

Social Network Analysis of International E-road Network

การวิเคราะห์เครือข่ายทางสังคมของเครือข่ายถนนระหว่างเมืองในทวีปยุโรป

พิชญ์พงศ์ พูลผล

ภาควิชาวิทยาการคอมพิวเตอร์, คณะวิทยาศาสตร์

มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร

peachapong.poolpol@swu.ac.th

บทคัดย่อ - เครือข่ายถนนที่เชื่อมต่อกันส่งผลให้การเดินทางระหว่างเมืองเชื่อมถึงกัน การวิเคราะห์เครือข่ายทางสังคมของเครือข่ายถนนระหว่างเมืองในทวีปยุโรปด้วยเครื่องมือ Neo4j ทำให้เห็นภาพรวมของข้อมูล การทำ Visualization ผลการวัดหา Centrality รวมทั้งการใช้อัลกอริทึมในการค้นหาชุมชน โดยพบว่าเครือข่ายถนนระหว่างเมืองในทวีปยุโรปนี้ประกอบด้วย 1174 Nodes และ 1417 Edges โดยมาตรวจวัดของ Centrality แต่ละแบบได้ผลที่แตกต่างกัน รวมทั้งอัลกอริทึมที่แตกต่างกันในการค้นหาชุมชน ก็ได้ผลที่แตกต่างกันด้วย ทั้งนี้ยังมีข้อจำกัดของการใช้อัลกอริทึมที่อาจยังพัฒนาไม่เต็มที่ รวมถึงข้อจำกัดของข้อมูลที่ขาดการถ่วงน้ำหนัก ทำให้อาจไม่สะท้อนภาพความเป็นจริงของเครือข่ายถนนในชีวิตจริง

คำสำคัญ - การวิเคราะห์เครือข่ายทางสังคม, การค้นหาชุมชน, Centrality, Neo4j

I. บทนำ

การวิเคราะห์เครือข่ายทางสังคมเป็นกระบวนการในการวิเคราะห์เครือข่ายผ่านการใช้ทฤษฎีกราฟและการทำ Visualization โดยประกอบด้วยโครงสร้างหลักคือ Node ซึ่งมีข้อมูลต่าง ๆ บรรจุอยู่ภายใน เชื่อมต่อกันด้วย Edge ซึ่งทำหน้าที่เชื่อมต่อหรือแสดงความสัมพันธ์ระหว่าง Node โดยการวิเคราะห์เครือข่ายทางสังคมสามารถใช้ได้ในงานที่หลากหลาย ทำให้เห็นปฏิสัมพันธ์ของข้อมูล รวมทั้งการส่งต่อข้อมูลไปยังที่ต่าง ๆ เช่น การศึกษาเรื่องนักการเมืองที่มีอิทธิพล การตรวจจับการโกง การค้นหาผู้ติดเชื้อจากโรคติดเชื้อไวรัสโคโรนา 2019 เป็นต้น

การขนส่งและการเดินทาง เป็นอีกหนึ่งบริบทที่สามารถใช้การวิเคราะห์เครือข่ายทางสังคมในการค้นหาหรือสืบค้นข้อมูลต่าง ๆ เช่น การเดินทางโดยรถสาธารณะ ว่าตัวเลือกได้เป็นตัวเลือกที่เหมาะสม การเดินทางโดยเครื่องบิน สามารถศึกษาหาเส้นทางที่ดีที่สุด ที่ใช้ระยะเวลาการเดินทางและการคายน้อยที่สุด หรือการศึกษาเกี่ยวกับโครงข่ายของถนนในประเทศ หรือในทวีปต่าง ๆ เพื่อค้นหาการเชื่อมโยงระหว่างแยก หรือการเชื่อมโยงระหว่างเมืองต่าง ๆ อีกทั้งยังสามารถใช้การค้นหาชุมชนว่า เมืองที่อยู่รวมกันอย่างหนาแน่นมันจะมีโครงข่ายถนนที่เชื่อมต่อกันมากด้วยเช่นกัน

ในบทความนี้จะศึกษาเกี่ยวกับเครือข่ายถนนนานาชาติระหว่างเมืองในทวีปยุโรป โดยจะศึกษาข้อมูลทั่วไป การวัดความสำคัญของเครือข่ายถนนในทวีปยุโรปโดยใช้ Centrality และตัวอย่างการค้นหาชุมชน โดยใช้เครื่องมือหลักในการวิเคราะห์คือ Neo4j

II. งานวิจัยที่เกี่ยวข้อง

A. US Road Network [1]

การศึกษาเครือข่ายของถนนในสหรัฐอเมริกาโดยใช้การวิเคราะห์เครือข่ายสังคม เป็นการศึกษาเครือข่ายของถนนใน 3 รัฐ ได้แก่ California, Pennsylvania และ Texas โดยในการศึกษานี้จะเน้นไปที่การจัดกลุ่มของเครือข่ายถนนในการสร้างชุมชน เครือข่ายถนนในการศึกษานี้เป็นเครือข่ายขนาดใหญ่ ประกอบด้วย Nodes ประมาณ 1-2 ล้านและ Edges ประมาณ 1.5-2.7 ล้าน ซึ่งถือว่าเป็นข้อมูลขนาดใหญ่ โดย Node แสดงถึงแยก และ Edge แสดงถึงถนนที่เชื่อมแยกต่าง ๆ ซึ่งพบว่าเครือข่ายถนนของทั้ง 3 รัฐมีลักษณะโครงสร้างที่คล้ายคลึงกัน คือ ถนนส่วนใหญ่กระจุกตัวอยู่ในเมืองอย่างหนาแน่น โดย Node ที่มี Degree สูงที่สุดคือ 24 โดยในการศึกษานี้ใช้

เครื่องมือในการวิเคราะห์คือ GraphX, Spark Streaming, Tableau และ MLib มีการจัดกลุ่มโดยใช้อัลกอริทึม k -means โดยให้ k เท่ากับ 3 และ 4

B. Centrality

1. Degree centrality [2]

Degree centrality เป็นอัลกอริทึมที่ง่ายที่สุดในการวัด centrality เป็นการวัดที่ตรงไปตรงมา โดยการหา Node จากจำนวนความสัมพันธ์ที่เข้าหรือออกจาก Node นั้น ซึ่งเปรียบเสมือนการหา Node ที่มีความนิยมสูงสุด ในที่นี้จะเป็นการค้นหาเมืองที่มีจำนวนถนนเชื่อมออกไปยังเมืองอื่น ๆ มากที่สุด

2. Betweenness centrality [3]

Betweenness centrality เป็นอัลกอริทึมที่ใช้ในการค้นหา Node ที่ทำหน้าที่เป็นผู้มีอิทธิพลในการเดินทางของข้อมูล เปรียบเสมือนเป็นเมืองทางผ่านหรือเป็นสะพานเชื่อม โดยการวัดว่าเส้นทางที่สั้นที่สุด หรือ Shortest path จาก Node หนึ่งไปยัง Node อื่น ๆ หรือในที่นี้ก็คือการค้นหาเส้นทางที่สั้นที่สุดเมืองหนึ่ง ๆ ไปยังเมืองอื่น ๆ จะต้องผ่านเมืองใดมากที่สุด อาจเปรียบได้ว่าเป็นเมืองที่มีการนำข้อมูลไปยังเมืองอื่น ๆ มากที่สุด

3. Closeness centrality [4]

Closeness centrality เป็นอัลกอริทึมในการวัดหา Node ที่สามารถแพร่กระจายข้อมูลไปได้อย่างมีประสิทธิภาพมากที่สุด เป็นการวัดหาค่าเฉลี่ยของความไกลหรือระยะทางไปยัง Node อื่น ๆ โดย Node ที่มีค่าสูงหมายถึงมีระยะทางสั้นที่สุดเพื่อไปยัง Node อื่น ๆ ในที่นี้จึงเป็นการวัดว่าเมืองที่อยู่ใกล้กับเมืองอื่นมากที่สุด คือ เมืองใดแสดงว่าจากเมืองนั้นสามารถเข้าถึงได้ไวที่สุด

4. Eigenvector centrality [5]

Eigenvector centrality เป็นการวัดที่สนใจคุณภาพของ Node ว่ามีความสำคัญหรือมีอิทธิพลมากเพียงใด โดยจะพิจารณา Node ข้างเคียงด้วยว่ามีการเชื่อมต่อกับ Node ที่เหลืออื่น ๆ อย่างไรบ้าง สำหรับในข้อมูลนี้จึงเป็นการดูคุณภาพของเมืองมีการพิจารณาถนนที่เชื่อมโยงจากเมืองข้างเคียงด้วย

5. PageRank centrality [6]

PageRank centrality เป็นอัลกอริทึมที่ใช้วัดความสำคัญของ แต่ละ Node ภายในกราฟ โดยอาศัยจำนวนของความสัมพันธ์ที่เข้า

มายัง Node นั้น โดยมีการให้คะแนนตามจำนวนความสัมพันธ์ที่เข้ามา โดยเป็น Centrality ที่ใช้ในการวัดสำหรับกราฟแบบมีทิศทาง ซึ่งเป็นเครื่องมือของ Google ทั้งนี้ในการศึกษานี้จะไม่มีการคำนวณ PageRank centrality เนื่องจากว่าเป็นกราฟแบบไม่มีทิศทาง

III. ระเบียบวิธีการวิจัย

A. ข้อมูล

การศึกษานี้ใช้ข้อมูลเปิด เป็นข้อมูลเครือข่ายถนนนานาชาติที่เชื่อมต่อระหว่างเมืองในทวีปยุโรป [7] ชื่อ Euroroad จาก http://konect.cc/networks/subelj_euroroad/ โดยมี Node เป็นชื่อของเมืองต่าง ๆ ในทวีปยุโรป และ Edge หรือเส้นเชื่อมเป็นถนนที่เชื่อมระหว่างเมือง โดยกราฟในบทความนี้เป็นกราฟแบบไม่มีทิศทาง และไม่มีการถ่วงน้ำหนัก (Undirected unweighted graph)

B. เครื่องมือที่ใช้

การศึกษานี้จะใช้การวิเคราะห์เครือข่ายทางสังคมผ่านเครื่องมือ Neo4j เป็นหลัก โดยใช้คำสั่ง Cypher Query Language (CQL) ในการส่งคำสั่ง ร่วมกับใช้ Graph Data Science Library ในการวิเคราะห์หา Centrality และการค้นหาชุมชน โดยอัลกอริทึมที่ใช้หา Centrality ประกอบด้วย Degree centrality, Betweenness centrality, Closeness centrality และ Eigenvector centrality และอัลกอริทึมที่ใช้ค้นหาชุมชนประกอบด้วย Louvain Method, Weakly Connected Components, Label Propagation, Triangle count และ Local clustering coefficient ทั้งนี้อัลกอริทึมสำหรับการค้นหาชุมชนยังมีอีกหลายชนิด แต่ผู้วิจัยเลือกใช้เฉพาะอัลกอริทึมชนิด Production-quality เนื่องจากมีการทดสอบแล้วว่ามีความคงที่สูง

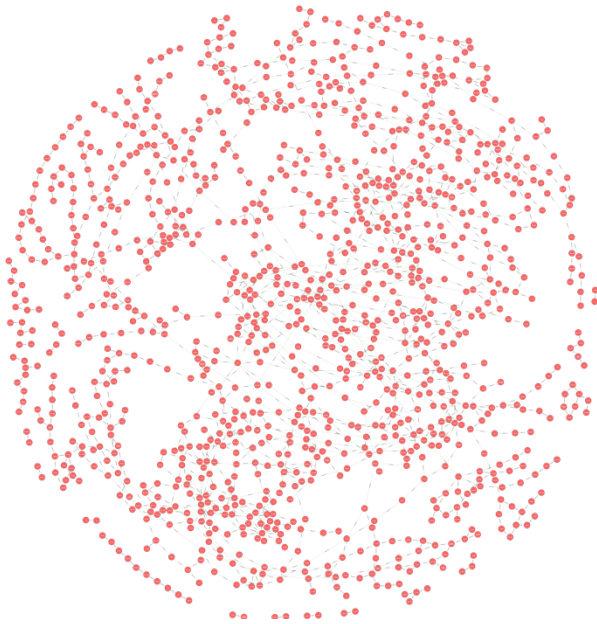
IV. ผลการศึกษา

A. ข้อมูลทั่วไป

การนำเข้าข้อมูลจะนำเข้าผ่านไฟล์ชนิด csv จำนวน 2 ไฟล์ ประกอบด้วยไฟล์ที่เป็น Node และ Edge list โดยมี Node เพียงชนิดเดียว ดังภาพที่ 1 คือ ชื่อของเมือง ซึ่งกราฟนี้จะประกอบด้วย 1174 Nodes และ 1417 Edges และแสดงแผนภาพของการเชื่อมโยงของเครือข่ายถนน ดังภาพที่ 2



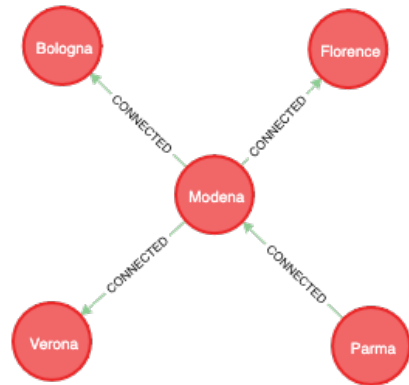
ภาพที่ 1 แสดงลักษณะของ Node และ relationship



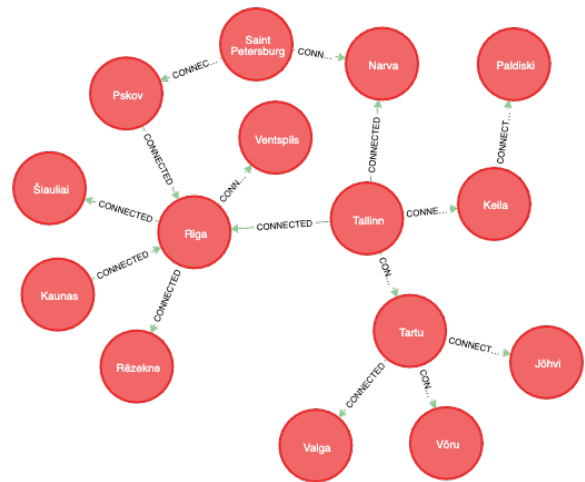
ภาพที่ 2 ตัวอย่างแผนภาพของการเชื่อมโยงของเครือข่ายถนนระหว่างเมืองในทวีปยุโรป

แสดงตัวอย่าง Adjacent node ที่อยู่ติดกัน เช่นเมือง Modena มี Adjacent nodes คือเมือง Bologna, เมือง Florence, เมือง Verona, เมือง Parma โดยเมือง Modena มีจำนวน Degree เท่ากับ 4 ดังภาพที่ 3

แสดงตัวอย่างการเดินทางจากเมืองหนึ่งไปยังเมืองใกล้เคียง โดยหากต้องการเดินทางจากเมือง Tallinn เพื่อไปเที่ยว 2 เมืองใกล้เคียงทางถนน สามารถเดินทางไปได้วยังเมืองต่าง ๆ ได้ดังภาพที่ 4



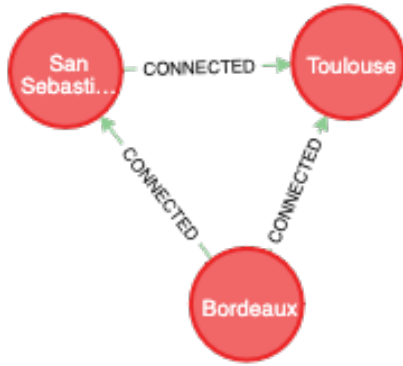
ภาพที่ 3 แสดงการเชื่อมต่อของถนนจากเมือง Modena ไปยังเมืองต่าง ๆ



ภาพที่ 4 แสดงตัวอย่างการเดินทางจากเมือง Tallinn ไปยัง 2 เมืองข้างเคียง

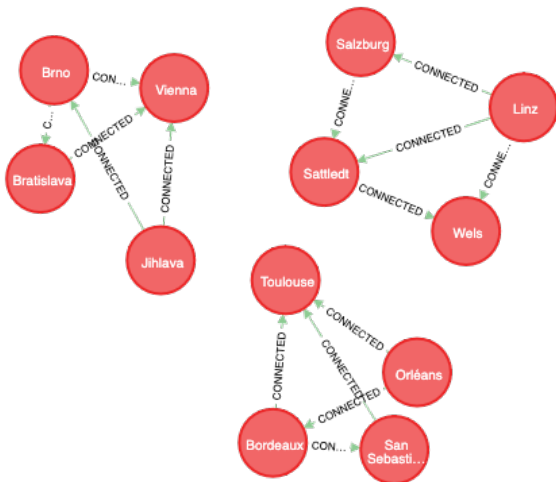
ในตัวอย่างข้อมูลชุดนี้จะไม่ข้อมูลของ Parallel edge ซึ่งหมายถึง Edge 2 เส้นเชื่อมที่มี Node คู่เดียวกันและ self-loop ซึ่งหมายถึง Edge ที่มี Node เป็นตัวเดียวกัน เนื่องจากเป็นข้อมูลจริงของถนนที่ถูกกำหนดว่าเป็นถนนเชื่อมต่อระหว่างเมือง จึงจัดว่าโครงข่ายของถนนนี้เป็น Simple graph

แสดงตัวอย่างของ Simple cycle ที่เป็น Triad ระหว่างเมือง San Sebastian, Bordeaux และ Toulouse ดังภาพที่ 5



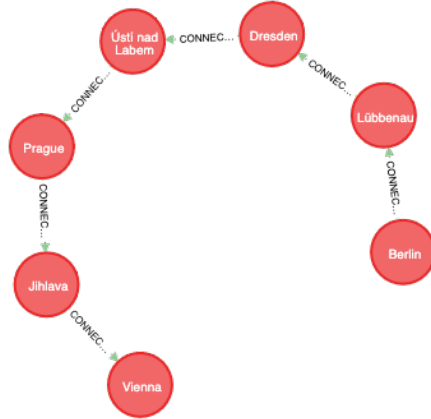
ภาพที่ 5 แสดงตัวอย่างของ Simple cycle

จากการสำรวจข้อมูลเพื่อหา Complete graph หรือกราฟที่มี Edge เชื่อมไปยังทุก Node ในกราฟนั้น พบว่าไม่มี complete graph หากพิจารณาตั้งแต่ 4 เมืองขึ้นไป โดยในข้อมูลชุดนี้ มีเพียงกราฟที่มีลักษณะใกล้เคียงกับ Complete graph สำหรับ 4 เมือง ดังภาพที่ 6



ภาพที่ 6 แสดงตัวอย่างใกล้เคียง Complete graph ระหว่าง 4 เมือง

การสำรวจเพื่อดู Reachability หรือการเข้าถึง Node อื่น ๆ จาก Node หนึ่ง โดยแสดงตัวอย่างในการค้นหาว่าเมือง Berlin สามารถเดินทางไปยังเมือง Vienna ได้ด้วยเส้นทางที่สั้นที่สุด คือต้องเดินทางจากเมือง Berlin ผ่านเมือง Lubbenau ผ่านเมือง Dresden ผ่านเมือง Usti nad Labem ผ่านเมือง Prague และผ่านเมือง Jihlava เพื่อไปยังเมือง Vienna ดังภาพที่ 7



ภาพที่ 7 แสดงตัวอย่างการเดินทางจากเมือง Berlin ไปยังเมือง Vienna ด้วยการผ่านเมืองที่น้อยที่สุด

B. Centrality

1. Degree centrality

พบว่าเมืองที่มี Degree centrality สูงที่สุดได้แก่เมือง Moscow คือมีถนนเชื่อมออกจำนวน 10 สาย โดยแสดง 10 ลำดับของเมืองที่มี Degree centrality สูงสุด หรือหมายถึงเมืองที่มีถนนเชื่อมต่อไปยังเมืองข้างเคียงมากที่สุด แสดงดังตารางที่ 1 และแสดงเมืองที่เชื่อมต่อกับเมือง Moscow ดังภาพที่ 8

ตารางที่ 1 แสดง 10 ลำดับของเมืองที่มี Degree centrality สูงที่สุด

เมือง	Degree centrality
1. Moscow	10
2. Paris, Liège, Munich, Berlin, Budapest	8
3. Metz, Prague, Bratislava, Warsaw, Vienna	7



ภาพที่ 8 แสดงการเชื่อมต่อของเมืองต่าง ๆ กับเมือง Moscow ที่มี Degree centrality สูงสุด

2. Betweenness centrality

พบว่าเมืองที่มี Betweenness centrality สูงที่สุดคือเมือง Brest รองลงมาคือเมือง Moscow และเมือง 10 ลำดับแรกที่มีค่า Betweenness centrality สูงที่สุด แสดงดังตารางที่ 2

ตารางที่ 2 แสดง 10 ลำดับของเมืองที่มี Betweenness centrality สูงที่สุด

เมือง	Betweenness centrality
1. Brest	147,644.87
2. Moscow	146,078.27
3. Saint Petersburg	134,358.41
4. Le Mans	120,368.00
5. Rennes	119,855.24
6. Minsk	119,155.66
7. Warsaw	110,015.70
8. Smolensk	107,544.26
9. Vyborg	105,952.67
10. Vaalimaa	105,160.33

3. Closeness centrality

พบว่าจำนวน 27 เมืองที่มีค่า closeness centrality สูงสุดคือเท่ากับ 1 ได้แก่เมือง Porto-Vecchio, Sassari, Å, Svolvær, Førde,

Usharal, Druzhba, Trosna, Glukhkov, Courtney, Troyes, Memmingen, Füssen, Krasnodar, Djoubga, Digoin, Chalon-sur-Saône, Lugo, Ilia, Spezzano Albanese, Sybaris, Sant'Eufemia Lamezia, Catanzaro, Alcamo, Trapani, Eleusina และ Thebes

4. Eigenvector centrality

พบว่าเมืองที่มี Eigenvector centrality สูงสุดได้แก่เมือง Paris และเมือง 10 ลำดับแรกที่มีค่า Eigenvector centrality สูงที่สุด แสดงดังตารางที่ 3

ตารางที่ 3 แสดง 10 ลำดับของเมืองที่มี Eigenvector centrality สูงที่สุด

เมือง	Eigenvector centrality
1. Paris	17.37
2. Metz	14.52
3. Reims	13.23
4. Vienna	12.43
5. Bratislava	11.67
6. Brussels	11.47
7. Brno	11.00
8. Liège	10.50
9. Munich	10.30
10. Orléans	10.13

C. การค้นหาชุมชน (Community detection)

ในการศึกษานี้เลือกใช้อัลกอริทึมที่ใช้ค้นหาชุมชนทั้งหมดจำนวน 5 อัลกอริทึม ได้แก่ Louvain Method, Weakly Connected Components, Label Propagation, Triangle count และ Local clustering coefficient

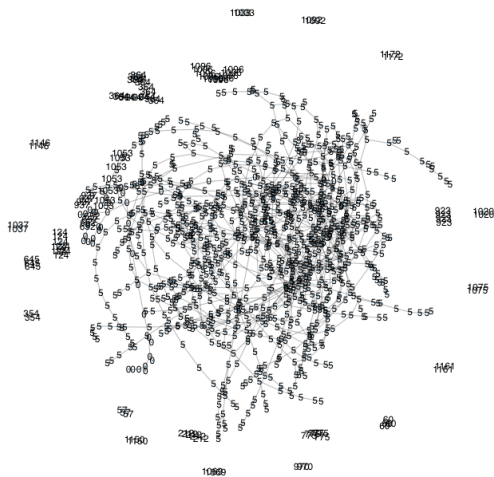
1. Louvain Method [8]

Louvain Method เป็นอัลกอริทึมที่ใช้ในการค้นหาชุมชนขนาดใหญ่ โดยทำการหาค่า Modularity score ที่มากที่สุดในแต่ละชุมชน โดยที่ค่า Modularity จะบอกคุณภาพของการจัดการ Node ในชุมชน เป็นการวัดว่าชุมชนอยู่กันหนาแน่นเพียงใด นั่นหมายถึงช่วง

เมืองใดที่มีถนนเชื่อมโยงกันอย่างหนาแน่น พบว่าจาก Louvain method นี้สามารถค้นหาชุมชนได้จำนวน 45 ชุมชน โดยชุมชนที่มีขนาดใหญ่ที่สุดพบว่า มีจำนวนเมืองในชุมชนนั้น 76 เมืองด้วยกัน โดยพบว่าไม่มี intermediate communities เลย ทั้งนี้เพราะว่า Louvain method เป็นอัลกอริทึม แบบการจัดกลุ่มเป็นลำดับขั้น (Hierarchical clustering) คือจะมีการจัดกลุ่มรวมกันเพิ่มขึ้นเรื่อย ๆ จนกว่าจะได้ Node ที่สมบูรณ์ ในบางกรณีอาจจะมียา intermediate communities ได้

2. Weakly Connected Components [9]

Weakly connected components (WCC) เป็นการค้นหาชุมชนที่เชื่อมต่อกันสำหรับกราฟที่ไม่มีทิศทาง โดยพบว่าเครือข่ายถนนในทวีปยุโรปนี้ไม่เป็น connectedness คือทั้งเครือข่ายไม่ได้เชื่อมต่อกันอย่างสมบูรณ์ เพราะสามารถแบ่งเป็น connected subgraph ทั้งหมดจำนวน 26 กราฟย่อย ซึ่งกราฟย่อยที่เป็น connected components ของเครือข่ายถนนนี้มีจำนวน Node มากที่สุด มีทั้งสิ้น 1039 Nodes โดยแสดงชุมชนที่เชื่อมต่อกันดังภาพที่ 9



ภาพที่ 9 แสดงชุมชนที่ทำการค้นหาด้วยวิธี Weakly connected components

3. Label Propagation [10]

Label Propagation (LPA) เป็นอัลกอริทึมที่อาศัยโครงสร้างของเครือข่ายเป็นต้นนำ โดยไม่สนใจข้อมูลอื่น ๆ ซึ่ง LPA ทำงานโดยการติดฉลากและเผยแพร่ Node ต่าง ๆ ไปตามโครงข่าย พบว่าได้

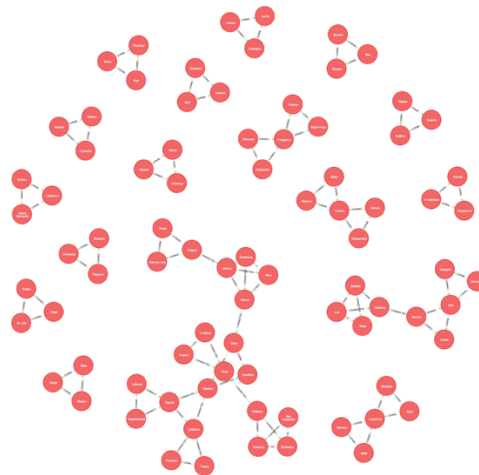
ชุมชนทั้งสิ้น 47 ชุมชน และชุมชนที่มีขนาดใหญ่ที่สุด ประกอบด้วย 316 เมือง โดยเปรียบเทียบกับอัลกอริทึมอื่น ๆ ดังตารางที่ 4

ตารางที่ 4 เปรียบเทียบจำนวนชุมชนและจำนวนเมืองสูงที่สุดในแต่ละชุมชนในแต่ละอัลกอริทึม

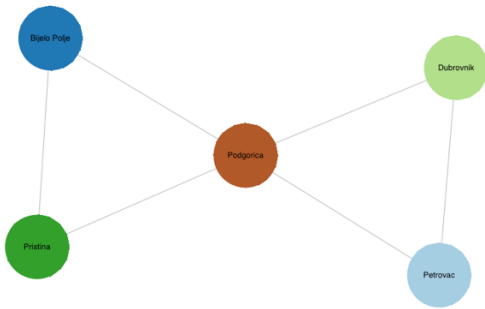
อัลกอริทึม	จำนวนชุมชน	จำนวนเมืองสูงที่สุด
Louvain Method	45	76
WCC	26	1039
LPA	47	316

4. Triangle count [11]

Triangle count เป็นอัลกอริทึมสำหรับนับจำนวนสามเหลี่ยมของแต่ละ Node โดยสามเหลี่ยมจะประกอบไปด้วย Node จำนวน 3 Nodes โดยที่แต่ละ Node มีความสัมพันธ์เชื่อมกับ 2 Nodes ที่เหลือ ในทฤษฎีของกราฟอาจเรียกว่า 3-clique ข้อดีของการใช้ triangle count ในฐานข้อมูลนี้คือ ทำให้ทราบได้ว่า มีเมือง 3 เมืองใดบ้างที่สามารถเชื่อมหากันได้ ซึ่งจะแสดงว่าทั้ง 3 เมืองมีความเป็นกลุ่มก้อนและความสำคัญพอกัน ซึ่งพบว่ามี 70 เมืองที่มีการเชื่อมกันเป็นสามเหลี่ยม 1 รูป และอีก 13 เมืองที่มีการเชื่อมกันเป็นสามเหลี่ยม 2 รูป แสดงดังภาพที่ 10 ในขณะที่เมืองที่เหลืออีก 1091 เมืองไม่มีถนนเชื่อมต่อกันเป็นสามเหลี่ยมเลย โดยแสดงตัวอย่างของเมืองที่มีการเชื่อมต่อกันเป็นสามเหลี่ยม 2 รูป เช่น Podgorica ซึ่งแสดงในภาพที่ 11 ทั้งนี้อัลกอริทึม Triangle count มีความแตกต่างจากอัลกอริทึมที่ 3 วิธีแรกที่กล่าวไปข้างต้น



ภาพที่ 10 แสดงเมืองที่มีการเชื่อมต่อกันเป็นสามเหลี่ยม



ภาพที่ 11 แสดงเมืองที่มีการเชื่อมต่อกันเป็นสามเหลี่ยมสองรูป

5. Local clustering coefficient [11]

Local clustering coefficient เป็นอัลกอริทึมที่คำนวณค่าของแต่ละ Node ในกราฟ คือความเป็นไปได้ที่ Node เพื่อนบ้านจะเชื่อมต่อกัน ซึ่งอาศัยการคำนวณจากสามเหลี่ยมของ Node ที่เกิดขึ้น และ Degree ของ Node โดยใช้สูตรดังนี้

$$C_n = T_n \frac{2T_n}{dn(dn-1)} \quad (1)$$

C_n คือค่า coefficient ของ Node n

T_n คือค่าของสามเหลี่ยม

dn คือค่าของ degree ของ Node

พบว่าจำนวน 4 เมืองที่มีค่า Local clustering coefficient สูงสุด เท่ากับ 1 ได้แก่เมือง Brindisi, Domokos, Vyšné Nemecké, Postojna และเมือง Spielfeld และอีก 22 เมือง มีค่า coefficient เท่ากับ 0.33

V. สรุปผลการศึกษา

การวัดโดยใช้ Centrality พบว่า มาตรวัดแต่ละชนิดในการศึกษาโครงข่ายทางสังคมก็ได้ผลที่แตกต่างกัน เนื่องจากความสำคัญในการวัด และรูปแบบของอัลกอริทึมในการหาแตกต่างกัน เช่นเดียวกับการค้นหาชุมชน อัลกอริทึมที่แตกต่างกัน ก็ได้ผลของการค้นหาชุมชนที่แตกต่างกันด้วย

Neo4j เป็น Graph platform ที่ถูกพัฒนาเพื่องานกราฟ โดยเฉพาะการทำให้เป็นฐานข้อมูลกราฟ มีการทำ Visualization ที่สวยงามสามารถทำ Graph analytics ได้เช่น การทำ Pathfinding and search การหา Centrality การค้นหาชุมชน (Community

detection) การทำ Link Prediction และการหา Similarity โดยการวิเคราะห์กราฟจะผ่าน Graph Data Science Library ซึ่งปัจจุบันยังมีข้อจำกัดบางประการในบางอัลกอริทึม ยังมีบางอัลกอริทึมที่ยังพัฒนาไม่สมบูรณ์

ในส่วนข้อจำกัดของเครือข่ายนี้คือ ไม่มีการถ่วงน้ำหนัก หมายถึงไม่ได้มีการบ่งบอกระยะทางของถนนจริง ๆ เป็นการบอกเพียงว่าเมืองต่าง ๆ มีการเชื่อมด้วยถนนอะไรบ้าง อาจส่งผลต่อการวัดต่าง ๆ ที่ควรจะต้องมีการถ่วงน้ำหนักเข้ามาเกี่ยวข้อง โดยเฉพาะระยะทางระหว่างเมืองต่าง ๆ หรือการหา Shortest path เป็นต้น

VI. เอกสารอ้างอิง

- [1] Mambou EN, Nlend S, Liu H, editors. Study of the US road network based on social network analysis. 2017 IEEE AFRICON; 2017 18-20 Sept. 2017.
- [2] Freeman LC. Centrality in social networks conceptual clarification. Social Networks. 1978;1(3):215-39.
- [3] Brandes U, Pich C. Centrality Estimation in Large Networks. International Journal of Bifurcation and Chaos. 2007;17(07):2303-18.
- [4] Borgatti SP. Centrality and network flow. Social Networks. 2005;27(1):55-71.
- [5] Bonacich P. Power and Centrality: A Family of Measures. American Journal of Sociology. 1987;92(5):1170-82.
- [6] Brin S, Page L. The anatomy of a large-scale hypertextual Web search engine. Computer Networks and ISDN Systems. 1998;30(1):107-17.
- [7] Šubelj L, Bajec M. Robust network community detection using balanced propagation. The European Physical Journal B. 2011;81.
- [8] Lu H, Halappanavar M, Kalyanaraman A, Choudhury S, editors. Parallel Heuristics for Scalable Community Detection. 2014 IEEE International Parallel &

Distributed Processing Symposium Workshops; 2014
19-23 May 2014.

- [9] Monge AE, Elkan C, editors. An Efficient Domain-Independent Algorithm for Detecting Approximately Duplicate Database Records. DMKD; 1997.
- [10] Raghavan U, Albert R, Kumara S. Near linear time algorithm to detect community structures in large-scale networks. Physical review E, Statistical, nonlinear, and soft matter physics. 2007;76 3 Pt 2:036106.
- [11] Becchetti L, Boldi P, Castillo C, Gionis A. Efficient semi-streaming algorithms for local triangle counting in massive graphs. Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining; Las Vegas, Nevada, USA: Association for Computing Machinery; 2008. p. 16–24.