วิทยานิพนธ์นี้เป็นการนำเสนอวิธีการกรองข้อมูล GPS ที่ผิดปกติหรือไม่สามารถนำมาช่วยในการประเมินสภาพจราจรได้ออกไป ซึ่งข้อมูลตัวอย่างในการทดสอบจะใช้ข้อมูลจากรถแท็กซี่ที่ติดตั้ง GPS ในบริเวณกรุงเทพมหานครและปริมณฑลซึ่งมีอยู่ประมาณ 3000-4000 คัน อัตราการส่งข้อมูล GPS มายังศูนย์วิทยุคือ 3 นาที 10 วินาที ซึ่งจากการวิเคราะห์ข้อมูล GPS ทั้งหมดพบว่ามีข้อมูลผิดปกติแบ่งออกเป็นสี่ประเภทหลักๆได้แก่ ข้อมูลจากรถแท็กซี่ที่ขับวนเพื่อรอผู้โดยสาร, ข้อมูลจากรถแท็กซี่ที่จอดรอผู้โดยสาร, ข้อมูลจากรถแท็กซี่ที่มีทิศทางเคลื่อนที่ของรถผิดไปจากรถคันอื่นในบริเวณนั้นและข้อมูลจากรถแท็กซี่ที่มีความเร็วที่คลาดเคลื่อนเนื่องจากข้อจำกัดของเครื่องรับ GPS และเพื่อเป็นการยืนยันว่าการกรองข้อมูลที่ผิดปกติออกไปแล้วจะทำให้ระบบมีประสิทธิภาพมากขึ้นจึงได้มีการวัดค่าการกระจายของข้อมูลด้วยสัมประสิทธิ์การแปรผันซึ่งค่าสัมประสิทธิ์การแปรผันที่ลดลงนั้นในทางสถิติกล่าวไว้ว่าหากข้อมูลใดๆที่มีค่าสัมประสิทธิ์การแปรผันน้อยนั่นหมายถึงค่าแนวโน้มเข้าสู่ศูนย์กลางของข้อมูลชุดนั้นจะเป็นตัวแทนที่ดีและน่าเชื่อถือกว่าข้อมูลชุดที่มีค่าสัมประสิทธิ์การแปรผันมากกว่า ในการทดลองพบว่าค่าสัมประสิทธิ์การแปรผันของข้อมูลหลังกำจัดความผิดปกติออกไปมีค่าลดลงสูงสุดร้อยละ 11.98 ทั้งนี้ขึ้นอยู่กับแต่ละพื้นที่ หลังจากการวิเคราะห์และแยกข้อมูลผิดปกติออกไปแล้วจะได้ทำการจำลองระบบการกรองข้อมูลด้วยนิวรอนเน็ตเวิร์คโมเดลเพื่อให้ระบบเรียนรู้การคัดกรองด้วยตนเองเพื่อเป็นการเพิ่มประสิทธิภาพและความเร็วในการกรองข้อมูลของระบบซึ่งผลที่ได้เมื่อเปรียบเทียบกับการกรองด้วยการวิเคราะห์ข้อมูลแล้วให้ค่าความถูกต้องสูงถึงร้อยละ 83.19

218461

This research examined an algorithm for filtering unusual GPS data from multiple GPS-equipped cars. The GPS data was periodically transmitted to a base station via a radio link every 3 minutes and 10 seconds by 3000-4000 taxis around Bangkok and suburb. From manually analysis of the data, there are four main problems that affect the accuracy of traffic estimation system: (1) taxi driver drive round one place for a long time, (2) taxi waits for passengers, (3) the direction of car does not flow along the street, (4) the measurement error of the GPS receivers. If these problems are filtered out before being sent to the traffic estimation system, the system will be more accurate. To validate the performance of the filtered GPS data, the coefficient of variation (C.V.) was applied to measure the variation of two groups of data: non-filtered data and filtered data. Low value of C.V. indicates consistent data, thus more accurate. The experiment indicated that the algorithm can reduce the C.V. of the data by an average of 11.98 percent depending on the locations. Then, a neural network model was implemented to model the filter process. The results showed 83.19 percent accuracy compared to the rule-based method.