่านหลายบิตแมปเวกเตอร์ก็จะทำให้ต้องใช้เวลามาก และเวลาที่ใช้ในการดำเนินการตรรกะระหว่างบิตแมปเวกเตอร์ ถ้าต้องดำเนินการหลายครั้งย่อมต้องใช้เวลามากเช่นกัน พิจารณาจำนวนบิตแมปเวกเตอร์ที่อ่านและจำนวนครั้งในการดำเนินการตรรกะระหว่างบิตแมปเวกเตอร์เมื่อมีการสอบถามแบบค่าเท่ากันของดัชนีบิตแมปทั้ง 5 ชนิด ดังตาราง 5-2

ตาราง 5-2 แสดงจำนวนบิตแมปเวกเตอร์ที่อ่านและจำนวนครั้งในการดำเนินการตรรกะระหว่างบิตแมปเวกเตอร์เมื่อมีการสอบถามแบบค่าเท่ากันของดัชนีบิตแมปทั้ง 5 ชนิด

ชนิดของดัชนีบิตแมป	จำนวนบิตแมปเวกเตอร์ที่อ่าน	จำนวนครั้งในการดำเนินการตรรกะระหว่างบิตแมปเวกเตอร์
ดัชนีบิตแมปแบบพื้นฐาน	1	0
ดัชนีบิตแมปแบบช่วง	2	2 (1AND,1NOT)
ดัชนีบิตแมปแบบกระจาย	2	1 (1AND)
ดัชนีบิตแมปแบบเข้ารหัส	$\lceil \log_2 C \rceil$	เทียบค่ากับตารางการเทียบค่า
ดัชนีบิตแมปแบบคู่กัน	2	1 (1AND)

จากตาราง 5-2 จะเห็นว่า ในการสอบแบบค่าเท่ากัน ดัชนีบิตแมปแบบพื้นฐานมีจำนวนบิตแมปเวกเตอร์ที่อ่านน้อยที่สุด คือ 1 บิตแมปเวกเตอร์ และไม่ต้องมีการดำเนินการตรรกะใด ๆ ดังนั้นเวลาที่ใช้ในการสอบถามแบบค่าเท่ากันบนดัชนีบิตแมปแบบพื้นฐานจึงประกอบด้วย เวลาที่ใช้ในการอ่านบิตแมปเวกเตอร์เพียงบิตแมปเวกเตอร์เดียวเท่านั้น ในทางตรงกันข้าม ดัชนีบิตแมปแบบเข้ารหัสมีจำนวนบิตแมปเวกเตอร์ที่อ่านมากที่สุด คือ $\lceil \log_2 C \rceil$ และใช้การเทียบค่ากับตารางการเทียบค่า เป็นผลให้เวลาที่ใช้ในการสอบถามแบบค่าเท่ากันบนดัชนีบิตแมปแบบเข้ารหัสประกอบด้วย เวลาในการอ่านบิตแมปเวกเตอร์ที่มากกว่าดัชนีบิตแมปชนิดอื่น และเวลาในการเทียบค่ากับตารางการเทียบซึ่งจะต้องเวลานานมากกว่ากันดำเนินการตรรกะระหว่างบิตแมปเวกเตอร์ทั่วไป ส่วนการสอบถามแบบค่าเท่ากันบนดัชนีบิตแมปแบบช่วง ดัชนีบิตแมปแบบกระจาย และดัชนีบิตแมปแบบคู่กัน มีจำนวนบิตแมปเวกเตอร์ที่อ่านเท่ากับ 2 จึงใช้เวลาในการอ่านบิตแมปเวกเตอร์เท่ากัน แต่จำนวนครั้งในการดำเนินการตรรกะไม่เท่ากัน กล่าวคือ กรณีแย่ที่สุดของการสอบถามแบบค่าเท่ากันบนดัชนีบิตแมปแบบช่วงต้องมีการดำเนินการตรรกะระหว่างบิตแมปเวกเตอร์ 2 ครั้ง ในขณะที่ดัชนีบิตแมปแบบกระจายและดัชนีบิตแมปแบบคู่กันมีการดำเนินการตรรกะระหว่างบิตแมปเวกเตอร์ 1 ครั้งเสมอ จึงเป็นผลให้ในบางกรณีการสอบถามแบบค่าเท่ากันบนดัชนีบิตแมปแบบช่วงต้องใช้เวลาในส่วนของการดำเนินการตรรกะมากกว่าดัชนีบิตแมปแบบกระจายและดัชนีบิตแมปแบบคู่กัน อย่างไรก็ตามในกรณีดีที่สุดของการสอบถามแบบค่าเท่ากันบนดัชนีบิตแมปแบบช่วงมีการอ่านเพียงบิตแมปเวกเตอร์เดียวและไม่ต้องมีการดำเนินการตรรกะใด ๆ ซึ่งกรณีนี้ดัชนีบิตแมปแบบช่วงจะใช้เวลาน้อยกว่าดัชนีบิตแมปแบบกระจายและดัชนีบิตแมปแบบคู่กัน

ที่กล่าวมานี้เป็นการเปรียบเทียบค่าใช้จ่ายเกี่ยวกับเวลาที่ใช้ในการสอบถามแบบค่าเท่ากันด้วยการวิเคราะห์ทางทฤษฎีโดยพิจารณาจากรูปแบบการสอบถามของดัชนีบิตแมปแต่ละชนิด ต่อไปเราจะทำการเปรียบเทียบค่าใช้จ่ายเกี่ยวกับเวลาที่ใช้ในการสอบถามแบบค่าเท่ากัน

จากผลการทดลอง โดยทำการทดลองบนเครื่องคอมพิวเตอร์รุ่น Intel(R) Celeron(R) ที่มีหน่วยประมวลผลกลางขนาด 1.69 GHz หน่วยความจำ 512 MB ระบบปฏิบัติการ Linux Red Hat 9.0 และใช้ตัวแปลภาษาซี(C Compiler) ในการเขียนโปรแกรม

ข้อมูลที่ใช้ในการทดสอบเป็นข้อมูลมาตรฐานจาก TPC-H Benchmark [25] โดยแอทริบิวต์ที่เลือกมาสร้างดัชนีมีคาร์ดินอลิตี้เท่ากับ 25, 50 และ 150 มีจำนวนเรคอร์ดเท่ากับ 1,000,000 เรคอร์ด

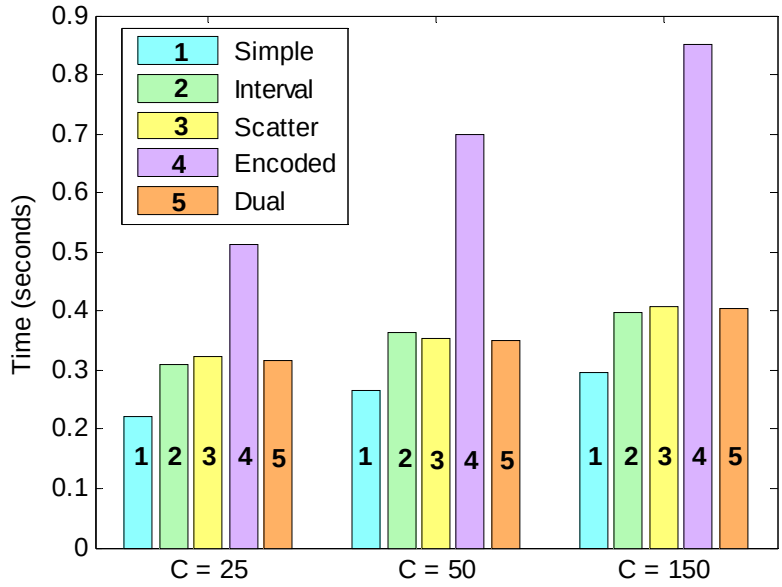
ในการทดลองกำหนดให้เวลาที่ใช้ในการสอบถามแบบค่าเท่ากันบนดัชนีบิตแมปแต่ละชนิด ประกอบด้วย การคำนวณเพื่อหาบิตแมปเวกเตอร์ที่ใช้ การอ่านบิตแมปเวกเตอร์ที่ใช้ การดำเนินการตรรกะระหว่างบิตแมปเวกเตอร์ และการบันทึกคำตอบลงในไฟล์ข้อมูล

การบันทึกผลการทดลอง ดำเนินการโดยรันโปรแกรมเพื่อสอบถามทีละค่าแล้วบันทึกเวลาที่ใช้ จนครบทุกค่า คือ ตั้งแต่ 0 ถึง $C-1$ จากนั้นหาค่าเฉลี่ยสำหรับการสอบถามค่าใด ๆ ซึ่งดำเนินการซ้ำทั้งหมด 3 รอบ ดังผลการทดลองในตาราง 5-3

ตาราง 5-3 แสดงเวลาที่ใช้ในการสอบถามแบบค่าเท่ากันบนดัชนีบิตแมปทั้ง 5 ชนิด เมื่อคาร์ดินอลิตี้มีค่าเท่ากับ 25, 50 และ 150

คาร์ดินอลิตี้(C) ครั้งที่	เวลาที่ใช้ในการสอบถาม (วินาที)				
	ดัชนีบิตแมปแบบพื้นฐาน	ดัชนีบิตแมปแบบช่วง	ดัชนีบิตแมปแบบกระจาย	ดัชนีบิตแมปแบบเข้ารหัส	ดัชนีบิตแมปแบบคู่กัน
C=25					
ครั้งที่ 1	0.2316	0.3016	0.3240	0.5188	0.3197
ครั้งที่ 2	0.2252	0.3112	0.3172	0.5041	0.3155
ครั้งที่ 3	0.2404	0.3158	0.3258	0.5156	0.3088
เฉลี่ย	0.2324	0.3095	0.3223	0.5128	0.3147
C=50					
ครั้งที่ 1	0.2746	0.3601	0.3548	0.7112	0.3491
ครั้งที่ 2	0.2602	0.3642	0.3496	0.6912	0.3494
ครั้งที่ 3	0.2648	0.3716	0.3597	0.6942	0.3519
เฉลี่ย	0.2665	0.3653	0.3547	0.6989	0.3501
C=150					
ครั้งที่ 1	0.2946	0.3899	0.4247	1.0438	0.4031
ครั้งที่ 2	0.2938	0.3872	0.4098	1.0159	0.4027
ครั้งที่ 3	0.2959	0.3933	0.4138	1.0293	0.4070
เฉลี่ย	0.2948	0.3901	0.4161	1.0297	0.4043

จากตาราง 5-3 นำค่าเฉลี่ยของเวลาที่ใช้ในการสอบถามมาเขียนกราฟได้ดังภาพประกอบ



ภาพประกอบ 5-4 กราฟแสดงเวลาที่ใช้ในการสอบถามแบบค่าเท่ากันบนดัชนีบิตแมปทั้ง 5 ชนิด เมื่อคาร์ดินอลิตี้มีค่าเท่ากับ 25, 50 และ 150

จากภาพประกอบ 5-4 ถ้าพิจารณาดัชนีบิตแมปแต่ละชนิด จะเห็นว่าเมื่อคาร์ดินอลิตี้มีค่าสูงขึ้น ต้องใช้เวลาในการสอบถามแบบค่าเท่ากันมากขึ้นด้วย แต่ถ้าพิจารณาแต่ละค่าของคาร์ดินอลิตี้ จะได้ว่า ดัชนีบิตแมปแบบพื้นฐาน(หมายเลข 1)ใช้เวลาในการสอบถามน้อยที่สุด ส่วนดัชนีบิตแมปแบบเข้ารหัส(หมายเลข 4)ใช้เวลาในการสอบถามมากที่สุด สำหรับดัชนีบิตแมปแบบช่วง(หมายเลข 2) ดัชนีบิตแมปแบบกระจาย(หมายเลข 3) และดัชนีบิตแมปแบบคู่กัน(หมายเลข 5) ใช้เวลาในการสอบถามใกล้เคียงกัน โดยดัชนีบิตแมปทั้ง 3 ชนิดใช้เวลามากกว่าดัชนีบิตแมปแบบพื้นฐาน แต่น้อยกว่าดัชนีบิตแมปแบบเข้ารหัส

เหตุผลที่ทำให้เวลาที่ใช้ในการสอบถามแบบค่าเท่ากันบนดัชนีบิตแมปแบบพื้นฐานและดัชนีบิตแมปแบบเข้ารหัสมีความแตกต่างไปจากดัชนีบิตแมปชนิดอื่นอย่างเห็นได้ชัดเจน คือ รูปแบบการสอบถามที่ต่างกับดัชนีบิตแมปชนิดอื่นอย่างสิ้นเชิง โดยเฉพาะจำนวนบิตแมปเวกเตอร์ที่อ่าน ดัชนีบิตแมปแบบพื้นฐานมีการอ่านเพียงบิตแมปเวกเตอร์เดียว และไม่มีการดำเนินการตรรกะใด ๆ จึงทำให้ใช้เวลาน้อย แต่ดัชนีบิตแมปแบบเข้ารหัสมีการอ่าน $\lceil \log_2 C \rceil$ บิตแมปเวกเตอร์ และใช้การเทียบค่ากับตารางการเทียบค่าในการหาคำตอบ จึงทำให้ใช้เวลามาก ส่วนดัชนีบิตแมปแบบช่วง ดัชนีบิตแมปแบบกระจาย และดัชนีบิตแมปแบบคู่กัน มีการอ่าน 2 บิตแมปเวกเตอร์เหมือนกัน และมีจำนวนครั้งในการดำเนินการตรรกะระหว่างบิตแมปเวกเตอร์ใกล้เคียงกัน คือ 0 ถึง 2 ครั้ง ซึ่งถือว่าไม่ต่างกันมาก จึงใช้เวลาในการสอบถาม

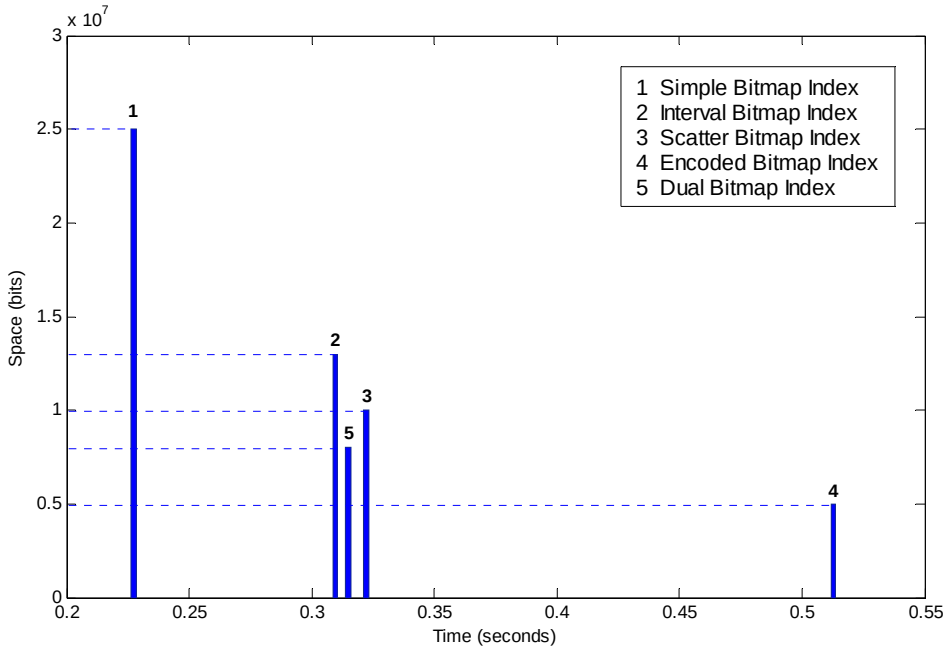
แบบค่าเท่ากันใกล้เคียงกัน และจะเห็นว่าไม่มีความแน่นอนในการเรียงลำดับเวลาที่ใช้ในการสอบถามแบบค่าเท่ากันของดัชนีบิตแมปทั้ง 3 ชนิดนี้ ทั้งนี้อาจขึ้นอยู่กับการจัดการของหน่วยความจำในการจัดเก็บดัชนีบิตแมปแต่ละชนิดด้วย เช่น ข้อมูลของดัชนีหนึ่งถูกเก็บไว้ในบล็อค(block)เดียวกัน ส่วนอีกของดัชนีหนึ่งถูกเก็บไว้คนละบล็อคกัน เป็นต้น

โดยสรุปแล้ว หากต้องการทำดัชนีบิตแมปมีประสิทธิภาพในการใช้เวลามากที่สุด ควรเลือกใช้ดัชนีบิตแมปแบบพื้นฐาน เพราะใช้เวลาในการสอบถามแบบค่าเท่ากันน้อยที่สุด

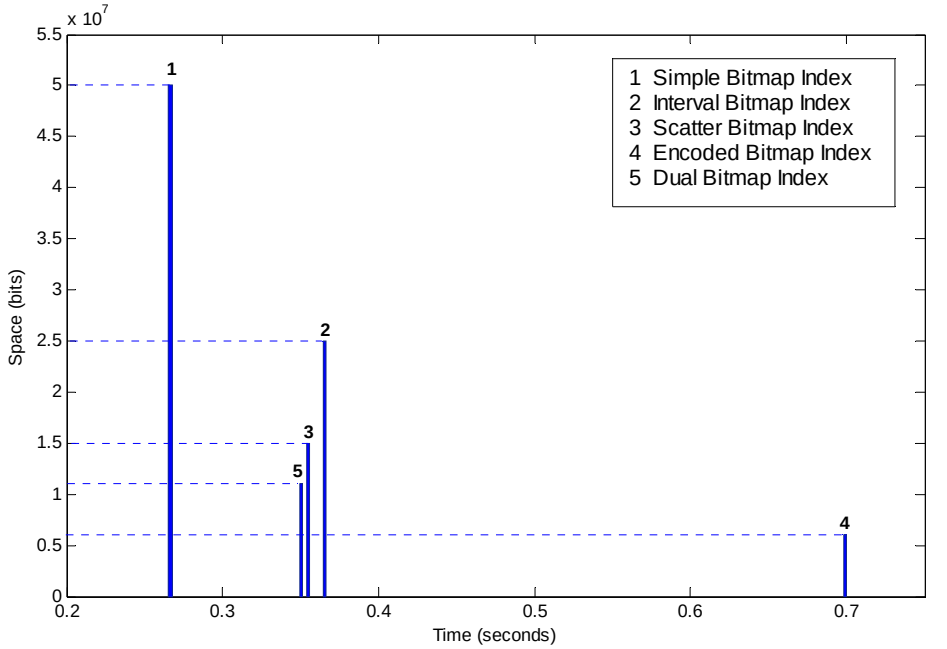
5.3 ประสิทธิภาพในแง่ Space-Time Trade-off

จากหัวข้อ 5.2 และ 5.3 จะเห็นว่า ดัชนีบิตแมปแบบเข้ารหัสมีประสิทธิภาพในแง่พื้นที่มากที่สุด แต่มีประสิทธิภาพในแง่เวลาน้อยที่สุด ส่วนดัชนีบิตแมปแบบพื้นฐานมีประสิทธิภาพในแง่เวลามากที่สุด แต่มีประสิทธิภาพในแง่พื้นที่น้อยที่สุด ดังนั้น ในการสร้างดัชนีบิตแมป หากคำนึงถึงประสิทธิภาพของพื้นที่เป็นสำคัญควรเลือกใช้ดัชนีบิตแมปแบบเข้ารหัส แต่หากคำนึงถึงประสิทธิภาพของเวลาเป็นสำคัญควรเลือกใช้ดัชนีบิตแมปแบบพื้นฐาน และอีกทางเลือกหนึ่งที่เป็นไปได้ คือ การแลกเปลี่ยนระหว่างประสิทธิภาพของพื้นที่กับเวลา หรือเรียกว่าประสิทธิภาพในแง่ Space-Time Trade-off เช่น ยอมให้มีการใช้พื้นที่มากขึ้นเพื่อให้เวลาในการสอบถามน้อยลง หรือ ยอมให้มีการใช้เวลามากขึ้นเพื่อให้ใช้พื้นที่น้อยลง เป็นต้น

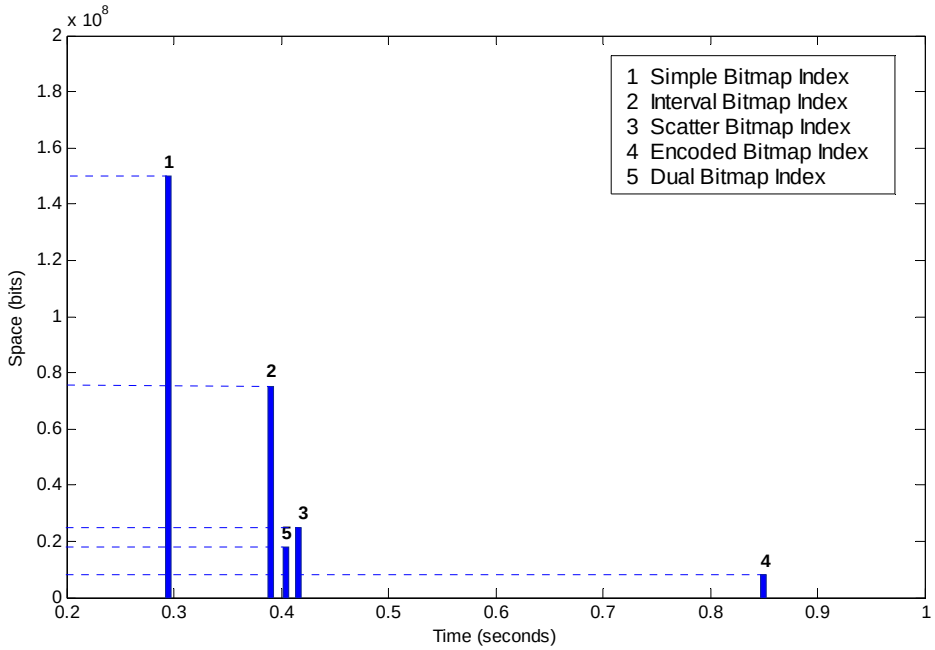
ในหัวข้อนี้ เราจะดำเนินการวิเคราะห์ประสิทธิภาพในแง่ Space-Time Trade-off ของดัชนีบิตแมปแต่ละชนิด โดยพิจารณาจากพื้นที่ใต้กราฟที่แสดงความสัมพันธ์ระหว่างพื้นที่ที่ใช้ในการสร้างดัชนีบิตแมปกับเวลาที่ใช้ในการสอบถามแบบค่าเท่ากันของดัชนีบิตแมปแต่ละชนิด ดังภาพประกอบ 5-5, 5-6 และ 5-7



ภาพประกอบ 5-5 กราฟแสดงความสัมพันธ์ระหว่างพื้นที่ที่ใช้ในการสร้างดัชนีบิตแมปและเวลาที่ใช้ในการสอบถามของดัชนีบิตแมปทั้ง 5 ชนิด เมื่อคาร์ดินอลิตี้มีค่าเท่ากับ 25



ภาพประกอบ 5-6 กราฟแสดงความสัมพันธ์ระหว่างพื้นที่ที่ใช้ในการสร้างดัชนีบิตแมปและเวลาที่ใช้ในการสอบถามของดัชนีบิตแมปทั้ง 5 ชนิด เมื่อคาร์ดินอลิตี้มีค่าเท่ากับ 50



ภาพประกอบ 5-7 กราฟแสดงความสัมพันธ์ระหว่างพื้นที่ที่ใช้ในการสร้างดัชนีบิตแมปและเวลาที่ใช้ในการสอบถามแบบค่าเท่ากันของดัชนีบิตแมปทั้ง 5 ชนิด เมื่อคาร์ดินอลิตี้มีค่าเท่ากับ 150

จากภาพประกอบ 5-5, 5-6 และ 5-7 เมื่อพิจารณาพื้นที่ใต้กราฟของดัชนีบิตแมปแบบพื้นฐาน จะเห็นว่ามีลักษณะเป็นรูปสี่เหลี่ยมที่มีความสูงมาก เนื่องจากใช้พื้นที่ในการจัดเก็บดัชนีมาก ส่งผลให้พื้นที่ใต้กราฟมาก ส่วนพื้นที่ใต้กราฟของดัชนีบิตแมปแบบเข้ารหัสมีลักษณะเป็นรูปสี่เหลี่ยมที่มีฐานกว้างมาก เนื่องจากใช้เวลาในการสอบถามแบบค่าเท่ากันมาก จึงทำให้พื้นที่ใต้กราฟมากด้วยเช่นกัน สำหรับพื้นที่ใต้กราฟของดัชนีบิตแมปอีก 3 ชนิด คือ ดัชนีบิตแมปแบบช่วง ดัชนีบิตแมปแบบกระจาย และดัชนีบิตแมปแบบคู่กัน มีลักษณะเป็นรูปสี่เหลี่ยมที่มีฐานใกล้เคียงกันและไม่สามารถเรียงลำดับความยาวฐานของพื้นที่ใต้กราฟของดัชนีบิตแมปทั้ง 3 ชนิด เนื่องจากดัชนีบิตแมปทั้ง 3 ชนิดใช้เวลาในการสอบถามแบบค่าเท่ากันใกล้เคียงกันและไม่สามารถเรียงลำดับได้ แต่เมื่อพิจารณาความสูงของพื้นที่ใต้กราฟของดัชนีบิตแมปทั้ง 3 ชนิด เราจะพบว่า ความสูงของพื้นที่ใต้กราฟของดัชนีบิตแมปแบบช่วงสูงที่สุด รองลงมาคือพื้นที่ใต้กราฟของดัชนีบิตแมปแบบกระจาย และพื้นที่ใต้กราฟของดัชนีบิตแมปแบบคู่กัน ตามลำดับ จึงได้ว่าพื้นที่ใต้กราฟของดัชนีบิตแมปทั้ง 3 ชนิด เรียงจากมากไปหาน้อยคือ พื้นที่ใต้กราฟของดัชนีบิตแมปแบบช่วง พื้นที่ใต้กราฟของดัชนีบิตแมปแบบกระจาย และ พื้นที่ใต้กราฟของดัชนีบิตแมปแบบคู่กัน จะเห็นได้ว่า เมื่อเปรียบเทียบพื้นที่ใต้กราฟดัชนีบิตแมปทั้ง 5 ชนิด แล้วพื้นที่ใต้กราฟ

ดัชนีบิตแมปแบบคู่กันน้อยที่สุด นั่นหมายความว่าดัชนีบิตแมปแบบคู่กันมีประสิทธิภาพในแง่ Space-Time Trade-off มากที่สุด