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

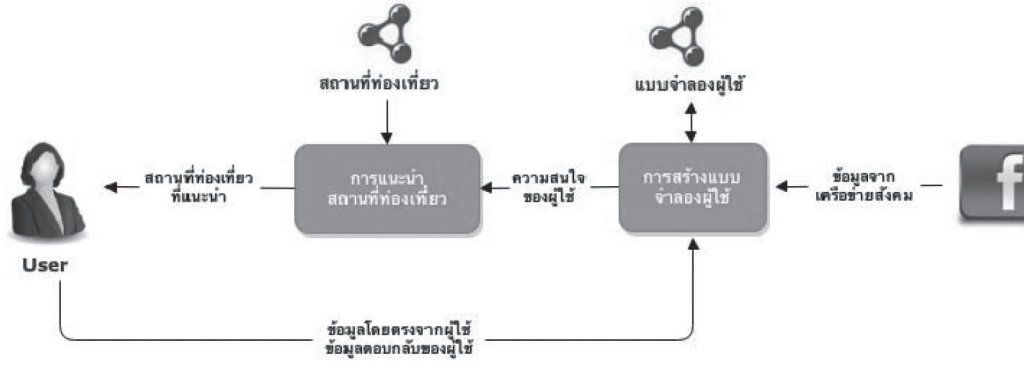
ในระบบแนะนำข้อมูลเฉพาะบุคคลนั้น แบบจำลองผู้ใช้ (User Model) เป็นส่วนประกอบสำคัญ [8] ที่จะต้องใช้ข้อมูลสำคัญจำนวนมากในการสกัดเพื่อหาความสนใจของผู้ใช้ ซึ่งเป็นหัวใจสำคัญในการแนะนำผลิตภัณฑ์หรือบริการให้ได้ตรงกับความต้องการของผู้ใช้อย่างแท้จริง อย่างไรก็ตาม การสร้างแบบจำลองผู้ใช้มีความท้าทายอย่างยิ่งสำหรับผู้พัฒนาระบบคือปัญหาโคลด์สตาร์ท (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

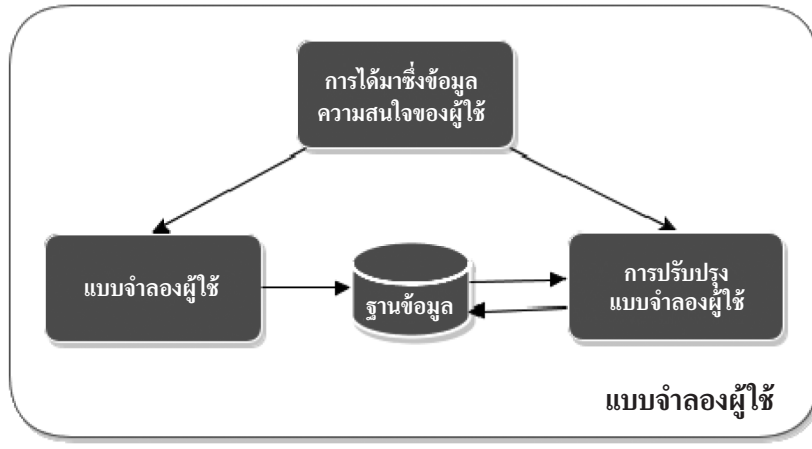
ระบบแนะนำข้อมูลการท่องเที่ยวเฉพาะบุคคลเป็นระบบที่นำเสนอข้อมูลการท่องเที่ยวในลักษณะต่างๆ เช่น ข้อมูลสถานที่ท่องเที่ยว แผนการเดินทาง ข้อมูลแพคเกจท่องเที่ยว เป็นต้น ให้กับนักท่องเที่ยวตามความสนใจที่แตกต่างกันไปในแต่ละบุคคล โดยระบบแนะนำข้อมูลการท่องเที่ยวเฉพาะบุคคลที่ผู้วิจัยกำลังพัฒนาเป็นการแนะนำข้อมูลสถานที่ท่องเที่ยวให้กับนักท่องเที่ยว 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

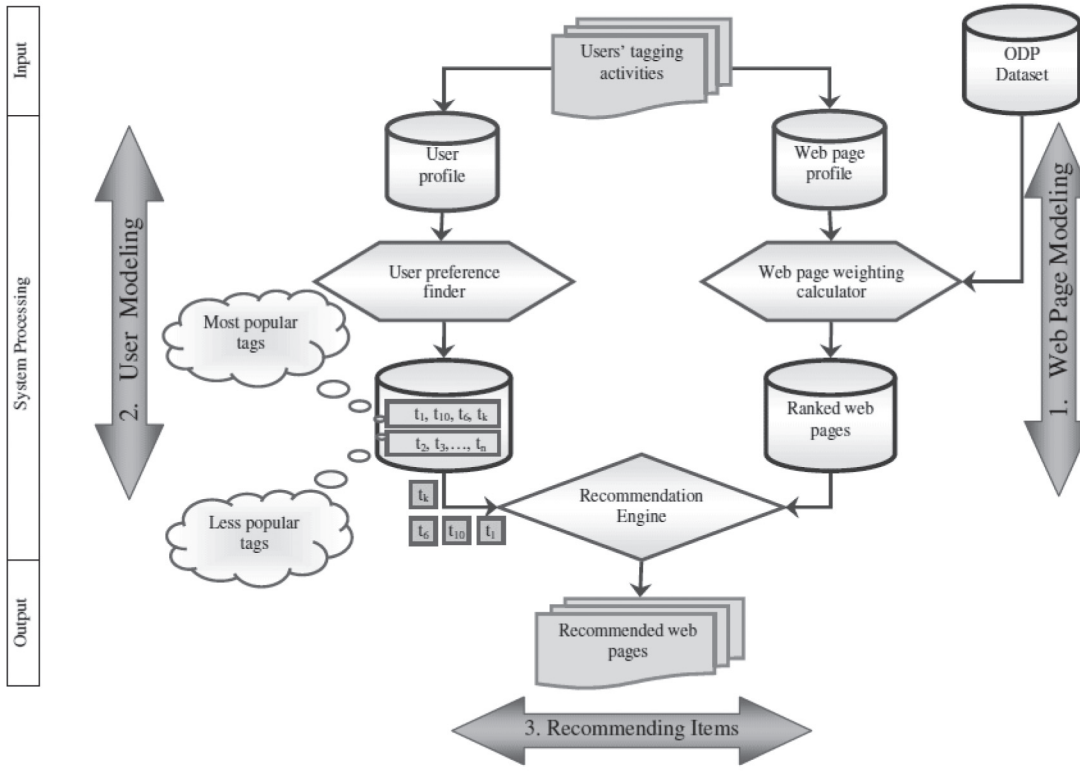
5. เทคนิคการพัฒนาแบบจำลองผู้ใช้โดยใช้เครือข่ายสังคม

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

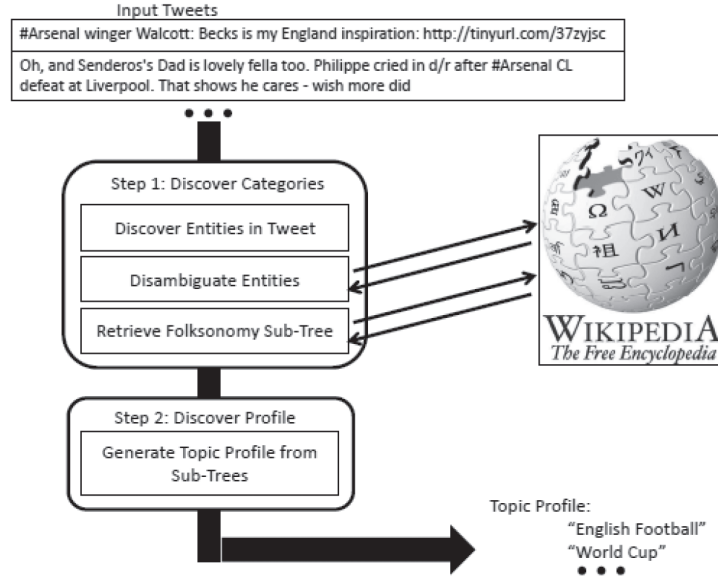


รูปที่ 3 ระบบแนะนำข้อมูลเว็บเพจโดยใช้การสกัดความสนใจจากแท็กในโซเชียลบุ๊กมาร์ก [28]

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]

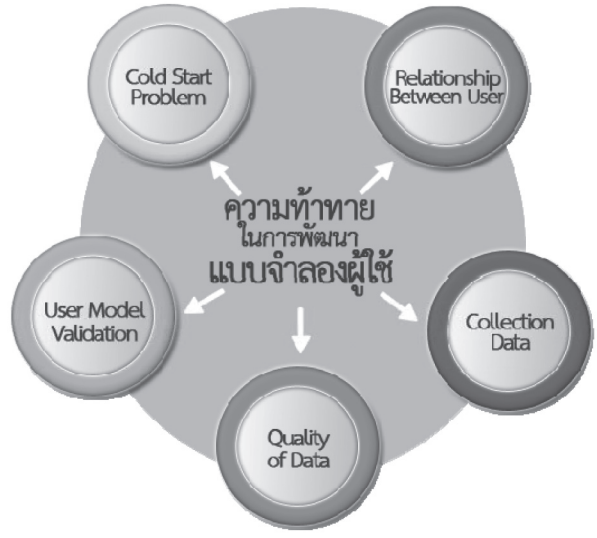
5.3 การพัฒนาแบบจำลองผู้ใช้โดยใช้การผสมผสานระหว่างเครือข่ายสังคมหลายแหล่ง

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