รายงานวิจัยฉบับสมบูรณ์

โครงการย่อยที่ 1

ระบบแนะนำสถานที่ท่องเที่ยวเฉพาะบุคคล

(Personalized POI Recommendation)

โดย

ผู้ช่วยศาสตราจารย์ ดร.ไกรศักดิ์ เกษร

กุมภาพันธ์ 2558

สัญญาเลขที่ RDG5650020

รายงานวิจัยฉบับสมบูรณ์

โครงการย่อยที่ 1

ระบบแนะนำสถานที่ท่องเที่ยวเฉพาะบุคคล

(Personalized POI Recommendation)

คณะผู้วิจัย	สังกัด
1. ผู้ช่วยศาสตราจารย์ ดร.ไกรศักดิ์ เกษร	มหาวิทยาลัยนเรศวร

ชุดโครงการ ระบบให้บริการข้อมูลการเดินทางเฉพาะบุคคลสำหรับการท่องเที่ยวไทย

สนับสนุนโดยสำนักงานคณะกรรมการวิจัยแห่งชาติ (วช.)

และสำนักงานกองทุนสนับสนุนการวิจัย (สกว.)

(ความเห็นในรายงานนี้เป็นของผู้วิจัย วช. สกว. ไม่จำเป็นต้องเห็นด้วยเสมอไป)

บทสรุปผู้บริหาร

อินเทอร์เน็ตมีบทบาทต่อการพัฒนาอุตสาหกรรมการท่องเที่ยวอย่างมากทั้งเว็บไซต์สำหรับให้บริการข้อมูลการท่องเที่ยวและพัฒนาขึ้นเป็นพาณิชย์อิเล็กทรอนิกส์ซึ่งเป็นเครื่องมือที่ช่วยเผยแพร่ข้อมูลการท่องเที่ยวได้อย่างมีประสิทธิภาพและเป็นผลดีต่อนักท่องเที่ยวที่สามารถหาข้อมูลการท่องเที่ยวได้หลากหลายขึ้น แต่ข้อมูลเกี่ยวกับการท่องเที่ยวจำนวนมากที่เกิดขึ้นนี้ทำให้นักท่องเที่ยวต้องใช้เวลาเพิ่มมากขึ้นในการค้นหาข้อมูลการท่องเที่ยวให้ได้ตรงกับความต้องการดังนั้นเพื่อลดเวลาการค้นหาข้อมูลของนักท่องเที่ยวระบบแนะนำข้อมูล (Recommendation System-RS) เป็นวิธีหนึ่งที่ช่วยแก้ไขปัญหาดังกล่าวได้โดยระบบแนะนำข้อมูลนั้นถูกนำมาใช้อย่างแพร่หลายในธุรกิจพาณิชย์อิเล็กทรอนิกส์ เช่น ระบบแนะนำข้อมูลหนังสือ วารสาร ข่าว รายการโทรทัศน์ เพลง เว็บเพ็จการท่องเที่ยว ฯลฯ ต่อมาได้มีการพัฒนาระบบแนะนำข้อมูลในรูปแบบการแนะนำข้อมูลเฉพาะบุคคล (Personalized Recommendation System-PRS) ที่จะช่วยให้ผู้ใช้สามารถค้นหาสินค้าและบริการที่ต้องการแบบเฉพาะบุคคลได้ในเวลาที่รวดเร็วขึ้น ในขอบเขตของการท่องเที่ยว ระบบแนะนำข้อมูลเฉพาะบุคคลได้รับความสนใจและนำมาใช้สนับสนุนนักท่องเที่ยวทั้งในการวางแผนการเดินทาง ช่วยจองบัตรโดยสารเครื่องบินในวันหยุดเทศกาล และแนะนำแพ็คเกจท่องเที่ยว เป็นต้น นอกจากนักท่องเที่ยวจะได้สินค้าและบริการที่ตรงกับความต้องการผู้ให้บริการยังได้ประโยชน์จากการเสนอสินค้าและบริการที่เหมาะสมซึ่งเป็นการสร้างกลยุทธ์ทางการตลาดเพื่อเพิ่มจำนวนผู้ใช้บริการให้มากยิ่งขึ้นและเพิ่มโอกาสในการตัดสินใจซื้อสิ้นค้าและบริการได้ง่ายขึ้นอีกด้วย

ในระบบแนะนำข้อมูลเฉพาะบุคคลนั้นการค้นหาความสนใจของผู้ใช้เป็นหัวใจสำคัญในการที่จะแนะนำผลิตภัณฑ์หรือบริการให้ได้ตรงกับความต้องการของผู้ใช้อย่างแท้จริงได้ ด้วยความสำคัญของปัญหาดังกล่าวงานวิจัยนี้จึงมีหลายแนวคิดที่จะพัฒนาเทคนิคการค้นหาความสนใจที่สามารถแก้ปัญหาข้างต้นและได้ข้อสรุปความสนใจของผู้ใช้ได้อย่างถูกต้องและมีประสิทธิภาพ เนื่องจากปัจจุบันเครือข่ายสังคม (Social Network) นอกจากจะช่วยโฆษณาประชาสัมพันธ์การท่องเที่ยวได้แล้วยังเป็นที่เก็บข้อมูลจำนวนมากจากผู้ใช้ประกอบกับการเจริญเติบโตของเครือข่ายสังคมที่เป็นไปอย่างรวดเร็วทำให้ข้อมูลทั้งที่เกิดจากการสนทนา แสดงความคิดเห็น หรือแบ่งปันมีจำนวนที่มากขึ้น ข้อมูลจำนวนมหาศาลเหล่านี้ยังไม่ถูกนำไปใช้ประโยชน์เท่าที่ควร อย่างไรก็ตามผู้วิจัยส่วนหนึ่งพยายามจะใช้ประโยชน์จากข้อมูลเหล่านี้โดยการนำข้อมูลจากเครือข่ายสังคมมาวิเคราะห์เพื่อหาความสนใจของผู้ใช้ โดยข้อมูลสำคัญที่นำมาวิเคราะห์คือข้อมูลการเช็คอิน (Check-in) ในเฟสบุ๊ค แล้วนำมาประมวลผลโดยใช้ค่าทางสถิติและหลังจากนั้นระบบจะพยายามแนะนำสถานที่ท่องเที่ยวให้กับนักท่องเที่ยวโดยการแนะนำจะพิจารณาความสนใจของผู้ใช้เป็นหลัก โดยระบบจะแนะนำสถานที่ท่องเที่ยวให้กับผู้โดย พิจารณาจากข้อมูลการเช็คอินของผู้ใช้ในสถานที่ต่างๆ ที่ผู้ใช้เคยเช็คอินจากอดีตที่ผ่านมา โดยดูจากความถี่ของสถานที่ที่ผู้ใช้เช็คอินและนำมาผนวกกับความนิยมของสถานที่ดังกล่าว หากสถานที่ใดมีจำนวนเช็คอินสูงแสดงว่าได้รับความนิยมจากนักท่องเที่ยวเป็นจำนวนมาก ในกรณีที่ผู้ใช้มีข้อมูลการเช็คอินไม่เพียงพอระบบจะไปหาข้อมูลเช็คอินจากเพื่อนๆ ของผู้ใช้คนนั้นๆ เนื่องจากมีสมมติฐานว่าเพื่อนกันมักจะมีความชอบคล้ายๆ กันโดยเฉพาะเรื่องของการท่องเที่ยว แต่อย่างไรก็ตามระบบยังรอบรับผู้ใช้ที่ไม่เป็นสมาชิกของเฟสบุ๊ค เนื่องจากระบบไม่ทราบว่าผู้ใช้เป็นใคร อายุ หรือเพศใด จึงไม่สามารถแนะนำเฉพาะส่วนบุคคลได้ ระบบจะแนะนำข้อมูลจากแบบสำรวจว่าสถานที่ใดในจังหวัดต่างๆ เป็นที่นิยม

ของนักท่องเที่ยวบ้าง ซึ่งข้อมูลได้มากจากแบบสอบถามและข้อมูลการเช็คอินของผู้ใช้สังคมออนไลน์ และนำค่าเฉลี่ยของสถานที่ต่างๆ มานำเสนอต่อผู้ใช้

บทสรุปของโครงการวิจัยนี้คือ ข้อมูลเครือข่ายสังคมออนไลน์ เช่น เฟสบุ๊ค นั้นเป็นสังคมออนไลน์อันดับหนึ่งของคนไทยที่ใช้เพื่อติดต่อสื่อสารกับเพื่อนๆ และข้อมูลที่มีอยู่อย่างมหาศาลในเฟสบุ๊คสามารถเป็นประโยชน์ต่อการนำไปใช้หาความสนใจของผู้ใช้ได้อย่างมีประสิทธิภาพ โดยพบว่าระบบแนะนำสถานที่ตรงกับความสนใจของผู้ใช้ได้อย่างถูกต้องสูงสุดถึง 76% และผู้ใช้ส่วนใหญ่เห็นว่าระบบดังกล่าวมีประโยชน์ช่วยในการแนะนำสถานที่ท่องเที่ยวอย่างมาก

บทคัดย่อ

ปัญหาสำคัญของนักท่องเที่ยวคือการหาข้อมูลของสถานที่ท่องเที่ยวต่างๆ ในเมืองหรือจังหวัดที่ตนเองสนใจอยากจะเดินทางไป และการหาข้อมูลของสถานที่ท่องเที่ยวต่างๆ นักท่องเที่ยวต้องใช้เวลาอย่างมากในการค้นหาข้อมูลซึ่งกระจายอยู่ตามเว็บไซต์ต่างๆ เพื่อกรองหาสถานที่ที่ตนเองสนใจอยากจะไป ดังนั้นงานวิจัยนี้จึงมุ่นเน้นในการแก้ปัญหาดังกล่าวโดยการใช้เทคโนโลยีทางวิทยาการคอมพิวเตอร์เข้ามาช่วยแนะนำสถานที่ท่องเที่ยวให้กับนักเดินทาง โดยอาศัยข้อมูลของผู้ใช้จำนวนมหาศาลที่อยู่บนเครือข่ายสังคมออนไลน์ (Online Social Network) โดยเฉพาะเกี่ยวกับการท่องเที่ยวของผู้ใช้ เช่น ข้อมูลการเช็คอิน (Check-in) ข้อมูลการกดไลค์ (Like) ต่างๆ เพื่อนำไปเป็นประโยชน์ในการคำนวณสรุปหาความสนใจของนักท่องเที่ยวแต่ละคนและแนะนำสถานที่ท่องเที่ยวเพื่อให้สอดคล้องกับความสนใจดังกล่าว นอกจากนี้ข้อมูลจากเครือข่ายสังคมยังสามารถนำไปใช้เพื่อแนะนำผู้ใช้คนอื่นๆ ที่ไม่เป็นสมาชิกของเครือข่ายสังคมออนไลน์ใดๆ เลยก็ตาม ซึ่งงานวิจัยนี้จะใช้ข้อมูลจากเฟสบุ๊ค (Facebook) เพื่อเป็นข้อมูลตัวอย่างในการทดลอง กระบวนการสำคัญของงานวิจัยนี้ประกอบด้วย 4 ขั้นตอนได้แก่ 1) การดึงข้อมูลจากเฟสบุ๊ค 2) การวิเคราะห์ความสนใจของผู้ใช้ 3) การแนะนำสถานที่ท่องเที่ยว และ 4) การตรวจจับการปรับเปลี่ยนความสนใจของนักท่องเที่ยว ผลจากการทดลองพบว่าข้อมูลจากเครือข่ายสังคมออนไลน์สามารถใช้เพื่อหาความสนใจของผู้ใช้และระบบสามารถแนะนำสถานที่ท่องเที่ยวต่างๆ ได้อย่างมีประสิทธิภาพโดยมีความถูกต้องถึง 76% โดยใช้เทคนิคการตอบกลับจากผู้ใช้

Abstract

A major obstacle of most tourists is to gather attractions information for desired visiting places. This task is great time consuming and laborious because those information is distributed into several websites over the Internet. As such, this research aims to resolve this problem by employing the computer technologies and methodologies to find the places that match user interests. With the emergence of social network, many useful data are available as check-in and Like data from Facebook. These data are employed for recommendation system regarding to user interest. In addition, they are also advantage to recommend the popular places for none Facebook user based on check-in statistic of Facebook users. There are four key processes in the proposed framework including: 1) Facebook data extraction, 2) user interest analysis, 3) attraction recommendation, and 4) user interest change detection. The experiments show that data from social network (Facebook) enables the user interest computation and can be used to enhance the recommendation system efficiency up to 76% of accuracy using the relevance feedback technique.

สารบัญ

บทสรุปผู้บริหาร ..... **ก**
บทคัดย่อ ..... **ค**
ABSTRACT ..... **ง**
สารบัญ ..... **จ**
สารบัญตาราง ..... **ช**
สารบัญรูป ..... **ช**

บทที่ 1 บทนำ **1**
1.1 ความเป็นมาหลักการเหตุผล 1
1.2 วัตถุประสงค์ของการวิจัย 2
1.3 ขอบเขตของการวิจัย 2
1.4 สมมติฐาน 3
1.5 ระยะเวลาทำการวิจัยและแผนการดำเนินงานตลอดโครงการวิจัย 3

บทที่ 2 เอกสารและงานวิจัยที่เกี่ยวข้อง **6**
2.1 ระบบแนะนำข้อมูลการท่องเที่ยวเฉพาะบุคคล 6
2.2 องค์ประกอบของระบบแนะนำข้อมูลเฉพาะบุคคล 7
2.3 การค้นหาความสนใจของผู้ใช้ (USER INTEREST DETECTION) 8
2.4 การสร้างแบบจำลองความสนใจของผู้ใช้ (USER INTEREST MODELING) 11
2.5 การแนะนำข้อมูล (RECOMMENDATION TECHNIQUES) 13
2.6 เครือข่ายสังคมออนไลน์กับการนำมาใช้ในระบบแนะนำข้อมูลเฉพาะบุคคล 16
2.7 เทคนิคตรวจสอบการเปลี่ยนแปลงความสนใจของผู้ใช้ (ข้อมูลตอบกลับของผู้ใช้) 23

บทที่ 3 วิธีดำเนินงานวิจัย **25**
3.1 กรอบแนวคิดงานวิจัย 25
3.2 กระบวนการทำงานของระบบสำหรับผู้ใช้แบบสมาชิก (MEMBER) 26
3.3 การสร้างแบบจำลองความสนใจของผู้ใช้และการปรับปรุงข้อมูลประวัติของผู้ใช้ (USER INTERESTS MODEL CONSTRUCTION AND USER PROFILE UPDATING) 27
3.4 วิเคราะห์ความสนใจของผู้ใช้จากข้อมูลในเครือข่ายสังคมออนไลน์ (USER INTERESTS ANALYSIS USING ONLINE SOCIAL NETWORK INFORMATION) 31
3.5 การแนะนำข้อมูลการท่องเที่ยว (RECOMMENDATION) 38
3.6 วิเคราะห์ความสนใจของผู้ใช้จากข้อมูลตอบกลับของผู้ใช้ (RELEVANCE FEEDBACK ANALYSIS) 41

3.7 ปรับเปลี่ยนแบบจำลองผู้ใช้โดยอัตโนมัติ (USER MODEL ADAPTATION) 42
3.8 วิเคราะห์ความสนใจของผู้ใช้จากข้อมูลโดยตรงจากผู้ใช้ 44

บทที่ 4 ผลการดำเนินการวิจัย 47
4.1 เครื่องมือที่ใช้ในการศึกษา 47
4.2 การวิเคราะห์ความต้องการของผู้ใช้ 49
4.3 ผลการออกแบบและพัฒนาระบบ 52
4.4 การประเมินประสิทธิภาพของระบบทดสอบระบบ 57

บทที่ 5 สรุปผลการวิจัยและข้อเสนอแนะ 71
5.1 สรุปผลการวิจัย 71
5.2 งานที่จะพัฒนาต่อในอนาคต 72

บรรณานุกรม 73
ภาคผนวก 75
ภาคผนวก ก : บทความตีพิมพ์ผลงาน 76

สารบัญตาราง

ตารางที่ 1 แผนการดำเนินงานวิจัย 3
ตารางที่ 2 ข้อดีของการใช้การได้มาซึ่งข้อมูลความสนใจของผู้ใช้ลักษณะต่างๆ 9
ตารางที่ 3 ข้อดีของการสร้างแบบจำลองความสนใจของผู้ใช้ด้วยรูปแบบการจัดเก็บแบบต่างๆ 13
ตารางที่ 4 ข้อดีในการเลือกใช้เทคนิคการแนะนำข้อมูลประเภทต่างๆ 15
ตารางที่ 5 ความหมายของแต่ละตารางและฟิลด์ข้อมูล 29
ตารางที่ 6 ข้อมูลเพื่อระบุตัวตนของผู้ใช้ 33
ตารางที่ 7 ข้อมูลการเช็คอินของผู้ใช้จำนวน 1 เช็คอิน 33
ตารางที่ 8 หมวดหมู่ (category_list) ของสถานที่ที่ถูกเช็คอิน 34
ตารางที่ 9 เกณฑ์ในการพิจารณาค่าระดับความสนใจและค่าน้ำหนักจากพฤติกรรมการตอบกลับของผู้ใช้ 42
ตารางที่ 10 เปรียบเทียบลำดับความสนใจในหมวดหมู่สถานที่ท่องเที่ยวที่แท้จริงกับลำดับที่วิเคราะห์โดยระบบของผู้ใช้ 1 คน 59
ตารางที่ 11 ลำดับความสนใจในหมวดหมู่สถานที่ท่องเที่ยวของผู้ใช้ A เปรียบเทียบกับกลุ่มเพื่อน 60
ตารางที่ 12 ลำดับความสนใจในหมวดหมู่ของสถานที่ท่องเที่ยวและเวลาในการประมวลผลของหนึ่งผู้ใช้ 61
ตารางที่ 13 ลักษณะของผู้ตอบแบบสอบถามความคิดเห็น 65
ตารางที่ 14 ข้อมูลสถานะภาพของผู้ตอบแบบสอบถามความคิดเห็น 65
ตารางที่ 15 ข้อมูลช่วงอายุของผู้ตอบแบบสอบถามความคิดเห็น 65
ตารางที่ 16 ข้อมูลระดับการศึกษาของผู้ตอบแบบสอบถามความคิดเห็น 66
ตารางที่ 17 ผลการประเมินความพึงพอใจของผู้ใช้ด้านส่วนติดต่อผู้ใช้งาน (User Interface) 67
ตารางที่ 18 ผลการประเมินความพึงพอใจของผู้ใช้ด้านการแนะนำสถานที่ท่องเที่ยว 68
ตารางที่ 19 ผลการประเมินความพึงพอใจของผู้ใช้ด้านการแนะนำแผนการเดินทาง 68
ตารางที่ 20 ข้อมูลความคิดเห็นเกี่ยวกับการใช้งานแอปพลิเคชั่นในการช่วยหาข้อมูลท่องเที่ยวครั้งต่อไป 69
ตารางที่ 21 ข้อมูลความคิดเห็นเกี่ยวกับการใช้งานแอปพลิเคชั่นในการวางแผนการท่องเที่ยวครั้งต่อไป 70
ตารางที่ 22 ข้อมูลความคิดเห็นเกี่ยวกับการแนะนำแอปพลิเคชั่นนี้กับบุคคลอื่น 70

สารบัญรูป

รูปที่ 1 แสดงแนวโน้มการค้นหาข้อมูลการท่องเที่ยว โดยใช้คำว่า “เที่ยวไทย” ค้นหาใน Google.com 1
รูปที่ 2 องค์ประกอบของระบบแนะนำข้อมูลการท่องเที่ยวเฉพาะบุคคล 8
รูปที่ 3 แบบจำลองฐานข้อมูลแบบ Flat File..... 11
รูปที่ 4 แบบจำลองฐานข้อมูลสัมพันธ์ 12
รูปที่ 5 ตัวอย่างข้อมูลประเภทข้อความจากทวิตเตอร์ 17
รูปที่ 6 ตัวอย่างข้อมูลประเภทแท็กจากอินสตาแกรม..... 18
รูปที่ 7 ตัวอย่างข้อมูลประเภทเช็คอินจากเฟสบุ๊ค..... 18
รูปที่ 8 การสกัดข้อมูลความสนใจของผู้ใช้จากข้อความบนทวิตเตอร์โดยใช้ฐานความรู้ของ Wikipedia..... 20
รูปที่ 9 ระบบแนะนำข้อมูลเว็บเพ็จโดยใช้การสกัดความสนใจจากแท็กในโซเชียลบุ๊กมาร์ก..... 21
รูปที่ 10 เครือข่ายสังคมที่ได้รับความนิยมในประเทศไทยในปี 2013..... 22
รูปที่ 11 ประเภทข้อมูลที่นิยมเผยแพร่ในเครือข่ายสังคมออนไลน์เฟสบุ๊ค..... 23
รูปที่ 12 ภาพรวมกรอบแนวคิดของโครงการวิจัย..... 25
รูปที่ 13 วิธีการระบุตัวตนผู้ใช้..... 26
รูปที่ 14 วิธีการสร้างโมเดลผู้ใช้และการปรับปรุงโมเดล..... 28
รูปที่ 15 ฐานข้อมูลที่ใช้จัดเก็บแบบจำลองความสนใจของผู้ใช้..... 29
รูปที่ 16 วิธีการวิเคราะห์ข้อมูลจากเครือข่ายสังคม 32
รูปที่ 17 ตัวอย่างฟิลด์ข้อมูลของผู้ใช้ที่สามารถดึงมาจากเฟสบุ๊ค..... 33
รูปที่ 18 การแนะนำข้อมูลการท่องเที่ยว..... 39
รูปที่ 19 การวิเคราะห์ข้อมูลตอบกลับของผู้ใช้ 41
รูปที่ 20 ปรับเปลี่ยนแบบจำลองผู้ใช้โดยอัตโนมัติ 43
รูปที่ 21 การวิเคราะห์ความสนใจจากข้อมูลโดยตรงจากผู้ใช้..... 45
รูปที่ 22 ระดับความสนใจในสถานที่ท่องเที่ยวของผู้ใช้..... 45
รูปที่ 23 ขั้นตอนการดำเนินงานวิจัย 49
รูปที่ 24 แหล่งค้นหาข้อมูลสถานที่ท่องเที่ยวต่างๆ 50
รูปที่ 25 สัดส่วนการใช้สังคมออนไลน์เพื่อหาข้อมูลการท่องเที่ยว..... 51
รูปที่ 26 ความสำคัญในการพัฒนาระบบการให้ข้อมูลการท่องเที่ยวเฉพาะบุคคล..... 52
รูปที่ 27 การคลิกเพื่อเลือกจุดหมายปลายทางและช่วงเวลา..... 52
รูปที่ 28 หน้าจอหน้าจอแนะนำสถานที่ท่องเที่ยว..... 53
รูปที่ 29 หน้าจอการปรับค่าความสนใจ..... 54
รูปที่ 30 เลือกสถานที่เพื่อสร้างแผนการเดินทาง 55
รูปที่ 31 ปุ่มเริ่มประมวลผลข้อมูลจากเฟสบุ๊ค 56
รูปที่ 32 ระบบแนะนำสถานที่ท่องเที่ยวไปยังผู้ใช้..... 56
รูปที่ 33 แบบสอบถามลำดับความสนใจ 58

รูปที่ 34 ค่าเฉลี่ยความถูกต้องของลำดับหมวดหมู่ความสนใจและเวลาเฉลี่ยในการประมวลผล 61
รูปที่ 35 ค่า precision ของการแนะนำเปรียบเทียบระหว่างการใช้เกณฑ์ที่ต่างกัน 62
รูปที่ 36 รายการแนะนำที่เปลี่ยนไปเมื่อนำพฤติกรรมการตอบกลับมาคำนวณคำแนะนำใหม่ 63
รูปที่ 37 ค่าพรีซิชั่นของการแนะนำเปรียบเทียบระหว่างการปรับปรุงคำแนะนำใหม่ 64

บทที่ 1

บทนำ

1.1 ความเป็นมาหลักการเหตุผล

การวางแผนการท่องเที่ยวเป็นกิจกรรมที่นักท่องเที่ยวส่วนใหญ่จะทำก่อนการเดินทาง แต่กิจกรรมดังกล่าวเป็นสิ่งที่ไม่ง่ายสำหรับนักท่องเที่ยว เนื่องจากมีความซับซ้อนและยุ่งยากเพราะนักท่องเที่ยวขาดประสบการณ์และไม่มีความรู้เกี่ยวกับแหล่งท่องเที่ยวที่ต้องการจะเดินทางไป ดังนั้นการวางแผนการท่องเที่ยวต้องอาศัยความรู้และความสามารถพิเศษในการวางแผน (Ambite et al., 2002) จากสถิติจาก Google (http://www.google.com/trends/) พบว่าผู้ใช้หรือนักท่องเที่ยวจำนวนมากทำการค้นหาข้อมูลการท่องเที่ยวผ่านทางระบบค้นคืนข้อมูล เช่น Google และมีแนวโน้มที่เพิ่มขึ้นทุกปี (แสดงดังรูปที่ 1) จะเห็นว่าอินเทอร์เน็ตมีบทบาทที่สำคัญในการช่วยนักท่องเที่ยวในการค้นหาข้อมูลที่เป็นประโยชน์ในการวางแผนต่างๆ

รูปที่ 1 แสดงแนวโน้มการค้นหาข้อมูลการท่องเที่ยว โดยใช้คำว่า “เที่ยวไทย” ค้นหาใน Google.com

ซึ่งสอดคล้องกับนโยบายของรัฐบาลที่ต้องการส่งเสริมการท่องเที่ยวของประเทศ และยุทธศาสตร์การวิจัยของสำนักงานคณะกรรมการวิจัยแห่งชาติ (วช.) โดยส่งเสริมการพัฒนากลยุทธ์เพื่อดึงดูดให้นักท่องเที่ยวเดินทางเข้ามาท่องเที่ยวในประเทศไทยมากขึ้น ประเด็นหนึ่งกล่าวว่า

“พัฒนาด้านการตลาดและประชาสัมพันธ์การท่องเที่ยว โดยสนับสนุนให้ภาคเอกชนเข้ามามีส่วนร่วมในการกำหนดนโยบายและกลยุทธ์ด้านการตลาด และกำหนดกลยุทธ์การประชาสัมพันธ์เพื่อสนับสนุนจุดขายที่มีความโดดเด่นของไทยให้เป็นศูนย์กลางระดับโลก...” (สำนักงานคณะกรรมการวิจัยแห่งชาติ, 2553)

จากการวิเคราะห์นโยบายดังกล่าวจะเห็นว่า ส่วนหนึ่งต้องมีกลยุทธ์ด้านการตลาดและการประชาสัมพันธ์ที่ดี ดังนั้นการประชาสัมพันธ์ที่ดีคือมีการให้ข้อมูลที่ตรงกับความสนใจของกลุ่มเป้าหมาย (By

Target Audience) การประชาสัมพันธ์จึงจะได้ผล ถึงแม้ปัจจุบันจะมีอินเทอร์เน็ตเป็นเครื่องมือสำคัญในการประชาสัมพันธ์การท่องเที่ยว แต่การประชาสัมพันธ์ที่ไม่ตรงกับความสนใจของผู้ซื้อก็คงไม่เกิดประโยชน์แต่อย่างใดหรือเกิดประโยชน์น้อย ดังนั้นนอกจากเราจะใช้อินเทอร์เน็ตเป็นเครื่องมือในการประชาสัมพันธ์เพื่อนำเสนอรูปแบบการท่องเที่ยวที่ในรูปแบบต่างๆ แล้ว ระบบการประชาสัมพันธ์นี้ควรจะมีความสามารถที่จะหาความสนใจของผู้ใช้และนำเสนอข้อมูลให้ตรงต่อรสนิยมและความต้องการของนักท่องเที่ยวได้อย่างถูกต้องแม่นยำ

ดังนั้นปัญหาที่สำคัญคือการจะได้มาหรือทราบความสนใจ (User's Interests) ของนักท่องเที่ยวที่ค้นหาข้อมูลเกี่ยวกับประเทศไทยผ่านอินเทอร์เน็ตเป็นเรื่องที่ท้าทายและเป็นประโยชน์อย่างยิ่งในการโฆษณาประชาสัมพันธ์ข้อมูลการท่องเที่ยว การประยุกต์ใช้คอมพิวเตอร์เข้ามาวิเคราะห์เพื่อหาความสนใจที่แท้จริงของนักท่องเที่ยวเป็นเรื่องที่น่าสนใจ และเมื่อทราบความสนใจของนักท่องเที่ยวแล้ว ระบบสามารถที่จะนำเสนอข้อมูลหรือประชาสัมพันธ์ข้อมูลเกี่ยวการท่องเที่ยวต่างๆ ตลอดจนให้คำแนะนำที่ตรงกับความสนใจของนักเที่ยว นอกจากนี้ระบบยังสามารถช่วยวางแผนการเดินทาง ซึ่งการนำเสนอนี้เป็นการนำเสนอข้อมูลที่ตรงกับความสนใจเฉพาะบุคคล (Personal Interests) ประโยชน์หลักของการศึกษาวิจัยนี้คือเมื่อเรานำเสนอข้อมูลได้ตรงกับความสนใจของนักท่องเที่ยวมากขึ้น ผลที่ตามมาคือการนำความสนใจนั้นไปใช้ในการแนะนำข้อมูลและวางแผนการเดินทางได้อย่างมีประสิทธิภาพ และส่งผลต่อรายได้ของประเทศที่เพิ่มขึ้นนั่นเอง

การศึกษาของโครงการวิจัยนี้มีความสอดคล้องกับยุทธศาสตร์การวิจัยแห่งชาติ (พ.ศ. 2555-2559) ยุทธศาสตร์การวิจัยที่ 2: การสร้างศักยภาพและความสามารถเพื่อการพัฒนาทางเศรษฐกิจ ในกลยุทธ์การวิจัยที่ 4: พัฒนาศักยภาพทางเศรษฐกิจการท่องเที่ยว

1.2 วัตถุประสงค์ของการวิจัย

1) เพื่อพัฒนาเทคนิควิธีการในการค้นหาความสนใจที่แท้จริงของนักท่องเที่ยว เพื่อนำเสนอข้อมูลการท่องเที่ยวที่ตรงกับความสนใจเฉพาะบุคคล
2) เพื่อการนำเสนอข้อมูลท่องเที่ยวที่ตรงกับความสนใจของผู้ใช้ได้อย่างมีประสิทธิภาพ
3) เพื่อให้ระบบคอมพิวเตอร์อื่นๆ นำความสนใจที่ได้รับจากระบบไปใช้ประโยชน์ในการวางแผนการเดินทางได้

1.3 ขอบเขตของการวิจัย

ขอบเขตด้านข้อมูล

1. ตัวอย่างข้อมูลการใช้งานของนักท่องเที่ยวทุกคนที่เข้ามาระบบผ่านเครือข่ายอินเทอร์เน็ตซึ่งได้จากอาสาสมัคร

ขอบเขตด้านการทำงานของระบบ

1. ระบบสามารถหาความสนใจของผู้ใช้ได้ โดยใช้ข้อมูลจากเครือข่ายสังคมออนไลน์
2. ระบบสามารถนำเสนอข้อมูลโดยเรียงลำดับข้อมูลสถานที่ท่องเที่ยวตามความสนใจของผู้ใช้จากมากไปน้อยได้

1.4 สมมติฐาน

การค้นหาความสนใจเฉพาะบุคคลเป็นส่วนสำคัญในการพัฒนาระบบแนะนำข้อมูลเฉพาะบุคคลโดยข้อมูลที่สกัดจากเครือข่ายสังคมที่มีอยู่จำนวนมากบนอินเทอร์เน็ตอาจใช้เป็นตัวแทนเพื่อแสดงความสนใจของแต่ละผู้ใช้ได้จึงนำไปสู่สมมติฐานที่ 1 และ 2 ดังนี้

สมมติฐานที่ 1 : ข้อมูลจากเครือข่ายสังคมสามารถนำมาวิเคราะห์และสกัดเพื่อหาความสนใจเฉพาะบุคคลของผู้ใช้ได้

สมมติฐานที่ 2 : ข้อมูลจากเครือข่ายสังคมสามารถนำมาแก้ไขปัญหาโคลด์สตาร์ทได้

การเปลี่ยนแปลงความสนใจของผู้ใช้ที่อาจเกิดขึ้นได้ตลอดเวลา เป็นหัวข้อหนึ่งที่สำคัญในระบบแนะนำข้อมูล การตรวจจับการเปลี่ยนแปลงที่เกิดขึ้นจะทำให้การวิเคราะห์ข้อมูลเพื่อแนะนำสถานที่ท่องเที่ยวมีประสิทธิภาพมากขึ้นได้ดังสมมติฐานที่ 3 ของงานวิจัยคือ

สมมติฐานที่ 3: การนำข้อมูลตอบกลับจากผู้ใช้มาใช้ประโยชน์ สามารถทำให้ระบบทราบการเปลี่ยนแปลงความสนใจของผู้ใช้และแนะนำสถานที่ท่องเที่ยวได้ถูกต้องมากยิ่งขึ้น

1.5 ระยะเวลาทำการวิจัยและแผนการดำเนินงานตลอดโครงการวิจัย

ในการทำโครงงานวิจัยนี้สามารถแบ่งเป็นขั้นตอนหลักๆ ในการทำวิจัยดังมีรายละเอียดในตารางที่ 1

ตารางที่ 1 แผนการดำเนินงานวิจัย

วัตถุประสงค์	กิจกรรม	ผลสำเร็จ %	เหตุผลกรณีล่าช้า
1. วิเคราะห์ข้อจำกัดของระบบแนะนำข้อมูลเฉพาะบุคคลแบบเดิม แนวทางการประยุกต์ใช้และตั้งสมมติฐานต่างๆ เพื่อใช้ในการพัฒนาระบบใหม่ที่สร้างขึ้น	1.1 ทำความเข้าใจถึงปัญหา ทำการวิเคราะห์ปัญหา (Critical analysis and problem identification) ที่เกิดขึ้นในระบบปัจจุบันและระบุถึงข้อจำกัดต่างๆ ของงานวิจัยเดิมที่มีอยู่	100	

วัตถุประสงค์	กิจกรรม	ผลสำเร็จ %	เหตุผลกรณีล่าช้า
	1.2 ทำการทบทวนวรรณกรรม (Literature survey) การรวบรวมวรรณกรรมที่เกี่ยวข้องนั้นสามารถทำได้ผ่านทางสื่ออินเทอร์เน็ตและระบบห้องสมุดอิเล็กทรอนิคส์ของมหาวิทยาลัยนเรศวร	100	
2. ระบุถึงเทคนิคที่เป็นไปได้ทั้งหมดที่จะนำมาใช้ในการศึกษาวิจัย โดยการคัดเลือกเทคนิคที่มีความเป็นไปได้ในการแก้ปัญหา และเลือกเทคนิคมากกว่า 1 วิธี โดยที่ลำดับที่ 1 คือวิธีการที่เป็นไปได้มากที่สุด หากในระหว่างการทดลองวิธีการที่ 1 ไม่สามารถแก้โจทย์ที่ต้องการได้ จะนำวิธีการที่ 2 มาใช้ซึ่งเป็นวิธีการสำรอง ทั้งนี้เพื่อลดความเสี่ยงในการล้มเหลวของการพัฒนาระบบ	2.1 ทำการคิดค้นเทคนิคเพื่อแก้ปัญหาและข้อจำกัดของระบบเดิมที่อยู่ 2.2 ติดต่อประสานงานจัดซื้อวัสดุ อุปกรณ์ที่จำเป็นต่อการทำวิจัย	100 100	
3. รวบรวมข้อมูลต่างๆ ที่เกี่ยวข้องทั้งหมด ที่จำเป็นต่อการศึกษาวิจัย	3.1 ทำการติดต่อประสานงานกับองค์กรต่างๆ เพื่อขอข้อมูลสถานที่ท่องเที่ยวและข้อมูลต่างๆ 3.2 เก็บรวบรวมข้อมูลสถานที่ท่องเที่ยวและข้อมูลต่างๆ ที่จำเป็นต่อการศึกษาวิจัย	100 100	
4. ทำการออกแบบเฟรมเวิร์คใหม่ (Design framework) เพื่อแก้ปัญหาที่ระบุไว้ (Software specification)	4.1 ทำการออกแบบเฟรมเวิร์คใหม่ (Design framework) และทำการระบุงานต่างๆ (Tasks decomposition) เพื่อดูว่าแต่ละงานรองรับวัตถุประสงค์ที่วางไว้ครบถ้วนหรือไม่ หากไม่ครบถ้วนก็จะทำการปรับปรุงเฟรมเวิร์คที่ได้ออกแบบไว้ทำเช่นนี้ไปเรื่อยๆ จนกว่าเฟรมเวิร์คจะรองรับวัตถุประสงค์ที่วางไว้ทุกข้อ และทุกส่วนทำงานเข้ากันได้	100	
5. พัฒนาโปรแกรมตามที่ได้ออกแบบไว้และควบคุมการพัฒนาระบบให้เป็นไปตามทิศทางที่ตั้งไว้	5.1 ควบคุมดูแลและตรวจสอบความถูกต้องของการพัฒนาโปรแกรม 5.2 ทำการสร้างระบบ (Implementation) ตามที่ได้ออกแบบไว้ โดยใช้ระบบคอมพิวเตอร์ส่วนบุคคล (PC) ของภาควิชาวิทยาการคอมพิวเตอร์ฯ คณะวิทยาศาสตร์ มหาวิทยาลัยนเรศวร	100 100	เนื่องจากระบบมีการทำงานที่ซับซ้อนกว่าที่คาดการณ์ไว้ และต้องมีการทำงานร่วมกับโครงการย่อยที่ 2 จึงทำให้เสีย

วัตถุประสงค์	กิจกรรม	ผลสำเร็จ %	เหตุผลกรณีล่าช้า
	5.3 ประสานงานและอำนวยความสะดวกต่างๆ ในการทำงานของนักพัฒนาโปรแกรม รวมถึงการเบิกจ่ายค่าตอบแทนในการพัฒนาโปรแกรม	100	เวลานานในการตรวจสอบการทำงานร่วมกันเพื่อการทำงานของระบบที่สมบูรณ์และเต็มรูปแบบมากที่สุด
6.ตรวจสอบความถูกและประเมินประสิทธิภาพของระบบที่ได้สร้างขึ้น	6.1 ประเมินระบบได้สร้างขึ้น (Evaluation) โดยใช้ตัวอย่างข้อมูลที่เก็บรวบรวมไว้ ทำการรวบรวมผลการทดลองที่ได้ วิเคราะห์ผลการทดลอง (Analysis) แปรผลการทดลอง (Interpret) และสรุปผลผลการทดลอง	80	เนื่องจากการทดสอบการทำงานของระบบมีความล่าช้าจึงส่งผลกระทบในการประเมินประสิทธิภาพของระบบโดยผู้ใช้ ทำให้เกิดความล่าช้าตามไปด้วย
7.คู่มือการใช้ระบบ (User Manual) และขอจดลิขสิทธิ์	7.1 เขียนคู่มือการใช้โปรแกรม 7.2 เตรียมเอกสารเพื่อยื่นขอจดลิขสิทธิ์	100 0	เนื่องจากไม่แน่ใจว่าลิบสิทธิ์ของระบบจะต้องเป็นของใครหากเป็นของผู้ให้ทุน ดังนั้น สกว. จะต้องยื่นขอจดลิขสิทธิ์ของระบบนี้ด้วยตนเอง
8. เขียนรายงานเสนอต่อ วช. เพื่อปิดโครงการ	8.1 เขียนคู่มือการใช้ระบบ 8.2 ส่งรายงานให้กับภาคส่วนต่างๆ ที่ได้ประโยชน์จากโครงการวิจัย	100	ต้องการรอการปรับแก้เป็นเวอร์ชั่นสุดท้ายที่ได้รับข้อแนะนำจากคณะกรรมการเรียบร้อยแล้ว จึงจะดำเนินการเผยแพร่รายงานไปยังภาคส่วนที่เกี่ยวข้อง เช่น ททท. หรือเนคเท็คต่อไป

บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

การพัฒนาระบบแนะนำข้อมูลการท่องเที่ยวเฉพาะบุคคลนี้ ผู้วิจัยได้ทำการศึกษาทฤษฎี เอกสารและงานวิจัยที่เกี่ยวข้อง ดังนี้

2.1 ระบบแนะนำข้อมูลการท่องเที่ยวเฉพาะบุคคล

2.2 องค์ประกอบของระบบแนะนำข้อมูลเฉพาะบุคคล

2.3 เครือข่ายสังคมออนไลน์กับการนำมาใช้ในระบบแนะนำข้อมูลเฉพาะบุคคล

2.4 เทคนิคตรวจสอบการเปลี่ยนแปลงความสนใจของผู้ใช้ (ข้อมูลตอบกลับของผู้ใช้)

2.1 ระบบแนะนำข้อมูลการท่องเที่ยวเฉพาะบุคคล

ในการทบทวนวรรณกรรมและศึกษางานวิจัยที่เกี่ยวข้อง ผู้วิจัยพบว่ามีนักวิจัยเสนอแนวคิดเกี่ยวกับการแนะนำข้อมูลการท่องเที่ยวหลายแนวคิดด้วยกัน เช่น (Adler and Blue, 1998) ได้นำเสนอแนวคิดเกี่ยวกับการค้นหาความสนใจของนักท่องเที่ยวโดยสร้างข้อมูลประวัติ (Profile) ของนักท่องเที่ยวและปรับเปลี่ยนแผนการเดินทางโดยขึ้นอยู่กับข้อมูลประวัติของผู้ใช้ (Poslad et al., 2001) ได้นำเสนอ“CRUMPET” ซึ่งเป็นระบบที่สามารถให้ข้อมูลการเดินทางกับนักท่องเที่ยว จุดเด่นของระบบนี้คือใช้วิธีการค้นหาความสนใจของผู้ใช้โดยอ้อม (Implicit) ซึ่งเป็นวิธีที่ระบบทำการเก็บข้อมูลและวิเคราะห์ความสนใจในขณะที่ผู้ใช้ใช้งานระบบอยู่ ทำให้ระบบนำเสนอเส้นทางการเดินทางแก่นักท่องเที่ยวโดยอัตโนมัติโดยที่ผู้ใช้ไม่ต้องกรอกข้อมูลใดๆ ให้กับระบบในขณะที่ (Soo and Liang, 2001) นำเสนอระบบที่ชื่อว่า “Zje system” ซึ่งใช้วิธีการตรงกันข้ามกับวิธีการก่อนหน้าคือการสอบถามข้อมูลโดยตรงจากผู้ใช้ (Explicit) และนำข้อมูลที่ได้ไปวางแผนการเดินทางให้แก่นักท่องเที่ยว ข้อเสียของวิธีการนี้คือผู้ใช้ต้องใช้เวลาในการกรอกข้อมูลมากเกินไปซึ่งผู้ใช้ส่วนมากไม่อยากเสียเวลาในขั้นตอนเหล่านี้มากนัก นอกจากนี้ยังมีระบบที่ชื่อว่า “Dietorecs” (Fesenmaier et al., 2003) เป็นระบบอัตโนมัติที่ให้บริการแนะนำแพ็คเก็จท่องเที่ยวต่างๆ หรือให้ข้อมูลเกี่ยวกับการเดินทางที่นักท่องเที่ยวต้องการ เช่น สถานที่พักรถเช่า เป็นต้น โดยเว็บไซต์ Expedia (www.expedia.com) ได้นำไปใช้งานร่วมกับระบบหลัก ซึ่งหลักการของ Dietorecs คือการเก็บข้อมูลตัวอย่างแผนการท่องเที่ยว (Case Base) ไว้ในระบบจากนั้นคำนวณค่าความเหมือน (Similarity) ของความสนใจระหว่างข้อมูลประวัติของผู้ใช้กับข้อมูลประวัติของตัวอย่างที่มี ข้อเสียของระบบนี้คือหากข้อมูลตัวอย่างไม่ตรงกับผู้ใช้ระบบก็ไม่สามารถแนะนำแผนการท่องเที่ยวที่ตรงกับความต้องการของผู้ใช้ได้ จะเห็นว่าระบบแนะนำข้อมูลเกี่ยวกับการท่องเที่ยวข้างต้นยังไม่มีระบบใดนำข้อมูลจากเครือข่ายสังคมออนไลน์เข้ามาช่วยในการค้นหาความสนใจของผู้ใช้เลย อย่างไรก็ตามมีการนำเครือข่ายสังคมออนไลน์มาช่วยในระบบแนะนำข้อมูลประเภทอื่นๆ ซึ่งจะกล่าวถึงในหัวข้อ 2.6

แนวทางที่สำคัญในการพัฒนาเว็บไซต์ที่มีการให้ข้อมูลหรือแนะนำข้อมูลต่างๆ ตลอดจน ระบบค้นคืนสารสนเทศ (Search engine) มีแนวโน้มที่จะพยายามให้ข้อมูลที่ตรงกับความต้องการของผู้ใช้แต่ละคน (Personalized Information Presentation and Retrieval-PIPR) มากที่สุด นั่นหมายความว่าผู้ใช้ที่ใช้คิวรี่

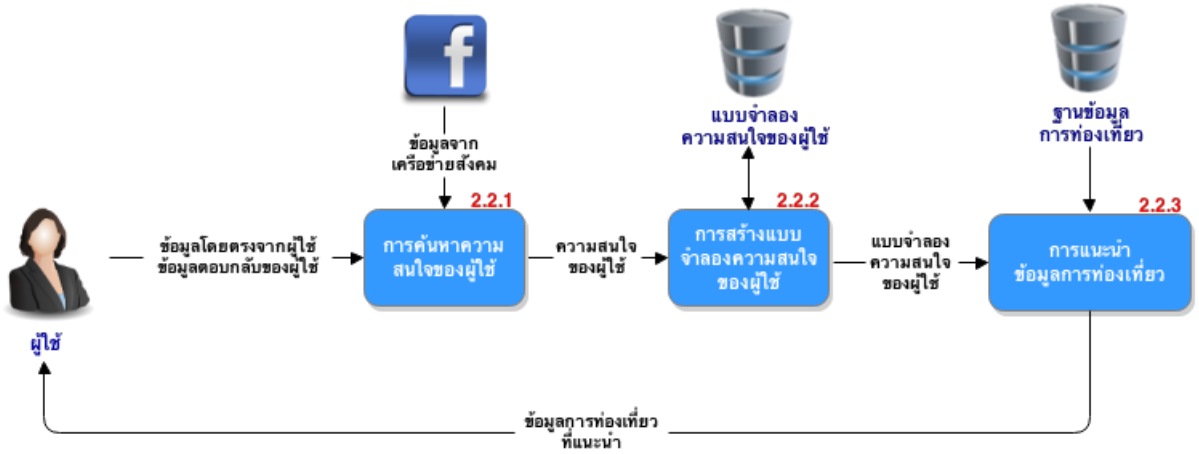
(Query) เหมือนกัน อาจจะมีความสนใจที่ต่างกัน เช่น นักท่องเที่ยวทั้งสองคนอยากมาท่องเที่ยวที่จังหวัดพิษณุโลก คนหนึ่งชอบท่องเที่ยวสถานที่ที่เป็นธรรมชาติแต่อีกคนกลับชอบสถานที่ที่เป็นสถาปัตยกรรมโบราณ ผลลัพธ์ที่ได้จากการค้นหาคำว่า “สถานที่ท่องเที่ยว พิษณุโลก”ควรจะออกมาไม่เหมือนกัน นั่นคือเว็บไซต์ควรจะมีความฉลาดขึ้นในการนำเสนอข้อมูลให้กับผู้ใช้ที่ต่างกัน เช่น นำเสนอข้อมูลเกี่ยวกับน้ำตก อุทยานแห่งชาติ ในจังหวัดพิษณุโลกให้กับผู้ใช้คนที่หนึ่ง ในขณะที่ระบบจะนำเสนอข้อมูลเกี่ยวกับอุทยานประวัติศาสตร์ให้กับผู้ใช้คนที่สองเป็นต้น ดังนั้นจุดประสงค์หลักของ PIPR คือปรับปรุงประสิทธิภาพการนำเสนอและค้นหาข้อมูลโดยนำความสนใจของผู้ใช้แต่ละคนเป็นเงื่อนไขในการค้นหาข้อมูลด้วย (Vallet et al., 2007) ซึ่งแนวคิดนี้ถือเป็นแนวคิดที่สำคัญอันหนึ่งในการพัฒนาระบบค้นคืนสารสนเทศ เว็บไซต์ที่มีการให้ข้อมูล รวมถึงเว็บไซต์ที่มีระบบแนะนำข้อมูลต่างๆ โดยระบบแนะนำข้อมูลที่นำหลักการ PIPR ไปใช้ เรียกว่าระบบแนะนำข้อมูลเฉพาะบุคคลซึ่งจะอธิบายในหัวข้อถัดไป

2.2 องค์ประกอบของระบบแนะนำข้อมูลเฉพาะบุคคล

ระบบแนะนำข้อมูลเฉพาะบุคคล เป็นระบบที่ใช้ความสนใจของแต่ละผู้ใช้เป็นเงื่อนไขหนึ่งในการพิจารณาข้อมูลเพื่อนำมาแนะนำให้กับผู้ใช้ เช่นเดียวกันกับระบบแนะนำข้อมูลการท่องเที่ยวเฉพาะบุคคลเป็นระบบที่นำเสนอข้อมูลการท่องเที่ยวในลักษณะต่างๆ เช่น ข้อมูลสถานที่ท่องเที่ยว แผนการเดินทาง ข้อมูลแพ็คเกจท่องเที่ยว เป็นต้น ให้กับนักท่องเที่ยวตามความสนใจที่แตกต่างกันไปในแต่ละบุคคล ลักษณะองค์ประกอบของระบบแนะนำข้อมูลการท่องเที่ยวเฉพาะบุคคลที่แต่ละผู้วิจัยได้พัฒนาขึ้นจะแตกต่างกันไปในรายละเอียด แต่โดยองค์ประกอบหลักแล้วจะประกอบไปด้วย 2 ส่วนที่สำคัญ คือ 1) การค้นหาความสนใจของผู้ใช้ และ 2) การแนะนำข้อมูลการท่องเที่ยวตามความสนใจของผู้ใช้

ระบบแนะนำข้อมูลการท่องเที่ยวเฉพาะบุคคลมีองค์ประกอบหลักดังรูปที่ 2 ประกอบด้วย 1) การค้นหาความสนใจของผู้ใช้ (User Interest Detection) เป็นการหาข้อสรุปความสนใจและปรับปรุงให้ทันต่อความสนใจที่เปลี่ยนไปของผู้ใช้ผ่านแหล่งที่มาหลายแหล่ง เช่น ข้อมูลที่ได้จากผู้ใช้โดยตรง ข้อมูลจากเครือข่ายสังคมออนไลน์ ข้อมูลตอบกลับ เป็นต้น 2) การสร้างแบบจำลองความสนใจของผู้ใช้ (User Interest Modeling) เป็นการการจัดเก็บข้อมูลความสนใจ ลงในฐานข้อมูล และ 3) การแนะนำสถานที่ท่องเที่ยว (Tourism Attractions Recommendation) เป็นการแนะนำสถานที่ท่องเที่ยวโดยใช้ข้อมูลความสนใจที่ได้มาจากแบบจำลองความสนใจของผู้ใช้

ทั้งนี้ผู้วิจัยได้ทำการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องกับองค์ประกอบหลักทั้งสาม โดยจะอธิบายเพิ่มเติมในหัวข้อที่ 2.3 2.4 และ 2.5 ตามลำดับ



รูปที่ 2 องค์ประกอบของระบบแนะนำข้อมูลการท่องเที่ยวเฉพาะบุคคล

2.3 การค้นหาความสนใจของผู้ใช้ (User Interest Detection)

การค้นหาความสนใจของผู้ใช้เป็นกระบวนการที่สำคัญของระบบแนะนำข้อมูลแบบเฉพาะบุคคลเป็นอย่างยิ่ง เนื่องจากในการที่จะแนะนำผลิตภัณฑ์หรือบริการให้ได้ตรงกับความต้องการของผู้ใช้จะต้องทราบถึงความสนใจที่มีของผู้ใช้ นอกจากนี้ยังต้องเรียนรู้ความสนใจที่เปลี่ยนไปของผู้ใช้ที่เกิดขึ้นได้ตลอดเวลา เพื่อให้ผู้ใช้ได้รับข้อมูลการแนะนำที่มีประสิทธิภาพมากยิ่งขึ้น โดยในกระบวนการนี้ประกอบไปด้วย 2 ส่วน คือ 1) การได้มาซึ่งข้อมูลความสนใจของผู้ใช้ (User Interest Acquisition) และ 2) การเรียนรู้ความสนใจของผู้ใช้ (User Interest Learning)

1) **การได้มาซึ่งข้อมูลความสนใจของผู้ใช้ (User Interest Acquisition)**

การได้มาซึ่งข้อมูลความสนใจของผู้ใช้ แบ่งเป็น 2 ลักษณะ คือ 1) การได้มาซึ่งข้อมูลความสนใจของผู้ใช้โดยตรง (Explicit) หมายถึงการที่ระบบได้รับข้อมูลจากผู้ใช้โดยตรง เช่น ข้อมูลรายละเอียดส่วนบุคคลที่ผู้ใช้กรอกให้กับระบบเมื่อลงทะเบียน การตอบคำถามของผู้ใช้ การทำแบบสอบถาม เป็นต้น สำหรับ 2) การได้มาซึ่งข้อมูลความสนใจของผู้ใช้โดยอ้อม (Implicit) คือการที่ระบบทำการเก็บข้อมูลจากการใช้งานหรือพฤติกรรมของผู้ใช้ที่ไม่เป็นการขอข้อมูลจากผู้ใช้โดยตรง เช่น การคลิกการโหวต การแสดงความคิดเห็นบนระบบ เป็นต้น โดยตัวอย่างระบบที่ใช้ข้อมูลโดยตรงจากผู้ใช้ เช่น Entrée (Burke, 2000) และ Triplehop's TripMatcher (Kabassi, 2010) ได้ใช้ประโยชน์จากข้อมูลที่สอบถามจากผู้ใช้โดยตรง แต่ปัญหาที่เกิดคือผู้ใช้อาจต้องสูญเสียเวลาในขั้นตอนการตอบคำถามเหล่านี้มากเกินไปซึ่งเป็นการรบกวนผู้ใช้ นอกจากนี้ผู้ใช้ยังอาจไม่สามารถอธิบายลักษณะความพึงพอใจของตนเองได้อย่างถูกต้องผ่านวิธีการนี้ ในขณะที่ PTA (Coyle and Cunningham, 2003) และ GUIDE (Cheverst et al., 2002) ใช้การได้มาซึ่งข้อมูลความสนใจของผู้ใช้โดยอ้อมผ่านแหล่งที่มาหลายแหล่ง วิธีการนี้สามารถขจัดปัญหาที่เกิดกับระบบที่ขอข้อมูลจากผู้ใช้โดยตรงได้ อย่างไรก็ตามยังมีปัญหาสำคัญคือ ผู้ใช้ที่เริ่มต้นใช้งานระบบจะยังไม่มีข้อมูลโดยอ้อมของผู้ใช้มากพอที่จะสกัดความสนใจเพื่อแนะนำข้อมูลให้ตรงกับความสนใจได้ ซึ่งเรียกปัญหานี้ว่า โคลด์สตาร์ท เพื่อลดปัญหาระหว่างการได้มาซึ่งข้อมูลความสนใจของผู้ใช้ทั้งสองวิธีข้างต้น จึงมีการ

ผนวกทั้งสองวิธีเข้าด้วยกันเช่นใน TravelPlanner (Chin and Porage, 2001) จะเลือกแบบสอบถามที่มีประโยชน์ที่สุดในการเรียนรู้ความสนใจให้กับผู้ใช้และใช้ข้อมูลอื่นๆ ที่ได้มาจากทางอ้อมด้วย เช่นเดียวกับ Zhu และคณะ (2012) ใช้ข้อมูลที่ผู้ใช้ทำการลงทะเบียนเพื่อสร้างแบบจำลองผู้ใช้เริ่มต้น จากนั้นจะใช้ข้อมูลโดยอ้อมผ่านพฤติกรรมผู้ใช้ เช่น พฤติกรรมการคลิก พฤติกรรมการโหวต พฤติกรรมการตอบและพูดคุย เป็นต้น เพื่อเรียนรู้ความสนใจของผู้ใช้ ในทำนองเดียวกันกับ SPETA (García-Crespo et al., 2009) เมื่อผู้ใช้เริ่มต้นใช้งานระบบ ระบบจะขอข้อมูลเกี่ยวกับประเภทสถานที่ที่ต้องการเข้าไปเยี่ยมชมและการให้คะแนนไปยังสถานที่ที่ผู้ใช้สนใจ จากนั้นก็จะเรียนรู้ความสนใจผ่านพฤติกรรมของผู้ใช้และข้อมูลที่ถูกสกัดจากเครือข่ายสังคมออนไลน์ที่ได้เป็นสมาชิกอยู่ จะเห็นว่าการผสมผสานการใช้ข้อมูลทั้งโดยตรงและโดยอ้อมสามารถช่วยลดทั้งปัญหาโคลด์สตาร์ทและการสูญเสียเวลาในการตอบคำถามของผู้ใช้ โดยปัญหาที่เกิดขึ้นในลักษณะการได้มาของข้อมูลด้วยวิธีต่างๆ แสดงดังตารางที่ 2

ตารางที่ 2 ข้อดีของการใช้การได้มาซึ่งข้อมูลความสนใจของผู้ใช้ลักษณะต่างๆ

ข้อดี	การได้มาซึ่งข้อมูลความสนใจของผู้ใช้ (User Interest Acquisition)		
	โดยตรง (Explicit)	โดยอ้อม (Implicit)	ผสมผสาน (Hybrid)
1. ผู้ใช้ใช้เวลาน้อยลงในการกรอกข้อมูลความสนใจให้กับระบบ	✗	✓	✓
2. แก้ปัญหาการอธิบายลักษณะหรือระดับความพึงพอใจของผู้ใช้ ในกรณีที่ใช้การขอข้อมูลจากผู้ใช้โดยตรง	✗	✓	✓
3. ระบบมีข้อมูลเพียงพอที่จะนำมาใช้สกัดหาความสนใจ เมื่อผู้ใช้เริ่มต้นใช้ระบบ	✓	✗	✓

ปัญหาโคลด์สตาร์ทเป็นปัญหาที่เกิดขึ้นเมื่อผู้ใช้เริ่มต้นใช้งานระบบแต่ระบบยังไม่มีข้อมูลของผู้ใช้มากพอที่จะสามารถสกัดความสนใจและให้การแนะนำสินค้าหรือบริการได้อย่างถูกต้องโดยเฉพาะอย่างยิ่งในระบบแนะนำข้อมูลเฉพาะบุคคล ซึ่งปัญหาโคลด์สตาร์ทนอกจากจะใช้วิธีการผสมผสานการใช้ข้อมูลจากโดยตรงและโดยอ้อมจากผู้ใช้ที่ได้กล่าวไปข้างต้นแล้ว ยังมีแนวคิดการนำข้อมูลจากผู้ใช้อื่นมาช่วยแก้ปัญหาอีกด้วย เช่น Rashid และคณะ (2008) ได้พยายามแก้ไขปัญหานี้โดยใช้ข้อมูลจากผู้ใช้อื่นๆ ที่มีต่อระบบแนะนำภาพยนตร์ เช่น ข้อมูลภาพยนตร์ที่กำลังเป็นที่นิยม ซึ่งคำนวณจากความถี่ในการเขาถึงของผู้ใช้ข้อมูล จำนวนของการแสดงความคิดเห็นข้อมูล การให้คะแนนภาพยนตร์ของผู้ใช้ เป็นต้น ภาพยนตร์ที่มีลำดับคะแนนที่สูงก็จะถูกแนะนำให้กับผู้ใช้ที่เริ่มต้นใช้งานระบบ จะเห็นว่าปัญหานี่ถูกแก้ไขในระดับหนึ่งโดยใช้ความสนใจของผู้ใช้อื่นโดยรวม ต่างจาก Hang และคณะ (2009) ที่ใช้ข้อมูลจากผู้ใช้อื่นมาพัฒนาอัลกอริทึมที่ใช้ในการวัดค่าความคล้ายระหว่างผู้ใช้และช่วยในการจับคู่ระหว่างผู้ใช้กับสิ่งที่สนใจ โดยมีหลักการว่าผู้เริ่มต้นใช้ระบบจะจับคู่กับสิ่งที่สนใจที่มีค่าสำคัญมากที่สุด (จาก

ผู้ใช้ที่มีความคล้ายคลึงกัน) ซึ่งวิธีการนี้ให้ผลที่น่าพอใจกับระบบที่เน้นไปที่การหาความสนใจรายบุคคลมากกว่าวิธีการแรก

อย่างไรก็ตามยังมีวิธีการที่น่าสนใจที่จะนำมาแก้ไขปัญหานี้อีกวิธีหนึ่งคือการนำข้อมูลมาจากเครือข่ายสังคมออนไลน์ของผู้ใช้ ซึ่งวิธีการนี้สามารถใช้ประโยชน์จากทั้งข้อมูลของผู้ใช้เองและผู้ใช้อื่นๆ รวมถึงไม่เป็นการรบกวนผู้ใช้มากเกินไปในการขอข้อมูลต่างๆ เพราะเพียงผู้ใช้ทำการอนุญาตให้ระบบเข้าถึงข้อมูล ระบบจะทำการดึงข้อมูลในเครือข่ายเหล่านี้โดยไม่ต้องให้ผู้ใช้กรอกข้อมูลไปยังระบบเอง โดยการนำข้อมูลลักษณะต่างๆ ในเครือข่ายสังคมออนไลน์ มาใช้ประโยชน์ในการค้นหาความสนใจนั้นจะอธิบายในหัวข้อที่ 2.6

2) การเรียนรู้ความสนใจของผู้ใช้ (User Interest Learning)

การเรียนรู้ความสนใจของผู้ใช้เป็นอีกหนึ่งหัวข้อหลักของแบบจำลองผู้ใช้ที่นักวิจัยพยายามพัฒนาเพื่อให้รองรับการเปลี่ยนแปลงความสนใจของผู้ใช้ที่สามารถเกิดขึ้นได้ตลอดเวลาเพื่อให้ระบบแนะนำข้อมูลสามารถแนะนำข้อมูลได้เหมาะสมกับความสนใจของผู้ใช้ในขณะนั้นมากที่สุด โดย Widyantoro และคณะ (1999) ได้เสนออัลกอริทึมในการเรียนรู้ความสนใจของผู้ใช้โดยแบ่งความสนใจของผู้ใช้เป็นแบบระยะสั้นเพื่อเรียนรู้ความสนใจในช่วงเวลาสั้นๆ (Short term) ขณะปัจจุบันและแบบระยะยาว (Long term) เพื่อเรียนรู้ความสนใจโดยรวม จากนั้นจะคำนวณค่าน้ำหนักไปที่แต่ละแบบและเลือกใช้ความสนใจที่มีค่าน้ำหนักมากที่สุด วิธีนี้สามารถเพิ่มความถูกต้องในการค้นหาความสนใจที่เปลี่ยนแปลงไปตามช่วงเวลาได้ และสามารถปรับปรุงความสนใจผู้ใช้ให้ถูกต้องขึ้น สำหรับการเปลี่ยนแปลงความสนใจของผู้ใช้แบบทันทียังไม่ถูกนำมากล่าวถึงมากนัก Zhu และคณะ (2012) ได้เสนอการเรียนรู้ความสนใจของผู้ใช้โดยนำข้อมูลโดยอ้อม เช่น พฤติกรรมการคลิก พฤติกรรมการโหวต เป็นต้น จากนั้นคำนวณหาค่าความสนใจในแต่ละเรื่องร่วมกับการคำนวณด้านเวลา (เช่นการกดโหวตเมื่อวานนี้ย่อมมีนัยสำคัญเกี่ยวกับความสนใจมากกว่าการโหวตเมื่อปีที่แล้ว) ค่าความสนใจทั้งหมดของผู้ใช้จะถูกเก็บไว้ในเมทริกซ์และเปลี่ยนแปลงค่าด้วยตัวเองเมื่อผู้ใช้เปลี่ยนแปลงพฤติกรรมต่างๆ ข้างต้น ค่าความสนใจที่ได้จากวิธีการนี้สามารถนำไปใช้แนะนำผลิตภัณฑ์หรือบริการได้ตามลำดับค่าความสนใจ นอกจากนี้การเปลี่ยนแปลงค่าความสนใจตามพฤติกรรมในปัจจุบันยังช่วยรองรับการเปลี่ยนแปลงความสนใจในแต่ละช่วงเวลาได้ดียิ่งขึ้น Mezhoudi (2013) ใช้ประโยชน์จากข้อมูลตอบกลับของผู้ใช้ (User Feedback) ปรับปรุงแบบจำลองผู้ใช้ให้กับระบบโดยใช้ข้อมูลตอบกลับจากการปฏิสัมพันธ์ของผู้ใช้ใน 2 แบบคือ 1) Emoticons Based Feedback ซึ่งเป็นการแสดงระดับความพึงพอใจของผู้ใช้ผ่านทางไอคอนแสดงอารมณ์ ตัวอย่างการนำไปประยุกต์ใช้ เช่น การตอบคำถามข้อเสนอแนะด้วยไอคอนแสดงอารมณ์ในแอพพลิเคชั่นบนโทรศัพท์มือถือ (Arhippainen et al., 2004) 2) Recommendation Frames เป็นการเก็บข้อมูลจากการปฏิสัมพันธ์โดยการแนะนำ การแนะนำอาจจะแสดงเป็นหน้าต่างป็อปอัพหรือแสดงบนหน้าเพ็จ ซึ่งใช้กันมากในพาณิชอิเล็กทรอนิคส์ (E-Commerce) เพื่อให้คำแนะนำแก่ลูกค้า การนำข้อมูลตอบกลับลักษณะดังกล่าวมาใช้ในการเรียนรู้ความสนใจของผู้ใช้เป็นส่วนเสริมในการประเมินความถูกต้องหรือพึงพอใจในการให้คำแนะนำของระบบและยังทำให้ระบบนำข้อมูลเหล่านี้ไปปรับเปลี่ยนความสนใจของผู้ใช้ให้ได้ทันท่วงทีอีกด้วย

จากกระบวนการค้นหาความสนใจของผู้ใช้ทั้งสองส่วนข้างต้น การเลือกใช้วิธีการได้มาซึ่งข้อมูลความสนใจของผู้ใช้โดยตรงนั้น สามารถลดปัญหาโคลด์สตาร์ทที่อาจเกิดขึ้นเมื่อใช้วิธีการได้มาซึ่งข้อมูลความสนใจของ

ผู้ใช้โดยอ้อม อย่างไรก็ตามการขอข้อมูลจากผู้ใช้โดยตรงบางวิธีจะรบกวนผู้ใช้มากเกินไปเพราะผู้ใช้ต้องเสียเวลากับขั้นตอนการกรอกข้อมูลเหล่านี้ ดังนั้นการผสมผสานทั้งสองวิธีการเข้าด้วยกันจึงเป็นอีกทางเลือกหนึ่งที่สามารถลดปัญหาที่เกิดขึ้นจากวิธีการทั้งสองได้ ส่วนการเรียนรู้ความสนใจของผู้ใช้ ซึ่งมีทั้งแนวคิดการแบ่งความสนใจออกเป็นระยะสั้นและระยะยาว การจับพฤติกรรมของผู้ใช้เพื่อปรับเปลี่ยนลักษณะความสนใจ และการใช้ข้อมูลตอบกลับ โดยการใช้ข้อมูลตอบกลับเป็นวิธีการที่มีประสิทธิภาพมากกว่าอีกสองวิธีที่ได้กล่าวไว้ ซึ่งนอกจากจะปรับปรุงความสนใจได้แล้วยังสามารถนำมาใช้ประเมินความถูกต้องหรือพึงพอใจในการให้คำแนะนำของระบบอีกด้วย

2.4 การสร้างแบบจำลองความสนใจของผู้ใช้ (User Interest Modeling)

การสร้างแบบจำลองผู้ใช้เป็นกระบวนการหลักของระบบเช่นกัน โดยหลังจากที่ระบบได้วิเคราะห์และเรียนรู้ความสนใจจากผู้ใช้แล้ว รูปแบบของการจัดเก็บความสนใจเหล่านี้ก็เป็นสิ่งที่มีความสำคัญ ซึ่งรูปแบบของการจัดเก็บข้อมูลที่นิยมใช้ในปัจจุบัน ดังเช่น

1) Flat File Database Model

แบบจำลองฐานข้อมูลแบบ Flat File หรือเรียกอีกอย่างหนึ่งว่า System File เป็นการจัดเก็บข้อมูลด้วยข้อความ (Plain Text) ซึ่งโดยมากจะใช้เครื่องหมายจุลภาค (Comma) การเว้นวรรค (Space or Tab) เพื่อแยกข้อมูลแต่ละคอลัมน์ออกจากกัน และใช้การขึ้นบรรทัดใหม่เพื่อแบ่งข้อมูลในแต่ละแถว ดังรูปที่ 3

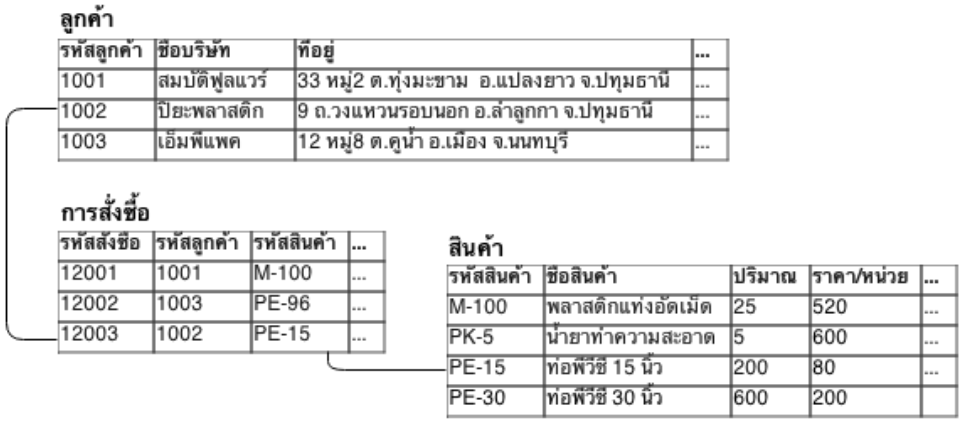
ข้อจำกัดของการจัดเก็บไฟล์ลักษณะนี้คือ จะไม่สามารถสร้างความสัมพันธ์ระหว่างข้อมูลแต่ละแถวได้ ไม่รองรับการเข้าใช้งานข้อมูลเดียวในเวลาเดียวกัน อาจเกิดความซ้ำซ้อนกันของข้อมูล ทำการค้นหาข้อมูลได้ยาก และแม้จะมีการจัดเก็บข้อมูลเป็นชุดข้อมูล แต่ก็ไม่สามารถนำมาประมวลผลจัดเรียงข้อมูลได้ อย่างไรก็ตาม Flat File เหมาะกับการจัดเก็บข้อมูลบางประเภทที่ต้องการความเร็วในการดึงข้อมูลแต่ไม่มีการเข้าใช้งานพร้อมกัน ซึ่งจะทำความความเร็วได้มากกว่าการจัดเก็บแบบจำลองฐานข้อมูลข้อมูลสัมพันธ์

รหัส	ชื่อ	เพศ
1	มงคล	ชาย
2	นาตยา	หญิง
3	ประภาพรรณ	หญิง
4	ธวัช	ชาย
5	ศิริพร	หญิง

รูปที่ 3 แบบจำลองฐานข้อมูลแบบ Flat File

2) Relational Database Model

แบบจำลองฐานข้อมูลสัมพันธ์ (Relational Database Model ย่อ RDM) เป็นการจัดเก็บข้อมูลในลักษณะที่เป็นกลุ่มของข้อมูล โดยนิยมเก็บเป็นตาราง (Table) ที่มีความสัมพันธ์กัน เพื่อให้ง่ายต่อการเข้าใจและการประยุกต์ใช้งาน ซึ่งในฐานข้อมูลหนึ่งๆ จะมีตารางตั้งแต่ 1 ตารางเป็นต้นไป และในแต่ละตารางนั้นเก็บข้อมูล 2 มิติที่ประกอบด้วยคอลัมน์ (Column) และแถว (Row) แสดงตัวอย่างดังรูปที่ 4



รูปที่ 4 แบบจำลองฐานข้อมูลสัมพันธ์

ข้อดีของการจัดเก็บข้อมูลลักษณะนี้คือสามารถสร้างความสัมพันธ์ระหว่างข้อมูลได้ ลดความซ้ำซ้อนกันของการเก็บข้อมูล ค้นหาหรือสอบถามข้อมูลได้ง่าย นอกจากนี้การใช้ RDBMS (Relational Database Model Management System) ยังช่วยในเรื่องความปลอดภัยในการเข้าถึงข้อมูลโดยการจัดลำดับสิทธิ์ (Permission Level) รองรับการเข้าใช้งานข้อมูลเดียวในเวลาเดียวกัน รวมถึงมีการสำรองข้อมูล (Backup) อย่างไรก็ตามในกลุ่มข้อมูลที่มีความสัมพันธ์จำนวนมาก ซึ่งต้องมีการเชื่อมโยงกันในหลายตาราง ทำให้อาจเสียเวลาในการดึงข้อมูลมากขึ้น

3) Ontology-based Model

ออนโทโลจี (ontology) เป็นเทคนิคหนึ่งในการสร้างองค์ความรู้ (Knowledge-based) โดยถูกนำมาใช้มากขึ้นในหลายสาขาทางวิทยาการคอมพิวเตอร์ เนื่องจากมีประสิทธิภาพในการวิเคราะห์ข้อมูล (Castells et al., 2007) โดยปัจจุบันเริ่มมีการนำออนโทโลจีมาใช้เป็นรูปแบบการจัดเก็บข้อมูล เช่น การจัดเก็บข้อมูลการเดินทางและสถานที่ท่องเที่ยว (นฤพนธ์ และจักรกฤษณ์, 2010) เพื่อให้เกิดความเชื่อมโยงระหว่างข้อมูล ซึ่งการจัดเก็บจะต้องอาศัยการจัดเก็บในภาษาที่มีความสามารถในการอธิบายถึงความหมายและความสัมพันธ์ระหว่างข้อมูล โดยส่วนใหญ่เป็นภาษาที่มีพื้นฐานมาจากภาษาเอ็กซ์เอ็มแอล (XML) เช่น RDF RDF-S OWL เป็นต้น ข้อดีของการจัดเก็บข้อมูลลักษณะนี้คือ ช่วยในเรื่องการเข้าใจความหมายของข้อมูล ลดความกำกวมในความหมายของข้อมูลได้ (ไกรศักดิ์, 2555) แต่ปัญหาสำคัญของออนโทโลจี คือ ยังไม่มีความยืดหยุ่น มีความปลอดภัยต่ำ และการ

เปลี่ยนโครงสร้างของข้อมูลทำได้ยาก ส่งผลให้เมื่อนำไปใช้จัดเก็บข้อมูลที่ต้องมีการเปลี่ยนแปลงจะทำได้ยากกว่าแบบจำลองฐานข้อมูลสัมพันธ์

จากรูปแบบของการจัดเก็บข้อมูลข้างต้น พบว่าการจัดเก็บข้อมูลแบบ Flat File ไม่เหมาะสมที่จะนำมาจัดเก็บข้อมูลแบบจำลองความสนใจของผู้ใช้ เนื่องจากไม่สามารถสร้างความสัมพันธ์ระหว่างข้อมูลได้ การค้นหาและการเปลี่ยนแปลงข้อมูลทำได้ยาก และไม่มีความปลอดภัยในข้อมูล ส่วนการจัดเก็บแบบออนโทโลจี แม้จะได้รับความสนใจมากขึ้นเรื่อยๆ เพราะสามารถอธิบายความหมายของข้อมูลได้ สืบค้นข้อมูลได้ โดยนิยมใช้จัดเก็บข้อมูลที่มีลักษณะเป็นโดเมน (Domain) เช่น ข้อมูลสถานที่ท่องเที่ยว เป็นต้น แต่การนำมาใช้ในการจัดเก็บข้อมูลที่ต้องมีการเปลี่ยนแปลง เช่น ความสนใจของผู้ใช้ที่ต้องมีการปรับเปลี่ยนบ่อยครั้ง การปรับเปลี่ยนจะทำได้ยาก และยังไม่มีระบบการจัดการ (Management System) มากนัก ทำให้อาจมีปัญหาเรื่องความปลอดภัยในข้อมูลได้ด้วย ส่วนการจัดเก็บแบบฐานข้อมูลสัมพันธ์นั้นเป็นรูปแบบการจัดเก็บที่ได้รับความนิยมในปัจจุบันรวมถึงมีความเหมาะสมในการจัดเก็บข้อมูลความสนใจของผู้ใช้ เนื่องจากสามารถเก็บข้อมูลที่มีความสัมพันธ์กัน สามารถสร้าง ค้นหา และเปลี่ยนแปลงข้อมูลได้ง่าย รวมถึงมีระบบการจัดการหลายระบบให้เลือกใช้ซึ่งสามารถเพิ่มความปลอดภัยในข้อมูลได้ โดยผู้วิจัยได้สรุปปัญหาหลักที่เกิด ดังตารางที่ 3

ตารางที่ 3 ข้อดีของการสร้างแบบจำลองความสนใจของผู้ใช้ด้วยรูปแบบการจัดเก็บแบบต่างๆ

ข้อดี	การสร้างแบบจำลองความสนใจของผู้ใช้ (User Interest Modeling)		
	Flat File	Relational	Ontology-based
1. สามารถสร้างความสัมพันธ์ระหว่างข้อมูลได้	✗	✓	✓
2. การค้นหาข้อมูลทำได้ง่าย	✗	✓	✓
3. การเปลี่ยนแปลงข้อมูลทำได้ง่าย	✗	✓	✗
4. มีความปลอดภัยของข้อมูล	✗	✓	✗

2.5 การแนะนำข้อมูล (Recommendation Techniques)

การแนะนำข้อมูลเป็นกระบวนการที่ช่วยค้นหาสินค้าหรือบริการที่ผู้ใช้กำลังสนใจ เปรียบกับการจับคู่สิ่งที่เหมาะสมกับความต้องการของผู้ใช้นั่นเอง ซึ่งเทคนิคที่ใช้ในการแนะนำข้อมูลมีอยู่หลายวิธีการ เช่น Content-based, Collaborative, Demographic, Knowledge-based และวิธีแบบผสม (Hybrid) เป็นต้น แต่วิธีที่นิยมนำมาใช้มากที่สุดคือวิธีการ Content-based และ Collaborative โดยวิธีการแบบ Content-based นั้นเป็นการแนะนำโดยการวิเคราะห์จากพฤติกรรมในอดีตของผู้ใช้เอง ในขณะที่วิธีการแบบ Collaborative เป็นการแนะนำโดยใช้ข้อมูลจากคำแนะนำที่ให้กับผู้ใช้อื่น ซึ่งวิธีการทั้งสองอธิบายเพิ่มเติมดังนี้

1) วิธีการคัดกรองความสนใจแบบ Content-based (Content-based Filtering Technique)

วิธีการแบบ Content-based นั้นจะแนะนำสินค้าหรือบริการที่มีลักษณะตามที่ผู้ใช้มีความต้องการหรือสนใจ โดยผู้ใช้อาจจะให้ข้อมูลความสนใจโดยตรงหรืออาจวิเคราะห์จากพฤติกรรมในอดีตของผู้ใช้ เช่น การซื้อสินค้าหรือบริการหรือการค้นหาข้อมูลในอดีต เป็นต้น ซึ่งระบบที่นำวิธีการนี้มาพัฒนาคือ ระบบ TripleHop's TripMatcher2 และ VacationCoachs ที่นำมาใช้โดยเว็บไซต์ Travelocity.com (Kabassi, 2010) โดยทั้งสองระบบใช้วิธีการแบบ content-based ที่ให้ผู้ใช้ใส่ข้อมูลความต้องการหรือความสนใจ จากนั้นระบบจะทำการจับคู่ความต้องการของผู้ใช้กับสินค้าที่อยู่ในรายการ ทั้งนี้จากการศึกษาพบว่าวิธีการจับคู่ความต้องการนี้ สามารถทำได้หลายรูปแบบ เช่น การเปรียบเทียบอย่างง่ายระหว่างคุณลักษณะของสินค้าและบริการกับลักษณะความสนใจของผู้ใช้ ลักษณะสินค้าที่ตรงกับความสนใจของผู้ใช้มากที่สุดก็จะถูกแนะนำ หรือระบบได้ใช้หลักของ rule-based reasoning หรือ Bayesian networks เพื่อที่จะปรับปรุงประสิทธิภาพในการแนะนำของระบบ เป็นต้น ข้อดีของวิธีการนี้คือ ระบบจะวิเคราะห์ข้อมูลตามที่เกิดขึ้นจริงของผู้ใช้แต่ละคน นั่นคือระบบสามารถแนะนำและจับการเปลี่ยนแปลงความต้องการของผู้ใช้แต่ละคนได้ มีเพียงบางกรณีที่วิธีการนี้อาจจะแนะนำผิดพลาด เช่น ผู้ใช้ค้นหาข้อมูลสำหรับเป็นของขวัญให้กับผู้อื่น ซึ่งระบบคิดว่าเป็นความสนใจและความต้องการของผู้ใช้เอง ทำให้ระบบอาจจะแนะนำสินค้าหรือบริการผิดพลาดได้เมื่อผู้ใช้เข้ามาใช้อีกครั้ง อย่างไรก็ตามปัญหานี้ไม่ค่อยเกิดขึ้นบ่อยนักในระบบแนะนำข้อมูลการท่องเที่ยว เพราะผู้ใช้ส่วนใหญ่มักจะซื้อสินค้าหรือบริการที่ต้องการสำหรับตัวเอง ไม่ใช่สำหรับเป็นของขวัญ ที่โดยมากจะเกิดขึ้นในระบบแนะนำของพาณิชย์อิเล็กทรอนิกส์ประเภทหนังสือหรือภาพยนตร์ (Montaner et al., 2003) อีกความท้าทายหนึ่งของวิธี content-based คือการปฏิสัมพันธ์ครั้งแรกของผู้ใช้กับระบบ ระบบอาจจะไม่มีข้อมูลหรือมีข้อมูลไม่มากพอเกี่ยวกับความต้องการของผู้ใช้นั้น (ปัญหาโคลด์สตาร์ท) ทำให้การแนะนำสินค้าหรือบริการไม่มีสิทธิภาพมากพอ

2) วิธีการคัดกรองความสนใจแบบ Collaborative (Collaborative Filtering Technique)

วิธีการแบบ Collaborative นั้นจะค้นหาความสนใจของผู้ใช้ที่มีต่อสินค้าหรือบริการ และจับคู่ผู้ใช้ไปยังผู้ใช้อื่น ๆ ที่มีความสนใจคล้ายกัน และแนะนำสินค้าหรือบริการที่เคยแนะนำให้กับผู้ใช้อื่นที่มีความสนใจคล้ายกันนั่นเอง ดังนั้นทุกครั้งที่ผู้ใช้ใหม่เข้าสู่ระบบ ระบบจะพยายามเชื่อมต่อสิ่งที่ผู้ใช้และผู้ใช้อื่นน่าจะมีความสนใจเหมือนกัน ซึ่งจะเรียกผู้ใช้เหล่านี้ว่า 'เพื่อนบ้าน' (Neighbors) ระบบที่ใช้วิธีการนี้ ตัวอย่างเช่น MAIS Project (Corallo et al., 2006) รวมถึง TripleHops (Kabassi, 2010) ที่ใช้วิธีการทางสถิติคำนวณหาผู้ใช้ที่คล้ายกันจากคิวรี่ (Query) ของผู้ใช้ วิธีการแบบ Collaborative มีจะมีข้อดีคือ เมื่อระบบได้พบ 'เพื่อนบ้าน' ของผู้ใช้ ระบบจะสามารถแนะนำสินค้าและบริการที่เป็นส่วนบุคคลได้ อย่างไรก็ตามความท้าทายของวิธีการนี้คือการหาเพื่อนบ้านของผู้ใช้ในกรณีที่ผู้ใช้ยังไม่มีปฏิสัมพันธ์อย่างเพียงพอกับระบบ รวมถึงการหา 'เพื่อนบ้านที่ดี' ของผู้ใช้ด้วยเพราะเพื่อนบ้านที่ได้มาอาจจะให้ข้อมูลที่ขัดแย้งและทำให้เกิดคำแนะนำที่แตกต่างกัน และผู้ใช้อาจจะมีลักษณะความสนใจพิเศษเมื่อเทียบกับเพื่อนบ้านอื่นๆ ทำให้ระบบไม่ได้แนะนำข้อมูลที่ควรจะต้องแนะนำสำหรับผู้ใช้ไป นอกจากนี้ยังมีข้อจำกัดอื่นอีก คือ ปัญหา 'new item' ซึ่งเกิดขึ้นในกรณีที่สินค้าหรือบริการถูกเพิ่มเข้าไปในระบบโดยยังไม่ได้มีการให้คะแนน (Rating) มากพอ จะทำให้ระบบไม่สามารถแนะนำสินค้าหรือบริการนี้ให้กับผู้ใช้ได้ ในทางกลับกันวิธีการแบบ Content-based ที่ไม่ได้ใช้การจัดอันดับของผู้ใช้อื่น ก็จะไม่พบกับปัญหานี้

3) วิธีการคัดกรองความสนใจแบบผสมผสาน (Hybrid Technique)

วิธีการแบบผสมผสาน เป็นวิธีที่นักวิจัยหลายท่านได้เสนอขึ้นเพื่อลดข้อจำกัดของ ทั้งวิธีการแบบ Content-based และ Collaborative วิธีการแรก คือการนำวิธี Content-based และ Collaborative รวมเข้าด้วยกัน เพื่อที่จะใช้ประโยชน์จากข้อได้เปรียบของและลดข้อบกพร่อง ตัวอย่างเช่น Pazzani (1999) เสนอกรอบการทำงานที่ผสมผสานระหว่าง Content-based และ Collaborative โดยใช้ Content-based สำหรับการสร้างโปรไฟล์ผู้ใช้และใช้ Collaborative คำนวณความคล้ายคลึงกันระหว่างความสนใจของผู้ใช้ Schiaffino และ Amandi (2009) นำเสนอซอฟท์แวร์เอเจนต์ชื่อว่า Traveller เอเจนต์นี้จะใช้การผสมผสานทั้งสองวิธีข้างต้นและใช้ข้อมูลทางประชากรเพื่อแนะนำแพคเกจท่องเที่ยว นอกจากนี้ยังมีการนำวิธีการ knowledge-based มาผสมผสานเช่น ระบบแนะนำร้านอาหารที่เสนอโดย Burke (2000) โดยแนะนำร้านอาหารในเมืองอื่นที่ผู้ใช้จะเดินทางไป จากประเภทร้านอาหารที่คล้ายคลึงกันกับที่ผู้ใช้ชอบและเคยไป เพื่อสรุปความคล้ายคลึงกันระหว่างร้านอาหาร ซึ่งช่วยแก้ปัญหาเรื่องความสนใจเฉพาะของผู้ใช้ได้ดีกว่าที่จะใช้วิธีการแบบ Collaborative อย่างเดียว รวมถึง Maw และคณะ (2006) ได้ใช้ Rule-based ร่วมกับ collaborative ในการแนะนำข้อมูลและปรับปรุงการแนะนำให้เหมาะกับความสนใจและความต้องการของผู้ใช้ ซึ่งสามารถลดข้อจำกัดของการใช้วิธีแบบ collaborative เพียงอย่างเดียวได้

จากเทคนิคในการแนะนำข้อมูลข้างต้น ปัญหาหลักที่เกิดขึ้นคือ การใช้เทคนิคแบบ Content-based ในกรณีผู้ใช้ใช้งานระบบครั้งแรก แล้วระบบไม่มีข้อมูลความสนใจของผู้ใช้ จะทำให้ระบบไม่สามารถแนะนำข้อมูลที่ถูกต้องให้กับผู้ใช้ได้ (เกิดปัญหาโคลด์สตาร์ท) แต่เทคนิคแบบ Collaborative ระบบจะพยายามเชื่อมต่อสิ่งที่ผู้ใช้และผู้ใช้อื่นน่าจะมีความสนใจเหมือนกัน และแนะนำสิ่งที่เคยแนะนำให้ผู้ใช้อื่นให้กับผู้ใช้ ทำให้ลดปัญหาโคลด์สตาร์ทไปได้ อย่างไรก็ตามวิธีการแบบ Collaborative มีข้อจำกัดคือในความเป็นจริงผู้ใช้แต่ละคน อาจจะมีความสนใจพิเศษเฉพาะ ทำให้ระบบอาจจะไม่ได้แนะนำข้อมูลที่ควรจะต้องแนะนำให้กับผู้ใช้ ซึ่งตรงกันข้ามกับแบบ Content-based ที่จะไม่เกิดปัญหานี้ เนื่องจากใช้ข้อมูลจากผู้ใช้เองเป็นหลัก โดยสรุปปัญหาในการเลือกใช้เทคนิคการแนะนำข้อมูล ดังตารางที่ 4

ตารางที่ 4 ข้อดีในการเลือกใช้เทคนิคการแนะนำข้อมูลประเภทต่างๆ

ข้อดี	เทคนิคการแนะนำข้อมูลให้ผู้ใช้ (Recommendation Techniques)		
	Content-based	Collaborative	Hybrid
1. กรณีที่เกิดปัญหาโคลด์สตาร์ท ระบบยังสามารถแนะนำข้อมูลได้	✗	✓	✓
2. ระบบสามารถแนะนำข้อมูลที่เฉพาะเจาะจง กรณีที่ผู้ใช้มีความสนใจพิเศษเฉพาะบุคคล	✓	✗	✓

2.6 เครือข่ายสังคมออนไลน์กับการนำมาใช้ในระบบแนะนำข้อมูลเฉพาะบุคคล

ข้อมูลในเครือข่ายสังคมออนไลน์ได้นำมาใช้ค้นหาความสนใจของผู้ใช้ในระบบแนะนำข้อมูลต่างๆ หลายวิธี โดยประโยชน์หลักของเครือข่ายสังคมออนไลน์คือสามารถนำไปสู่ข้อสรุปของความสนใจของผู้ใช้ได้ ซึ่งในหัวข้อนี้จะกล่าวถึงลักษณะของเครือข่ายสังคมออนไลน์ประเภทต่างๆ และเทคนิคที่นักวิจัยนำเครือข่ายสังคมออนไลน์มาใช้เพื่อค้นหาความสนใจของผู้ใช้

2.6.1 เครือข่ายสังคมออนไลน์

เครือข่ายสังคมออนไลน์ มีการเจริญเติบโตเป็นอย่างมากในไม่กี่ปีผ่านมา เนื่องจากได้กลายเป็นเครื่องมือสื่อสารสำคัญของบุคคลในปัจจุบัน เครือข่ายเหล่านี้ให้พื้นที่ในการติดต่อสื่อสารการสร้างและแบ่งปันข้อมูลระหว่างเพื่อนและบุคคลอื่นๆ ทำให้เกิดข้อมูลจำนวนมากมายบนเครือข่ายเหล่านี้ เช่น ข้อมูลความคิดเห็น ข่าว ภาพ วีดิโอ แม้กระทั่งบทวิจารณ์ในผลิตภัณฑ์และบริการ เป็นต้น ทำให้เกิดเว็บไซต์เครือข่ายสังคมออนไลน์มากมาย เช่น เฟสบุ๊ค (Facebook.com) ทวิตเตอร์ (Twitter.com) อินสตาแกรม (Instargram.com) โฟร์สแควร์ (Foursquare.com) ฟลิคเกอร์ (flickr.com) ทริปแอดไวเซอร์ (Tripadvisor.com) เป็นต้น โดยลักษณะที่สำคัญของเครือข่ายสังคมออนไลน์เหล่านี้มีดังนี้

1) เฟสบุ๊ค (Facebook.com) เป็นเครือข่ายสังคมออนไลน์ที่นิยมมากที่สุดในประเทศไทยในปัจจุบัน เฟสบุ๊คเป็นเครื่องมือที่ใช้ติดต่อสื่อสารกันระหว่างบุคคลหรือกลุ่มบุคคลโดยผ่านทั้งการแชท การส่งข้อความ การแสดงความคิดเห็นในหน้าเฟสบุ๊ค (Timeline) การกดชื่นชอบ การแชร์รูปภาพ ลิงค์และวีดิโอ การเช็คอิน การเชิญชวนให้มาร่วมกิจกรรม เป็นต้น นอกจากนี้ยังได้มีการใช้งานส่วนอื่นๆ อีกมากมาย โดยที่น่าสนใจ เช่น การสร้างกลุ่ม สร้างหน้าเพจ และเกม เป็นต้น (Hogan, 2011)

2) ทวิตเตอร์ (Twitter.com) เป็นหนึ่งในเครือข่ายสังคมออนไลน์ที่ได้รับความนิยมในประเทศไทยเช่นกัน โดยในทวิตเตอร์สามารถให้ผู้ใช้รายงานข่าว แสดงความคิดเห็น ฯลฯ ผ่านการโพสต์ข้อความสั้นไม่เกิน 140 ตัวอักษรหรือเรียกว่า ทวีต (Tweet) (Barash and Golder, 2011) ซึ่งในทวีตนอกจากข้อความแล้วยังสามารถสร้างแท็กและใส่ลิงค์ได้ และผู้ใช้สามารถติดตามข้อความจากผู้ใช้อื่นๆ ที่สนใจโดยการกดติดตาม (Follow)

3) อินสตาแกรม (Instargram.com) เป็นเครือข่ายสังคมออนไลน์แบบ Social Photo Sharing คือเป็นเครือข่ายสังคมออนไลน์ของการแบ่งปันภาพถ่าย แต่ต่อมาได้พัฒนาให้แบ่งปันวิดีโอสั้นๆ ได้ โดยผู้ใช้สามารถที่จะใส่ข้อความหรือแท็ก เพื่ออธิบายภาพหรือวิดีโอเหล่านั้น รวมถึงสามารถแสดงความคิดเห็นและกดชื่นชอบ นอกจากนี้ยังมีระบบติดตาม (Follow) ให้สามารถติดตามชมรูปภาพ ความเคลื่อนไหวการใช้งานของผู้ใช้อื่น เช่นเดียวกับทวิตเตอร์ และสามารถแชร์รูปภาพและวิดีโอไปยังเครือข่ายอื่นๆ ได้ เช่น เฟสบุ๊ค ทวิตเตอร์ เป็นต้น

4) โฟร์สแควร์ (Foursquare.com) เป็นเครือข่ายสังคมออนไลน์แบบ “Location-Based” คือการผสมผสานกันระหว่างเครือข่ายสังคมออนไลน์และสถานที่ ฟังก์ชันหลักของโฟร์สแควร์คือการเช็คอิน (Check-in) เพื่อบอกว่าว่าผู้ใช้อยู่ที่ไหน และแชร์ข้อมูลนี้ไปยังเพื่อนๆ นอกจากนี้ยังสามารถแชร์การเช็คอินนี้ไปยังเครือข่ายที่ได้รับความนิยมอย่างเฟสบุ๊คและทวิตเตอร์ได้อีกด้วย

5) ฟลิคเกอร์ (flickr.com) เป็นเครือข่ายสังคมออนไลน์ช่วยให้ผู้ใช้อัพโหลด จัดหมวดหมู่หรือการจัดการต่างๆเกี่ยวกับภาพถ่ายและวิดีโอ โดยผู้ใช้สามารถใส่ข้อความ แท็ก เพื่ออธิบายภาพหรือวิดีโอเหล่านั้น และทำการแชร์ไปยังเพื่อนๆ ได้ นอกจากนี้ฟลิคเกอร์ยังเป็นเครื่องมือหนึ่งสำหรับผู้ที่ชื่นชอบการถ่ายภาพที่สามารถแสดงความคิดสร้างสรรค์ผ่านภาพที่โพสต์ได้อีกด้วย (Rodrigues and Milic-Frayling, 2011)

6) ทริปแอดไวเซอร์ (Tripadvisor.com) เป็นเว็บไซต์ท่องเที่ยวที่ใหญ่ที่สุดในโลกแห่งหนึ่ง เป็นชุมชนที่ให้ข้อมูลการท่องเที่ยว เปิดให้ผู้ใช้สามารถวางแผนการเดินทาง รวมถึงส่วนเสริมอื่นๆ เช่น การค้นหาเที่ยวบิน การจองที่พัก นอกจากนี้ยังมีข้อมูลการให้คะแนนและคำแนะนำจากผู้ใช้อื่นที่เคยเดินทางไปสถานที่นั้นๆ ด้วย

จากการศึกษาข้อมูลเครือข่ายสังคมที่น่าสนใจทั้ง 6 เครือข่าย พบว่าทั้งหมดรองรับการให้ข้อมูลแบบเฉพาะบุคคล ยกเว้นเพียงทริปแอดไวเซอร์ ที่แม้ว่าจะเป็นเครือข่ายสังคมที่เกี่ยวกับการท่องเที่ยวโดยตรงที่น่าจะนำมาค้นหาความสนใจในการท่องเที่ยวได้ แต่เซอร์วิสที่ให้บริการข้อมูลไม่สามารถให้ข้อมูลเฉพาะรายบุคคลได้ จะใช้ได้เพียงข้อมูลโดยรวม เช่น ข้อมูลความคิดเห็นในสถานที่ท่องเที่ยวของผู้ใช้ทั่วไป เป็นต้น นอกจากนี้ยังมีขั้นตอนที่ค่อนข้างยุ่งยากในการขอข้อมูล ทริปแอดไวเซอร์จึงไม่ใช่ทางเลือกที่ดีนักในการพัฒนาระบบแบบเฉพาะบุคคล เครือข่ายสังคมที่รองรับการให้ข้อมูลแบบเฉพาะบุคคลอีก 5 เครือข่ายที่เหลือมีบริการ (Service) ที่ให้ข้อมูลแบบเฉพาะบุคคลและมีขั้นตอนการขอไม่ยุ่งยาก ผู้วิจัยจึงได้ศึกษาประเภทของข้อมูลที่มีการเผยแพร่ในแต่ละเครือข่าย พบว่า มี 3 ประเภทข้อมูลที่มีความเป็นไปได้ในการนำมาค้นหาความสนใจแบบเฉพาะบุคคล ได้แก่

1) ข้อความ (Message) คือข้อความต่างๆที่เกิดขึ้นบนเครือข่ายสังคม เช่น ข้อความในการแสดงความคิดเห็น (Comment) ข้อความสถานะต่างๆ (Status Message) หรือทวีต (Tweet) เป็นต้น ตัวอย่างข้อมูลประเภทข้อความแสดงดังรูปที่ 5

รูปที่ 5 ตัวอย่างข้อมูลประเภทข้อความจากทวิตเตอร์

2) แท็ก (Tag) คือข้อความที่ใช้จัดกลุ่มข้อมูลบนเครือข่ายสังคม เช่น การแท็กบนโซเชียลบุ๊กมาร์ก การติดแฮชแท็ก (Hashtag) บนเฟสบุ๊คและทวิตเตอร์ เป็นต้น ตัวอย่างข้อมูลประเภทแท็กแสดงดังรูปที่ 6

รูปที่ 6 ตัวอย่างข้อมูลประเภทแท็กจากอินสตาแกรม

3) เช็คอิน (Check-in) คือการแชร์ข้อมูลของสถานที่บนเครือข่ายสังคม เพื่อบอกว่าผู้ใช้อยู่ที่ไหน เช่น การเช็คอินบนเฟสบุ๊ค โฟร์สแควร์ เป็นต้น ตัวอย่างข้อมูลประเภทเช็คอินแสดงดังรูปที่ 7

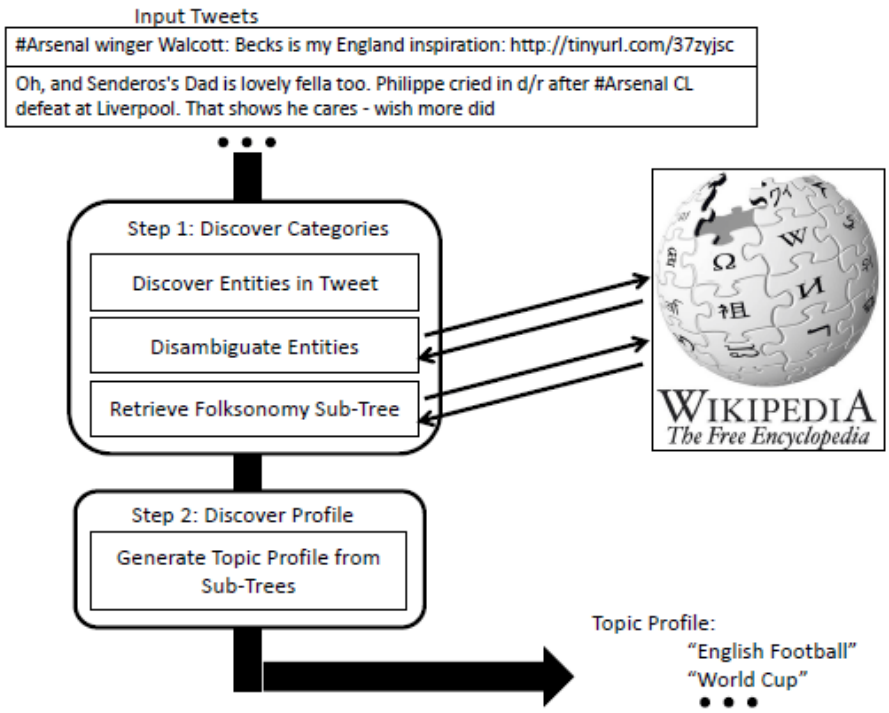
รูปที่ 7 ตัวอย่างข้อมูลประเภทเช็คอินจากเฟสบุ๊ค

จากนั้นได้ศึกษาเทคนิคและวิธีการในการที่จะสามารถสรุปความสนใจของผู้ใช้จากประเภทข้อมูลต่างๆ เหล่านี้ ดังนี้

2.6.2 การค้นหาความสนใจของผู้ใช้โดยใช้ข้อความบนเครือข่ายสังคมออนไลน์

ในปัจจุบันเครือข่ายสังคมออนไลน์ข้อมูลประเภท ภาพ วีดิโอ มีแนวโน้มที่จะถูกใช้มากขึ้น เพราะสามารถที่จะกระตุ้นความสนใจของผู้ใช้อื่นได้ดีกว่าข้อความ อย่างไรก็ตามพบว่าข้อความยังคงถูกนำมาใช้ควบคู่กันเพื่ออธิบายภาพ วีดิโอเหล่านั้น ข้อความจึงเป็นแหล่งข้อมูลที่มีผู้วิจัยนำไปใช้ในการสร้างแบบจำลองผู้ใช้มากกว่า เนื่องจากสามารถเก็บข้อมูลและนำไปประมวลผลได้ง่ายกว่านั้นเอง ทวิตเตอร์เป็นเครือข่ายสังคมออนไลน์ที่ให้ผู้ใช้สามารถรายงานข่าวแสดงความคิดเห็น ฯลฯ โดยการโพสต์ข้อความสั้นไม่เกิน 140 ตัวอักษรหรือที่เรียกกันว่า “ทวีต” จะเห็นว่าในทวิตเตอร์ผู้ใช้ไม่ให้คำสำคัญ (Keyword) แต่เป็นเพียงข้อความสั้นที่ผู้ใช้ทวีตลงไป เช่นเดียวกับเฟสบุ๊คที่เป็นเครือข่ายสังคมออนไลน์ที่ให้ผู้ใช้สามารถโพสต์ข้อความรูปภาพหรือวีดิโอลงบนพื้นที่ของ

ตนเองและเพื่อนๆ หรือที่เรียกว่า “Timeline” และสามารถแสดงความคิดเห็นต่างๆ ได้ทำให้เกิดปัญหาเช่นเดียวกับทวิตเตอร์นั่นคือมีข้อความมากมายเกิดขึ้นแต่ไม่ได้มีการระบุคำสำคัญดังนั้นจึงจำเป็นจะต้องทำการกรองหรือจับคำสำคัญก่อนที่จะใช้เป็นข้อมูลสำหรับสร้างแบบจำลองของผู้ใช้ดังเช่น Castillo และคณะ (2011) พยายามที่จะจำแนกทวีตออกเป็นสองแบบคือ1) ทวีตเกี่ยวกับข่าวและ2) การพูดคุยเพื่อที่จะประเมินความน่าเชื่อถือของข้อมูลบนทวิตเตอร์ Abel และคณะ (2011) ได้ทำการสกัดหัวข้อ (Topic) บนข้อความทวีต เช่น เป็นทวีตของบุคคลเหตุการณ์หรือผลิตภัณฑ์เป็นต้นเพื่อใช้สำหรับการแนะนำบทความข่าวให้กับแต่ละบุคคลดังนั้นการใช้ข้อมูลจากข้อความของทั้งทวิตเตอร์และเฟสบุ๊คเพื่อหาข้อความที่แสดงถึงความสนใจของผู้ใช้จะต้องผ่านวิธีการที่สำคัญคือการกรองและสกัดข้อมูลมากมายเหล่านั้นออกมาเป็นคำสำคัญที่สามารถเชื่อมโยงกับความสนใจได้ดังเช่นการนำคำสำคัญที่สกัดมาได้จับคู่ไปยังฐานความรู้ของเวิร์ดเน็ต (Wordnet) หรือวิกิพีเดีย (Wikipedia) เพื่อตัดคำที่ไม่ได้เกี่ยวข้องกับความสนใจออกไป (Wasim et al., 2011; Lim and Datta, 2013) ซึ่งในรูปที่ 8 เป็นไดอะแกรมที่แสดงตัวอย่างวิธีการค้นหาหัวข้อความสนใจของผู้ใช้จากข้อความบนทวิตเตอร์ที่นำเสนอโดย Michelson และคณะ (2010) โดยจะแบ่งเป็น 2 ขั้นตอนคือ 1) Discover Categories เป็นการค้นหาหมวดหมู่ความสนใจของผู้ใช้ โดยขั้นแรก Discover Entities in Tweet จะเลือกคำ (Entity) จากข้อความทวีตที่ขึ้นต้นด้วยตัวอักษรพิมพ์ใหญ่และไม่เป็นคำที่ไม่สำคัญ (Stop-words) ต่อมาขั้นที่สอง Disambiguate Entities นำคำที่ได้ในขั้นแรกไปหาความหมายที่แท้จริงโดยจับคู่คำไปยังฐานความรู้วิกิพีเดียเพื่อหาความหมายของคำที่เป็นไปได้ทั้งหมด (Candidate Entities) จากนั้นจะเลือกความหมายของคำที่แท้จริงโดยพิจารณาจากคำอื่นๆ ในทวีตว่าปรากฏมากน้อยเพียงใดในหน้าเพ็จวิกิพีเดียของแต่ละคำที่เป็นไปได้ทั้งหมดเมื่อเลือกความหมายของคำได้แล้วสุดท้าย Retrieve Folksonomy Sub-Tree วิกิพีเดียจะส่งหมวดหมู่ของคำนั้นๆ ในรูปแบบโครงสร้างแบบต้นไม้กลับไปยังระบบ2) Discover Profile เป็นการเลือกหัวข้อความสนใจของผู้ใช้โดยใช้ค่าลำดับความสนใจ (Rank) ซึ่งคำนวณจากค่าความถี่ในการทวีตข้อความที่เป็นความสนใจในหมวดหมู่ที่ได้เลือกไว้ในขั้นตอนที่1 คูณกับค่าน้ำหนักของคำในโครงสร้างแบบต้นไม้โดยคำที่ปรากฏในลำดับชั้นสูงจะมีค่าน้ำหนักมากกว่าชั้นที่ต่ำกว่าอย่างไรก็ตามวิธีการนี้ในขั้นตอนการจับคู่คำไปยังฐานความรู้วิกิพิเดียเพื่อหาความหมายของคำยังอาจทำเกิดข้อผิดพลาดได้อยู่ เช่น ระบบแปลความหมาย “Apple” ว่าเป็นผลไม้ชนิดหนึ่งแต่ในความเป็นจริงผู้ใช้ต้องการสื่อถึง“Apple” ที่เป็นแบรนด์สินค้า เป็นต้น สิ่งสามารถก่อให้เกิดการแนะนำข้อมูลที่ไม่ตรงกับความสนใจของผู้ใช้ที่แท้จริงได้นอกจากนี้ผู้วิจัยเล็งเห็นว่า การใช้ข้อมูลที่เป็นลักษณะข้อความในระบบที่จะพัฒนาขึ้น ระบบต้องใช้เวลามากกับการกรองคำในข้อความที่มีจำนวนมาก ทำให้อาจใช้เวลาในการสร้างแบบจำลองตั้งต้นนานเกินไป ผู้ที่เข้ามาใช้งานจริงในระบบคงไม่พึงพอใจนักที่ต้องเสียเวลารอคอย

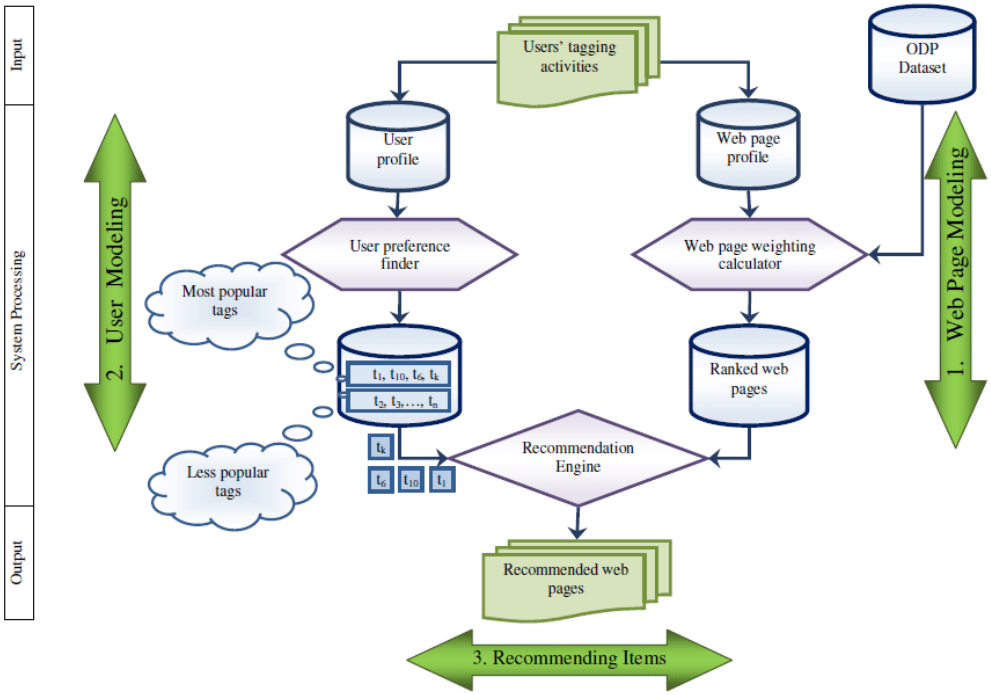


รูปที่ 8 การสกัดข้อมูลความสนใจของผู้ใช้จากข้อความบนทวิตเตอร์โดยใช้ฐานความรู้ของ Wikipedia (Michelson and Macskassy, 2010)

2.6.3 การค้นหาความสนใจของผู้ใช้โดยใช้แท็กบนเครือข่ายสังคมออนไลน์

ในเริ่มต้นจุดประสงค์หลักในการสร้างแท็กคือเพียงเพื่อจัดกลุ่มข้อความที่มีความคล้ายกัน อย่างไรก็ตามปัจจุบันในเครือข่ายสังคมออนไลน์ก็ยังนิยมนำแท็กมาใช้เพื่อแสดงอารมณ์หรือเป็นตัวช่วยเน้นย้ำข้อความให้กับผู้ใช้ เช่น เครือข่ายสังคมออนไลน์ที่กำลังนิยมอย่าง เฟสบุ๊ค (Facebook.com) ทวิตเตอร์ (Twitter.com) เป็นต้น โซเชียลบุ๊กมาร์ก (Social Bookmarking) ซึ่งเป็นเครือข่ายสังคมออนไลน์ประเภทหนึ่งที่นำแท็กมาใช้เช่นเดียวกันแต่เพื่อจุดประสงค์ในการจัดกลุ่มบุ๊กมาร์กของผู้ใช้เอง โดยโซเชียลบุ๊กมาร์กเป็นบริการที่ให้ผู้ใช้สามารถบุ๊กมาร์ก (Bookmarks) เว็บเพ็จต่างๆและแบ่งปันไปยังเพื่อนหรือผู้ใช้อื่นๆได้โดยผู้ใช้แต่ละคนจะทำการบุ๊กมาค์กเว็บเพ็จและสร้างแท็ก (Tag) ติดไปกับบุ๊กมาร์กด้วยคำที่แตกต่างกันทำให้คำที่ใช้สร้างแท็กเหล่านี้ถูกตีความว่าเป็นสิ่งที่ผู้ใช้สนใจนอกจากนี้ในหนึ่งเว็บเพ็จอาจจะถูกแต่ละผู้ใช้บุ๊กมาค์กด้วยแท็กที่ต่างกันทำให้แท็กเหล่านั้นที่ติดไปกับเว็บเพ็จเดียวกันถูกมองว่ามีความสัมพันธ์กันในเชิงความหมายตัวอย่างของการนำแท็กมาใช้ในการสร้างแบบจำลองผู้ใช้เช่น Michlmayr และคณะ (2007) นำเสนอวิธีการ Add-a-Tag ที่ทำการสร้างแบบจำลองผู้ใช้โดยแท็กจากเว็บไซต์โซเชียลบุ๊กมาร์ก Delicious.com ในลักษณะเดียวกัน Firan และคณะ (2007) ได้นำเสนอการสร้างแบบจำลองผู้ใช้แบบ Tags-Based โดยใช้แท็กจากเว็บไซต์ Last.fm และนำแบบจำลองนี้ไปใช้ในการสร้างคำแนะนำในระบบค้นหาเพลงแต่ข้อจำกัดของวิธีนี้คือการใช้ตัวแทนของความสนใจจากผู้ใช้แบบโดยรวมซึ่งคำนวณจากแท็กของผู้ใช้ทั้งหมด (Global Tag) ทำให้อาจจะยังไม่ตอบโจทย์ในการหาความสนใจเป็นรายบุคคลเพื่อใช้ในระบบแนะนำข้อมูลเฉพาะบุคคลมากนักซึ่งหากปรับปรุงโดยใช้แท็กของผู้ใช้เฉพาะบุคคลร่วมด้วยจะได้ข้อมูลความสนใจที่ถูกต้อง

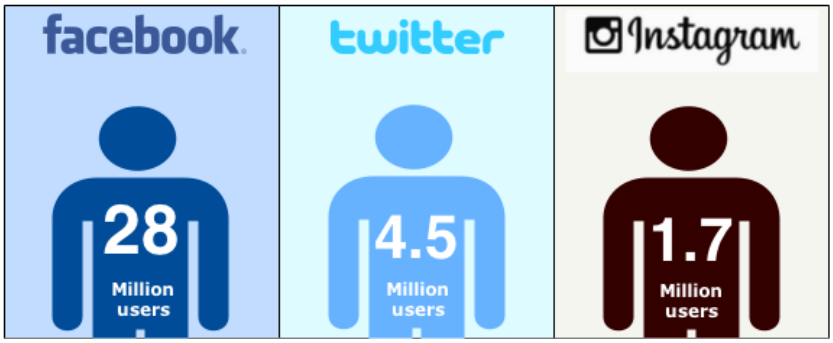
มากยิ่งขึ้นดังเช่น Durao และคณะ (2009) ได้คำนวณหาค่าความสัมพันธ์ระหว่างแท็กกับผู้ใช้โดยใช้ค่าความถี่ที่ผู้ใช้แต่ละคนทำการสร้างแท็กนั้นๆ นั่นคือแท็กที่ผู้ใช้สร้างบ่อยจะมีค่าความสัมพันธ์ที่สูงซึ่งสามารถนำมาลำดับความสนใจของผู้ใช้ได้ในลักษณะเดียวกัน Davoodi และคณะ (2012) นำเสนอระบบแนะนำข้อมูลเว็บเพ็จแบบเฉพาะบุคคลโดยสกัดความสนใจของผู้ใช้จากแท็กในโซเชียลบุ๊กมาร์กซึ่งแสดงภาพรวมของระบบดังรูปที่ 9 โดยระบบแบ่งการทำงานออกเป็น 3 ส่วนคือ1) Web Page Modeling 2) User Modeling 3) Recommending Item ในส่วนแรกคือ Web Page Modeling เป็นการจัดหมวดหมู่และจัดอันดับเว็บเพ็จโดยการจัดหมวดหมู่จะใช้หมวดหมู่ของแท็กที่ผู้ใช้สร้างไปยังเว็บเพ็จนั้นๆ ประกอบกับการใช้ ODP (Open Directory Project) ซึ่งเป็นเว็บไดเรกทอรีในการจัดหมวดหมู่เว็บเพ็จแบบลำดับชั้นการจัดอันดับจะใช้ค่าความถี่ของแท็กที่ทุกผู้ใช้สร้างไปยังหน้าเพ็จนั้นแท็กใดที่มีค่าความถี่สูงก็จะถูกจัดอยู่ในลำดับต้นๆของหมวดหมู่นั้นต่อมาส่วนที่สอง User Modeling เป็นการหาลำดับความสนใจของผู้ใช้โดยใช้ความถี่ของแท็กที่ผู้ใช้นั้นๆสร้างขึ้นแท็กที่ถูกผู้ใช้สร้างมากที่สุดจะเป็นหมวดหมู่ความสนใจที่ผู้ใช้สนใจมากที่สุดส่วนสุดท้าย Recommending Item เป็นการแนะนำเว็บเพ็จไปยังผู้ใช้โดยจะจับคู่แท็กที่ผู้ใช้แต่ละคนสนใจไปยังเว็บเพ็จที่ได้จัดหมวดหมู่และจัดลำดับแล้วในส่วนที่สองอย่างไรก็ตามในประเทศไทยการใช้โซเชียลบุ๊กมาร์กไม่ได้รับความนิยมมากนักทำให้ระบบที่จะพัฒนาขึ้นไม่มีข้อมูลเพื่อสร้างแบบจำลองผู้ใช้แบบเฉพาะบุคคลได้ อีกทั้งเครือข่ายสังคมออนไลน์อื่นที่ได้รับความนิยม ส่วนมากสร้างแท็กเพื่อแสดงอารมณ์หรือเป็นตัวช่วยเน้นย้ำข้อความให้กับผู้ใช้ ซึ่งแม้แท็กจะช่วยกรองคำและลดปริมาณข้อมูลจากการสร้างแบบจำลองผู้ใช้ด้วยข้อความก็ตาม แต่คำที่ได้มาจากแท็กก็ไม่มีคุณภาพมากพอในการค้นหาความสนใจของผู้ใช้ได้



รูปที่ 9 ระบบแนะนำข้อมูลเว็บเพ็จโดยใช้การสกัดความสนใจจากแท็กในโซเชียลบุ๊กมาร์ก (Davoodi and Fatemi, 2012)

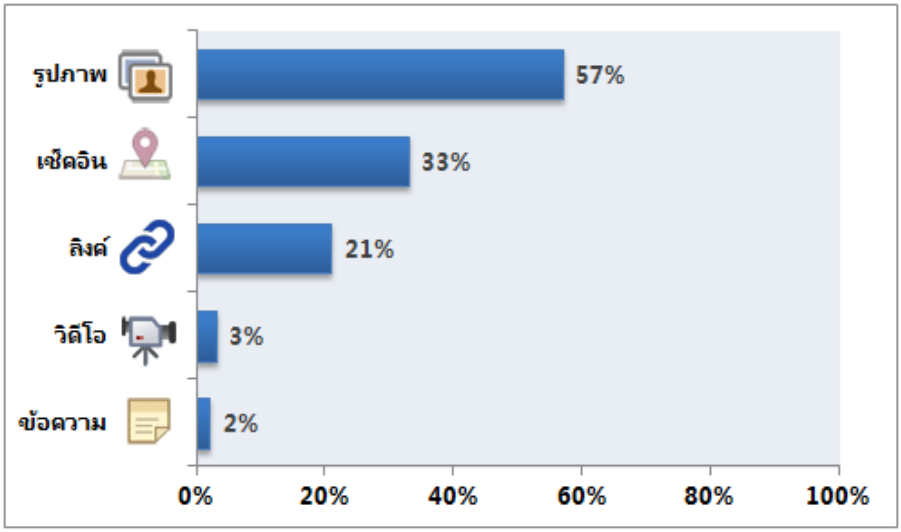
2.6.4 การเลือกใช้เครือข่ายสังคมออนไลน์และประเภทของข้อมูลในการพัฒนาระบบ

เนื่องด้วยระบบที่จะพัฒนาเป็นระบบที่ใช้แนะนำสถานที่ และจากข้อดีข้อเสียของการใช้ข้อมูลประเภทต่างๆ ในเครือข่ายสังคมออนไลน์ ประกอบกับปริมาณของข้อมูลเช็คอินซึ่งน้อยกว่าการใช้ข้อมูลประเภทข้อความหรือแท็ก แต่มีคุณภาพในการบ่งบอกถึงลักษณะความสนใจในสถานที่มากกว่า ผู้วิจัยจึงมีแนวคิดที่จะนำข้อมูลสถานที่ในการเช็คอินต่างๆ มาค้นหาความสนใจซึ่งเกี่ยวข้องกับสถานที่ และศึกษาความเป็นไปได้ที่จะนำข้อมูลประเภทนี้จากเครือข่ายสังคมออนไลน์มาใช้ ซึ่งจากข้อมูลการจัดอันดับเครือข่ายสังคมออนไลน์ที่ได้รับความนิยมในประเทศไทยในปี 2013 โดย ZocialRank (zocialinc.com) 3 อันดับเครือข่ายสังคมออนไลน์ที่มีผู้ใช้มากที่สุด ได้แก่ 1) เฟสบุ๊ค 2) ทวิตเตอร์และ 3) อินสตาแกรม (Instagram) ตามลำดับ ดังรูปที่ 10



รูปที่ 10 เครือข่ายสังคมที่ได้รับความนิยมในประเทศไทยในปี 2013
(Zocial inc. Online Analysis, 2014)

ผู้วิจัยได้ศึกษาประเภทข้อมูลที่นิยมเผยแพร่ในแต่ละเครือข่ายสังคมออนไลน์ พบว่าทวิตเตอร์นิยมทวีตด้วยข้อความที่ติดแท็ก (Hashtag) การเช็คอินไม่ได้ถูกทวีตมากนัก ส่วนอินสตาแกรมกิจกรรมที่ผู้ใช้นิยมคือการโพสรูปภาพและวีดิโอที่ติดแท็ก แม้ปัจจุบันอินสตาแกรมจะเพิ่มฟังก์ชั่นการเช็คอิน ทำให้การโพสภาพพร้อมกับเช็คอินได้รับความนิยมเพิ่มขึ้น แต่ถ้าเปรียบเทียบกับเฟสบุ๊คที่ผู้ใช้นิยมการโพสรูปภาพและการเช็คอินเช่นเดียวกันแล้ว จำนวนผู้ใช้งานในเฟสบุ๊คที่มีมากกว่าอินสตาแกรมกว่า 16 เท่า ทำให้ระบบสามารถรองรับการเข้าถึงของผู้ใช้ได้มากกว่า ดังรูปที่ 11 ผู้วิจัยจึงเลือกที่จะพัฒนาระบบที่ใช้ข้อมูลจากเครือข่ายสังคมออนไลน์อย่างเฟสบุ๊คโดยกรอบแนวคิดในการพัฒนาจะนำเสนอในหัวข้อที่ 3.1



รูปที่ 11 ประเภทข้อมูลที่นิยมเผยแพร่ในเครือข่ายสังคมออนไลน์เฟสบุ๊ค

2.7 เทคนิคตรวจสอบการเปลี่ยนแปลงความสนใจของผู้ใช้ (ข้อมูลตอบกลับของผู้ใช้)

จากการหัวข้อการเรียนรู้ความสนใจของผู้ใช้ที่กล่าวไปแล้วข้างต้นว่าเป็นส่วนที่สำคัญในการรองรับการเปลี่ยนแปลงความสนใจของผู้ใช้ วิธีการที่ควรจะนำมาพัฒนาคือการใช้ข้อมูลตอบกลับ เพราะสามารถปรับเปลี่ยนความสนใจของผู้ใช้ได้ทันท่วงทีและยังใช้เป็นส่วนเสริมในการประเมินความถูกต้องหรือพึงพอใจในการให้คำแนะนำของระบบได้

ข้อมูลตอบกลับของผู้ใช้แบ่งเป็น 2 ลักษณะคือ 1) การตอบกลับแบบปริยาย (Implicit Feedback) เป็นการตรวจสอบจากการสนใจสินค้าและบริการที่ระบบแนะนำให้แก่ผู้ใช้เช่น การคลิกเข้าไปดูรายละเอียดของสินค้าเป็นต้น ซึ่งผู้ใช้จะไม่ทราบว่ามีการเก็บข้อมูล อย่างไรก็ตามวิธีการตอบกลับลักษณะนี้ควรจะมีขั้นตอนในการตรวจสอบว่าการคลิกเข้าไปดูนั้นผู้ใช้มีความสนใจจริงหรือไม่ หรือเพียงคลิกผิดพลาดเท่านั้น 2) การตอบกลับแบบชัดเจน (Explicit Feedback) อีกวิธีการหนึ่งที่ได้รับความสนใจ โดยวิธีการนี้จะได้ความเห็นจากผู้ใช้ว่าข้อมูลที่ระบบแนะนำมีความถูกต้องหรือตรงกับความสนใจของผู้ใช้เพียงใด (ไกรศักดิ์, 2555) เช่น การกดปฏิเสธสินค้าการให้คะแนนสินค้าข้อความแนะนำจากผู้ใช้ เป็นต้น แม้จะเป็นวิธีการที่ต้องมีการรบกวนผู้ใช้แต่หากออกแบบส่วนติดต่อที่สามารถเสริมสร้างแรงจูงใจให้ผู้ใช้รู้สึกว่าไม่เป็นการรบกวนข้อมูลตอบกลับนี้จะช่วยปรับปรุงความสนใจของผู้ใช้ได้มีประสิทธิภาพมากกว่าและได้ยังสามารถปรับปรุงทันท่วงทีได้อีกด้วย

ผู้วิจัยจึงได้ศึกษาวิธีการที่จะนำข้อมูลตอบกลับไปใช้ในการปรับปรุงความสนใจ โดย พบว่าในระบบค้นคืนสารสนเทศ (Information Retrieval) ข้อมูลตอบกลับจากผู้ใช้ได้นำมาใช้อย่างแพร่หลายในการปรับปรุงประสิทธิภาพในการค้นคืน ทั้งการค้นคืนเว็บไซต์ รูปภาพ เป็นต้น โดยอัลกอริทึมที่นิยมคือ Rocchio (Salton, 1971) ซึ่งเป็นอัลกอริทึมที่ช่วยปรับปรุงประสิทธิภาพของการสอบถาม (Query) ตั้งต้น โดยใช้ข้อมูลตอบกลับจากผู้ใช้เพื่อให้สามารถค้นคืนเอกสารที่มีความเกี่ยวข้องได้ดียิ่งขึ้น ซึ่งใช้วิธีการให้ค่าน้ำหนักไปยังรายการของเอกสารที่ผู้ใช้ให้กับระบบว่าเกี่ยวข้อง (relevant) หรือไม่เกี่ยวข้อง (non-relevant) แล้วนำผลรวมของค่าทั้งสองไปปรับปรุงประสิทธิภาพของการสอบถาม ดังสมการที่ (1) อัลกอริทึมนี้มีข้อเสียคือไม่มีทฤษฎีมารองรับสูตรการคำนวณ อย่างไรก็ตามผลที่ได้จากการทดลองพบว่าเมื่อใช้อัลกอริทึมนี้ ประสิทธิภาพของการค้นคืนก็ดีขึ้นอย่างมีนัยสำคัญ

(Pazzani and Billsus, 2007) โดยในระบบแนะนำข้อมูล สามารถนำหลักการของอัลกอริทึมนี้ไปประยุกต์ใช้ได้ โดยให้ผู้ใช้แบ่งรายการการแนะนำเป็นสิ่งที่ผู้ใช้สนใจ หรือไม่สนใจ เพื่อเรียนรู้ความสนใจของผู้ใช้และปรับเปลี่ยนการแนะนำให้ตรงกับสิ่งที่ผู้ใช้สนใจได้ดียิ่งขึ้น

$$Q_{i+1} = \alpha Q_i + \beta \sum_{rel} \frac{D_i}{|D_i|} - \gamma \sum_{nonrel} \frac{D_i}{|D_i|} \quad (1)$$

Q_i คือ การสอบถาม (Query) ของผู้ใช้ ครั้งที่ i
α, β, γ คือ ค่าน้ำหนักของแต่ละการกระทำ
rel คือ เอกสาร (Document-D) ที่มีความเกี่ยวข้อง
$nonrel$ คือ เอกสาร (Document-D) ที่ไม่มีความเกี่ยวข้อง

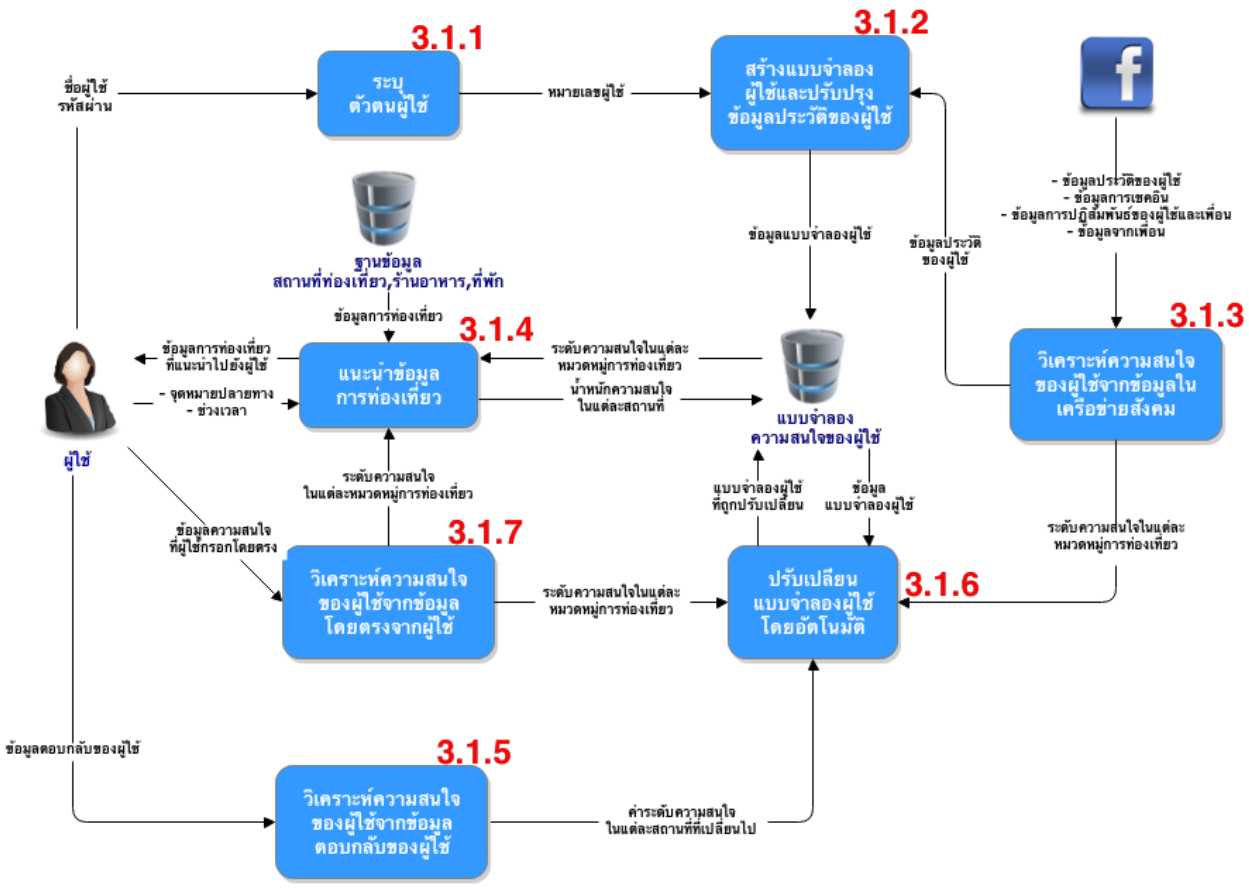
จากทฤษฎีต่างๆ ที่อธิบายมาในบทนี้จะถูกนำไปประยุกต์ใช้ในระบบที่นำเสนอในส่วนต่างๆ ซึ่งจะอธิบายรายละเอียดในบทถัดไป

บทที่ 3

วิธีดำเนินการวิจัย

ในบทนี้จะอธิบายถึงกรอบแนวคิดงานวิจัยซึ่งประกอบด้วย 7 ส่วนสำคัญแสดงดังรูปที่ 12 ซึ่งจะอธิบายรายละเอียดในแต่ละส่วนตามหัวข้อต่างๆ ดังต่อไปนี้

3.1 กรอบแนวคิดงานวิจัย

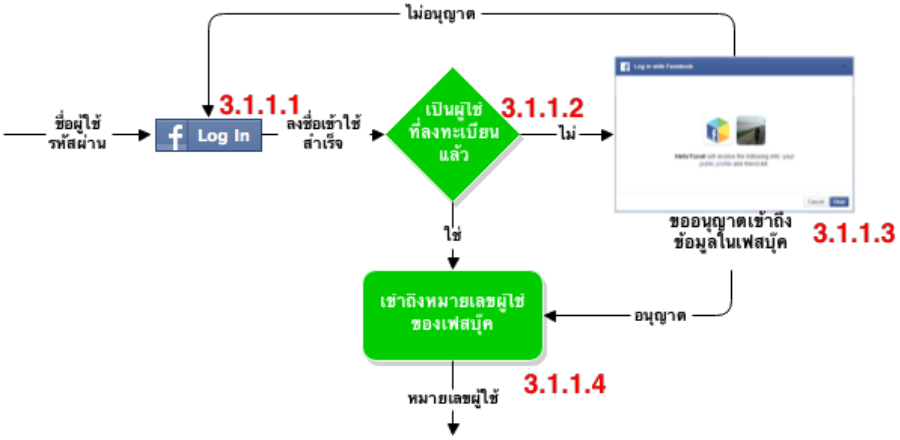


รูปที่ 12 ภาพรวมกรอบแนวคิดของโครงการวิจัย

จากกรอบแนวคิดในรูปที่ 12 ผู้ใช้เข้าถึงผ่านหน้าเว็บไซต์การท่องเที่ยวที่มีระบบแนะนำข้อมูลส่วนบุคคลสำหรับข้อมูลการท่องเที่ยว แนวคิดหลักของระบบคือช่วยแนะนำข้อมูลสถานที่ท่องเที่ยวตามความสนใจของนักท่องเที่ยวที่เข้ามาใช้งานในเว็บไซต์ โดยลักษณะการเข้าถึงระบบจะแบ่งผู้ใช้ออกเป็น 2 แบบ คือ

1) ผู้ใช้แบบผู้มาเยือน (Guest) คือ ผู้ใช้ที่ไม่ได้ทำการล็อกอินเข้าสู่ระบบ โดยระบบจะไม่จัดเก็บข้อมูลของผู้ใช้ลงในฐานข้อมูล ซึ่งระบบยังคงให้การแนะนำข้อมูลสถานที่ท่องเที่ยวได้ได้ แต่ผู้ใช้จะต้องระบุความสนใจด้วยตัวเอง จากกรอบแนวคิดในรูปที่ 12 ผู้ใช้แบบผู้มาเยือนจะใช้เพียงกระบวนการที่ 3.1.4 และ 3.1.7

2) ผู้ใช้แบบสมาชิก (Member) คือ ผู้ใช้ที่ทำการล็อกอินเข้าสู่ระบบ โดยระบบจะใช้เทคนิคการวิเคราะห์ข้อมูลจากเครือข่ายสังคมในการหาความสนใจและจัดเก็บเป็นแบบจำลองความสนใจลงในฐานข้อมูล จากนั้นจะใช้ความสนใจในการแนะนำข้อมูลสถานที่ท่องเที่ยวไปยังผู้ใช้ ซึ่งเทคนิคในการใช้เครือข่ายสังคมนี้สามารถช่วยลดปัญหาโคลสตาร์ทที่มักเกิดในระบบแนะนำข้อมูลที่ยังไม่มีข้อมูลความสนใจของผู้ใช้เพียงพอที่จะแนะนำข้อมูลการท่องเที่ยวที่ตรงกับความต้องการของผู้ใช้ได้ ส่วนการระบุตัวตนของผู้ใช้แบบสมาชิกนั้นใช้การล็อกอินผ่านเครือข่ายสังคมเพื่อลดขั้นตอนการลงทะเบียนที่ยุ่งยาก นอกจากนี้ระบบยังใช้ข้อมูลตอบกลับจากผู้ใช้เพื่อช่วยในการปรับปรุงข้อมูลความสนใจของผู้ใช้ให้มีความถูกต้องมากขึ้นเพื่อที่จะได้แนะนำข้อมูลการท่องเที่ยวได้ตรงกับความต้องการของผู้ใช้มากยิ่งขึ้น จากกรอบแนวคิดในรูปที่ 13 ผู้ใช้แบบสมาชิก ระบบจะใช้กระบวนการที่ 3.1.1 - 3.1.6



รูปที่ 13 วิธีการระบุตัวตนผู้ใช้

3.2 กระบวนการทำงานของระบบสำหรับผู้ใช้แบบสมาชิก (Member)

สำหรับผู้ใช้งานแบบสมาชิกนั้น ระบบใช้กระบวนการหลายขั้นตอน ซึ่งมีประสิทธิภาพในการแนะนำข้อมูลที่ดีกว่าผู้ใช้งานแบบผู้มาเยือน โดยกระบวนการจะทำงานตามลำดับดังนี้

3.2.1 การระบุตัวตนผู้ใช้ (User Identification)

ผู้ใช้เริ่มใช้งานระบบโดยการล็อกอินเพื่อระบุตัวตนของผู้ใช้ โดยงานวิจัยนี้จะใช้เทคนิคการระบุตัวตน (รูปที่ 13) โดยให้ผู้ใช้ล็อกอินผ่านเฟสบุ๊คด้วย Facebook API (3.1.1.1) เพื่อลดขั้นตอนในการกรอกข้อมูลเพื่อสมัคร

เข้าใช้งานระบบ จากนั้นระบบจะตรวจสอบว่าเป็นผู้ใช้ที่เข้าใช้งานครั้งแรกหรือได้เคยเข้าใช้ระบบแล้ว (3.1.1.2) โดยลักษณะของผู้ใช้ที่เข้าใช้งานทั้งสองแบบ ระบบจะแบ่งกรทำงานออกเป็น 2 กรณี คือ

1) ผู้ใช้ใหม่ที่เข้าใช้งานครั้งแรก (New Member) ระบบจะทำการขออนุญาตเข้าถึงข้อมูลในเฟสบุ๊คของผู้ใช้ (3.1.1.3) เพื่อนำข้อมูลเหล่านี้ไปวิเคราะห์หาความสนใจของผู้ใช้ ซึ่งจะทำการขออนุญาตเพียงครั้งแรกที่เข้าใช้เท่านั้น จากนั้นระบบจะดึงข้อมูลหมายเลขผู้ใช้ของเฟสบุ๊คหรือ Facebook ID (3.1.1.4) เพื่อใช้เป็นหมายเลขผู้ใช้ (User ID) สำหรับระบุตัวตนของผู้ใช้ในการสร้าง การจับเก็บและการปรับปรุงแบบจำลองความสนใจ เพื่อให้ระบบทำการแนะนำข้อมูลในขั้นตอนต่อไป

2) ผู้ใช้เดิมที่เคยเข้าใช้งานระบบแล้ว (Existing Member) ระบบจะทำการดึงหมายเลขผู้ใช้ของเฟสบุ๊คสำหรับระบุตัวตน ในการดึงแบบจำลองความสนใจที่เก็บไว้ในฐานข้อมูลมาใช้หรือปรับปรุง เพื่อให้ระบบทำการแนะนำข้อมูลในขั้นตอนต่อไป

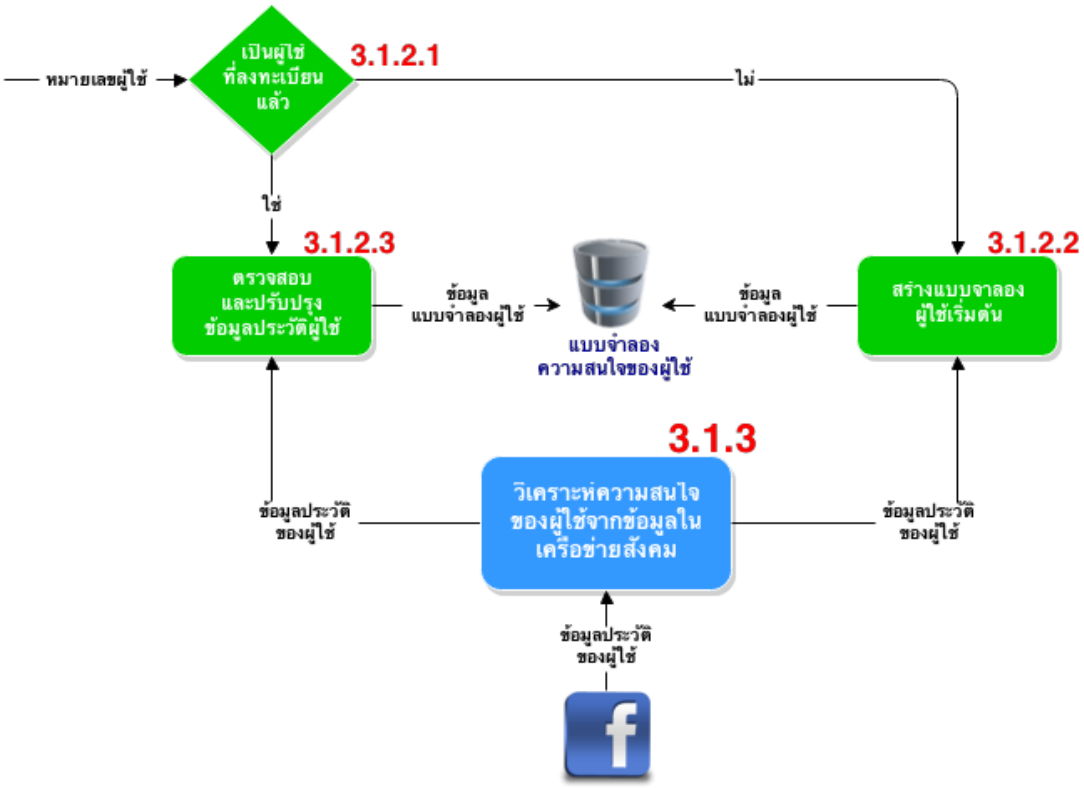
เมื่อระบบได้หมายเลขผู้ใช้แล้ว ก็จะส่งไปยังกระบวนการที่ 3.1.2 ซึ่งมีการทำงานดังนี้

3.3 การสร้างแบบจำลองความสนใจของผู้ใช้และการปรับปรุงข้อมูลประวัติของผู้ใช้ (User Interests Model Construction and User Profile Updating)

แบบจำลองความสนใจของผู้ใช้ (User Interests Model) เป็นส่วนที่ใช้จัดเก็บข้อมูลที่แสดงถึงความสนใจของผู้ใช้ (User Interests) และข้อมูลประวัติผู้ใช้ (User Profile) โดยข้อมูลความสนใจของผู้ใช้เป็นการจัดเก็บระดับความสนใจในแต่ละหมวดหมู่ของสถานที่ท่องเที่ยวและสถานที่ท่องเที่ยวที่ผู้ใช้สนใจ ในส่วนข้อมูลประวัติจะเก็บข้อมูลที่แสดงถึงรายละเอียดของผู้ใช้ โดยกระบวนการสร้างแบบจำลองความสนใจของผู้ใช้ตั้งต้นและปรับปรุงในส่วนของข้อมูลประวัติผู้ใช้แสดงดังรูปที่ 14 โดยระบบจะดึงข้อมูลข้อมูลประวัติผู้ใช้จากเฟสบุ๊ค ซึ่งได้ผ่านการตรวจสอบข้อมูล (Data Cleansing) มาจากกระบวนการที่ 3.1.3 ซึ่งจะกล่าวเป็นกระบวนการถัดไป โดยในกระบวนการนี้จะแบ่งการทำงานเป็น 2 กรณี เช่นกัน คือ

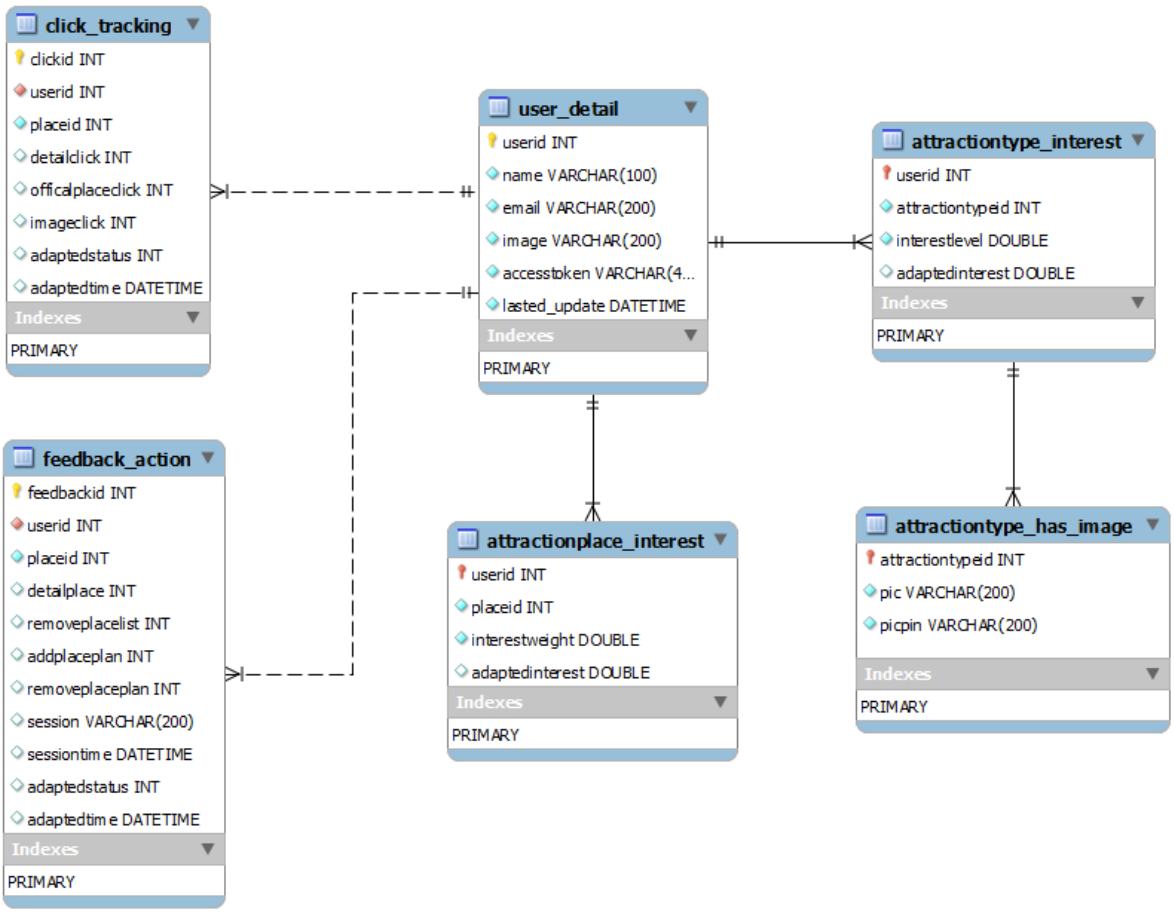
1) ผู้ใช้ใหม่ที่เข้าใช้งานครั้งแรก ระบบจะสร้างแบบจำลองความสนใจของผู้ใช้ตั้งต้น และทำการดึงข้อมูลประวัติผู้ใช้จากเฟสบุ๊คเพื่อใส่ข้อมูลประวัติของผู้ใช้บนแบบจำลอง (3.1.2.2)

2) ผู้ใช้เดิมที่เคยเข้าใช้งานระบบแล้ว ระบบจะตรวจสอบว่ามีการเปลี่ยนแปลงข้อมูลประวัติในเฟสบุ๊คหรือไม่ ถ้ามีก็จะเปลี่ยนแปลงไปยังแบบจำลองผู้ใช้ (3.1.2.3)



รูปที่ 14 วิธีการสร้างโมเดลผู้ใช้และการปรับปรุงโมเดล

จากนั้นจะนำแบบจำลองจัดเก็บลงฐานข้อมูล โดยการจัดเก็บจะใช้รูปแบบฐานข้อมูลสัมพันธ์ (Relational Database) เพื่อความสะดวกในการจัดการข้อมูล ซึ่งมีประสิทธิภาพมากกว่าการจัดเก็บแบบ Flat File เช่น การจัดหมวดหมู่หรือการเก็บข้อมูลที่มีความสัมพันธ์กัน การค้นหาที่รวดเร็วกว่า เป็นต้น นอกจากนี้ยังมีความปลอดภัยในการจัดเก็บข้อมูลที่ดีกว่าอีกด้วย รวมถึงมีความเหมาะสมกว่าการจัดเก็บแบบออนโทโลจีในเรื่องของความง่ายในการแก้ไขข้อมูลแบบอัตโนมัติโดยระบบ โดยผู้วิจัยได้ออกแบบฐานข้อมูลที่ใช้จัดเก็บแบบจำลองความสนใจของผู้ใช้ดังรูปที่ 15



รูปที่ 15 ฐานข้อมูลที่ใช้จัดเก็บแบบจำลองความสนใจของผู้ใช้

ความหมายของแต่ละตารางและฟิลด์ข้อมูลในฐานข้อมูลแสดงดังตารางที่ 5

ตารางที่ 5 ความหมายของแต่ละตารางและฟิลด์ข้อมูล

ตาราง		ฟิลด์ข้อมูล	
ชื่อตาราง	**ความหมาย**	**ชื่อฟิลด์**	**ความหมาย**
user_detail	แสดงรายละเอียดของผู้ใช้	userid name email image accesstoken lasted_update	หมายเลขผู้ใช้ ชื่อผู้ใช้ ที่อยู่อิเมลล์ รูปภาพประจำตัวผู้ใช้ token ในการเข้าถึงข้อมูลจากเฟสบุ๊ค การเข้าใช้งานระบบล่าสุด
attractiontype_interest	แสดงระดับความสนใจของผู้ใช้ในแต่ละหมวดหมู่ของสถานที่	userid attractiontypeid	หมายเลขผู้ใช้ หมายเลขหมวดหมู่ของสถานที่ท่องเที่ยว

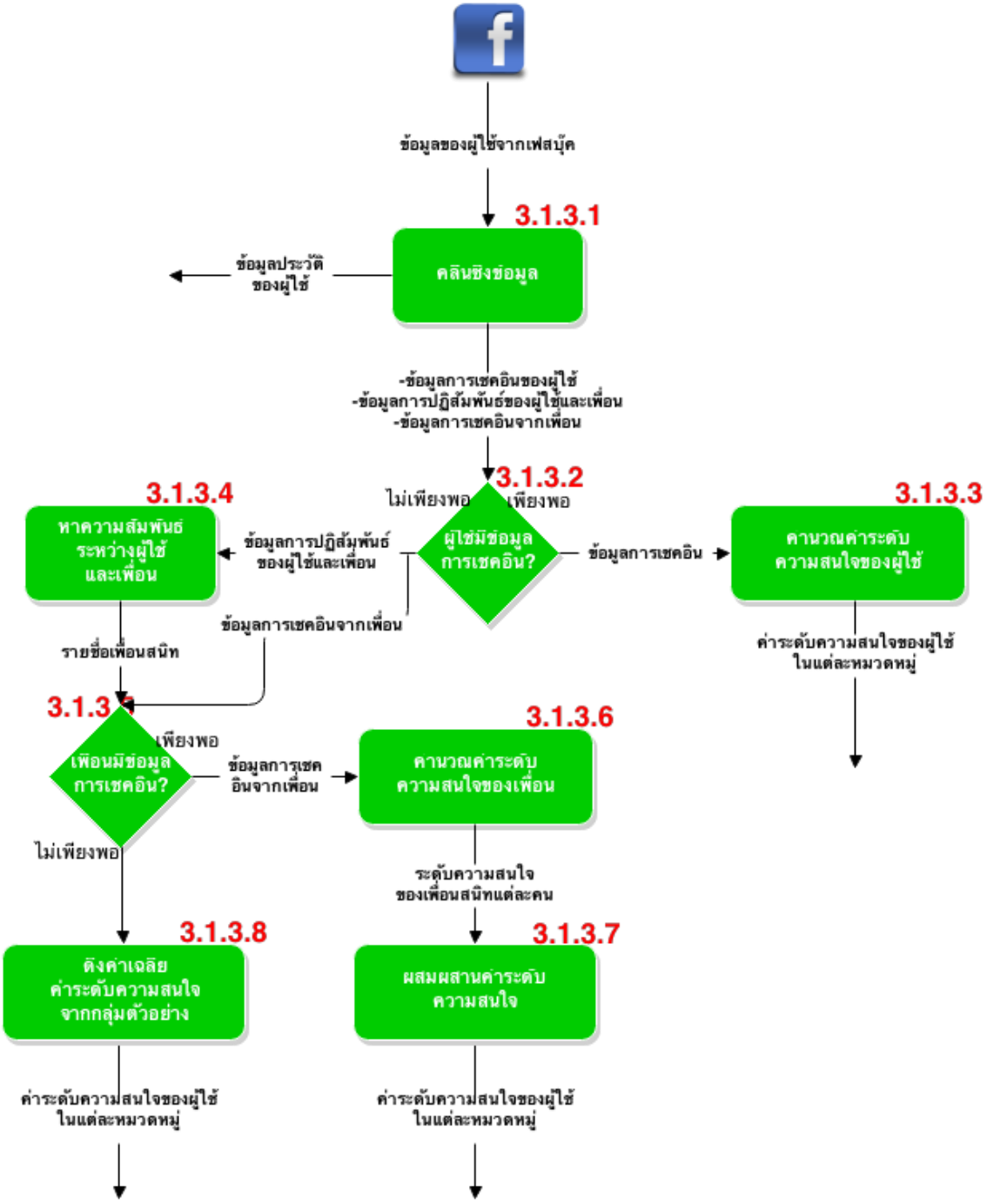
ตาราง		ฟิลด์ข้อมูล	
ชื่อตาราง	**ความหมาย**	**ชื่อฟิลด์**	**ความหมาย**
	ท่องเที่ยว	interestlevel	ระดับความสนใจในหมวดหมู่การท่องเที่ยว
		adaptedinterest	ค่าความสนใจในหมวดหมู่การท่องเที่ยวที่เปลี่ยนไป
attractiontype_has_image	แสดงรูปภาพไอคอนและรูปภาพที่ใช้ปักหมุดบนแผนที่ของหมวดหมู่สถานที่ท่องเที่ยว	attractiontypeid	หมายเลขหมวดหมู่ของสถานที่ท่องเที่ยว
		pic	รูปภาพไอคอนหมวดหมู่ของสถานที่ท่องเที่ยว
		picpin	รูปภาพปักหมุดบนแผนที่ของหมวดหมู่ของสถานที่ท่องเที่ยว
attractionplace_interest	แสดงค่าน้ำหนักของแต่ละสถานที่ของผู้ใช้	userid	หมายเลขผู้ใช้
		placeid	หมายเลขสถานที่
		interestweight	ค่าน้ำหนักของสถานที่
		adaptedinterest	ค่าน้ำหนักของสถานที่ที่เปลี่ยนแปลงไป
click_tracking	การคลิกส่วนต่างๆในหน้าแสดงรายละเอียดของแต่ละสถานที่ท่องเที่ยวของแต่ละผู้ใช้	clickid	หมายเลขการคลิก
		userid	หมายเลขผู้ใช้
		placeid	หมายเลขสถานที่
		detailclick	ผู้ใช้คลิกดูรายละเอียดของสถานที่
		officalplaceclick	ผู้ใช้คลิกไปยังเว็บไซต์หลักของสถานที่
		imageclick	ผู้ใช้คลิกดูรูปภาพของสถานที่
		adaptedstatus	สถานะในการนำข้อมูลการคลิกไปประมวลผลเพื่อปรับปรุงข้อมูล โดย 1 = นำไปใช้แล้ว และ 0 = ยังไม่ได้นำไปใช้
		adaptedtime	นำข้อมูลไปประมวลผลเพื่อปรับปรุงข้อมูลเมื่อใด
feedback_action	แสดงข้อมูลตอบกลับที่ผู้ใช้กระทำไปยังสถานที่ที่ระบบแนะนำข้อมูล	feedbackid	หมายเลขข้อมูลตอบกลับ
		userid	หมายเลขผู้ใช้
		placeid	หมายเลขสถานที่
		detailplace	ผู้ใช้เข้าไปดูข้อมูลของสถานที่เพิ่มเติม (ดูจากข้อมูลในตาราง click_tracking)
		removeplacelist	

ตาราง		ฟิลด์ข้อมูล	
ชื่อตาราง	**ความหมาย**	**ชื่อฟิลด์**	**ความหมาย**
		addplaceplan	ผู้ใช้ลบสถานที่ออกจากการแนะนำ เพื่อให้ระบบแนะนำใหม่
		removeplaceplan	ผู้ใช้เพิ่มไปยังสถานที่ที่ต้องการสร้างแผนการเดินทาง
		session	ผู้ใช้ลบออกจากสถานที่ที่ต้องการสร้างแผนการเดินทาง
		sessiontime	หมายเลขเซสชั่นขณะกระทำข้อมูลตอบกลับ
		adaptedstatus	สร้างเซสชั่นนี้เมื่อเวลาใด
		adaptedtime	สถานะในการนำข้อมูลตอบกลับไปประมวลผลเพื่อปรับปรุงข้อมูล โดย 1 = นำไปใช้แล้ว และ 0 = ยังไม่ได้นำไปใช้ นำข้อมูลตอบกลับไปประมวลผลเพื่อปรับปรุงข้อมูลเมื่อใด

เมื่อทำการจัดเก็บแบบจำลองความสนใจลงฐานข้อมูลแล้ว กระบวนการต่อไปเป็นกระบวนการที่สำคัญในการจัดการปัญหาโคลด์สตาร์ทคือ การวิเคราะห์ความสนใจจากข้อมูลในเครือข่ายสังคม

3.4 วิเคราะห์ความสนใจของผู้ใช้จากข้อมูลในเครือข่ายสังคมออนไลน์ (User Interests Analysis using Online Social Network Information)

ขั้นตอนนี้เป็นการวิเคราะห์ข้อมูลจากเครือข่ายสังคมเพื่อค้นหาความสนใจของผู้ใช้ดังรูปที่ 16 โดยกระบวนการนี้ออกแบบมาเพื่อแก้ปัญหาโคลด์สตาร์ททที่มักเกิดขึ้นกับผู้ใช้ที่เริ่มต้นใช้งานระบบ ซึ่งระบบจะใช้ข้อมูลตามลำดับจากเครือข่ายสังคม ถ้าลำดับก่อนหน้าไม่มีข้อมูลเพียงพอให้วิเคราะห์ความสนใจได้ก็ใช้ข้อมูลถัดไป ลำดับของข้อมูลที่จะใช้ คือ 1) ใช้ข้อมูลในเครือข่ายสังคมของผู้ใช้ 2) ใช้ข้อมูลในเครือข่ายสังคมของเพื่อนผู้ใช้ 3) ใช้ข้อมูลค่าเฉลี่ยของกลุ่มตัวอย่างในเครือข่ายสังคม



รูปที่ 16 วิธีการวิเคราะห์ข้อมูลจากเครือข่ายสังคม

ในการนำข้อมูลจากเครือข่ายสังคมมาใช้จะใช้การดึงข้อมูลมาจากเฟสบุ๊คผ่าน Facebook API ซึ่งจะได้ข้อมูลเป็นแบบ JSON โดยข้อมูลที่สามารถดึงมาจากเฟสบุ๊คประกอบด้วยหลายฟิลด์ (Field) ข้อมูล ดังรูปที่ 17

fields	languages
about	last_name
address	link
age_range	locale
bio	location
birthday	meeting_for
can_send_to_mobile	middle_name
cover	name
currency	name_format
devices	payment_engagement
education	payment_mobile_pricepoints
email	payment_pricepoints
favorite_athletes	payment_test_group
favorite_teams	political
first_name	quotes
gender	relationship_status

รูปที่ 17 ตัวอย่างฟิลด์ข้อมูลของผู้ใช้ที่สามารถดึงมาจากเฟสบุ๊ค

แต่อย่างไรก็ตามระบบจะเลือกใช้เฉพาะฟิลด์ที่มีประโยชน์ต่อการระบุตัวตนและวิเคราะห์ความสนใจของผู้ใช้ โดยฟิลด์ข้อมูลที่ผู้วิจัยนำมาใช้ภายในระบบประกอบด้วย

1) ข้อมูลเพื่อระบุตัวตนของผู้ใช้ แสดงดังตารางที่ 6 คือ

1.1) user id (รหัสผู้ใช้) โดยใช้ฟิลด์ id (หมายเลขผู้ใช้ของเฟสบุ๊ค)

1.2) username (ชื่อผู้ใช้) โดยใช้ฟิลด์ first_name (ชื่อ) last_name (นามสกุล)

1.3) email (อิเมล์) โดยใช้ฟิลด์ email (อิเมล์)

โดยข้อมูลเพื่อระบุตัวตนของผู้ใช้เหล่านี้เป็นข้อมูลประวัติของผู้ใช้ ที่จะส่งไปยังกระบวนการ 3.3 เพื่อนำข้อมูลประวัติผู้ใช้บนเครือข่ายสังคมไปสร้างหรือปรับปรุงข้อมูลประวัติผู้ใช้ในแบบจำลอง ซึ่งได้กล่าวมาแล้วก่อนหน้านี้

ตารางที่ 6 ข้อมูลเพื่อระบุตัวตนของผู้ใช้

ข้อมูลที่ได้จากเฟสบุ๊ค	ความหมาย

<pre>{
"id": "100000965654095",</pre>
	หมายเลขผู้ใช้ของเฟสบุ๊ค ใช้ระบุตัวตนในการเข้าถึงระบบ

<pre>"first_name": "TiDz",
"last_name": "PohSkidjai",</pre>
	ชื่อและนามสกุล ใช้แสดงบนส่วนติดต่อกับผู้ใช้ (UI)

<pre>"email": "tidz.pohskidjai@gmail.com"
}</pre>
	อิเมลล์ ใช้สำหรับแจ้งข้อมูลต่างๆ ของระบบ

2) ข้อมูลการเช็คอินในแต่ละสถานที่ นำมาจากฟิลด์ location โดยแสดงรายละเอียดดังตารางที่ 7 และข้อมูลรายละเอียดของสถานที่ที่ถูกเช็คอินดังตารางที่ 8 เพื่อนำมาใช้ในการวิเคราะห์ความสนใจของผู้ใช้

ตารางที่ 7 ข้อมูลการเช็คอินของผู้ใช้จำนวน 1 เช็คอิน

ข้อมูลที่ได้จากเฟสบุ๊ค	ความหมาย
{ "place": {	การเช็คอินในสถานที่จำนวน 1 เช็คอิน
"id": "196230743801741",	หมายเลขไอดีของสถานที่ ที่ทำการเช็คอิน
"name": "วัดพระศรีรัตนมหาธาตุราชวรวิหาร",	ชื่อของสถานที่ ที่ทำการเช็คอิน
"location": {	ตำแหน่งที่ตั้งของสถานที่ ที่ทำการเช็คอิน
"latitude": 17.429238993972,	ละติจูดของสถานที่ ที่ทำการเช็คอิน
"longitude": 99.811871548508,	ลองจิจูดของสถานที่ ที่ทำการเช็คอิน
"street": "อ.ศรีสัชนาลัย" }	ข้อมูลที่อยู่ของสถานที่ ที่ทำการเช็คอิน
}, "id": "750723874966970",	หมายเลขไอดีของโพสที่ทำการเช็คอิน
"created_time": "2013-04-13T09:21:03+0000" }	เวลาในการสร้างโพสที่ทำการเช็คอิน

ตารางที่ 8 หมวดหมู่ (category_list) ของสถานที่ที่ถูกเช็คอิน

ข้อมูลที่ได้จากเฟสบุ๊ค	ความหมาย
{ "id": "196230743801741",	หมายเลขไอดีของสถานที่ ที่ทำการเช็คอิน
"name": "วัดพระศรีรัตนมหาธาตุราชวรวิหาร",	ชื่อของสถานที่ ที่ทำการเช็คอิน
"category_list": [	รายการหมวดหมู่ของสถานที่ที่จัดแบ่งบนเฟสบุ๊ค
{ "id": "124861974254366",	หมายเลขไอดีหมวดหมู่ที่ 1 ของสถานที่
"name": "Tours & Sightseeing" },	ชื่อหมวดหมู่ที่ 1 ของสถานที่
{ "id": "197097220301977",	หมายเลขไอดีหมวดหมู่ที่ 2 ของสถานที่
"name": "Historical Place" }	ชื่อหมวดหมู่ที่ 2 ของสถานที่
] }	

โดยข้อมูลที่นำมาใช้ในการวิเคราะห์ความสนใจ คือ 1. Place -> id (รหัสสถานที่) 2. created_time (เวลาที่เช็คอิน) และ 3. category_list (หมวดหมู่ของสถานที่) เป็นต้น

เมื่อได้ข้อมูลของผู้ใช้ตามฟิลด์ข้อมูลที่กล่าวข้างต้นแล้ว ผู้วิจัยจะทำการตรวจสอบข้อมูลการเช็คอินอีกครั้ง โดยจะเลือกใช้ข้อมูลการเช็คอินที่มีจุดประสงค์เพื่อการท่องเที่ยว เนื่องจากการศึกษาพบว่าผู้ใช้ในเครือข่ายสังคมมักจะเช็คอินในสถานที่หนึ่งๆ เป็นจำนวนมากเหมือนๆ กัน เช่น ห้างสรรพสินค้า โรงภาพยนตร์ บ้านของผู้ใช้ เป็นต้น ซึ่งสถานที่นี้ส่วนมากเป็นกิจกรรมในการดำเนินชีวิตทั่วไป อย่างไรก็ตามมีบางกรณีที่สถานที่เหล่านี้อาจจัดว่าเป็นสถานที่ท่องเที่ยวได้ ดังนั้นผู้วิจัยจึงเพิ่มการกรองสถานที่ที่ไม่ใช่จุดประสงค์เพื่อการท่องเที่ยวออก โดยใช้มิติของเวลาเข้ามาคำนวณ โดยสถานที่มีการเช็คอินมากกว่าหนึ่งครั้งต่อสัปดาห์จะไม่ถูกนำมาพิจารณา อีกทั้งยังใช้ค่าของเวลาลดความซ้ำซ้อนของการเช็คอินที่อาจเกิดขึ้นหลายครั้งในวันเดียวกัน ยกตัวอย่างเช่น การโพสรูปภาพหลายๆ รูปภาพที่ติดกับเช็คอินในสถานที่เดียวกันในวันเดียวกัน สถานที่ที่ซ้ำซ้อนเหล่านี้จะถูกตัดออกให้เหลือเพียงหนึ่งครั้งในหนึ่งวัน

จากนั้นนำสถานที่ที่เช็คอินเหล่านี้จัดลงในแต่ละหมวดหมู่ที่ผู้วิจัยกำหนดไว้ อย่างไรก็ตามพบว่า category_list ที่มีในเฟสบุ๊คนั้นมีจำนวนมาก ประกอบกับฟังก์ชันของเฟสบุ๊คที่ให้ผู้ใช้สามารถสร้างหมวดหมู่เองได้ ผู้วิจัยจึงเลือกใช้ category_list ที่เป็นพื้นฐานของเฟสบุ๊คและ category_list ที่เก็บจากตัวอย่าง โดยผู้วิจัยได้เลือกตัวอย่างเป็นผู้ใช้งานเฟสบุ๊คจำนวน 1027 บัญชีผู้ใช้ ที่มีจำนวนการเช็คอินไม่ต่ำกว่า 10 ครั้ง โดยเก็บข้อมูล category_list จากการเช็คอินของผู้ใช้เหล่านี้ จากนั้นนำมาจำแนกลงหมวดหมู่ที่กำหนดไว้ เช่น สถานที่ท่องเที่ยวแบ่งเป็น 6 หมวด หมู่ ซึ่งปรับปรุงมาจากหมวดหมู่หลักของการท่องเที่ยวแห่งประเทศไทย (ททท) ได้แก่

1) สถานที่ท่องเที่ยวทางวัฒนธรรม หมายถึงแหล่งท่องเที่ยวที่ทำให้นักท่องเที่ยวได้รู้จักวิถีชีวิต วัฒนธรรม และความเป็นอยู่ของผู้คนในชุมชน ในสถานที่ที่จัดนิทรรศการซึ่งแสดงเรื่องราวและงานศิลปหัตถกรรมของผู้คนในชุมชน เช่น งานประเพณี การแสดงศิลปวัฒนธรรม งานแสดงสินค้าพื้นเมือง เป็นต้น โดยจัด category_list ในเฟสบุ๊คให้อยู่ในหมวดหมู่นี้ ยกตัวอย่างเช่น Cultural_Gifts_Store, Fairground, Shopping District เป็นต้น

2) สถานที่ท่องเที่ยวทางประวัติศาสตร์ หมายถึงสถานที่ท่องเที่ยวที่แสดงถึงร่องรอยของความรุ่งเรืองในอดีตของสถานที่สำคัญหลายแห่ง ซึ่งบางแห่งได้รับคัดเลือกให้เป็นแหล่งมรดกโลกอีกด้วย เช่น โบราณสถาน อุทยานประวัติศาสตร์ วัด ศาสนสถาน สิ่งก่อสร้างที่มีคุณค่าทางศิลปะและสถาปัตยกรรม เป็นต้น โดยจัด category_list ในเฟสบุ๊คให้อยู่ในหมวดหมู่นี้ ยกตัวอย่างเช่น Historical_Place, Buddhist_Temple, Religious_Organization เป็นต้น

3) สถานที่ท่องเที่ยวเพื่อความบันเทิง หมายถึงสถานที่ท่องเที่ยวเชิงสันทนาการและบันเทิง ซึ่งมีหลายรูปแบบ เช่น สวนสัตว์ สวนสนุก ย่านบันเทิง สวนสาธารณะ สิ่งปลูกสร้างต่างๆ เป็นต้น โดยจัด category_list ในเฟสบุ๊คให้อยู่ในหมวดหมู่นี้ ยกตัวอย่างเช่น Market, Park, Art&Entertaimment, Outdoor, Zoo เป็นต้น

4) สถานที่ท่องเที่ยวทางธรรมชาติ เหมายถึงสถานที่ท่องเที่ยวที่เกิดจากแหล่งธรรมชาติหลากหลายประเภท ตั้งแต่ ภูเขา ผืนป่า น้ำตก โถงถ้ำ ทะเลสาบ ทุ่งดอกไม้ น้ำพุร้อน ฯลฯ โดยผู้วิจัยได้เลือก category_list ในเฟสบุ๊คให้อยู่ในหมวดหมู่นี้ยกตัวอย่างเช่น National_Park, Mountain, River, Rock Climbing, Lake เป็นต้น

5) สถานที่ท่องเที่ยวเพื่อการศึกษา หมายถึงสถานที่ท่องเที่ยวเชิงวิชาการสำหรับผู้ที่สนใจใฝ่หาความรู้ อาทิ ห้องสมุด พิพิธภัณฑ์เพื่อการศึกษา ศูนย์ฝึกอบรม เป็นต้น โดยจัด category_list ในเฟสบุ๊คให้อยู่ในหมวดหมู่นี้ ยกตัวอย่างเช่น Museum, Education เป็นต้น

6) สถานที่ท่องเที่ยวทางวิถีชีวิตชุมชน เป็นสถานที่ท่องเที่ยวที่แสดงถึงวิถีชีวิตของคนในชุมชน เช่น การท่องเที่ยวในรูปแบบโฮมสเตย์ หรือจะเป็นตลาด ตลาดน้ำ ที่เป็นแหล่งท่องเที่ยวที่เต็มไปด้วยสีสันของวิถีชีวิตไทยดั้งเดิมและอุดมด้วยอาหารไทย หลากหลาย โดยจัด category_list ในเฟสบุ๊คให้อยู่ในหมวดหมู่นี้ ยกตัวอย่างเช่น Region, Farmers Market เป็นต้น

เมื่อได้ข้อมูลการเช็คอินที่จัดลงหมวดหมู่เรียบร้อยแล้ว ระบบจะตรวจสอบว่าข้อมูลเพียงพอที่จะนำมาวิเคราะห์ความสนใจหรือไม่ (3.1.3.2) โดยหากข้อมูลการเช็คอินที่ผ่านการตรวจสอบแล้ว มีไม่ต่ำกว่า 10 เช็คอิน

ระบบจะใช้ข้อมูลการเช็คอินของผู้ใช้เอง หากมีน้อยกว่าจะใช้ข้อมูลการเช็คอินของเพื่อนสนิท หากข้อมูลจากเพื่อนไม่เพียงพออีก ก็จะใช้ข้อมูลจากกลุ่มตัวอย่าง ดังนี้

1) ใช้ข้อมูลการเช็คอินของผู้ใช้เอง (3.1.3.3) นำข้อมูลการเช็คอินมาวิเคราะห์หาความสนใจของผู้ใช้ โดยใช้วิธีการคำนวณหาค่าระดับความสนใจ (Interest Level - I) ซึ่งใช้เทคนิคการแสดงค่าความสนใจด้วยตัวเลข (Numerical Representations) ที่ระบบ Sigtur/E-Destination (Moreno et al., 2013) ได้นำเสนอขึ้น แต่ในระบบที่ผู้วิจัยกำลังพัฒนานั้นจะใช้ข้อมูลจากเครือข่ายสังคมเพิ่มเติมเข้ามา โดยข้อมูลจากเครือข่ายสังคมจะใช้ข้อมูลจากการเช็คอิน ซึ่งลักษณะของเทคนิคนี้ ค่าระดับความสนใจ จะถูกคำนวณไปยังแต่ละหมวดหมู่ของการท่องเที่ยวจากการใช้ข้อมูลข้างต้น โดยค่าระดับความสนใจ ให้มีค่า 0.0 ถึง 1.0 โดย 0 หมายถึงไม่มีความสนใจ และถึง 1 หมายถึงมีความสนใจระดับสูงสุด

วิธีการคำนวณค่าระดับความสนใจ ใช้สูตรสถิติพื้นฐานเพื่อความรวดเร็วในการประมวลผล คือ การใช้ความถี่ของการเช็คอินที่ปรากฏในแต่ละหมวดหมู่ของแต่ละผู้ใช้ ดังเช่นสถานที่ท่องเที่ยวที่ถูกแบ่งออกเป็น 6 หมวดหมู่ดังกล่าวข้างต้น โดยค่าระดับความสนใจคำนวณดังสมการที่ (2)

$$I(c) = \frac{n_c}{\sum_{k=1}^{6} n_k} \quad , 1 \leq c \leq 6 \tag{2}$$

$I(c)$ คือค่าระดับความสนใจ (Interest Level - I) ที่มีต่อหมวดหมู่ (Category - c)
n_c คือจำนวนของการเช็คอินที่ปรากฏในหมวดหมู่ c
n_k คือจำนวนของการเช็คอินที่ปรากฏ ในทุกๆ หมวดหมู่ k

ค่าระดับความสนใจในแต่ละหมวดหมู่ของผู้ใช้แสดงดังสมการที่ (3)

$$I = \{I(c_1), I(c_2), I(c_3), I(c_4), I(c_5), I(c_6)\} \tag{3}$$

2) ใช้ข้อมูลของเพื่อนผู้ใช้ โดยจะใช้ในกรณีที่ผู้ใช้ไม่มีข้อมูลการเช็คอินเพียงพอ โดยตั้งสมมติฐานว่าผู้ใช้อื่นที่มีการปฏิสัมพันธ์กับผู้ใช้ในระดับสูง จะมีความสนใจสิ่งหนึ่งที่คล้ายกับผู้ใช้ ผู้วิจัยจึงใช้ข้อมูลจากเพื่อนเพื่อแก้ไขปัญหาโคลด์สตาร์ทนี้ โดยทำการวิเคราะห์เพื่อนของผู้ใช้ทั้งหมดในเฟสบุ๊คเพื่อหาเพื่อนที่มีปฏิสัมพันธ์กับผู้ใช้ในระดับที่สูงที่สุด 5 อันดับ (3.1.3.4) โดยประยุกต์ใช้อัลกอริทึมชื่อ ay-fb-friend-rank (gajus, 2012) และนำข้อมูลการเช็คอินของเพื่อนเหล่านั้นคำนวณหาความสนใจของผู้ใช้แทน

ay-fb-friend-rank อาศัยหลักการที่ชื่อว่า EdgeRank (Widman, n.d.) โดย EdgeRank เป็นอัลกอริทึมที่เฟสบุ๊คใช้คำนวณเพื่อเลือกโพสที่จะแสดงในหน้า newsfeed ของแต่ละผู้ใช้ โดยปัจจุบันเฟสบุ๊คได้ปรับปรุงอัลกอริทึมใหม่ซึ่งมีความซับซ้อนมากยิ่งขึ้น อย่างไรก็ตามอัลกอริทึมที่ปรับปรุงใหม่นั้นก็มีพื้นฐานมาจาก EdgeRank นั่นเอง โดย EdgeRank นั้นประกอบด้วย 3 ส่วน คือ 1) Affinity Score คือค่าการปฏิสัมพันธ์ที่มีระหว่างผู้ใช้ 2) Edge Weight คือค่าน้ำหนักที่ให้กับ Affinity Score แต่ละประเภท และ 3) Time Decay คือ

ค่าระยะเวลาที่โพสถูกสร้างขึ้น โดยอัลกอริทึม ay-fb-friend-rank นั้นใช้เพียง 2 ส่วนประกอบ คือ 1) Affinity Score และ 2) Edge Weight เนื่องจากต้องการหลีกเลี่ยงการคำนวณ Time Decay ที่ต้องใช้เวลานาน โดยอัลกอริทึม ay-fb-friend-rank อธิบายเพิ่มเติมได้ดังนี้

2.1) Affinity Score คือค่าการปฏิสัมพันธ์ที่มีระหว่างผู้ใช้และเพื่อน โดยในอัลกอริทึม ay-fb-friend-rank คำนวณจากการปฏิสัมพันธ์ 4 ประเภท ดังนี้

1. User's feed entries เช่น feed_like, feed_comment, feed_add ressed เป็นต้น คือจะให้ค่าคะแนนกับเพื่อนที่มีการกดชื่นชอบ แสดงความเห็นลงบนหน้าเฟสบุคของผู้ใช้ หรือติดแท็ก (tagging) ไปยังผู้ใช้

2. Photos เช่น photo_tagged_user_by_friend, photo_tagged_friend_by_user, photo_like, photo_comment เป็นต้น คือจะให้ค่าคะแนนกับเพื่อนที่มีการติดแท็กที่รูปภาพของผู้ใช้ หรือผู้ใช้ถูกติดแท็กไปยังรูปภาพของเพื่อน รวมถึงการที่เพื่อนหรือผู้ใช้กดชื่นชอบหรือแสดงความคิดเห็นบนรูปภาพเหล่านั้นด้วย

3. Friends เช่น friend_mutual เป็นต้น คือจะให้ค่าคะแนนไปยังเพื่อนที่มีเพื่อนร่วมกัน

4. Inbox เช่น inbox_in_conversation, inbox_chat เป็นต้น คือจะให้ค่าคะแนนไปยังเพื่อนที่มีการสนทนาระหว่างบุคคลกับผู้ใช้ (chat) หรือสนทนาเป็นกลุ่ม (conversation)

2.2) Edge Weights คือค่าน้ำหนักที่ให้กับแต่ละ Affinity Score โดยค่าน้ำหนักนี้แสดงถึงความสำคัญของปฏิสัมพันธ์แต่ละประเภทของ Affinity Score ซึ่งกำหนดให้ Affinity Score ที่มีค่า Edge Weights สูงกว่ามีความสำคัญมากกว่า อย่างไรก็ตามอัลกอริทึม ay-fb-friend-rank ให้ค่าน้ำหนักกับทุกประเภทเป็น 1 ทั้งหมด

เมื่อได้รายชื่อเพื่อนสนิทจากการคำนวณโดยอัลกอริทึมข้างต้นและกำหนดค่าปฏิสัมพันธ์ระหว่างผู้ใช้กับเพื่อนสนิททั้ง 5 แล้ว ระบบตรวจสอบว่ามีข้อมูลจากเพื่อนเพียงพอหรือไม่ (3.1.3.5) โดยหาก 3 ใน 5 มีข้อมูลเพียงพอ จะใช้ข้อมูลเช็คอินจากเพื่อนคำนวณค่าระดับความสนใจ (3.1.3.6) โดยวิธีการเช่นเดียวกับการคำนวณไปยังผู้ใช้ในสมการที่(2)และ(3) สุดท้ายใช้สมการที่ (4) คำนวณหาค่าระดับความสนใจของผู้ใช้จากค่าระดับความสนใจของเพื่อน (3.1.3.7)

$$I(c) = \frac{\sum_{i=1}^{5} F_i I_i(c)}{\sum_{i=1}^{5} F_i} \qquad (4)$$

$I(c)$ คือค่าระดับความสนใจ (Interest Level - I) ที่มีต่อหมวดหมู่ (Category - c)

F_i คือค่าปฏิสัมพันธ์ระหว่างผู้ใช้และเพื่อนคนที่ i

$I_i(c)$ คือค่าระดับความสนใจ (Interest Level - I) ที่มีต่อหมวดหมู่ (Category - c) ของเพื่อนคนที่ i

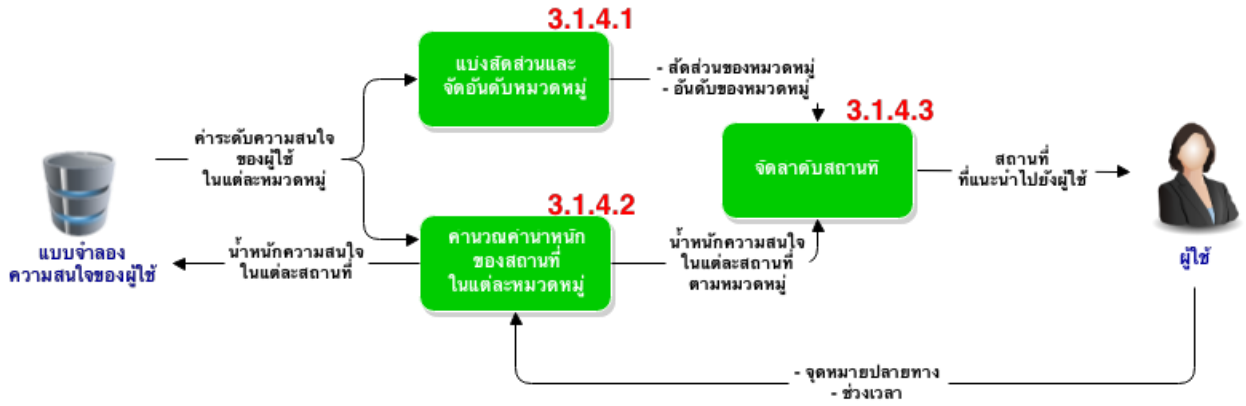
3) ใช้ข้อมูลของกลุ่มตัวอย่าง โดยจะใช้ในกรณีที่ข้อมูลจากเพื่อนไม่เพียงพอ จะใช้ค่าเฉลี่ยของค่าระดับความสนใจในแต่ละหมวดหมู่ที่คำนวณจากข้อมูลการเช็คอินของผู้ใช้กลุ่มตัวอย่าง โดยตัวอย่างจะเลือกจากผู้ใช้ที่มีเช็คอินที่ผ่านการตรวจสอบมากกว่า 10 เช็คอิน

โดยวิธีคำนวณค่าระดับความสนใจมีการทำงานดังอัลกอริทึมที่ 1 และค่าระดับความสนใจที่คำนวณได้ จะส่งไปยังกระบวนการถัดไป เพื่อใช้ในการแนะนำสถานที่ท่องเที่ยว

อัลกอริทึมที่ 1: วิธีคำนวณค่าระดับความสนใจของผู้ใช้จากการใช้ข้อมูลเช็คอิน

```
1 : Get check-in places of user.
2 : Delete redundant check-in places.
3 : Delete check-in places if not for travel purpose.
4 : Classify check-in places in category.
5 :   if all check-in places of user > 10
6 :       Calculate interest level from frequency of user check-in place
          appeared in each category.
7 :   else
8 :       Rank friends of user.
9 :       Get check-in places of top five closed friends.
10 :      if three friends of top five closed friends have check-in places > 10
11 :          Calculate interest level of top five closed friends from
frequency
12 :          of place appeared in each category.
13 :      else
14 :          Get average interest level from population.
Return  interest level.
```

3.5 การแนะนำข้อมูลการท่องเที่ยว (Recommendation)



รูปที่ 18 การแนะนำข้อมูลการท่องเที่ยว

กระบวนการในการแนะนำข้อมูล (รูปที่ 18) แบ่งออกเป็น 3 ขั้นตอน

1) การแบ่งสัดส่วนและจัดอันดับหมวดหมู่ ขั้นตอนนี้นำหลักการของ Content-based มาประยุกต์ใช้ โดยจะทำการแนะนำสถานที่ตามที่ผู้ใช้สนใจ แต่จะแบ่งสัดส่วนในการแนะนำตามค่าระดับความสนใจที่คำนวณได้ในกระบวนการก่อนหน้า หมวดหมู่ที่ผู้ใช้มีความสนใจสูงก็จะถูกแนะนำในสัดส่วนที่มากและถูกจัดอันดับไว้สูงกว่าหมวดหมู่ที่ผู้ใช้มีความสนใจน้อยกว่าโดยการคำนวณสัดส่วนแสดงดังสมการที่ (5)

$$R(c) = \frac{I(c)}{\sum_{k=1}^{6} I(k)} N \ , \ 1 \leq c \leq 6 \qquad (5)$$

$R(c)$ คือค่าสัดส่วนการแนะนำ (Ratio -R) ที่มีต่อหมวดหมู่ (Category - c)
$I(c)$ คือค่าระดับความสนใจ (Interest Level - I) ที่มีต่อหมวดหมู่ (Category - c)
N คือจำนวนของสถานที่ที่ต้องการแสดงไปยังผู้ใช้

2) คำนวณค่าน้ำหนักของสถานที่ในแต่ละหมวดหมู่ เป็นการคำนวณค่าน้ำหนักของสถานที่ที่อยู่ในแต่ละหมวดหมู่จากฐานข้อมูลสถานที่ท่องเที่ยว โดยเลือกพิจารณาเฉพาะปลายทางที่ผู้ใช้ระบุ เช่น ปลายทางคือจังหวัดพิษณุโลก ก็จะนำสถานที่ที่อยู่ในจังหวัดพิษณุโลกมาคำนวณเท่านั้น ซึ่งเกณฑ์การคำนวณแบ่งเป็น 3 ส่วน คือ 1) ความนิยมของสถานที่ 2) เป็นสถานที่ที่เพื่อนๆ เคยไป และ 3) เป็นช่วงเวลาที่นิยมไป ดังสมการที่ (6)

$$W(p) = [\alpha P(p) + \beta F(p) + \gamma T(p)] \prod IL(p, c_m) \qquad (6)$$

$W(p)$ คือค่าน้ำหนักของสถานที่ (Place Weight -W) ที่มีต่อสถานที่ (Place - p)
$P(p)$ คือค่าความนิยม (Popularity -P) ของสถานที่ (Place - p)

$F(p)$ คือค่าคะแนนของสถานที่ (Place - p) ที่เพื่อนสนิทในเฟสบุ๊คของผู้ใช้เคยไป

$T(p)$ คือค่าช่วงเวลาที่เหมาะสม (Time -T) ที่มีต่อสถานที่ (Place - p) โดยมีค่า =1 เมื่อมีสถานที่ (Place - p) เหมาะสมที่จะไปในช่วงเวลาที่ผู้ใช้กำหนด และมีค่า = 0 เมื่อไม่เหมาะสม

α คือค่าน้ำหนักของค่าความนิยม

β คือค่าน้ำหนักของการคำนวณสถานที่จากเพื่อน

γ คือค่าน้ำหนักของค่าเวลา

$I(p, c_m)$ คือค่าความสนใจ (Interest Level - I) จากสถานที่ (Place - p) ตามลำดับชั้นไปยังหมวดหมู่หลัก (Main Category - c_m)

ทั้งนี้ค่าความนิยม $P(p)$ คิดจากจำนวนการเช็คอินและการกดไลค์จากผู้ใช้ทั้งหมดบนเฟสบุ๊ค ซึ่งคำนวณดังสมการที่ (7)

$$P(p) = \alpha_{ch} \frac{n_{ch,p}}{\max(n_{ch,k})} + \alpha_{li} \frac{n_{li,p}}{\max(n_{li,k})} \quad (7)$$

$P(p)$ คือค่าความนิยม (Popularity -P) ของสถานที่ (Place - p)

α_{ch}, α_{li}คือค่าน้ำหนักของการเช็คอิน (Check-in - ch) หรือการกดไลค์ (Like - li) ที่มีต่อสถานที่ (Place - p)

$n_{ch,p}, n_{li,p}$คือจำนวนของการเช็คอิน (Check-in - ch) หรือการกดไลค์ (Like - li) ที่มีต่อสถานที่ (Place - p)

$n_{ch,k}, n_{li,k}$คือจำนวนของการเช็คอิน (Check-in - ch) หรือการกดไลค์ (Like - li) ที่มีต่อทุกๆ สถานที่ k

และค่าคะแนนของสถานที่ที่เพื่อนสนิทในเฟสบุ๊คของผู้ใช้เคยไป $F(p)$ คำนวณดังสมการที่ (8) โดยจะไม่ใช่ค่านี้คำนวณร่วมในสมการที่ (6) ในกรณีที่ใช้การคำนวณค่า $I(c)$ จากกลุ่มตัวอย่าง (3.1.3.8)

$$F(p) = \frac{\sum_{i=1}^{5} F_i I_i}{\sum_{i=1}^{5} F_i} \quad (8)$$

$F(p)$ คือ ค่าคะแนนของสถานที่ (Place - p) ที่เพื่อนสนิทในเฟสบุ๊คของผู้ใช้เคยไป

F_i คือ ค่าปฏิสัมพันธ์ระหว่างผู้ใช้และเพื่อนคนที่ i

I_i มีค่า =1 เมื่อมีสถานที่(Place - p) ปรากฏในการเช็คอินของเพื่อนคนที่ i และมีค่า = 0 เมื่อไม่ปรากฏ

กรณีที่เป็นผู้ใช้เดิมจะใช้ค่าน้ำหนักของสถานที่ $W(p)$ ที่เก็บไว้ในแบบจำลองความสนใจของผู้ใช้ที่คำนวณไว้แล้ว

3) จัดลำดับสถานที่ เป็นการจัดลำดับ (Rank) สถานที่ที่จะแนะนำไปยังผู้ใช้ โดยใช้ค่าสัดส่วนของหมวดหมู่ $R(C)$ และค่าน้ำหนัก $W(p)$ ที่คำนวณไปยังทุกสถานที่โดย ใช้กฎดังต่อไปนี้

1) ถ้า $R(c_1)$ มีค่าสูงกว่า $R(c_2)$ ดังนั้น c_1จะมีลำดับสูงกว่า c_2

2) ถ้าไม่สามารถทำตามข้อ 1) ได้แล้ว จะทำการสุ่มลำดับ

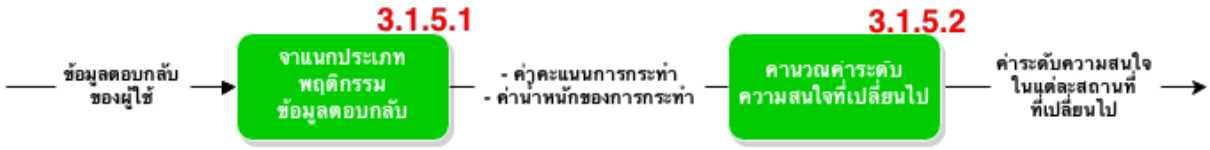
3) ในแต่ละ C ถ้า $W(p_1)$ มีค่าสูงกว่า $W(p_2)$ ดังนั้น p_1 จะมีลำดับสูงกว่า p_2

4) ถ้าไม่สามารถทำตามข้อ 3) ได้แล้ว จะทำการสุ่มลำดับ

กระบวนการแนะนำข้อมูลการท่องเที่ยวใช้อัลกอริทึมที่ 2 ดังนี้

อัลกอริทึมที่ 2 : การแนะนำสถานที่ท่องเที่ยว	
1 :	Calculate ratio of each category.
2 :	Calculate weight of each place from popularity, friend and time score.
3 :	popularity calculate from population check-in of this place.
4 :	friend calculate from friend check-in of this place.
5 :	time calculate from appropriate time to visit this place.
6 :	Rank places in each category from weight.
7 :	Recommend places from ratio of each category and places rank.

3.6 วิเคราะห์ความสนใจของผู้ใช้จากข้อมูลตอบกลับของผู้ใช้ (Relevance Feedback Analysis)



รูปที่ 19 การวิเคราะห์ข้อมูลตอบกลับของผู้ใช้

เมื่อระบบได้แนะนำสถานที่ท่องเที่ยวไปยังผู้ใช้แล้ว ระบบจะวิเคราะห์ข้อมูลตอบกลับสำหรับปรับปรุงแบบจำลองความสนใจของผู้ใช้ เพื่อแนะนำสถานที่ใหม่ที่เหมาะสม โดยกระบวนการนี้ (รูปที่ 19) จะวิเคราะห์ข้อมูลตอบกลับจากผู้ใช้หลายประเภททั้งแบบ Implicit และ Explicit เช่น การคลิกเข้าไปดูรายละเอียด (Implicit) การเลือกเป็นสถานที่ที่สนใจ (Explicit) เป็นต้น ซึ่งพฤติกรรมที่เป็นข้อมูลตอบกลับจากผู้ใช้ จะมีค่าคะแนนการกระทำและความเชื่อมั่นการกระทำ โดยค่าคะแนนการกระทำจะจัดกลุ่มเป็นการกระทำที่แสดงถึงความสนใจและไม่สนใจในสถานที่ที่แนะนำ ซึ่งประยุกต์จากอัลกอริทึม Rocchio กำหนดการกระทำที่แสดงถึงความสนใจจะให้ค่าคะแนนการกระทำเป็นบวก การกระทำที่แสดงถึงความไม่สนใจจะใช้ค่าคะแนนการกระทำเป็นลบ และประยุกต์ใช้เกณฑ์ค่าความเชื่อมั่นการกระทำจากระบบ Sigtur / E-Destination (Moreno et al., 2013) โดยกำหนดพฤติกรรมการตอบกลับที่เป็นแบบ Explicit จะมีค่าสูงกว่า Implicit ดัง

ตารางที่ 9

โดยเริ่มต้นจะจำแนกพฤติกรรมการตอบกลับว่าเป็นประเภทการกระทำแบบใด (3.1.5.1) เพื่อหาค่าคะแนนการกระทำและค่าความเชื่อมั่นของการกระทำนั้น

ตารางที่ 9 เกณฑ์ในการพิจารณาค่าระดับความสนใจและค่าน้ำหนักจากพฤติกรรมการตอบกลับของผู้ใช้

พฤติกรรมข้อมูลตอบกลับ	ค่าคะแนน การกระทำ (S)	ค่าความเชื่อมั่น การกระทำ (C)
1. การคลิกเข้าไปดูรายละเอียดของสิ่งที่แนะนำ (Implicit) เช่น ดูรายละเอียดของสถานที่ รูปภาพ เว็บไซต์ เป็นต้น	1.0	0.2
2. การเลือกสิ่งที่แนะนำเป็นสิ่งที่สนใจ (Explicit) เช่น การเลือกสถานที่เพื่อสร้างแผนการเดินทาง เป็นต้น	1.0	0.5
4. การเลือกไม่สนใจสิ่งที่แนะนำและขอให้แนะนำใหม่ (Explicit) เช่น การลบสถานที่ออกจากรายการสร้างแผนการเดินทาง การลบออกจากรายการแนะนำ เป็นต้น	-1.0	0.5

จากนั้นนำค่าทั้งสองมาคำนวณหาความสนใจที่เปลี่ยนไป (4.5.2) โดยคำนวณจากทุกๆ การกระทำไปยังแต่ละสถานที่ p ที่ระบบแนะนำไปยังผู้ใช้ ดังสมการที่ (9)

$$A(p) = \sum_{i=1}^{n_p} \frac{S_i C_i}{n_p} \qquad (9)$$

$A(p)$ คือค่าความสนใจที่เปลี่ยนไป (Adaptation Interest Score -A) ของผู้ใช้ (User - u) ที่มีต่อสถานที่ (Place - p)

n_p คือจำนวนของการกระทำของผู้ใช้ที่มีต่อสถานที่ (Place - p)

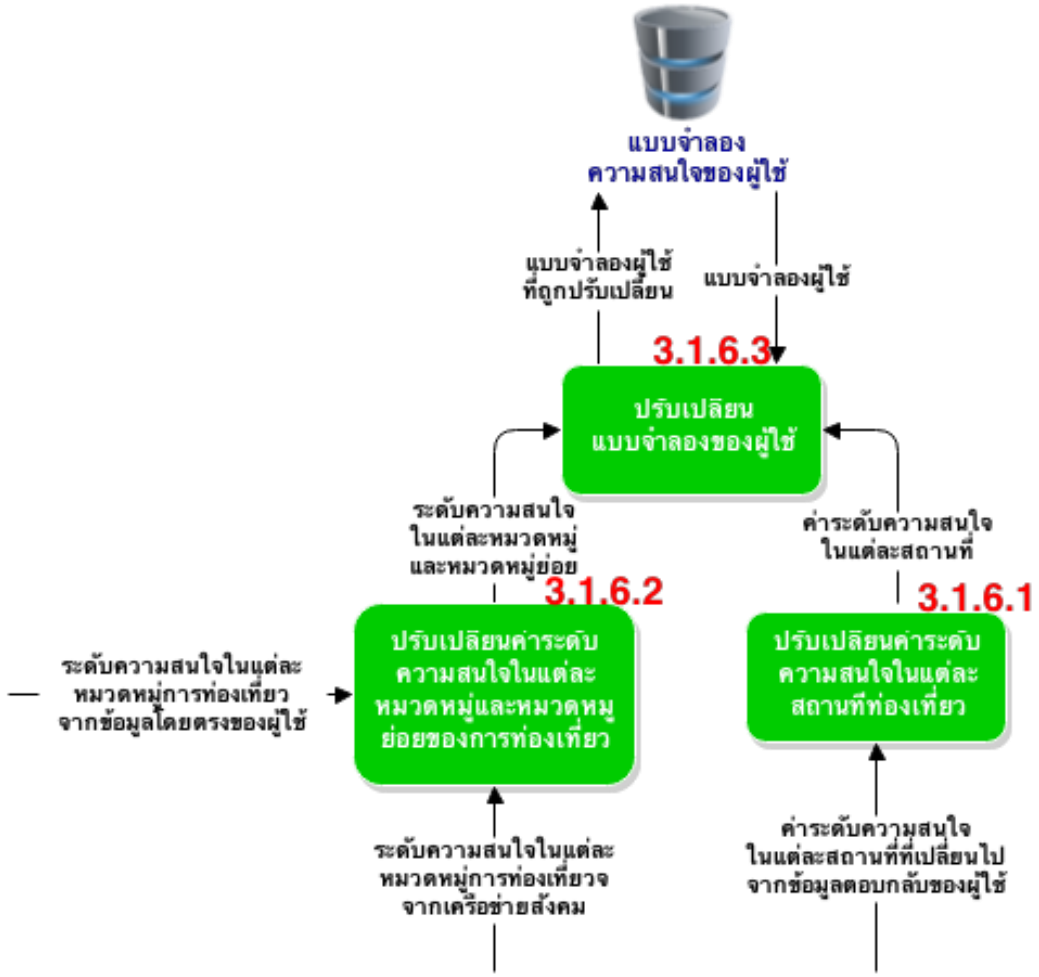
S_i คือค่าคะแนนการกระทำ (Action Score - S)

C_i คือค่าความเชื่อมั่นการกระทำ (Action Confidence - C)

เมื่อคำนวณค่าความสนใจที่เปลี่ยนไป $A(p)$ เรียบร้อยแล้ว ก็จะทำการปรับปรุงแบบจำลองความสนใจของผู้ใช้โดยอัตโนมัติในกระบวนการที่ 3.1.6 ดังนี้

3.7 ปรับเปลี่ยนแบบจำลองผู้ใช้โดยอัตโนมัติ (User Model Adaptation)

ขั้นตอนนี้เป็นกระบวนการปรับปรุงแบบจำลองผู้ใช้แสดงดังรูปที่ 20 โดยใช้การวิเคราะห์ความสนใจของผู้ใช้จากข้อมูลตอบกลับของผู้ใช้ในกระบวนการก่อนหน้า และจะปรับปรุงการแนะนำและความสนใจของผู้ใช้ทันทีเมื่อได้รับข้อมูลตอบกลับจากผู้ใช้ ซึ่งใช้หลักการของอัลกอริทึม Rocchio (Salton, 1971) ในการแบ่งกลุ่มพฤติกรรมตอบกลับที่ผู้ใช้มีต่อระบบ ว่าสนใจหรือไม่สนใจในสถานที่ที่แนะนำ และคำนวณผลรวมของค่าน้ำหนักของแต่ละกลุ่มพฤติกรรม ซึ่งได้ถูกคำนวณมาแล้วในขั้นตอนที่ 3.1.5 (วิเคราะห์ความสนใจของผู้ใช้จากข้อมูลตอบกลับของผู้ใช้) จากนั้นจะประยุกต์ใช้อัลกอริทึม Rocchio ในการปรับปรุงค่าไปยังคำแนะนำตั้งต้น



รูปที่ 20 ปรับเปลี่ยนแบบจำลองผู้ใช้โดยอัตโนมัติ

โดยในการปรับปรุงคำแนะนำจะใช้การปรับเปลี่ยนค่าน้ำหนักของสถานที่ (3.1.6.1) $W(p)$ ดังสมการที่ (10) และปรับปรุงความสนใจโดยการปรับเปลี่ยนค่าระดับความสนใจ (3.1.6.2) $I(C)$ ดังสมการที่ (11) เมื่อมีพฤติกรรมตอบกลับไปยังสถานที่นั้นๆ และหมวดหมู่ที่สถานที่นั้นอยู่ ซึ่งสำหรับค่าระดับความสนใจนั้นจะคำนวณขึ้นไปถึงระดับหมวดหมู่หลัก และปรับปรุงไปยังฐานข้อมูลแบบจำลองความสนใจของผู้ใช้ (3.1.6.3)

$$W'(p) = W(p) + A(p) \tag{10}$$

$W'(p)$ คือค่าน้ำหนักของสถานที่ (Place Weight -W) ของสถานที่ (Place - p) ที่ปรับปรุงใหม่

$W(p)$ คือค่าน้ำหนักของสถานที่ (Place Weight -W) ของสถานที่ (Place - p) เดิม

$A(p)$ คือค่าความสนใจที่เปลี่ยนไป (Adaptation Interest Level -A) ของผู้ใช้ที่มีต่อสถานที่ (Place - p)

$$I'(c) = I^u(c) + \alpha_c[A(p)\prod I(p,c)] \quad (11)$$

$I'(c)$ คือค่าระดับความสนใจ (Interest Level - I) ของผู้ใช้ที่มีต่อหมวดหมู่ (Main Category - c) ที่ปรับปรุงใหม่

$I^u(c)$ คือค่าระดับความสนใจ (Interest Level - I) ของผู้ใช้ที่มีต่อหมวดหมู่ (Main Category - c) เดิม

$A(p)$ คือค่าความสนใจที่เปลี่ยนไป (Adaptation Interest Level -A) ของผู้ใช้ที่มีต่อสถานที่ (Place - p)

α_c คือค่าน้ำหนักของหมวดหมู่ (Category - c)

$I(p,c)$คือค่าความสนใจ (Interest Level - I) ของผู้ใช้จากสถานที่ (Place - p) ตามลำดับชั้นไปยังหมวดหมู่ (Category - c)

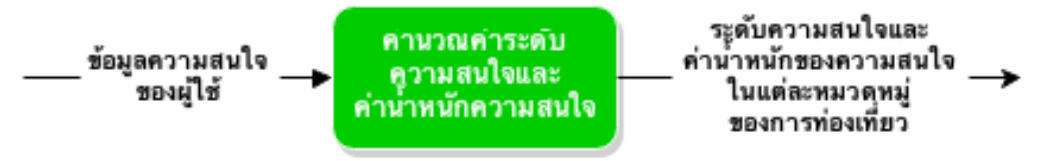
จากนั้นระบบจะไปยังกระบวนการแนะนำข้อมูลการท่องเที่ยว (3.1.4) อีกครั้งเพื่อแนะนำสถานที่ใหม่ไปยังผู้ใช้ โดยกระบวนการปรับเปลี่ยนแบบจำลองผู้ใช้โดยอัตโนมัติ แสดงดังอัลกอริทึมที่ 3

อัลกอริทึมที่ 3 : การปรับเปลี่ยนแบบจำลองผู้ใช้โดยอัตโนมัติ	
1 :	Calculate adaptation interest score from all user feedback of places.
2 :	Adapt weight of places and interest level of categories from adaptation interest score.
3 :	Rank places from new weight of places.
4 :	Calculate new ratio of category from new interest level of categories.
5 :	Recommend places from new ratio of each category and new places rank.

กระบวนการทำงานของระบบสำหรับผู้ใช้แบบผู้มาเยือน (Guest)

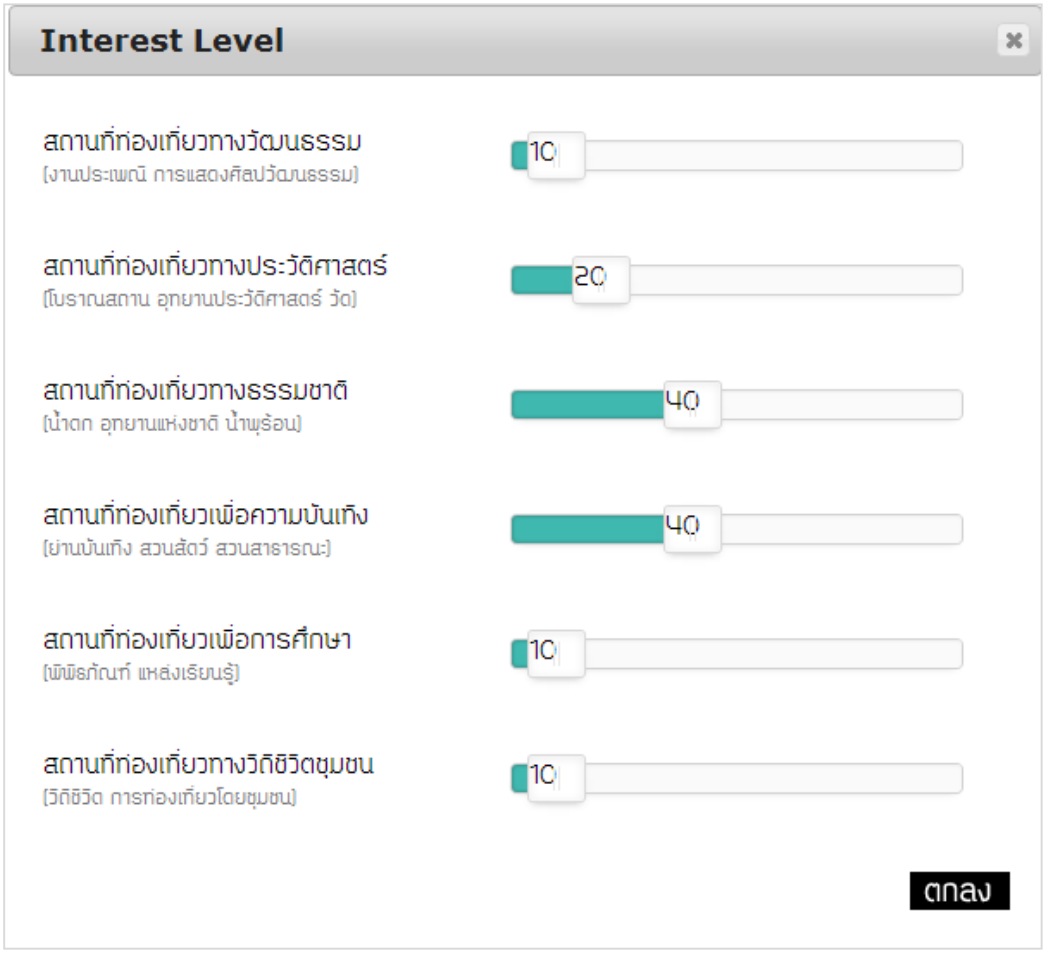
สำหรับผู้ใช้งานแบบผู้มาเยือนนั้น ระบบใช้กระบวนการ 2 ขั้นตอน โดยกระบวนการจะทำงานตามลำดับดังนี้

3.8 วิเคราะห์ความสนใจของผู้ใช้จากข้อมูลโดยตรงจากผู้ใช้



รูปที่ 21 การวิเคราะห์ความสนใจจากข้อมูลโดยตรงจากผู้ใช้

ผู้ใช้งานแบบผู้มาเยือนนั้นจะต้องให้ข้อมูลความสนใจกับระบบเอง จากนั้นคำนวณค่าระดับความสนใจในแต่ละหมวดหมู่ของสถานที่ท่องเที่ยว เพื่อส่งไปยังกระบวนการถัดไป ดังรูปที่ 21 โดยผู้ใช้จะต้องเลือกระดับความสนใจในสถานที่ท่องเที่ยวในรูปที่ 22 ซึ่งค่า 0 จะหมายถึงไม่ให้ความสนใจ แล้วไล่ระดับไปจนถึง 100 ซึ่งจะหมายถึงให้ความสนใจมากที่สุด ซึ่งหมวดหมู่ที่จะนำมาให้ผู้ใช้เลือกนั้นจะใช้หมวดหมู่ของสถานที่ท่องเที่ยว 6 หมวดหมู่ตามข้างต้น



รูปที่ 22 ระดับความสนใจในสถานที่ท่องเที่ยวของผู้ใช้

การวิเคราะห์ความสนใจจะใช้วิธีการเดียวกับการวิเคราะห์จากเครือข่ายสังคมโดยคำนวณหาค่าระดับความสนใจ ในกระบวนการนี้ ค่าระดับความสนใจคำนวณได้จากข้อมูลที่ผู้ใช้ให้มาโดยตรง เช่น หมวดหมู่ “สถานที่ท่องเที่ยวทางธรรมชาติ” ผู้ใช้เลือกเป็น 100% ค่าระดับความสนใจ = 1 หมวดหมู่ “สถานที่ท่องเที่ยวเพื่อความบันเทิง”ผู้ใช้เลือกเป็น 80% ค่าระดับความสนใจ = 0.8 เป็นต้น

เมื่อได้ค่าระดับความสนใจในแต่ละหมวดหมู่การท่องเที่ยวแล้ว จะส่งต่อไปยังกระบวนการ 3.1.4 ในรูปที่ 12 เพื่อแนะนำข้อมูลการท่องเที่ยวต่อไป ซึ่งจะใช้ขั้นตอนทั้ง 3 ขั้นตอนเช่นเดียวกับผู้ใช้แบบสมาชิก แต่ในการคำนวณค่าน้ำหนักของสถานที่ในแต่ละหมวดหมู่ในสมการที่ (6) จะไม่ใช้ค่า $F(p)$ ซึ่งเป็นค่าคะแนนที่คิดจากเพื่อนในเฟสบุ๊ค นอกจากนี้ผู้ใช้สามารถปรับเปลี่ยนระดับความสนใจใหม่โดยเลื่อนสไลด์บาร์ เพื่อให้ระบบเปลี่ยนแปลงการแนะนำได้

จากวิธีการต่างๆ ที่นำเสนอในบทนี้ผลลัพธ์ที่ได้คือระบบแนะนำสถานที่ท่องเที่ยวที่ตรงกับความสนใจของผู้ใช้ ในบทถัดไปจะทำการประเมินประสิทธิภาพในการแนะนำสถานที่ท่องเที่ยวต่างๆ ให้กับผู้ใช้

บทที่ 4

ผลการดำเนินการวิจัย

ในบทนี้จะอธิบายถึงการดำเนินการวิจัยโดยประกอบด้วย 3 ส่วน คือ 1) เครื่องมือที่ใช้ในการศึกษา 2) การวิเคราะห์ความต้องการของผู้ใช้ 3) ผลการออกแบบและพัฒนาระบบ และ 4) การประเมินประสิทธิภาพของระบบ ซึ่งในแต่ละส่วนสามารถอธิบายรายละเอียดได้ดังต่อไปนี้

4.1 เครื่องมือที่ใช้ในการศึกษา

เครื่องมือที่ใช้ในการศึกษาวิจัยในโครงการนี้ประกอบด้วย เครื่องมือในการสอบถามความต้องการของผู้ใช้ เครื่องมือในการประเมินผล และเครื่องมือที่ใช้ในการพัฒนาระบบ

4.1.1 เครื่องมือในการสอบถามความต้องการของผู้ใช้

เครื่องมือที่ใช้ในการสอบถามความต้องการของผู้ใช้งานแบ่งออกเป็น 2 รูปแบบ ได้แก่

1) การสัมภาษณ์เชิงลึก นำมาใช้เพื่อ
   - สอบถามความต้องการของนักท่องเที่ยวจากผู้เชี่ยวชาญด้านการท่องเที่ยว
   - วิเคราะห์และออกแบบระบบ
   - นำข้อมูลมาใช้พัฒนาแบบสอบถาม
2) การสำรวจโดยใช้แบบสอบถาม (Questionnaire survey)
   นำมาใช้เพื่อเก็บข้อมูลเชิงปริมาณสำหรับความต้องการของนักท่องเที่ยวชาวไทยต่อระบบที่กำลังพัฒนา (ระบบวางแผนการท่องเที่ยวและระบบนำเที่ยวบนอุปกรณ์พกพา)

4.1.2 เครื่องมือในการประเมินผล

เครื่องมือที่ใช้ในการประเมินผลการทำงานของระบบระบบวางแผนการเดินทางแบ่งออกเป็น 3 รูปแบบ ได้แก่

1) การใช้ตัววัดมาตรฐานที่ใช้ในการทดสอบประสิทธิภาพของระบบได้แก่ ค่าพรีซีชั่น
   ค่าพรีซีชั่น (Precision) เป็นค่าใช้วัดความถูกต้องของการแนะนำสถานที่ท่องเที่ยวว่าตรงกับความสนใจของผู้ใช้หรือไม่
2) แบบสอบถามหลังการใช้งานระบบ
   นำมาใช้เก็บข้อมูลความพึงพอใจหลังการทดลองใช้งานระบบ โดยแบ่งเป็นแบบสอบถามสำหรับการใช้งานเว็บแอปพลิเคชั่น

3) การสัมภาษณ์เชิงลึกและความคิดเห็นจากผู้เชี่ยวชาญด้านการท่องเที่ยวหลังการใช้งานเว็บแอปพลิเคชั่น

4.1.3 เครื่องมือที่ใช้ในการพัฒนาระบบ

เครื่องมือที่ใช้ในการดำเนินการพัฒนาระบบครั้งนี้ ประกอบไปด้วยฮาร์ดแวร์และซอฟต์แวร์ ดังนี้

1) ฮาร์ดแวร์
   - เครื่องคอมพิวเตอร์ส่วนตัว (Personal computer)
   - เครื่องเซิร์ฟเวอร์ (Server)
2) ซอฟต์แวร์
   - ภาษาที่ใช้ในการพัฒนาระบบ: PHP
   - โปรแกรมฐานข้อมูลเชิงพื้นที่ : PostgreSQL + PostGIS plugin
   - โปรแกรมช่วยออกแบบฐานข้อมูล : DBDesigner
   - โปรแกรมช่วยออกแบบเว็บไซต์ : Dreamweaver
   - โปรแกรมช่วยในการพัฒนาแอปพลิเคชั่นบนระบบปฏิบัติการแอนดรอยด์ : Eclipse + Android Development Tools (ADT) plugin
   - โปรแกรมเว็บเซิร์ฟเวอร์ : Apache Tomcat

4.1.4 วิธีการดำเนินการวิจัย

การพัฒนาระบบวางแผนการเดินทางและระบบนำเที่ยวบนอุปกรณ์พกพา มีขั้นตอนในการพัฒนาระบบทั้งหมด 6 ขั้นตอน ได้แก่ 1) วิเคราะห์ความต้องการของผู้ใช้ 2) การรวบรวมข้อมูล 3) วิเคราะห์และออกแบบระบบ 4) พัฒนาโปรแกรมและแอปพลิเคชั่น 5) ทดสอบการใช้งานและแก้ไขข้อบกพร่อง และ 6) ประเมินและสรุปการวิจัย แสดงดังรูปที่ 23

รูปที่ 23 ขั้นตอนการดำเนินงานวิจัย

4.2 การวิเคราะห์ความต้องการของผู้ใช้

4.2.1 วิธีการเก็บรวบรวมความต้องการ

การเก็บรวบรวมความต้องการของผู้ใช้ในงานวิจัยนี้กระทำโดยให้ผู้ใช้ตอบแบบสอบถามออนไลน์ (Online questionnaire) โดยมีจำนวนผู้ตอบแบบสอบถามทั้งสิ้นเป็นจำนวน 450 คน แบ่งเป็นเพศชาย 191 คน (42.2%) และเพศหญิง 259 คน (57.8%) โดยผู้ตอบแบบสอบถามส่วนใหญ่มีช่วงอายุระหว่าง 31-45 ปีคิดเป็น 42.7% และอายุระหว่าง 18-30 ปี คิดเป็น (38.7%) ซึ่งเป็นช่วงอายุที่ใช้คอมพิวเตอร์และอินเทอร์เน็ตในชีวิตประจำวันทุกวัน สำหรับผู้ตอบแบบสอบถามที่มีอายุเกิน 60 ปี มีจำนวนเพียง 6.7% เท่านั้นเนื่องจากคนในช่วงอายุดังกล่าวอาจจะไม่ใช้อินเทอร์เน็ตในชีวิตประจำวันมากนัก

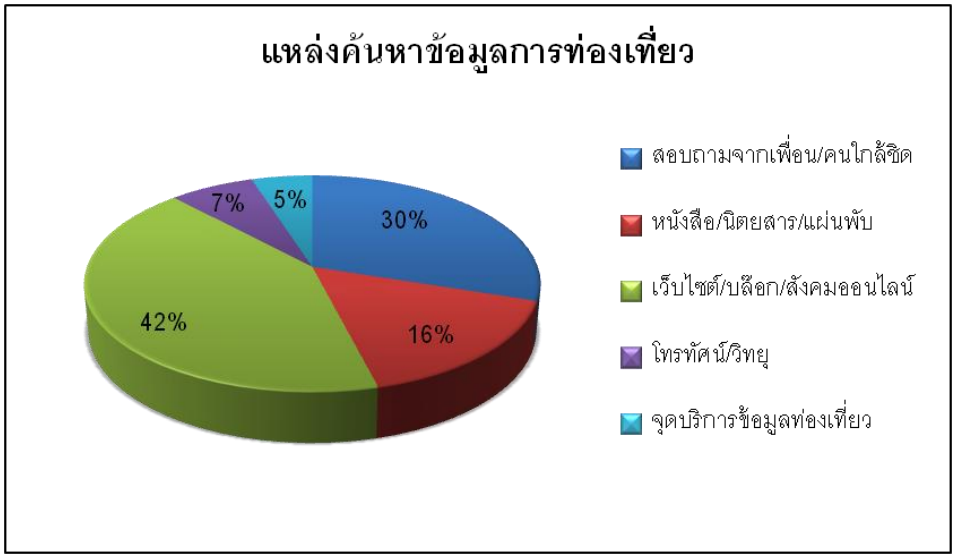
4.2.2 วิเคราะห์ความต้องการของผู้ใช้

จากการเก็บรวบรวมข้อมูลจากแบบสอบถามสามารถสรุปได้ดังต่อไปนี้

4.2.2.1 แหล่งข้อมูลการท่องเที่ยวที่ใช้ในการค้นหาข้อมูล

จากข้อมูลที่รวบรวมได้จะพบว่าแหล่งค้นหาข้อมูลที่สำคัญของผู้ตอบแบบสอบถามคือแหล่งข้อมูลบนเครือข่ายอินเทอร์เน็ต (รูปที่ 24) เช่น เว็บไซต์ต่างๆ บล็อก (Blog) และสังคมออนไลน์ ทั้งนี้เนื่องจากมีความสะดวก รวดเร็ว และให้ข้อมูลที่หลากหลายต่อนักท่องเที่ยว นอกจากนี้จากการสำรวจยังพบอีกว่า นักท่องเที่ยวมากกว่าร้อยละ 59% จะใช้เครือข่ายอินเทอร์เน็ตในการจองโรงแรม ซื้อตั๋วเครื่องบิน หรือธุรกรรมการท่องเที่ยวอื่นๆ ดังนั้นจึงสรุปได้ว่าเครือข่ายอินเทอร์เน็ตเป็นแหล่งข้อมูลที่สำคัญในการให้ข้อมูลข่าวสารด้านการท่องเที่ยวแก่นักเดินทางทั้งในและต่างประเทศ

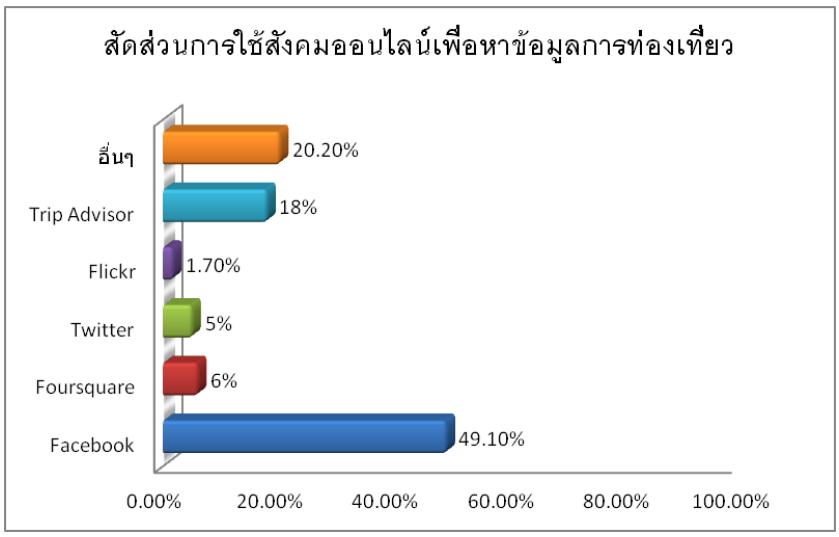
ข้อมูลที่น่าสนใจอีกประการหนึ่งคือ นอกจากแหล่งข้อมูลบนเครือข่ายอินเทอร์เน็ตแล้ว นักท่องเที่ยวส่วนมากมักจะถามข้อมูลการท่องเที่ยวจากเพื่อนๆ ที่เคยไปสถานที่ท่องเที่ยวต่างๆ มาก่อน (29.5%) ดังนั้นจะเห็นว่าข้อมูลจากเพื่อนที่เคยไปเที่ยวตามแหล่งท่องเที่ยวต่างๆ มาก่อนเป็นส่วนสำคัญในการตัดสินใจเดินทางไปสถานที่ต่างๆ ของตัวนักท่องเที่ยวเองเช่นกัน



รูปที่ 24 แหล่งค้นหาข้อมูลสถานที่ท่องเที่ยวต่างๆ

4.2.2.2 การใช้สังคมออนไลน์เพื่อเป็นแหล่งข้อมูลการท่องเที่ยว

จากการสอบถามกลุ่มตัวอย่างเกี่ยวกับการใช้สื่อสังคมออนไลน์ พบว่าผู้ตอบแบบสอบถามส่วนใหญ่ (84.7%) ใช้สังคมออนไลน์เพื่อหาข้อมูลท่องเที่ยว และในจำนวนสื่อสังคมออนไลน์ต่างๆ พบว่า เฟสบุ๊ค เป็นสังคมออนไลน์อันดับหนึ่งที่มีผู้ใช้มากที่สุด ประมาณ 49% (รูปที่ 25) รองลงมาคือ Trip Advisor (18%) ดังนั้นจึงสรุปได้ว่าเฟสบุ๊คจะเป็นแหล่งข้อมูลที่มีประโยชน์สำหรับการให้ข้อมูลการท่องเที่ยวกับผู้ใช้ได้



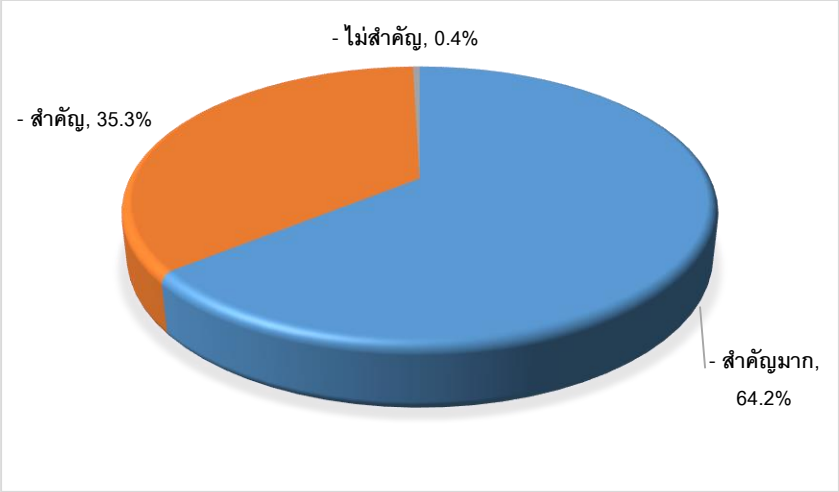
รูปที่ 25 สัดส่วนการใช้สังคมออนไลน์เพื่อหาข้อมูลการท่องเที่ยว

นอกจากนี้ผู้ตอบแบบสอบถามส่วนใหญ่ (92.7%) ยังสนใจสถานที่ท่องเที่ยวต่างๆ ที่เพื่อนในกลุ่มเคยไปมาและสนใจที่จะไปสถานที่ดังกล่าว ดังนั้นหากสามารถตรวจสอบได้ว่าเพื่อนๆ ของนักท่องเที่ยวเคยไปที่ไหนบ้างและนำมาแนะนำแก่นักท่องเที่ยว จะทำให้ระบบสามารถแนะนำข้อมูลได้ตรงกับความต้องการของนักท่องเที่ยวมากยิ่งขึ้น ซึ่งสถานที่ที่เพื่อนๆ เคยไปนี้สามารรถช่องทางหนึ่งที่สามารถตรวจสอบได้คือผ่านช่องทางสื่อสังคมออนไลน์นั่นเอง และสถิตินี้มีความสอดคล้องกับหัวข้อ 2.1 นั่นคือข้อมูลที่ได้จากเพื่อนๆ เป็นข้อมูลสำคัญที่นักท่องเที่ยวใช้เพื่อตัดสินใจว่าจะไปสถานที่ต่างๆ หรือไม่

4.2.2.3 การแนะนำข้อมูลสถานที่ท่องเที่ยวเฉพาะบุคคล

นอกจากนี้ผู้ตอบแบบสอบถามยังมีความเห็นว่า ระบบที่ใช้ข้อมูลการท่องเที่ยวควรสามารถให้ข้อมูลที่ตรงกับความต้องการของนักท่องเที่ยวมากที่สุด โดยเห็นว่าระบบดังกล่าวมีความสำคัญมากถึง 64.2% และมีเพียง 0.4% ที่คิดว่าระบบดังกล่าวไม่สำคัญ (รูปที่ 26

รูปที่ 26) ดังนั้นจึงสรุปได้ว่า ผู้ใช้เห็นว่าการพัฒนาระบบให้ข้อมูลการท่องเที่ยวเฉพาะบุคคลมีความสำคัญและเป็นประโยชน์ต่อนักท่องเที่ยวอย่างยิ่ง



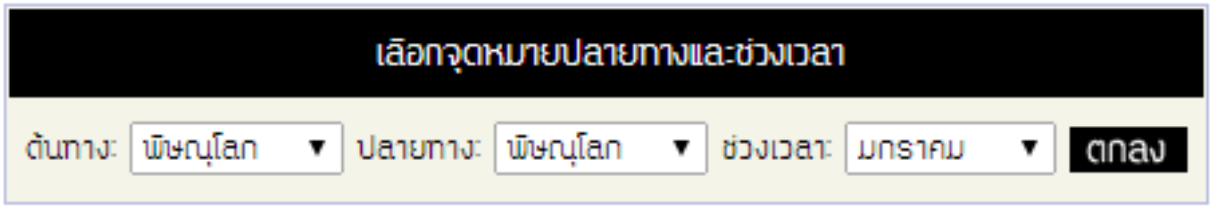
รูปที่ 26 ความสำคัญในการพัฒนาระบบการให้ข้อมูลการท่องเที่ยวเฉพาะบุคคล

4.3 ผลการออกแบบและพัฒนาระบบ

การแนะนำสถานที่จากระบบจะแบ่งการทำงานออกเป็น 2 แบบ ได้แก่ ระบบผู้มาเยือนและระบบสมาชิก

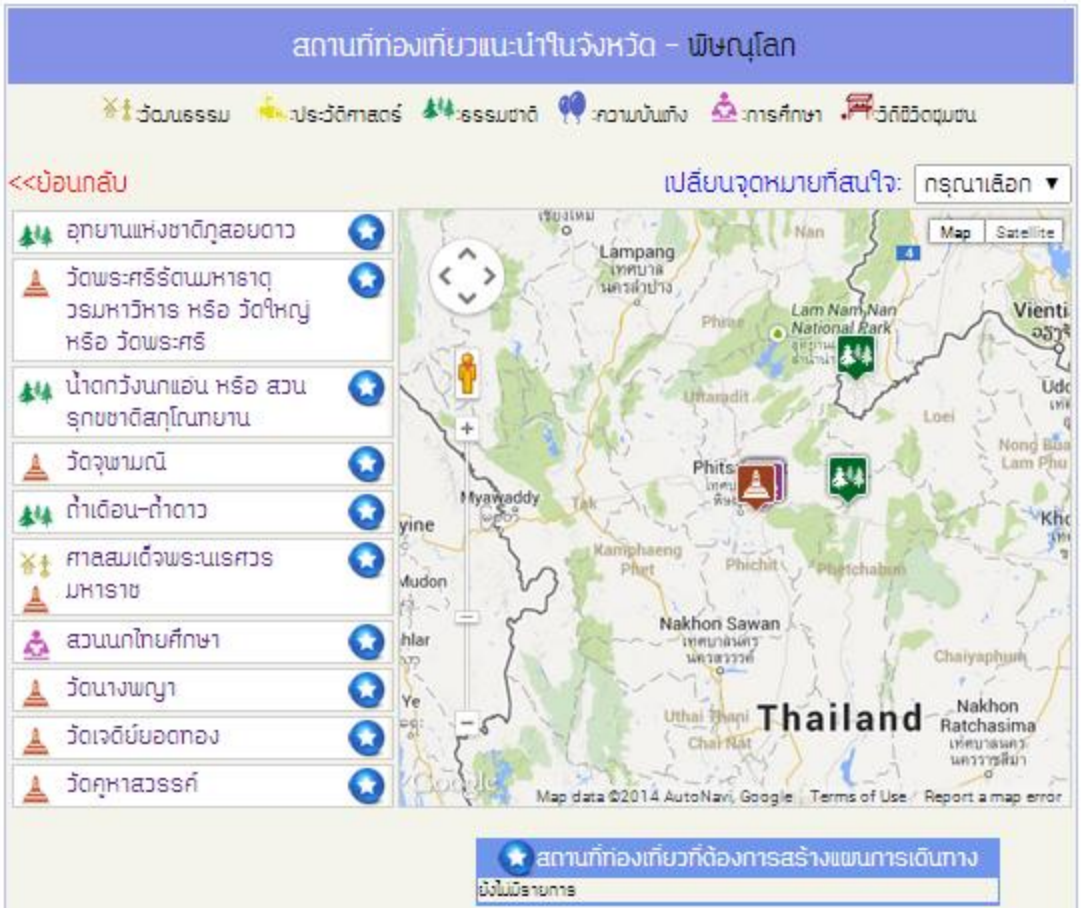
4.3.1 การแนะนำสถานที่จากระบบสำหรับผู้ใช้แบบผู้มาเยือน (Guest)

ผู้ใช้สามารถใช้ระบบแนะนำสถานที่ได้ โดยไม่ต้องล็อกอิน โดยระบบจะเก็บข้อมูลการแนะนำไว้จนกระทั่งผู้ใช้ปิดเว็บบราวเซอร์ ยกตัวอย่างเช่น เริ่มต้นคือเลือกจุดหมายปลายทางและช่วงเวลา ดังรูปที่ 27



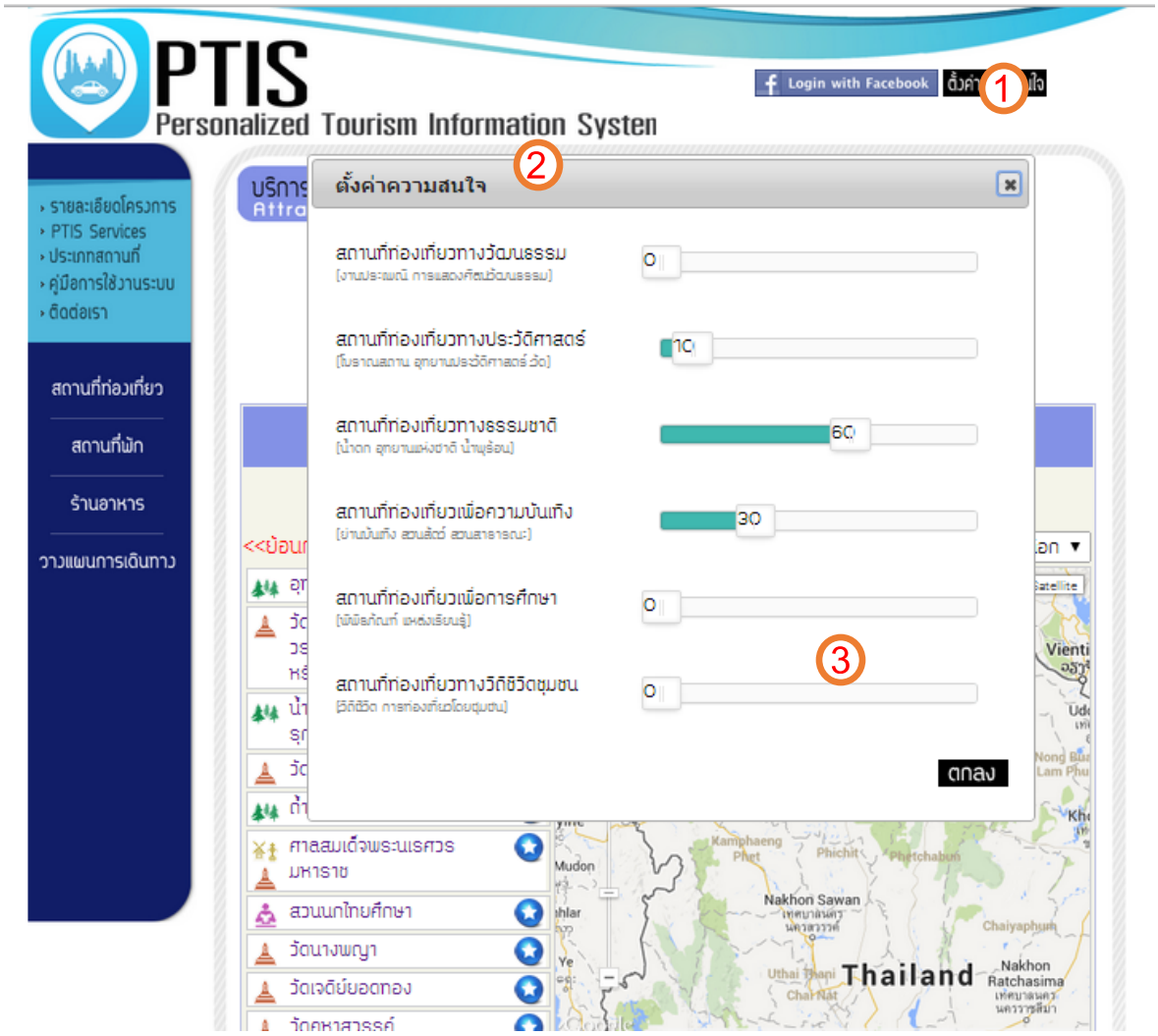
รูปที่ 27 การคลิกเพื่อเลือกจุดหมายปลายทางและช่วงเวลา

จากนั้นระบบจะแนะนำสถานที่โดยวิเคราะห์ความสนใจจากค่าเฉลี่ยของกลุ่มตัวอย่างและ จุดหมายปลายทางและช่วงเวลาที่ระบุจากขั้นตอนก่อนหน้า ดังรูปที่ 28



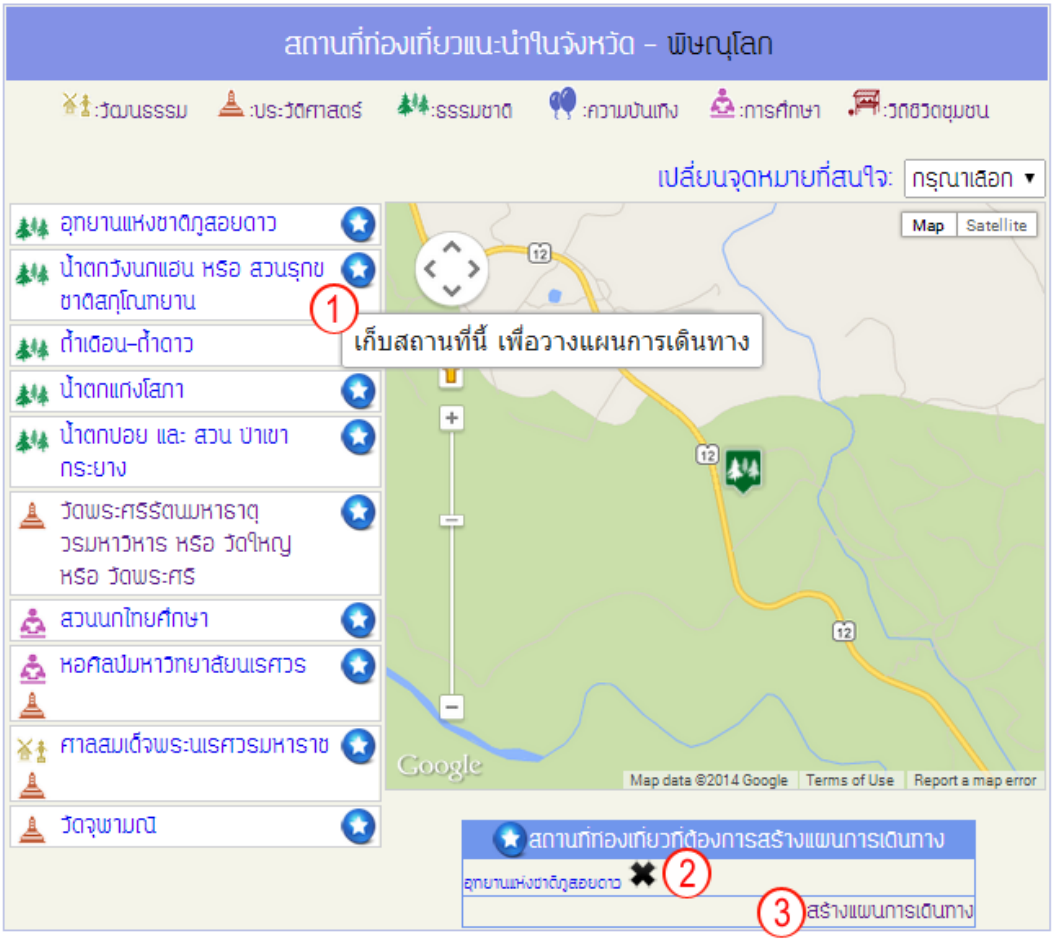
รูปที่ 28 หน้าจอหน้าจอแนะนำสถานที่ท่องเที่ยว

โดยผู้ใช้สามารถปรับค่าความสนใจ เพื่อให้ระบบวิเคราะห์และแนะนำใหม่ได้ โดยคลิกที่ “ตั้งค่าความสนใจ” จากนั้นทำการปรับเปลี่ยนความสนใจในสไลด์บาร์ และคลิก “ตกลง” ดังรูปที่ 29



รูปที่ 29 หน้าจอการปรับค่าความสนใจ

นอกจากนี้ผู้ใช้สามารถเลือกสถานที่เพื่อนำไปสร้างแผนการเดินทางได้ ดังรูปที่ 30 โดยคลิกที่ปุ่มรูปดาวสถานที่ที่ต้องการสร้างแผนการเดินทางจะปรากฏในตารางด้านล่าง และสามารถลบออกได้โดยคลิกปุ่มกากบาทในสถานที่ที่ต้องการนำออก

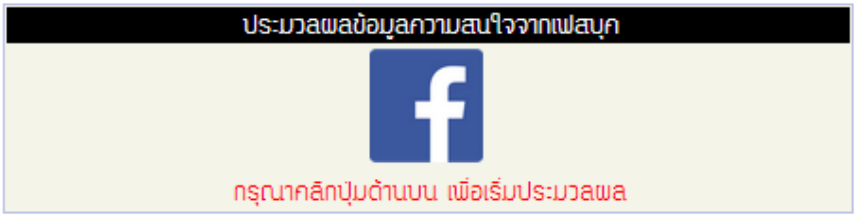


รูปที่ 30 เลือกสถานที่เพื่อสร้างแผนการเดินทาง

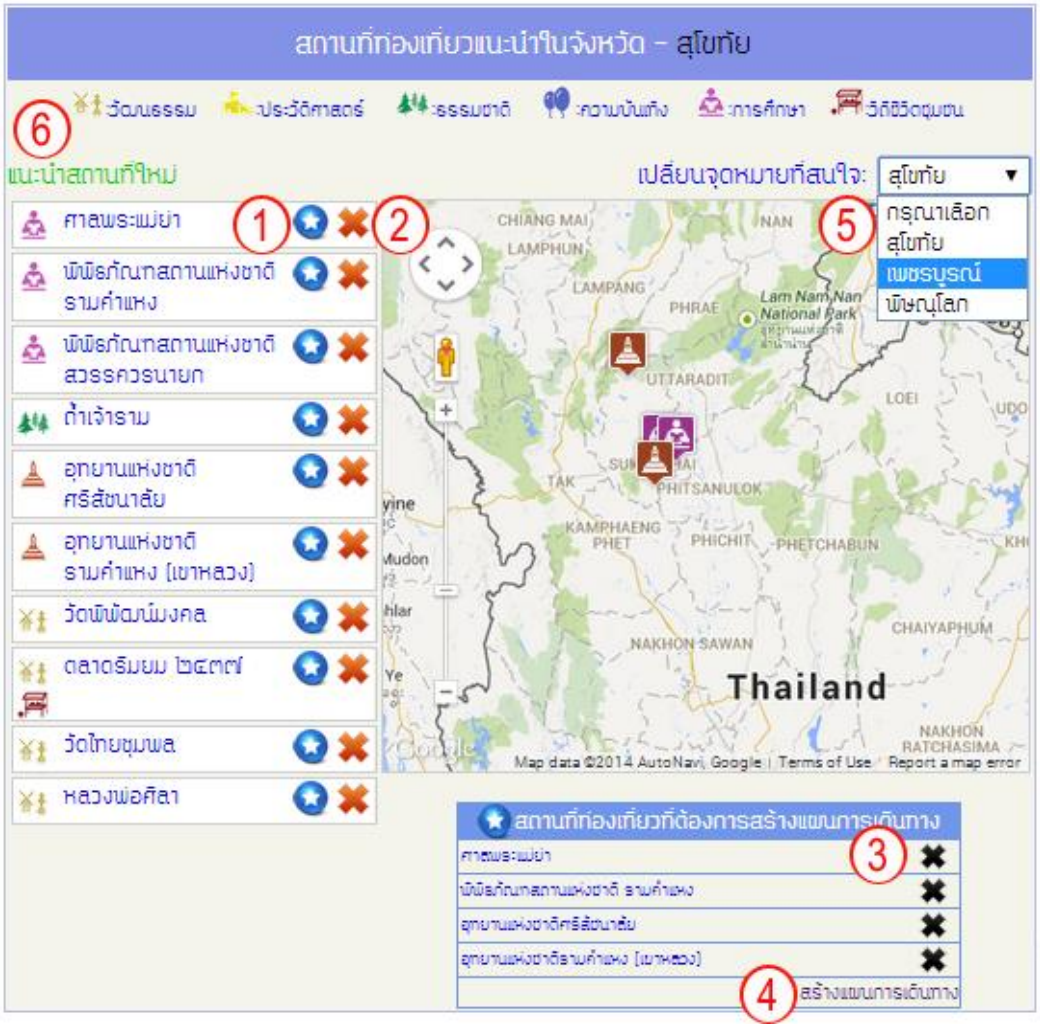
4.3.1.1 การแนะนำสถานที่จากระบบสำหรับผู้ใช้แบบสมาชิก (Member)

ผู้ใช้ที่มีบัญชีเฟสบุ๊คและล็อกอินเข้าสู่ระบบ ระบบจะให้คำแนะนำสถานที่โดยสมบูรณ์ โดยระบบจะวิเคราะห์ความสนใจของผู้ใช้จากเครือข่ายสังคมออนไลน์ จากนั้นจะใช้ข้อมูลตอบกลับที่ผู้ใช้ทำกับระบบ เช่น การคลิกดูรายละเอียด การเลือกเป็นสถานที่ที่ต้องการสร้างแผนการเดินทาง เป็นต้น ปรับปรุงคำแนะนำและความสนใจของผู้ใช้ใหม่เพื่อให้ตรงกับความต้องการของผู้ใช้มากยิ่งขึ้น ข้อมูลความสนใจต่างๆ ระบบจะเก็บไปยังแบบจำลองผู้ใช้ เมื่อผู้ใช้กลับมาล็อกอินเข้าสู่ระบบอีกครั้งระบบจะดึงข้อมูลจากแบบจำลองผู้ใช้มาแนะนำสถานที่ได้ทันที

ในกรณีที่ผู้ใช้เริ่มต้นใช้งานระบบครั้งแรก ระบบจะให้ผู้ใช้ประมวลผลข้อมูลจากเฟสบุ๊คเพื่อวิเคราะห์ระดับความสนใจในรูปที่ 31 เมื่อประมวลผลเรียบร้อยแล้ว ทำการเลือกจุดหมายปลายทางและช่วงเวลา ที่ต้องการให้แนะนำ จากนั้นระบบจะแนะนำสถานที่โดยวิเคราะห์ความสนใจจากระดับความสนใจที่วิเคราะห์มาจากเฟสบุ๊ค ดังรูปที่ 32



รูปที่ 31 ปุ่มเริ่มประมวลผลข้อมูลจากเฟสบุ๊ค



รูปที่ 32 ระบบแนะนำสถานที่ท่องเที่ยวไปยังผู้ใช้

ทั้งนี้ผู้ใช้สามารถเปลี่ยนแปลงค่าระดับความสนใจเหล่านี้ใหม่ได้ โดยคลิกที่ปุ่ม “ตั้งค่าความสนใจ” เช่นกัน ซึ่งการเปลี่ยนแปลงค่าก็จะทำให้แบบจำลองเปลี่ยนไป

นอกจากนี้ผู้ใช้สามารถเลือกสถานที่เพื่อสร้างแผนการเดินทางโดยคลิกปุ่ม favorite (รูปดาว) หรือกดปุ่ม no (รูปกากบาท) หมายถึง ไม่สนใจในสถานที่นี้ ต้องการให้ระบบแนะนำสถานที่อื่นๆ โดยพฤติกรรมที่ผู้ใช้มีต่อระบบเหล่านี้เป็นข้อมูลตอบกลับที่จะนำไปคำนวณเพื่อปรับปรุงค่าความสนใจในแต่ละสถานที่ของผู้ใช้ และแนะนำสถานที่ใหม่ให้ได้ตรงกับความต้องการของผู้ใช้ยิ่งขึ้น

4.4 การประเมินประสิทธิภาพของระบบทดสอบระบบ

ผู้วิจัยได้ออกแบบการทดสอบระบบเพื่อพิสูจน์สมมติฐานที่ตั้งไว้ โดยใช้การทดสอบทั้งหมด 4 แบบดังนี้

4.4.1 การทดสอบความสนใจเฉพาะบุคคลของผู้ใช้ (User Interest Evaluation)

การทดสอบในส่วนนี้ เป็นการทดสอบว่าระบบสามารถใช้เครือข่ายสังคมเพื่อวิเคราะห์หาความสนใจของแต่ละบุคคลได้หรือไม่ มีความถูกต้องในการคาดการณ์ความสนใจมากน้อยเพียงใด

ผู้วิจัยออกแบบการทดลองเพื่อประเมินความถูกต้องของระบบในการลำดับความสนใจในหมวดหมู่ของสถานที่ท่องเที่ยว 6 หมวดหมู่ โดยใช้ข้อมูลลำดับความสนใจแท้จริงที่ได้จากผู้ใช้ 40 คน ที่ล็อกอินเข้ามาใช้งานในระบบ เปรียบเทียบกับลำดับความสนใจที่ระบบวิเคราะห์ได้จากข้อมูลในเครือข่ายสังคมของผู้ใช้เอง

ข้อมูลลำดับความสนใจแท้จริง ได้จากการขอข้อมูลจากผู้ใช้โดยตรง โดยให้แต่ละผู้ใช้กรอกข้อมูลผ่านแบบสอบถามออนไลน์

PTIS
Personalized Travelling Information Services

ความคิดเห็นเกี่ยวกับความชื่นชอบในสถานที่ท่องเที่ยว

1.ท่านชอบไปสถานที่ท่องเที่ยวใดมากที่สุดคะ
เรียงลำดับสถานที่ที่ท่านชอบมากที่สุด (ชอบมากที่สุด = 1)

ประเภทสถานที่ท่องเที่ยว	ลำดับความชื่นชอบ					
	1	2	3	4	5	6
ศูนย์/ตลาดสินค้าพื้นเมือง	○	○	○	○	○	○
วัด โบราณสถาน	○	○	○	○	○	○

รูปที่ 33 แบบสอบถามลำดับความสนใจ

จากนั้นนำข้อมูลลำดับความสนใจแท้จริง เปรียบเทียบกับลำดับความสนใจที่ระบบวิเคราะห์ได้จากข้อมูลในเครือข่ายสังคมของผู้ใช้เอง โดยตารางที่ 10 เป็นการเปรียบเทียบของผู้ใช้ 1 คน

ตารางที่ 10 เปรียบเทียบลำดับความสนใจในหมวดหมู่สถานที่ท่องเที่ยวที่แท้จริงกับลำดับที่วิเคราะห์โดยระบบของผู้ใช้ 1 คน

หมวดหมู่สถานที่ท่องเที่ยว	ลำดับความสนใจ		ความถูกต้องในการลำดับความสนใจ
	ข้อมูลที่วิเคราะห์โดยระบบ	ข้อมูลจากผู้ใช้จริง	
สถานที่ท่องเที่ยวทางวัฒนธรรม	5	4	✗
สถานที่ท่องเที่ยวทางประวัติศาสตร์	2	2	✓
สถานที่ท่องเที่ยวเพื่อความบันเทิง	3	3	✓
สถานที่ท่องเที่ยวทางธรรมชาติ	1	1	✓
สถานที่ท่องเที่ยวเพื่อการศึกษา	4	6	✗
สถานที่ท่องเที่ยวทางวิถีชีวิตชุมชน	5	5	✓

จากการทดลองพบว่าระบบสามารถวิเคราะห์หาความสนใจของผู้ใช้จากข้อมูลในเครือข่ายสังคมของผู้ใช้ โดยได้ค่าเฉลี่ยความถูกต้องเท่ากับ 67.91% นอกจากนี้ยังพบว่ามีความถูกต้องในการลำดับความสนใจใน 3 อันดับแรก 77.5% อย่างไรก็ตามปัจจัยหนึ่งที่มีผลต่อความถูกต้องในการลำดับความสนใจคือการใช้ข้อมูลจากเครือข่ายสังคมตลอดช่วงเวลาตั้งแต่เริ่มต้นใช้เครือข่ายสังคมจนถึงปัจจุบัน ไม่ได้เลือกช่วงเวลาเฉพาะที่ผู้ใช้ต้องการไปยังสถานที่ท่องเที่ยว ซึ่งอาจทำให้ค่าความถูกต้องสูงขึ้นได้

4.4.2 การทดสอบการแก้ไขปัญหาโคลด์สตาร์ท (Cold Start Problem Evaluation)

การทดสอบนี้เป็นการวัดประสิทธิภาพในการแก้ไขปัญหาโคลด์สตาร์ท หรือกรณีที่ผู้ใช้มีข้อมูลการเช็คอินไม่เพียงพอสำหรับวิเคราะห์ความสนใจ ระบบจะใช้ข้อมูลการเช็คอินจากเพื่อนของผู้ใช้ในเฟสบุ๊คที่มีการปฏิสัมพันธ์กับผู้ใช้ในระดับสูงแทน โดยวัดประสิทธิภาพจากค่าความถูกต้อง (Accuracy) ของลำดับความสนใจที่วิเคราะห์ได้จากเพื่อนผู้ใช้ เปรียบเทียบกับข้อมูลของผู้ใช้เอง

ในการทดลอง จะใช้ผู้ใช้ที่มีข้อมูลเพียงพอให้ระบบวิเคราะห์ความสนใจได้ จำนวน 20 บัญชีผู้ใช้ โดยให้ระบบลำดับความสนใจในหมวดหมู่ของสถานที่ท่องเที่ยว 6 หมวดหมู่จากข้อมูลของผู้ใช้เองในเครือข่ายสังคม จากนั้นได้ลบข้อมูลในเครือข่ายสังคมออกเพื่อให้มีข้อมูลไม่เพียงพอและเกิดปัญหาโคลด์สตาร์ทขึ้น ทำให้ระบบต้องใช้ข้อมูลของเพื่อนผู้ใช้ในการวิเคราะห์เพื่อสร้างลำดับความสนใจแทน โดยตารางที่ 11 แสดงลำดับความสนใจในหมวดหมู่สถานที่ท่องเที่ยวของผู้ใช้ A เปรียบเทียบกับกลุ่มเพื่อน (1 หมายถึงมีความสนใจมากที่สุด ลำดับไปจนถึง 6 หมายถึงมีความสนใจน้อยที่สุด)

ตารางที่ 11 ลำดับความสนใจในหมวดหมู่สถานที่ท่องเที่ยวของผู้ใช้ A เปรียบเทียบกับกลุ่มเพื่อน

ลำดับหมวดหมู่สถานที่ท่องเที่ยว	ลำดับความสนใจ		ความถูกต้องในการลำดับความสนใจ
	ข้อมูลจากผู้ใช้A	ข้อมูลจากกลุ่มเพื่อนผู้ใช้A	
สถานที่ท่องเที่ยวทางวัฒนธรรม	3	3	✓
สถานที่ท่องเที่ยวทางประวัติศาสตร์	1	1	✓
สถานที่ท่องเที่ยวเพื่อความบันเทิง	5	4	✗
สถานที่ท่องเที่ยวทางธรรมชาติ	2	2	✓
สถานที่ท่องเที่ยวเพื่อการศึกษา	4	5	✗
สถานที่ท่องเที่ยวทางวิถีชีวิตชุมชน	6	6	✓

จากการทดลองพบว่าระบบสามารถวิเคราะห์หาความสนใจของผู้ใช้จากข้อมูลของกลุ่มเพื่อนได้ค่าเฉลี่ยความถูกต้องเท่ากับ 65.83% นอกจากนี้ยังพบว่ามีความถูกต้องในการลำดับความสนใจใน 3 อันดับแรก 75% อย่างไรก็ตามปัจจัยหนึ่งที่มีผลต่อความถูกต้องในการลำดับความสนใจคือเกณฑ์การเลือกกลุ่มเพื่อนสนิทจากอัลกอริทึม ay-fb-friend-rank (gajus, 2012) ที่ไม่ได้นำค่า Time Decay มาใช้ เนื่องจากข้อจำกัดด้านเวลาในการประมวลผลของระบบ ดังนั้นการคำนวณเพื่อเลือกกลุ่มเพื่อนสนิทจึงใช้การพิจารณาปฏิสัมพันธ์ระหว่างกันตลอดช่วงเวลาตั้งแต่เริ่มต้นใช้เครือข่ายสังคมจนถึงปัจจุบัน ไม่ได้เลือกกลุ่มเพื่อนจากที่มีการปฏิสัมพันธ์มากขณะปัจจุบันหรือช่วงเวลาเฉพาะที่ผู้ใช้ต้องการไปยังสถานที่ท่องเที่ยว ซึ่งอาจทำให้ค่าความถูกต้องสูงขึ้นได้

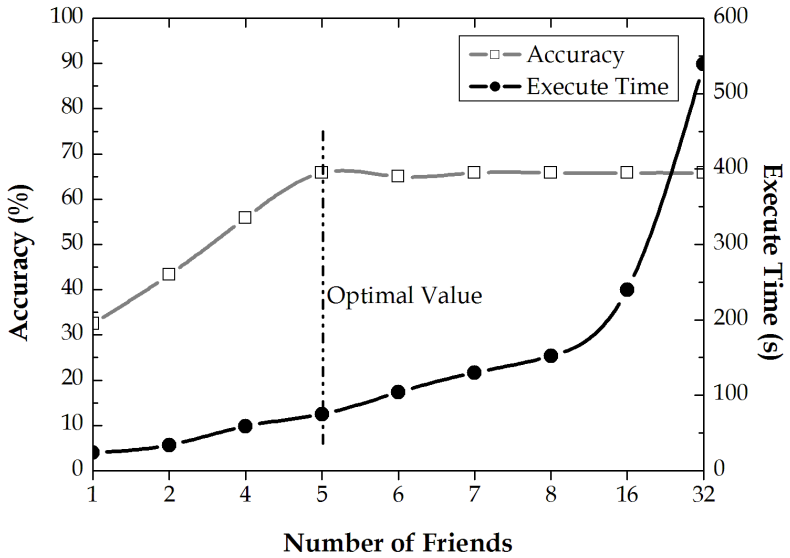
นอกจากนี้ในส่วนของการใช้ข้อมูลของเพื่อนเพื่อแก้ไขปัญหาโคลด์สตาร์ทในการทดสอบที่ 1 ผู้วิจัยได้ทดลองถึงการเลือกใช้จำนวนของเพื่อนที่เหมาะสมเพื่อนำมาวิเคราะห์ความสนใจแทนข้อมูลจากผู้ใช้ โดยพิจารณาจากมีความถูกต้องมากที่สุดในการลำดับหมวดหมู่ความสนใจสถานที่ท่องเที่ยว และเวลาใช้เวลาน้อยที่สุดในการประมวลผลเพื่อหาลำดับหมวดหมู่ความสนใจ

เนื่องจากไม่มีวิธีการอย่างเป็นทางการในการเลือกจำนวนเพื่อนที่เหมาะสมข้างต้น ผู้วิจัยจึงใช้วิธี grid search (Cao and Tay, 2003) ซึ่งจะกำหนดค่าจำนวนเพื่อนเป็นแบบเอกซ์โพเนนเชียล โดยใช้ค่าตั้งแต่ 2^0 - 2^5 การทดลองทำให้ทราบว่าลำดับความสนใจในหมวดหมู่ของสถานที่ท่องเที่ยวมีลำดับที่ต่างกันเมื่อกำหนดค่าจำนวนเพื่อนเท่ากับ 2^0 - 2^2 แต่จะได้ลำดับความสนใจที่เหมือนกันเมื่อกำหนดค่าจำนวนเพื่อนเท่ากับ 2^3 - 2^5 จากนั้นผู้วิจัยได้ทดลองเพิ่มเติมโดยใช้ค่าจำนวนเพื่อนระหว่าง 2^2 และ 2^3 คือ 5 6 และ 7 ซึ่งผลการทดลองแสดงดังตารางที่ 12 โดยการทดสอบของผู้ใช้หนึ่งคน จะเห็นว่าลำดับหมวดหมู่ความสนใจสถานที่ท่องเที่ยวมีลำดับที่เหมือนกันเมื่อเลือกใช้ข้อมูลจากเพื่อนจำนวน 5 คน ขึ้นไป เมื่อพิจารณาเวลาที่ใช้ในการประมวลผล จะเห็นว่าเมื่อใช้เพื่อนจำนวนมากขึ้นก็จะใช้เวลาในการประมวลผลนานขึ้น

ตารางที่ 12 ลำดับความสนใจในหมวดหมู่ของสถานที่ท่องเที่ยวและเวลาในการประมวลผลของหนึ่งผู้ใช้

ลำดับหมวดหมู่สถานที่ท่องเที่ยว	จำนวนเพื่อนที่ใช้คำนวณลำดับหมวดหมู่ความสนใจ								
	2^0	2^1	2^2	5	6	7	2^3	2^4	2^5
สถานที่ท่องเที่ยวทางวัฒนธรรม	5	6	6	6	6	6	6	6	6
สถานที่ท่องเที่ยวทางประวัติศาสตร์	2	2	3	3	3	3	3	3	3
สถานที่ท่องเที่ยวเพื่อความบันเทิง	1	1	1	1	1	1	1	1	1
สถานที่ท่องเที่ยวทางธรรมชาติ	3	3	2	2	2	2	2	2	2
สถานที่ท่องเที่ยวเพื่อการศึกษา	4	4	4	5	5	5	5	5	5
สถานที่ท่องเที่ยวทางวิถีชีวิตชุมชน	5	5	5	4	4	4	4	4	4
ค่าเฉลี่ยเวลาที่ใช้ประมวลผล (วินาที)	13	22	42	61	120	165	176	262	659

จากนั้นผู้วิจัยได้ทดสอบในผู้ใช้ 20 ตัวอย่าง ได้ค่าเฉลี่ยความถูกต้องของลำดับหมวดหมู่ความสนใจและเวลาเฉลี่ยในการประมวลผลดังรูปที่ 34 จะเห็นว่าการเลือกใช้เพื่อน 5 คนมีความเหมาะสมที่สุดเนื่องจากมีค่าความถูกต้องมากที่สุดและใช้เวลาในการประมวลผลน้อยที่สุด โดยพบว่าปัจจัยที่ส่งผลต่อความถูกต้องคือค่าปฏิสัมพันธ์ระหว่างผู้ใช้และเพื่อน เนื่องจากในจำนวนเพื่อนที่สูงขึ้น ค่าปฏิสัมพันธ์ F_i ของเพื่อนที่เพิ่มเข้ามา ทำให้การคำนวณค่าระดับความสนใจ $I(c)$ ถูกเฉลี่ยไปยังเพื่อนหลายคนตามค่าปฏิสัมพันธ์ แตกต่างจากเลือกใช้เพื่อนในจำนวนที่ต่ำ อย่างเช่นการเลือกใช้เพื่อน 1 คนค่าระดับความสนใจจะขึ้นอยู่กับเพื่อนเพียงคนเดียว ซึ่งอธิบายได้ว่าผู้ใช้และเพื่อนสนิทที่สุดที่เลือกมาเพียง 1 คนยังคงมีความแตกต่างในความสนใจ ทำให้ได้ค่าความถูกต้องที่ไม่สูงนัก อย่างไรก็ตามเมื่อเลือกใช้กลุ่มเพื่อนที่สนิทที่สุด (มากกว่า 1 คน) กลับได้ค่าความถูกต้องที่เพิ่มขึ้น ดังนั้นกลุ่มเพื่อนสนิทบนเครือข่ายสังคมจึงเป็นตัวแปรสำคัญที่มีผลต่อความสนใจของผู้ใช้มากกว่าเพื่อนสนิทเพียงคนเดียว



รูปที่ 34 ค่าเฉลี่ยความถูกต้องของลำดับหมวดหมู่ความสนใจและเวลาเฉลี่ยในการประมวลผล

4.4.3 การทดสอบประสิทธิภาพในการแนะนำสถานที่ท่องเที่ยว (Recommendation Evaluation)

การทดสอบนี้จะวัดประสิทธิภาพของระบบว่าสามารถแนะนำสถานที่ท่องเที่ยวได้ตรงกับที่ผู้ใช้สนใจหรือค้นหาอยู่ โดยใช้ข้อมูลสถานที่ท่องเที่ยวที่ระบบแนะนำไปยังผู้ใช้จำนวน 10 รายการ (Top-10 recommendation) เปรียบเทียบกับสถานที่ที่ผู้ใช้มีความสนใจจริง ซึ่งจะใช้วิธีการที่นิยมในระบบแนะนำข้อมูลอย่าง precision@10 คำนวณได้จากอัตราส่วนของสถานที่ที่ระบบแนะนำและผู้ใช้มีความสนใจต่อสถานที่ที่ระบบแนะนำทั้งหมด ดังสมการที่ 12 ระบบจะระบุสถานที่ที่ผู้ใช้มีความสนใจจากการที่ผู้ใช้คลิกเพื่อค้นหารายละเอียดเพิ่มเติมในสถานที่นั้นๆ หรือการคลิกเพื่อกำหนดเป็นสถานที่ที่สนใจ

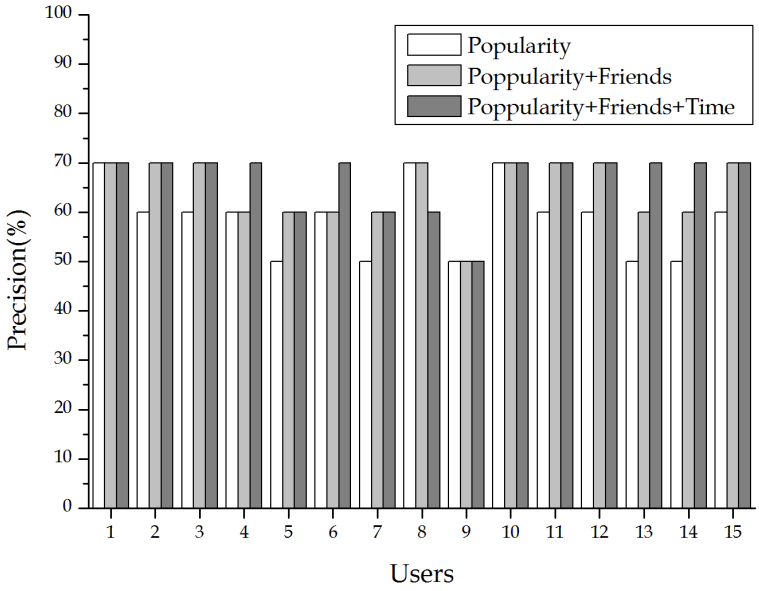
$$precision = \frac{interest \cap recommended}{recommended} \times 100 \qquad (12)$$

โดยการทดสอบของผู้ใช้ตัวอย่างจำนวน 15 ผู้ใช้เปรียบเทียบกับการใช้เกณฑ์ในการคำนวณค่าน้ำหนักของสถานที่เป็น 3 แบบ คือ

1) ใช้ค่าความนิยม (Popularity -P) ในสถานที่นั้นๆ เพียงอย่างเดียว โดยค่าความนิยมคำนวณจากจำนวนการเช็คอินในเฟสบุ๊คของสถานที่นั้น

2) ใช้ค่าความนิยม (Popularity -P) รวมกับค่าคะแนนจากเพื่อน (Friend - F) ซึ่งคำนวณจากของสถานที่ที่เพื่อนสนิทในเฟสบุ๊คของผู้ใช้เคยไป

3) ใช้ค่าความนิยม (Popularity -P) รวมกับค่าคะแนนจากเพื่อน (Friend - F) และค่าช่วงเวลาที่เหมาะสม (Time -T) ในการไปเยี่ยมชมสถานที่

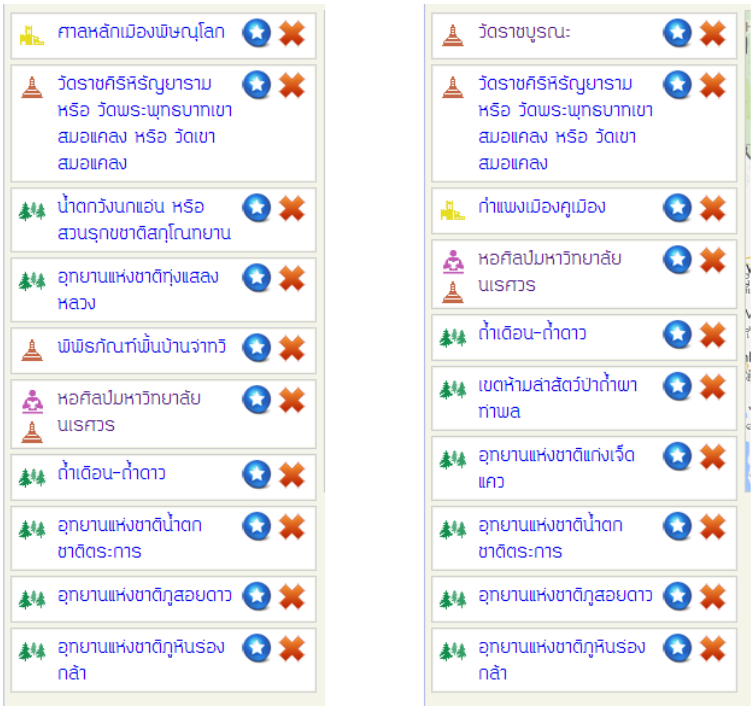


รูปที่ 35 ค่า precision ของการแนะนำเปรียบเทียบระหว่างการใช้เกณฑ์ที่ต่างกัน

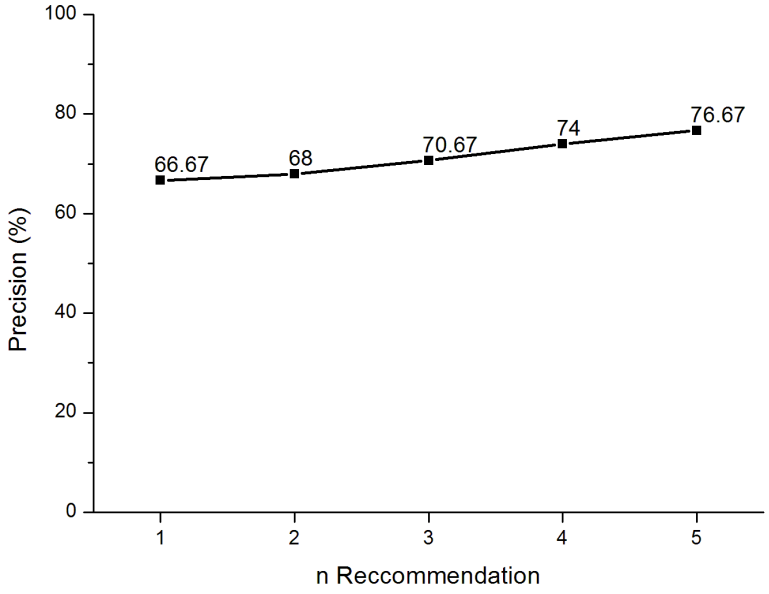
ผลการทดสอบพบว่าเมื่อใช้ค่าความนิยมเพียงอย่างเดียวจะให้ค่า precision โดยเฉลี่ยต่ำที่สุดคือ 59% แต่เมื่อเพิ่มการใช้ข้อมูลของเพื่อนและค่าของเวลาที่เหมาะสมจะได้ค่า precision โดยเฉลี่ยที่สูงขึ้นเป็น 64.66% และ 66.67% ตามลำดับ ดังนั้นสถานที่เพื่อนเคยไปและค่าของเวลาที่เหมาะสมในการไปเยี่ยมชมสถานที่ จึงมีผลต่อการเลือกสถานที่ที่ผู้ใช้สนใจหรือค้นหาอยู่ ซึ่งสามารถเพิ่มประสิทธิภาพการแนะนำสถานที่ท่องเที่ยวให้ดีขึ้นได้ อย่างไรก็ตามพบว่าในบางผู้ใช้ จากรูปที่ 35 เช่น ผู้ใช้ที่ 8 การใช้ค่าของเวลาที่เหมาะสมเพิ่มเข้ามา ทำให้ค่า precision ต่ำลง เนื่องมาจากผู้ใช้ระบุช่วงเวลาไม่ตรงกับที่ผู้ใช้สนใจจริง เช่น ระบุช่วงเวลาหน้าร้อนให้กับระบบ แต่ช่วงเวลาจริงที่จะไปเป็นหน้าหนาว ทำให้ระบบแนะนำน้ำตกแทนที่จะเป็นภูเขา นอกจากนี้ผู้ใช้ที่ 1 9 และ 10 การใช้เกณฑ์ทั้งสามแบบให้ค่าเท่ากัน เนื่องมาจากเพื่อนสนิทในเฟสบุ๊คมีข้อมูลการเยี่ยมชมสถานที่น้อย ไม่ส่งผลต่อการเปลี่ยนแปลงคำแนะนำ รวมถึงผู้ใช้มีความสนใจในสถานที่ไม่มีช่วงเวลาพิเศษ เช่น โบราณสถานหรือวัดบางแห่ง ค่าช่วงเวลาจึงไม่ส่งผลต่อการเปลี่ยนแปลงคำแนะนำเช่นกัน

4.4.4 การทดสอบการตรวจจับการเปลี่ยนแปลงความสนใจ (User Interest Change Evaluation)

ในการทดสอบนี้ เป็นการตรวจสอบว่าการนำข้อมูลตอบกลับจากผู้ใช้ สามารถทำให้ระบบทราบการเปลี่ยนแปลงความสนใจของผู้ใช้และแนะนำสถานที่ท่องเที่ยวได้ถูกต้องมากยิ่งขึ้นได้หรือไม่ โดยในงานวิจัยนี้ใช้ข้อมูลตอบกลับของผู้ใช้จำนวน 15 ผู้ใช้ในการปรับปรุงคำแนะนำ ซึ่งจะคำนวณค่าพรีซิชั่นในคำแนะนำใหม่เปรียบเทียบกับคำแนะนำเดิม โดยใช้สมการที่ 6



รูปที่ 36 รายการแนะนำที่เปลี่ยนไปเมื่อนำพฤติกรรมการตอบกลับมาคำนวณคำแนะนำใหม่



รูปที่ 37 ค่าพรีซิชั่นของการแนะนำเปรียบเทียบระหว่างการปรับปรุงคำแนะนำใหม่

รูปที่ 36 แสดงให้เห็นว่าระบบมีการเปลี่ยนรายการแนะนำสถานที่ท่องเที่ยวของผู้ใช้คนเดิมตามความสนใจของผู้ใช้ที่เปลี่ยนไป ผลการทดสอบพบว่าเมื่อปรับปรุงคำแนะนำใหม่ค่าพรีซิชั่นมีแนวโน้มที่สูงขึ้น ดังรูปที่ 37 จะเห็นว่าการปรับปรุงคำแนะนำครั้งที่ 1 2 3 4 และ 5 ให้ค่าพรีซีชั่นเท่ากับ 66.67% 68% 70.67% 74% และ 76.67% ตามลำกับ ดังนั้นการปรับปรุงคำแนะนำจากการตอบกลับของผู้ใช้สามารถแนะนำสถานที่ท่องเที่ยวได้ตรงกับความสนใจของผู้ใช้มากขึ้น อย่างไรก็ตามพบว่าในบางผู้ใช้ การปรับปรุงคำแนะนำกลับให้ค่า precisionเท่าเดิม เนื่องมาจากพฤติกรรมการตอบกลับบางรูปแบบอาจจะทำให้ค่าน้ำหนักของสถานที่เปลี่ยนแปลง แต่ไม่ได้ทำให้การลำดับเปลี่ยนแปลงไป

4.4.5 ความพึงพอใจในการใช้งานระบบจากผู้ใช้

การประเมินความพึงพอใจในการใช้งานระบบ ได้ใช้แบบสอบถามความคิดเห็นการใช้งานเว็บแอปพลิเคชั่น PTIS แบบออนไลน์ โดยให้ผู้ใช้งานจริงทำการประเมินความพึงพอใจหลังจากใช้งานระบบจำนวน 57 คน จากนั้นนำข้อมูลที่ได้มาหาค่าเฉลี่ยและเพื่อประเมินความพึงพอใจของผู้ใช้ที่มีต่อการใช้งานระบบ โดยจากการสอบถามจะแบ่งการประเมินออกเป็น 3 ส่วน คือ 1) ข้อมูลทั่วไปของผู้ตอบแบบสอบถาม 2) ความพึงพอใจด้านต่าง ๆ ของการใช้งานเว็บแอปพลิเคชั่น และ 3) ความคิดเห็นเกี่ยวกับแอปพลิเคชั่น ดังนี้

1. ข้อมูลทั่วไปของผู้ตอบแบบสอบถาม

จากผู้ตอบแบบสอบถามความคิดเห็นทั้งหมด พบว่าส่วนมากเป็นเพศหญิง โดยมีจำนวน 37 คน คิดเป็นร้อยละ 68.52 เป็นเพศชาย จำนวน 17 คน คิดเป็นร้อยละ 31.48 ดังแสดงในตารางที่ 13

ตารางที่ 13 ลักษณะของผู้ตอบแบบสอบถามความคิดเห็น

เพศ	จำนวน	ร้อยละ	แปลผล
ชาย	17	31.48	น้อย
หญิง	37	68.52	มาก
รวม	**54**	**100.00**	**หญิงมากกว่าชาย**

จากผู้ตอบแบบสอบถามความคิดเห็นทั้งหมด พบว่าส่วนมากมีสถานะภาพโสด โดยมีจำนวน 48 คน คิดเป็นร้อยละ 88.89 มีสถานะภาพสมรส จำนวน 6 คน คิดเป็นร้อยละ 11.11 ดังแสดงในตารางที่ 14

ตารางที่ 14 ข้อมูลสถานะภาพของผู้ตอบแบบสอบถามความคิดเห็น

สถานะภาพ	จำนวน	ร้อยละ	แปลผล
โสด	48	88.89	มาก
สมรส	6	11.11	น้อย
รวม	**54**	**100.00**	**โสดมากกว่าสมรส**

จากผู้ตอบแบบสอบถามความคิดเห็นทั้งหมด พบว่าช่วงอายุการศึกษาของผู้ตอบแบบสอบถามส่วนมาก คือ มีช่วงอายุ 18 - 30 ปี จำนวน 43 คน คิดเป็นร้อยละ 79.63 รองลงมาช่วงอายุ 31 - 45 ปี จำนวน 11 คน คิดเป็นร้อยละ 20.37 ดังแสดงในตารางที่ 15

ตารางที่ 15 ข้อมูลช่วงอายุของผู้ตอบแบบสอบถามความคิดเห็น

ช่วงอายุ	จำนวน	ร้อยละ	แปลผล
น้อยกว่า 18 ปี	0	0.00	ไม่มี
18 - 30 ปี	43	79.63	มาก
31 - 45 ปี	11	20.37	น้อย
46 - 60 ปี	0	0.00	ไม่มี
มากกว่า 60 ปี	0	0.00	ไม่มี
รวม	**54**	**100.00**	**มีช่วงอายุ 18 - 30 ปี เป็นส่วนมาก**

จากผู้ตอบแบบสอบถามความคิดเห็นทั้งหมด พบว่าระดับการศึกษาของผู้ตอบแบบสอบถามส่วนมาก คือปริญญาตรีและปริญญาโท โดยมีระดับการศึกษาปริญญาตรี จำนวน 30 คน คิดเป็นร้อยละ 55.56 มีระดับการศึกษาปริญญาโท จำนวน 22 คน คิดเป็นร้อยละ 40.74 ดังแสดงในตารางที่ 16

ตารางที่ 16 ข้อมูลระดับการศึกษาของผู้ตอบแบบสอบถามความคิดเห็น

ระดับการศึกษา	จำนวน	ร้อยละ	แปลผล
มัธยมศึกษา / อาชีวศึกษา	0	0.00	ไม่มี
ประกาศนียบัตร / อนุปริญญา / ปวส.	0	0.00	ไม่มี
ปริญญาตรี	30	55.56	มาก
ปริญญาโท	22	40.74	มาก
ปริญญาเอก	2	3.70	น้อย
รวม	**54**	**100.00**	**ระดับปริญญาตรีและโทเป็นส่วนมาก**

2. ความพึงพอใจด้านต่าง ๆ ของการใช้งานเว็บแอปพลิเคชั่น

การประเมินความพึงพอใจของระบบแบ่งออกเป็น 3 ด้าน คือ 1) ด้านส่วนติดต่อผู้ใช้งาน (User Interface) 2) ด้านการแนะนำสถานที่ท่องเที่ยว 3) ด้านการแนะนำแผนการเดินทาง โดยในแบบประเมินความพึงพอใจ แบ่งความพึงพอใจออกเป็น 5 ระดับ ซึ่งมีความหมายดังนี้

มากที่สุด	หมายถึงผู้ตอบแบบสอบถามมีความพึงพอใจมากที่สุด
มาก	หมายถึงผู้ตอบแบบสอบถามมีความพึงพอใจมาก
ปานกลาง	หมายถึงผู้ตอบแบบสอบถามมีความพึงพอใจปานกลาง
น้อย	หมายถึงผู้ตอบแบบสอบถามมีความพึงพอใจน้อย
น้อยที่สุด	หมายถึงผู้ตอบแบบสอบถามมีความพึงพอใจน้อยที่สุด

การให้ค่าคะแนนเป็นไปตามเกณฑ์ดังนี้

มากที่สุด	มีค่าเท่ากับ	5
มาก	มีค่าเท่ากับ	4
ปานกลาง	มีค่าเท่ากับ	3
น้อย	มีค่าเท่ากับ	2
น้อยที่สุด	มีค่าเท่ากับ	1

และเกณฑ์ในการประเมินความพึงพอใจดังนี้

ค่าเฉลี่ยระดับความพึงพอใจ	การประเมินผล
4.21 - 5.00	ระบบมีประสิทธิภาพดีมาก
3.41 - 4.20	ระบบมีประสิทธิภาพดี
2.61 - 3.40	ระบบมีประสิทธิภาพปานกลาง
1.81 - 2.60	ระบบมีประสิทธิภาพพอใช้
1.00 - 1.80	ระบบมีประสิทธิภาพควรปรับปรุง

การประเมินประสิทธิภาพจากความพึงพอใจของผู้ใช้ที่มีต่อระบบด้านส่วนติดต่อผู้ใช้งาน (User Interface) ในเรื่องความชัดเจนและเข้าใจง่ายในขั้นตอนของการใช้งานระบบพบว่า ในส่วนติดต่อผู้ใช้งานในการค้นหาสถานที่ท่องเที่ยว โดยผู้ใช้มีระดับความพึงพอใจต่อการใช้งานในส่วนนี้ค่าเฉลี่ยอยู่ที่ 4.22 ส่วนการวางแผนการท่องเที่ยวและนำเสนอแผนการเดินทาง ผู้ใช้มีระดับความพึงพอใจต่อการใช้งานในส่วนนี้ที่ค่าเฉลี่ย 4.13 และ 4.16 ตามลำดับ เมื่อคิดค่าเฉลี่ยระดับความพึงพอใจของผู้ใช้ที่มีต่อระบบด้านส่วนติดต่อผู้ใช้งาน (User Interface) จะได้ค่าเฉลี่ยระดับความพึงพอใจที่ 4.17 จึงสรุปได้ว่าผู้ใช้มีความพึงพอใจในส่วนติดต่อผู้ใช้งาน (User Interface)ในระดับดี ดังแสดงในตารางที่ 17

ตารางที่ 17 ผลการประเมินความพึงพอใจของผู้ใช้ด้านส่วนติดต่อผู้ใช้งาน (User Interface)

	ระดับความพึงพอใจ						
ด้านที่ประเมิน	**5**	**4**	**3**	**2**	**1**	**ค่าเฉลี่ย**	**แปลผล**
ด้านส่วนติดต่อผู้ใช้งาน (User Interface)							
1. การค้นหาสถานที่ท่องเที่ยวมีขั้นตอนชัดเจนและเข้าใจง่าย	35.18	55.56	5.56	3.70	0.00	4.22	ดีมาก
2. การวางแผนการท่องเที่ยวมีขั้นตอนชัดเจนและเข้าใจง่าย	31.48	51.85	14.82	1.85	0.00	4.13	ดี
3. การนำเสนอแผนการเดินทางและตารางเวลาการเดินทางมีความชัดเจนและเข้าใจง่าย	57.04	50.00	5.56	7.40	0.00	**4.16**	ดี
ค่าเฉลี่ยด้านส่วนติดต่อผู้ใช้งาน						**4.17**	ดี

การประเมินประสิทธิภาพจากความพึงพอใจของผู้ใช้ที่มีต่อระบบด้านการแนะนำสถานที่ท่องเที่ยว ในเรื่องความถูกต้องในการวิเคราะห์ระดับความสนใจ ความแม่นยำในการแนะนำและประโยชน์ของสถานที่แนะนำต่อการวางแผนการท่องเที่ยวของระบบพบว่า ระบบมีประสิทธิภาพดีในความถูกต้องของการวิเคราะห์ค่าระดับความสนใจในประเภทสถานที่ท่องเที่ยวจากเครือข่ายสังคมเฟสบุ๊คและความแม่นยำในการแนะนำสถานที่ท่องเที่ยวให้ตรงกับความสนใจ โดยมีค่าเฉลี่ยระดับความพึงพอใจที่ 4.20 และ 4.16 ตามลำดับ ส่วนความเป็นประโยชน์ของสถานที่แนะนำที่มีต่อการนำไปวางแผนการท่องเที่ยว พบว่า สถานที่ท่องเที่ยวแนะนำมีประโยชน์ต่อการนำไปวางแผนการท่องเที่ยวดีมาก โดยมีระดับความพึงพอใจที่ 4.32 ที่พักและร้านอาหารมีประโยชน์ในระดับดี โดยมีระดับความพึงพอใจที่ 4.09 และ 4.11 ตามลำดับ เมื่อคิดค่าเฉลี่ยระดับความพึงพอใจของผู้ใช้ที่มีต่อระบบด้านการแนะนำสถานที่ท่องเที่ยว จะได้ค่าเฉลี่ยระดับความพึงพอใจที่ 4.18 จึงสรุปได้ว่าผู้ใช้มีความพึงพอใจระบบในด้านการแนะนำสถานที่ท่องเที่ยวในระดับดี ดังแสดงในตารางที่ 18

ตารางที่ 18 ผลการประเมินความพึงพอใจของผู้ใช้ด้านการแนะนำสถานที่ท่องเที่ยว

ด้านที่ประเมิน	ระดับความพึงพอใจ 5	4	3	2	1	ค่าเฉลี่ย	แปลผล
ด้านการแนะนำสถานที่ท่องเที่ยว							
1. ความถูกต้องของค่าระดับความสนใจในประเภทสถานที่ท่องเที่ยวที่วิเคราะห์จากข้อมูล Facebook	31.48	57.41	11.11	0	0	4.20	ดี
2. ความแม่นยำของสถานที่ท่องเที่ยวแนะนำที่ตรงกับความสนใจของผู้ใช้	31.48	57.41	9.26	0	1.85	**4.16**	ดี
3. สถานที่ท่องเที่ยวแนะนำมีประโยชน์ต่อการวางแผนการท่องเที่ยวมากน้อยเพียงใด	46.30	40.74	11.11	1.85	0	**4.32**	ดีมาก
4. ที่พักแนะนำมีประโยชน์ต่อการวางแผนการท่องเที่ยวมากน้อยเพียงใด	31.48	53.71	11.11	3.70	0	**4.09**	ดี
5. ร้านอาหารแนะนำมีประโยชน์ต่อการวางแผนการท่องเที่ยวมากน้อยเพียงใด	33.34	50.00	12.96	1.85	1.85	4.11	ดี
ค่าเฉลี่ยด้านการแนะนำสถานที่ท่องเที่ยว						**4.18**	ดี

การประเมินประสิทธิภาพจากความพึงพอใจของผู้ใช้ที่มีต่อระบบด้านการแนะนำแผนการเดินทาง พบว่า ผู้ใช้มีความพึงพอใจในการแสดงสถานที่แนะนำได้ตรงกับความสนใจของผู้ใช้ โดยมีค่าเฉลี่ยระดับความพึงพอใจที่ 4.11 และมีความพึงพอใจในการลำดับสถานที่ท่องเที่ยวได้อย่างเหมาะสม โดยมีค่าเฉลี่ยอยู่ที่ 3.98 มีนอกจากนี้ผู้ใช้มีความพึงพอใจในการแสดงตารางเวลาที่ใช้เวลาใกล้เคียงกับประสบการณ์ท่องเที่ยวจริงเหมาะสม และความรวดเร็วในการสร้างแผนการเดินทาง โดยมีค่าเฉลี่ยระดับความพึงพอใจที่ 4.07 และมีความพึงพอใจในประโยชน์ของรายการแผนการเดินทางแนะนำในการวางแผนการท่องเที่ยวของผู้ใช้ โดยมีค่าเฉลี่ยระดับความพึงพอใจ 4.16 เมื่อคิดค่าเฉลี่ยระดับความพึงพอใจของผู้ใช้ที่มีต่อระบบด้านการแนะนำแผนการเดินทาง จะได้ค่าเฉลี่ยระดับความพึงพอใจ 4.08 ซึ่งอยู่ในระดับดี ดังแสดงในตารางที่ 19

ตารางที่ 19 ผลการประเมินความพึงพอใจของผู้ใช้ด้านการแนะนำแผนการเดินทาง

ด้านที่ประเมิน	ระดับความพึงพอใจ 5	4	3	2	1	ค่าเฉลี่ย	แปลผล
ด้านการแนะนำแผนการเดินทาง							
1. สถานที่แนะนำที่ปรากฏในแผนการเดินทางตรงกับความสนใจของผู้ใช้	25.93	59.26	14.81	0	0	4.11	ดี

2. แผนการเดินทางแนะนำมีลำดับสถานที่ท่องเที่ยวที่เหมาะสม	24.07	55.56	14.81	5.56	0	3.98	ดี
3. ตารางเวลาที่แนะนำเหมาะสม (ใช้เวลาใกล้เคียงกับประสบการณ์ท่องเที่ยวจริง)	31.48	48.15	16.67	3.70	0	4.07	ดี
4. ความสะดวก รวดเร็วในการวางแผนการเดินทางแต่ละครั้ง	24.07	62.97	9.26	3.70	0	4.07	ดี
5. รายการแผนการเดินทางแนะนำมีประโยชน์ต่อการวางแผนการท่องเที่ยวของผู้ใช้มากน้อยเพียงใด	35.19	48.15	14.82	1.85	0	4.16	ดี
ค่าเฉลี่ยด้านการแนะนำแผนการเดินทาง						**4.08**	ดี

3. ความคิดเห็นเกี่ยวกับแอปพลิเคชั่น

จากผู้ตอบแบบสอบถามความคิดเห็นทั้งหมด พบว่าผู้ใช้จะใช้ระบบเพื่อช่วยหาข้อมูลการท่องเที่ยวในครั้งต่อไป จำนวน 48 คน คิดเป็นร้อยละ 88.89 ไม่ใช้ระบบในการช่วยหาข้อมูลการท่องเที่ยวในครั้งต่อไป จำนวน 6 คน คิดเป็นร้อยละ 11.11 ดังนั้นผู้ใช้ส่วนมากยังคงใช้ระบบในการช่วยหาข้อมูลการท่องเที่ยวอีกในครั้งต่อไป ดังแสดงในตารางที่ 20

ตารางที่ 20 ข้อมูลความคิดเห็นเกี่ยวกับการใช้งานแอปพลิเคชั่นในการช่วยหาข้อมูลท่องเที่ยวครั้งต่อไป

ใช้ช่วยหาข้อมูลการท่องเที่ยวในครั้งต่อไป	จำนวน	ร้อยละ	แปลผล
ใช่	48	88.89	มาก
ไม่ใช่	6	11.11	น้อย
รวม	**54**	**100.00**	**ส่วนมากใช้ช่วยหาข้อมูลการท่องเที่ยวในครั้งต่อไป**

จากผู้ตอบแบบสอบถามความคิดเห็นทั้งหมด พบว่าผู้ใช้จะใช้ระบบเพื่อช่วยวางแผนการท่องเที่ยวในครั้งต่อไป จำนวน 47 คน คิดเป็นร้อยละ 87.04 ไม่ใช้ระบบในการช่วยวางแผนการท่องเที่ยวในครั้งต่อไป จำนวน 7 คน คิดเป็นร้อยละ 12.96 ดังนั้นผู้ใช้ส่วนมากยังคงใช้ระบบในการช่วยวางแผนการท่องเที่ยวอีกในครั้งต่อไป ดังแสดงในตารางที่ 21

ตารางที่ 21 ข้อมูลความคิดเห็นเกี่ยวกับการใช้งานแอปพลิเคชั่นในการวางแผนการท่องเที่ยวครั้งต่อไป

ใช้วางแผนการท่องเที่ยวในครั้งต่อไป	จำนวน	ร้อยละ	แปลผล
ใช่	47	87.04	มาก
ไม่ใช่	7	12.96	น้อย
รวม	**54**	**100.00**	**ส่วนมากใช้วางแผนการท่องเที่ยวในครั้งต่อไป**

จากผู้ตอบแบบสอบถามความคิดเห็นทั้งหมด พบว่าผู้ใช้จะแนะนำแอปพลิเคชั่นนี้ให้กับบุคคลอื่น จำนวน 50 คน คิดเป็นร้อยละ 92.59 ไม่แนะนำแอปพลิเคชั่นนี้ให้กับบุคคลอื่น จำนวน 4 คน คิดเป็นร้อยละ 7.41 ดังนั้นผู้ใช้ส่วนมากจะแนะนำแอปพลิเคชั่นนี้ให้กับบุคคลอื่น ดังแสดงในตารางที่ 22

ตารางที่ 22 ข้อมูลความคิดเห็นเกี่ยวกับการแนะนำแอปพลิเคชั่นนี้กับบุคคลอื่น

แนะนำแอปพลิเคชั่นนี้กับบุคคลอื่น	จำนวน	ร้อยละ	แปลผล
ใช่	50	92.59	มาก
ไม่ใช่	4	7.41	น้อย
รวม	**54**	**100.00**	**ส่วนมากแนะนำแอปพลิเคชั่นนี้กับบุคคลอื่น**

บทที่ 5

สรุปผลการวิจัยและข้อเสนอแนะ

5.1 สรุปผลการวิจัย

ระบบแนะนำข้อมูลการท่องเที่ยวเฉพาะบุคคลที่ศึกษาวิจัยนี้ พยายามที่จะวิเคราะห์ความสนใจของผู้ใช้เพื่อแนะนำข้อมูลการท่องเที่ยวให้ต้องกับความต้องการของแต่ละบุคคล โดยแบ่งการแนะนำออกเป็น 3 ประเภท คือ 1. ข้อมูลสถานที่ท่องเที่ยวและงานเทศกาล 2. ข้อมูลสถานที่ร้านอาหาร และ 3. ข้อมูลสถานที่ที่พัก ระบบถูกออกแบบโดยพยายามแก้ปัญหาที่เกิดขึ้นในระบบแนะนำข้อมูลคือ ปัญหาโคลสตาร์ท เวลาที่ผู้ใช้ต้องสูญเสียในการกรอกข้อมูลให้กับระบบและความยุ่งยากในการลงทะเบียนใช้งาน ดังนั้นระบบที่นำเสนอจึงใช้เทคนิคการผสานข้อมูลที่มาจากเครือข่ายสังคมออนไลน์ (Online Social network) และข้อมูลที่ได้จากผู้ใช้โดยตรง ข้อดีของวิธีดังกล่าวคือระบบทราบข้อมูลเบื้องต้นของผู้ใช้และยังสามารถนำข้อมูลดังกล่าวไปคำนวณเพื่อหาความสนใจของผู้ใช้ได้ ซึ่งเป็นแนวคิดใหม่ของงานวิจัยนี้ นอกจากนี้ระบบมีฟังก์ชันการทำงานสำหรับผู้ที่ไม่มีข้อมูลในสังคมออนไลน์ นั่นคือผู้ใช้ไม่จำเป็นต้องล๊อกอินเข้าใช้ระบบก็สามารถใช้งานได้เช่นกัน แต่ความถูกต้องของข้อมูลที่นำเสนออาจจะไม่แม่นยำมากนัก เนื่องจากระบบไม่มีข้อมูลมากพอที่จะแนะนำข้อมูลท่องเที่ยวให้กับนักท่องเที่ยว ระบบจะใช้ข้อมูลเบื้องต้นที่ผู้ใช้กรอกให้กับระบบเท่านั้น เพื่อคำนวณหาความสนใจของผู้ใช้ อย่างไรก็ตามระบบก็ยังสามารถให้ข้อมูลได้อย่างมีประสิทธิภาพเมื่อเทียบกับระบบค้นหาข้อมูล (search engine) ทั่วๆ ไป

โดยผลการทดลองของระบบ มีดังนี้

1. จากการวิเคราะห์ความต้องการของผู้ใช้ พบว่าแหล่งข้อมูลบนอินเทอร์เน็ตอย่างเว็บไซต์ บล็อก และเครือข่ายสังคมออนไลน์ ถูกใช้ในการค้นหาข้อมูลการท่องเที่ยวในสัดส่วนที่สูง โดยเครือข่ายสังคมออนไลน์ที่ผู้ใช้นิยมใช้หาข้อมูลการท่องเที่ยวมากที่สุดคือ เฟสบุ๊ค ซึ่งเป็นเครือข่ายสังคมออนไลน์ที่มีจำนวนผู้ใช้ในประเทศไทยมากเช่นกัน นอกจากนี้ผลของการวิเคราะห์ความสนใจเบื้องต้นโดยใช้ข้อมูลเช็คอิน จากเฟสบุ๊ค สามารถแสดงถึงความสนใจที่แตกต่างกันระหว่างบุคคลได้ ดังนั้นจึงเป็นไปตามสมมติฐานที่ 1 : “ข้อมูลจากเครือข่ายสังคมออนไลน์สามารถนำมาวิเคราะห์เพื่อหาความสนใจเฉพาะบุคคลของผู้ใช้ได้”

2. จากการวิเคราะห์ความต้องการของผู้ใช้อีกประเด็นหนึ่งพบว่านักท่องเที่ยวส่วนมากมักสอบถามข้อมูลการท่องเที่ยวจากเพื่อนๆ ที่เคยไปสถานที่ท่องเที่ยวนั้นมาก่อน และยังมีความสนใจสถานที่ท่องเที่ยวต่างๆ ที่เพื่อนๆ เคยไปมา ข้อมูลจากเพื่อนๆ จึงเป็นปัจจัยหนึ่งที่สำคัญในวิเคราะห์เพื่อแนะนำข้อมูลให้กับนักท่องเที่ยวได้ดียิ่งขึ้น ดังนั้นจึงเป็นไปตามสมมติฐานที่ 2: “ข้อมูลจากเพื่อนของผู้ใช้สามารถนำมาใช้ประโยชน์ในการแนะนำสถานที่ท่องเที่ยวให้กับผู้ใช้และเรียงลำดับสถานที่ท่องเที่ยวที่น่าสนใจได้”

3. ระบบสามารถใช้ข้อมูลการตอบกลับจากผู้ใช้มาวิเคราะห์ถึงการเปลี่ยนแปลงความสนใจของผู้ใช้ได้ และสามารถแนะนำสถานที่ท่องเที่ยวที่ตรงกับการเปลี่ยนแปลงความสนใจของผู้ใช้ได้อย่างมีประสิทธิภาพ

อย่างไรก็ตามระบบที่พัฒนาขึ้นยังมีข้อจำกัดอยู่หลายประการ และจากการสอบถามผู้ใช้สามารถสรุปข้อเสนอแนะได้ดังต่อไปนี้

1) ยังไม่มีเวอร์ชั่นภาษาอังกฤษ ในอนาคตควรมีเวอร์ชั่นภาษาอังกฤษด้วย
2) การดึงข้อมูลจากเฟสบุ๊คยังทำงานได้ช้า โดยแนวทางแก้ไขคือการจำกัดระยะเวลาของข้อมูลการเช็คอินของผู้ใช้ เช่น ดึงข้อมูลเฉพาะช่วงเวลา 1 ปีที่ผ่านมาเท่านั้น ข้อมูลที่เก่ากว่านั้นจะไม่ถูกนำมาพิจารณาความสนใจ เป็นต้น
3) เว็บไซต์ยังขาดความสวยงามและความน่าสนใจ ซึ่งสามารถแก้ไขได้โดยการเพิ่มอนิเมชั่นต่างๆ ทำให้หน้าสนใจมากยิ่งขึ้น

5.2 งานที่จะพัฒนาต่อในอนาคต

แนวทางการพัฒนาระบบต่อไปในอนาคตคือการเพิ่มฟังชั่นต่างๆ ในการค้นหาข้อมูลสถานที่ท่องเที่ยวต่างๆ แต่ระบบการค้นหาข้อมูลดังกล่าวนี้มีลักษณะพิเศษคือ สามารถรองรับคำถามที่เป็นภาษาธรรมชาติได้ เช่น “พิษณุโลกมีสถานที่อะไรน่าสนใจบ้าง” ซึ่งข้อคำถามดังกล่าวระบบทั่วไปยังไม่สามารถเข้าใจความหมายของข้อคำถามได้ ดังนั้นเทคนิคหนึ่งที่จะนำมาใช้เพื่ออำนวยความสะดวกในการวิเคราะห์คำถามคือ ออนโทโลยี คุณสมบัติเด่นของออนโทโลยีคือการค้นหาข้อมูลเชิงความหมายได้ และสามารถที่จะลดความกำกวมของคำถามได้อย่างมีประสิทธิภาพ จึงเป็นที่นิยมในหมู่นักวิจัยทางด้านวิทยาการคอมพิวเตอร์ และนำมาประยุกต์ใช้ในการค้นหาข้อมูลเป็นสำคัญ แต่งานวิจัยทางด้านการท่องเที่ยวนั้น ยังไม่ปรากฏการนำเทคโนโลยีนี้ไปใช้มากนัก ดังนั้นจึงเป็นแนวทางหนึ่งในการพัฒนาต่อยอดงานวิจัยนี้ นอกจากนี้ยังควรพัฒนาระบบสอบถามนี้ให้อยู่ในรูปแบบของแอปพลิเคชั่นบนมือถือซึ่งจะมีความสะดวกในการใช้งานมากกว่าเวอร์ชั่นที่ทำงานบนเครื่องคอมพิวเตอร์ส่วนบุคคลทั่วไป ซึ่งแอปพลิเคชั่นนั้นควรมีทั้งเวอร์ชั่นบนระบบปฏิบัติการ Android และ iOS

ควรมีการสนับสนุนให้นักวิจัยทางด้านวิทยาการคอมพิวเตอร์พัฒนาแอปพลิเคชั่นบนมือถือต่างๆ และสามารถเรียกใช้บริการข้อมูลจากโครงการวิจัยที่นำเสนอนี้ เพื่อให้เกิดความหลากหลายและเป็นประโยชน์ต่อนักท่องเที่ยวในอนาคตต่อไป ทั้งนี้แอปพลิเคชั่นต่างๆ ควรรองรับความต้องการของนักท่องเที่ยวชาวต่างๆ ชาติด้วย ซึ่งควรจะรองรับการแสดงผลในรูปแบบของภาษาต่างๆ เช่น จีน ญี่ปุ่น เกาหลี อังกฤษ ฝรั่งเศส เยอรมัน หรือภาษาอื่นๆ ที่เป็นกลุ่มนักท่องเที่ยวหลักในเมืองไทย เป็นต้น

บรรณานุกรม

สำนักงานคณะกรรมการวิจัยแห่งชาติ. (2553). *คู่มือการประเมินผลข้อเสนอการวิจัยของหน่วยงานภาครัฐที่เสนอของบประมาณประจำปีงบประมาณ 2555 ตามมติคณะรัฐมนตรี*.

Adler, J. L., and Blue, and V. J. (1998). Toward the design of intelligent traveler information systems. *Transportation Research Part C: Emerging Technologies*, 6(3), pp. 157-172.

Ambite, J. L., Barish, G., Knoblock, C. A., Muslea, M., Oh, J., & Minton, S. (2002). Getting from here to there: interactive planning and agent execution for optimizing travel. In *Proceedings of the 14th Conference on Innovative Applications of Artificial Intelligence* (Vol. 1, pp. 862–869). Edmonton, Alberta, Canada.

Berka, T., & Plnig, M. (2004). Designing recommender systems for tourism. In *The 11th International Conference on Information Technology in Travel and Tourism, ENTER*. Cairo, Egypt.

Castells, P., Fernandez, M., & Vallet, D. (2007). An Adaptation of the Vector-Space Model for Ontology-Based Information Retrieval. *IEEE Transactions on Knowledge and Data Engineering*, *19*(2), 261–272.

Castells, P., Fernández, M., Vallet, D., Mylonas, P., & Avrithis, Y. (2005). Self-tuning Personalized Information Retrieval in an Ontology-Based Framework. In *OTM Workshops on the Move to Meaningful Internet Systems* (pp. 977–986).

gajus. (2012, November 16). *ay-fb-friend-rank*. *GitHub.com*. Retrieved January 12, 2014, from https://github.com/gajus/facebook-friend-rank

Gauch, S., Chaffee, J., & Pretschner, A. (2003). Ontology-based personalized search and browsing. *Web Intelligence and Agent Systems*, *1*, 3–4.

Gnter, G. E., Erbach, G., Neumann, G., & Uszkoreit, H. (1997). Multilingual Indexing, Navigation and Editing Extensions for the World-Wide Web. In *Proceedings of the 3rd DELOS workshop - Cross-Language Information Retrieval* (pp. 22–28).

Hofmann, T. (1999). Probabilistic Latent Semantic Indexing. In *Proceedings of the 22nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 50–57).

Kesorn, K., Liang, Z., & Poslad, S. (2009). Use of Granularity and Coverage in a User Profile Model to Personalise Visual Content Retrieval. In *Proceedings of the 2nd International*

Conference on Advances in Human-Oriented and Personalized Mechanisms, Technologies, and Services (pp. 79–84).

Li, Y., & Zhong, N. (2006). Mining Ontology for Automatically Acquiring Web User Information Needs. *IEEE Transactions on Knowledge and Data Engineering*, *18*(4), 554–568.

Liu, F., Yu, C., & Meng, W. (2004). Personalized Web Search For Improving Retrieval Effectiveness. *IEEE Transaction on Knowledge and Data Engineering*, *16*(1), 28–40.

Moreno, A., Valls, A., Isern, D., Marin, L., & Borràs, J. (2013). SigTur/E-Destination: Ontology-based personalized recommendation of Tourism and Leisure Activities. *Engineering Applications of Artificial Intelligence*, *26*(1), 633–651.

Mylonas, P., Vallet, D., Castells, P., Fernandez, M., & Avrithis, Y. (2008). Personalized Information Retrieval Based on Context and Ontological Knowledge. *The Knowledge Engineering Review*, *23*(Special Issue), 73–100.

Ricci, F. (2002). Travel Recommender Systems. *IEEE Intelligent Systems*, *17*(6), 55–57.

Salguero, A., Araque, F., & Delgado, C. (2008). Application of business intelligence methods for personalizing tourist services. *WSEAS TRANSACTIONS on SYSTEMS*, *7*(10), 1176–1185.

Salton, G. (1971). *The SMART Retrieval System—Experiments in Automatic Document Processing*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc.

Stabb, S., Werther, H., Ricci, F., Zipf, A., Gretzel, U., Fesenmaier, D. R., … Knoblock, C. (2002). Intelligent Systems for Tourism. *IEEE Intelligent Systems*, *17*(6), 53– 66.

Tao, X., & Li, Y. (2009). Concept-Based, Personalized Web Information Gathering: A Survey. In *Proceedings of the 3rd International Conference on Knowledge Science, Engineering and Management* (pp. 215–228).

Tao, X., Li, Y., Zhong, N., & Nayak, R. (2007). Ontology Mining for PersonalizedWeb Information Gathering. In *Proceedings of the IEEE/WIC/ACM International Conference on Web Intelligence* (pp. 351–358).

Vallet, D., Castells, P., Fernandez, M., Mylonas, P., & Avrithis, Y. (2007). Personalized Content Retrieval in Context Using Ontological Knowledge. *IEEE Transactions on Circuits and Systems for Video Technology*, *17*(3), 336–346.

Widman, J. (n.d.). *EdgeRank*. *edgerank.net*. Retrieved July 2, 2014, from http://edgerank.net/

ภาคผนวก

ภาคผนวก ก : บทความตีพิมพ์ผลงาน

ในโครงการวิจัยที่ 1 นี้ได้ทำการเผยแพร่แนวคิดเกี่ยวกับการใช้ประโยชน์จากเครือข่ายสังคมออนไลน์สำหรับแนะนำข้อมูลการท่องเที่ยวในบทความวิชาการ เรื่อง *"ทบทวนวิธีการสร้างแบบจำลองผู้ใช้และการใช้ประโยชน์จากเครือข่ายสังคมสำหรับระบบแนะนำข้อมูลการท่องเที่ยว"* ใน วารสารวิทยาศาสตร์ มศว. ดังรายละเอียดต่อไปนี้

วนารัตน์ จุฬาพันธ์ทอง และไกรศักดิ์ เกษร, *"ทบทวนวิธีการสร้างแบบจำลองผู้ใช้และการใช้ประโยชน์จากเครือข่ายสังคมสำหรับระบบแนะนำข้อมูลการท่องเที่ยว"*, วารสารวิทยาศาสตร์ มศว., ปีที่ 30, ฉบับที่ 1, 2557 หน้า 209-228.

บทความวิชาการ

ทบทวนวิธีการสร้างแบบจำลองผู้ใช้และการใช้ประโยชน์จากเครือข่ายสังคมสำหรับระบบแนะนำข้อมูลการท่องเที่ยว

วนารัตน์ จุฬพันธ์ทอง และ ไกรศักดิ์ เกษร*

บทคัดย่อ

ระบบแนะนำข้อมูลการท่องเที่ยวเป็นประโยชน์ต่อทั้งนักท่องเที่ยวและผู้ประกอบการ ซึ่งจะช่วยทำให้นักท่องเที่ยวเสียเวลาน้อยลงในการค้นหาข้อมูลการท่องเที่ยวที่ต้องการ สำหรับผู้ประกอบการ ระบบจะช่วยสร้างแรงจูงใจให้นักท่องเที่ยวเข้ามาใช้บริการของตนเองมากขึ้นหากมีเครื่องมือเพื่อแนะนำข้อมูลที่เป็นประโยชน์ การแนะนำข้อมูลให้นักท่องเที่ยวได้อย่างมีประสิทธิภาพนั้น แบบจำลองผู้ใช้ (User Model) ถือเป็นส่วนสำคัญของระบบแนะนำข้อมูลที่มีหน้าที่ในการหาข้อสรุปความสนใจของผู้ใช้ อย่างไรก็ตามการสร้างแบบจำลองผู้ใช้มีความท้าทายอย่างยิ่งสำหรับระบบแนะนำข้อมูลการท่องเที่ยวกล่าวคือ มีข้อมูลไม่เพียงพอในการสรุปความสนใจสำหรับผู้ใช้ที่เริ่มต้นใช้งานระบบหรือเรียกว่า ปัญหาโคลด์สตาร์ท (Cold Start Problem) ด้วยความสำคัญของปัญหาดังกล่าวจึงมีหลายแนวคิดที่จะพัฒนาแบบจำลองผู้ใช้ที่สามารถแก้ปัญหาและได้ข้อสรุปความสนใจของผู้ใช้ได้อย่างถูกต้องและมีประสิทธิภาพ ปัจจุบันการใช้เครือข่ายสังคมที่มีข้อมูลของผู้ใช้ที่หลากหลาย ข้อมูลเหล่านี้ถูกนำมาสกัดเพื่อใช้เป็นตัวแทนของความสนใจของผู้ใช้ได้ ในบทความนี้ได้รวบรวมแนวคิดต่างๆ ในการใช้ประโยชน์จากเครือข่ายสังคมและการสกัดความสนใจจากผู้ใช้เครือข่ายสังคมสำหรับระบบแนะนำข้อมูล (Recommendation System) รวมถึงวิเคราะห์และชี้ให้เห็นถึงความท้าทายด้านต่างๆ ที่มีสำหรับการสร้างแบบจำลองผู้ใช้ในระบบแนะนำข้อมูลโดยใช้ข้อมูลจากเครือข่ายสังคมเป็นหลัก

คำสำคัญ: ปัญหาโคลด์สตาร์ท การสกัดความสนใจของผู้ใช้ เครือข่ายสังคม ระบบแนะนำข้อมูล

ภาควิชาวิทยาการคอมพิวเตอร์และเทคโนโลยีสารสนเทศ คณะวิทยาศาสตร์ มหาวิทยาลัยนเรศวร

*ผู้นิพนธ์ประสานงาน, e-mail: kraisakk@nu.ac.th

An Overview of User Modeling Approaches and Their Social Network Information Exploration for Tourism Recommendation Systems

Wanarat Juraphanthong and Kraisak Kesorn*

ABSTRACT

The Tourism Recommendation System (TRS) is useful for tourists in itinerary planning and for tourism stakeholders in advertising. TRS allows tourists to spend less time searching for information and enables easier decision making to buy services from tourism providers. The key for TRS is a user model to store user preference information in such a way that it is easily retrievable for analysis. However, constructing a user model is very challenging. For example, there is inadequate information for a new user; this is called the "cold start problem". Therefore, several researchers are trying to overcome such problems by exploiting useful information from social networks freely available on the Internet. In this paper, we have surveyed and summarized the ideas for user interest extraction from various social networks. In addition, several challenges for constructing a user model for TRS are addressed.

Keywords: cold start problem, user interests extraction, social network, recommendation system

Department of Computer Science and Information Technology, Faculty of Science, Naresuan University
*Corresponding author: e-mail: kraisakk@nu.ac.th

1. บทนำ

อุตสาหกรรมการท่องเที่ยวถือได้ว่าเป็นภาคส่วนที่สำคัญในประเทศที่กำลังพัฒนา โดยเฉพาะในเอเชียตะวันออกเฉียงใต้ เนื่องจากเป็นอุตสาหกรรมที่สามารถสร้างรายได้หลักให้กับแต่ละประเทศ โดยในประเทศไทยนอกจากเป็นประเทศหลักที่นักท่องเที่ยวต้องการเดินทางเข้ามาท่องเที่ยวในเอเชียตะวันออกเฉียงใต้แล้ว [1] จากข้อมูลพบว่าในปี 2011 อุตสาหกรรมการท่องเที่ยวยังสร้างรายได้เป็นมูลค่ากว่าสองหมื่นหกพันล้านดอลลาร์สหรัฐซึ่งคิดเป็นร้อยละ 8.2 ของผลิตภัณฑ์มวลรวมของประเทศ [2] การพัฒนาอุตสาหกรรมการท่องเที่ยวจึงเป็นสิ่งที่ภาคส่วนต่างๆ ทั้งรัฐบาลและเอกชนได้ดำเนินการพัฒนาในหลายด้านอย่างต่อเนื่อง อินเทอร์เน็ตมีบทบาทต่อการพัฒนาอุตสาหกรรมการท่องเที่ยวอย่างมาก ทั้งเว็บไซต์สำหรับให้บริการข้อมูลการท่องเที่ยวและพัฒนาขึ้นเป็นพาณิชย์อิเล็กทรอนิกส์ ซึ่งเป็นเครื่องมือที่ช่วยเผยแพร่ข้อมูลการท่องเที่ยวได้อย่างมีประสิทธิภาพและเป็นผลดีต่อนักท่องเที่ยวที่สามารถหาข้อมูลการท่องเที่ยวได้หลากหลายขึ้น แต่ข้อมูลเกี่ยวกับการท่องเที่ยวจำนวนมากที่เกิดขึ้นนี้ทำให้นักท่องเที่ยวต้องใช้เวลาเพิ่มมากขึ้นในการค้นหาข้อมูลการท่องเที่ยวให้ได้ตรงกับความต้องการ ดังนั้นเพื่อลดเวลาการค้นหาข้อมูลของนักท่องเที่ยว ระบบแนะนำข้อมูลเป็นวิธีหนึ่งที่ช่วยแก้ไขปัญหาดังกล่าวได้ โดยระบบแนะนำข้อมูลนั้นถูกนำมาใช้อย่างแพร่หลายในธุรกิจพาณิชย์อิเล็กทรอนิกส์ [3] เช่น ระบบแนะนำข้อมูลหนังสือ วารสาร ข่าว รายการโทรทัศน์ เพลง เว็บเพจ การท่องเที่ยว ฯลฯ [4] ต่อมาได้มีการพัฒนาระบบแนะนำข้อมูลในรูปแบบการแนะนำข้อมูลเฉพาะบุคคล (Personalized Recommendation System) ที่จะช่วยให้ผู้ใช้สามารถค้นหาสินค้าและบริการที่ต้องการแบบเฉพาะบุคคลได้ในเวลาที่รวดเร็วขึ้น ในขอบเขตของการท่องเที่ยว ระบบแนะนำข้อมูลเฉพาะบุคคลได้รับความสนใจและนำมาใช้สนับสนุนนักท่องเที่ยวทั้งในการวางแผนการเดินทาง [5] ช่วยจองบัตรโดยสารเครื่องบินในวันหยุดเทศกาล [6] แนะนำแพคเกจท่องเที่ยว [7] เป็นต้น นอกจากนักท่องเที่ยวจะได้สินค้าและบริการที่ตรงกับความต้องการ ผู้ให้บริการยังได้ประโยชน์จากการเสนอสินค้าและบริการที่เหมาะสม ซึ่งเป็นการสร้างกลยุทธ์ทางการตลาดเพื่อเพิ่มจำนวนผู้ใช้บริการให้มากยิ่งขึ้น

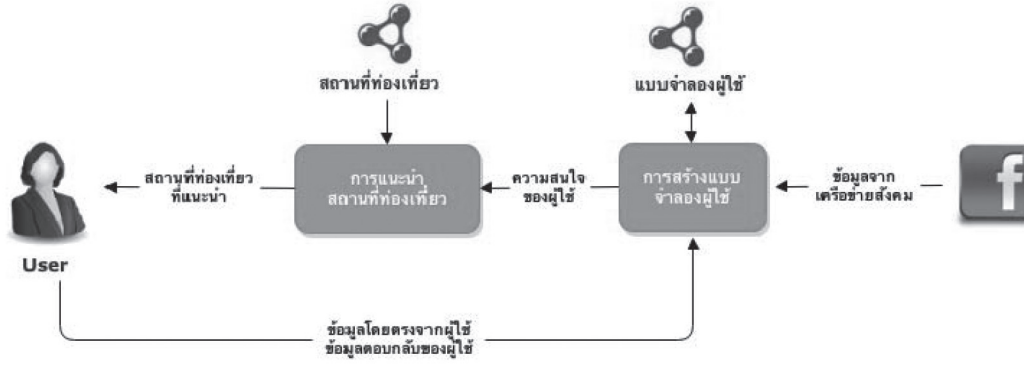
ในระบบแนะนำข้อมูลเฉพาะบุคคลนั้น แบบจำลองผู้ใช้ (User Model) เป็นส่วนประกอบสำคัญ [8] ที่จะต้องใช้ข้อมูลสำคัญจำนวนมากในการสกัดเพื่อหาความสนใจของผู้ใช้ ซึ่งเป็นหัวใจสำคัญในการแนะนำผลิตภัณฑ์หรือบริการให้ได้ตรงกับความต้องการของผู้ใช้อย่างแท้จริง อย่างไรก็ตาม การสร้างแบบจำลองผู้ใช้มีความท้าทายอย่างยิ่งสำหรับผู้พัฒนาระบบคือปัญหาโคลด์สตาร์ท (Cold Start Problem) กล่าวคือมีข้อมูลไม่เพียงพอในการสรุปความสนใจสำหรับผู้ใช้ที่เริ่มต้นใช้งานระบบ ด้วยความสำคัญของปัญหาดังกล่าวจึงมีหลายแนวคิดที่จะพัฒนาแบบจำลองผู้ใช้ที่สามารถแก้ปัญหาและได้ข้อสรุปความสนใจของผู้ใช้ได้อย่างถูกต้องและมีประสิทธิภาพ เครือข่ายสังคม (Social Network) นอกจากจะช่วยโฆษณา ประชาสัมพันธ์การท่องเที่ยวได้แล้ว ยังเป็นที่เก็บข้อมูลจำนวนมากจากผู้ใช้ ประกอบกับการเจริญเติบโตของเครือข่ายสังคมที่เป็นไปอย่างรวดเร็ว ทำให้ข้อมูลทั้งที่เกิดจากการสนทนา แสดงความคิดเห็น แบ่งปัน มีจำนวนที่มากขึ้น ข้อมูลจำนวนมหาศาลเหล่านี้ยังไม่ถูกนำไปใช้ประโยชน์เท่าที่ควร อย่างไรก็ตาม นักวิจัยส่วนหนึ่งพยายามจะใช้ประโยชน์จากข้อมูลเหล่านี้โดยการนำข้อมูลจากเครือข่ายสังคมมาวิเคราะห์เพื่อหาความสนใจของผู้ใช้ เช่น การใช้แท็ก (Tag) ในโซเชียลบุ๊กมาร์ก (Social Bookmarking) หาความสนใจของผู้ใช้เพื่อสร้าง

คำแนะนำในระบบค้นหาเพลง [9] การใช้ข้อความทวีต (Tweet) ในทวิตเตอร์ (Twitter.com) ค้นหาความสนใจของผู้ใช้เพื่อสร้างระบบแนะนำข่าว [10] เป็นต้น เครือข่ายสังคมจึงเป็นแหล่งข้อมูลที่สามารถนำไปสู่ข้อสรุปความสนใจและสามารถนำมาประยุกต์ใช้ในการพัฒนาแบบจำลองผู้ใช้สำหรับระบบแนะนำข้อมูลการท่องเที่ยวแบบเฉพาะบุคคลอย่างมีประสิทธิภาพได้นอกจากนี้ยังมีแนวโน้มที่จะสามารถแก้ปัญหาโคลด์สตาร์ทและลดภาระของผู้ใช้ที่ต้องทำการกรอกข้อมูลความสนใจด้วยตนเองในระบบเดิมอีกด้วย

ในบทความนี้ได้ทำการทบทวนวรรณกรรมและรวบรวมแนวคิดของการใช้ประโยชน์จากเครือข่ายสังคมเพื่อนำมาสร้างระบบแนะนำการท่องเที่ยว โดยแบ่งออกเป็น 7 หัวข้อดังนี้ 1) บทนำ 2) ระบบแนะนำข้อมูลการท่องเที่ยว 3) การสร้างแบบจำลองผู้ใช้ 4) เครือข่ายสังคมกับการประยุกต์ใช้ด้านการท่องเที่ยว 5) เทคนิคการพัฒนาแบบจำลองผู้ใช้โดยใช้เครือข่ายสังคม 6) ความท้าทายในการพัฒนาแบบจำลองผู้ใช้ และ 7) บทสรุป

2. ระบบแนะนำข้อมูลการท่องเที่ยว

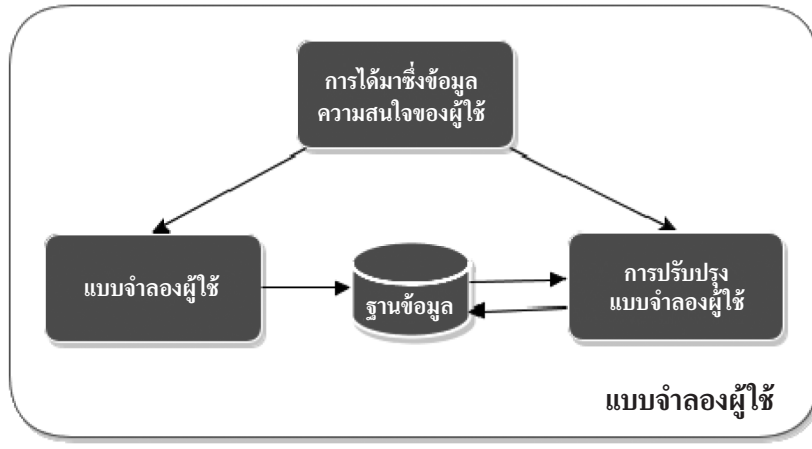
ระบบแนะนำข้อมูลการท่องเที่ยวเฉพาะบุคคลเป็นระบบที่นำเสนอข้อมูลการท่องเที่ยวในลักษณะต่างๆ เช่น ข้อมูลสถานที่ท่องเที่ยว แผนการเดินทาง ข้อมูลแพคเกจท่องเที่ยว เป็นต้น ให้กับนักท่องเที่ยวตามความสนใจที่แตกต่างกันไปในแต่ละบุคคล โดยระบบแนะนำข้อมูลการท่องเที่ยวเฉพาะบุคคลที่ผู้วิจัยกำลังพัฒนาเป็นการแนะนำข้อมูลสถานที่ท่องเที่ยวให้กับนักท่องเที่ยว ซึ่งมีโครงสร้างหลักดังรูปที่ 1 ประกอบด้วย 1) การสร้างแบบจำลองผู้ใช้ (User Modeling) เป็นการหาข้อสรุปความสนใจและปรับปรุงให้ทันต่อความสนใจที่เปลี่ยนไปของผู้ใช้ผ่านแหล่งที่มาหลายแหล่ง เช่น ข้อมูลที่ได้จากผู้ใช้โดยตรง ข้อมูลจากเครือข่ายสังคม ข้อมูลตอบกลับ เป็นต้น ซึ่งผู้วิจัยได้ทำการทบทวนวรรณกรรมเกี่ยวกับการสร้างแบบจำลองผู้ใช้ในหัวข้อที่ 3 2) การแนะนำสถานที่ท่องเที่ยว (Tourism Attractions Recommendation) เป็นการแนะนำสถานที่ท่องเที่ยวโดยใช้ข้อมูลความสนใจที่ได้มาจากแบบจำลองผู้ใช้ ซึ่งในบทความนี้จะยังไม่ขอกล่าวถึงส่วนนี้



รูปที่ 1 โครงสร้างหลักของระบบแนะนำข้อมูลการท่องเที่ยวเฉพาะบุคคลที่ผู้วิจัยพัฒนา

3. การสร้างแบบจำลองผู้ใช้

ในระบบแนะนำข้อมูลการท่องเที่ยวเฉพาะบุคคลนั้น ระบบต้องการข้อมูลที่เป็นข้อสรุปเกี่ยวกับความสนใจของผู้ใช้เพื่อนำไปใช้ในการแนะนำข้อมูลการท่องเที่ยวที่ตรงกับความต้องการของผู้ใช้ ข้อมูลความสนใจเหล่านี้จะถูกจัดเก็บอยู่ในแบบจำลองผู้ใช้ โดยการสร้างแบบจำลองผู้ใช้ประกอบด้วย 3 องค์ประกอบสำคัญ คือ 1) ประเภทของแบบจำลองผู้ใช้ (Type of User Modeling) 2) การได้มาซึ่งข้อมูลความสนใจของผู้ใช้ (Information Acquisition) และ 3) การปรับปรุงแบบจำลองผู้ใช้ (User Model Adaptation) ซึ่งการสร้างแบบจำลองผู้ใช้ในแต่ละระบบจะต้องวิเคราะห์ว่าจะใช้วิธีการได้มาซึ่งข้อมูลความสนใจของผู้ใช้ด้วยวิธีการใด จากนั้นข้อมูลที่ได้มาโดยวิธีการที่เลือกข้างต้นจะถูกส่งไปยังแบบจำลองผู้ใช้ สำหรับผู้ใช้ที่เริ่มต้นใช้งานระบบเพื่อเลือกประเภทและสร้างแบบจำลองผู้ใช้เริ่มแรก และข้อมูลดังกล่าวจะถูกส่งไปยังการปรับปรุงแบบจำลองผู้ใช้ สำหรับผู้ใช้ที่ได้ใช้งานระบบไปแล้วเพื่อปรับเปลี่ยนความสนใจของผู้ใช้ให้ถูกต้องยิ่งขึ้น นอกจากนี้แบบจำลองผู้ใช้และการปรับปรุงแบบจำลองผู้ใช้จะติดต่อกับฐานข้อมูลเพื่อทำการจัดเก็บข้อมูลแบบจำลองผู้ใช้ โดยทั้งสามองค์ประกอบแสดงดังรูปที่ 2



รูปที่ 2 องค์ประกอบในการสร้างแบบจำลองผู้ใช้

3.1 ประเภทของแบบจำลองผู้ใช้ (Type of User Modeling)

การสร้างแบบจำลองผู้ใช้แบ่งออกเป็น 2 ลักษณะ คือ การสร้างแบบจำลองผู้ใช้แบบกลุ่ม (Group Modeling) และการสร้างแบบจำลองผู้ใช้แบบเฉพาะบุคคล (Personalization Modeling) สำหรับการสร้างแบบจำลองผู้ใช้แบบกลุ่มเป็นการจัดผู้ใช้ที่มีลักษณะคล้ายกันไว้ในกลุ่มหรือหมวดหมู่เดียวกันเพื่อแนะนำสิ่งที่สนใจเหมือนกันไปยังกลุ่ม ส่วนการสร้างแบบจำลองผู้ใช้แบบเฉพาะบุคคลเป็นการสร้างแบบจำลองผู้ใช้เพื่อหาความสนใจของผู้ใช้เป็นรายบุคคลเพื่อแนะนำสิ่งที่สนใจของแต่ละบุคคล ตัวอย่างของแบบจำลองผู้ใช้แบบกลุ่ม เช่น TravelPlanner [11] ใช้กลุ่มแอตทริบิวต์ (Stereo-type) ที่กำหนดลักษณะของผู้ใช้ เช่น ทักษะ ความต้องการ และระดับความรู้ เป็นต้น ในการจัดผู้ใช้ให้อยู่เป็นกลุ่มเดียวกัน และทำการแนะนำข้อมูลสถานที่ท่องเที่ยวที่ตรงกับความสนใจของกลุ่มผู้ใช้ วิธีการสร้างแบบจำลองผู้ใช้แบบกลุ่มนี้มี

ประโยชน์กับระบบแนะนำข้อมูลที่ผู้ใช้เริ่มต้นใช้งานและยังไม่มีข้อมูลของผู้ใช้มากพอที่จะสกัดความสนใจเพื่อแนะนำข้อมูลให้ตรงกับความสนใจได้ ซึ่งเรียกปัญหานี้ว่า โคลด์สตาร์ท การจัดกลุ่มผู้ใช้และแนะนำข้อมูลเช่นเดียวกับกลุ่มที่ผู้ใช้อยู่เป็นวิธีการหนึ่งที่จะสามารถจัดการปัญหานี้ได้ อย่างไรก็ตาม ข้อเสียของวิธีนี้คือแต่ละผู้ใช้อาจจะมีความคล้ายกันในกลุ่มแอตทริบิวต์ที่นำมาใช้จัดกลุ่มผู้ใช้ แต่ในความเป็นจริงผู้ใช้แต่ละบุคคลนั้นยังคงมีความแตกต่างกันในสิ่งที่สนใจ

TripleHop's TripMatcher (ski-europe.com) [4] และ GUIDE [12] สร้างแบบจำลองผู้ใช้เป็นรายบุคคล โดยการสร้างบัญชีผู้ใช้และใส่ข้อมูลที่เป็นรายละเอียดส่วนบุคคลเช่น ชื่อ อายุ เพศ รวมทั้งความสนใจของผู้ใช้ (เช่น อาหารที่ชื่นชอบ) ที่อยู่ ณ ปัจจุบัน สถานที่สำคัญต่างๆ ที่ผู้ใช้เข้าเยี่ยมชมและท่องเที่ยว เป็นต้น วิธีการสร้างแบบจำลองผู้ใช้ลักษณะนี้ช่วยระบุความสนใจที่แตกต่างกันในแต่ละบุคคลได้ดีกว่าการสร้างแบบจำลองผู้ใช้แบบกลุ่มเนื่องจากพิจารณากลุ่มคลาสเป็นแบบรายบุคคล อย่างไรก็ตาม วิธีนี้มีข้อเสียคือเมื่อผู้ใช้เริ่มต้นใช้ระบบ ก็อาจจะเกิดปัญหาโคลด์สตาร์ทได้เช่นกัน

3.2 การได้มาซึ่งข้อมูลความสนใจของผู้ใช้ (Information Acquisition)

การได้มาซึ่งข้อมูลความสนใจของผู้ใช้ แบ่งเป็น 2 ลักษณะ คือ การได้มาซึ่งข้อมูลความสนใจของผู้ใช้โดยตรง (Explicit) ซึ่งคือการที่ระบบขอข้อมูลจากผู้ใช้โดยตรง เช่น ข้อมูลรายละเอียดส่วนบุคคลที่ผู้ใช้กรอกให้กับระบบเมื่อลงทะเบียน การตอบคำถามของผู้ใช้ การทำแบบสอบถาม เป็นต้น สำหรับการได้มาซึ่งข้อมูลความสนใจของผู้ใช้โดยอ้อม (Implicit) คือการที่ระบบทำการเก็บข้อมูลจากการใช้งานหรือพฤติกรรมของผู้ใช้ที่ไม่เป็นการขอข้อมูลจากผู้ใช้โดยตรง เช่น การคลิก การโหวต การแสดงความคิดเห็นบนระบบ เป็นต้น ซึ่ง Entrée [13] และ TripleHop's TripMatcher ใช้ประโยชน์จากข้อมูลที่สอบถามจากผู้ใช้โดยตรง แต่ปัญหาที่เกิดคือผู้ใช้อาจต้องสูญเสียเวลาในขั้นตอนการตอบคำถามเหล่านี้มากเกินไปซึ่งเป็นการรบกวนผู้ใช้ นอกจากนี้ผู้ใช้ยังอาจไม่สามารถอธิบายลักษณะความพึงพอใจของตนเองได้อย่างถูกต้องผ่านวิธีการนี้ ในขณะที่ PTA [6] และ GUIDE [12] ใช้การได้มาซึ่งข้อมูลความสนใจของผู้ใช้โดยอ้อมผ่านแหล่งที่มาหลายแหล่ง วิธีการนี้สามารถขจัดปัญหาที่เกิดกับระบบที่ขอข้อมูลจากผู้ใช้โดยตรงได้ อย่างไรก็ตาม ยังมีปัญหาสำคัญนั่นคือเวลาที่ไม่เพียงพอให้ระบบสกัดความสนใจจากข้อมูลโดยอ้อมได้อย่างถูกต้อง เพื่อลดปัญหาระหว่างการได้มาซึ่งข้อมูลความสนใจของผู้ใช้ทั้งสองวิธีข้างต้น จึงมีการผนวกทั้งสองวิธีเข้าด้วยกัน เช่นใน TravelPlanner จะเลือกแบบสอบถามที่มีประโยชน์ที่สุดในการเรียนรู้ความสนใจให้กับผู้ใช้และใช้ข้อมูลอื่นๆ ที่ได้มาจากทางอ้อมด้วย เช่นเดียวกับ Zhu และคณะ [8] ใช้ข้อมูลที่ผู้ใช้ได้ให้กับระบบเมื่อทำการลงทะเบียนเพื่อสร้างแบบจำลองผู้ใช้เริ่มต้น จากนั้นจะใช้ข้อมูลโดยอ้อมผ่านพฤติกรรมผู้ใช้ เช่น พฤติกรรมการคลิก พฤติกรรมการโหวต พฤติกรรมการตอบและพูดคุย เป็นต้น เพื่อเรียนรู้ความสนใจในลำดับต่อมา ในทำนองเดียวกันกับ SPETA [14] เมื่อผู้ใช้เริ่มต้นใช้งานระบบ ระบบจะขอข้อมูลเกี่ยวกับความสนใจ ประเภทสถานที่ที่ต้องการเข้าไปเยี่ยมชม และการให้คะแนนให้กับสถานที่ที่ผู้ใช้น่าสนใจ จากนั้นก็จะเรียนรู้ความสนใจผ่านพฤติกรรมของผู้ใช้และข้อมูลที่ถูกสกัดจากเครือข่ายสังคมที่ได้เป็นสมาชิกอยู่

3.3 การปรับปรุงแบบจำลองผู้ใช้ (User Model Adaptation)

การเปลี่ยนแปลงความสนใจของผู้ใช้เป็นอีกหนึ่งหัวข้อหลักของแบบจำลองผู้ใช้ที่นักวิจัยพยายามพัฒนาเพื่อให้รองรับการเปลี่ยนแปลงความสนใจของผู้ใช้ที่สามารถเกิดขึ้นได้ตลอดเวลา เพื่อให้ระบบแนะนำข้อมูลสามารถแนะนำข้อมูลได้เหมาะสมกับความสนใจของผู้ใช้ในขณะนั้นมากที่สุด โดย Widyantoro และคณะ [15] ได้เสนออัลกอริทึมในการเรียนรู้ความสนใจของผู้ใช้ โดยแบ่งความสนใจของผู้ใช้เป็นแบบระยะสั้นเพื่อเรียนรู้ความสนใจในช่วงเวลาสั้นๆ ขณะปัจจุบัน และแบบระยะยาวเพื่อเรียนรู้ความสนใจโดยรวม จากนั้นจะคำนวณค่าน้ำหนักไปที่แต่ละแบบและเลือกใช้ความสนใจที่มีค่าน้ำหนักมากที่สุด วิธีนี้สามารถเพิ่มความถูกต้องในการค้นหาความสนใจที่เปลี่ยนแปลงไปตามช่วงเวลาได้ และสามารถปรับปรุงแบบจำลองผู้ใช้ให้ถูกต้องตรงกับความสนใจได้ดีขึ้น สำหรับการเปลี่ยนแปลงความสนใจของผู้ใช้แบบทันที ยังไม่ถูกนำมากล่าวถึงมากนัก Zhu และคณะ [8] ได้เสนอการสร้างแบบจำลองความสนใจของผู้ใช้โดยนำข้อมูลโดยอ้อม เช่น พฤติกรรมการคลิก พฤติกรรมการโหวต เป็นต้น จากนั้นคำนวณหาค่าความสนใจในแต่ละเรื่องร่วมกับการคำนวณด้านเวลา (เช่น การกดโหวตเมื่อวานนี้ย่อมมีนัยสำคัญเกี่ยวกับความสนใจมากกว่าการโหวตเมื่อปีที่แล้ว) ค่าความสนใจทั้งหมดของผู้ใช้จะถูกเก็บไว้ในเมทริกซ์และเปลี่ยนแปลงค่าด้วยตัวเองเมื่อผู้ใช้เปลี่ยนแปลงพฤติกรรมต่างๆ ข้างต้น ค่าความสนใจที่ได้จากวิธีการนี้สามารถนำไปใช้แนะนำผลิตภัณฑ์หรือบริการได้ตามลำดับค่าความสนใจ นอกจากนี้การเปลี่ยนแปลงค่าความสนใจตามพฤติกรรมในปัจจุบันยังช่วยรองรับการเปลี่ยนแปลงความสนใจในแต่ละช่วงเวลาได้ดียิ่งขึ้น Mezhoudi [16] ใช้ประโยชน์จากข้อมูลตอบกลับของผู้ใช้ (User Feedback) ปรับปรุงแบบจำลองผู้ใช้ให้กับระบบ โดยใช้ข้อมูลตอบกลับจากการปฏิสัมพันธ์ของผู้ใช้ใน 2 แบบ คือ 1) Emoticons Based Feedback ซึ่งเป็นการแสดงระดับความพึงพอใจของผู้ใช้ผ่านทางไอคอนแสดงอารมณ์ ตัวอย่างการนำไปประยุกต์ใช้ เช่น การตอบคำถามข้อเสนอแนะด้วยไอคอนแสดงอารมณ์ในแอพลิเคชันบนโทรศัพท์มือถือ [17] 2) Recommendation Frames เป็นการเก็บข้อมูลจากการปฏิสัมพันธ์โดยการแนะนำ การแนะนำอาจจะแสดงเป็นหน้าต่างป็อปอัพหรือแสดงบนหน้าเพจ ซึ่งใช้กันมากในอีคอมเมิร์ซเพื่อให้คำแนะนำแก่ลูกค้า การนำข้อมูลตอบกลับลักษณะดังกล่าวมาใช้ในการปรับเปลี่ยนแบบจำลองความสนใจของผู้ใช้ เป็นส่วนเสริมในการประเมินความถูกต้องหรือพึงพอใจในการให้คำแนะนำของระบบ และยังทำให้ระบบนำข้อมูลเหล่านี้ไปปรับเปลี่ยนแบบจำลองผู้ใช้ให้ตรงกับความสนใจของผู้ใช้ได้ทันท่วงที

จากทั้ง 3 องค์ประกอบข้างต้น ในองค์ประกอบแรกพบว่าการเลือกใช้การสร้างแบบจำลองผู้ใช้แบบกลุ่มนั้นสามารถช่วยแก้ปัญหาโคลด์สตาร์ทได้เนื่องจากสามารถใช้ข้อมูลที่เคยแนะนำไปยังผู้ใช้ที่อยู่ในกลุ่มเดียวกันมาแนะนำให้กับผู้ใช้ที่ระบบยังไม่มีข้อมูลมากพอ แต่การแนะนำข้อมูลแบบกลุ่มนี้ไม่รองรับกับความสนใจที่แตกต่างกันในแต่ละบุคคล ในทางกลับกันการสร้างแบบจำลองผู้ใช้แบบเฉพาะบุคคลสามารถตอบสนองความสนใจที่แตกต่างกันระหว่างบุคคลได้ แต่ปัญหาสำคัญที่เกิดตามมาคือปัญหาโคลด์สตาร์ทที่เกิดขึ้นเมื่อผู้ใช้เริ่มต้นใช้งานระบบนั่นเอง องค์ประกอบต่อมาคือ การเลือกใช้วิธีการได้มาซึ่งข้อมูลความสนใจของผู้ใช้โดยตรงสามารถลดปัญหาโคลด์สตาร์ทที่อาจเกิดขึ้นเมื่อใช้วิธีการได้มาซึ่งข้อมูลความสนใจของผู้ใช้โดยอ้อม อย่างไรก็ตาม การขอข้อมูลจากผู้ใช้โดยตรงบางวิธีจะรบกวนผู้ใช้มากเกินไป เพราะผู้ใช้ต้องเสียเวลากับขั้นตอนการกรอกข้อมูลเหล่านี้ ดังนั้นการผสมผสานทั้งสองวิธีการเข้าด้วยกันจึงเป็นอีกทางเลือกหนึ่งที่สามารถลดปัญหาที่เกิดขึ้นจากวิธีการทั้งสองได้ และองค์ประกอบสุดท้ายคือ การปรับปรุงแบบ

จำลองผู้ใช้ ซึ่งมีทั้งแนวคิดการแบ่งความสนใจออกเป็นระยะสั้นและระยะยาว การจับพฤติกรรมของผู้ใช้เพื่อปรับเปลี่ยนลักษณะความสนใจ การใช้ข้อมูลตอบกลับเพื่อเป็นส่วนเสริมในการประเมินความถูกต้องหรือพึงพอใจในการให้คำแนะนำของระบบ นอกจากนี้การสร้างแบบจำลองผู้ใช้ยังมีส่วนที่ต้องคำนึงถึงอื่นๆ อีก เช่น ลักษณะการตรวจสอบแบบจำลองผู้ใช้และลักษณะการจัดเก็บข้อมูลซึ่งจะกล่าวถึงในหัวข้อที่ 6.3 และ 6.4 ตามลำดับ ลักษณะและองค์ประกอบ ในระบบแนะนำข้อมูลแบบต่างๆ แสดงดังตารางที่ 1

ตารางที่ 1 ลักษณะและองค์ประกอบในระบบแนะนำข้อมูลแบบต่างๆ

ระบบแนะนำข้อมูล	ประเภทของแบบจำลอง	การได้มาซึ่งข้อมูล	การปรับปรุงแบบจำลอง	การเก็บข้อมูล	ตรวจสอบแบบจำลอง	ปัญหาโคลด์สตาร์ท
Entrée [13]	กลุ่ม	โดยตรง	ไม่มี	ไม่มี	ข้อมูลตอบกลับ	ถูกแก้ไข
TripleHop's TripMatcher [4]	บุคคล	โดยตรง	ไม่ได้ระบุ	ไม่ได้ระบุ	ไม่ได้ระบุ	ถูกแก้ไข
GUIDE [4]	บุคคล	โดยอ้อม	พฤติกรรมผู้ใช้	ไม่ได้ระบุ	ไม่ได้ระบุ	ไม่ได้ถูกแก้ไข
CAPA [18]	บุคคล	ผสมผสาน	ระยะสั้นและระยะยาว	ไม่ได้ระบุ	ข้อมูลตอบกลับ	ถูกแก้ไข
PTS [19]	บุคคล	โดยอ้อม	พฤติกรรมผู้ใช้	ฐานข้อมูล	การคลิกเลือก	ไม่ได้ถูกแก้ไข
UMT [20]	กลุ่ม	โดยอ้อม	ไม่มี	ฐานข้อมูล	ไม่ได้ระบุ	ถูกแก้ไข
SPETA [14]	บุคคล	ผสมผสาน	พฤติกรรมผู้ใช้	ออนโทโลยี	ไม่ได้ระบุ	ถูกแก้ไข
Zhu และคณะ [8]	บุคคล	ผสมผสาน	พฤติกรรมผู้ใช้	ไม่ได้ระบุ	ข้อมูลตอบกลับ	ถูกแก้ไข

จะเห็นว่าปัญหาโคลด์สตาร์ทเป็นปัญหาสำคัญในระบบแนะนำข้อมูลแบบเฉพาะบุคคลด้านต่างๆ รวมถึงการแนะนำข้อมูลการท่องเที่ยวที่ผู้วิจัยพยายามพัฒนาขึ้น ซึ่งการใช้ประโยชน์จากข้อมูลที่มีในเครือข่ายสังคมนั้นเป็นวิธีที่สามารถลดปัญหาโคลด์สตาร์ทและลดการรบกวนผู้ใช้ได้ (อธิบายเพิ่มเติมในหัวข้อที่ 6.1) โดยใช้การผสมระหว่างการใช้ข้อมูลสาธารณะ (Public) และการขออนุญาตใช้ข้อมูลส่วนตัว (Private) จากเครือข่ายสังคมที่ผู้ใช้ใช้งานอยู่ ซึ่งเครือข่ายสังคมและบทบาทของเครือข่ายสังคมในการพัฒนาสารสนเทศด้านการท่องเที่ยวที่ผ่านมาจะอธิบายในหัวข้อถัดไป

4. เครือข่ายสังคมกับการประยุกต์ใช้ด้านการท่องเที่ยว

ในปัจจุบันเครือข่ายสังคม (Social Network) มีการเจริญเติบโตขึ้นอย่างรวดเร็ว เนื่องจากได้กลายเป็นเครื่องมือสื่อสารสำคัญของบุคคล โดยเครือข่ายเหล่านี้ให้พื้นที่ในการติดต่อสื่อสาร การสร้างและแบ่งปันข้อมูลระหว่างเพื่อนและบุคคลอื่นๆ ทำให้เกิดข้อมูลจำนวนมากมายบนเครือข่ายเหล่านี้ เช่น ข้อมูล ความคิดเห็น ข่าว ภาพ วีดิโอ แม้กระทั่งบทวิจารณ์ในผลิตภัณฑ์และบริการ เป็นต้น และถูกเผยแพร่ข้อมูลได้อย่างรวดเร็ว

ในด้านการท่องเที่ยว พบว่าเครือข่ายสังคมมีความสำคัญต่อธุรกิจการท่องเที่ยวอย่างมาก ทั้งกับผู้ประกอบการ และนักท่องเที่ยว โดย Miguéns [21] ได้ศึกษาการทำการตลาดของโรงแรมที่ตั้งอยู่ที่เมืองลิสบอนบนเว็บไซต์ TripAdvisor.com ซึ่งผลการวิจัยพบว่าผู้ประกอบการโรงแรมใช้เครือข่ายสังคมเป็นเครื่องมือช่วยทำการตลาดในระดับที่ต่ำ จึงส่งผลต่อการจัดอันดับโรงแรมบนเว็บไซต์ การศึกษานี้เป็นการยืนยันถึงความสำคัญของเครือข่ายสังคมที่มีต่อทั้งผู้ประกอบการและสถานที่ท่องเที่ยวเหล่านั้น ซึ่งจะเห็นว่าในส่วนของผู้ประกอบการหากทราบและเข้าใจถึงบทบาทของเครือข่ายสังคมและพัฒนาเทคโนโลยีเหล่านี้ก็จะสามารถปรับปรุงตำแหน่งทางการตลาดให้ดีขึ้นได้

โดยในปัจจุบันจึงพบผู้ประกอบการจำนวนมากใช้เครือข่ายสังคมช่วยในการทำการตลาด ทั้งการโฆษณาผลิตภัณฑ์และบริการ การจัดกิจกรรมส่งเสริมการขาย หรือแม้แต่ใช้ในการขายผลิตภัณฑ์และบริการโดยตรงผ่านเครือข่ายสังคม เช่น The BALSAMS Grand Resort Hotel ใช้เว็บไซต์และเครือข่ายสังคม เช่น เฟสบุค ทวิตเตอร์ ยูทูป (YouTube.com) ฟลิคเกอร์ (Flickr.com) บอกเล่าเรื่องราวที่เป็นเอกลักษณ์ของโรงแรมผ่านแคมเปญที่ชื่อว่า “Resorter InnBedded” เป็นแคมเปญที่หาผู้ที่สนใจจะใช้ชีวิตที่โรงแรมระหว่างเดือนกรกฎาคมถึงสิงหาคม พร้อมโพสภาพถ่ายวีดีโอ และข้อมูลที่โรงแรมต้องการประชาสัมพันธ์ในทุกๆ วันตลอดช่วงเวลาของแคมเปญ การทำการตลาดในแคมเปญนี้ส่งผลให้มียอดจองห้องพักเพิ่มขึ้นถึง 20% ในเดือนสิงหาคม [22]

ในส่วนของนักท่องเที่ยว เครือข่ายสังคมถูกใช้เป็นแหล่งข้อมูลที่ช่วยให้นักท่องเที่ยววางแผนการเดินทางและตัดสินใจที่จะท่องเที่ยว จะเห็นได้จากการเกิดขึ้นของเว็บไซต์เครือข่ายสังคมที่เกี่ยวกับการท่องเที่ยว เช่น Tripbod (Tripbod.com), VirtualTourist (Virtualtourist.com) และ Trippy (Trippy.com) เป็นเว็บเครือข่ายสังคมที่ให้นักท่องเที่ยวได้แบ่งบันประสบการณ์การท่องเที่ยวระหว่างกัน TripIt (Tripit.com) เป็นเครือข่ายสังคมสำหรับช่วยนักท่องเที่ยววางแผนการเดินทาง AirBnB (Airbnb.com) ให้สมาชิกสามารถรีวิวห้องพัก บ้าน อพาร์ทเม้น เพื่อเปิดให้สมาชิกท่านอื่นๆ เข้ามาจองได้ [23, 24] Tripadvisor (Tripadvisor.com) เป็นเว็บไซต์เครือข่ายสังคมที่นอกจากจะช่วยวางแผนการเดินทาง ยังเป็นชุมชนการท่องเที่ยวขนาดใหญ่ที่มีทั้งการรีวิวสถานที่ท่องเที่ยว โรงแรม ร้านอาหาร ฯลฯ รวมถึงข้อมูลความคิดเห็นและประสบการณ์การท่องเที่ยวของนักท่องเที่ยวท่านอื่น

นอกจากนี้จากการที่เครือข่ายทางสังคมได้รับความนิยมอย่างต่อเนื่องยังทำให้นักท่องเที่ยวมีอำนาจต่อรองเพิ่มขึ้นในเชิงธุรกิจ [25] เนื่องจากความคิดเห็นและประสบการณ์การท่องเที่ยวในเครือข่ายสังคมที่มีความน่าเชื่อถือเป็นตัวแทนสำคัญที่ส่งผลต่อการตัดสินใจที่จะเลือกใช้บริการเกี่ยวกับการท่องเที่ยว อย่างไรก็ตามก็มีผู้ประกอบการจำนวนไม่น้อยใช้ประโยชน์จากความคิดเห็นเชิงบวกที่ผู้ประกอบการสร้างขึ้นช่วยทำการตลาดให้ตนเอง

ดังนั้นจะเห็นว่าเครือข่ายสังคมมีบทบาทมากกว่าการติดต่อสื่อสารระหว่างกลุ่มเพื่อนหรือบุคคลในครอบครัว แต่ยังถูกนำมาใช้ในระบบสารสนเทศด้านการท่องเที่ยว เพื่อส่งเสริมการปฏิบัติงานในด้านการท่องเที่ยวให้มีประสิทธิภาพที่ดียิ่งขึ้น ในบทความนี้ผู้วิจัยจะขอเน้นเฉพาะการนำข้อมูลของเครือข่ายสังคมมาวิเคราะห์เพื่อหาความสนใจส่วนตัวของผู้ใช้ โดยมีนักวิจัยหลายท่านได้นำเสนอแนวคิดในการพัฒนาเทคนิคในการสร้างแบบจำลองดังกล่าว ซึ่งจะอธิบายในหัวข้อถัดไป

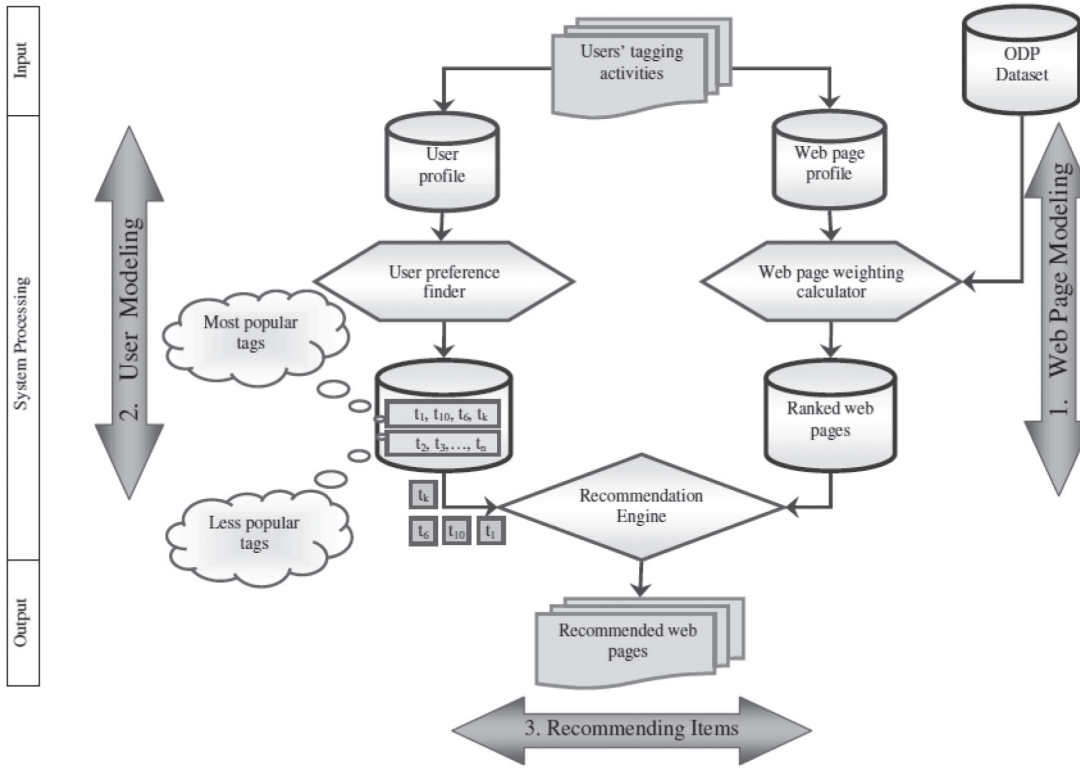
5. เทคนิคการพัฒนาแบบจำลองผู้ใช้โดยใช้เครือข่ายสังคม

เครือข่ายสังคมนอกจากจะนำไปใช้ในระบบสารสนเทศด้านการท่องเที่ยวที่กล่าวมาแล้วในหัวข้อที่ 4 ยังได้ถูกนำมาพัฒนาแบบจำลองผู้ใช้และระบบแนะนำข้อมูลต่างๆ หลายเทคนิควิธี โดยประโยชน์หลักของเครือข่ายสังคมที่มีต่อการพัฒนาแบบจำลองผู้ใช้คือสามารถนำไปสู่ข้อสรุปของความสนใจของผู้ใช้ได้ และยังนำแบบจำลองความสนใจของผู้ใช้นี้ไปประยุกต์ใช้ในระบบแนะนำข้อมูลการท่องเที่ยว ซึ่งในหัวข้อนี้จะกล่าวถึงเทคนิคต่างๆ ที่นักวิจัยนำมาใช้เพื่อสร้างแบบจำลองผู้ใช้โดยใช้เครือข่ายสังคม

5.1 การพัฒนาแบบจำลองผู้ใช้โดยใช้แท็กจากโซเชียลบุ๊กมาร์ก

โซเชียลบุ๊กมาร์ก (Social Bookmarking) เป็นบริการที่ให้ผู้ใช้สามารถบุ๊กมาร์ก (Bookmarks) เว็บเพจต่างๆ และแบ่งปันไปยังเพื่อนหรือผู้ใช้อื่นๆ ได้ โดยผู้ใช้แต่ละคนจะทำการบุ๊กมาร์กเว็บเพจและสร้างแท็ก (Tag) ไปยังบุ๊กมาร์กที่แตกต่างกัน ทำให้แท็กเหล่านี้ถูกตีความว่าเป็นสิ่งที่ผู้ใช้สนใจ นอกจากนี้ในหนึ่งเว็บเพจอาจจะถูกแต่ละผู้ใช้บุ๊กมาร์กด้วยแท็กที่ต่างกัน ทำให้แท็กเหล่านั้นถูกมองว่ามีความสัมพันธ์กันในเชิงความหมาย ตัวอย่างของการนำแท็กมาใช้ในการสร้างแบบจำลองผู้ใช้ เช่น Michlmayr และคณะ [26] นำเสนอวิธีการ Add-a-Tag ที่ทำการสร้างแบบจำลองผู้ใช้โดยแท็กจากเว็บไซต์โซเชียลบุ๊กมาร์ก Delicious.com ในลักษณะเดียวกัน Firan และคณะ [9] ได้นำเสนอการสร้างแบบจำลองผู้ใช้แบบ Tags-Based โดยใช้ แท็กจากเว็บไซต์ Last.fm และนำแบบจำลองนี้ไปใช้ในการสร้างคำแนะนำในระบบค้นหาเพลง แต่ข้อจำกัดของวิธีนี้คือการใช้ตัวแทนของความสนใจจากผู้ใช้แบบโดยรวมซึ่งคำนวณจากแท็กของผู้ใช้ทั้งหมด (Global Tag) ทำให้อาจจะยังไม่ตอบโจทย์ในการหาความสนใจเป็นรายบุคคลเพื่อใช้ในระบบแนะนำข้อมูลเฉพาะบุคคลมากนัก ซึ่งหากปรับปรุงโดยใช้แท็กของผู้ใช้เฉพาะบุคคลร่วมด้วยจะได้ข้อมูลความสนใจที่ถูกต้องมากยิ่งขึ้น ดังเช่น Durao และคณะ [27] ได้คำนวณหาค่าความสัมพันธ์ระหว่างแท็กกับผู้ใช้โดยใช้ค่าความถี่ที่ผู้ใช้แต่ละคนทำการสร้างแท็กนั้นๆ นั่นคือแท็กที่ผู้ใช้สร้างบ่อยจะมีค่าความสัมพันธ์ที่สูง ซึ่งสามารถนำมาลำดับความสนใจของผู้ใช้ได้ ในลักษณะเดียวกัน Davoodi และคณะ [28] นำเสนอระบบแนะนำข้อมูลเว็บเพจแบบเฉพาะบุคคลโดยสกัดความสนใจของผู้ใช้จากแท็กในโซเชียลบุ๊กมาร์ก ซึ่งแสดงภาพรวมของระบบดังรูปที่ 3 โดยระบบแบ่งการทำงานออกเป็น 3 ส่วน คือ 1) Web Page Modeling 2) User Modeling 3) Recommending Item ในส่วนแรกคือ Web Page Modeling เป็นการจัดหมวดหมู่และจัดอันดับเว็บเพจ โดยการจัดหมวดหมู่จะใช้หมวดหมู่ของแท็กที่ผู้ใช้สร้างไปยังเว็บเพจนั้นๆ ประกอบกับการใช้ ODP (Open Directory Project) ซึ่งเป็นเว็บไดเรกทอรีในการจัดหมวดหมู่เว็บเพจแบบลำดับชั้น การจัดอันดับจะใช้ค่าความถี่ของแท็กที่ทุกผู้ใช้สร้างไปยังหน้าเพจนั้น แท็กใดที่มีค่าความถี่สูงก็จะถูกจัดอยู่ในลำดับต้นๆ ของหมวดหมู่นั้น ต่อมาส่วนที่สอง User Modeling เป็นการหาลำดับความสนใจของผู้ใช้โดยใช้

ความถี่ของแท็กที่ผู้ใช้นั้นๆ สร้างขึ้น แท็กที่ถูกผู้ใช้สร้างมากที่สุดจะเป็นหมวดหมู่ความสนใจที่ผู้ใช้สนใจมากที่สุด ส่วนสุดท้าย Recommending Item เป็นการแนะนำเว็บเพจไปยังผู้ใช้ โดยจะจับคู่แท็กที่ผู้ใช้แต่ละคนสนใจไปยังเว็บเพจที่ได้จัดหมวดหมู่และจัดลำดับแล้วในส่วนที่สอง อย่างไรก็ตามการใช้แท็กของผู้ใช้เฉพาะบุคคลของผู้ใช้ที่เริ่มต้นใช้โซเชียลบุ๊กมาร์ก ข้อมูลแท็กอาจจะยังไม่เพียงพอในการที่จะหาข้อมูลความสนใจของผู้ใช้ที่แท้จริงได้หรืออาจทำให้เกิดปัญหาโคลด์สตาร์ทนั่นเอง

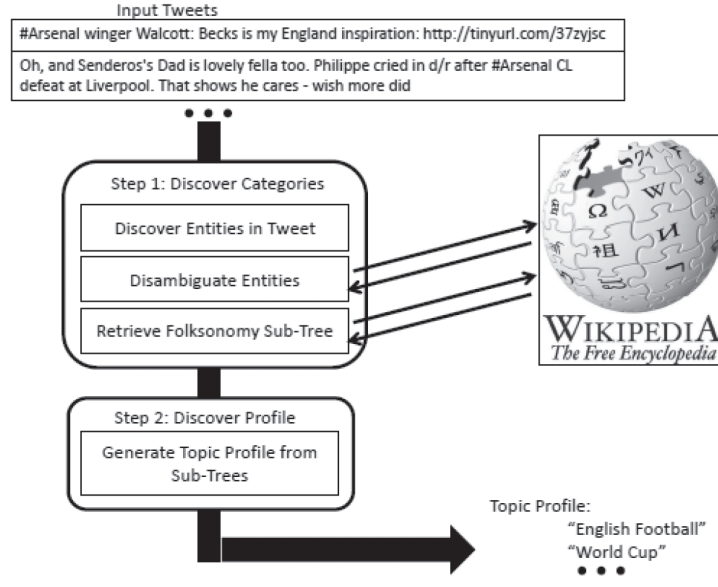


รูปที่ 3 ระบบแนะนำข้อมูลเว็บเพจโดยใช้การสกัดความสนใจจากแท็กในโซเชียลบุ๊กมาร์ก [28]

5.2 การพัฒนาแบบจำลองผู้ใช้โดยข้อความบนทวิตเตอร์หรือเฟสบุค

ทวิตเตอร์เป็นเครือข่ายสังคมที่ให้ผู้ใช้สามารถรายงานข่าว แสดงความคิดเห็น ฯลฯ โดยการโพสต์ข้อความสั้นไม่เกิน 140 ตัวอักษร หรือที่เรียกกันว่า “ทวีต” จะเห็นว่าในทวิตเตอร์ผู้ใช้ไม่ให้คำสำคัญ (Keyword) แต่เป็นเพียงข้อความสั้นที่ผู้ใช้ทวีตลงไปเช่นเดียวกับเฟสบุคที่เป็นเครือข่ายสังคมที่ให้ผู้ใช้สามารถโพสข้อความ รูปภาพ หรือวีดิโอลงบนพื้นที่ของตนเองและเพื่อนๆ หรือที่เรียกว่า “Timeline” และสามารถแสดงความคิดเห็นต่างๆ ได้ ทำให้เกิดปัญหาเช่นเดียวกับทวิตเตอร์นั่นคือมีข้อความมากมายเกิดขึ้นแต่ไม่ได้มีการระบุคำสำคัญ ดังนั้นจึงจำเป็นจะต้องทำการกรองหรือจับคำสำคัญก่อนที่จะใช้เป็นข้อมูลสำหรับสร้างแบบจำลองของผู้ใช้ ดังเช่น Castillo และคณะ [29] พยายามที่จะจำแนกทวีต ออกเป็นสองแบบคือ 1) ทวีตเกี่ยวกับข่าว และ 2) การพูดคุย เพื่อที่จะประเมินความน่าเชื่อถือของข้อมูลบนทวิตเตอร์ Abel

และคณะ [10] ได้ทำการสกัดหัวข้อ (Topic) บนข้อความทวีต เช่นเป็นทวีตของบุคคล เหตุการณ์ หรือผลิตภัณฑ์ เป็นต้น เพื่อใช้สำหรับการแนะนำบทความข่าวให้กับแต่ละบุคคล ดังนั้นการใช้ข้อมูลจากข้อความของทั้งทวิตเตอร์และเฟสบุคเพื่อหาข้อความที่แสดงถึงความสนใจของผู้ใช้จะต้องผ่านวิธีการที่สำคัญคือการกรองและสกัดข้อมูลมากมายเหล่านั้นออกมาเป็นคำสำคัญที่สามารถเชื่อมโยงกับความสนใจได้ ดังเช่นการนำคำสำคัญที่สกัดมาได้จับคู่ไปยังฐานความรู้ของเวิร์ดเน็ต (Wordnet) หรือวิกิพีเดีย (Wikipedia) เพื่อตัดคำที่ไม่ได้เกี่ยวข้องกับความสนใจออกไป [30, 31] ซึ่งในรูปที่ 4 เป็นไดอะแกรมที่แสดงตัวอย่างวิธีการค้นหาหัวข้อความสนใจของผู้ใช้จากข้อความบนทวิตเตอร์ที่นำเสนอโดย Michelson และคณะ [32] โดยจะแบ่งเป็น 2 ขั้นตอน คือ 1) Discover Categories เป็นการค้นหาหมวดหมู่ความสนใจของผู้ใช้ โดยขั้นแรก Discover Entities in Tweet จะเลือกคำ (Entity) จากข้อความทวีตที่ขึ้นต้นด้วยตัวอักษรพิมพ์ใหญ่และไม่เป็นคำที่ไม่สำคัญ (Stop-words) ต่อมาขั้นที่สอง Disambiguate Entities นำคำที่ได้ในขั้นแรกไปหาความหมายที่แท้จริง โดยจับคู่คำไปยังฐานความรู้วิกิพีเดียเพื่อหาความหมายของคำที่เป็นไปได้ทั้งหมด (Candidate Entities) จากนั้นจะเลือกความหมายของคำที่แท้จริงโดยพิจารณาจากคำอื่นๆ ในทวีตว่าปรากฏมากน้อยเพียงใดในหน้าเพจวิกิพีเดียของแต่ละคำที่เป็นไปได้ทั้งหมด เมื่อเลือกความหมายของคำได้แล้วสุดท้าย Retrieve Folksonomy Sub-Tree วิกิพีเดียจะส่งหมวดหมู่ของคำนั้นๆ ในรูปแบบโครงสร้างแบบต้นไม้กลับไปยังระบบ 2) Discover Profile เป็นการเลือกหัวข้อความสนใจของผู้ใช้โดยใช้ค่าลำดับความสนใจ (Rank) ซึ่งคำนวณจากค่าความถี่ในการทวีตข้อความที่เป็นความสนใจในหมวดหมู่ที่ได้เลือกไว้ในขั้นตอนที่ 1 คูณกับค่าน้ำหนักของคำในโครงสร้างแบบต้นไม้ โดยคำที่ปรากฏในลำดับชั้นสูงจะมีค่าน้ำหนักมากกว่าชั้นที่ต่ำกว่า อย่างไรก็ตาม วิธีการนี้ในขั้นตอนการจับคู่คำไปยังฐานความรู้วิกิพิเดียเพื่อหาความหมายของคำยังอาจทำให้เกิดข้อผิดพลาดได้อยู่ เช่น ระบบแปลความหมาย “Apple” ว่าเป็นผลไม้ชนิดหนึ่ง แต่ในความเป็นจริงผู้ใช้ต้องการสื่อถึง “Apple” ที่เป็นแบรนด์สินค้า เป็นต้น ซึ่งสามารถก่อให้เกิดการแนะนำข้อมูลที่ไม่ตรงกับความสนใจของผู้ใช้ที่แท้จริงได้



รูปที่ 4 การสกัดข้อมูลความสนใจของผู้ใช้จากข้อความบนทวิตเตอร์โดยใช้ฐานความรู้ของ Wikipedia [32]

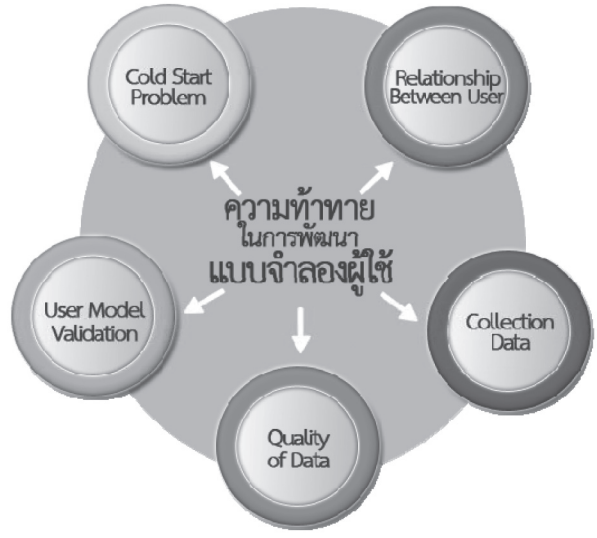
5.3 การพัฒนาแบบจำลองผู้ใช้โดยใช้การผสมผสานระหว่างเครือข่ายสังคมหลายแหล่ง

แนวคิดนี้เป็นการใช้ประโยชน์จากการผสมผสานเครือข่ายสังคมหลายแหล่งเพื่อลดข้อจำกัดที่เกิดขึ้นเมื่อใช้เครือข่ายสังคมเพียงแหล่งเดียว Saito และคณะ [33] ได้เสนอวิธีการสกัดความสนใจของผู้ใช้บนทวิตเตอร์ร่วมกับโซเชียลบุ๊กมาร์กโดยสร้างลำดับชั้นความสนใจจากแท็กของผู้ใช้ทั้งหมดในโซเชียลบุ๊กมาร์ก จากนั้นใช้ค่าความถี่ของคำที่ปรากฏในทวีตของผู้ใช้ร่วมกับค่าระยะทาง (Distance) บนลำดับชั้นของแท็ก เพื่อคำนวณหาคำที่ผู้ใช้นั้นกำลังสนใจเพื่อนำไปใช้กับระบบแนะนำข้อมูล วิธีการนี้ใช้ประโยชน์จากเครือข่ายสังคมหลายแหล่งเพื่อลดข้อเสียของกันและกัน การสร้างลำดับชั้นของแท็กช่วยให้ทราบถึงความสนใจของบุคคลโดยทั่วไปว่าขณะนี้กำลังสนใจสิ่งใดเรียงเป็นลำดับชั้นของความสนใจ ซึ่งจะเป็นประโยชน์ต่อผู้ใช้ที่ยังไม่มีแท็กมากนัก และการใช้คำที่ปรากฏในทวีตบ่อยๆ (รวมถึงคำใน รีทวีต และทวีตของผู้ที่เราติดตาม) ซึ่งแสดงถึงความสนใจของผู้ใช้ในขณะนั้น ในการเลือกคำในลำดับชั้นของแท็กจะช่วยให้ได้ข้อมูลความสนใจของผู้ใช้ที่ตรงกับความต้องการที่แท้จริงของผู้ใช้แต่ละบุคคลมากขึ้น

ในการสร้างและพัฒนาแบบจำลองผู้ใช้ที่ได้ทบทวนวรรณกรรมไปแล้วในหัวข้อที่ 3 และ 5 พบว่ายังมีปัญหาสำคัญที่เป็นความท้าทายยิ่งอีกหลายประการ แต่ในบทความนี้จะขอกล่าวถึงปัญหาหลัก 5 ปัญหา ซึ่งจะอธิบายในหัวข้อถัดไป

6. ความท้าทายในการพัฒนาแบบจำลองผู้ใช้

การสร้างและการพัฒนาแบบจำลองผู้ใช้ในแต่ละวิธีเกิดปัญหาและข้อจำกัดที่แตกต่างกันไป บางปัญหาได้กล่าวถึงไปแล้วในหัวข้อก่อนหน้า แต่ก็จะนำมาอธิบายเพิ่มเติมในหัวข้อนี้ นอกจากนี้ยังมีความท้าทายอื่นๆ ที่ผู้พัฒนาควรจะคำนึงถึง โดยได้รวบรวมเป็น 5 ปัญหา ที่เป็นความท้าทายอย่างยิ่ง ได้แก่ 1) ปัญหาโคลด์สตาร์ท 2) การหาความสัมพันธ์ระหว่างผู้ใช้ 3) การตรวจสอบแบบจำลองผู้ใช้ 4) การจัดเก็บข้อมูลแบบจำลองผู้ใช้ และ 5) คุณภาพของข้อมูล ซึ่งสามารถแสดงได้ดังรูปที่ 5



รูปที่ 5 ความท้าทายในการพัฒนาแบบจำลองผู้ใช้

6.1 ปัญหาโคลด์สตาร์ท (Cold Start Problem)

ปัญหาโคลด์สตาร์ทเป็นปัญหาที่เกิดขึ้นเมื่อผู้ใช้เริ่มต้นใช้งานระบบ แต่ระบบยังไม่มีข้อมูลของผู้ใช้มากพอที่จะสามารถสกัดความสนใจและให้การแนะนำสินค้าหรือบริการได้อย่างถูกต้อง โดยเฉพาะอย่างยิ่งในระบบแนะนำข้อมูลเฉพาะบุคคล ซึ่ง Rashid และคณะ [34] ได้พยายามแก้ไขปัญหานี้โดยใช้ข้อมูลจากผู้ใช้อื่นที่มีต่อระบบแนะนำภาพยนตร์ เช่น ข้อมูลภาพยนตร์ที่กำลังเป็นที่นิยมซึ่งคำนวณจากความถี่ในการเข้าถึงของผู้ใช้ ข้อมูลจำนวนของการแสดงความคิดเห็น ข้อมูลการให้คะแนนภาพยนตร์ของผู้ใช้ เป็นต้น ภาพยนตร์ที่มีลำดับคะแนนที่สูงก็จะถูกแนะนำให้กับผู้ใช้ที่เริ่มต้นใช้งานระบบ จะเห็นว่าปัญหานี้ถูกแก้ไขในระดับหนึ่งโดยใช้ความสนใจของผู้ใช้อื่นโดยรวม ต่างจาก Hang และคณะ [35] ที่ใช้ข้อมูลจากผู้ใช้อื่นมาพัฒนาอัลกอริทึมที่ใช้ในการวัดค่าความคล้ายระหว่างผู้ใช้และช่วยในการจับคู่ระหว่างผู้ใช้กับสิ่งที่สนใจ โดยมีหลักการว่าผู้เริ่มต้นใช้ระบบจะจับคู่กับสิ่งที่สนใจที่มีค่าสำคัญมากที่สุด (จากผู้ใช้ที่มีความคล้ายคลึงกัน) ซึ่งวิธีการนี้ให้ผลที่น่าพอใจกับระบบที่เน้นไปที่การหาความสนใจรายบุคคลมากกว่าวิธีการแรก นอกจากนี้ยังมีวิธีการที่ใช้ข้อมูลของผู้ใช้เอง เช่น Park และคณะ [36] ที่ใช้ข้อมูลประวัติของผู้ใช้รวมถึงรายการข้อมูลต่างๆ อย่างไรก็ตาม วิธีการนี้ก็ยังไม่สามารถลดการรบกวนผู้ใช้ในการขอข้อมูลส่วนบุคคลต่างๆ ได้ วิธีการที่น่าสนใจที่จะนำมาแก้ไขปัญหานี้อีกวิธีหนึ่งคือ การนำข้อมูลมาจากเครือข่ายสังคมของผู้ใช้ เช่น ข้อมูลประวัติ การแสดงความคิดเห็น การคลิกชื่นชอบ การเช็คอินในสถานที่ต่างๆ และข้อมูลจากเพื่อนที่มีความสนใจคล้ายกัน เป็นต้น เนื่องจากเครือข่ายสังคมกำลังเป็นที่นิยมอย่างมากในปัจจุบัน ทำให้ผู้ใช้มีข้อมูลที่สามารถเป็นตัวแทนของความสนใจและมีการอัพเดตมากมายบนเครือข่ายเหล่านั้น นอกจากนี้การใช้ข้อมูลจากเครือข่ายสังคมยังไม่เป็นการรบกวนผู้ใช้มากเกินไปในการขอข้อมูลต่างๆ เพราะเพียงผู้ใช้ทำการอนุญาตให้ระบบเข้าถึงข้อมูล ระบบจะทำการดึงข้อมูลในเครือข่ายเหล่านี้โดยไม่ต้องให้ผู้ใช้กรอกข้อมูลไปยังระบบเอง อย่างไรก็ตาม ก็ยังมีปัญหาในการใช้ข้อมูลของเพื่อนในเครือข่าย เพราะจำเป็นจะต้องหาเพื่อนที่มีความสนใจคล้ายกันหรือเพื่อนที่มีความสัมพันธ์สำคัญกับผู้ใช้ ซึ่งจะกล่าวในหัวข้อถัดไป

6.2 การหาความสัมพันธ์ระหว่างผู้ใช้ (Relationship Between User)

การหาความสัมพันธ์ระหว่างผู้ใช้เป็นปัญหาที่เกิดเนื่องจากการนำข้อมูลจากเพื่อนในเครือข่ายสังคมมาใช้ในการพัฒนาระบบ โดยจะมีวิธีการใดที่จะระบุได้ว่าเพื่อนคนใดที่มีความสนใจคล้ายคลึงกัน หรือกำลังพูดหรือสนใจในสิ่งเดียวกันกับผู้ใช้ ซึ่งในเฟสบุค การใช้ข้อมูลการใช้งานที่มีร่วมกันต่างๆ เช่น การกดถูกใจหน้าเพจเดียวกัน การโพสสถานะที่ระบุความชอบในสิ่งเดียวกัน เป็นต้น เป็นวิธีที่ถูกนำมาใช้หาความสัมพันธ์ระหว่างผู้ใช้โดย Ho และคณะ [37] อย่างไรก็ตามยังมีการใช้งานอื่นๆ ที่ไม่ได้ถูกกล่าวถึง เช่น การกดไลค์ในข้อความสถานะเดียวกัน การถูกติดป้ายบนข้อความสถานะ การเช็คอินร่วมกันสถานที่เดียวกัน เป็นต้น อาจเป็นอีกวิธีที่สามารถนำมาคำนวณหาค่าความสัมพันธ์ระหว่างผู้ใช้ได้ นอกจากนี้ในการเลือกผู้ใช้อื่นที่มีลักษณะความสนใจคล้ายคลึงกันมาใช้ในการค้นหาความสนใจของผู้ใช้อาจจะใช้ผู้ใช้อื่นที่มีค่าความคล้ายคลึงมากที่สุด หรืออาจใช้ผู้ใช้อื่นมากกว่าหนึ่งคนเพื่อหาลำดับความสนใจที่เหมาะสม อย่างไรก็ตาม เครือข่ายสังคมอย่างเฟสบุคนั้นอาจสามารถใช้ประโยชน์จากหน้าฟีดข่าว (News Feed) เป็นเครื่องมือในการช่วยหาความสัมพันธ์ระหว่างผู้ใช้ได้ เนื่องจากหน้าฟีดข่าวได้ใช้อัลกอริทึมที่มีค่าความสัมพันธ์ระหว่างผู้ใช้ (Affinity) เป็นตัวแปรในการแสดงความสัมพันธ์ระหว่างผู้ใช้ จะเห็นได้จากเพื่อนที่ผู้ใช้มีปฏิสัมพันธ์มากบนเฟสบุค สถานะของเพื่อนนั้นจะแสดงที่หน้าฟีดข่าวบ่อยครั้ง

6.3 การตรวจสอบแบบจำลองผู้ใช้ (User Model Validation)

เป็นการตรวจสอบแบบจำลองผู้ใช้ว่าได้สรุปข้อมูลความสนใจมีความถูกต้องมากเพียงใด เพื่อที่จะได้ปรับปรุงการแนะนำให้มีประสิทธิภาพดียิ่งขึ้น ซึ่งในระบบแนะนำข้อมูลจะใช้ข้อมูลการตอบกลับจากผู้ใช้ช่วยในการตรวจสอบ โดยแบ่งการตอบกลับเป็น 2 ลักษณะ คือ 1) การตอบกลับแบบปริยาย (Implicit Feedback) เป็นการตรวจสอบจากการเลือกหรือไม่เลือกสินค้าและบริการที่ระบบแนะนำให้แก่ผู้ใช้ ซึ่งผู้ใช้จะไม่ทราบว่ามีการเก็บข้อมูล อย่างไรก็ตาม สามารถวิธีการตอบกลับลักษณะนี้ ผู้ใช้ไม่ได้มุ่งหวังในการปรับปรุงประสิทธิภาพของระบบ แต่เพื่อความพึงพอใจของผู้ใช้เป็นหลัก 2) การตอบกลับแบบชัดเจน (Explicit Feedback) อีกวิธีการหนึ่งที่ได้รับความสนใจ โดยวิธีการนี้จะได้ความเห็นจากผู้ใช้ว่าข้อมูลที่ระบบแนะนำมีความถูกต้องหรือตรงกับความสนใจของผู้ใช้เพียงใด [38] เช่น การกดปฏิเสธสินค้า การให้คะแนนสินค้า ข้อความแนะนำจากผู้ใช้ เป็นต้น แม้จะเป็นวิธีการที่ต้องมีการรบกวนผู้ใช้ แต่หากออกแบบส่วนติดต่อที่สามารถเสริมสร้างแรงจูงใจให้ผู้ใช้รู้สึกว่าไม่เป็นการรบกวน ข้อมูลตอบกลับนี้จะช่วยให้ปรับปรุงแบบจำลองผู้ใช้ได้มีประสิทธิภาพมากกว่าและยังสามารถปรับปรุงได้ทันท่วงทีอีกด้วย

6.4 การจัดเก็บข้อมูลแบบจำลองผู้ใช้ (Collection Data)

รูปแบบของการจัดเก็บข้อมูลความสนใจในแบบจำลองผู้ใช้ที่ผ่านมาจะใช้การจัดเก็บเป็นข้อความลงในฐานข้อมูล แต่ในปัจจุบันเริ่มมีการนำออนโทโลยี (Ontology) มาใช้เป็นรูปแบบการจัดเก็บข้อมูล เช่น การจัดเก็บข้อมูลการเดินทางและการท่องเที่ยวเพื่อให้เกิดความเชื่อมโยงระหว่างข้อมูลต่างๆ [39] โดยการใช้ออนโทโลยีนั้นจะช่วยในเรื่องความหมายของคำ ลดความกำกวมในความหมายของคำได้ [38] นอกจากนี้ยังสามารถพัฒนารูปแบบข้อมูลนี้เพื่อความสะดวกในการนำไปใช้ในด้านอื่นๆ ทั้งนี้มีความเป็นไปได้ที่จะสามารถทำให้อยู่ในรูปแบบของเว็บเซอร์วิส หรือสร้าง API (Application Programming Interface) เพื่อประโยชน์แก่การนำข้อมูลความสนใจของผู้ใช้เหล่านี้ไปประยุกต์ใช้ต่อ

6.5 คุณภาพของข้อมูล (Quality of data)

การนำข้อมูลจากเครือข่ายสังคมมาใช้ ผู้พัฒนาระบบจะต้องคำนึงสิ่งต่างๆ 2 ประการ ประการแรกคือคุณภาพ (Quality) ของข้อมูลเป็นสิ่งสำคัญประการหนึ่งที่ผู้พัฒนาระบบที่จะนำข้อมูลจากเครือข่ายสังคมมาใช้ควรพึงระวังเป็นพิเศษในหลายๆ ด้าน เช่น ความถูกต้องของข้อมูลและความน่าเชื่อถือของข้อมูล เป็นต้น ประการที่สองคือ ปริมาณ (Quantity) ของข้อมูล เครือข่ายสังคมส่วนใหญ่จะมีข้อมูลจำนวนมากที่เป็นประโยชน์ต่อการนำมาประมวลผลเพื่อแนะนำข้อมูลกับผู้ใช้คนอื่นๆ อยู่แล้ว เนื่องจากมีผู้ใช้ทั่วโลก ดังนั้นเนื้อหาในส่วนนี้จะขอกล่าวถึงเฉพาะประเด็นของคุณภาพของข้อมูลเท่านั้น คุณภาพของข้อมูลสามารถพิจารณาได้จาก 4 มิติ เช่น โครงสร้างของข้อมูล เนื้อหาของข้อมูล ความถูกต้องของข้อมูล และในมิติที่เกี่ยวข้องกับเวลา

มิติแรกคือโครงสร้างของข้อมูล หมายถึงข้อมูลที่ได้จากเครือข่ายสังคมควรถูกจัดเก็บอย่างมีโครงสร้าง เพื่อที่จะสามารถนำไปประมวลผลต่อได้ง่าย ตัวอย่างเช่น XML JSON ซึ่งเป็นรูปแบบของข้อมูลที่มีคุณภาพมากกว่าข้อมูลที่ไม่มีโครงสร้าง เช่น ข้อความธรรมดา (Pain Text) เพราะข้อมูลที่มีโครงสร้างเหล่านั้นสามารถอธิบายถึงข้อมูลที่จัดเก็บอยู่ได้ อย่างไรก็ตาม ข้อมูลที่มีโครงสร้างบางประเภทยังมี

ข้อจำกัดบางประการ เช่น ใน JSON แม้จะมีรูปแบบการเขียนแท็กที่เรียบง่ายและใช้อักขระน้อย แต่พบว่าในข้อมูลดิบที่มีจำนวนแถว (Record) มาก JSON จะสร้างแท็กเหล่านั้นซ้ำ ซึ่งทำให้ข้อมูลที่อยู่ในรูปแบบของ JSON มีขนาดใหญ่เกินความจำเป็น นอกจากนี้ลักษณะโครงสร้างที่ต่างกันยังเป็นผลต่อคุณภาพของข้อมูล เนื่องจากโครงสร้างบางลักษณะไม่สามารถอธิบายข้อมูลได้อย่างชัดเจน ซึ่งอาจทำให้ระบบใช้ข้อมูลเหล่านั้นได้ไม่ถูกต้อง ตัวอย่างเช่น ข้อมูลเช็คอินจากเครือข่ายสังคม จะมีคุณภาพมากขึ้นเมื่อใช้ออนโทโลยีปรับโครงสร้าง เพื่อช่วยลดความกำกวมของคำ

มิติที่สอง เนื้อหาของข้อมูลสามารถบ่งบอกถึงคุณภาพของข้อมูลได้เช่นกัน เพราะสามารถแยกระดับความสำคัญของข้อมูลเหล่านั้นได้โดยวิธีต่างๆ เช่น 1) การมีเนื้อหาของข้อมูลตรงกับที่ระบบต้องการ 2) การมีเนื้อหาที่มีความสัมพันธ์กับผู้ใช้ และ 3) เนื้อหานั้นมีผู้ใช้งานสูง [40] ซึ่งในเครือข่ายสังคมจะพิจารณาจากเนื้อหาของความคิดเห็น การคลิก การชื่นชอบ การโหวต เป็นต้น จะเห็นว่าทั้ง 3 วิธี ยังต้องอาศัยปริมาณของข้อมูลในการค้นหาคุณภาพของข้อมูล ข้อมูลที่มีปริมาณเหมาะสมจะช่วยสนับสนุนให้สามารถเลือกข้อมูลที่มีคุณภาพได้ถูกต้องยิ่งขึ้น

มิติที่สามคือความถูกต้องของข้อมูล ซึ่งเป็นมิติที่สำคัญในการพิจารณาคุณภาพของข้อมูล [41] โดยมีหลายคำนิยามสำหรับความถูกต้องของข้อมูล เช่น Wang และ Strong [42] ได้นิยามความถูกต้องไว้ว่า “ขอบเขตที่ข้อมูลมีความถูกต้อง มีความน่าเชื่อถือ และได้รับการรับรอง” Ballou และ Pazer [43] ระบุว่าข้อมูลจะมีความถูกต้องเมื่อค่าของข้องมูลที่เก็บในฐานข้อมูลมีความสอดคล้องกับค่าจริง เป็นต้น ข้อมูลที่ได้มาจากเครือข่ายสังคม บางข้อมูลยังพบความไม่ถูกต้องของข้อมูล เช่น ข้อมูลสถานที่ท่องเที่ยวที่มีชื่อเรียกหลายชื่อ (วัดใหญ่หรือวัดพระศรีรัตนมหาธาตุวรมหาวิหาร) การสะกดหรือเขียนผิดไปจากชื่อจริง (พิษณุโลก พิดโลก) เป็นต้น ซึ่งอาจส่งผลให้การประมวลผลของระบบไม่ถูกต้องตามไปด้วย การเพิ่มคุณภาพของข้อมูลในมิติของความถูกต้องจึงจำเป็นอย่างยิ่ง โดยการผ่านกระบวนการการทำข้อมูลให้ถูกต้อง (Data Cleansing) ก่อนที่จะนำข้อมูลเหล่านั้นไปใช้ประโยชน์ต่อไป

มิติที่สี่ มิติที่เกี่ยวข้องกับเวลา ในข้อมูลที่มีการบันทึกเวลาลงไปด้วย มีความสัมพันธ์กับคุณภาพของข้อมูลในหลายมุม ทั้งในมุมของความเป็นปัจจุบัน (Up to date) ของข้อมูลและการเลือกข้อมูลไปใช้ในแต่ละช่วงเวลา เช่น ในระบบที่ต้องการหาความสนใจของผู้ใช้ บางระบบต้องการข้อมูลจากเครือข่ายที่ทันต่อเหตุการณ์ ตัวอย่างเช่น ข้อมูลการเช็คอินใน 1 วัน หรือ 1 สัปดาห์ที่ผ่านมา เพื่อวิเคราะห์ความสนใจของผู้ใช้ในปัจจุบัน แต่ในบางระบบอาจต้องการวิเคราะห์ข้อมูลในช่วงเวลาที่แตกต่างกัน เช่น ไตรมาส ฤดู หรือ ปี เพื่อทำเหมืองข้อมูลหาความสนใจในแต่ละช่วงเวลา เป็นต้น ดังนั้นหากข้อมูลที่ดึงมาจากเครือข่ายสังคมเหล่านั้นไม่มีเวลากำกับมาด้วย ระบบจะทำการวิเคราะห์ในเชิงลึกเพื่อหาความสนใจในมิติที่น่าสนใจอื่นๆ ได้ยาก

จากมิติคุณภาพของข้อมูลทั้งสี่มิติข้างต้น ทำให้ทราบถึงความสำคัญของคุณภาพของข้อมูล ซึ่งเป็นสิ่งที่สำคัญอย่างยิ่งที่ควรนำมาพิจารณาในการเลือกใช้แหล่งข้อมูล อย่างไรก็ตาม นักพัฒนาระบบควรที่จะคำนึงถึงปริมาณข้อมูลของแหล่งข้อมูลต่างๆ เนื่องจากหากมีปริมาณของข้อมูลที่ไม่เพียงพอ อาจจะมีผลต่อความถูกต้องของการวิเคราะห์ข้อมูล ดังนั้นคุณภาพของข้อมูลควรมาพร้อมกับปริมาณของข้อมูลที่เหมาะสมหรือเพียงพอที่จะนำมาใช้งานด้วยเช่นกัน

7. สรุป

บทความนี้ชี้ให้เห็นถึงศักยภาพของข้อมูลในเครือข่ายสังคมที่สามารถจะนำมาใช้ในระบบสารสนเทศด้านการท่องเที่ยว แต่ในบทความนี้เน้นการนำข้อมูลจากเครือข่ายสังคมมาใช้ในระบบแนะนำข้อมูลการท่องเที่ยว โดยมีนักวิจัยหลายๆ ท่านนำเสนอระบบแนะนำข้อมูลโดยใช้เทคนิคต่างๆ ดังที่กล่าวมาแล้วในหัวข้อที่ 5 บทความนี้ยังชี้ให้เห็นถึงความท้าทายหรือสิ่งที่นักพัฒนาระบบสารสนเทศควรคำนึงถึงในการพัฒนาระบบที่อยู่บนพื้นฐานของการใช้ข้อมูลจากเครือข่ายสังคม 5 ประการ ได้แก่ 1) ความท้าทายของปัญหาโคลด์สตาร์ทที่อาจเกิดขึ้นกับผู้ใช้ที่เพิ่งเริ่มต้นใช้งานระบบ 2) การหาความสัมพันธ์ระหว่างผู้ใช้ 3) วิธีการตรวจสอบแบบจำลองผู้ใช้เพื่อเพิ่มประสิทธิภาพของแบบจำลอง 4) วิธีการจัดเก็บข้อมูลแบบจำลองผู้ใช้เพื่อความสะดวกในการนำไปใช้ และ 5) คุณภาพของข้อมูลที่นำมาสกัดความสนใจในระบบ ประเด็นปัญหาเหล่านี้ยังรอการวิจัยและพัฒนาเทคนิคต่างๆ จากนักวิจัย เพื่อที่จะสามารถนำข้อมูลจากเครือข่ายสังคมที่มีอยู่อย่างมหาศาลมาใช้เพื่อก่อให้เกิดประโยชน์อย่างแท้จริง

เอกสารอ้างอิง

1. Harun, A. 2012. Thailand Tourism Industry: The Impact of Tourism Sector to Thai's Gross Domestic Product (GDP). Proceedings of the 2nd International Conference on Business, Economics, Management and Behavioral Sciences. 13-14 October 2012. Bali. Indonesia. Bali. PSRC. p. 90-95.
2. World Economic Forum. 2012. Fostering Prosperity and Regional Integration Through Travel and Tourism. *The ASEAN Travel & Tourism Competitiveness Report* 2012: 1-2.
3. Shahabi, C., and Chen, Y. 2003. An Adaptive Recommendation System without Explicit Acquisition of User Relevance Feedback. *Distributed and Parallel Databases* 14(2): 173-192.
4. Kabassi, K. 2010. Personalizing Recommendations for Tourists. *Telematics and Informatics* 27(1): 51-66.
5. Rabanser, U., and Ricci, F. 2005. Recommender Systems : Do They Have a Viable Business Model in e-Tourism?. *Information and Communication Technologies in Tourism* 2005. 160-171.
6. Coyle, L., and Cunningham, P. 2003. Exploiting Re-ranking Information in a Case-Based Personal Travel Assistant. In: Aha, D., Editors. Proceedings of the 5th International Conference on Case-Based Reasoning. 24 June 2003. Trondheim. Norway. London. Springer-Verlag. p. 11-20.
7. Schiaffino, S., and Amandi, A. 2009. Building an Expert Travel Agent as a Software Agent. *Expert Systems with Applications: An International Journal* 36(2): 1291-1299.

8. Zhu, Y., He, L., and Wang, X. 2012. User Interest Modeling and Self-Adaptive Update Using Relevance Feedback Technology. *Procedia Engineering* 29(5): 721-725.
9. Firan, C. S., Nejdl, W., and Paiu, R. 2007. The Benefit of Using Tag-based Profiles. Proceedings of the 2007 Latin American Web Conference. 31 October-2 November 2007. Santiago. U.S.A. Washington. IEEE Computer Society. p. 32-41.
10. Abel, F., Gao, Q., Houben, G.-J., and Tao, K. 2011. Analyzing User Modeling on Twitter for Personalized News Recommendations. Proceedings of the 19th International Conference on User Modeling, Adaption, and Personalization. 11-15 July 2011. Girona. Spain. Heidelberg. Springer-Verlag Berlin. p. 1-12.
11. Chin, D. N., and Porage, A. 2001. Acquiring User Preferences for Product Customization. Proceedings of the 8th International Conference on User Modeling. 13-17 July 2001. Sonthofen. Germany. London. Springer-Verlag. p. 95-104.
12. Cheverst, K., Mitchell, K., and Davies, N. 2002. The Role of Adaptive Hypermedia in a Context-Aware Tourist GUIDE. *Communications of the ACM* 45(5): 47-51.
13. Burke, R. 2000. Knowledge-Based Recommender Systems. *Encyclopedia of Library and Information Systems* 69(32): 180-200.
14. García-Crespo, A., Chamizo, J., Rivera, I., Mencke, M., Colomo-Palacios, R., and Gómez-Berbís, J. M. 2009. SPETA: Social Pervasive e-Tourism Advisor. *Telematics and Informatics* 26(3): 306-315.
15. Widyantoro, D. H., Ioerger, T. R., and Yen, J. 1999. An Adaptive Algorithm for Learning Changes in User Interests. Proceedings of the 8th International Conference on Information and Knowledge Management. 2-6 November 1999. Missouri. U.S.A. p. 405-412.
16. Mezhoudi, N. 2013. User Interface Adaptation Based on User Feedback and Machine Learning. Proceedings of the Companion Publication of the 2013 International Conference on Intelligent User Interfaces Companion. 19-22 March 2013. California. U.S.A. p. 25-28.
17. Arhippainen, L., Rantakokko, T., and Tähti, M. 2004. Mobile Feedback Application for Emotion and User Experience Collection. In: Lindén, G., Editors. Proceedings of Proactive Computing Workshop. 25-26 November 2004. Helsinki. Finland. Helsinki. Helsinki University Press. p. 77-81.
18. Tung, H.-W., and Soo, V.-W. 2004. A Personalized Restaurant Recommender Agent for Mobile e-Service. Proceedings of the 2004 IEEE International Conference on e-Technology, e-Commerce and e-Service. 28-31 March 2004. Taipei. Taiwan. Washington. IEEE Computer Society. p. 259-262.
19. Srivihok, A., and Sukonmanee. 2005. E-Commerce Intelligent Agent: Personalization Travel Support Agent Using Q Learning. the 7th International Conference on Electronic Commerce. 19-21 September 2005. Sanda. Japan. p. 287-292.

20. Yang, Y. and Marques, N. C. 2005. User Group Profile Modeling Based on User Transactional Data for Personalized Systems. Proceedings of the 12th Portuguese Conference on Progress in Artificial Intelligence. 5-8 December 2005. Covilhã. Portugal. Berlin. Springer-Verlag. p. 337-347.
21. Miguéns, J., Baggio, R., and Costa, C. 2008. Social Media and Tourism Destinations: TripAdvisor Case Study. Proceedings of the IASK Advances in Tourism Research 2008. 26-28 May 2008. Aveiro. Portugal. Aveiro. University of Aveiro. p. 194-199.
22. Leyl, M. B. 2011. 5 Smart Social PR Campaigns to Learn From. Available from URL: http://mashable.com/2011/03/08/social-pr-campaigns/. 27 June 2013.
23. Urmann, D. 2012. 7 Cool Social Media Trends in the Travel Industry. Available from URL: http://socialfresh.com/travel-social-media-tools/. 27 June 2013.
24. Walker, L. 2013. 8 Social Travel Networks Reviewed. Available from URL: http://personalweb.about.com/od/industry/tp/Social-Travel.htm. 16 August 2013.
25. Hays, S., John Page, S., and Buhalisb, D. 2013. Social Media as a Destination Marketing Tool: Its Use by National Tourism Organisations. *Current Issues in Tourism* 16(3): 211-239.
26. Michlmayr, E., Cayzer, S., and Shabajee, P. 2007. Add-A-Tag: Learning Adaptive User Profiles from Bookmark Collections. Proceedings of the 1st Collections International Conference on Weblogs and Social Media. 26-28 March 2007. Boulder, Colorado. U.S.A.
27. Durao, F., and Dolog, P. 2009. A Personalized Tag-Based Recommendation in Social Web Systems. *Adaptation and Personalization for Web* 2.0: 40.
28. Davoodi, F. G., and Fatemi, O. 2012. Tag Based Recommender System for Social Bookmarking Sites. Proceedings of the 2012 International Conference on Advances in Social Networks Analysis and Mining. 26-29 August 2012. Istanbul. Turkish. p. 934-940.
29. Castillo, C., Mendoza, M., and Poblete, B. 2011. Information Credibility on Twitter. Proceedings of the 20th International Conference on World Wide Web. 28 March-1 April 2011. Hyderabad. India. p. 675-684.
30. Wasim, M., Shahzadi, I., Ahmad, Q., and Mahmood, W. 2011. Extracting and Modeling User Interests Based on Social Media. Proceedings of the 2011 IEEE 14th International Conference on Multitopic Conference. 22-24 December 2011. Karachi. Pakistan. p. 284-289.
31. Lim, K. H., and Datta, A. 2013. Interest Classification of Twitter Users Using Wikipedia. Proceedings of the 9th International Symposium on Open Collaboration. 5-7 August 2013. Hong Kong. China. p. 1-2.

32. Michelson, M., and Macskassy, S. 2010. A. Discovering Users' Topics of Interest on Twitter: A First Look. Proceedings of the Fourth Workshop on Analytics for Noisy Unstructured Text Data. 26-30 October 2010. Oronto. Canada. p. 73-80.
33. Saito, J., and Yukawa, T. 2011. Extracting User's Interest Based on Social Bookmark Tags. *'Groundbreaking Research Workshop' at the KES2010 Conference* 1(1): 7-12.
34. Rashid, A. M., Karypis, G., and Riedl, J. 2008. Learning Preferences of New Users in Recommender Systems: an Information Theoretic Approach. *Web Mining and Web Usage Analysis* 10(2): 90-100.
35. Hang, Y., Guiran, C., and Xingwei, W. 2009. A Cold-Start Recommendation Algorithm Based on New User's Implicit Information and Multi-attribute Rating Matrix. Proceedings of the 9th International Conference on Hybrid Intelligent Systems. 12-14 August 2009. Shenyang. China. Washington. IEEE Computer Society. p. 353-358.
36. Park, S.-T., and Chu, W. 2009. Pairwise Preference Regression for Cold-Start Recommendation. Proceedings of the 3rd ACM Conference on Recommender Systems. 22-25 October 2009. New York. ACM. p. 21-28.
37. Ho, Q., Yan, R., Raina, R., and Xing, E. P. 2012. Understanding the Interaction Between Interests, Conversations and Friendships in Facebook. Available from URL: http://www.ml.cmu.edu/research/dap-papers/dap_ho.pdf.
38. ไกรศักดิ์ เกษร. 2555. ระบบค้นคืนสารสนเทศ: แนวคิดและแนวทางการพัฒนาในอนาคต. พิษณุโลก. โฟกัสพริ้นติ้ง. หน้า 134-135
39. นฤพนธ์ พนาวงศ์ และ จักรกฤษณ์ เสน่ห์. ระบบค้นหาสถานที่ท่องเที่ยวในประเทศไทยด้วยหลักการออนโทโลจีและเนมแมทชิ่ง. *Journal of Information Science and Technology* 1(2): 60-69.
40. Eugene, A., Carlos, C., Debora, D., Aristides, G., and Gilad, M. 2008. Finding High-Quality Content in Social Media. Proceedings of the 2008 International Conference on Web Search and Data Mining. 11-12 February 2008. California. U.S.A. p. 183-194.
41. Fisher, C., Lauria, E., Chengalur-Smith, S., and Wang, R. 2012. Introduction to Information Quality. U.S.A. AuthorHouse.
42. Wang, R. Y. and Strong, D. M. 1996. Beyond Accuracy: What Data Quality Means to Data Consumers. *Journal of Management Information Systems* 12(4): 5-33.
43. Ballou, D. P., and Pazer, H. L. 1985. Modeling Data and Process Quality in Multi-Input, Multi-Output Information Systems. *Management Science* 31(2): 150-162.

ได้รับบทความวันที่ 11 ตุลาคม 2556
ยอมรับตีพิมพ์วันที่ 20 กุมภาพันธ์ 2557