

การวิจัยเชิงสำรวจในการทำนายการจราจร โดยใช้การวิเคราะห์ด้วย Big Data

A Survey on Traffic Prediction based on Big Data Analytics

นายวุฒิชัย โนนสาธุ (Woottichai Nonsakhoo)¹ และนายปิยณัฐ ศิริสวัสดิ์ (Piyanat Sirisawat)¹

¹ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น

nonsakhoo@kkumail.com, piyanat_siri@kkumail.com

บทคัดย่อ

การเพิ่มขึ้นของข้อมูลในปัจจุบัน ทำให้มีการนำเอาเทคโนโลยีการเก็บข้อมูลขนาดใหญ่ (Big Data) มาร่วมใช้งานในหลากหลายแขนง เช่น การแพทย์ การผลิต และการโทรคมนาคม ซึ่งมีการนำเอาการวิเคราะห์ข้อมูลขนาดใหญ่มาใช้งาน ทำให้ข้อมูลที่มีความแม่นยำ และสามารถทำนายได้มากกว่าที่ผ่านมา และในปัจจุบันได้มีการนำมาใช้งานด้านการพยากรณ์และทำนายสภาพการจราจร โดยมีการศึกษาที่ผ่านมาได้นำเอา Big Data เข้ามาใช้งาน ทั้ง Machine learning, Data mining และ Neural networks โดยใช้แนวทางที่หลากหลายและจะเป็นประโยชน์ต่อการนำไปใช้งานในอนาคตต่อไป

คำสำคัญ: ข้อมูลขนาดใหญ่, การทำนายการจราจร, การเรียนรู้เชิงลึก, โครงข่ายประสาทเทียม

Abstract

Nowadays, big data are widely applied to different field. Many industries such as healthcare, manufacturing, telecommunications, and other applications market imports the technology of Big Data Analysis. This provides the advantage of valuable information that is hardly to perform in the past. To forecast and predict traffic situation, the various methods of big historical traffic data analytics. With the used of Big Data Analytics, Machine learning, Data mining, and Neural networks. The result are thoroughly studied and concisely summarized with comparison table. This can improve the possible usage of Big Data Analytics in traffic prediction.

Keyword: Big Data, Traffic Prediction, Deep learning, Neural Network.

1. บทนำ

ในปัจจุบันมีรถยนต์เพิ่มขึ้นจำนวนมาก การใช้งานการจราจรจึงเพิ่มมากขึ้น ปัญหาหนึ่งที่มาคือการจราจรติดขัด ทำให้การเดินทางไม่สะดวกและสูญเสียเวลา โดยเฉพาะอย่างยิ่งการจราจรในช่วงเวลาเร่งด่วน และการจราจรบนเส้นทางที่มีการใช้งานจำนวนมาก ซึ่งจะพบบ่อยบนถนนในเมืองขนาดใหญ่ จึงทำให้เกิดการเฝ้าระวังการตรวจสอบการจราจรแบบ Real-Time เกิดขึ้น ทำให้สามารถบอกถึงเส้นทางที่มีการจราจรจำนวนมากในเวลานั้น ๆ ได้ ช่วยทำให้สามารถเดินทางได้สะดวกโดยหลีกเลี่ยงเส้นทางที่มีการใช้งานมากในขณะนั้น แต่ถึงอย่างไรก็ตาม การจะวางแผนการเดินทาง [1] จำเป็นต้องทราบถึงโอกาสที่เส้นทางที่เราเดินทางอยู่ หรือเส้นทางที่เราวางแผนจะเดินทางนั้นจะเกิดการจราจรติดขัดได้ ดังนั้น จึงเกิดการสร้างการทำนายการจราจร โดยใช้ข้อมูลการเดินทางต่าง ๆ ที่เก็บเอาไว้จำนวนมาก ทั้งจาก Sensor ต่าง ๆ บนตัวรถยนต์หรือระบบ GPS/GPRS แต่เพราะข้อมูลมีจำนวนมากจึงนำเอาเทคโนโลยี Big Data มาใช้งาน เพื่อช่วยในการเก็บข้อมูล และนำข้อมูลมาใช้งานได้อย่างสะดวกและรวดเร็ว [2] ซึ่งหนึ่งในโปรแกรมที่มีการใช้งานอย่างกว้างขวางคือ Hadoop ที่สามารถจัดเก็บข้อมูล หรือคำนวณข้อมูลพร้อม ๆ กันได้ในเวลาเดียวกัน

2. เทคโนโลยี Big Data

2.1 ความหมายของ Big Data

คำว่า Big Data เป็นการเรียกข้อมูลที่มีขนาดใหญ่ ทั้งที่มีโครงสร้างและไม่มีโครงสร้าง ซึ่งประกอบด้วย 3V คือ Volume (ขนาด), Velocity (ความเร็ว) และ Variety (ความหลากหลาย)

ซึ่งบางหน่วยงานได้เพิ่ม Variability (ความแปรปรวน) Veracity (ความจริง) และ Complexity (ความซับซ้อน) เพิ่มเติมตามรูปแบบที่ต้องการใช้งาน

2.2 การวิเคราะห์ข้อมูลใน Big Data

การวิเคราะห์ข้อมูลใน Big Data นั้นจะประกอบไปด้วยวิธีการที่หลากหลายแต่โดยหลัก ๆ แล้วจะประกอบไปด้วย

Machine learning เป็นกระบวนการให้คอมพิวเตอร์เรียนรู้โดยไม่ต้องโปรแกรมอย่างชัดเจน เน้นการสร้าง Algorithms ที่สามารถเรียนรู้และทำนายจากข้อมูลได้ ช่วยในการทำนายและตัดสินใจจากข้อมูลที่มี

Data mining เป็นกระบวนการเรียนรู้จากข้อมูลขนาดใหญ่เป็นการนำหนึ่งในกระบวนการ Machine learning มาใช้งานเพื่อทำการคัดแยกข้อมูลออกมาและเลือกเฉพาะข้อมูลที่สามารถนำไปใช้ต่อไป ประกอบด้วย Cluster Analysis, Anomaly Detection และ Association Rule Mining

Cluster Analysis หรือ Clustering เป็นการจัดกลุ่มข้อมูล (Cluster) โดยนำเอาข้อมูลที่มีลักษณะคล้ายกันอย่างไร้แบบหนึ่งมาจัดกลุ่มรวมกัน การจัดกลุ่มจะมี Algorithm ที่แตกต่างกันไปตามรูปแบบที่ต้องการให้มีการจัดกลุ่ม

Anomaly Detection หรือ Outlier Detection เป็นการค้นหารูปแบบข้อมูลที่ไม่มีความสัมพันธ์กับข้อมูลที่อยู่ใน Dataset ที่มีอยู่ ทำให้การวิเคราะห์ข้อมูลเกิดความผิดพลาดขึ้นมาได้ในบาง Algorithm ของ Cluster Analysis ก็อาจจะสามารถค้นเจอความผิดปกตินี้เป็น Clustering ขนาดเล็ก ๆ ได้เรียกข้อมูลแบบนี้อีกอย่างว่า Noise

Association Rule เป็นกระบวนการค้นหาสิ่งที่สัมพันธ์ระหว่างตัวแปรต่าง ๆ ในกลุ่มข้อมูลขนาดใหญ่โดยใช้รูปแบบต่าง ๆ ที่มีความชัดเจนในการค้นหาถึงความสัมพันธ์ที่มี โดยปกติมักจะใช้ในกระบวนการทางการตลาด

Neural Networks เป็นกระบวนการคำนวณโดยใช้ Node หรือ Neural จำนวนมาก ๆ เชื่อมต่อกัน และสามารถกระตุ้นหรือระงับการทำงานของกันและกันได้ในแต่ละ Neural จะมี Function การทำงานที่ต้องทำก่อนส่งผ่านไปยัง Neural ตัวอื่น ๆ ทำให้เกิดการเรียนรู้ได้โดยไม่ต้องทำการโปรแกรมอย่างเจาะจงค่าผลลัพธ์ที่ออกมาจาก Neural แต่ละตัวจะมีค่า 0.0 ถึง 1

2.3 Algorithms ที่มักจะใช้งานใน Big Data

Naive Bayes เป็นการแบ่งกลุ่มข้อมูลออกจากกันโดยใช้ Bayes' Theorem โดยการอ้างอิงถึงความน่าจะเป็นที่คล้ายกันของข้อมูลเป็นตัวชี้วัดในการแบ่งกลุ่ม

k-Nearest Neighbors (KNN) Classification เป็นกระบวนการทางสถิติโดยใช้เพื่อนบ้านรอบ ๆ ข้อมูล และใช้ค่า k ในการหาระยะห่างและแบ่งกลุ่มตามระยะห่างนั้น

Density-based spatial clustering of applications with noise (DBSCAN) เป็นกระบวนการแบ่งกลุ่มโดยใช้จุด ๆ หนึ่ง และคำนวณข้อมูลที่อยู่ใกล้จุดนั้นอยู่ในกลุ่มเดียวกัน ทำให้ข้อมูลที่ผิดปกติไปจากกลุ่มนั้นอยู่ห่างไกลไปจากกลุ่ม

Fuzzy C-means Clustering มีลักษณะคล้าย k-mean คือเลือกจำนวนกลุ่ม จากนั้นคำนวณจุดศูนย์กลางของกลุ่มและเลือกสมาชิกในกลุ่ม ซึ่งจะแตกต่างจาก k-mean คือมีการเพิ่มตัวแปรในการแบ่งกลุ่ม ทำให้กลุ่มที่ได้นั้นละเอียดมากกว่า

2.4 การจัดเก็บข้อมูล Big Data

ในการทำงานกับ Big Data เนื่องจากข้อมูลมีขนาดใหญ่ ทำให้การรวบรวมเร็วในการเข้าถึงข้อมูลเป็นสิ่งสำคัญในการจัดเก็บและวิเคราะห์ข้อมูล โดยโปรแกรมหลัก ๆ ที่ใช้งาน Big Data มีดังนี้

Hadoop เป็นโปรแกรม platform โดยสามารถควบคุมการจัดการคลัสเตอร์ จัดเก็บข้อมูล และประมวลผล ตัวโปรแกรมจะสามารถสร้างกลุ่ม Computer หรือ Cluster โดยมี Master สำหรับควบคุม และ Slave สำหรับใช้งานในการจัดเก็บข้อมูล และประมวลผลพร้อมกันแบบกระจาย (Parallel) ได้ ทำให้มีความรวดเร็วทั้งการจัดเก็บข้อมูลและการประมวลผลข้อมูล

Cloud ในรูปแบบต่าง ๆ เป็น platform ที่สามารถเรียกใช้งาน โดยไม่จำเป็นต้องทำการติดตั้ง Cluster ด้วยตนเอง สามารถทำการจัดเก็บข้อมูล หรือวิเคราะห์ข้อมูลได้ และสามารถเรียกใช้งานจากที่ใดก็ได้ หลัก ๆ จะประกอบไปด้วย Google Cloud, Oracle Cloud และ Amazon Web Services

3. ข้อมูลการจราจรและ Big Data

ข้อมูลที่เก็บมาจากการใช้งานจราจรสามารถเก็บได้หลายรูปแบบ เช่น จาก Sensor ที่ติดกับตัวรถยนต์ หรือจาก Sensor ที่ติดอยู่ตามถนนต่าง ๆ เพื่อหาจำนวนรถยนต์ ดังนั้น ข้อมูล

เหล่านี้จึงมีการปรับปรุงในเวลาอันสั้น เช่น ในเวลา 1 นาที จำนวนรถในถนนก็จะเปลี่ยนไป ทำให้การเก็บข้อมูลเหล่านี้มีจำนวนมาก และในเวลา 1 วันก็จะได้ข้อมูลในแต่ละ Sensor มาหลายพัน Record หากทำการนำข้อมูลเหล่านั้นมารวมกันหลายวันก็จะทำให้ข้อมูลที่ได้มีจำนวนมาก และการจะวิเคราะห์ข้อมูลเหล่านั้นก็จะทำได้ยากตามมา

3.1 การวิเคราะห์ข้อมูลการจราจร

จากการศึกษาของ Schimbinschi และคณะ [3] ซึ่งได้ทำการวิเคราะห์ข้อมูลการจราจรที่เก็บมานานกว่า 6 ปี และมีขนาดมากกว่า 700 กิกะไบต์ โดยนำเอา Algorithm ของ Machine Learning มาใช้งานกับข้อมูลนี้โดยสุ่มมาจำนวน 15% พบว่า Algorithm บางตัวมีการใช้เวลาในการทำงานนาน ในขณะที่บางตัวใช้เวลาในการทำงานสั้นกว่า โดยประสิทธิภาพที่ได้แตกต่างกันไม่มากนัก ทำให้ทราบว่า Algorithm บางตัวนั้น ไม่เหมาะสมที่จะนำมาใช้งานการทำการจราจรแบบ Real-Time เนื่องจากใช้เวลาคำนวณมากกว่า Algorithm แบบอื่น

เพื่อเพิ่มประสิทธิภาพของผลลัพธ์ที่ได้จากการวิเคราะห์ได้ มีการนำเครือข่ายประสาทเทียมมาทดสอบแบบ Stacked Autoencoders (SAE) และ Deep learning [4] กลุ่มข้อมูลที่ใช้ทำงานเป็นข้อมูลที่ได้จาก Sensor ทุก ๆ 30 วินาที จำนวนกว่า 15,000 จุด เป็นเวลา 3 เดือน ผลลัพธ์ที่ได้พบว่าเมื่อเทียบกับ Algorithm ของเครือข่ายประสาทเทียมแบบอื่น ๆ แล้ว SAE มีประสิทธิภาพสูงกว่าประมาณ 10%

3.2 การนำเอา Hadoop มาใช้งาน

เพื่อให้การจัดเก็บข้อมูลเป็นไปอย่างรวดเร็วและการคำนวณมีประสิทธิภาพสูงขึ้น จึงมีการนำ Hadoop เขามาใช้งานร่วมกับ Big Data โดยใช้ KNN [5] ร่วมกับการเรียงข้อมูลในแนวตั้ง เพื่อทำนายหาสภาพการจราจรในแต่ละช่วงเวลาของแต่ละวัน โดยใช้ข้อมูลสภาพการจราจรเป็นเวลา 4 เดือน จากการศึกษาพบว่าสามารถให้ค่า R-Squared ซึ่งมีความถูกต้องเกินกว่า 75%

ใน Platform Hadoop ที่นำมาใช้งาน นั้นสามารถทำการกระจายการประมวลผล หรือ Parallel Computing ได้ จึงมีการเขียน Algorithm โดยใช้ KNN บน MapReduce [6] โดยใช้งานบนข้อมูลที่ได้มาจาก Sensor บน Taxi จำนวน 12,000 คัน เป็นเวลา 15 วัน โดยผลที่ได้พบว่ามีความแม่นยำสูงที่สุดถึง 90.7% และประมวลผลได้รวดเร็วกว่า

ในเรื่องของความรวดเร็ว นอกจากการวิเคราะห์โดยใช้งาน Hadoop ยังมีการใช้งานโปรแกรม Spark เพื่อวิเคราะห์ข้อมูลที่เก็บเอาไว้บน Hadoop ซึ่งสามารถทำงานได้รวดเร็วกว่า โดยนำ DBSCAN, K-Means, Fuzzy C-Means มาใช้ในการทดสอบ [7] และพยายามที่จะนำ Hybrid Algorithm ซึ่งนำเอา Algorithm ปกติมาใช้งานกับการจัดเรียงข้อมูลแนวตั้ง เพื่อช่วยในการทำนายว่าในเวลาที่กำหนดในแต่ละวันมีแนวโน้มของจำนวนรถยนต์อย่างไร โดยผลลัพธ์ที่ได้มีความแม่นยำ 97% สูงกว่ารูปแบบปกติที่มีความแม่นยำ 94% [8]

3.3 ระบบ Cloud และ Parallel Computing

ระบบ Cloud เป็นอีกระบบหนึ่งที่สามารถนำมาใช้งานเก็บข้อมูลและคำนวณ Big Data ได้ โดย Cloud นั้นไม่จำเป็นต้องทำการติดตั้งระบบด้วยตนเอง แต่จำเป็นต้องทำการจ่ายค่า “เช่า” ระบบ เพื่อนำมาใช้งานตามรูปแบบที่เราต้องการได้ การศึกษาของ Dongbo Zhang และคณะ [9] ได้นำเอาการทำ Map-matching algorithm มาใช้งานบน MapReduce ของ Cloud โดยใช้ข้อมูลการวิ่งของรถ Taxi จำนวน 11,800 คัน ผลจากการศึกษาพบว่า การใช้งาน Parallel Computing บนระบบ Distributed Applications สามารถทำได้โดยไม่สูญเสียความแม่นยำเมื่อเทียบกับระบบปกติ (81.66% vs 83.52%)

3.4 การทำนายอื่น ๆ ที่ใช้ข้อมูลการจราจร

ในฐานะข้อมูลของการคมนาคมของ United Kingdom (UK) หรือ อังกฤษ ได้ทำการเก็บข้อมูลการเกิดอุบัติเหตุ และข้อมูลการเสียชีวิตจากอุบัติเหตุ นอกเหนือจากข้อมูลสภาพการจราจรแบบปกติ โดยได้นำเอาข้อมูลนี้มาใช้งานเพื่อหาสาเหตุสำคัญของการเสียชีวิตจากการใช้งานท้องถนน ด้วย Data Mining โดยใช้ Algorithm Naive Bayes พบว่า พฤติกรรมมนุษย์เกี่ยวข้องกับการเกิดอุบัติเหตุมากที่สุด ซึ่งความแม่นยำของการทำนายอยู่ที่ประมาณ 84% จึงเป็นไปได้ที่จะใช้ข้อมูลนี้ร่วมการทำนายการจราจรและเตือนภัยในการใช้งานถนนที่มีความเสี่ยงสูงได้ [10]

4. เปรียบเทียบการใช้งาน Big Data รูปแบบต่าง ๆ

Big Data นั้นจะหมายถึงทุกข้อมูลที่มีความหมายตาม 3V ที่ได้กล่าวในหัวข้อที่ 2 และไม่ได้หมายถึงรูปแบบการจัดเก็บข้อมูล หรือลักษณะการจัดเก็บข้อมูล ดังนั้นการจัดเก็บและใช้

งาน Big Data จึงทำได้หลากหลายรูปแบบ ซึ่งในแต่ละรูปแบบ เปรียบเทียบการแต่ละการศึกษาว่าได้ผลลัพธ์อย่างไร ดังที่แสดง
 ก็ได้ให้ผลลัพธ์ที่แตกต่างกันออกไป ดังที่ได้กล่าวถึงในหัวข้อที่ ในตารางที่ 1
 3 เพื่อให้เกิดความเข้าใจที่ง่ายขึ้นจึงได้จัดทำตารางเพื่อ

ผู้ทำการศึกษา	เทคโนโลยี	อัลกอริทึม	ข้อมูล	สิ่งที่ศึกษา	ผลการศึกษา
Yin-Zhen Yan และคณะ (2015)	Big Data, Hadoop, Spark	DBSCAN, K- Means, Fuzzy C- Means	ข้อมูลการวิ่งรถบัส มากกว่า 1000 คัน	Real-time traffic situation โดยข้อมูลจาก Open data, GPS, GPRS และ Cloud	Spark ทำงานได้เร็วกว่า Hadoop
DAWEN XIA และคณะ (2015)	Big Data, Hadoop	k-nearest neighbor บน MapReduce	ข้อมูลการเดินทาง ของ Taxi 12,000 คัน	ใช้ parallel k-nearest neighbor optimization (parKNNO) ในการ แบ่งกลุ่มบน MapReduce	ผลลัพธ์ที่ได้ดีกว่า โดยรวมเทียบกับวิธีอื่น 90.07%
Hamzah Al Najada และ Imad Mahgoub (2016)	Data Mining, H2O, WEKA	Naive Bayes	ข้อมูลการเกิด อุบัติเหตุ และข้อมูล การเสียชีวิตจาก อุบัติเหตุ	ทำนายการเกิดอุบัติเหตุ	พฤติกรรมมนุษย์ เกี่ยวข้องกับการเกิด อุบัติเหตุมากที่สุด
Dongbo Zhang และคณะ (2016)	Cloud Computing Environment	Map-matching algorithm บน MapReduce	ข้อมูลการวิ่งของรถ Taxi จำนวน 11,800 คัน	Near real-time Map- matching algorithm สำหรับ FCD ขนาดใหญ่	สามารถใช้งาน Parallel Computing ได้โดยไม่ สูญเสียความแม่นยำ
Seungwoo Jeon และคณะ (2014)	Big Data, RHive, Hadoop	Algorithm ใหม่ สำหรับ prediction map	ข้อมูลการจราจร นานกว่า 4 เดือน จำนวน 8,769 ไฟล์	วิเคราะห์รูปแบบการจราจร เพื่อคาดการณ์สภาพ การจราจร	ให้ค่า R-Squared ซึ่งมี ความถูกต้องเกินกว่า 75%
Florin Schimbinschi และคณะ (2014)	Big Data, Machine Learning	Algorithms ของ Machine Learning จำนวน 8 ตัว	ข้อมูลการจราจร เวลานานกว่า 6 ปี	นำ Big Data และ Machine Learning มาใช้ในการทำ Traffic Prediction	Algorithms บางตัวไม่ เหมาะสำหรับ Real- Time
Yisheng Lv และคณะ (2015)	Stacked Autoencoder (SAE)	Deep-learning- based	ข้อมูลจราจรทุก 30 วินาทีใน California	นำ Deep architecture model แบบ SAE มาใช้ใน การทำ Traffic Prediction	Model ในการศึกษา ให้ประสิทธิภาพสูงกว่า แบบอื่น
Zhongsheng Hou และ Xingyi Li (2016)	Big Data	Machine Learning Hybrid algorithms	ข้อมูลการจราจร แบบ Real-Time จำนวน 82 วัน	ทำ Traffic Prediction ทั้ง แบบ short-term และ long- term	Repeatability Based Long-term Forecasting มีความแม่นยำประมาณ 97%

ตารางที่ 1 เปรียบเทียบการศึกษาที่ใช้ Big Data ร่วมกับข้อมูลการจราจร

5. อภิปรายผลการศึกษา

การทำ Traffic Prediction โดยใช้ Big data จะสามารถทำนายโดยนำข้อมูลทั้งแบบ Real-Time และแบบประวัติมาร่วมในการทำนาย ทำให้โอกาสที่ความแม่นยำของการทำนายจะถูกต้องมีประสิทธิภาพสูงขึ้น และ Algorithms ที่ใช้สำหรับการ Prediction มีผลต่อความเร็วในการคำนวณและความแม่นยำในการทำ Traffic Prediction ทำให้บาง Algorithms ไม่สามารถใช้ในการทำ Real-Time Prediction ได้

ในการใช้งาน Hadoop สำหรับการทำ Traffic Prediction การใช้งาน MapReduce และการ Clustering มีผลต่อความแม่นยำของการทำนายและความรวดเร็วของการทำนายแบบ Real-Time โดยข้อมูลที่เข้าสู่การทำนาย นอกจากจะต้องมีขนาดข้อมูลที่ใหญ่แล้ว ยังต้องเลือกเฉพาะข้อมูลที่เกี่ยวข้องกับการทำ Traffic Prediction เพื่อความรวดเร็ว ลดการใช้งานทรัพยากร และเพิ่มความแม่นยำของการทำนายเพราะไม่มีข้อมูลที่ไม่เกี่ยวข้องเข้ามารวมนำในการทำนายด้วย

6. สรุปการศึกษา

การนำเอา Big Data หรือเทคโนโลยีด้าน Data Mining และ Machine Learning มาใช้งานเพื่อทำนายการจราจรจะมีประสิทธิภาพสูงรวมถึงสามารถทำนายแบบ Real-Time ได้ตามประสิทธิภาพของตัว Algorithms ที่ใช้งาน

การทำ Parallel Computing สามารถทำได้บน Hadoop หรือ Environment อื่น ๆ ซึ่งใช้เทคโนโลยี Big Data ในการเก็บข้อมูล ทำให้การคำนวณข้อมูลสามารถทำกับข้อมูลที่มีขนาดใหญ่มากขึ้นได้โดยไม่สูญเสียประสิทธิภาพ เมื่อเทียบกับการคำนวณแบบปกติ ความเร็วและความแม่นยำของการทำ Traffic Prediction นั้นขึ้นกับ Algorithms ที่ใช้ ดังนั้นอาจจำเป็นต้องทำ การพัฒนา Algorithms ที่มีความเร็ว, ความแม่นยำ และ

International Symposium on Quality of Service (IWQoS), 2016, pp. 1–2.

- [2] Y. Liu, X. Chen, C. Chen, and X. Guan, “Traffic big data analysis supporting vehicular network access recommendation,” in *2016 IEEE International Conference on Communications (ICC)*, 2016, pp. 1–6.
- [3] F. Schimbschi, X. V. Nguyen, J. Bailey, C. Leckie, H. Vu and R. Kotagiri, “Traffic forecasting in complex urban networks: Leveraging big data and machine learning”, *2015 IEEE International Conference on Big Data*, pp. 1019-1024, Nov 2015.
- [4] Y. Lv, Y. Duan, W. Kang, Z. Li, and F. Wang, “Traffic Flow Prediction With Big Data: A Deep Learning Approach”, *Transactions on Intelligent Transportation Systems*, Vol. 16, No. 2, pp. 865-873, April 2015.
- [5] S. Jeon, B. Hong and B. Kim, “Big Data Processing for Prediction of Traffic Time Based on Vertical Data Arrangement”, *2014 IEEE 6th International Conference on Cloud Computing Technology and Science*, pp. 326-333, Dec 2014.
- [6] D. Xia, H. Li, B. Wang, Y. Li, and Z. Zhang, “A MapReduce-Based Nearest Neighbor Approach for Big-Data-Driven Traffic Flow Prediction”, *IEEE Access*, Vol. 4, pp. 2920-2934, Jun 2016.
- [7] Y. Yan, R. Liu, C. Yang, and S. Chen, “Cloud City Traffic State Assessment System Using a Novel Architecture of Big Data”, *International Conference on Cloud Computing and Big Data*, pp. 252-259, Nov 2015.
- [8] Z. Hou, and X. Li, “Repeatability and Similarity of Freeway Traffic Flow and Long-Term Prediction Under Big Data”, *Transactions on Intelligent Transportation Systems*, Vol. 17, No. 6, pp. 1786-1796, June 2016.
- [9] D. Zhang, Y. Shou and J. Xu, “The modeling of big traffic data processing based on cloud computing”, *2016 12th World Congress on Intelligent Control and Automation (WCICA)*, pp. 2394-2399, Jun 2016.
- [10] H. Al Najada and I. Mahgoub, “Big vehicular traffic Data mining: Towards accident and congestion prevention”, *2016 International Wireless Communications and Mobile Computing Conference (IWCMC)*, pp. 256-261, Sep 20

เอกสารอ้างอิง

- [1] M. Zhu, X. Y. Liu, M. Qiu, R. Shen, W. Shu, and M. Y. Wu, “Traffic big data based path planning strategy in public vehicle systems,” in *2016 IEEE/ACM 24th*