

## กระบวนการจัดการข้อมูลเพื่อใช้วิเคราะห์ผลด้วยวิธีกระบวนการเหมืองข้อมูล

### The process of data management for analyzing data process mining

เอนก นามขันธ์<sup>1</sup>, เสาวภา เมืองแก่น<sup>2</sup>

คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยสยาม กรุงเทพมหานคร<sup>1</sup>

สาขาวิชาคอมพิวเตอร์ธุรกิจ คณะบริหารธุรกิจ มหาวิทยาลัยธนบุรี กรุงเทพมหานคร<sup>2</sup>

Anake Nammakhunt<sup>1</sup>, Saowapha Muangkaen<sup>2</sup>

Faculty of Information Technology, Siam University, BKK<sup>1</sup>

Department of Business Computer, Faculty of Administration, Thonburi University, BKK<sup>2</sup>

E-mail: anake\_cc@thonburi-u.ac.th<sup>1</sup>, jaww\_muangkan@hotmail.com<sup>2</sup>

#### บทคัดย่อ

บทความฉบับนี้นำเสนอกระบวนการบริหารและจัดการข้อมูลโดยการใช้วิเคราะห์ผลด้วยวิธีกระบวนการเหมืองข้อมูลเพื่อวิเคราะห์พฤติกรรมของกระบวนการการทำงานในระบบสารสนเทศ ด้วยโปรแกรม DISCO หรือ ProM ซึ่งการส่งข้อมูลเข้าไปประมวลผลใน DISCO หรือ ProM ต้องจัดรูปแบบข้อมูลให้เหมาะสมในการประมวลผล ดังนั้นรูปแบบข้อมูลที่เหมาะสมต่อการประมวลผล ควรจัดให้อยู่ในรูปแบบ CSV และในปัจจุบันองค์กรส่วนใหญ่ได้นำระบบสารสนเทศเข้ามาช่วยสนับสนุนการดำเนินงานทางธุรกิจอย่างแพร่หลายองค์กรใดสามารถบริหารจัดการข้อมูลมหาศาล (Big data) ทางธุรกิจของตนเองได้ดี และสามารถวิเคราะห์จุดได้เปรียบ เสียเปรียบ หรือแนวทางปรับกระบวนการการทำงานได้ถึงระดับการทำงานเชิงพฤติกรรมโดยใช้ข้อมูลเชิงประวัติ (Event logs) จะทำให้องค์กรมีความได้เปรียบและสามารถแข่งขันกับคู่แข่งรวมทั้งก้าวทันเทคโนโลยีที่เปลี่ยนแปลงอย่างรวดเร็วหรือก้าวสู่ยุค Thailand 4.0 ดังนั้นการจัดการข้อมูลให้เหมาะสมต่อการวิเคราะห์ผลจึงเป็นจุดเริ่มต้นที่สำคัญ ผู้วิจัยจึงได้ทดลองเขียนชุดคำสั่ง SQL Command และ Visual Basic.Net พบว่าสามารถจัดเตรียมตามวิธีที่ผู้วิจัยทดลองได้ และนำไปใช้วิเคราะห์พบว่าคนหนึ่งคน ในหนึ่งวัน ใช้เวลาเดินโดยเฉลี่ยเท่าไร ใช้เวลารั้งโดยเฉลี่ยเท่าไร อยู่เฉย ๆ ใช้เวลาเท่าไร และยังสามารถทราบถึงพฤติกรรมของผู้ใช้มือถือด้วยว่าส่วนใหญ่มีกิจกรรมใดในแต่ละวัน และนอกจากนี้

**คำสำคัญ:** กระบวนการเหมืองข้อมูล, ระบบสารสนเทศ, การวิเคราะห์พฤติกรรมกระบวนการทำงาน

#### ABSTRACT

This article presents the process of managing and managing data by analyzing the results by method of the process mining process to analyze the behavior of the process in the information system. The program DISCO or ProM were used properly and formatted data for processing. There are the appropriate data format is easy to process. It should be used in CSV format and at present most organizations have adopted the information system to support business operations widely. In which organization can manage large data (Big Data) business very well of themselves. And it can analyze the advantages, disadvantages, or process adaptability to the level of behavioral, by using historical information (Event Logs) to make the organization more competitive and competitive with the pace of technology changes. Therefore, the management of data to the analysis of the results is an important starting point. The researchers tried to

write a set of SQL Command and Visual Basic.Net that could be prepared according to the method that the researchers did and to analyze how much one day a person walks on average. How long it to take run on average and also be aware of the behavior of mobile users that most often do activities on a daily basis. In addition, the researcher has experimented with the process of selling one commodity.

**Keywords:** Data process mining, Information system, Analysis of work process behavior

## บทนำ

ปัจจุบันความก้าวหน้าทางเทคโนโลยีและระบบสารสนเทศมีการพัฒนาและเติบโตอย่างรวดเร็วทำให้องค์กรธุรกิจหรือองค์กรต่าง ๆ มีการขยายตัวและเพิ่มขึ้นอย่างต่อเนื่อง จำเป็นต้องมีการจัดเก็บข้อมูลต่าง ๆ อย่างมากมายและอยู่ในรูปแบบที่หลากหลายและแตกต่างกัน ทำให้มีปริมาณข้อมูลจำนวนมากมายนับมหาศาล หรือจะเรียกเป็นบิ๊กดาต้า (Big data) และหากองค์กรใดให้ความสำคัญกับ Big data เหล่านี้ก็จะยิ่งรู้จักรักที่จะใช้หรือเลือกสรรข้อมูลเหล่านี้มาใช้ในการบริหารจัดการและตัดสินใจเพื่อให้ออกไปเกิดประโยชน์สูงสุดต่อองค์กร พร้อมทำนายสิ่งต่าง ๆ ที่จะเกิดขึ้นในอนาคต การผสมผสานเทคโนโลยีระดับสูงเพื่อสนับสนุนนวัตกรรมและเทคโนโลยีการบริหารจัดการข้อมูลขนาดใหญ่และเพื่อตอบสนองเทคโนโลยีที่มีการพัฒนาขึ้นอย่างรวดเร็วและด้วยขนาดข้อมูลที่อัตราการเพิ่มขึ้นอย่างรวดเร็วนี้ จึงมีการพัฒนานวัตกรรมแบบใหม่เพื่อการบริหารจัดการข้อมูลแบบต่าง ๆ และเพิ่มประสิทธิภาพการทำงานร่วมกับโปรแกรมประยุกต์หลากหลายรูปแบบ โดยสามารถปรับเปลี่ยนเทคโนโลยีการใช้งาน ได้อย่างยืดหยุ่น รวดเร็ว และมีประสิทธิภาพสูงสุด ซึ่งมีงานวิจัยที่ออกมาแล้วเหมือนกันว่าองค์กรที่ให้ความสำคัญกับข้อมูลและใช้ข้อมูลแปลงเป็นสารสนเทศเพื่อใช้ในการบริหารจัดการและตัดสินใจหรือเป็นลักษณะ Data Driven นั้นจะสร้างความได้เปรียบเชิงการแข่งขันและมีผลการดำเนินงานที่ดีกว่าองค์กรที่ไม่ได้ให้ความสำคัญกับข้อมูล [1] ผู้วิจัยจึงศึกษาค้นคว้าและนำเสนอกระบวนการบริหารและจัดการข้อมูลโดยการใช้วิเคราะห์ผลด้วยวิธีการขบวนการเหมืองข้อมูล (Process Mining) เพื่อวิเคราะห์พฤติกรรมของกระบวนการการทำงานในระบบสารสนเทศ ด้วยโปรแกรม DISCO หรือ ProM ซึ่งการส่งข้อมูลเข้าไปประมวลผลใน DISCO หรือ ProM ต้องจัดรูปแบบข้อมูลให้เหมาะสมในการประมวลผล ดังนั้นรูปแบบข้อมูลที่เหมาะสม ง่ายต่อการประมวลผล ควรจัดให้อยู่ในรูปแบบ CSV และในปัจจุบันองค์กรส่วนใหญ่ได้นำระบบสารสนเทศเข้ามาช่วยสนับสนุนการดำเนินงานทางธุรกิจของตนเองอย่างแพร่หลาย เพื่อให้สามารถแข่งขันกับคู่แข่ง รวมทั้งก้าวทันเทคโนโลยีที่เปลี่ยนแปลงอย่างรวดเร็ว หรือก้าวสู่ยุค Thailand 4.0 นั้นเอง

## ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

### Big data

Big data หมายถึง ปริมาณข้อมูลที่มีขนาดใหญ่มหาศาลเกินกว่าขีดความสามารถในการประมวลผลของระบบฐานข้อมูลธรรมดาที่จะรองรับได้ปริมาณข้อมูลที่มีขนาดใหญ่มาก ๆ จะมีอัตราการเพิ่มข้อมูลได้อย่างรวดเร็วมากและจะมีรูปแบบที่ไม่มีโครงสร้างหรือกึ่งโครงสร้าง ซึ่งไม่สามารถอยู่ในระบบฐานข้อมูลที่จะจัดเก็บข้อมูลได้ [2] ประโยชน์ Big data มีหลายประการ แต่ที่เห็นได้ชัดเจนมีอยู่ 2 ประการ คือ 1) การวิเคราะห์ (Analytical Use) คือ การวิเคราะห์ข้อมูลที่มีปริมาณมหาศาลเหล่านั้นเพื่อให้เห็นความรู้ที่ซ่อนอยู่ลึก เช่น ข้อมูลสภาพอากาศ จากเครื่องตรวจวัดจำนวนมากบนโลก ทั้งดาวเทียม เรดาร์ บอลลูน ยานตรวจอากาศ และ ทุ่นลอยในมหาสมุทร ข้อมูลจำนวนมหาศาลที่ได้รับเหล่านี้นำมาสู่การพยากรณ์อากาศที่แม่นยำเป็นรายชั่วโมงในด้านพันธุศาสตร์ นักวิทยาศาสตร์พยายามถอดรหัสพันธุกรรม เพื่อทำแผนที่ของสิ่งมีชีวิตต่าง ๆ บนโลก ซึ่งนำไปสู่การค้นพบยาใหม่ ๆ เป็นต้น และ 2) การเกิดผลิตภัณฑ์ใหม่ (Enabling New Products) โดยสามารถนำ Big data ที่มีปริมาณข้อมูลจำนวนมากมายนับมหาศาลมาใช้สร้างผลิตภัณฑ์ หรือพัฒนาการบริการเพื่อให้เหมาะสมตามความต้องการของ ลูกค้าหรือ

ผู้ใช้ เช่น การบริการจัดส่งสินค้าให้กับลูกค้าโดยที่ลูกค้าซื้อของผ่านระบบเครือข่ายบริการการจองหรือจำหน่ายตัวผ่านระบบเครือข่ายหรือการบริการโอนเงินผ่านระบบเครือข่าย เป็นต้น [3]

### Process Mining

Process Mining เป็นกระบวนการที่เกี่ยวข้องกับการตรวจสอบกิจกรรมทางธุรกิจ การบริหารจัดการกระบวนการทางธุรกิจ และการพัฒนากระบวนการทางธุรกิจ Process mining เป็นกระบวนการทางธุรกิจที่ใช้บันทึกเหตุการณ์ (Event Log) ที่เป็นพื้นฐานการทำเหมืองกระบวนการ ทำให้สามารถมองเห็นกระบวนการทางธุรกิจได้อย่างชัดเจนมากยิ่งขึ้นเห็นถึง workflow ที่มีประสิทธิภาพและข้อมูลขององค์กร ใช้การทำเหมืองกระบวนการเพื่อเพิ่มความสามารถให้สอดคล้องกับปัจจุบัน โดยนำโมเดลที่ได้มาเปรียบเทียบกับบันทึกเหตุการณ์การดำเนินงานล่าสุด โดยผลที่ออกมาจะเน้นที่ความแตกต่างของโมเดลและบันทึกเหตุการณ์นั้น ๆ ดังนั้นการทำ Process mining จึงเป็นเทคนิคหนึ่งในกระบวนการทางธุรกิจที่จะดึงข้อมูลจากบันทึกเหตุการณ์ของระบบและนำมาวิเคราะห์เป็นกระบวนการในการบริหารงานของธุรกิจ โดยมีวัตถุประสงค์เพื่อจะได้มาซึ่งกระบวนการดำเนินงานใหม่ ๆ ที่ดีเพื่อนำมาเปรียบเทียบกับรูปแบบเดิมที่มีอยู่และนำมาปรับปรุงการทำ Mining จากบันทึกเหตุการณ์จะทำให้ได้ข้อมูลที่มีประโยชน์ซึ่งหาไม่ได้จากวิธีการอื่น ๆ โดยสรุป Process mining คือ เทคนิคการดึงข้อมูลกระบวนการต่าง ๆ จากบันทึกเหตุการณ์นำมาวิเคราะห์และแสดงพฤติกรรมหลักของเหตุการณ์ออกมาเป็น Model ที่เรียกว่า Fuzzy Miner ซึ่งเป็นกระบวนการทำงานหรือเส้นทางการเดินทางของงานผ่านเครื่องมือที่เรียกว่า DISCO หรือ ProM [4]

### DISCO

DISCO เป็นเครื่องมือสำหรับการทำเหมืองกระบวนการที่สามารถใช้งานได้ง่าย เป็นเครื่องมือที่ช่วยจัดการกับข้อมูลเหตุการณ์ที่มีขนาดใหญ่และมีความซับซ้อนรวมทั้งสามารถ Filter ข้อมูลด้วยโปรแกรม DISCO ได้อีกด้วย อาทิ เช่น Timeframe เพื่อการค้นหาช่วงเวลาที่ต้องการอย่างมีประสิทธิภาพมากที่สุด Variation เพื่อหาข้อผิดพลาดและพฤติกรรมที่ไม่เหมาะสมในองค์กร Performance เพื่อการค้นหากระบวนการที่ยาวนานอย่างรวดเร็ว End points เพื่อสร้าง Model ที่ถูกต้องตามผังการทำงานขององค์กร Attribute เพื่อการจัดการประเภทข้อมูลให้ตรงตามความต้องการมากที่สุด และ Follower เพื่อหาเส้นทางการประมวลผลของข้อมูลที่ตรงตามความต้องการได้ [5]

### การดำเนินงาน

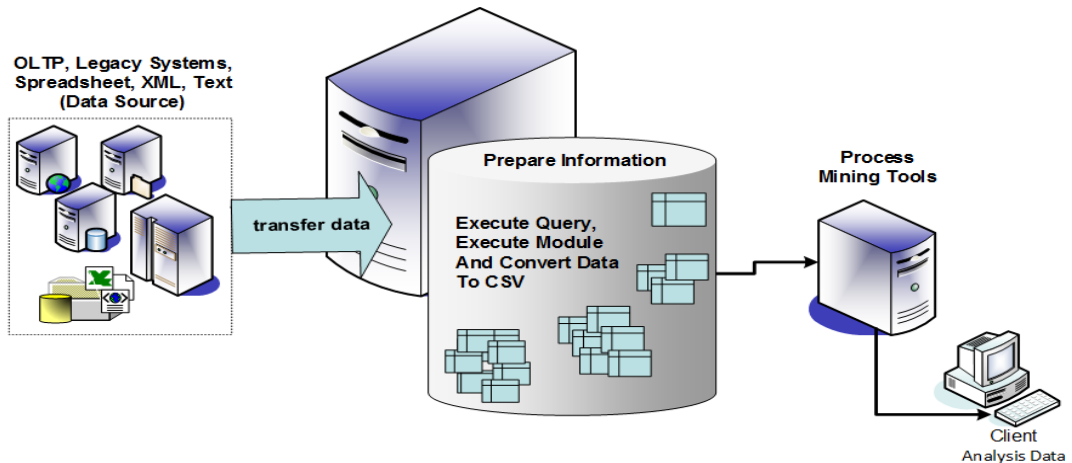
#### การจัดการข้อมูลด้วยกระบวนการเหมืองข้อมูล

การจัดการข้อมูลด้วยกระบวนการเหมืองข้อมูลเพื่อการวิเคราะห์พฤติกรรมในกระบวนการทำงานของระบบสารสนเทศ ปัจจุบันมีระบบสารสนเทศมากมายที่สามารถนำมาสนับสนุนการดำเนินงานทางธุรกิจได้ เช่น ระบบบริหารจัดการกระบวนการดำเนินงาน (Work Flow Management Systems หรือ WMS) ระบบการบริหารลูกค้าสัมพันธ์ (Customer Relationship Management หรือ CRM) ระบบการวางแผนทรัพยากรทางธุรกิจขององค์กรโดยรวม (Enterprise Planning หรือ ERP) ดังนั้นองค์กรใดสามารถบริหารจัดการข้อมูลทางธุรกิจของตนเองได้ดี และสามารถวิเคราะห์จุดได้เปรียบ เสียเปรียบ หรือแนวทางปรับกระบวนการการทำงาน ได้ถึงระดับการทำงานเชิงพฤติกรรม โดยใช้ข้อมูลเชิงประวัติ (Event Logs) จะทำให้องค์กรมีความได้เปรียบคู่แข่งสามารถต่อสู้กับคู่แข่งทางการค้า ลดต้นทุนการดำเนินงาน และเพิ่มผลกำไรให้กับธุรกิจ รวมถึงประโยชน์ด้านวางแผนพัฒนาธุรกิจ แผนกลยุทธ์ของการดำเนินงานทางธุรกิจขององค์กรได้อย่างเหมาะสม ดังนั้นการจัดการข้อมูลให้เหมาะสมต่อการวิเคราะห์ผลจึงเป็นจุดเริ่มต้นที่สำคัญ

ผู้วิจัยจึงได้ทดลองเขียนชุดคำสั่ง SQL Command และ Visual Basic.Net และขอแนะนำเสนอการจัดการข้อมูลด้วยกระบวนการเหมืองข้อมูล ในประเด็น การจัดเตรียมข้อมูล และการนำข้อมูลออกเพื่อการวิเคราะห์ผลพร้อมตัวอย่างข้อมูลที่ผ่านกระบวนการเหมืองข้อมูล ดังนี้

### การจัดเตรียมข้อมูล

การจัดเตรียมข้อมูลเพื่อวิเคราะห์ผลด้วยวิธีกระบวนการเหมืองข้อมูล การบริหารจัดการธุรกิจหรือองค์การต่าง ๆ มีการใช้ระบบสารสนเทศ และฐานข้อมูล ที่มีความแตกต่างและหลากหลาย ดังนั้นจึงมีความจำเป็นในการจัดหา หรือสร้างเครื่องมือในการจัดเตรียมข้อมูลเพื่อไม่ให้เกิดผลกระทบต่อการทำงานของผู้ปฏิบัติงานในแต่ละช่วงเวลา โดยใช้วิธีการนำเข้าข้อมูลจากฐานข้อมูลขององค์การใช้งาน เข้าสู่เครื่องมือที่จัดเตรียมไว้ ในที่นี้ผู้วิจัยได้เลือกเครื่องมือชื่อว่า Microsoft SQL Server เป็นเครื่องมือที่ผู้วิจัยจัดซื้อลิขสิทธิ์ถูกต้องตามกฎหมาย รวมทั้งโครงสร้างคำสั่งของ Microsoft SQL Server ยังสามารถนำไปประยุกต์ใช้กับเครื่องมือบริหารจัดการฐานข้อมูลอื่น ๆ ได้ แสดงดังรูปที่ 1

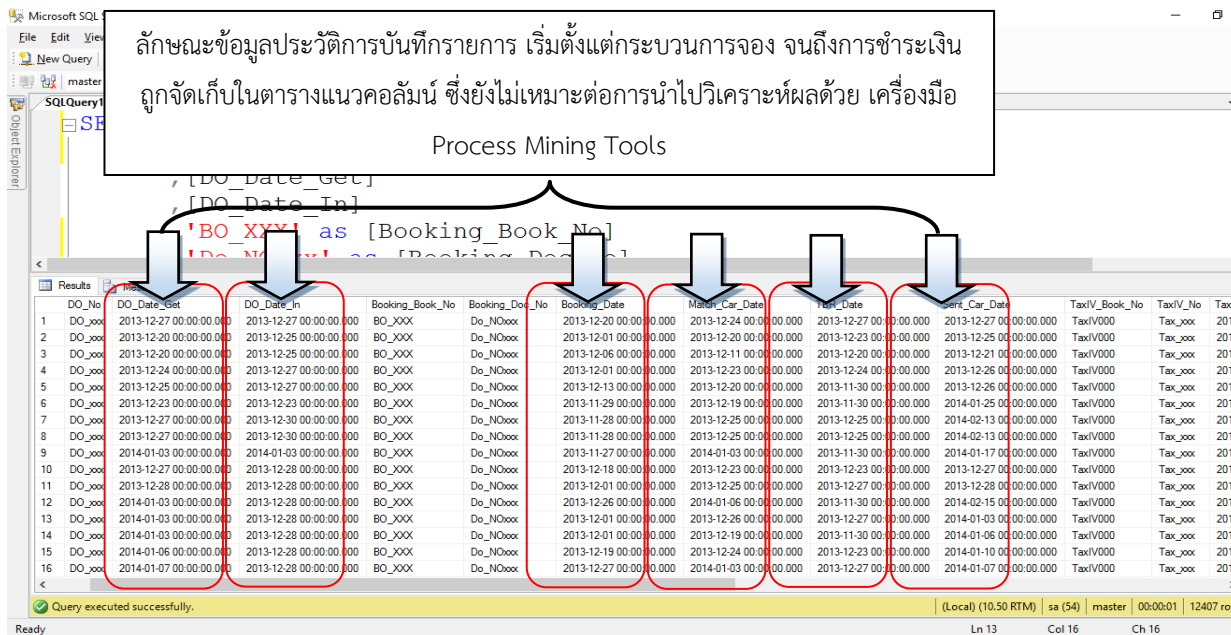


รูปที่ 1 โครงสร้างกระบวนการสำหรับการจัดเตรียมข้อมูลเพื่อวิเคราะห์ผลด้วยวิธีกระบวนการเหมืองข้อมูล (เอนก นามจันทร์, 2559)

จากรูปที่ 1 โครงสร้างกระบวนการสำหรับการจัดเตรียมข้อมูลเพื่อวิเคราะห์ผลด้วยวิธีกระบวนการเหมืองข้อมูล โดยผู้วิจัยได้เลือกเครื่องมือชื่อว่า Microsoft SQL Server เป็นเครื่องมือที่ผู้วิจัยจัดซื้อลิขสิทธิ์ถูกต้องตามกฎหมาย รวมทั้งโครงสร้างคำสั่งของ Microsoft SQL Server ยังสามารถนำไปประยุกต์ใช้กับเครื่องมือบริหารจัดการฐานข้อมูลอื่น ๆ ได้ ในการจัดเตรียมนั้น ต้องดำเนินการนำเข้าข้อมูลจากฐานข้อมูลที่ต้องการวิเคราะห์ เข้าสู่ฐานข้อมูลใหม่ที่สร้างขึ้น พร้อมทั้งศึกษา Work Flow การดำเนินงานของกระบวนการการทำงานในแต่ละขั้นตอนของระบบ ศึกษาโครงสร้างของฐานข้อมูล โครงสร้างของตาราง และความสัมพันธ์ของตารางข้อมูล (Relationship) ซึ่งรูปแบบข้อมูลที่สามารถนำเข้าเพื่อวิเคราะห์ผลด้วยโปรแกรม DISCO หรือ ProM ที่เหมาะสมและง่ายที่สุด คือ ข้อมูลให้อยู่ในรูปแบบของ CSV File และใน CSV File ควรมีข้อมูลแสดงให้เห็นถึง กิจกรรม (Activity) ต่าง ๆ ที่เกิดขึ้น และกรณี (Case) โดยที่ กรณี หรือ บางครั้งเรียกว่ากระบวนการ ณ ขณะนั้น (Process instance) คือสิ่งที่เราต้องการบริหารจัดการ เช่น กระบวนการจองรถ กระบวนการตัดรถ กระบวนการฉีก DO เป็นต้น นอกจากนี้ข้อมูลใน CSV File ควรสามารถแสดงให้เห็นถึงเวลาที่มีการกระทำของแต่ละกรณีด้วย เพื่อให้สามารถแบ่งช่วงเวลาบ่งชี้การเกิดขึ้นของกิจกรรมต่าง ๆ ได้ และสามารถวิเคราะห์ผลเชิงพฤติกรรมของผู้ใช้งานระบบ หรือกระบวนการของระบบได้ จะทำให้ทราบถึงกระบวนการการทำงานที่ก่อให้เกิดประโยชน์ต่อการดำเนินงาน รวมทั้งทำให้ทราบถึงจุดบกพร่องการกระบวนการการทำงานด้วย

**การนำข้อมูลออกเพื่อการวิเคราะห์ผลพร้อมตัวอย่างข้อมูลที่ผ่านกระบวนการเหมืองข้อมูล**

ในขั้นตอนการนำข้อมูลออกเพื่อการวิเคราะห์ผล ผู้วิจัยได้ทดลองจัดเตรียมข้อมูลเพื่อให้นำไปวิเคราะห์ผลด้วย เครื่องมือ Process Mining Tools โดยเครื่องมือที่นำมาทดลองจัดเตรียมข้อมูลคือ โปรแกรม Microsoft SQL Server และ Microsoft Visual Studio รายละเอียดปรากฏตามรูปที่ 2 ตัวอย่างข้อมูลก่อนผ่านกระบวนการจัดเตรียม ชุดที่ 1 ข้อมูลจากระบบขายรถยนต์ของบริษัทแห่งหนึ่ง และรูปที่ 4 ตัวอย่างข้อมูลก่อนผ่านกระบวนการจัดเตรียม ชุดที่ 2 ข้อมูลประวัติการเคลื่อนไหวของผู้ใช้โทรศัพท์แบบพกพา



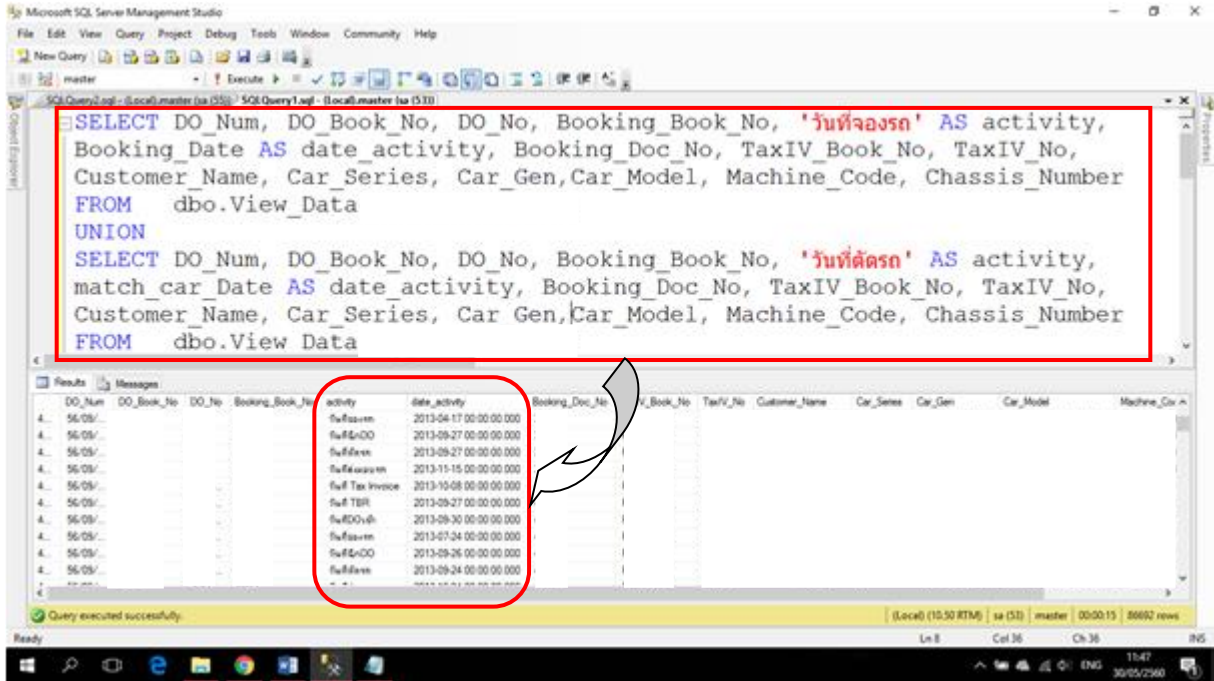
รูปที่ 2 ตัวอย่างข้อมูลก่อนผ่านกระบวนการจัดเตรียม ชุดที่ 1 ข้อมูลจากระบบขายรถยนต์ของบริษัทแห่งหนึ่ง

รูปที่ 2 ตัวอย่างข้อมูลก่อนผ่านกระบวนการจัดเตรียม ชุดที่ 1 ข้อมูลจากระบบขายรถยนต์ของบริษัทแห่งหนึ่ง จากรูปที่ 2 สามารถปรับข้อมูลที่จัดเก็บในตารางแนวคอลัมน์ ให้แสดงผลในแนวแถวได้ด้วยการเขียนคำสั่ง SQL Command แบ่งข้อมูลของแต่ละกิจกรรมออกจากกัน แล้วนำข้อมูลเหล่านั้นมารวมกันด้วยคำสั่ง Union โดยมีกิจกรรมดังต่อไปนี้

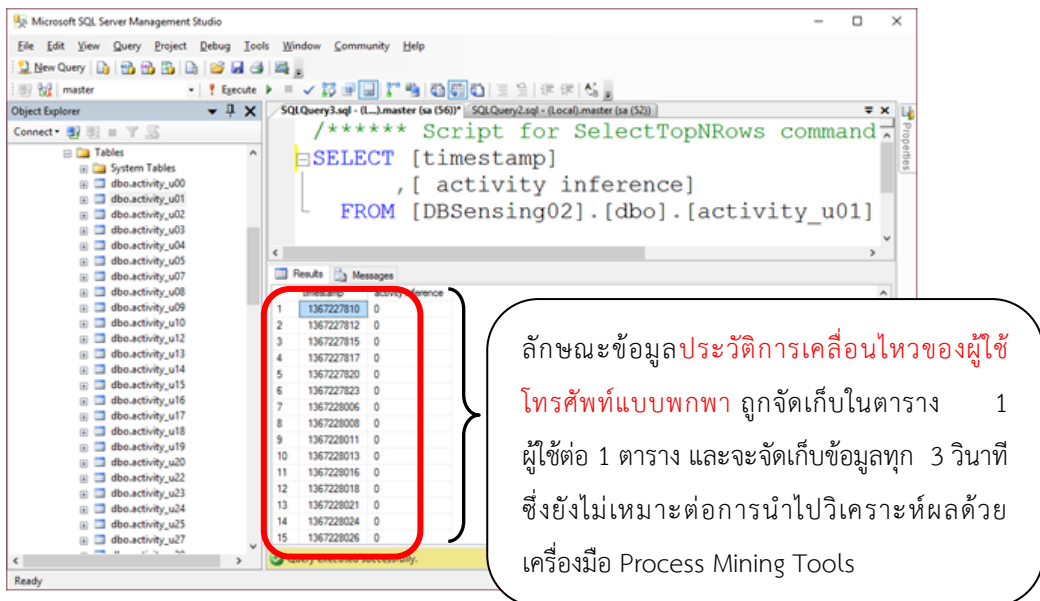
**ตารางที่ 1** กิจกรรมที่เกิดขึ้นในระบบขายรถยนต์ของบริษัทแห่งหนึ่ง

ลำดับ	ขั้นตอน	รายละเอียด
1	วันที่จองรถ	วันที่ลูกค้าตัดสินใจจะซื้อรถยนต์ใหม่จึงทำการจอง
2	วันที่ตัดรถ	สาขาแจ้งตัดรถไปสำนักงานใหญ่ เพื่อถือครองให้ลูกค้า
3	วันที่ใบรับเช็ค	ลูกค้าทำการจ่ายเงินค่ารถยนต์ใหม่เต็มจำนวนหรือบางส่วน
4	วันที่สั่งอุปกรณ์	พนักงานขายทำการสั่งอุปกรณ์เพิ่มเติมติดรถยนต์ใหม่
5	วันที่ฉีก DO	ทางสาขาออกเอกสารการส่งมอบรถยนต์ใหม่ (DO)
6	วันที่ DO เข้า	เอกสาร DO ส่งเข้าสัญญา ฝ่ายการตลาดเป็นผู้รับ
7	วันที่ตัดรถจากคลัง	ฝ่ายการตลาดแจ้งการขายเพื่อตัดรถจากคลัง
8	วันที่ส่งมอบรถ	วันที่ลูกค้ามารับรถยนต์ใหม่
9	วันที่ Tax Invoice	วันที่ออกใบเสร็จ และใบกำกับภาษีรถยนต์
10	วันที่ ส่งจดทะเบียน	วันที่ส่งเอกสารลูกค้าเพื่อทำการจดทะเบียน

จากกิจกรรมในตารางที่ 1 จะเห็นได้ว่าในระบบขายรถยนต์ 1 ระบบจะมีกิจกรรมย่อยที่เกิดขึ้นมากกว่า 1 กิจกรรม ซึ่งกิจกรรมเหล่านี้ จะต้องทำงานเรียงลำดับเป็นขั้นตอนจากลำดับที่ 1 ถึงลำดับที่ 10 แต่ข้อมูลที่ถูกจัดเก็บไว้ในตารางที่ 1 จะเห็น ว่าถูกจัดเก็บในลักษณะเรียงเป็นคอมลัมน์ ซึ่งทำให้การวิเคราะห์ผลด้วย Process Mining ไม่สามารถวิเคราะห์ เรียงลำดับจาก ลำดับที่ 1 ถึงลำดับที่ 10 ได้ ผู้วิจัยจึงได้ทดลองปรับข้อมูลจากแนวคอลัมน์ให้จัดเก็บในลักษณะเป็นแถว และแสดงข้อมูลหนึ่ง ไบขาย มีกิจกรรมเกิดขึ้นเรียงลำดับจากลำดับที่ 1 ถึงลำดับที่ 10 ในทุก ๆ ไบขาย ดังตัวอย่างคำสั่งในรูปที่ 3



รูปที่ 3 ตัวอย่างคำสั่งเพื่อสร้างข้อมูลโดยแปลงข้อมูลแนวคอลัมน์ให้แสดงผลแนวแถว จากข้อมูลตัวอย่าง ชุดที่ 1 ข้อมูลจากระบบขายรถยนต์ของบริษัทแห่งหนึ่ง



ลักษณะข้อมูลประวัติการเคลื่อนไหวของผู้ใช้โทรศัพท์แบบพกพา ถูกจัดเก็บในตาราง 1 ผู้ใช้ต่อ 1 ตาราง และจะจัดเก็บข้อมูลทุก 3 วินาที ซึ่งยังไม่เหมาะต่อการนำไปวิเคราะห์ผลด้วยเครื่องมือ Process Mining Tools

รูปที่ 4 ตัวอย่างข้อมูลก่อนผ่านกระบวนการจัดเตรียม ชุดที่ 2 ข้อมูลประวัติการเคลื่อนไหวของผู้ใช้โทรศัพท์แบบพกพา

จากรูปที่ 4 ตัวอย่างข้อมูลก่อนผ่านกระบวนการจัดเตรียม ชุดที่ 2 ข้อมูลประวัติการเคลื่อนไหวของผู้ใช้โทรศัพท์แบบพกพาซึ่งมูลตัวอย่างนี้เป็นข้อมูลที่บันทึกไว้ในอุปกรณ์แบบพกพา ซึ่งปัจจุบันหากผู้ใช้งานได้เปิดฟังก์ชันการทำงานเกี่ยวกับสุขภาพบนอุปกรณ์แบบพกพา อุปกรณ์ดังกล่าวจะจัดเก็บข้อมูล การเดิน การวิ่ง อยู่เฉย ๆ หรือไม่ทราบการกระทำ ไว้ในหน่วยความจำของอุปกรณ์แบบพกพา โดยจัดเก็บข้อมูลทุก 3 วินาที ดังนั้นเมื่อส่งข้อมูลออกมาแล้วจะได้ข้อมูลตัวอย่างดังรูปที่ 4 ตัวอย่างข้อมูลก่อนผ่านกระบวนการจัดเตรียม ชุดที่ 2 ข้อมูลประวัติการเคลื่อนไหวของผู้ใช้โทรศัพท์แบบพกพา จะเห็นได้ว่าข้อมูลของผู้ใช้อุปกรณ์แบบพกพานั้นถูกจัดเก็บ 1 ตารางต่อหนึ่งผู้ใช้ และข้อมูลบันทึกเวลาถูกจัดเก็บข้อมูลชนิดตัวเลข ที่เรียกว่า Bigint ดังนั้นจึงจำเป็นต้องแปลงชนิดข้อมูลดังกล่าวให้อยู่ในรูปแบบ วัน/เดือน/ปี เวลา (DateTime) และนำข้อมูลของผู้ใช้ที่ต้องการวิเคราะห์ผลทุกคนมารวมอยู่ในตารางเดียวกัน ในกรณีตัวอย่างนี้ผู้วิจัยจึงใช้วิธีการเขียนคำสั่ง SQL Command ร่วมกับการสร้าง Module แปลงข้อมูลด้วยโปรแกรม Microsoft Visual Studio โดยมีขั้นตอนดังนี้

1. เขียนคำสั่ง SQL Command เพื่อเพิ่มข้อมูล UserID และรวมข้อมูลผู้ใช้ทั้งหมดไว้ในตารางเดียวกัน ดังตัวอย่างคำสั่งต่อไปนี้

```
SELECT DISTINCT timestamp AS DateTimeStamp, [ activity inference] AS ActivityInference,
'u00' AS UserID FROM dbo.activity_u00
UNION
SELECT DISTINCT timestamp AS DateTimeStamp, [ activity inference] AS ActivityInference,
'u01' AS UserID FROM dbo.activity_u01 ....
```

เมื่อประมวลผลจะได้ข้อมูลดังตัวอย่างในรูปที่ 5 ลำดับที่ 2

2. เมื่อรวมข้อมูลของผู้ทุกคนเป็นที่เรียบร้อยแล้ว ในขั้นตอนนี้จะดำเนินการแปลงข้อมูลจากข้อมูลชนิดตัวเลข (Bigint) ให้อยู่ในรูปแบบ วัน/เดือน/ปี เวลา (DateTime) ตัวอย่างคำสั่ง

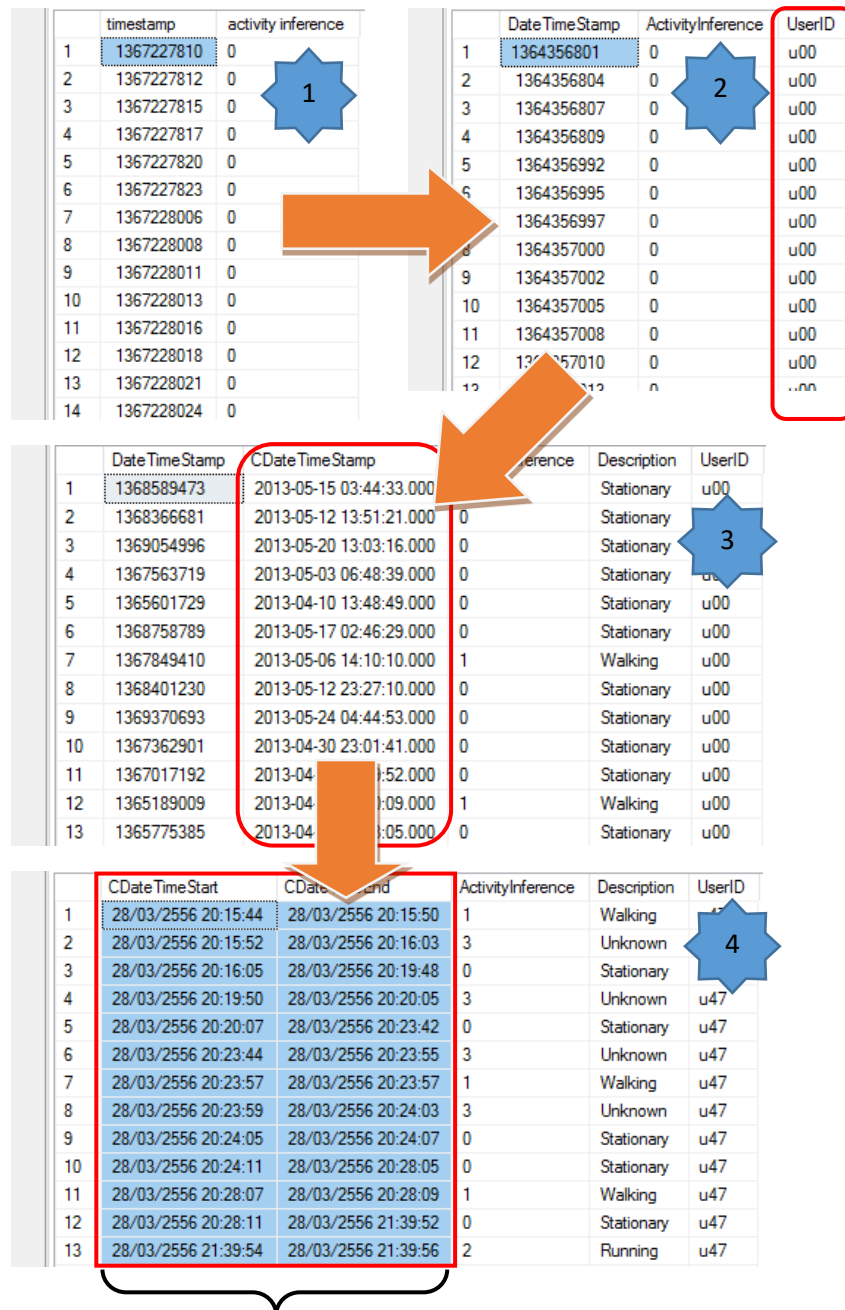
```
SELECT dbo.tb_Event.DateTimeStamp, DATEADD(SECOND, dbo.tb_Event.DateTimeStamp,
'1970-01-01 00:00:00') AS CDateTimeStamp, dbo.tb_Event.ActivityInference,
dbo.tb_Activity.Description, dbo.tb_Event.UserID
FROM dbo.tb_Activity RIGHT OUTER JOIN
dbo.tb_Event ON dbo.tb_Activity.InferenceID = dbo.tb_Event.ActivityInference
ORDER BY dbo.tb_Event.UserID
```

เมื่อประมวลผลจะได้ข้อมูลดังตัวอย่างในรูปที่ 5 ลำดับที่ 3

3. หลังจากทีแปลงวัน/เดือน/ปี:เวลา (DateTime) เป็นที่เรียบร้อยแล้ว จึงนำข้อมูลตามลำดับที่ 3 ไปประมวลผลผ่าน Module โปรแกรมเปรียบเทียบช่วงเวลา เพื่อหาเวลากิจกรรมที่เกิดขึ้นก่อนหลัง และหาเวลาเริ่มต้นกิจกรรม และสิ้นสุดกิจกรรม ได้ข้อมูลดังตัวอย่างในรูปที่ 5 ตามลำดับที่ 4

4. ดำเนินการส่งออกข้อมูลในรูปแบบ CSV File เพื่อนำไปวิเคราะห์ผลด้วยเครื่องมือ Process Mining Tools





นำวันเวลาของแต่ละกิจกรรมมาเปรียบเทียบเพื่อหา  
วันเวลาเริ่มกิจกรรมและสิ้นสุดกิจกรรม

รูปที่ 5 แสดงตัวอย่างข้อมูลระหว่างการประมวลผลในแต่ละขั้นตอน

ดังนั้นจะเห็นได้ว่าข้อมูลที่จัดเก็บไว้ในฐานข้อมูลของระบบสารสนเทศใด ๆ ที่ใช้งานองค์การสามารถนำมาใช้เพื่อวิเคราะห์ผลด้วยกระบวนการเหมืองข้อมูลได้ด้วยหลักการจัดเตรียมข้อมูล หรือแนวคิดการจัดเตรียมข้อมูลข้างต้นซึ่งแต่ละองค์การสามารถปรับแนวคิดการจัดเตรียมข้อมูลได้ตามความเหมาะสมของกระบวนการการทำงานขององค์การ และตามความต้องการขององค์การภายใต้ข้อมูลที่จัดเก็บไว้ในฐานข้อมูลที่แท้จริง ต้องคำนึงถึงความถูกต้องเป็นหลัก ไม่บิดเบือนข้อมูลหรือปรับ แต่งข้อมูลจนทำให้ข้อมูลเป็นเท็จ และคำนึงเสมอว่าการวิเคราะห์กระบวนการของระบบเพื่อหาแนวทางพัฒนากระบวนการ การทำงานให้ดีขึ้น



มากยิ่งขึ้น มุ่งสู่ความสำเร็จขององค์การเป็นหลักไม่ใช้วิเคราะห์เพื่อจับผิดกระบวนการการทำงานของผู้ปฏิบัติงาน รวมทั้งในการจัดเตรียมข้อมูลนั้นควรดำเนินการแยกจากฐานข้อมูลที่มีการใช้งานตามปกติ เพื่อไม่ให้มีผลกระทบต่อการประมวลผลและหลีกเลี่ยงการหยุดชะงักของระบบ

## สรุปผล

กระบวนการบริหารและจัดการข้อมูลโดยใช้วิเคราะห์ผลด้วยวิธีกระบวนการเหมืองข้อมูล (Process Mining) เพื่อวิเคราะห์พฤติกรรมของกระบวนการการทำงานในระบบสารสนเทศ ด้วยโปรแกรม DISCO หรือ ProM ซึ่งการส่งข้อมูลเข้าไปประมวลผลใน DISCO หรือ ProM ต้องจัดรูปแบบข้อมูลให้เหมาะสมในการประมวลผล ดังนั้นรูปแบบข้อมูลที่เหมาะสมง่ายต่อการประมวลผล ควรจัดให้อยู่ในรูปแบบ CSV

สำหรับขั้นตอนของการจัดเตรียมข้อมูลเพื่อใช้วิเคราะห์ผลด้วยวิธีกระบวนการเหมืองข้อมูล (Process Mining) เป็นขั้นตอนที่มีความซับซ้อนเนื่องด้วยหลายองค์การได้พัฒนาระบบสารสนเทศ โดยไม่คำนึงถึงการบันทึกประวัติการใช้งานในแต่ละขั้นตอนของระบบสารสนเทศไว้ในระบบฐานข้อมูล จึงทำให้ข้อมูลที่สามารถนำมาประมวลผลผ่านวิธีกระบวนการเหมืองข้อมูลทำได้ยาก และต้องใช้ผู้ใช้ที่มีประสบการณ์ มีความรู้ในระบบงานนั้น ๆ เป็นผู้จัดเตรียมข้อมูล หรือเป็นผู้ร่วมในการจัดเตรียมข้อมูล จึงจะได้ข้อมูลที่ถูกต้องตรงตามวัตถุประสงค์ที่จะวิเคราะห์ผล จากที่ผู้วิจัยได้ทดลองจัดเตรียมข้อมูลจากข้อมูลประวัติการเคลื่อนที่ของผู้ใช้โทรศัพท์มือถือ โดยจัดเก็บสถานะการเดินทาง การวิ่ง อยู่เฉย ๆ หรือ ไม่ทราบสถานะ พบว่า สามารถจัดเตรียมตามวิธีที่ผู้วิจัยได้ทดลอง และนำไปใช้วิเคราะห์ พบว่า คนหนึ่งคน ในหนึ่งวัน ใช้เวลาเดินโดยเฉลี่ยเท่าไร ใช้เวลาวิ่งโดยเฉลี่ยเท่าไร อยู่เฉย ๆ ใช้เวลาเท่าไร และยังสามารถทราบถึงพฤติกรรมของผู้ใช้มือถือด้วยว่าส่วนใหญ่จะทำกิจกรรมใดในแต่ละวัน นอกจากนี้ ผู้วิจัยได้ทดลองกับข้อมูลกระบวนการขายสินค้าชนิดหนึ่ง พบว่าวิธีดังกล่าวสามารถใช้ได้จริง

ดังนั้นองค์การหรือธุรกิจใดสามารถบริหารจัดการข้อมูลมหาศาลด้วยวิธีกระบวนการเหมืองข้อมูลได้ดีและสามารถวิเคราะห์จุดได้เปรียบ เสียเปรียบ หรือแนวทางปรับกระบวนการการทำงานได้ถึงระดับการทำงานเชิงพฤติกรรมโดยใช้ข้อมูลเชิงประวัติจะทำให้้องค์การมีความได้เปรียบคู่แข่งสามารถต่อสู้กับคู่แข่งทางการค้า ลดต้นทุนการดำเนินงาน และเพิ่มผลกำไรให้กับธุรกิจ รวมถึงประโยชน์ด้านวางแผนพัฒนาธุรกิจ แผนกลยุทธ์ของการดำเนินงานทางธุรกิจขององค์การได้อย่างเหมาะสม รวมทั้งก้าวทันเทคโนโลยีที่เปลี่ยนแปลงอย่างรวดเร็วหรือก้าวสู่ยุค Thailand 4.0

## เอกสารอ้างอิง

- [1] พสุ เดชะรินทร์ (2556,) *Big data หรือ อภิมหาข้อมูล*. (ออนไลน์). สืบค้นเมื่อ 15 มกราคม 2560, จาก (<http://library.acc.chula.ac.th/PageController.php?page=FindInformation/ArticleACC/2556/Pasu/BangkokBiznews/B2901131>)
- [2] Dumbill, E. (2012, January 11). *What is big data? An introduction to the big data landscape*. Retrieved July 24, 2012, from <http://radar.oreilly.com/2012/01/What-is-big-data.html>.
- [3] Panasas. (2012, 13 March). *Big Data*. Retrieved July 27, 2012, from <http://www.panasas.com/solutions/big-data>.
- [4] Nammakhunt, A., Romsaiyud, W., Porouhan, P. & Premchaiswadi, W. (2012). *Process mining: Converting data from MS-Access Database to MXML format*. In ICT and Knowledge Engineering (ICT & Knowledge Engineering 2012), 2012 10th International Conference on (pp. 205-212). IEEE.
- [5] Arpasat, P., Porouhan, P., & Premchaiswadi, W. (2015). *Improvement of call center customer service in a thai bank using disco fuzzy mining algorithm*. In ICT and Knowledge Engineering (ICT & Knowledge Engineering 2015), 2015 13th International Conference on (pp. 90-96). IEEE.