

# การวิเคราะห์เครือข่ายสังคมออนไลน์: การสำรวจแนวทางการวิจัย ในวิทยาการคอมพิวเตอร์

สิทธิพงษ์ เหล่าไก่

ภาคการศึกษาที่ 2 ปีการศึกษา 2563

ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์

มหาวิทยาลัยศรีนครินทรวิโรฒ

sitdhibong.laokok@g.swu.ac.th

**บทคัดย่อ—บทคัดย่อ**

**คำสำคัญ**—เครือข่ายสังคมออนไลน์, การสำรวจ, การวิจัยด้านคอมพิวเตอร์, การเรียนรู้ของเครื่อง, ระบบ

## I. บทนำ

อินเทอร์เน็ตนั้นพัฒนาตัวเองมาขึ้นเรื่อยๆ มากกว่าที่เราคาดถึง ช่วยเชื่อมต่อเครื่องคอมพิวเตอร์ทั่วทุกมุมโลก ต่างเชื้อชาติ และวัฒนธรรมเข้าด้วยกัน มากไปกว่านั้นอินเทอร์เน็ตได้ช่วยสร้างกลุ่มสังคมประหนึ่งโลกเสมือนจริงอีกใบหนึ่งขึ้นมา ด้วยการเชื่อมโยงที่เรียกว่าเครือข่ายสังคมออนไลน์ ผ่านแอปพลิเคชันของผู้ให้บริการที่พัฒนาขึ้นมาโดยเฉพาะ เช่น *Twitter*, *Google+*, และ *Facebook* ที่มีผู้ใช้งานรวมกันกว่าหลายร้อยล้านคนเปรียบได้กับประเทศขนาดย่อยบนโลกออนไลน์ก็ได้ ซึ่งจากการเก็บข้อมูลของผู้ให้บริการ พบว่าเวลาการใช้งานระบบเหล่านี้เพิ่มขึ้นเรื่อย เมื่อมีผู้ใช้งานที่หลากหลายทั้งด้วยพฤติกรรมการใช้งาน วัฒนธรรม การศึกษา รวมไปถึงคุณภาพการดำรงชีวิตมาประกอบเข้าด้วยกันแล้วนั้น หากเราสามารถทำความเข้าใจบริบทที่หลากหลายเหล่านี้ได้ทั้งหมด ตลอดจนเข้าใจผลกระทบหรือการเปลี่ยนแปลงทั้งหมดเหล่านี้ได้ ย่อมเป็นเรื่องที่ทำหายเป็นอย่างมาก ซึ่งในรายงานนี้สรุปความมาจากเอกสารวิชาการชื่อว่า *Online Social Network Analysis: A Survey of Research Applications in Computer Science* [?] ได้สรุปรวมแนวทางที่น่าสนใจ และความท้าทายด้านต่างๆ ที่น่าสนใจสำหรับงานวิจัยรวมเข้าไว้ด้วยกัน พร้อมทั้งให้ตัวอย่างงานวิจัยที่สนใจประกอบไปด้วยด้วยกัน ดังนี้

## II. หัวข้อในการศึกษาเกี่ยวกับเครือข่ายสังคมออนไลน์

ในหัวข้อนี้เป็นการอธิบายถึงความสำคัญของการศึกษาเครือข่ายสังคมออนไลน์ และทำไมการศึกษาในหัวข้อนี้ถึงได้สำคัญมากขึ้นเรื่อยๆ ข้อมูลที่น่าสนใจ รวมไปถึงเครื่องมือที่น่าสนใจที่มักจะใช้ในการศึกษาวิจัยเพื่อวิเคราะห์ข้อมูลเครือข่ายสังคมออนไลน์

### A. ทำไมเราต้องศึกษาเครือข่ายสังคมออนไลน์

เพื่อเข้าใจถึงปัญหาและความสำคัญของการศึกษานั้น ได้จัดกลุ่มของความน่าสนใจไว้ดังนี้

- **ข้อมูลที่ปรากฏอยู่ในเครือข่าย** จากที่ได้กล่าวมาแล้วในข้างต้นว่า เครือข่ายสังคมออนไลน์ นั้นมีแนวโน้มการใช้งานเพิ่มมากขึ้นเรื่อยๆ อย่างเช่น *Twitter* นั้นมีแนวโน้มมากขึ้นเรื่อยๆ จากข้อมูลพบว่าในปี 2011 นั้นมีข้อมูลทวีตรายวันอยู่ที่ 50 ล้านทวีต และเพิ่มขึ้นมาเป็น 500 ล้านทวีตต่อวัน ในปี 2013 [?] ถือเป็นเรื่องท้าทายส่วนหนึ่งในการจัดการข้อมูลที่มีขนาดใหญ่เช่นนี้ และเหมาะแก่การศึกษาในหัวข้อ ข้อมูลมหัต (Big Data)
- **ผู้ใช้งานที่หลากหลาย** จากที่กล่าวว่าผู้ใช้งานที่หลากหลาย ซึ่งเป็นลักษณะเฉพาะของเครือข่ายสังคมออนไลน์ รวมถึงหลายผู้ให้บริการ ดังนั้นข้อมูลที่เข้ามาในระบบนั้นย่อมมีมากขึ้นเป็นเงาตามตัว ซึ่งด้วยลักษณะของผู้ใช้งานนั้นมีความหลากหลาย จึงส่งผลให้ข้อมูลนั้นย่อมหลากหลายตามไปด้วย ดังนั้นเครือข่ายสังคมออนไลน์ จึงถือเป็นแหล่งข้อมูลที่ดีมากสำหรับข้อความ

เพื่อใช้ในงานประมวลผลภาษาธรรมชาติ (Natural Language Processing – NLP)

- **การปฏิสัมพันธ์ในนามของผู้ใช้งาน** ผู้ใช้งานนั้นมีสิทธิใช้งานกำหนดลักษณะการใช้งานเครือข่ายสังคมออนไลน์ได้อย่างเต็มที่ และสามารถเลือกได้ว่าจะสร้างความสัมพันธ์กับผู้ใช้รายอื่นอย่างไรบ้าง ซึ่งรูปแบบความสัมพันธ์ที่เกิดขึ้นนั้นเป็นประเด็นที่น่าสนใจในการศึกษาวิจัยได้
- **การเปลี่ยนแปลงชั่วคราว** การใช้งานเครือข่ายสังคมออนไลน์ นั้นผู้ใช้งานย่อมมีความสนใจที่หลากหลาย และเปลี่ยนไปได้เรื่อยๆ เสมอ ซึ่งการเปลี่ยนแปลงนี้เป็นเรื่องที่น่าสนใจในการวิจัย เพื่อหาแนวโน้มของการเปลี่ยนแปลง รวมถึงหัวข้อที่น่าจะเกิดขึ้นต่อไป
- **รวดเร็ว** ในเครือข่ายสังคมออนไลน์นั้น ตอบสนองต่อเหตุการณ์อย่างรวดเร็วทั้งภายในและภายนอกระบบ
- **ใช้งานได้ทั่วไป** เมื่อเทคโนโลยีนั้นพัฒนาขึ้นไปเรื่อยๆ ส่งผลให้ผู้ใช้งานสามารถใช้งานได้เกือบทุกที่ทุกเวลา ส่งผลให้ข้อมูลนั้นสร้างได้ทุกเมื่อ สิ่งก็ตามมานั้นคือข้อมูลเชิงภูมิศาสตร์จะปรากฏเข้ามาด้วยในเป็นอีกฟีเจอร์หนึ่งในข้อมูลเครือข่ายสังคมออนไลน์ ซึ่งเป็นอีกแนวทางหนึ่งที่น่าสนใจเพื่อวิเคราะห์

### B. เครือข่ายที่ใช้สำหรับข้อมูล

วิเคราะห์จาก 2 หัวข้อด้วยกัน นั่นคือ ปริมาณการใช้งานในเครือข่าย ว่ามีผู้ใช้งานมากน้อยขนาดไหน และความง่ายในการเข้าถึง ซึ่ง ณ เวลาที่ดำเนินการวิจัยนั้นพบว่าเครือข่ายสังคมออนไลน์ที่มีขนาดใหญ่ที่สุดนั้นคือ *Facebook* ด้วยจำนวนผู้ใช้งานกว่าพันล้านราย ทั้งนี้แบ่งกลุ่มเครือข่ายที่น่าสนใจดังนี้

- *Twitter* เป็นเครือข่ายที่น่าสนใจที่ใช้งานโดยการสื่อสารกันด้วยข้อความขนาดสั้นๆ มากกว่านั้นยังเตรียมช่องทางการเชื่อมต่อให้กับนักพัฒนาไว้อย่างมาก ซึ่งข้อดีของเครือข่ายนี้นอกจากการเข้าถึงได้ง่ายแล้ว ยังเป็นแหล่งข้อมูลที่หลากหลาย และเปลี่ยนแปลงอย่างรวดเร็วอีกด้วย
- *Facebook* เครือข่ายขนาดใหญ่ที่เหมาะกับกรวิเคราะห์ด้านความสัมพันธ์กันของบุคคลที่มีความหลากหลายทางด้าน เชื้อชาติและวัฒนธรรม
- *YouTube* แหล่งรวมข้อมูลสื่อที่สามารถนำข้อมูลเหล่านั้นมาวิเคราะห์ถึงแนวโน้มความสนใจของผู้ใช้งาน ในเวลาต่างๆ กันได้
- *Flickr* เครือข่ายในลักษณะเดียวกับ *YouTube* ที่รวมเอาสื่อด้านรูปภาพไว้ด้วยกัน สร้างกลุ่มความสัมพันธ์พิเศษของบุคคลที่มีทักษะและความสนใจในแนวเดียวกันไว้

### C. เครื่องมือสำหรับช่วยวิเคราะห์

ข้อมูลเครือข่ายในเครือข่ายสังคมออนไลน์นั้น นั้นมีความซับซ้อนมาก และการทำงานของเครือข่ายสังคมออนไลน์ส่วนใหญ่นั้น จะวิเคราะห์ข้อมูลทั้งหมดบนพื้นฐานของกราฟ (Graph-based) ซึ่งเครื่องมือที่มักจะใช้งานเพื่อวิเคราะห์ข้อมูลอยู่แล้วหลายรายการด้วยกัน โดยแบ่งเป็นกลุ่มๆ ได้ดังนี้

- **ฐานข้อมูลกราฟ (Graph database):** สำหรับจัดเก็บข้อมูลของเครือข่ายที่จัดเก็บไว้เพื่อวิเคราะห์ข้อมูล โดยฐานข้อมูลกราฟนี้จะมีคุณสมบัติ

พิเศษมากกว่าฐานข้อมูลเชิงสัมพันธ์ (Relational Database) ทั่วไป เพราะสามารถจัดเก็บคุณสมบัติต่างๆ ของกราฟได้เป็นอย่างดี โดยมีตัวอย่างที่มักใช้งานกันอยู่นั้นคือ AllegroGraph และ Neo4j

- **วาดกราฟ (Graph drawing):** แสดงลักษณะเฉพาะพิเศษบางอย่างเพื่อวิเคราะห์เบื้องต้น หรือ จัดเตรียมข้อมูลด้วยโครงสร้างข้อมูลสำหรับการแสดงผลกราฟ เช่น *Graphviz*, *Tulip*
- **วิเคราะห์ (Analysis):** เมื่อเครือข่ายมีขนาดใหญ่ขึ้น การค้นหาคุณลักษณะสำคัญบางอย่างนั้นย่อมทำได้ยากขึ้น เครื่องมือที่สามารถช่วยวิเคราะห์เครือข่ายได้อย่างมีประสิทธิภาพนั้น *NetworkX*, *igraph*
- **แสดงผล (Visualization):** เพื่อการสื่อสารเกี่ยวกับโครงสร้างของกราฟนั้นชัดเจนมากยิ่งขึ้นเครื่องมือสำหรับวาดกราฟข้างต้นนั้น จะทำได้เพียงแค่รูปแบบพื้นฐานเพียงแค่ว่า ขาว-ดำ เท่านั้น แต่เพื่อให้สามารถเข้าใจถึงโครงสร้างและคุณลักษณะพิเศษของกราฟมากยิ่งขึ้น จึงต้องใช้เครื่องมือเพิ่มเติมเพื่อกำหนดสี สัญลักษณ์แทนโหนด หรือข้อมูลอื่นๆ ประกอบ เพื่อให้การแสดงผลชัดเจนมากยิ่งขึ้นได้ ยกตัวอย่างเช่น *Gephi*, *Cytoscape*

#### D. หัวข้อการศึกษา

เนื่องจากแนวทางการศึกษาวิจัยในเครือข่ายสังคมออนไลน์นั้นมีหลากหลาย ดังนั้นเพื่อให้ง่ายต่อการเข้าใจ จึงได้แบ่งกลุ่มการศึกษามาเป็น 3 กลุ่มด้วยกัน ได้แก่

- 1) การวิเคราะห์โครงสร้าง: kkkkkk
- 2) การวิเคราะห์ข้อมูลเชิงสังคม:
- 3) การวิเคราะห์ข้อมูลปฏิสัมพันธ์เชิงสังคม: