ระบบแนะนำสถานที่ท่องเที่ยวโดยใช้ข้อมูลจากเครือข่ายสังคม

Attraction Recommendation System Using Social Network Information

วนารัตน์ จุฬ์พันธ์ทอง และไกรศักดิ์ เกษร*

Wanarat Juraphanthong and Kraisak Kesorn*

ภาควิชาวิทยาการคอมพิวเตอร์และเทคโนโลยีสารสนเทศ คณะวิทยาศาสตร์ มหาวิทยาลัยนเรศวร

Department of Computer Science and Information Technology, Faculty of Science, Naresuan University

บทคัดย่อ

ระบบแนะนำข้อมูลการท่องเที่ยวเฉพาะบุคคล (Personalized Tourism Recommendation System-PTRS) เป็นเครื่องมือที่ช่วยให้นักท่องเที่ยว ได้รับข้อมูลที่ต้องการค้นหาได้รวดเร็วขึ้น โดยระบบพยายามที่จะแนะนำข้อมูลการท่องเที่ยวให้ตรงกับความสนใจ อย่างไรก็ตามความท้าทายสำหรับผู้วิจัยคือปัญหาโคลด์สตาร์ท (Cold-Start) ซึ่งเกิดจากการขาดข้อมูลที่เพียงพอมาวิเคราะห์หาความสนใจของนักท่องเที่ยว ดังนั้นบทความวิจัยนี้จึงได้นำเสนอระบบแนะนำข้อมูลการท่องเที่ยวเฉพาะบุคคล โดยใช้ประโยชน์จากข้อมูลบนบริการเครือข่ายสังคมเฟสบุ๊คมาช่วยวิเคราะห์หาความสนใจของผู้ใช้เพื่อแก้ปัญหาข้างต้น โดยมุ่งเน้นไปที่ปัญหาโคล์ดสตาร์ทที่เกิดกับผู้ที่ใช้งานระบบครั้งแรกและภาระของผู้ใช้ในการกรอกข้อมูลความสนใจด้วยตนเองในระบบแบบเดิม ซึ่งจากผลการวิจัยพบว่าการใช้ประโยชน์จากข้อมูลบนเครือข่ายสังคมของผู้ใช้สามารถช่วยลดปัญหาโคลด์สตาร์ทและแนะนำสถานที่ท่องเที่ยวได้อย่างมีประสิทธิภาพบนพื้นฐานความต้องการของผู้ใช้ระบบได้

คำสำคัญ : ปัญหาโคลด์สตาร์ท การวิเคราะห์ความสนใจของผู้ใช้ บริการเครือข่ายสังคม ระบบแนะนำข้อมูล

Abstract

Personalized Tourism Recommendation System (PTRS) is a useful tool for offering tourism information which is relevant to tourist interests. The key challenge of this research is that the information used to construct a user preference model for new user is inadequate, so called "cold start problem". Therefore, this research aims to present a new method for extracting user interest from social network services using Facebook. This technique is able to solve the key problem and allows user to spend less time to input what they are interested to the system. The experimental results showed that extracted from social network resolved such a problem and improved recommendation based on users need.

Keywords : Cold-Start problem, User Interest Extracting, Social Network Services, Recommendation System

*Corresponding author. E-mail : kraisakk@nu.ac.th

บทนำ

ปัจจุบันเทคโนโลยีอินเทอร์เน็ตมีบทบาทต่อสารสนเทศด้านการท่องเที่ยวเป็นอย่างมาก ทำให้เกิดเว็บไซต์สำหรับให้บริการข้อมูลการท่องเที่ยวและพัฒนาเป็นพาณิชย์อิเล็กทรอนิกส์ (E-Commerce) ต่าง ๆ ซึ่งเป็นผลดีต่อทั้งนักท่องเที่ยวและผู้ประกอบการ โดยนักท่องเที่ยวมีแหล่งข้อมูลในการค้นหาข้อมูลการท่องเที่ยวมากขึ้น และผู้ประกอบการมีเครื่องมือที่ช่วยนำเสนอข้อมูลสินค้าและบริการ อย่างไรก็ตามพบว่าข้อมูลเกี่ยวกับการท่องเที่ยวจำนวนมาก ทำให้นักท่องเที่ยวต้องใช้เวลามากในการค้นหาข้อมูลการท่องเที่ยวให้ได้ตรงกับความต้องการ และการนำเสนอข้อมูลสินค้าและบริการไม่ตรงกับความสนใจของผู้ซื้อก็ทำให้ผู้ประกอบการมีโอกาสในการขายสินค้าและบริการการท่องเที่ยวได้น้อยลง ระบบแนะนำข้อมูลการท่องเที่ยวเฉพาะบุคคลเป็นเครื่องมือหนึ่งที่มีประสิทธิภาพที่สามารถแก้ไขปัญหาดังกล่าวได้ โดยระบบจะช่วยนำเสนอข้อมูลการท่องเที่ยวตามความสนใจแบบเฉพาะบุคคล ซึ่งจากงานวิจัยที่ผ่านมา นักวิจัยได้พัฒนาระบบแนะนำข้อมูลการท่องเที่ยวเฉพาะบุคคลเพื่อนำมาใช้สนับสนุนนักท่องเที่ยวทั้งในด้านการวางแผนการเดินทาง (Rabanser & Ricci, 2005) ช่วยจองบัตรโดยสารเครื่องบินในวันหยุดเทศกาล (Coyle & Cunningham, 2003) แนะนำแพ็คเกจท่องเที่ยว (Schiaffino & Amandi, 2009) เป็นต้น ซึ่งระบบเหล่านี้ ทำให้นักท่องเที่ยวได้สินค้าและบริการที่ตรงกับความต้องการโดยใช้เวลาในการค้นหาน้อยลง สำหรับผู้ประกอบการนั้น ก็ได้ประโยชน์จากการเสนอสินค้าและบริการที่เหมาะสม นอกจากจะเป็นการสร้างกลยุทธ์ทางการตลาดเพื่อเพิ่มจำนวนผู้ใช้บริการให้มากยิ่งขึ้นแล้ว ยังเพิ่มโอกาสในการตัดสินใจซื้อสินค้าและบริการได้ง่ายขึ้นอีกด้วย

ในการแนะนำสินค้าและบริการโดยใช้ระบบแนะนำข้อมูลแบบเฉพาะบุคคลดังกล่าว การค้นหาความสนใจของนักท่องเที่ยวนั้นเป็นหัวใจสำคัญ (Zhu, He & Wang, 2012) โดยข้อมูลความสนใจที่วิเคราะห์ออกมาจะเก็บอยู่ในแบบจำลองความสนใจของผู้ใช้ (User Interest Model) และนำไปใช้ในการแนะนำข้อมูลการท่องเที่ยวต่อไป อย่างไรก็ตามการค้นหาความสนใจมีความท้าทายอย่างยิ่งสำหรับผู้พัฒนาระบบ คือการไม่มีข้อมูลเพียงพอในการสรุปความสนใจสำหรับผู้ที่เริ่มต้นใช้งานระบบ หรือเรียกว่าปัญหาโคลด์สตาร์ท (Cold Start Problem) ซึ่งมักจะเกิดขึ้นในระบบที่ใช้การวิเคราะห์ความสนใจจากข้อมูลที่ได้โดยอ้อม (Implicit) จากการใช้งานหรือพฤติกรรมของผู้ใช้บนระบบที่ต้องใช้เวลามากพอจึงจะได้ข้อสรุปความสนใจ ทำให้ระบบไม่สามารถให้การแนะนำที่ดีกับผู้ที่เริ่มต้นใช้งานระบบได้ จึงมีนักวิจัยหลายท่านได้นำเสนอแนวคิดในการพัฒนาระบบเพื่อแก้ปัญหาข้างต้น เช่น ระบบ Entrée (Burke, 2000) ได้ใช้การสอบถามข้อมูลความสนใจจากผู้ใช้โดยตรง (Explicit) เมื่อเริ่มต้นใช้งานระบบเพื่อสร้างแบบจำลองความสนใจของผู้ใช้ แต่ปัญหาที่เกิดคือผู้ใช้ต้องเสียเวลาในขั้นตอนการตอบคำถามเหล่านี้มากเกินไป แนวคิดต่อมาคือระบบ TravelPlanner (Chin & Porage, 2001) ซึ่งพยายามลดขั้นตอนในการตอบคำถามโดยการใช้คำถามน้อยลงและเลือกคำถามที่มีประโยชน์สำหรับการวิเคราะห์ความสนใจ อีกแนวคิดหนึ่งคือระบบ SPETA (García-Crespo *et al*., 2009) จะขอข้อมูลจากผู้ใช้ เช่น ประเภทสถานที่ที่ต้องการให้แนะนำ ระดับความสนใจในประเภทสถานที่นั้น เป็นต้น จากนั้นจะใช้ข้อมูลจากเครือข่ายสังคมที่ใช้มาตรฐานเอพีไอของ OpenSocial (opensocial.org) ในการวิเคราะห์ความสนใจของสมาชิก การผสมผสานข้อมูลข้างต้น จุดประสงค์ก็เพื่อลดขั้นตอนในการตอบคำถามเช่นกัน จะเห็นว่าผู้พัฒนาระบบให้ความสำคัญในการแก้ปัญหาโคลด์สตาร์ทและการสูญเสียเวลาตอบคำถามของผู้ใช้ ในระบบแนะนำข้อมูลแบบเฉพาะบุคคลประเภทอื่นๆ ยังมีแนวคิดในการนำข้อมูลจากผู้ใช้อื่นมาแก้ไขปัญหา เช่น Rashid และคณะ (2008) ใช้ข้อมูลของผู้ใช้อื่นที่มีต่อระบบ เช่น ความถี่ในการเข้าชมข้อมูลภาพยนตร์ ข้อมูลแสดงความคิดเห็นและการให้คะแนนภาพยนตร์ เป็นต้น ซึ่งภาพยนตร์ที่มีความนิยมในหมู่ผู้ใช้ในระบบก็จะถูกแนะนำไปยังผู้ใช้ที่เริ่มต้นใช้ระบบด้วย Hang และคณะ (2009) ที่ใช้วิธีการแบบ Collaborative โดยวัดค่าความคล้ายระหว่างผู้ใช้และช่วยจับคู่ระหว่างผู้ใช้กับสิ่งที่สนใจ โดยผู้ที่เริ่มต้นใช้งาน

ระบบจะจับคู่กับสิ่งที่สนใจที่มีค่ามากที่สุดจากผู้ใช้ที่มีความคล้ายคลึงกัน นอกจากนี้การนำข้อมูลจากเครือข่ายสังคม (Social Network) มาช่วยในระบบแนะนำข้อมูล เป็นแนวคิดที่ได้รับความสนใจเช่นกัน เนื่องจากในปัจจุบันการพัฒนาของเครือข่ายสังคมที่เป็นไปอย่างรวดเร็ว ทำให้ข้อมูลทั้งที่เกิดจากการสนทนา แสดงความคิดเห็นและแบ่งปัน มีจำนวนที่มากขึ้น จึงมีนักวิจัยพยายามจะใช้ประโยชน์จากข้อมูลเหล่านี้มาวิเคราะห์เพื่อหาความสนใจของผู้ใช้ เช่น การใช้แท็ก (Tag) ในโซเชียลบุ๊กมาร์ก (Social Bookmarking) วิเคราะห์ความสนใจของผู้ใช้เพื่อสร้างคำแนะนำในระบบค้นหาเพลง (Firan, Nejdl & Paiu, 2007) การใช้ข้อความทวีต (Tweet) ในทวิตเตอร์ (Twitter.com) ค้นหาความสนใจของผู้ใช้เพื่อสร้างระบบแนะนำข่าว (Abel, Gao, Houben & Tao, 2011) แนวคิดการนำข้อความบนเฟสบุ๊คมาช่วยในการปรับปรุงแบบจำลองความสนใจสำหรับระบบแนะนำเพคเกจการท่องเที่ยว (Lorenzi *et al*., 2014) นอกจากนี้ยังมีงานวิจัยที่ใช้เครือข่ายสังคมมุ่งเน้นแก้ปัญหาโคลด์สตาร์ท เช่น การใช้ page likes บนเฟสบุ๊คเพิ่มประสิทธิภาพการแนะนำหนังสือไปยังผู้ใช้ที่ไม่มีข้อมูลประวัติการซื้อหนังสือก่อนหน้า (Sedhain *et al*., 2014) การนำข้อมูลสถานที่การเช็คอินจากโฟรสแควร์ (Foursquare.com) หาความสัมพันธ์ระหว่างผู้ใช้จากนั้นคาดการณ์ความสนใจโดยอาศัยหลักการความน่าจะเป็น และแนะนำข้อมูลตามหัวข้อความสนใจที่คาดการณ์ (Castillejo, Almeida, & López-de-Ipiña, 2012) จะเห็นว่าเครือข่ายสังคมเป็นแหล่งข้อมูลที่สามารถนำไปสู่ข้อสรุปความสนใจ มีแนวโน้มช่วยแก้ไขปัญหาโคลด์สตาร์ท นอกจากนี้ยังลดภาระของผู้ใช้ที่ต้องกรอกข้อมูลความสนใจด้วยตนเองในระบบเดิมอีกด้วย

แต่อย่างไรก็ตามในโดเมนของการท่องเที่ยวการแก้ปัญหาโคลด์สตาร์ทโดยใช้ข้อความทวีตหรือข้อความบนเฟสบุ๊คยังมีประสิทธิภาพต่ำเนื่องจากข้อความที่เป็นภาษาธรรมชาติ (Natural language) ที่มนุษย์สร้างขึ้นมีความคลุมเครือ (Ambiguous) จึงทำให้การสกัดความสนใจจากข้อความมีความซับซ้อน ใช้เวลานาน ซึ่งส่งผลให้ผู้ใช้งานเสียเวลารอระบบแนะนำสถานที่ท่องเที่ยวนานเกินไป ส่วนข้อมูลประเภทแท็ก แม้จะช่วยกรองคำและลดปริมาณของข้อมูลประเภทข้อความ แต่คำที่ได้มาจากแท็กก็ไม่มีคุณภาพมากพอในการวิเคราะห์ความสนใจของผู้ใช้ เนื่องจากผู้ใช้ส่วนมากสร้างแท็กขึ้นเพื่อแสดงอารมณ์เท่านั้น จึงไม่สามารถนำมาวิเคราะห์ความสนใจได้ โดยเฉพาะอย่างยิ่งในโดเมนของการท่องเที่ยว ด้วยความสำคัญของปัญหาดังกล่าว ผู้วิจัยจึงมีแนวคิดในการพัฒนาระบบแนะนำข้อมูลการท่องเที่ยวเฉพาะบุคคล เพื่อช่วยนำเสนอข้อมูลสถานที่ท่องเที่ยวให้ได้ตรงกับความสนใจของผู้ใช้อย่างแท้จริง และนำเสนอแนวคิดในการพัฒนาเทคนิคการค้นหาความสนใจที่สามารถช่วยลดปัญหาโคลด์สตาร์ทให้กับผู้ใช้ใหม่และมีประสิทธิภาพโดยใช้ประโยชน์จากข้อมูลเช็คอินบนบริการเครือข่ายสังคมที่ได้รับความนิยมอย่างเฟสบุ๊ค เนื่องจากสามารถใช้ข้อมูลสถานที่เช็คอินบ่งบอกถึงลักษณะความสนใจในสถานที่มากกว่าข้อมูลประเภทอื่นและมีความชัดเจนเกี่ยวกับสถานที่ท่องเที่ยวมากกว่าข้อความทวีต นอกจากนี้ปริมาณของข้อมูลที่มีน้อยกว่า ซึ่งจะช่วยให้ระบบใช้เวลาน้อยลงในการประมวลผล และสามารถสรุปความสนใจเกี่ยวกับสถานที่ท่องเที่ยวได้อย่างถูกต้องมากยิ่งขึ้น

วิธีการวิจัย

1. ประชากรและกลุ่มตัวอย่าง

ประชากรในการวิจัยนี้คือนักท่องเที่ยวที่มีความสนใจหรือวางแผนจะท่องเที่ยวในจังหวัดต้นแบบ และเป็นกลุ่มที่นิยมใช้แหล่งข้อมูลบนอินเทอร์เน็ตในการหาข้อมูลเพื่อวางแผนการท่องเที่ยว การเลือกกลุ่มตัวอย่างของประชากรใช้วิธีการสุ่มแบบ

เจาะจง (Purposive Sampling) เพื่อให้ได้กลุ่มตัวอย่างที่ตรงกับวัตถุประสงค์ของการทดลองแต่ละแบบ โดยเลือกตัวอย่างจำนวน 42 ตัวอย่างที่มีบัญชีผู้ใช้เฟสบุ๊คและมีการใช้งานอย่างน้อย 1 ครั้งต่อสัปดาห์

2. การเก็บรวบรวมข้อมูลสถานที่ท่องเที่ยว

ผู้วิจัยใช้ข้อมูลสถานที่ท่องเที่ยว 5 จังหวัด ได้แก่ พิษณุโลก อุตรดิตถ์ ตาก สุโขทัย และ เพชรบูรณ์ โดย Harfield (2014) ร่วมกับข้อมูลสถานที่ท่องเที่ยวจากการท่องเที่ยวแห่งประเทศไทย (ททท.) เป็นต้นแบบพื้นที่การแนะนำของระบบ

3. การออกแบบและพัฒนาระบบ

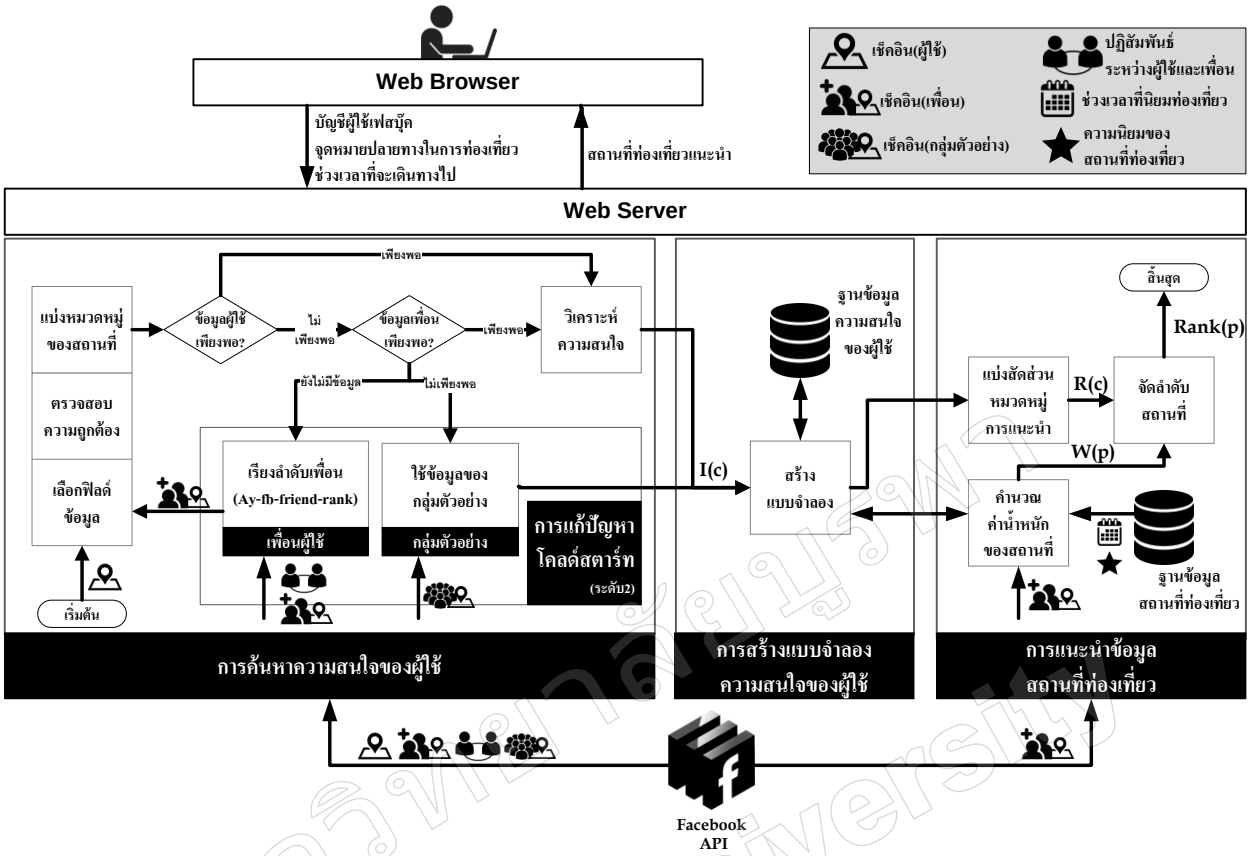
ผู้วิจัยได้ออกแบบและพัฒนาระบบแนะนำข้อมูลการท่องเที่ยวเฉพาะบุคคลบนพื้นฐานของเทคโนโลยีเว็บ ซึ่งสามารถใช้งานผ่านเว็บบราวเซอร์ โดยมุ่งเน้นไปที่การใช้ประโยชน์จากข้อมูลเช็คอินบนเครือข่ายสังคมมาช่วยแก้ไขปัญหาโคลด์สตาร์ทที่มักเกิดขึ้นเมื่อผู้ใช้เข้าใช้งานระบบครั้งแรก และช่วยปรับปรุงประสิทธิภาพในการแนะนำข้อมูลให้ดียิ่งขึ้น นอกจากนี้ยังคำนึงถึงปัญหาสำคัญในส่วนของการใช้งาน คือ ความยุ่งยากในการให้ข้อมูลความสนใจจากการตอบคำถามหรือการทำแบบสอบถาม และความรวดเร็วในการประมวลของระบบ

ในการค้นหาความสนใจและแนะนำข้อมูลไปยังผู้ใช้ มีวิธีที่ได้รับความนิยมคือ Content-based และ Collaborative โดยวิธีการ Content-based นั้นเป็นการแนะนำตามความต้องการหรือความสนใจที่วิเคราะห์จากพฤติกรรมในอดีตของผู้ใช้เอง ส่วนวิธีการ Collaborative เป็นการแนะนำโดยอาศัยผู้ใช้อื่นที่มีความสนใจคล้ายกัน (Kabassi, 2010) ซึ่งในงานวิจัยนี้ได้ใช้วิธีการแบบผสมผสาน (Hybrid) ทั้งสองวิธีข้างต้น เนื่องจากในการศึกษาพฤติกรรมการใช้งานบนเครือข่ายสังคมเฟสบุ๊ค พบว่าส่วนหนึ่งของผู้ใช้ที่เข้าใช้งานเฟสบุ๊คเป็นประจำ ไม่ได้ทำเช็คอินในทุกครั้งที่ใช้งาน ซึ่งอาจทำให้ผู้ใช้ไม่มีข้อมูลประวัติการเช็คอินที่มากเพียงพอในการบอกถึงความสนใจโดยใช้วิธีการแบบ Content-based ได้ หรืออาจเกิดปัญหาโคล์ดสตาร์ทขึ้นอีกนั่นเอง ดังนั้นผู้วิจัยจึงได้แบ่งปัญหาโคลด์สตาร์ทที่เกิดกับผู้ใช้ที่เริ่มต้นใช้งานระบบเป็น 2 ระดับ และใช้วิธีการทั้งสองแบบผสมผสานกันเพื่อแก้ไขปัญหาทั้ง 2 ระดับ คือ

1) โคลด์สตาร์ทระดับที่ 1 กรณีที่ระบบไม่มีข้อมูลความสนใจของผู้ใช้ที่เริ่มใช้งานระบบ โดยใช้หลักการ User-based Content Filtering ซึ่งจะข้อมูลการเช็คอินในอดีตของผู้ใช้ในการค้นหาความสนใจและแนะนำข้อมูลไปยังผู้ใช้

2) โคลด์สตาร์ทระดับที่ 2 กรณีที่ไม่สามารถใช้ข้อมูลการเช็คอินของผู้ใช้ในโคลด์สตาร์ทในระดับที่ 1 ได้ โดยใช้หลักการ Friend-based Collaborative Filtering ซึ่งจะใช้ข้อมูลการเช็คอินจากเพื่อนที่มีพฤติกรรม หรือความสนใจคล้ายกัน

ผู้วิจัยได้ออกแบบระบบ โดยแบ่งการทำงานออกเป็น 4 ส่วน มีองค์ประกอบหลักดังภาพที่ 1 ประกอบด้วย 1) การค้นหาความสนใจของผู้ใช้ (User Interest Detection) เป็นการหาข้อสรุปความสนใจโดยใช้ข้อมูลจากเครือข่ายสังคม 2) การแก้ไขปัญหาโคลด์สตาร์ท (Cold-Start Problem Method) เป็นวิธีการแก้ปัญหากรณีที่ผู้ใช้มีข้อมูลจากเครือข่ายสังคมไม่เพียงพอ 3) การสร้างแบบจำลองความสนใจของผู้ใช้ (User Interest Modeling) เป็นการการจัดเก็บข้อมูลความสนใจลงในฐานข้อมูล และ 4) การแนะนำข้อมูลสถานที่ท่องเที่ยว (Tourism Attractions Recommendation) เป็นการแนะนำสถานที่ท่องเที่ยวโดยใช้ข้อมูลความสนใจที่ได้มาจากแบบจำลองความสนใจของผู้ใช้



ภาพที่ 1 องค์ประกอบของระบบแนะนำข้อมูลการท่องเที่ยวเฉพาะบุคคลที่ผู้วิจัยพัฒนา

3.1 การค้นหาความสนใจของผู้ใช้

กระบวนการนี้เป็นการวิเคราะห์ข้อมูลจากเครือข่ายสังคมเพื่อค้นหาความสนใจของผู้ใช้ โดยใช้วิธีการแบบผสมผสาน ซึ่งมีกระบวนการค้นหาความสนใจของผู้ใช้เป็นดังนี้

3.1.1 การเลือกฟิลด์ข้อมูล

ตารางที่ 1 ข้อมูลการเช็คอินของผู้ใช้จำนวน 1 เช็คอิน

ข้อมูลที่ได้จากเฟสบุ๊ค	ความหมาย
{ "id": "196230743801741",	หมายเลขไอดีของสถานที่ ที่เช็คอิน
"name": "วัดพระศรีรัตนมหาธาตุราชวรวิหาร",	ชื่อของสถานที่ ที่เช็คอิน
"category_list": [	รายการหมวดหมู่ของสถานที่ที่จัดแบ่ง
{ "id": "124861974254366",	หมายเลขไอดีหมวดหมู่ที่ 1 ของสถานที่
"name": "Tours & Sightseeing"},	ชื่อหมวดหมู่ที่ 1 ของสถานที่
{"id": "197097220301977",	หมายเลขไอดีหมวดหมู่ที่ 2 ของสถานที่
"name": "Historical Place" }]	ชื่อหมวดหมู่ที่ 2 ของสถานที่
"created_time": "2013-04-13T09:21:03+0000"}	เวลาในการสร้างโพสที่เช็คอิน

ข้อมูลเช็คอินจากเฟสบุ๊ค ใช้การดึงข้อมูลผ่าน Facebook API ซึ่งได้ข้อมูลอยู่ในรูปแบบ JSON โดยข้อมูลที่สามารถดึงมาจากเฟสบุ๊คประกอบด้วยหลายฟิลด์ข้อมูล อย่างไรก็ตามผู้วิจัยเลือกใช้เฉพาะฟิลด์ที่มีประโยชน์ต่อการวิเคราะห์ความสนใจของผู้ใช้ โดยฟิลด์ข้อมูลที่นำมาใช้ภายในระบบประกอบด้วย 1.id (รหัสสถานที่) 2.name (ชื่อสถานที่) 3.created_time (เวลาที่เช็คอิน) และ 4.category_list (หมวดหมู่ของสถานที่) ดังตารางที่ 1

3.1.2 การตรวจสอบความถูกต้อง (Cleansing)

เมื่อได้ข้อมูลของผู้ใช้ตามฟิลด์ข้อมูลที่กล่าวข้างต้นแล้ว ผู้วิจัยจะตรวจสอบความถูกต้องของข้อมูลการเช็คอิน โดยเลือกใช้ข้อมูลการเช็คอินที่มีจุดประสงค์เพื่อการท่องเที่ยวเท่านั้น เนื่องจากการศึกษาพบว่าผู้ใช้ในเครือข่ายสังคมมักจะเช็คอินในสถานที่หนึ่ง ๆ เป็นจำนวนมากเหมือน ๆ กัน เช่น ห้างสรรพสินค้า โรงภาพยนตร์ บ้านของผู้ใช้ เป็นต้น ซึ่งสถานที่นี้ส่วนมากเป็นกิจกรรมในการดำเนินชีวิตทั่วไป ซึ่งหากวิเคราะห์ความสนใจจากข้อมูลที่ไม่ผ่านการตรวจสอบความถูกต้องเหล่านี้ ความสนใจที่วิเคราะห์ได้อาจไม่สอดคล้องกับความสนใจของผู้ใช้ในสถานที่ท่องเที่ยวอย่างแท้จริง อย่างไรก็ตามมีบางกรณีที่สถานที่เหล่านี้อาจจัดว่าเป็นสถานที่ท่องเที่ยวได้ ดังนั้นผู้วิจัยจึงใช้วิธีการพิจารณามิติของเวลาเข้ามาแทนการตัดสถานที่ที่ไม่ใช่จุดประสงค์เพื่อการท่องเที่ยวออกไปทันที โดยจะเลือกตัดสถานที่มีการเช็คอินมากกว่าหนึ่งครั้งต่อสัปดาห์ อีกทั้งยังใช้ค่าของเวลาลดความซ้ำซ้อนของการเช็คอินที่อาจเกิดขึ้นหลายครั้งในวันเดียวกัน ยกตัวอย่างเช่น การโพสรูปภาพหลาย ๆ รูปภาพที่ติดกับเช็คอินในสถานที่เดียวกันในวันเดียวกัน สถานที่ที่ซ้ำซ้อนเหล่านี้จะถูกตัดออกให้เหลือเพียงหนึ่งครั้งในหนึ่งวัน

3.1.3 การประเมินความน่าเชื่อถือของข้อมูล

ผู้วิจัยได้ประเมินความน่าเชื่อถือของข้อมูลเช็คอินโดยสอบถามกลุ่มตัวอย่างถึงสถานที่ที่ได้ไปจริงจากสถานที่ที่ได้เช็คอินบนเฟสบุ๊ค แสดงผลดังภาพที่ 2

ภาพที่ 2 สถานที่ที่ผู้ใช้ได้ไปจริงจากสถานที่เช็คอิน

จากกลุ่มตัวอย่างที่มีการเช็คอินทั้งหมดจำนวน 2494 ครั้ง พบว่า 80.3% เช็คอินไปยังสถานที่ที่เคยไปจริง อย่างไรก็ตามเมื่อกรองการเช็คอินเป็นเช็คอินที่เป็นจุดประสงค์เพื่อการท่องเที่ยว พบว่าผู้ใช้ได้ไปยังสถานที่ที่ได้เช็คอินจริง 97.6%

3.1.4 การแบ่งหมวดหมู่ของสถานที่เช็คอิน

การแบ่งหมวดหมู่ (Classification) ของสถานที่เช็คอินจะเริ่มกระบวนการเมื่อการเช็คอินผ่านการตรวจสอบความถูกต้องเรียบร้อยแล้ว โดยตรวจสอบว่าสถานที่ท่องเที่ยวที่ได้จากการเช็คอินนั้นอยู่ในหมวดหมู่ใด ใน 6 หมวดหมู่ที่จัดไว้ในฐานข้อมูล category_list โดยอัตโนมัติ

ในการสร้างฐานข้อมูล category_list นั้น ใช้ category_list พื้นฐานของเฟสบุ๊ค category_list ที่เก็บจากกลุ่มตัวอย่าง และ category_list จากผู้ใช้เฟสบุ๊คทั่วไปที่ตั้งสถานะการเช็คอินเป็นแบบ public ใช้วิธีการแบ่งหมวดหมู่

จากคำสำคัญที่เกี่ยวข้องกับหมวดหมู่นั้น อย่างไรก็ตามพบว่าเฟสบุ๊คมีฟังก์ชันให้ผู้ใช้สามารถสร้าง category_list เองได้ ทำให้เกิดกรณีชื่อ category_list ต่างกัน แต่หมายถึงสิ่งเดียวกัน หรือความเฉพาะเจาะจงของชื่อ ซึ่งในระบบต้นแบบนี้ผู้วิจัยได้จัดการกับปัญหาข้างต้นโดยวิธีการแบบ Knowledge-based ซึ่งฐานข้อมูล category_list จะแบ่งตามกลุ่มความสนใจในสถานที่ท่องเที่ยวออกเป็น 6 หมวดหมู่ โดยปรับปรุงมาจากหมวดหมู่หลักของการท่องเที่ยวแห่งประเทศไทย (tourismthailand.org) ได้แก่

1) **สถานที่ท่องเที่ยวทางวัฒนธรรม** (Culture) หมายถึง แหล่งท่องเที่ยวที่ทำให้นักท่องเที่ยวได้รู้จักวิถีชีวิต วัฒนธรรม และความเป็นอยู่ของผู้คนในชุมชน ในสถานที่ที่จัดนิทรรศการซึ่งแสดงเรื่องราวและงานศิลปหัตถกรรมของผู้คนในชุมชน เช่น งานประเพณี การแสดงศิลปวัฒนธรรม งานแสดงสินค้าพื้นเมือง เป็นต้น โดยจัด category_list ในเฟสบุ๊คให้อยู่ในหมวดหมู่นี้ ตัวอย่างเช่น Cultural_Gifts_Store, Fairground, Shopping_District เป็นต้น

2) **สถานที่ท่องเที่ยวทางประวัติศาสตร์** (Historical) หมายถึง สถานที่สำคัญในการท่องเที่ยวที่แสดงถึงร่องรอยของความรุ่งเรืองในอดีตหลายแห่ง ซึ่งบางแห่งได้รับคัดเลือกให้เป็นแหล่งมรดกโลก เช่น โบราณสถาน อุทยานประวัติศาสตร์ วัด ศาสนสถาน สิ่งก่อสร้างที่มีคุณค่าทางศิลปะและสถาปัตยกรรม เป็นต้น โดยจัด category_list ในเฟสบุ๊คให้อยู่ในหมวดหมู่นี้ ตัวอย่างเช่น Historical_Place, Buddhist_Temple, Religious_ Organization เป็นต้น

3) **สถานที่ท่องเที่ยวเพื่อความบันเทิง** (Entertainment) หมายถึง สถานที่ท่องเที่ยวเชิงสันทนาการและบันเทิง ซึ่งมีหลายรูปแบบ เช่น สวนสัตว์ สวนสนุก ย่านบันเทิง สวนสาธารณะ สิ่งปลูกสร้างต่างๆ เป็นต้น โดยจัด category_list ในเฟสบุ๊คให้อยู่ในหมวดหมู่นี้ ตัวอย่างเช่น Market, Park, Art&Entertaimment, Outdoor, Zoo เป็นต้น

4) **สถานที่ท่องเที่ยวทางธรรมชาติ** (Nature) หมายถึง สถานที่ท่องเที่ยวที่เป็นแหล่งธรรมชาติหลากหลายประเภท เช่น ภูเขา ผืนป่า น้ำตก โถงถ้ำ ทะเลสาบ ทุ่งดอกไม้ น้ำพุร้อน ฯลฯ โดยผู้วิจัยได้เลือก category_list ในเฟสบุ๊คให้อยู่ในหมวดหมู่นี้ ตัวอย่างเช่น National_Park, Mountain, River, Rock Climbing, Lake เป็นต้น

5) **สถานที่ท่องเที่ยวเพื่อการศึกษา** (Educational) หมายถึง สถานที่ท่องเที่ยวเชิงวิชาการสำหรับผู้ที่สนใจใฝ่หาความรู้ อาทิ ห้องสมุด พิพิธภัณฑ์เพื่อการศึกษา ศูนย์ฝึกอบรม เป็นต้น โดยจัด category_list ในเฟสบุ๊คให้อยู่ในหมวดหมู่นี้ ตัวอย่างเช่น Museum, Education เป็นต้น

6) **สถานที่ท่องเที่ยวทางวิถีชีวิตชุมชน** (Community-based) เป็นสถานที่ท่องเที่ยวที่แสดงถึงวิถีชีวิตของคนในชุมชน เช่น การท่องเที่ยวในรูปแบบโฮมสเตย์ ตลาด ตลาดน้ำ ที่เป็นแหล่งท่องเที่ยวที่เต็มไปด้วยสีสันของวิถีชีวิตไทยดั้งเดิมและอุดมด้วยอาหารไทยหลากหลาย โดยจัด category_list ในเฟสบุ๊คให้อยู่ในหมวดหมู่นี้ตัวอย่างเช่น Region, Farmers Market เป็นต้น

3.1.5 การวิเคราะห์ความสนใจ

ขั้นตอนนี้เป็นการวิเคราะห์ความสนใจจากพฤติกรรมการท่องเที่ยวในอดีตของผู้ใช้ ซึ่งหมวดหมู่สถานที่ท่องเที่ยวที่ผู้ใช้ไปท่องเที่ยวบ่อย จะถูกมองว่าเป็นหมวดหมู่สถานที่ที่ผู้ใช้สนใจมากกว่าหมวดหมู่สถานที่ที่ผู้ใช้ไม่ค่อยได้ เดินทางไปท่องเที่ยว ดังนั้นระบบจะวัดความสนใจจากความถี่ (Item Frequency) ที่ผู้ใช้ได้ไปหมวดหมู่สถานที่นั้น และแสดงเป็นค่าระดับความสนใจ (Interest Level - I) ของผู้ใช้ในแต่ละหมวดหมู่ของสถานที่ท่องเที่ยว โดยใช้เทคนิคการแสดงค่าความสนใจด้วยตัวเลข (Numerical Representations) (Moreno *et al*., 2013) โดยลักษณะของเทคนิคนี้ ค่าระดับความสนใจ จะคำนวณ

ไปยังแต่ละหมวดหมู่ของสถานที่ท่องเที่ยว โดยให้ค่าระดับความสนใจมีค่า 0.0 ถึง 1.0 โดย 0 หมายถึงไม่มีความสนใจ และถึง 1 หมายถึงมีความสนใจระดับสูงสุด

3.2 การแก้ไขปัญหาโคลด์สตาร์ท

กระบวนการนี้เป็นวิธีการใช้พฤติกรรมการท่องเที่ยวในอดีต ผ่านการเช็คอิน เพื่อแก้ไขปัญหาโคล์ดสตาร์ท 2 ระดับ ดังนี้

1) **โคลด์สตาร์ทระดับที่** 1 ผู้ใช้มีข้อมูลการเช็คอิน (ข้อมูลการท่องเที่ยวในอดีต) เพียงพอให้ระบบวิเคราะห์ความสนใจของผู้ใช้ใหม่ได้ โดยจะวัดความเพียงพอของข้อมูลจากจำนวนการเช็คอิน ซึ่งหากข้อมูลการเช็คอินที่ผ่านการตรวจสอบความถูกต้องแล้ว มีไม่ต่ำกว่า 10 เช็คอิน จะถือว่าผู้ใช้มีข้อมูลเพียงพอ (จากการทดลองที่ 1 ให้ค่าความถูกต้องในการลำดับความสนใจที่ยอมรับได้) ระบบจะใช้ข้อมูลเช็คอินของผู้ใช้เองในการวิเคราะห์ความสนใจ โดยนับจำนวนของการเช็คอินที่ปรากฏในแต่ละหมวดหมู่การท่องเที่ยวของแต่ละผู้ใช้ กำหนดให้ $I(c)$ คือค่าระดับความสนใจของแต่ละผู้ใช้ในแต่ละหมวดหมู่การท่องเที่ยว (Category - c) คำนวณดังสมการที่ 1

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

n_c คือจำนวนของการเช็คอินที่ปรากฏในหมวดหมู่ c และ n_k คือจำนวนของการเช็คอินที่ปรากฏ ในทุก ๆ หมวดหมู่ k

2) **โคลด์สตาร์ทระดับที่** 2 ผู้ใช้มีข้อมูลการเช็คอิน (ข้อมูลการท่องเที่ยวในอดีต) ไม่เพียงพอให้ระบบวิเคราะห์ความสนใจของผู้ใช้ใหม่ได้ โดยเกิดขึ้นกรณีผู้ใช้มีการเช็คอินน้อยกว่า 10 เช็คอิน ระบบจะใช้ข้อมูลเช็คอินจากเพื่อนของผู้ใช้ในเฟสบุ๊คมาพิจารณาแทน ซึ่งเพื่อนมักจะมีความสนใจหลายอย่างเหมือนกันเมื่อมีพฤติกรรมที่คล้ายกัน (Ye, Yin, Lee, & Lee, 2011) เช่น เพื่อนที่ไปรับประทานอาหาร ชมภาพยนตร์ หรือ ท่องเที่ยวด้วยกัน ผู้วิจัยจึงพิจารณาถึงตัวแปรที่บอกถึงพฤติกรรมที่คล้ายกันบนเครือข่ายสังคมนั่นคือการทำกิจกรรมบางอย่างร่วมกันหรือการมีปฏิสัมพันธ์กัน ดังนั้นเพื่อนที่มีการปฏิสัมพันธ์กับผู้ใช้ในระดับสูง จะมีความสนใจที่คล้ายกับผู้ใช้มากกว่าเพื่อนที่มีการปฏิสัมพันธ์กับผู้ใช้ในระดับต่ำ ดังนั้นผู้วิจัยจึงใช้ข้อมูลจากเพื่อนที่มีความคล้ายคลึงกันเพื่อแก้ไขปัญหาโคลด์สตาร์ทนี้ โดยระบบเรียงลำดับเพื่อนตามระดับปฏิสัมพันธ์ที่มีกับผู้ใช้จากการประยุกต์ใช้อัลกอริทึมชื่อ ay-fb-friend-rank (gajus, 2012) ซึ่งอาศัยหลักการที่ชื่อว่า EdgeRank (Widman, 2014) โดย EdgeRank นั้นประกอบด้วย 3 ส่วน คือ 1) Affinity Score คือค่าที่คำนวณจากการปฏิสัมพันธ์ประเภทต่าง ๆ ที่มีระหว่างผู้ใช้ เช่น การกดไลค์ การแชร์ และการแสดงความคิดเห็นของเพื่อนบนโพสที่มีร่วมกัน การเช็คอินร่วมกัน การมีเพื่อนร่วมกัน เป็นต้น 2) Edge Weight คือค่าน้ำหนักที่ให้กับ Affinity Score แต่ละประเภท เช่น การแสดงความคิดเห็นจะมีค่าน้ำหนักสูงกว่าการกดไลค์ เป็นต้น และ 3) Time Decay คือค่าระยะเวลาที่โพสถูก สร้างขึ้น เช่น โพสที่ถูกสร้างล่าสุดจะมีค่าสูงกว่าโพสที่สร้างเมื่อสัปดาห์ที่แล้ว เป็นต้น โดยอัลกอริทึม ay-fb-friend-rank นั้นใช้เพียง 2 ส่วนประกอบ คือ 1) Affinity Score และ 2) Edge Weight เนื่องจากต้องการหลีกเลี่ยงการคำนวณ Time Decay ที่ต้องใช้เวลานาน

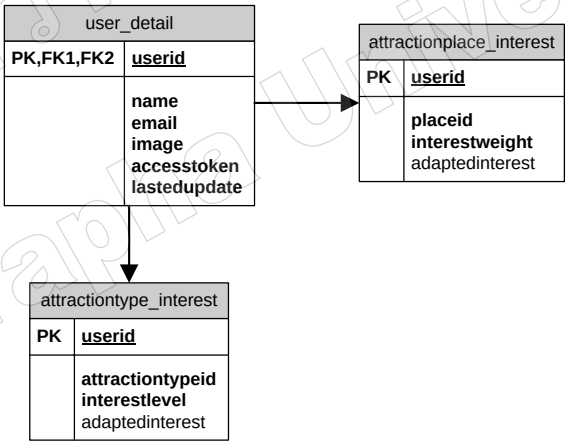
เมื่อได้รายชื่อเพื่อนสนิทและค่าปฏิสัมพันธ์ระหว่างผู้ใช้กับเพื่อนจากการคำนวณโดยอัลกอริทึมข้างต้น ซึ่งจากผลการทดลองที่ 2 พบว่าการเลือกใช้เพื่อน 5 อันดับแรก (กลุ่มเพื่อน) มีความเหมาะสมที่สุด (ให้ค่าความถูกต้องในการลำดับ

ความสนใจมากและใช้เวลาประมวลผลน้อย) จากนั้นระบบจะตรวจสอบเช่นเดียวกับกรณีโคลด์สตาร์ทระดับที่ 1 คือตรวจสอบข้อมูลความเพียงพอของข้อมูลเช็คอินของเพื่อนจากจำนวนเช็คอิน โดยหากมีไม่ต่ำกว่า 10 เช็คอิน จะถือว่าข้อมูลเพียงพอ (จากการทดลองที่ 1 ให้ค่าความถูกต้องในการลำดับความสนใจที่ยอมรับได้) ซึ่งจะพิจารณาเลือกเพื่อน 5 อันดับแรกที่มีค่าปฏิสัมพันธ์สูงสุดและมีข้อมูลเช็คอินเพียงพอ โดยหากยังมีเพื่อน 3 ใน 5 ที่เป็นไปตามเงื่อนไขข้างต้น ระบบใช้ข้อมูลเช็คอินจากเพื่อนคำนวณค่าระดับความสนใจของเพื่อนแต่ละคนโดยวิธีการเช่นเดียวกับการคำนวณไปยังผู้ใช้ในสมการที่ 1 และให้น้ำหนักค่าระดับความสนใจของเพื่อนแต่ละคนตามความคล้ายคลึงกับผู้ใช้ โดยเพื่อนที่มีค่าปฏิสัมพันธ์ที่สูงกว่าจะมีความสนใจที่คล้ายกับผู้ใช้มากกว่า ดังนั้นจึงใช้ค่าปฏิสัมพันธ์เพื่อให้น้ำหนักไปยังค่าระดับความสนใจของเพื่อนแต่ละคน ดังสมการที่ 2

$$I(c) = \frac{\sum_{i=1}^{n} F_i \; I_i(c)}{\sum_{i=1}^{n} F_i}, 1 \le n \le 5 \qquad (2)$$

โดย $I_i(c)$ คือค่าระดับความสนใจที่มีต่อแต่ละหมวดหมู่การท่องเที่ยวของเพื่อนคนที่ i และ F_i คือค่าปฏิสัมพันธ์ระหว่างผู้ใช้และเพื่อนคนที่ i กรณีที่ข้อมูลของผู้ใช้และกลุ่มเพื่อนไม่เพียงพอ จะใช้ $I(c)$ ที่ระบบกำหนดซึ่งเป็นค่าเฉลี่ยที่คำนวณจากกลุ่มตัวอย่าง

3.3 การสร้างแบบจำลองความสนใจของผู้ใช้



ภาพที่ 3 ฐานข้อมูลที่ใช้จัดเก็บแบบจำลองความสนใจของผู้ใช้

การสร้างแบบจำลองความสนใจของผู้ใช้ (User Interests Model) เป็นกระบวนการหลักของระบบเช่นกัน โดยหลังจากที่ระบบได้วิเคราะห์ และค้นหาความสนใจจากผู้ใช้แล้ว การจัดเก็บความสนใจเหล่านี้ก็เป็นสิ่งที่มีความสำคัญ ซึ่งแบบจำลองความสนใจของผู้ใช้นี้จะจัดเก็บข้อมูลที่แสดงถึงความสนใจของผู้ใช้ (User Interests) และข้อมูลส่วนบุคคล (Personal Information) โดยข้อมูลความสนใจของผู้ใช้เป็นการจัดเก็บระดับความสนใจในแต่ละหมวดหมู่ของสถานที่ท่องเที่ยวและสถานที่ท่องเที่ยวที่ผู้ใช้สนใจ ในส่วนข้อมูลส่วนบุคคลจะเก็บข้อมูลที่แสดงถึงรายละเอียดของผู้ใช้ การจัดเก็บจะใช้รูปแบบฐานข้อมูลสัมพันธ์ (Relational Database) เพื่อความสะดวกในการจัดการข้อมูล ซึ่งมีประสิทธิภาพมากกว่าการจัดเก็บ

แบบ Flat File เช่น การจัดหมวดหมู่หรือการเก็บข้อมูลที่มีความสัมพันธ์กัน การค้นหาที่รวดเร็วกว่า เป็นต้น นอกจากนี้ยังมีความปลอดภัยในการจัดเก็บข้อมูลที่ดีกว่าอีกด้วย รวมถึงมีความเหมาะสมกว่าการจัดเก็บแบบออนโทโลจีในเรื่องของความง่ายในการจัดการข้อมูลแบบอัตโนมัติโดยระบบ โดยผู้วิจัยได้ออกแบบฐานข้อมูลที่ใช้จัดเก็บแบบจำลองความสนใจของผู้ใช้ดังภาพที่ 3

3.4 การแนะนำข้อมูลสถานที่ท่องเที่ยว

แบบจำลองความสนใจที่วิเคราะห์ได้ ไม่ว่าจะจากวิธี User-based Content Filtering หรือ Friend-based Collaborative Filtering จะใช้การเลือกสถานที่ท่องเที่ยวเพื่อแนะนำไปยังผู้ใช้ในลักษณะเดียวกัน โดยแบ่งออกเป็น 3 ขั้นตอน

3.4.1 การแบ่งสัดส่วนหมวดหมู่ในการแนะนำ จากการวิเคราะห์ค่าระดับความสนใจ พบว่ากรณีที่ผู้ใช้มีค่าระดับความสนใจในหมวดหมู่ไม่ต่างกันมากนัก การเลือกแนะนำเฉพาะสถานที่ในหมวดหมู่ที่มีความสนใจมากที่สุด จะทำให้ระบบไม่ได้แนะนำสถานที่ในอีกหมวดหมู่หนึ่งที่มีความสนใจมากเช่นกัน (แต่ไม่ได้มากที่สุด) ดังนั้นการใช้ค่าระดับความสนใจมากำหนดจำนวนสถานที่แนะนำในหมวดหมู่นั้นๆ จะสามารถแก้ไขปัญหาดังกล่าวได้ โดยหมวดหมู่ที่ผู้ใช้มีความสนใจสูง (มีค่า $I(c)$ สูง) ก็จะแนะนำในสัดส่วนที่มากกว่าหมวดหมู่ที่ผู้ใช้มีความสนใจน้อยกว่า กำหนดให้ $R(c)$ คือค่าสัดส่วนการแนะนำที่มีต่อแต่ละหมวดหมู่การท่องเที่ยว ซึ่งคำนวณดังสมการที่ 3

$$R(c) = \frac{I(c)}{\sum_{k=1}^{6} I(k)} \times N,\ 1 \le c \le 6 \qquad (3)$$

โดย $I(c)$ คือค่าระดับความสนใจที่คำนวณจากหัวข้อ 3.2 และ N คือจำนวนของสถานที่ที่ต้องการแสดงไปยังผู้ใช้ ในงานวิจัยนี้ต้องการแนะนำสถานที่ Top10 กำหนด $N = 10$

ค่าสัดส่วนการแนะนำ $R(c)$ นำมาใช้เลือกจำนวนของสถานที่ในแต่ละหมวดหมู่การท่องเที่ยว เพื่อแนะนำไปยังผู้ใช้ เช่นหมวดหมู่สถานที่ท่องเที่ยวทางวัฒนธรรมมีค่า $R(c) = 2$ หมายความว่า ระบบจะเลือกแสดงสถานที่ท่องเที่ยวทางวัฒนธรรม 2 สถานที่ เป็นต้น เมื่อได้ค่าสัดส่วนการแนะนำแต่ละหมวดหมู่แล้ว ก็จะคำนวณค่าน้ำหนักไปยังสถานที่ท่องเที่ยวที่อยู่ในแต่ละหมวดหมู่นั้น เพื่อเลือกสถานที่ท่องเที่ยวที่เหมาะสมในหัวข้อถัดไป

2.4.2 คำนวณค่าน้ำหนักของสถานที่ในแต่ละหมวดหมู่ เป็นการคำนวณค่าน้ำหนักของสถานที่ที่อยู่ในแต่ละหมวดหมู่จากฐานข้อมูลสถานที่ท่องเที่ยว โดยเลือกพิจารณาเฉพาะปลายทางที่ผู้ใช้ระบุ เช่น ปลายทางคือจังหวัดพิษณุโลก ก็จะนำสถานที่ที่อยู่ในจังหวัดพิษณุโลกมาคำนวณเท่านั้น ซึ่งในการเลือกสถานที่แนะนำจะเลือกสถานที่ที่มีค่าน้ำหนักที่สูงสุดเป็นลำดับไป โดยจากการศึกษาพบว่านักท่องเที่ยวมักจะเลือกไปยังสถานที่ที่มีความนิยม ซึ่งเป็นสถานที่ท่องเที่ยวที่นักท่องเที่ยวส่วนมากจะไป ดังนั้นจึงคำนวณความนิยมของสถานที่จากจำนวนการเช็คอิน และกดไลค์ในสถานที่นั้นดังสมการที่ 5 นอกจากนี้ในเก็บรวบรวมความต้องการของนักท่องเที่ยวผ่านแบบสอบถามออนไลน์ พบว่า 92% ผู้ใช้สนใจที่จะไปในสถานที่ท่องเที่ยวที่เพื่อนเคยไป ผู้วิจัยจึงได้พิจารณาตัวแปรสถานที่ที่เพื่อนเคยไปมาใช้ในการคำนวณ โดยใช้สถานที่ที่กลุ่มเพื่อนสนิท (5 คนที่มีปฏิสัมพันธ์สูงสุด) ในเฟสบุ๊คของผู้ใช้เคยไป ดังสมการที่ 6 ตัวแปรสุดท้ายคือช่วงเวลา ซึ่งหากระบุ

ช่วงเวลาที่น่าท่องเที่ยวของสถานที่นั้น จะช่วยให้ระบบแนะนำสถานที่ได้ถูกต้องมากขึ้น เช่น กรณีที่ผู้ใช้มีความสนใจในสถานที่ท่องเที่ยวที่เป็นธรรมชาติ ระบบควรจะแนะนำการเที่ยวภู ดอย มากกว่าน้ำตก เมื่อผู้ใช้จะไปในหน้าหนาว ดังนั้นระบบจึงให้ผู้ใช้ระบุช่วงเวลาที่วางแผนจะเดินทาง และนำมาใช้เป็นเกณฑ์ในการคำนวณน้ำหนักด้วย ซึ่งค่าช่วงเวลาที่เหมาะสมของสถานที่คำนวณดังสมการที่ 7

จากเหตุผลข้างต้น ผู้วิจัยจึงใช้ตัวแปรในการคำนวณค่าน้ำหนัก 3 ส่วน คือ 1) ความนิยมของสถานที่ 2) เป็นสถานที่ที่เพื่อนๆ เคยไป และ 3) เป็นช่วงเวลาที่นิยมไป โดยใช้ Linear Regression Analysis วิเคราะห์ความสัมพันธ์ระหว่างตัวแปร กำหนดให้ $W(p)$ คือค่าน้ำหนักของสถานที่ p และ และ $0<\alpha+\beta+\gamma<1$ คำนวณดังสมการที่ 4

$$W(p)=\left[\alpha P(p)+\beta F(p)+\gamma T(p)\right] \quad (4)$$

โดย $P(p)$ คือค่าความนิยมของสถานที่คำนวณจากจำนวนการเช็คอินและกดไลค์ในสถานที่นั้น ดังสมการที่ 5

$$P(p)=\frac{n_{ch,p}}{\max(n_{ch,p})}+\frac{n_{li,p}}{\max(n_{li,p})} \quad (5)$$

$F(p)$ คือค่าคะแนนของสถานที่ที่เพื่อนสนิทในเฟสบุ๊คของผู้ใช้เคยไป โดย C_i มีค่า =1 เมื่อมีสถานที่ปรากฏในการเช็คอินของเพื่อนคนที่ i และมีค่า = 0 เมื่อไม่ปรากฏ คำนวณดังสมการที่ 6

$$F(p)=\frac{\sum_{i=1}^{5}F_iC_i}{\sum_{i=1}^{5}F_i},\ 1\le n\le 5 \quad (6)$$

$T(p)$ คือค่าช่วงเวลาที่เหมาะสมของสถานที่ คำนวณดังสมการที่ 7

$$T(p)=\begin{Bmatrix}1 & if\ special\ time\\ 0 & otherwise\end{Bmatrix} \quad (7)$$

โดยในแต่ละหมวดหมู่ c นั้น ระบบจะเลือกสถานที่ที่มีค่าน้ำหนักของสถานที่ $W(p)$ สูงที่สุดตามลำดับเป็นจำนวนเท่ากับ $R(c)$ จากนั้นจัดลำดับสถานที่เป็นลำดับถัดไปดังนี้

2.4.3 **จัดลำดับสถานที่** เป็นการจัดลำดับ (Rank) สถานที่ที่จะแนะนำไปยังผู้ใช้ โดยลำดับตามความสนใจและมีความเหมาะสมกับผู้ใช้ (มีความนิยม เพื่อนเคยไปท่องเที่ยว และช่วงเวลาเหมาะสมกับที่จะวางแผนเดินทางไป) ซึ่งใช้ค่า $R(c)$ และ $W(p)$ ดังสมการที่ 8 โดยกำหนดให้ ถ้าค่า $Rank(p_1)$ ค่าสูงกว่า $Rank(p_2)$ แล้ว p_1 จะมีลำดับสูงกว่า p_2

$$Rank(p)=\frac{R(c)\times W(p)}{N} \quad (8)$$

โดย $R(c)$ คือค่าสัดส่วนการแนะนำของสถานที่จากสมการ 3 $W(p)$ คือค่าน้ำหนักของสถานที่จากสมการที่ 4 และ N คือจำนวนของสถานที่ที่ต้องการแสดงไปยังผู้ใช้

ผลการวิจัยและวิจารณ์ผล

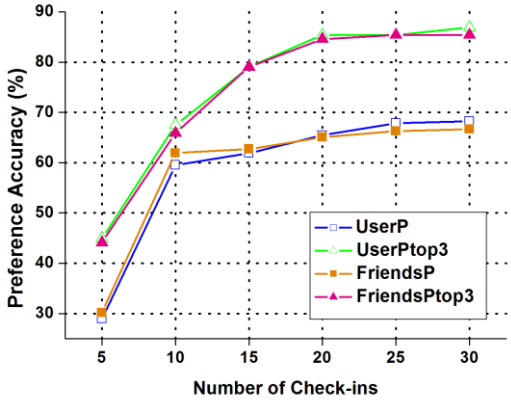
ผู้วิจัยได้พัฒนาระบบตามวิธีวิจัย และประเมินผลโดยใช้ข้อมูลของกลุ่มตัวอย่างที่นำเสนอไปแล้วข้างต้น โดยออกแบบการทดสอบเป็น 3 แบบ ดังนี้

1. **การทดสอบประสิทธิภาพในการแก้ไขปัญหาโคลด์สตาร์ท**

การทดสอบนี้เป็นการวัดประสิทธิภาพในการแก้ไขปัญหาโคลด์สตาร์ทสองระดับ โดยวัดประสิทธิภาพจากค่าความถูกต้อง (Accuracy) ของลำดับความสนใจในหมวดหมู่การท่องเที่ยวที่วิเคราะห์ได้จากระบบ เปรียบเทียบกับลำดับความสนใจจริงที่ได้จากผู้ใช้ โดยภาพที่ 4 (ซ้าย) แสดงวิธีการวัดความถูกต้องในการลำดับความสนใจของผู้ใช้ 1 คน ซึ่งจะพบว่าระบบมีความถูกต้องในการเรียงลำดับ 4 หมวดหมู่ จากทั้งหมด 6 หมวดหมู่ ได้ความถูกต้องของการเรียงลำดับ 66.67% และเมื่อพิจารณาลำดับความสนใจใน 3 อันดับแรก พบว่าเรียงลำดับได้ถูกต้องทั้งหมด ได้ความถูกต้องของการเรียงลำดับ 100%

จากนั้นได้ทดลองกับกลุ่มตัวอย่างจำนวน 42 บัญชีผู้ใช้ ที่มีข้อมูลการเช็คอินที่ตรวจสอบความถูกต้องแล้วมากกว่า 30 เช็คอิน เพื่อพิจารณาปริมาณของข้อมูลเช็คอินที่ส่งผลต่อความถูกต้องของการลำดับความสนใจ จากนั้นจำลองการเกิดปัญหาโคลด์สตาร์ททั้งสองระดับโดย 1.เริ่มต้นระบบไม่มีข้อมูลความสนใจ เกิดโคลด์สตาร์ทระดับ 1 และ 2.เริ่มต้นระบบไม่มีข้อมูลความสนใจของผู้ใช้และผู้ใช้มีเช็คอินไม่เพียงพอ เกิดโคลด์สตาร์ทระดับ 2 ในกรณีนี้ผู้วิจัยใช้การลบข้อมูลการเช็คอินของผู้ใช้ในเครือข่ายสังคมออกเพื่อให้มีข้อมูลไม่เพียงพอ ทำให้ระบบต้องใช้ข้อมูลของเพื่อนผู้ใช้ในการวิเคราะห์เพื่อสร้างลำดับความสนใจแทน จากนั้นเปรียบเทียบกับลำดับความสนใจแท้จริงที่ได้จากการสอบถามผู้ใช้ และลำดับความสนใจที่คำนวณจากระบบโดยใช้จำนวนเช็คอินล่าสุดที่ต่างกัน (กำหนดจำนวนเช็คอินล่าสุดเท่ากับ 5 10 15 20 25 30) โดยภาพที่ 4 (ขวา) แสดงค่าความถูกต้องในการลำดับความสนใจกรณีโคลด์สตาร์ทระดับ 1 (UserP) และ 2 (FriendsP) และแสดงค่าความถูกต้องในการลำดับความสนใจ 3 อันดับแรกในกรณีโคลด์สตาร์ทระดับ 1 (UserPtop3) และ 2 (FriendsPtop3)

ลำดับหมวดหมู่สถานที่ท่องเที่ยว	ลำดับความสนใจ		ความถูกต้อง
	ผู้ใช้จริง	วิเคราะห์โดยระบบ	
สถานที่ท่องเที่ยวทางวัฒนธรรม	3	3	✓
สถานที่ท่องเที่ยวทางประวัติศาสตร์	1	1	✓
สถานที่ท่องเที่ยวเพื่อความบันเทิง	5	4	✗
สถานที่ท่องเที่ยวทางธรรมชาติ	2	2	✓
สถานที่ท่องเที่ยวเพื่อการศึกษา	4	5	✗
สถานที่ท่องเที่ยวทางวิถีชีวิตชุมชน	6	6	✓



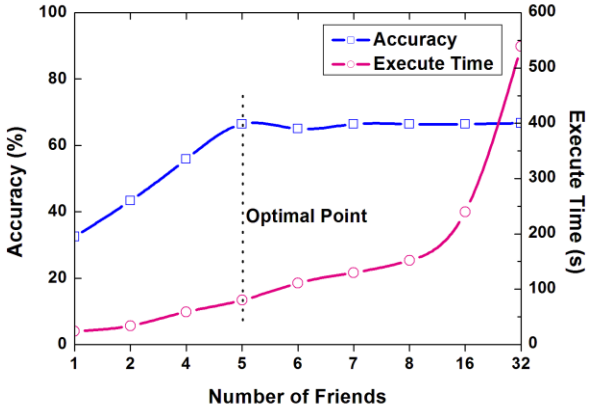
ภาพที่ 4 ความถูกต้องของลำดับหมวดหมู่ความสนใจของผู้ใช้ 1 คน (ซ้าย) และค่าเฉลี่ยความถูกต้องของลำดับหมวดหมู่ความสนใจ (ขวา)

จากการทดลองพบว่าทั้งในกรณีโคลด์สตาร์ทระดับ 1 และ 2 ค่าเฉลี่ยความถูกต้องของการลำดับความสนใจในหมวดหมู่การท่องเที่ยวมีแนวโน้มสูงขึ้นเมื่อมีจำนวนการเช็คอินที่เพิ่มขึ้น โดยโคลด์สตาร์ทระดับ 1 ซึ่งใช้ข้อมูลเช็คอินของผู้ใช้เอง (UserP) มีค่าเฉลี่ยความถูกต้องสูงสุดเท่ากับ 68.25% และมีค่าเฉลี่ยความถูกต้องในการลำดับความสนใจใน 3 อันดับแรก (UserPtop3) สูงสุดเท่ากับ 86.91% เมื่อใช้เช็คอินล่าสุดจำนวน 30 เช็คอิน และโคลด์สตาร์ทระดับ 2 ซึ่งใช้ข้อมูลเช็คอินจากกลุ่มเพื่อนสนิท (FriendsP) มีค่าเฉลี่ยความถูกต้องสูงสุดเท่ากับ 66.67% และมีค่าเฉลี่ยความถูกต้องในการลำดับ 3 อันดับแรก (FriendsPtop3) สูงสุดเท่ากับ 85.40% เมื่อใช้เช็คอินล่าสุดจำนวน 30 เช็คอิน นอกจากนี้ผู้วิจัยได้พิจารณาถึงการใช้จำนวนเช็คอินที่มีผลต่อความถูกต้องเป็นเกณฑ์ในการกำหนดความเพียงพอของข้อมูล โดยให้ความถูกต้องในการลำดับความสนใจที่ยอมรับได้เป็น 60% ซึ่งจากกราฟในภาพที่ 4 (ขวา) จะให้ค่าเฉลี่ยความถูกต้องน้อยกว่า 60% หากผู้ใช้ที่มีจำนวนการเช็คอินที่ตรวจสอบความถูกต้องแล้วน้อยกว่า 10 ในทั้งสองระดับ (UserP และ FriendsP) ดังนั้นผู้ใช้จะถือว่ามีข้อมูลไม่พอเพียงเมื่อมีจำนวนการเช็คอินน้อยกว่า 10 เช็คอิน

2. การทดสอบหาจำนวนของเพื่อนที่เหมาะสม ในกรณีใช้ข้อมูลจากเพื่อนในการหาความสนใจ

นอกจากนี้ในส่วนของการใช้ข้อมูลของเพื่อนเพื่อแก้ไขปัญหาโคลด์สตาร์ทระดับที่ 2 ผู้วิจัยได้ทดลองถึงการเลือกใช้จำนวนของเพื่อนที่เหมาะสมเพื่อนำมาวิเคราะห์ความสนใจแทนข้อมูลจากผู้ใช้ โดยพิจารณาจากมีความถูกต้องมากที่สุดในการลำดับหมวดหมู่ความสนใจสถานที่ท่องเที่ยว และเวลาใช้เวลาน้อยที่สุดในการประมวลผลเพื่อหาลำดับหมวดหมู่ความสนใจ

เนื่องจากไม่มีวิธีการอย่างเป็นทางการในการเลือกจำนวนเพื่อนที่เหมาะสมข้างต้น ผู้วิจัยจึงใช้วิธี grid search (Cao & 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 โดยการทดสอบของหนึ่งผู้ใช้ พบว่าลำดับหมวดหมู่ความสนใจสถานที่ท่องเที่ยวมีลำดับที่เหมือนกันเมื่อเลือกใช้ข้อมูลจากเพื่อนจำนวน 5 คน ขึ้นไป เมื่อพิจารณาเวลาที่ใช้ในการประมวลผล จะเห็นว่าเมื่อใช้เพื่อนจำนวนมากขึ้นก็จะใช้เวลาในการประมวลผลนานขึ้น โดยผู้วิจัยได้ทดสอบในผู้ใช้ 42 ตัวอย่าง ได้ค่าเฉลี่ยความถูกต้องของลำดับหมวดหมู่ความสนใจและเวลาเฉลี่ยในการประมวลผลดังภาพที่ 5 จะเห็นว่าการเลือกใช้เพื่อน 5 คนมีความเหมาะสมที่สุดเนื่องจากมีค่าความถูกต้องมากที่สุดและใช้เวลาในการประมวลผลน้อยที่สุด โดยพบว่าปัจจัยที่ส่งผลต่อความถูกต้องคือค่าปฏิสัมพันธ์ระหว่างผู้ใช้และเพื่อน เนื่องจากในจำนวนเพื่อนที่สูงขึ้น ค่าปฏิสัมพันธ์ F_i ของเพื่อนที่เพิ่มเข้ามา ทำให้การคำนวณค่าระดับความสนใจ $I(c)$ ถูกเฉลี่ยไปยังเพื่อนหลายคนตามค่าปฏิสัมพันธ์ แตกต่างจากเลือกใช้เพื่อนในจำนวนที่ต่ำอย่างเช่นการเลือกใช้เพื่อน 1 คนค่าระดับความสนใจจะขึ้นอยู่กับเพื่อนเพียงคนเดียว ซึ่งอธิบายได้ว่าผู้ใช้และเพื่อนสนิทที่สุดที่เลือกมาเพียง 1 คนยังคงมีความแตกต่างในความสนใจ ทำให้ได้ค่าความถูกต้องที่ไม่สูงนัก อย่างไรก็ตามเมื่อเลือกใช้กลุ่มเพื่อนที่สนิทที่สุด (มากกว่า 1 คน) กลับได้ค่าความถูกต้องที่เพิ่มขึ้น ดังนั้นกลุ่มเพื่อนสนิทบนเครือข่ายสังคมจึงเป็นตัวแปรสำคัญที่มีผลต่อความสนใจของผู้ใช้มากกว่าเพื่อนสนิทเพียงคนเดียว



ภาพที่ 5 ค่าเฉลี่ยความถูกต้องของลำดับหมวดหมู่ความสนใจและเวลาเฉลี่ยในการประมวลผล

3. การทดสอบประสิทธิภาพการแนะนำสถานที่ท่องเที่ยวไปยังผู้ใช้

การทดสอบนี้จะวัดประสิทธิภาพของระบบว่าสามารถแนะนำสถานที่ท่องเที่ยวได้ตรงกับที่ผู้ใช้สนใจหรือค้นหาอยู่ โดยใช้ข้อมูลสถานที่ท่องเที่ยวที่ระบบแนะนำไปยังผู้ใช้จำนวน 10 รายการ (Top-10 Recommendation) เปรียบเทียบกับสถานที่ที่ผู้ใช้มีความสนใจจริง ซึ่งจะใช้วิธีการที่นิยมในระบบแนะนำข้อมูลอย่าง Precision@10 คำนวณได้จากอัตราส่วนของสถานที่ที่ระบบแนะนำและผู้ใช้มีความสนใจต่อสถานที่ที่ระบบแนะนำทั้งหมด ดังสมการที่ 9 นอกจากนี้ยังใช้ Rankscore_{10} ในสมการที่ 10 ซึ่งวัดประสิทธิภาพในการลำดับสถานที่แนะนำ โดยกำหนด half life (α) เท่ากับ 5 ซึ่งจะให้ค่าสูงขึ้นหากผู้ใช้สนใจสถานที่แนะนำในลำดับต้นๆ

$$precision = \frac{\text{interest} \cap \text{recommended}}{\text{recommended}} \times 100 \qquad (9)$$

$$rankscore = \frac{\sum_{i \in \text{interest} \cap \text{recommended}} \frac{1}{2^{\frac{rank(i)-1}{\alpha}}}}{\sum_{i \in \text{interest}} \frac{1}{2^{\frac{idx(i)-1}{\alpha}}}} \qquad (10)$$

การทดสอบใช้ตัวอย่างจำนวน 42 ผู้ใช้ ทั้งในกรณีโคล์ดสตาร์ทระดับที่ 1 ซึ่งแนะนำจากความสนใจของผู้ใช้เอง (User-based) และระดับที่ 2 (Friends-based) ซึ่งแนะนำจากความสนใจของกลุ่มเพื่อนสนิท โดยเปรียบเทียบกับการแนะนำในหลายลักษณะ ซึ่งใช้เกณฑ์ในการคำนวณค่าน้ำหนักของสถานที่ในสมการที่ 4 แตกต่างกันไป 5 แบบ คือ

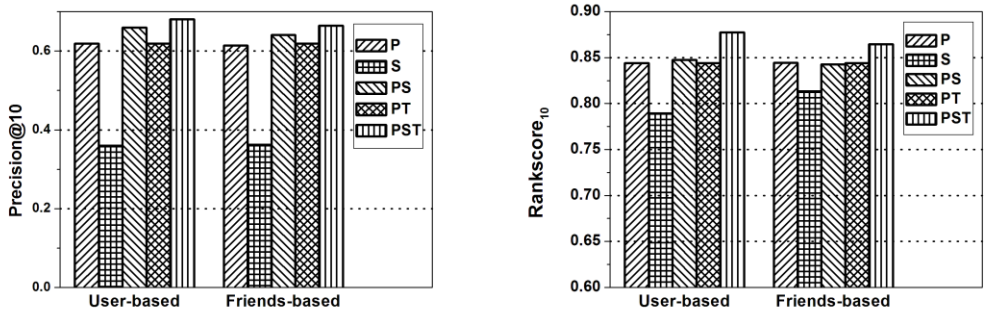
1) Popularity Weight (P) ใช้ค่าความนิยมในสถานที่เพียงอย่างเดียว โดยกำหนดค่า $\beta = 0$ และ $\gamma = 0$

2) Social Weight (S) พิจารณาจากสถานที่ที่เพื่อนสนิทในเฟสบุ๊คของผู้ใช้เคยไป โดยกำหนดค่า $\alpha = 0$ และ $\gamma = 0$

3) Popularity/Social Weight (PS) พิจารณา Popularity Weight (P) ร่วมกับ Social Weight (S)

4) Popularity/Time Weight (PT) พิจารณา Popularity Weight (P) ร่วมกับ ค่าช่วงเวลาที่เหมาะสม (Time -T)

5) Popularity/Social/Time Weight (PST) พิจารณาทั้ง 3 ค่าร่วมกัน คือ ค่าความนิยม สถานที่ที่เพื่อนสนิทในเฟสบุ๊คของผู้ใช้เคยไป และค่าช่วงเวลาที่เหมาะสม (Time -T) ในการไปเยี่ยมชมสถานที่

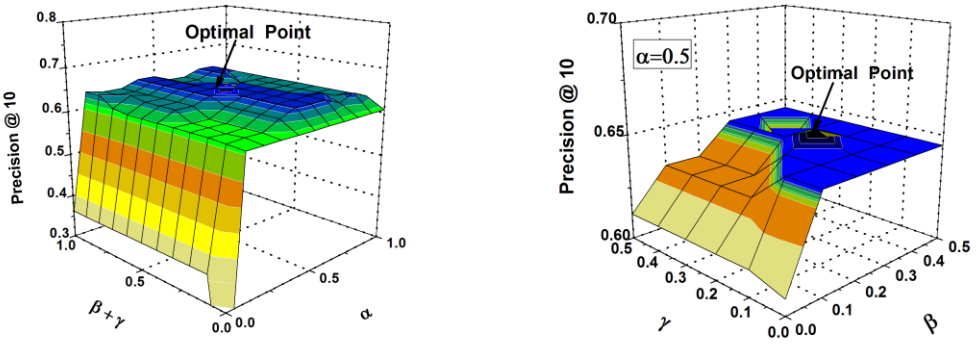


ภาพที่ 6 precision@10 และ Rankscore10 ของการแนะนำเปรียบเทียบระหว่างการใช้เกณฑ์ที่ต่างกัน

ผลการทดสอบพบว่าการแนะนำทั้งแบบ User-based และ Friends-based การใช้เกณฑ์ในการคำนวณค่าน้ำหนักของสถานที่แบบ PST จะมีค่า Precision โดยเฉลี่ยสูงที่สุด 0.6810 และ 0.6643 ตามลำดับ เช่นเดียวกับค่า Rankscore การใช้เกณฑ์แบบ PST จะให้ค่าเฉลี่ยสูงสุด โดย User-based มีค่าเฉลี่ย Rankscore เท่ากับ 0.8775 และแบบ Friends-based เท่ากับ 0.8647 ดังกราฟในภาพที่ 6 ดังนั้นสถานที่เพื่อนเคยไปและค่าของเวลาที่เหมาะสมในการไปเยี่ยมชมสถานที่ จึงมีผลต่อการเลือกสถานที่ที่ผู้ใช้สนใจหรือค้นหาอยู่ ซึ่งสามารถเพิ่มประสิทธิภาพการแนะนำสถานที่ท่องเที่ยวให้ดีขึ้นได้ และจากค่า Rankscore ที่ได้ ทำให้ทราบว่าผู้ใช้มีความสนใจในสถานที่ในลำดับต้นๆ ซึ่งสอดคล้องกับค่าความถูกต้องของการลำดับความสนใจใน 3 อันดับแรกที่ให้ค่าที่สูงกว่าความถูกต้องในการลำดับความสนใจในหมวดหมู่ทั้งหมด อย่างไรก็ตามพบว่าในบางผู้ใช้การใช้ค่าของเวลาที่เหมาะสมเพิ่มเข้ามาในทั้งแบบ PT และ PTS ทำให้ค่า Precision และ Rankscore ต่ำลง เนื่องมาจากผู้ใช้ระบุช่วงเวลาไม่ตรงกับที่ผู้ใช้สนใจจริง เช่น ระบุช่วงเวลาหน้าร้อนให้กับระบบ แต่ช่วงเวลาจริงที่จะไปเป็นหน้าหนาว ทำให้ระบบแนะนำน้ำตกแทนที่จะเป็นภูเขา นอกจากนี้ในบางตัวอย่างก็ให้ค่า Precision และ Rankscore เท่ากันในการพิจารณาเกณฑ์ทั้งห้าแบบ เนื่องมาจากเพื่อนสนิทในเฟสบุ๊คมีข้อมูลการเยี่ยมชมสถานที่น้อย ไม่ส่งผลต่อการเปลี่ยนแปลงคำแนะนำ รวมถึงผู้ใช้มีความสนใจในสถานที่ไม่มีช่วงเวลาพิเศษ เช่น โบราณสถานหรือวัดบางแห่ง ค่าช่วงเวลาจึงไม่ส่งผลต่อการเปลี่ยนแปลงคำแนะนำเช่นกัน

4. การทดสอบหาค่าตัวแปรที่เหมาะสม ในการคำนวณค่าน้ำหนักของสถานที่

การทดสอบนี้เป็นการหาค่า α β และ γ ที่เหมาะสมในการแนะนำสถานที่โดยใช้เกณฑ์ 5 แบบข้างต้น (สมการที่ 4) พิจารณาจากค่าเฉลี่ย Precision สูงสุดที่ได้รับเมื่อใช้ค่า α β และ γ แตกต่างกัน กำหนดให้ $0<\alpha+\beta+\gamma<1$ โดยทดลองเพิ่มค่าทีละ 0.1



ภาพที่ 7 ค่า α β และ γ ที่เหมาะสมในการแนะนำแบบ PST

จากกลุ่มตัวอย่าง 42 ผู้ใช้ พบว่าการแนะนำแบบ PST ซึ่งมีค่า Precision สูงที่สุดในทั้งห้าแบบนั้น ได้ค่า α β และ γ ที่เหมาะสมคือ 0.5 0.3 และ 0.2 ตามลำดับ ดังภาพที่ 7 ซึ่งจะทำให้ได้ค่า Precision มากที่สุดในการใช้ค่า α β และ γ แบบอื่น ส่วนการแนะนำโดยใช้เกณฑ์แบบอื่น ได้ค่า α β และ γ ที่เหมาะสม แสดงดังตารางที่ 2

ตารางที่ 2 ค่า α β และ γ ที่เหมาะสมในการแนะนำโดยใช้เกณฑ์แต่ละแบบ

	α	β	γ
Popularity Weight (P)	(0,1]	-	-
Social Weight (S)	-	(0,1]	-
Popularity/Social Weight (PS)	0.6	0.4	-
Popularity/Time Weight (PT)	0.8	-	0.2
Popularity/Social/Time Weight (PST)	0.5	0.3	0.2

สรุปผลการวิจัย

บทความวิจัยนี้ได้ออกแบบและพัฒนาระบบแนะนำข้อมูลการท่องเที่ยวเฉพาะบุคคลบนพื้นฐานการใช้งานเว็บไซต์โดยมีแนวคิดในการนำข้อมูลจากบริการเครือข่ายสังคมที่เป็นที่นิยมอย่างเฟสบุ๊ค มาแก้ปัญหาโคลด์สตาร์ทที่มักเกิดขึ้นกับผู้ใช้ที่เริ่มต้นใช้งานระบบ ซึ่งระบบได้ใช้ทั้งข้อมูลของผู้ใช้เองและข้อมูลจากกลุ่มเพื่อนที่มีปฏิสัมพันธ์กับผู้ใช้ในระดับสูง มาช่วยวิเคราะห์หาหมวดหมู่ความสนใจและคำนวณค่าน้ำหนักเพื่อเลือกสถานที่ท่องเที่ยวและแนะนำไปยังผู้ใช้แบบเฉพาะบุคคล ซึ่งการทดสอบประสิทธิภาพนั้นยังใช้กลุ่มตัวอย่างจำนวนไม่มาก แต่ผลการวิจัยก็สามารถแสดงให้เห็นว่าข้อมูลการเช็คอินบนเฟสบุ๊คสามารถนำมาใช้ประโยชน์ในระบบแนะนำข้อมูลการท่องเที่ยวเฉพาะบุคคลได้ จากทั้งการใช้ข้อมูลจากผู้ใช้และกลุ่มเพื่อนที่มีปฏิสัมพันธ์ที่สูงของผู้ใช้มาวิเคราะห์ความสนใจ รวมถึงการใช้ข้อมูลสถานที่ที่เพื่อนเคยไปบนเฟสบุ๊คก็ช่วยเพิ่มประสิทธิภาพการแนะนำได้ อย่างไรก็ตามงานที่จะพัฒนาในขั้นต่อไปคือ การวิเคราะห์ความสนใจจากเครือข่ายสังคมโดยแบ่งพิจารณาเป็นช่วงเวลา เช่น ความสนใจสถานที่ท่องเที่ยวในฤดูที่แตกต่างกัน การนำข้อมูลตอบกลับของผู้ใช้ (User Feedback) มาช่วยปรับปรุงแบบจำลองความสนใจ และเพิ่มประสิทธิภาพของการแนะนำสถานที่ท่องเที่ยวให้ดียิ่งขึ้น

มหาวิทยาลัยบูรพา Burapha University

รวมถึงการทดสอบกับกลุ่มตัวอย่างที่มากขึ้น และครอบคลุมเพื่อวิเคราะห์ปัจจัย เช่น เพศ อายุ การศึกษา เป็นต้น ที่อาจมีผลต่อการแนะนำ

เอกสารอ้างอิง

Abel, F., Gao, Q., Houben, G.-J., & Tao, K. (2011). Analyzing User Modeling on Twitter for Personalized News Recommendations. In J. A. Konstan, R. Conejo, J. L. Marzo, & N. Oliver (Eds.), *User Modeling, Adaption and Personalization* (pp. 1–12). Heidelberg: Springer-Verlag.

Burke, R. (2000). Knowledge-Based Recommender Systems. In *Encyclopedia of Library and Information Sciences* (p. 2000). New York: Marcel Dekker.

Cao, L. J., & Tay, F. E. H. (2003). Support vector machine with adaptive parameters in financial time series forecasting. *IEEE Transactions on Neural Networks*, *14*(6), 1506–1518.

Castillejo, E., Almeida, A., & López-de-Ipiña, D. (2012). Social Network Analysis Applied to Recommendation Systems: Alleviating the Cold-User Problem. In J. Bravo, D. López-de-Ipiña, & F. Moya (Eds.), *Ubiquitous Computing and Ambient Intelligence* (pp. 306–313). Heidelberg: Springer-Verlag.

Chin, D. N., & Porage, A. (2001). Acquiring User Preferences for Product Customization. In M. Bauer, P. J. Gmytrasiewicz, & J. Vassileva (Eds.), *User Modeling 2001* (pp. 95–104). London: Springer-Verlag.

Coyle, L., & Cunningham, P. (2003). Exploiting Re-ranking Information in a Case-Based Personal Travel Assistent. In *In Workshop on Mixed-Initiative Case-Based Reasoning at the 5th International Conference on Case-Based Reasoning*. (pp. 11-20). n.p.

Firan, C. S., Nejdl, W., & Paiu, R. (2007). The Benefit of Using Tag-Based Profiles. In *Proceedings of the 2007 Latin American Web Conference* (pp. 32–41). Washington, DC, USA: IEEE Computer Society.

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

García-Crespo, A., Chamizo, J., Rivera, I., Mencke, M., Colomo-Palacios, R., & Gómez-Berbís, J. M. (2009). SPETA: Social pervasive e-Tourism advisor. *Telematics and Informatics*, *26*(3), 306–315.

Hang, Y., Guiran, C., & Xingwei, W. (2009). A Cold-Start Recommendation Algorithm Based on New User's Implicit Information and Multi-attribute Rating Matrix. In *Ninth International Conference on Hybrid Intelligent Systems.* (pp. 353–358). USA:IEEE.

Harfield, A. (2014). Tourism Database for lower north of Thailand. Retrieved January 17, 2014, from http://www.mobcomlab.com

Kabassi, K. (2010). Personalizing recommendations for tourists. *Telematics and Informatics*, *27*(1), 51–66.

Lorenzi, F., Fontanella, B., Prestes, E., & Peres, A. (2014). How to Improve Multi-Agent Recommendations Using Data from Social Networks? In *Proceedings of the Twenty-Seventh International Florida Artificial Intelligence Research Society Conference*. (pp. 63–68). Florida, USA: AAAI.

มหาวิทยาลัยบูรพา Burapha University

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.

Rashid, A. M., Karypis, G., & Riedl, J. (2008). Learning Preferences of New Users in Recommender Systems: An Information Theoretic Approach. *ACM SIGKDD Explorations Newsletter*, *10*(2), 90–100.

Schiaffino, S., & Amandi, A. (2009). Building an expert travel agent as a software agent. *Expert Systems with Applications*, *36*(2, Part 1), 1291–1299.

Sedhain, S., Sanner, S., Braziunas, D., Xie, L., & Christensen, J. (2014). Social Collaborative Filtering for Cold-start Recommendations. In *Proceedings of the 8th ACM Conference on Recommender Systems* (pp. 345–348). New York, USA: ACM.

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

Ye, M., Yin, P., Lee, W.-C., & Lee, D.-L. (2011). Exploiting Geographical Influence for Collaborative Point-of-interest Recommendation. In *Proceedings of the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 325–334). New York, USA: ACM.

Zhu, Y., He, L., & Wang, X. (2012). User Interest Modeling and Self-Adaptive Update Using Relevance Feedback Technology. *Procedia Engineering*, *29,* 721–725.

มหาวิทยาลัยบูรพา Burapha University