

ปัจจุบันข้อมูลถูกสร้างขึ้นและถูกเผยแพร่เป็นจำนวนมาก การเผยแพร่ข้อมูลส่วนบุคคลสำหรับการทำงานหรือข้อมูลจำเป็นต้องมีการรักษาความเป็นส่วนตัวของข้อมูลเนื่องจากข้อมูลที่ถูกระบุสามารถระบุตัวบุคคลอีกครั้งได้โดยการเชื่อมต่อกับข้อมูลที่ถูกเผยแพร่อื่นๆ ดังนั้นจึงจำเป็นต้องทำการแปลงข้อมูลที่เผยแพร่ให้มีการรักษาความเป็นส่วนตัว

ขั้นตอนวิธีที่ใช้รักษาความเป็นส่วนตัวของข้อมูลมีหลายวิธี ในงานวิจัยนี้ได้เลือกขั้นตอนวิธี MCCRT (Minimum Classification Correction Rate Transformation) มาใช้ในการพัฒนา เนื่องจากเป็นขั้นตอนวิธีที่มีประสิทธิภาพและประสิทธิผลสูงสำหรับการแปลงค่าข้อมูลเพื่อรักษาความเป็นส่วนตัวของข้อมูล ในกรณีที่ข้อมูลซึ่งถูกแปลงค่าแล้วจะถูกนำไปใช้ในการจำแนกแบบความสัมพันธ์ (Associative Classification)

อย่างไรก็ตามขั้นตอนวิธี MCCRT เป็นขั้นตอนวิธีที่อยู่ในลักษณะหนึ่งการประมวลผลต่อหนึ่งงาน (One-Time Fashion) กล่าวคือเป็นขั้นตอนวิธีที่ต้องเริ่มต้นกระบวนการใหม่ทั้งหมดอีกครั้งเมื่อมีข้อมูลนำเข้าหรือข้อมูลมีการเปลี่ยนแปลง ในงานวิจัยนี้ได้พัฒนาขั้นตอนวิธี MCCRT เป็นขั้นตอนวิธีการแปลงค่าข้อมูลแบบเพิ่มขึ้นซึ่งมีประสิทธิภาพมากขึ้นเมื่ออยู่ในสถานการณ์ที่ข้อมูลนำเข้าถูกเพิ่มขึ้น และข้อมูลที่เป็นผลลัพธ์ของขั้นตอนวิธีแปลงค่าข้อมูลแบบเพิ่มขึ้นนี้เหมือนกับข้อมูลที่เป็นผลลัพธ์ของขั้นตอนวิธี MCCRT

จากผลการทดลองพบว่ามีเพียงกรณีเดียวที่ขั้นตอนวิธีแปลงค่าข้อมูลแบบเพิ่มขึ้นมีประสิทธิภาพน้อยกว่าหรือใช้เวลามากกว่าขั้นตอนวิธีแบบ MCCRT คือเมื่อการกำหนดค่า k มีค่าน้อยสุดในการทดลองนี้คือ k เท่ากับ 2 เนื่องจากข้อมูลนำเข้าของกระบวนการมีคุณสมบัติตามที่กำหนดไว้ในเทคนิค k -Anonymity กล่าวคือเป็นข้อมูลนำเข้าไม่จำเป็นต้องผ่านกระบวนการก็เป็นข้อมูลผลลัพธ์อยู่แล้วนั่นเอง และในกรณีดังกล่าวขั้นตอนวิธี MCCRT เร็วกว่าขั้นตอนวิธีการแปลงค่าข้อมูลแบบเพิ่มขึ้นเพียง 10% เท่านั้น ส่วนการทดลองในกรณีอื่นๆนั้นผลการทดลองปรากฏว่าขั้นตอนวิธีแปลงค่าข้อมูลแบบเพิ่มขึ้นใช้เวลาเพียง 30% ของขั้นตอนวิธี MCCRT เท่านั้น ดังนั้นจากการทดลองสามารถสรุปได้ว่าขั้นตอนวิธีแปลงค่าข้อมูลแบบเพิ่มขึ้นมีประสิทธิภาพโดยรวมสูงกว่าขั้นตอนวิธี MCCRT

Present, huge data was created and published in the Internet. Publishing privacy data for data mining purpose need to preservation before release. Because the released data can be re-identified by being joined with other public databases. Thus, data transformation to ensure privacy preservation is needed.

Now, we have many algorithm and method for privacy preservation. In this research interest MCCRT (Minimum Classification Correction Rate Transformation) algorithm for develop in our research, because MCCRT has good efficient and good effectiveness of data privacy preserving. And its result is good for associative classification.

However MCCRT is One-Time Fashion algorithm. It must re-run when input data was changed. In this research we improve it to Incremental algorithm. The Incremental algorithm has more efficient than MCCRT when input data was increased. But it not changed the result and effectiveness from MCCRT.

From experiment we found only case that the Incremental algorithm is worst than MCCRT. The case is we set k be 2 for k -Anonymity. It is the less k possible of input data. The input data are satisfying k -Anonymity before run the process. However the MCCRT is fast than the Incremental algorithm only 10%. For another case the Incremental algorithm uses time only 30% of MCCRT used. So we can conclusion the Incremental algorithm has more efficient than MCCRT.