

สารบัญ	หน้า
กิตติกรรมประกาศ	ค
บทคัดย่อภาษาไทย	ง
บทคัดย่อภาษาอังกฤษ	ฉ
สารบัญตาราง	ญ
สารบัญภาพ	ฎ
บทที่ 1 บทนำ	1
1.1 ปัญหาและที่มาของการศึกษา	1
1.2 แนวทางการแก้ปัญหา	12
1.3 สรุปสาระสำคัญจากเอกสารที่เกี่ยวข้อง	12
1.4 วัตถุประสงค์ของการศึกษา	14
1.5 ประโยชน์ที่ได้รับจากการศึกษา เจริญทฤษฎี และเชิงประยุกต์	14
1.6 ขอบเขตการทำวิจัย	14
1.7 วิธีการทำวิจัย	15
1.8 เครื่องมือในการพัฒนา	15
บทที่ 2 ทฤษฎีที่เกี่ยวข้อง	16
2.1 เทคนิค k -Anonymity	16
2.2 เทคนิค (α, k) -Anonymity	22
2.3 ขั้นตอนวิธี MCCRT	22
2.4 เทคนิค IncSpan	25
บทที่ 3 ปัญหาและวิธีการแก้ปัญหา	33
3.1 นิยามเบื้องต้น	34
3.2 ขั้นตอนวิธี MCCRT (Minimum Classification Correction Rate Transformation algorithm)	36
3.3 นิยามของปัญหา	38

3.4	ผลกระทบของข้อมูลจากกระบวนการการเพิ่มเข้ามา ของข้อมูลระเบียบใหม่กับขั้นตอนวิธี MCCRT	39
3.5	ขั้นตอนวิธีแปลงข้อมูลแบบเพิ่มขึ้น (Incremental Algorithm)	57
บทที่ 4	การทดลองและผลการทดลอง	68
4.1	การเตรียมการทดลอง	68
4.1.1	การทดลองเปรียบเทียบความเร็วในการทำงานของขั้นตอนวิธี MCCRT กับขั้นตอนวิธีการแปลงค่าข้อมูลแบบเพิ่มขึ้นเมื่อค่า k ถูกเปลี่ยนแปลง	70
4.1.2	การทดลองเปรียบเทียบความเร็วในการทำงานของขั้นตอนวิธี MCCRT กับขั้นตอนวิธีการแปลงค่าข้อมูลแบบเพิ่มขึ้นเมื่อ จำนวนคอลัมน์ที่เป็นตัวบ่งชี้ข้อมูลทางอ้อมมีการเปลี่ยนแปลง	71
4.1.3	การทดลองเปรียบเทียบความเร็วในการทำงานของขั้นตอนวิธี MCCRT กับขั้นตอนวิธีการแปลงค่าข้อมูลแบบเพิ่มขึ้นเมื่อ ขนาดของข้อมูลที่เพิ่มขึ้นมีการเปลี่ยนแปลง	71
4.2	ผลการทดลอง	72
4.2.1	ผลการทดลองเปรียบเทียบความเร็วในการทำงานของขั้นตอนวิธี MCCR กับขั้นตอนวิธีการแปลงค่าข้อมูลแบบเพิ่มขึ้นเมื่อค่า k ถูก เปลี่ยนแปลง	72
4.2.2	ผลการทดลองเปรียบเทียบความเร็วในการทำงานของขั้นตอนวิธี MCCRT กับขั้นตอนวิธีการแปลงค่าข้อมูลแบบเพิ่มขึ้นเมื่อค่า จำนวนคอลัมน์ที่เป็นตัวบ่งชี้บุคคลทางอ้อม ($ \Delta Q $) ในข้อมูลนำเข้า เข้าถูกเปลี่ยนแปลง	74
4.2.3	ผลการทดลองเปรียบเทียบความเร็วในการทำงานของขั้นตอนวิธี MCCRT กับขั้นตอนวิธีการแปลงค่าข้อมูลแบบเพิ่มขึ้นเมื่อขนาด ของข้อมูลที่เพิ่มขึ้น $ \Delta D $ มีการเปลี่ยนแปลง	80
4.3	สรุปผลการทดลอง	84
บทที่ 5	สรุปผลการวิจัย	85
5.1	สรุปสาระสำคัญของงานวิจัย	85

5.2	ส่วนสำคัญของงานวิจัย	86
5.3	วิเคราะห์งานวิจัย	86
5.4	ข้อเสนอแนะและงานวิจัยในอนาคต	86
เอกสารอ้างอิง		87
ประวัติผู้เขียน		89



ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
 Copyright© by Chiang Mai University
 All rights reserved

สารบัญตาราง

ตาราง	หน้า
1.1 โครงสร้างข้อมูลของแฟ้มที่ 2	2
1.2 ตัวอย่างการลบข้อมูลที่เป็นตัวบ่งชี้บุคคล	2
1.3 ตัวอย่างข้อมูลจากประกันสังคม	3
1.4 ตัวอย่างการให้รหัส (ยังไม่ใช่เป็นทางการ)	4
1.5 โครงสร้างข้อมูลแฟ้มที่ 11 มาตรฐานแฟ้มข้อมูลการเงิน	5
2.1 ตัวอย่างข้อมูลการวินิจฉัยโรค	17
2.2 ตารางข้อมูลหลังการแปลงค่าโดยระดับการเงินเนอรัลไลเซชัน GL เท่ากับ $\langle (BirthDate, 1), (Sex, 0), (Weight, 0), (Height, 0), (Career, 0) \rangle$	20
2.3 ตารางข้อมูลหลังการเปลี่ยนแปลงค่าซึ่งมีคุณสมบัติ 2-Anonymity โดย GL เท่ากับ $\langle (BirthDate, 1), (Sex, 1), (Weight, 1), (Height, 1), (Career, 0) \rangle$	20
2.4 ตารางข้อมูลหลังการเปลี่ยนแปลงค่าซึ่งมีคุณสมบัติ 2-Anonymity โดย GL เท่ากับ $\langle (BirthDate, 1), (Sex, 0), (Weight, 1), (Height, 1), (Career, 1) \rangle$	21
2.5 ตัวอย่างข้อมูลเพื่อใช้ในการทำ MCCRT	25
2.6 ข้อมูลที่มีคุณสมบัติ 2-Anonymity โดย GL เท่ากับ $\langle (Career, 1), (Sex, 0) \rangle$	25
2.7 ตัวอย่างการเติมข้อมูล	29
3.1 ตัวอย่างข้อมูล D	63
3.2 ตารางข้อมูลหลังการแปลงข้อมูล D' ซึ่งมีคุณสมบัติ 2-Anonymity โดย GL เท่ากับ $\langle (BirthDate, 2), (Sex, 1), (Weight, 1), (Height, 1), (Career, 0) \rangle$	64
3.3 ตารางข้อมูลเพิ่มขึ้นจำนวน 2 ระเบียบ	64
3.4 ตัวอย่างข้อมูลเดิมรวมกับข้อมูลเพิ่มขึ้น	64
3.5 ตารางข้อมูลเพิ่มขึ้นหลังการแปลงข้อมูล $\Delta D'$ โดยระดับการเงินเนอรัลไลเซชันเดิม GL เท่ากับ $\langle (BirthDate, 2), (Sex, 1), (Weight, 1), (Height, 0), (Career, 0) \rangle$	65
3.6 ข้อมูลรวมหลังการแปลง $(D' + \Delta D)'$	65
3.7 ข้อมูลรวมหลังการแปลงตามระดับการเงินเนอรัลไลเซชันใหม่ GL' เท่ากับ $\langle (BirthDate, 2), (Weight, 1), (Sex, 0), (Height, 0), (Career, 0) \rangle$	66
3.8 ข้อมูลรวมหลังการแปลงตามระดับการเงินเนอรัลไลเซชันใหม่ GL' เท่ากับ $\langle (BirthDate, 2), (Weight, 2), (Sex, 0), (Height, 0), (Career, 0) \rangle$	67

รูป	หน้า
1.1 ตัวอย่างการระบุตัวตนอีกครั้งเพื่อเชื่อมโยงข้อมูลจากการวิเคราะห์ความสัมพันธ์	6
1.2 ตัวอย่างการทำให้ข้อมูลมีคุณสมบัติ k -Anonymity	8
1.3 ผลกระทบของค่า k ต่อค่า CCR [5]	9
1.4 ผลกระทบของจำนวนคอลัมน์ที่เชื่อมช้อนกันต่อเวลาการทำงาน [6]	10
1.5 ผลกระทบค่า k ต่อค่า CCR [6]	10
2.1 ขั้นตอนการแปลงค่าข้อมูลตามลำดับชั้นของคอลัมน์ BirthDate	17
2.2 ขั้นตอนการแปลงค่าข้อมูลตามลำดับชั้นของคอลัมน์ Sex	18
2.3 ขั้นตอนการแปลงค่าข้อมูลตามลำดับชั้นของคอลัมน์ Weight	18
2.4 ขั้นตอนการแปลงค่าข้อมูลตามลำดับชั้นของคอลัมน์ Height	19
2.5 ขั้นตอนการแปลงค่าข้อมูลตามลำดับชั้นของคอลัมน์ Career	19
2.6 ขั้นตอนการหารูปแบบลำดับจากข้อมูลอนุกรมเวลา	28
2.7 ขั้นตอนการสร้างโครงสร้างรูปต้นไม้	29
2.8 การเปลี่ยนโครงสร้างข้อมูลรูปต้นไม้ต้นไม้นี้ในกรณี (2)	30
2.9 การเปลี่ยนโครงสร้างข้อมูลรูปต้นไม้ในกรณี (4)	31
2.10 การโปรเจกชันโดย b	31
3.1 รหัสเทียมขั้นตอนวิธี MCCRT	38
3.2 ตัวอย่างการเพิ่มเข้ามาของข้อมูลที่เป็นกรณีค่าข้อมูลซ้ำเดิมหรือมีคุณสมบัติ k -Anonymity	42
3.3 ตัวอย่างการเพิ่มเข้ามาของข้อมูลที่เป็นกรณีค่าข้อมูลเป็นค่าใหม่และมีคุณสมบัติ k -Anonymity	44
3.4 ตัวอย่างการเพิ่มเข้ามาของข้อมูลที่เป็นกรณีค่าข้อมูลเป็นค่าใหม่และไม่มีคุณสมบัติ k -Anonymity	46
3.5 ตัวอย่างการสลับลำดับก่อน CGP	49
3.6 ตัวอย่างการสลับลำดับหลัง CGP	51
3.7 รหัสเทียมขั้นตอนวิธีการแปลงข้อมูลแบบเพิ่มขึ้น	60
4.1 ผลกระทบค่า k ต่อเวลาการทำงานของข้อมูล CRX	72
4.2 ผลกระทบค่า k ต่อเวลาการทำงานของข้อมูล Breast	73
4.3 ผลกระทบของค่า $ \Delta Q $ ต่อเวลาการทำงานของข้อมูล CRX โดยค่า $k=2$	74

4.4	ผลกระทบของค่า $ \Delta Q $ ต่อเวลาการทำงาน ของข้อมูล CRX โดยค่า $k=5$	75
4.5	ผลกระทบของค่า $ \Delta Q $ ต่อเวลาการทำงาน ของข้อมูล CRX โดยค่า $k=10$	76
4.6	ผลกระทบของค่า $ \Delta Q $ ต่อเวลาการทำงาน ของข้อมูล Breast โดยค่า $k=2$	77
4.7	ผลกระทบของค่า $ \Delta Q $ ต่อเวลาการทำงาน ของข้อมูล Breast โดยค่า $k=5$	78
4.8	ผลกระทบของค่า $ \Delta Q $ ต่อเวลาการทำงาน ของข้อมูล Breast โดยค่า $k=10$	78
4.9	ผลกระทบค่า $ \Delta D $ ต่อเวลาในการทำงาน ของข้อมูล CRX ที่ $k=2$ และ $ \Delta Q =9$	80
4.10	ผลกระทบค่า $ \Delta D $ ต่อเวลาในการทำงาน ของข้อมูล CRX ที่ $k=5$ และ $ \Delta Q =9$	81
4.11	ผลกระทบค่า $ \Delta D $ ต่อเวลาในการทำงาน ของข้อมูล CRX ที่ $k=10$ และ $ \Delta Q =9$	81
4.12	ผลกระทบค่า $ \Delta D $ ต่อเวลาในการทำงาน ของข้อมูล Breast ที่ $k=2$ และ $ \Delta Q =9$	82
4.13	ผลกระทบค่า $ \Delta D $ ต่อเวลาในการทำงาน ของข้อมูล Breast ที่ $k=5$ และ $ \Delta Q =9$	83
4.14	ผลกระทบค่า $ \Delta D $ ต่อเวลาในการทำงาน ของข้อมูล Breast ที่ $k=10$ และ $ \Delta Q =9$	83