

บทที่ 4

ใช้โปรแกรมทดสอบเพื่อหาขนาดของรหัสที่เหมาะสม

เพื่อทดสอบว่าการจับกลุ่มต่าง ๆ ที่กล่าวไว้ในบทก่อนจะให้ผลการใช้งานเป็นอย่างไร และการใช้ขนาดรหัสเท่าไรจึงจะเหมาะสม จำเป็นต้องสร้างโปรแกรมขึ้นเพื่อทดสอบ ตัวอย่างโปรแกรมที่ใช้ทดสอบแสดงไว้ในภาคผนวก สำหรับการเลือกวิธีจับกลุ่มแบบใดและความยาวรหัสควรเป็นเท่าไร จะพิจารณาตามหลักเกณฑ์และวัตถุประสงค์ดังนี้ คือ

- 1) เพื่อหาขนาดของรหัสที่ทำให้ได้ข้อมูลออกมา 1 ข้อถูกต้องมากที่สุด
- 2) เพื่อหาขนาดของรหัสที่ทำให้ได้โอกาสค้นหาข้อมูลไม่พบน้อย
- 3) เพื่อหาขนาดของรหัสที่ทำให้ข้อมูลที่สะกดต่างกันแต่ใกล้เคียงกันมีน้อย
นั่นคือ มีความซ้ำซ้อนของรหัสน้อย
- 4) เพื่อหาขนาดของรหัสที่จะประหยัดค่าใช้จ่ายไปมาก

ในการพิจารณาความเหมาะสมของขนาดความยาวของรหัสตามหลักเกณฑ์ดังกล่าว เราไม่สามารถที่จะพิจารณาค่าหนึ่งค่าใดโดยเฉพาะ เนื่องจากถ้าใช้ขนาดรหัสสั้นโอกาสค้นหาข้อมูลไม่พบจะมีน้อย หรือแทบไม่มีเลย และโอกาสที่ได้ข้อมูลออกมา 1 ข้อถูกต้องก็จะน้อยด้วย แต่จะมีรหัสซ้ำซ้อนกันมาก ถึงแม้ว่าจะสิ้นเปลืองค่าใช้จ่ายน้อย ดังนั้นในการพิจารณาขนาดของรหัสที่เหมาะสมนั้นจะพิจารณาขนาดของรหัสที่ให้รรถประโยชน์สูงสุดตามหลักเกณฑ์หรือวัตถุประสงค์ดังกล่าวข้างต้น

4.1 การทดสอบและผลจากการทดสอบ

การทดสอบทั้งการจึกกลุ่มที่ 1 และการจึกกลุ่มที่ 2 ทำโดยใช้ชื่อ 220 ชื่อ เข้าค้นหาข้อมูลในแฟ้มข้อมูลซึ่งมีจำนวน 2,925 ชื่อ ขนาดความยาวของรหัสที่ใช้ทดสอบ ใช้ตั้งแต่ 3 ถึง 10 ในการทดสอบทั้งสองกลุ่มจะใช้รูปแบบในการพิจารณา ดังนี้

- 1) โอกาสค้นหาข้อมูลไม่พบ
- 2) โอกาสค้นหาข้อมูลพบ 1 ชื่อและถูกต้อง
- 3) โอกาสค้นหาข้อมูลและถูกต้องแต่มีจำนวนชื่อซ้ำกันออกมามากกว่า 1 ชื่อ
- 4) จำนวนชื่อซ้ำที่ออกมามากที่สุด และจำนวนชื่อซ้ำโดยเฉลี่ยที่ออกมา

4.1.1 การทดสอบโดยใช้รหัสกลุ่มที่ 1

ขนาดของรหัส 3

- 1) โอกาสที่ค้นหาข้อมูลไม่พบประมาณ 0.45% หรือจากการค้นหาข้อมูล 220 ชื่อ มีข้อมูลที่ค้นหาไม่พบ 1 ชื่อ
- 2) โอกาสที่ค้นหาข้อมูลได้ 1 ชื่อและถูกต้อง 9.5% หรือจากการค้นหาข้อมูล 220 ชื่อ มีข้อมูลที่ออกมา 1 ชื่อและถูกต้อง 21 ชื่อ
- 3) โอกาสค้นหาข้อมูลพบและถูกต้องแต่ออกมามากกว่า 1 ชื่อ 90% หรือจากการค้นหาข้อมูล 220 ชื่อ มีจำนวนชื่อที่ออกมามากกว่า 1 ชื่อ 198 ชื่อ
- 4) จำนวนชื่อที่มากที่สุดที่ออกมา 58 ชื่อ (รหัสซ้ำซ้อนมากที่สุด)
จำนวนชื่อโดยเฉลี่ยที่ออกมา 11.627 ชื่อหรือประมาณ 12 ชื่อ

ขนาดของรหัส 4

- 1) โอกาสที่ค้นหาข้อมูลไม่พบประมาณ 3.64% หรือจากการค้นหาข้อมูล 220 ชื่อ มีข้อมูลที่ค้นหาไม่พบ 8 ชื่อ
- 2) โอกาสที่ค้นหาข้อมูลได้ 1 ชื่อและถูกต้อง 37.73% หรือจากการค้นหาข้อมูล 220 ชื่อ มีข้อมูลที่ออกมา 1 ชื่อและถูกต้อง 83 ชื่อ

- 3) โอกาสค้นหาข้อมูลพบและถูกต้องแต่ออกมามากกว่า 1 ข้อ 58.636%
หรือจากการค้นหาข้อมูล 220 ข้อมีจำนวนที่ออกมามากกว่า 1 ข้อ 129 ข้อ
- 4) จำนวนข้อมากที่สุดที่ออกมา 42 ข้อ
จำนวนข้อโดยเฉลี่ยที่ออกมา 4.509 ข้อหรือประมาณ 5 ข้อ

ขนาดของรหัส 5

- 1) โอกาสที่ค้นหาข้อมูลไม่พบประมาณ 8.18% หรือจากการค้นหาข้อมูล 220 ข้อมีข้อมูลที่ค้นหาไม่พบ 18 ข้อ
- 2) โอกาสที่ค้นหาข้อมูลได้ 1 ข้อและถูกต้อง 62.3% หรือจากการค้นหาข้อมูล 220 ข้อมีข้อมูลที่ออกมา 1 ข้อและถูกต้อง 137 ข้อ
- 3) โอกาสค้นหาข้อมูลพบและถูกต้องแต่ออกมามากกว่า 1 ข้อ 29.545%
หรือจากการค้นหาข้อมูล 220 ข้อมีจำนวนที่ออกมามากกว่า 1 ข้อ 65 ข้อ
- 4) จำนวนข้อมากที่สุดที่ออกมา 11 ข้อ
จำนวนข้อโดยเฉลี่ยที่ออกมา 1.682 ข้อหรือประมาณ 2 ข้อ

ขนาดของรหัส 6

- 1) โอกาสที่ค้นหาข้อมูลไม่พบประมาณ 10.45% หรือจากการค้นหาข้อมูล 220 ข้อมีข้อมูลที่ค้นหาไม่พบ 23 ข้อ
- 2) โอกาสที่ค้นหาข้อมูลได้ 1 ข้อและถูกต้อง 80.7% หรือจากการค้นหาข้อมูล 220 ข้อมีข้อมูลที่ออกมา 1 ข้อและถูกต้อง 178 ข้อ
- 3) โอกาสที่ค้นหาข้อมูลพบและถูกต้องแต่ออกมามากกว่า 1 ข้อ 8.636%
หรือจากการค้นหาข้อมูล 220 ข้อมีจำนวนข้อที่ออกมามากกว่า 1 ข้อ 19 ข้อ
- 4) จำนวนข้อมากที่สุดที่ออกมา 5 ข้อ
จำนวนข้อโดยเฉลี่ยที่ออกมา 1.041 ข้อหรือประมาณ 1 ข้อ

ขนาดของรหัส 7

- 1) โอกาสที่ค้นหาข้อมูลไม่พบประมาณ 13.18% หรือจากการค้นหาข้อมูล
220 ซึ่งมีข้อมูลที่ค้นหาไม่พบ 29 ข้อ
- 2) โอกาสที่ค้นหาข้อมูลได้ 1 ข้อและถูกต้อง 85.5% หรือจากการ
ค้นหาข้อมูล 220 ซึ่งมีข้อมูลที่ออกมา 1 ข้อและถูกต้อง 188 ข้อ
- 3) โอกาสค้นหาข้อมูลพบและถูกต้องแต่ออกมามากกว่า 1 ข้อ 1.363%
หรือจากการค้นหาข้อมูล 220 ซึ่งมีจำนวนข้อที่ออกมามากกว่า 1 ข้อ 3 ข้อ
- 4) จำนวนข้อที่มากที่สุดที่ออกมา 5 ข้อ
จำนวนข้อโดยเฉลี่ยที่ออกมา 0.895 ข้อหรือประมาณ 1 ข้อ

ขนาดของรหัส 8

- 1) โอกาสที่ค้นหาข้อมูลไม่พบประมาณ 19.55% หรือจากการค้นหาข้อมูล
220 ซึ่งมีข้อมูลที่ค้นหาไม่พบ 43 ข้อ
- 2) โอกาสที่ค้นหาข้อมูลได้ 1 ข้อและถูกต้อง 80% หรือจากการค้นหา
ข้อมูล 220 ซึ่งมีข้อมูลที่ออกมา 1 ข้อและถูกต้อง 176 ข้อ
- 3) โอกาสค้นหาข้อมูลพบและถูกต้องแต่ออกมามากกว่า 1 ข้อ 0.455%
หรือจากการค้นหาข้อมูล 220 ซึ่งมีจำนวนข้อที่ออกมามากกว่า 1 ข้อเพียง 1 ข้อเท่านั้น
- 4) จำนวนข้อที่มากที่สุดที่ออกมา 2 ข้อ
จำนวนข้อโดยเฉลี่ยที่ออกมา 0.809 ข้อหรือประมาณ 1 ข้อ

ขนาดของรหัส 9

- 1) โอกาสที่ค้นหาข้อมูลไม่พบประมาณ 21.36% หรือจากการค้นหาข้อมูล
220 ซึ่งมีข้อมูลที่ค้นหาไม่พบ 47 ข้อ
- 2) โอกาสที่ค้นหาข้อมูลได้ 1 ข้อและถูกต้อง 78.6% หรือจากการ
ค้นหาข้อมูล 220 ซึ่งมีข้อมูลที่ออกมา 1 ข้อและถูกต้อง 173 ข้อ

- 3) โอกาสที่ค้นหาข้อมูลพบและถูกต้อง แต่ออกมามากกว่า 1 ชื่อไม่มีเลย
- 4) จำนวนชื่อโดยเฉลี่ยที่ออกมา 0.786 ชื่อหรือประมาณ 1 ชื่อ

ขนาดของรหัส 10

- 1) โอกาสที่ค้นหาข้อมูลไม่พบประมาณ 22.72% หรือจากการค้นหาข้อมูล 220 ชื่อมีข้อมูลที่ค้นหาไม่พบ 50 ชื่อ
- 2) โอกาสที่ค้นหาข้อมูลได้ 1 ชื่อและถูกต้อง 77.3% หรือจากการค้นหาข้อมูล 220 ชื่อมีข้อมูลที่ออกมา 1 ชื่อและถูกต้อง 170 ชื่อ
- 3) โอกาสที่ค้นหาข้อมูลพบและถูกต้อง แต่ออกมามากกว่า 1 ชื่อไม่มีเลย
- 4) จำนวนชื่อโดยเฉลี่ยที่ออกมา 0.773 ชื่อหรือประมาณ 1 ชื่อ

จะเห็นได้ว่ายิ่งขนาดของรหัสเพิ่มมากขึ้นเท่าใด จำนวนชื่อที่ค้นหาไม่พบก็จะมียากขึ้นด้วย แต่การที่จะได้ข้อมูลออกมามากกว่า 1 ชื่อมีน้อยลง

จากการพิจารณา ชื่อต่าง ๆ ที่ค้นหาไม่พบ เกิดจากเหตุผลต่อไปนี้

- 1) การใช้พยัญชนะต้นคำต่างกัน คือ

ชิตพงษ์ ตระกูลบุญ	ในแฟ้มข้อมูลเป็น	ชิตพงษ์ ตระกูลบุญ
พรศ เหล่าบุรินทร์	ในแฟ้มข้อมูลเป็น	กรศ เหล่าบุรินทร์
- 2) การใช้พยัญชนะควบกล้ำ และไม่ไ้ไ้ใช้ คือ

ทินกร ลีทันกร	ในแฟ้มข้อมูลเป็น	ทินกร ศรีทันกร
สมิตรา ยมะสมิต	ในแฟ้มข้อมูลเป็น	สมิตา ยมะสมิต
สมบัติ จันทรสุรวงศ์	ในแฟ้มข้อมูลเป็น	สมบัติ จันสุรวงษ์
กาลี่ ตรียศุต	ในแฟ้มข้อมูลเป็น	การี่ ตรียศุต
ศิริพจน์ วรรณาสมิต	ในแฟ้มข้อมูลเป็น	ศิริพจน์ วรรณารดสมิต
พัชรินทร์ รินทรกุล	ในแฟ้มข้อมูลเป็น	พัชรินทร์ รินทรกุล
วิทยา กายนันท์	ในแฟ้มข้อมูลเป็น	วิทยา กายนันท์

3) การที่เครื่องไม่สามารถแยกได้ว่า พยัญชนะตัวใดเป็นพยัญชนะต้นพยางค์ หรือ พยัญชนะสะกด คือ

ปรียาวที บุญญวิบูลย์ ในแฟ้มข้อมูลเป็น ปรียาวที บุญญวิบูลย์
4) การใช้ ใ และ ัย คือ

ชราไส โพธิ์ทศ ในแฟ้มข้อมูลเป็น ชราศัย โพธิ์ทศ

5) กรณีที่ใช้ รร และไม้ไ้ใช้ คือ

ลคาวรัลย์ บุญจิรังสรรค ในแฟ้มข้อมูลเป็น ลคาวรรณ บุญจิรังสรรค

นิสา เพชรวิวิธน์ ในแฟ้มข้อมูลเป็น นิสา เพชรวิวรรธน์

จิราภรณ์ กীরคิสรพ ในแฟ้มข้อมูลเป็น จิราภรณ์ กীরคิศัพท์

4.1.2 การทดสอบโดยวิธีกลุ่มที่ 2

ขนาดของรหัส 3

1) โอกาสที่ค้นหาข้อมูลไม่พบประมาณ 0.45% หรือจากการค้นหาข้อมูล
220 ชื่อมีข้อมูลที่ค้นหาไม่พบ 1 ชื่อ

2) โอกาสที่ค้นหาข้อมูลได้ 1 ชื่อและถูกต้อง 5% หรือจากการค้นหา
ข้อมูล 220 ชื่อมีข้อมูลที่ออกมา 1 ชื่อและถูกต้อง 11 ชื่อ

3) โอกาสที่ค้นหาข้อมูลและถูกต้อง แต่ออกนามากกว่า 1 ชื่อ
94.545% หรือจากการค้นหาข้อมูล 220 ชื่อมีจำนวนชื่อที่ออกนามากกว่า 1 ชื่อ
208 ชื่อ

4) จำนวนชื่อที่มากที่สุดที่ออกมา 58 ชื่อ

จำนวนชื่อโดยเฉลี่ยที่ออกมา 12.641 ชื่อหรือประมาณ 13 ชื่อ

ขนาดของรหัส 4

- 1) โอกาสที่ค้นหาข้อมูลไม่พบประมาณ 2.7% หรือจากการค้นหาข้อมูล
220 ชื่อมีข้อมูลที่ค้นหาไม่พบ 6 ชื่อ
- 2) โอกาสที่ค้นหาข้อมูลได้ 1 ชื่อและถูกต้อง 37.3% หรือจากการ
ค้นหาข้อมูล 220 ชื่อมีจำนวนชื่อที่ออกมา 1 ชื่อและถูกต้อง 82 ชื่อ
- 3) โอกาสค้นหาข้อมูลพบและถูกต้อง แต่ออกมามากกว่า 1 ชื่อ 60%
หรือจากการค้นหาข้อมูล 220 ชื่อมีจำนวนชื่อที่ออกมามากกว่า 1 ชื่อ 132 ชื่อ
- 4) จำนวนชื่อที่มากที่สุดที่ออกมา 42 ชื่อ
จำนวนชื่อโดยเฉลี่ยที่ออกมา 4.827 ชื่อหรือประมาณ 5 ชื่อ

ขนาดของรหัส 5

- 1) โอกาสที่ค้นหาข้อมูลไม่พบประมาณ 6.8% หรือจากการค้นหาข้อมูล
220 ชื่อมีข้อมูลที่ค้นหาไม่พบ 15 ชื่อ
- 2) โอกาสที่ค้นหาข้อมูลได้ 1 ชื่อและถูกต้อง 63.6% หรือจากการ
ค้นหาข้อมูล 220 ชื่อมีข้อมูลที่ออกมา 1 ชื่อและถูกต้อง 140 ชื่อ
- 3) โอกาสที่ค้นหาข้อมูลพบและถูกต้อง แต่ออกมามากกว่า 1 ชื่อ
29.545% หรือจากการค้นหาข้อมูล 220 ชื่อมีจำนวนชื่อที่ออกมามากกว่า 1 ชื่อ
65 ชื่อ
- 4) จำนวนชื่อที่มากที่สุดที่ออกมา 11 ชื่อ
จำนวนชื่อโดยเฉลี่ยที่ออกมา 1.695 ชื่อหรือประมาณ 2 ชื่อ

ขนาดของรหัส 6

- 1) โอกาสที่ค้นหาข้อมูลไม่พบประมาณ 9.55% หรือจากการค้นหาข้อมูล
220 ชื่อมีข้อมูลที่ค้นหาไม่พบ 21 ชื่อ

- 2) โอกาสที่ค้นหาข้อมูลได้ 1 ข้อและถูกต้อง 81.8% หรือจากการค้นหา
ข้อมูล 220 ข้อมีข้อมูลที่ออกมา 1 ข้อและถูกต้อง 180 ข้อ
- 3) โอกาสที่ค้นหาข้อมูลพบและถูกต้อง แต่ออกมามากกว่า 1 ข้อ
8.636% หรือจากการค้นหาข้อมูล 220 ข้อมีจำนวนข้อที่ออกมามากกว่า 1 ข้อ 19 ข้อ
- 4) จำนวนข้อมากที่สุดที่ออกมา 5 ข้อ
จำนวนข้อโดยเฉลี่ยที่ออกมา 1.05 ข้อหรือประมาณ 1 ข้อ

ขนาดของรหัส 7

- 1) โอกาสที่ค้นหาข้อมูลไม่พบประมาณ 12.27% หรือจากการค้นหาข้อมูล
220 ข้อมีข้อมูลที่ค้นหาไม่พบ 27 ข้อ
- 2) โอกาสที่ค้นหาข้อมูลได้ 1 ข้อและถูกต้อง 86.4% หรือจากการ
ค้นหาข้อมูล 220 ข้อมีข้อมูลที่ออกมา 1 ข้อและถูกต้อง 190 ข้อ
- 3) โอกาสค้นหาข้อมูลพบและถูกต้อง แต่ออกมามากกว่า 1 ข้อ 1.363%
หรือจากการค้นหาข้อมูล 220 ข้อมีจำนวนข้อที่ออกมามากกว่า 1 ข้อ 3 ข้อ
- 4) จำนวนข้อที่มากที่สุดที่ออกมา 5 ข้อ
จำนวนข้อโดยเฉลี่ยที่ออกมา 0.905 ข้อหรือประมาณ 1 ข้อ

ขนาดของรหัส 8

- 1) โอกาสที่ค้นหาข้อมูลไม่พบประมาณ 18.64% หรือจากการค้นหาข้อมูล
220 ข้อมีข้อมูลที่ค้นหาไม่พบ 41 ข้อ
- 2) โอกาสที่ค้นหาข้อมูลได้ 1 ข้อและถูกต้อง 80.9% หรือจากการค้นหา
ข้อมูล 220 ข้อมีข้อมูลที่ออกมา 1 ข้อและถูกต้อง 178 ข้อ
- 3) โอกาสค้นหาข้อมูลพบและถูกต้อง แต่ออกมามากกว่า 1 ข้อ 0.455%
หรือจากการค้นหาข้อมูล 220 ข้อมีข้อมูลที่ออกมามากกว่า 1 ข้อเพียง 1 ข้อเท่านั้น
- 4) จำนวนข้อมากที่สุดที่ออกมา 2 ข้อ
จำนวนข้อโดยเฉลี่ยที่ออกมา 0.818 ข้อหรือประมาณ 1 ข้อ

ขนาดของรหัส 9

- 220 1) โอกาสที่ค้นหาข้อมูลไม่พบประมาณ 20.45% หรือจากการค้นหาข้อมูล
 220 ที่มีข้อมูลที่ค้นหาไม่พบ 45 ชื่อ
- 2) โอกาสที่ค้นหาข้อมูลได้ 1 ชื่อและถูกต้อง 79.5% หรือจากการค้นหา
 ข้อมูล 220 ที่มีข้อมูลที่ออกมา 1 ชื่อและถูกต้อง 175 ชื่อ
- 3) โอกาสที่ค้นหาข้อมูลพบและถูกต้อง แต่ออกมามากกว่า 1 ชื่อไม่มีเลย
- 4) จำนวนชื่อโดยเฉลี่ยที่ออกมา 0.795 ชื่อหรือประมาณ 1 ชื่อ

ขนาดของรหัส 10

- 220 1) โอกาสที่ค้นหาข้อมูลไม่พบประมาณ 21.82% หรือจากการค้นหาข้อมูล
 220 ที่มีข้อมูลที่ค้นหาไม่พบ 48 ชื่อ
- 2) โอกาสที่ค้นหาข้อมูลได้ 1 ชื่อและถูกต้อง 78.2% หรือจากการค้นหา
 ข้อมูล 220 ที่มีข้อมูลที่ออกมา 1 ชื่อและถูกต้อง 172 ชื่อ
- 3) โอกาสที่ค้นหาข้อมูลพบและถูกต้อง แต่ออกมามากกว่า 1 ชื่อไม่มีเลย
- 4) จำนวนชื่อโดยเฉลี่ยที่ออกมา 0.782 ชื่อหรือประมาณ 1 ชื่อ

เนื่องจากว่าในการจัดกลุ่มที่ 2 นี้เรานำพยานะต้นคำมาจัดกลุ่มด้วย ดังนั้นจึงสามารถแก้ปัญหของการจัดกลุ่มที่ 1 ในกรณีที่พยานะต้นคำต่างกันแล้วทำให้การค้นหา
 นั้นไม่พบไปได้ สำหรับปัญหาอื่นที่พบนั้นก็ยังสามารถแก้ไขได้เช่นเดียวกันดังที่ได้แนะนำวิธีแก้ไข
 ไว้แล้วในบทก่อน

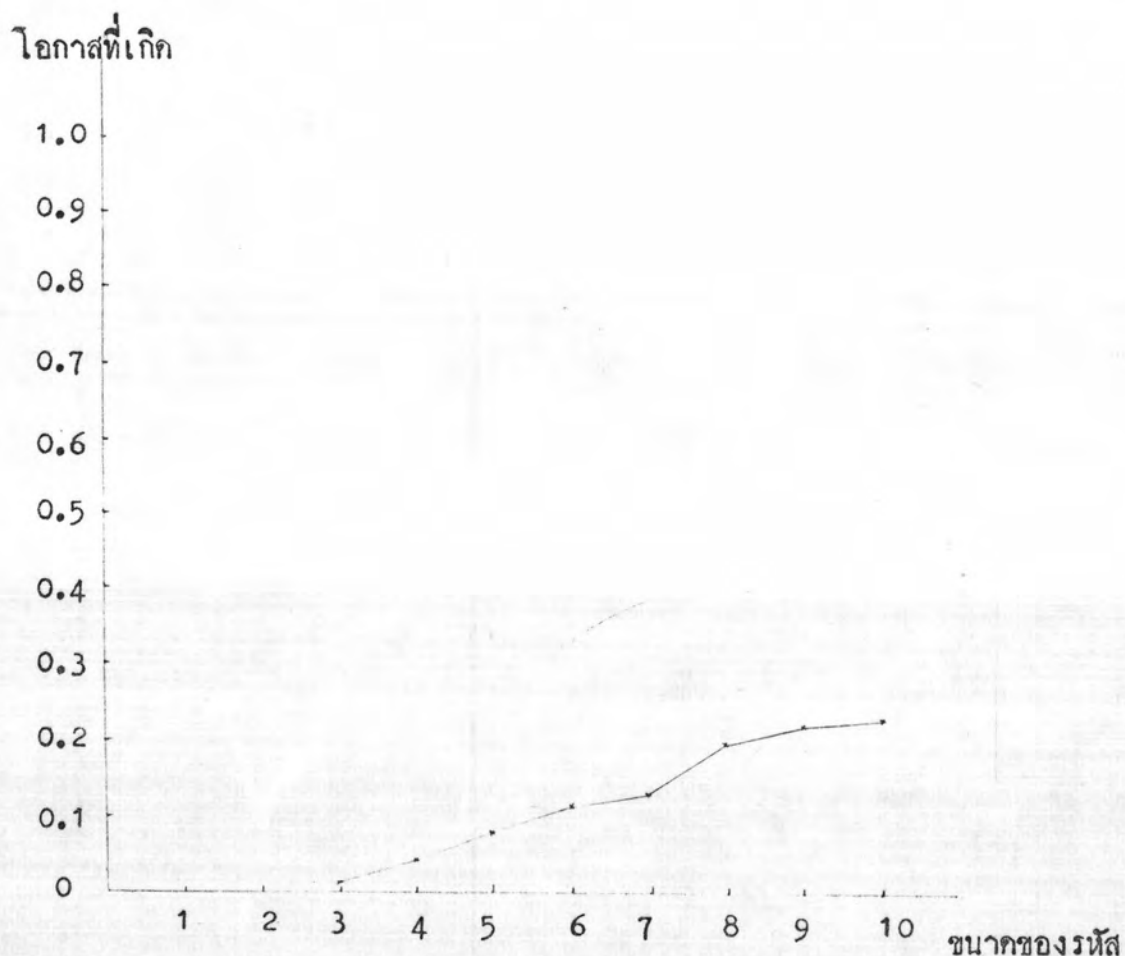
4.2 พิจารณาขนาดของรหัสที่เหมาะสม

วิธีการพิจารณาขนาดของรหัสที่เหมาะสม พิจารณาจากโอกาสของข้อมูลที่ค้นหาไม่พบ โอกาสของการค้นหาข้อมูลได้ 1 ชื่อและถูกต้อง ร่วมกับโอกาสของข้อมูลที่ไคร้รหัสซ้ำกัน

จากตัวอย่างการทดสอบไคร้การค้นหาชื่อ 220 ชื่อจากแฟ้มข้อมูลที่มีชื่อทั้งหมด 2,925 ชื่อ มีโอกาสที่ค้นพบข้อมูล โอกาสที่ค้นหาข้อมูลได้ 1 ชื่อและถูกต้อง และโอกาสของข้อมูลที่ไคร้รหัสซ้ำกันออกมา เมื่อค้นหาข้อมูลโดยใช้ขนาดของรหัสต่าง ๆ กัน โดยแสดงเป็นตารางและกราฟ เพื่อเปรียบเทียบหาขนาดของรหัสที่เหมาะสม ทั้งของการจัดกลุ่มที่ 1 และการจัดกลุ่มที่ 2 ตามตารางที่ 4.1 และ ตารางที่ 4.2 ตามลำดับ

ขนาดของรหัส	โอกาสที่ค้นหาข้อมูลไม่พบ	โอกาสที่ค้นหาข้อมูลได้ชื่อ 1 ชื่อถูกต้อง	โอกาสที่ชื่อซ้ำกันไคร้รหัสซ้ำกัน
3	0.0045	0.0955	0.9000
4	0.0364	0.3773	0.5864
5	0.0818	0.6227	0.2955
6	0.1045	0.8091	0.0864
7	0.1318	0.8545	0.0136
8	0.1955	0.8000	0.0045
9	0.2136	0.7864	0.0000
10	0.2273	0.7727	0.0000

ตารางที่ 4.1 โอกาสที่ค้นหาข้อมูลไม่พบ โอกาสที่ค้นหาข้อมูลได้ชื่อ 1 ชื่อและถูกต้อง และโอกาสที่ชื่อซ้ำกันไคร้รหัสซ้ำกันจากการค้นหาข้อมูลโดยอาศัยวิธีการจัดกลุ่มที่ 1



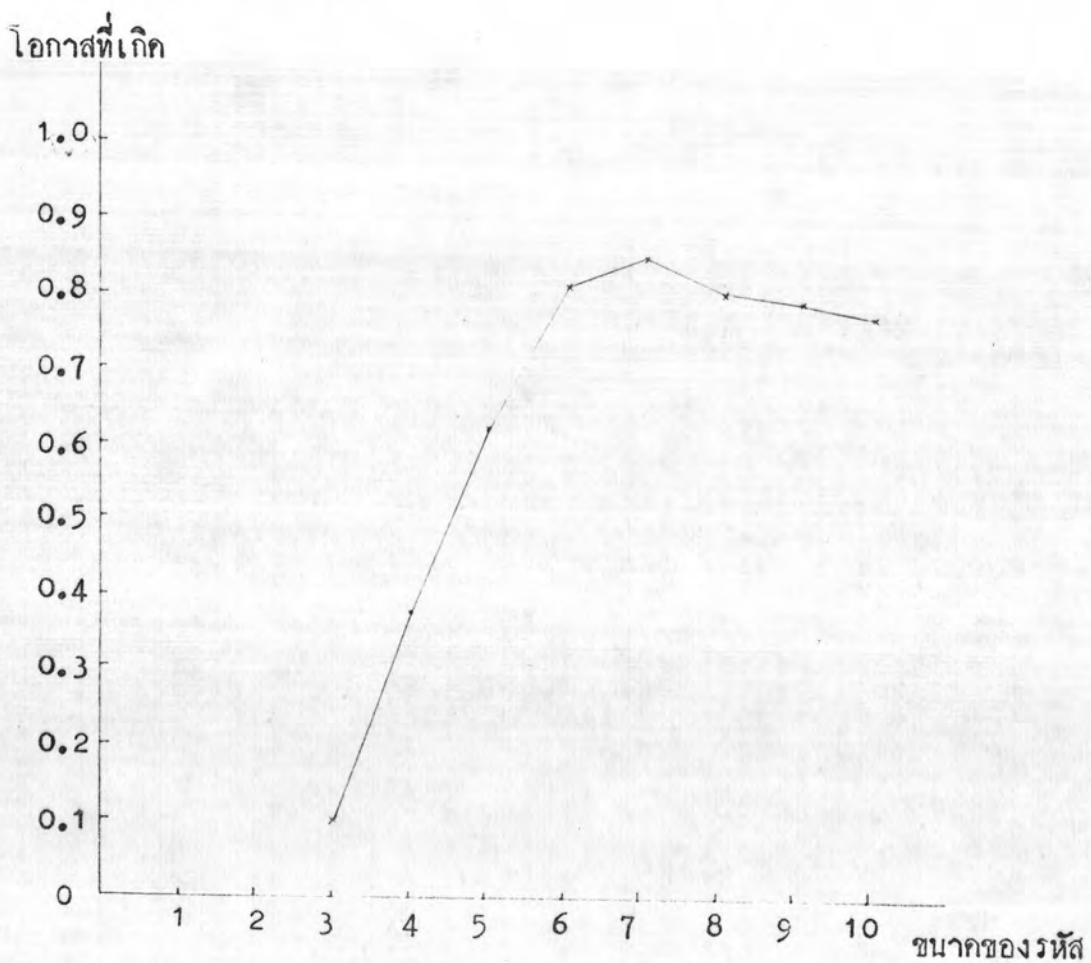
รูปที่ 4.1 ก โอกาสที่ค้นหาข้อมูลไม่พบของการจิกกลุ่มที่ 1

จากรูปที่ 4.1 ก ทำการเปรียบเทียบโอกาสที่ค้นหาข้อมูลไม่พบที่ขนาดของรหัสเป็น 8 9 และ 10 จะเห็นว่าเมื่อขนาดของรหัสเป็น 9 และ 10 จะต้องเสียค่าใช้จ่ายมากกว่าที่ขนาดของรหัสเป็น 8 และมีโอกาสที่ค้นหาข้อมูลไม่พบมากกว่าเมื่อขนาดของรหัสเป็น 8 ดังนั้นจะศึกษาขนาดของรหัสเป็น 9 และ 10 ออกไปไม่น่ามาคิด

เปรียบเทียบที่ขนาดของรหัสเป็น 7 และ 8 จะเห็นว่าถ้าขนาดของรหัสเป็น 8 โอกาสที่ค้นหาข้อมูลไม่พบจะมากกว่าเมื่อใช้ขนาดของรหัสเป็น 7 เกือบ 7%

และถ้าใช้ขนาดของรหัสเป็น 8 จะเสียค่าใช้จ่ายมากกว่าเมื่อใช้ขนาดของรหัสเป็น 7
ดังนั้นจะศึกษาขนาดของรหัสเป็น 8 ออกไปไม่น่ามาคิด

พิจารณาที่ขนาดของรหัสเป็น 3 4 5 6 และ 7 จะเห็นว่า
สัดส่วนของการค้นไม่พบข้อมูลจะเป็นสัดส่วนที่ค่อนข้างจะแน่นอนซึ่งขึ้นกับขนาดของรหัส ถ้า
ขนาดของรหัสยาวขึ้นก็จะมีโอกาสที่ค้นหาข้อมูลไม่พบมากขึ้น ดังนั้นในการพิจารณาว่า จะใช้
ขนาดของรหัสเท่าใดในขนาด 3 4 5 6 และ 7 นั้นก็ต้องอาศัยส่วนประกอบ
อย่างอื่นมาช่วยในการพิจารณาด้วย คือโอกาสที่ค้นพบข้อมูลได้ 1 ชื่อและถูกต้อง ที่ขนาด
ความยาวของรหัสต่าง ๆ กัน



รูปที่ 4.1 ข โอกาสที่ค้นหาข้อมูลได้ 1 ชื่อและถูกต้องของการจัดกลุ่มที่ 1



จากรูปที่ 4.1 ข ทำการเปรียบเทียบโอกาสที่ไอซ์ออกมา 1 ชื่อและ ถูกต้องที่ขนาดของรหัสเป็น 7 8 9 และ 10 จะเห็นว่าถ้าขนาดของรหัส เป็น 8 9 และ 10 นั้นมีโอกาสที่ไอซ์ออกมาถูกต้องน้อยกว่าเมื่อใช้ขนาดของรหัส เป็น 7 แต่จะต้องเสียค่าใช้จ่ายมากกว่าเมื่อใช้ขนาดของรหัสเป็น 7 ดังนั้นจึงคัดเลือกขนาดของรหัสเป็น 8 9 และ 10 ออกไปไม่นำมาคิด

พิจารณาเปรียบเทียบที่ขนาดของรหัสเป็น 3 4 5 6 และ 7 จะเห็นว่าถึงแม้ว่าเมื่อขนาดของรหัสเป็น 3 4 และ 5 จะเสียค่าใช้จ่ายน้อยกว่า เมื่อใช้ขนาดของรหัสเป็น 6 และ 7 ก็ตามแต่โอกาสที่ค้นพบข้อมูลออกมาถูกต้องนั้น น้อยกว่าเมื่อใช้ขนาดของรหัสเป็น 6 และ 7 มากดังนั้นจึงคัดเลือกขนาดของรหัสเป็น 3 4 และ 5 ออกไปไม่นำมาคิด

พิจารณาเปรียบเทียบที่ขนาดของรหัสเป็น 6 และ 7 ต้องพิจารณาร่วมกันจากรูปที่ 4.1 ก จะเห็นว่าเมื่อใช้ขนาดของรหัสเป็น 6 มีโอกาสที่ค้นหาข้อมูล ไม่พบน้อยกว่าเมื่อใช้ขนาดของรหัสเป็น 7 เพียง 0.0273 แต่จากรูปที่ 4.1 ข นั้น เมื่อใช้ขนาดของรหัสเป็น 7 มีโอกาสที่ไอซ์ออกมา 1 ชื่อและถูกต้องมากกว่าเมื่อใช้ขนาดของรหัสเป็น 6 ถึง 0.454 ดังนั้นเมื่อพิจารณาาร่วมกันทั้งจากรูปที่ 4.1 ก และรูปที่ 4.1 ข จะเห็นว่าถึงแม้ขนาดของรหัสเป็น 6 จะเสียค่าใช้จ่ายและมีโอกาส ค้นไม่พบข้อมูลน้อยกว่า เมื่อใช้ขนาดของรหัสเป็น 7 ก็ตาม แต่โอกาสที่ไอซ์ออกมา 1 ชื่อและถูกต้องเมื่อใช้ขนาดของรหัสเป็น 7 จะมีมากกว่า

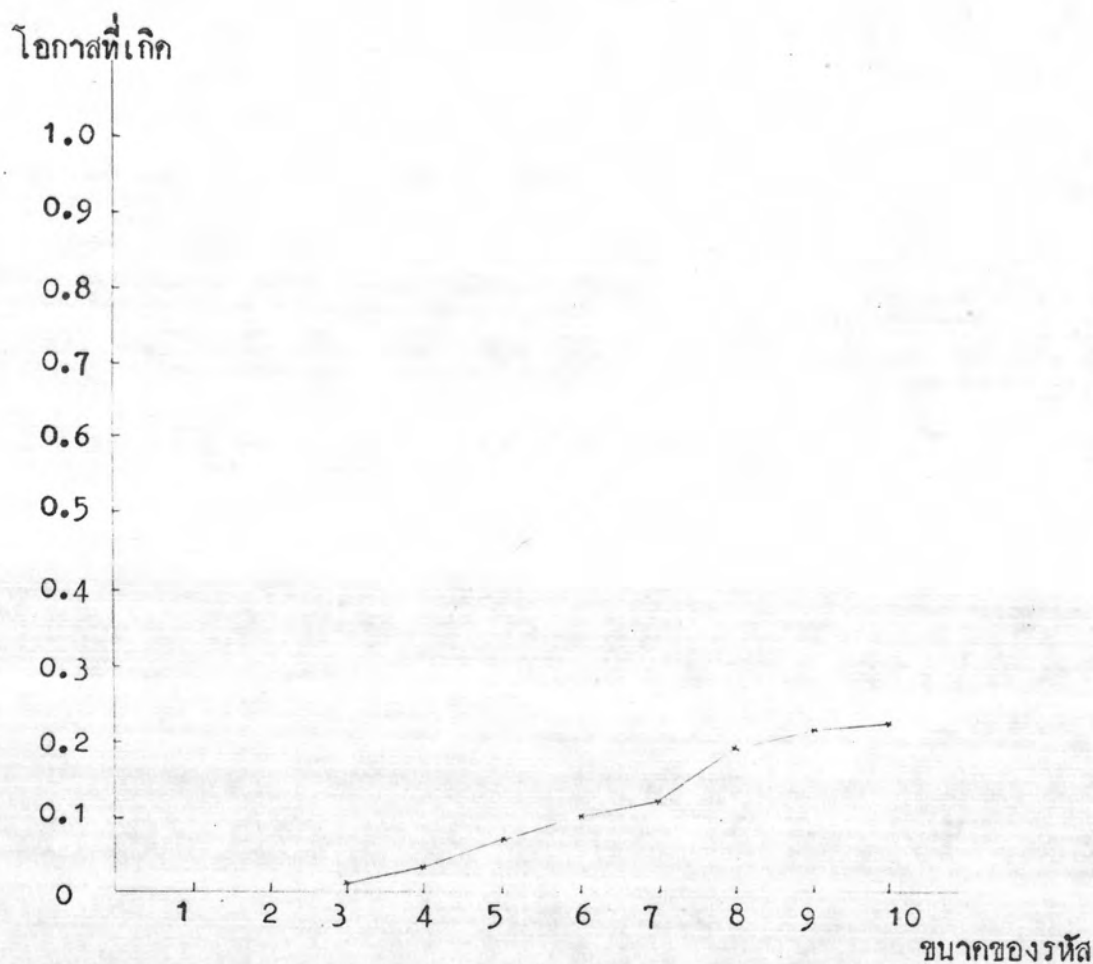
ดังนั้น จากการพิจารณาร่วมกันจากรูปที่ 4.1 ก และ รูปที่ 4.1 ข นั้นจะเลือกใช้ขนาดของรหัสเป็น 7 เมื่อใช้วิธีการจัดกลุ่มที่ 1

ขนาดของรหัส	โอกาสที่ค้นหาข้อมูลไม่พบ	โอกาสที่ค้นหาข้อมูลไค้ชื่อ 1 ชื่อ ถูกต้อง	โอกาสที่ชื่อต่างกันไค้รหัสซ้ำกัน
3	0.0045	0.0500	0.9455
4	0.0273	0.3727	0.6000
5	0.0682	0.6364	0.2955
6	0.0955	0.8182	0.0864
7	0.1227	0.8636	0.0136
8	0.1864	0.8091	0.0045
9	0.2045	0.7955	0.0000
10	0.2182	0.7818	0.0000

ตารางที่ 4.2 โอกาสที่ค้นหาข้อมูลไม่พบ โอกาสที่ค้นหาข้อมูลไค้ชื่อ 1 ชื่อ และถูกต้อง และโอกาสที่ชื่อต่างกันไค้รหัสซ้ำกันจากกรรค้นหาข้อมูลโดยอาศัยวิธีการจิกกลุ่มที่ 2

จากรูปที่ 4.2 ก ทำการเปรียบเทียบโอกาสที่ค้นหาข้อมูลไม่พบที่ขนาดของรหัส เป็น 8 9 และ 10 แล้วจะเห็นว่าเมื่อขนาดของรหัสเป็น 9 และ 10 จะต้องเสียค่าใช้จ่ายมากกว่าที่ขนาดของรหัสเป็น 8 และมีโอกาสที่ค้นหาข้อมูลไม่พบมากกว่า เมื่อขนาดของรหัสเป็น 8 กังนั้นจะคักขนาดของรหัสเป็น 9 และ 10 ออกไป ไม่น่ามาคัก

เปรียบเทียบที่ขนาดของรหัสเป็น 7 และ 8 จะเห็นว่าถ้าขนาดของรหัส เป็น 8 โอกาสที่ค้นหาข้อมูลไม่พบจะมากกว่าเมื่อใช้ขนาดของรหัสเป็น 7 เกือบ 7%

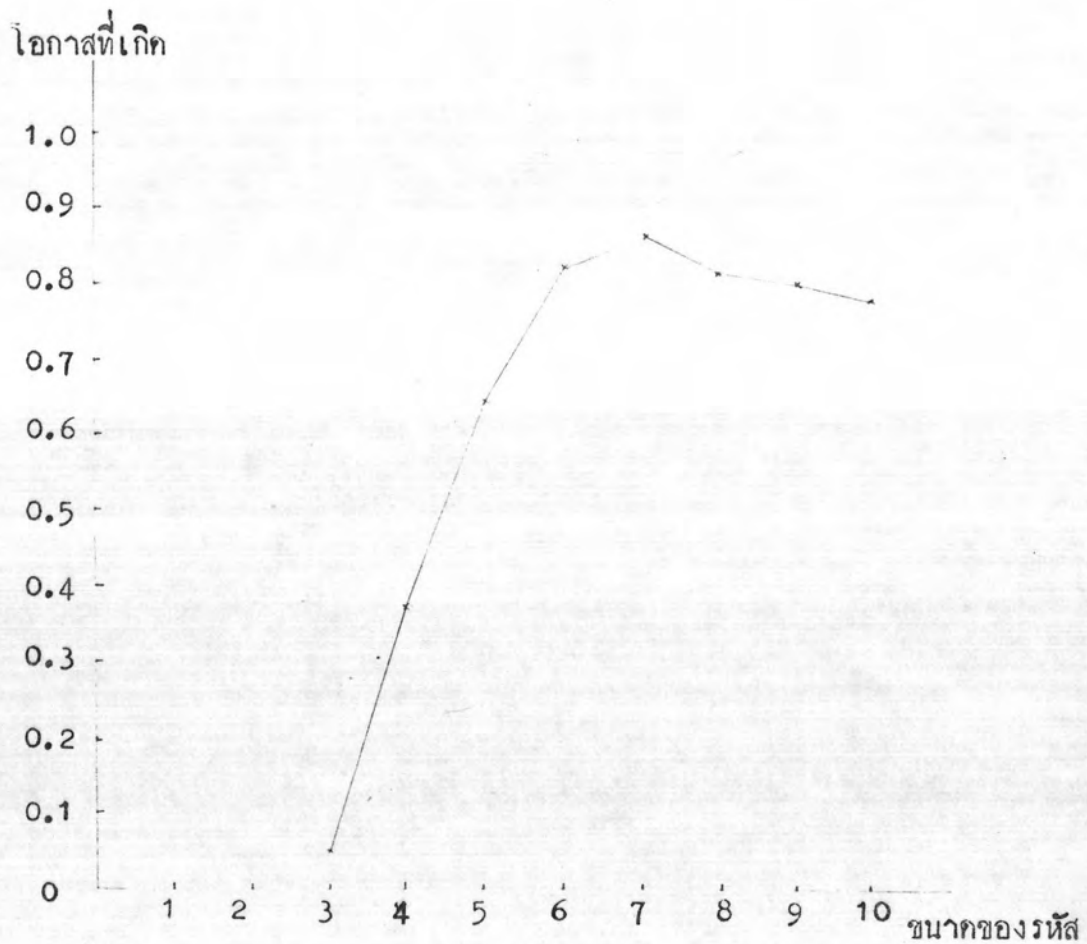


รูปที่ 4.2 ก โอกาสที่ค้นหาข้อมูลไม่พบของการจัดกลุ่มที่ 2

และถ้าใช้ขนาดของรหัสเป็น 8 จะเสียค่าใช้จ่ายมากกว่าขนาดของรหัสเป็น 7 ดังนั้นจะคัดขนาดของรหัสเป็น 8 ออกไปไม่น่ามาคิด

พิจารณาที่ขนาดของรหัสเป็น 3 4 5 6 และ 7 จะเห็นว่าสัดส่วนของการค้นหาข้อมูลไม่พบจะเป็นสัดส่วนที่ค่อนข้างจะแน่นอนซึ่งขึ้นกับขนาดของรหัส ถ้าขนาดของรหัสยาวขึ้นก็จะมีโอกาสที่ค้นหาข้อมูลไม่พบมากขึ้น ดังนั้นในการพิจารณาว่าจะใช้ขนาดของรหัสเท่าใดใน 3 4 5 6 และ 7 นั้นก็ต้องอาศัยส่วนประกอบอย่างอื่นมาช่วยในการ

พิจารณาด้วยคือ โอกาสที่ค้นหาข้อมูลได้ 1 ชื่อและถูกต้อง ที่ขนาดความยาวของรหัสต่าง ๆ กัน



รูปที่ 4.2 ข โอกาสที่ค้นหาข้อมูลได้ 1 ชื่อและถูกต้องของการจัดกลุ่มที่ 2

จากรูปที่ 4.2 ข ทำการเปรียบเทียบโอกาสที่ได้ออกมา 1 ชื่อและถูกต้อง ที่ขนาดของรหัสเป็น 7 8 9 และ 10 จะเห็นว่าถ้าขนาดของรหัสเป็น 8 9 และ 10 นั้นมีโอกาสที่ได้ออกมา 1 ชื่อและถูกต้องน้อยกว่าเมื่อใช้ขนาดของรหัสเป็น 7 แต่จะต้องเสียค่าใช้จ่ายมากกว่าเมื่อใช้ขนาดของรหัสเป็น 7 ทั้งนี้

จึงศึกษาขนาดของรหัสเป็น 8 9 และ 10 ออกไปไม่นำมาคิด

พิจารณาเปรียบเทียบที่ขนาดของรหัสเป็น 3 4 5 6 และ 7 จะเห็นว่าถึงแม้ว่าเมื่อขนาดของรหัสเป็น 3 4 และ 5 จะเสียค่าใช้จ่ายน้อยกว่าเมื่อใช้ขนาดของรหัสเป็น 6 และ 7 ก็ตามแต่โอกาสที่ค้นหาข้อมูลได้ 1 ชื่อและถูกต้องนั้นน้อยกว่าเมื่อใช้ขนาดของรหัสเป็น 6 และ 7 มากถึงนั้นจึงศึกษาขนาดของรหัสเป็น 3 4 และ 5 ออกไปไม่นำมาคิด

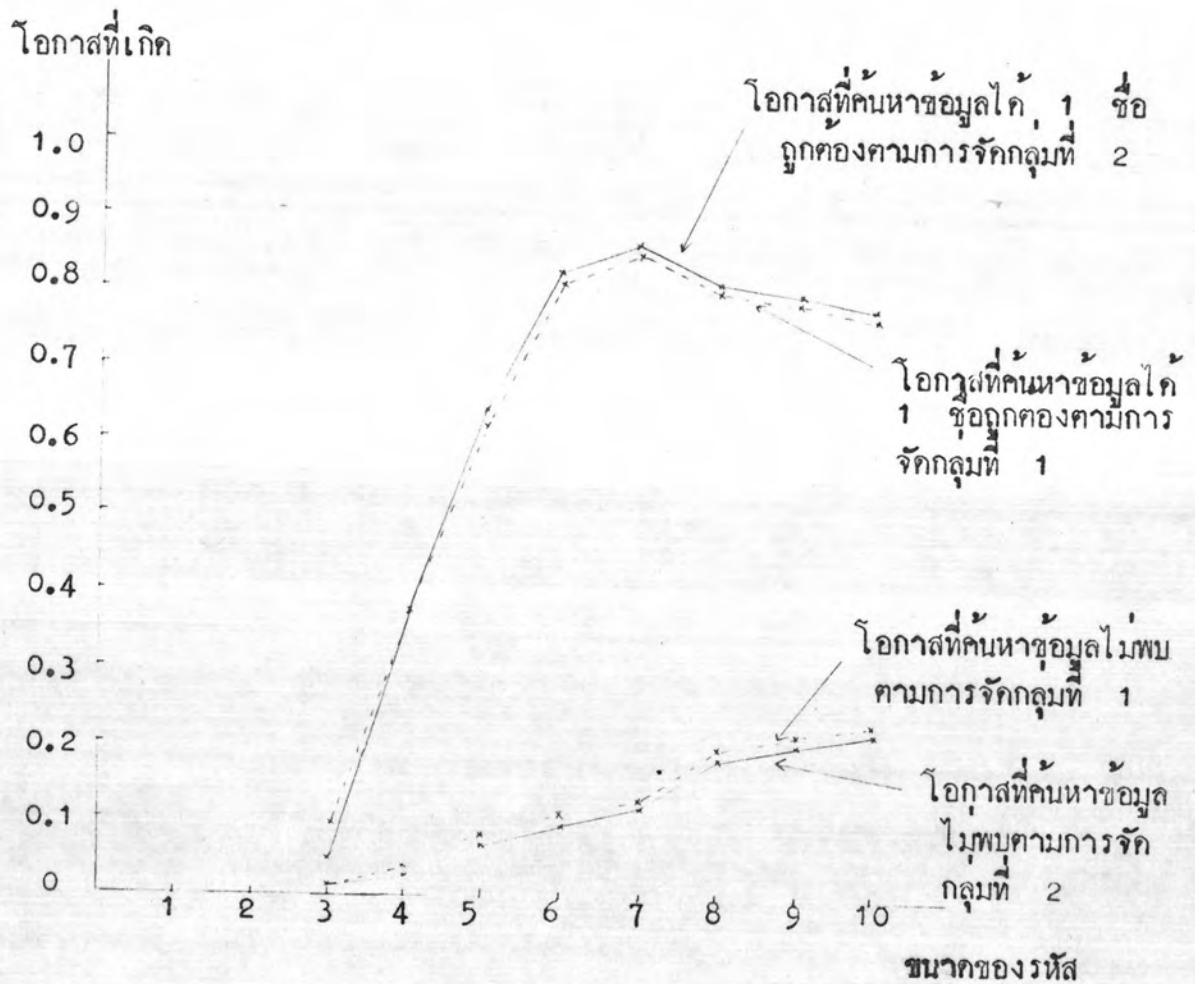
พิจารณาเปรียบเทียบที่ขนาดของรหัสเป็น 6 และ 7 ต้องพิจารณาร่วมกัน จากรูปที่ 4.2 ก จะเห็นว่าเมื่อใช้ขนาดของรหัสเป็น 6 มีโอกาสค้นหาข้อมูลไม่พบ น้อยกว่าเมื่อใช้ขนาดของรหัสเป็น 7 เพียง 0.0273 แต่จากรูปที่ 4.2 ข นั้น เมื่อใช้ขนาดของรหัสเป็น 7 มีโอกาสที่ค้นหาข้อมูลได้ชื่อออกมา 1 ชื่อและถูกต้องมากกว่าเมื่อใช้ขนาดของรหัสเป็น 6 ถึง 0.0454 ทั้งนี้เมื่อพิจารณาร่วมกันทั้งจากรูปที่ 4.2 ก และรูปที่ 4.2 ข จะเห็นว่าถึงแม้ขนาดของรหัสเป็น 6 จะเสียค่าใช้จ่าย และมีโอกาสค้นหาข้อมูลไม่พบน้อยกว่าเมื่อใช้ขนาดของรหัสเป็น 7 ก็ตามแต่โอกาสที่ค้นหาข้อมูลได้ชื่อออกมา 1 ชื่อและถูกต้องเมื่อใช้ขนาดของรหัสเป็น 7 จะมีมากกว่า

ดังนั้นจากการพิจารณาจากรูปที่ 4.2 ก และรูปที่ 4.2 ข นั้นจะเลือกใช้ขนาดของรหัสเป็น 7 เมื่อใช้วิธีการจัดกลุ่มที่ 2

จากการพิจารณาขนาดของรหัสตามที่กล่าวมา ไม่ว่าจะเป็นการจัดกลุ่มที่ 1 หรือการจัดกลุ่มที่ 2 ก็ตามจะใช้ขนาดของรหัสที่เหมาะสมคือ ขนาดของรหัสเป็น 7 ทั้ง 2 กรณี ทั้งนี้ จะเห็นได้ว่าตามการวิจัยนี้ขนาดของรหัสที่เหมาะสมจากการทดสอบโดยใช้ ตัวอย่างชื่อ 220 ชื่อมาค้นหาข้อมูลจากแฟ้มข้อมูลที่มี 2,925 ชื่อนั้นจะใช้ขนาดของรหัสที่เหมาะสมและใช้โค้ดที่สุก คือขนาดของรหัสเป็น 7

เมื่อใช้ขนาดของรหัสที่เหมาะสมแล้วจะทำการเปรียบเทียบการจัดกลุ่มทั้ง 2 วิธี เพื่อจะทราบว่า การจัดกลุ่มวิธีใดมีประสิทธิภาพได้ดีกว่า ซึ่งกราฟของการเปรียบเทียบนี้ แสดงไว้ในรูปที่ 4.3

4.3 เปรียบเทียบการจักกลุ่มที่ 1 และการจักกลุ่มที่ 2



รูปที่ 4.3 แสดงการเปรียบเทียบการจักกลุ่มที่ 1 และ การจักกลุ่มที่ 2

จากรูปที่ 4.3 ซึ่งเป็นการเปรียบเทียบวิธีการจักกลุ่มที่ 1 และวิธีการจักกลุ่มที่ 2 ว่าการจักกลุ่มแบบใดจะมีประสิทธิภาพดีกว่ากัน จะเห็นว่าโอกาสที่ค้นหาข้อมูลไม่พบของการจักกลุ่มที่ 2 จะน้อยกว่าโอกาสที่ค้นหาข้อมูลไม่พบของการจักกลุ่มที่ 1 เสมอไม่ว่าจะใช้ขนาดของรหัสเป็นเท่าไรก็ตาม

แต่เมื่อพิจารณาโอกาสที่ไอช้อมูลออกมา 1 ซึ่งถูกต้องจะเห็นว่าการจัดกลุ่มที่ 2 มีโอกาสที่ไอช้อมูลออกมา 1 ชื่อและถูกต้อง ที่ขนาดของรหัสเป็น 5 6 7 8 9 และ 10 มากกว่าการจัดกลุ่มที่ 1 เนื่องจากว่าจากการพิจารณาขนาดของรหัสที่เหมาะสมเห็นว่าควรจะใช้ขนาดของรหัสเป็น 7 ไม่ว่าจะเป็นการจัดกลุ่มที่ 1 หรือการจัดกลุ่มที่ 2 ก็ตาม

ดังนั้น มาพิจารณาเปรียบเทียบวิธีการจัดกลุ่มทั้ง 2 วิธีที่ขนาดของรหัสเป็น 7 จะเห็นว่าที่ขนาดของรหัสเป็น 7 ในการจัดกลุ่มที่ 1 มีโอกาสค้นหาข้อมูลไม่พบ 0.1318 ซึ่งมากกว่าในการจัดกลุ่มที่ 2 ซึ่งมีโอกาสค้นหาข้อมูลไม่พบ 0.1227 และที่ขนาดของรหัสเป็น 7 ในการจัดกลุ่มที่ 1 มีโอกาสที่ไอช้อมูลออกมา 1 ชื่อและถูกต้อง 0.8545 ซึ่งน้อยกว่าในการจัดกลุ่มที่ 2 ซึ่งมีโอกาสที่ไอช้อมูลออกมา 1 ชื่อและถูกต้อง 0.8636 ดังนั้นเมื่อเปรียบเทียบการจัดกลุ่มทั้ง 2 วิธีที่ขนาดของรหัสที่เหมาะสมแล้วจะเห็นว่าการจัดกลุ่มที่ 2 จะให้ประสิทธิภาพดีกว่าการจัดกลุ่มที่ 1 นั่นคือการจัดกลุ่มที่ 2 และเมื่อขนาดของรหัสเป็น 7 จะให้ประสิทธิภาพที่ดีที่สุด