Measurement and Evaluation of Transport Network Performance: A Data Driven Approach

Tao FENG, Qi HAN

Department of the Built Environment, Eindhoven University of Technology, The Netherlands

Abstract: This paper attempts to propose a method to measure and evaluate the network performance by identifying regions with serious mismatch problems between travel demand and transport network services. Performance measures that include temporal, spatial and travel time reliability are generated from travel information in the form of GPS data. Results show that the proposed approach produces better and more accurate results than the ones in the literature. The results of standard deviation and coefficient variation of the identified regions based on the variation of average speed, travel time, and the amount of trips yield a higher reliability for the proposed approach.

Keywords: Network Performance; Network Reliability; GPS; Big Data

1. INTRODUCTION

The trend of population mitigation has never stopped. As predicted by WHO, in 2050, around 70% of the world’s population will live in cities. The enormous population moving into urban areas will lead to serious consequences on future infrastructures, e.g. housing, roads, parking, etc. Apart from available resources, over-populating in an already “dense” country will also cause social problems, e.g. high cost of housing and job availability. The decision in the trend of residential and work choices will highly interact with the mobility choices, which affects the productivity and the wellbeing of people. The reliability of transportation network plays a vital role in planning and individuals’ choices.

Network performance can be analyzed in different ways and from different perspectives. Researchers, traffic planners and policy makers have showed much interest in the evaluation of travel time reliability (Mahmassani et al., 2013). However, traditional network reliability evaluation based on models suffers from the availability of fine-grained data support. In recent, with the popularity of emerging technology like GPS in new travel survey methods and the increasing availability of open data source, new opportunities of a data driven approach become promising to better understanding of network performance and improving the analysis approach.

GPS data has been well accepted in transportation due to its high accuracy, high resolution of spatial and temporal information and light burden for respondents in assistance of traditional travel surveys. By utilizing GPS recording, the movements of vehicles or individual persons can automatically be tracked. Researchers have used GPS data to represent travel patterns across a network by constructing the travel demand. The origin-destination (OD) matrix built using GPS data can have multiple
facets of information, including the time of day, day of week etc., providing necessary information for travel time reliability analysis.

This paper, therefore, aims at developing a method to evaluate the performance of transport networks using performance measures that relate to traffic congestion. These measures are formulated using the extracted information from GPS data. The proposed method will provide decision makers with a better understanding of current travel demand and transport infrastructure problems as well as help them in creating effective solutions.

The remainder of this paper is organized as follows. Section 2 will review the relevant works in network performance evaluation. Section 3 will introduce the proposal approach. Data and results will be presented in Section 4 and Section 5, respectively. The paper will be summarized and concluded with some future research directions in Section 6.

2. LITERATURE REVIEW

A transport network can be described as the basis of all movement within a certain area, and is basically a network of infrastructure designed to realize vehicular movement or flow of some commodity from one location to another. The performance of transport networks can be measured in different ways depending on different characteristics of the network itself. Several advised steps have been proposed, identifying measures, strategies, assess strategies using measurements and defining the necessary data.

A number of existing studies have been conducted to better understand transportation networks and evaluate their performance. In general, three particular performance measures from temporal, spatial and accessible aspects have been proposed (Cui, et al., 2016-A). The temporal travel efficiency measures focus on the temporal aspect of the travel in the network and includes travel speed, travel times and congestion situations. The spatial travel efficiency measures compares the ratio between the actual travel distance between a pair of origin and destination in the network and the Euclidean distance of these two locations. For this measure, circuity or route directness normally plays a vital role. The accessibility measures take into account both the spatial and temporal travel efficiencies as well as the distribution of land-use and activity locations across the urban area. These measures investigate the ease and extent to which the combination of land-use and transport systems enable commuters to reach activities and destinations by means of certain transport modes.

Additionally, the literature also revealed the increasing importance of travel time reliability as a network performance descriptor for both travelling public and traffic managers and policy makers (Mahmassani, Hou and Saberi, 2013). Travel time describes the time it takes an average commuter to move from one location to the other location (Elefteriadou and Cui, 2005). Travel time reliability is a performance measure which indicates how dependable the travel time on a given transport network is (Lyman, 2007). Although this measure’s definition can vary in different contexts, each definition closely relates to the variation of travel time.

Apart from the performance measurement itself, studies also differ according to different research objectives. For this reason, most studies have used different transport network performance strategies that represents the methodology researchers have used to evaluate the transport network performance.
Table 1 Identified performance strategies and measures

<table>
<thead>
<tr>
<th>Performance strategy</th>
<th>Performance measures</th>
<th>Assessment (found in literature)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transport network conditions</td>
<td>Temporal efficiency measures</td>
<td>Origin-Destination (OD) prediction (Lu and Li, 2014)</td>
</tr>
<tr>
<td></td>
<td>Temporal and spatial efficiency measures</td>
<td>Identifying mismatch between travel demand and transport network services (Cui, et al., 2016-A)</td>
</tr>
<tr>
<td></td>
<td>Temporal, spatial and accessibility efficiency measures</td>
<td>Detecting network accessibility (Cui, et al., 2016-B)</td>
</tr>
<tr>
<td>Transport network reliability</td>
<td>Travel time reliability</td>
<td>Assessing transport network reliability and vulnerability (Oliveira, Portugal and Junior, 2016)</td>
</tr>
<tr>
<td></td>
<td>Travel time reliability</td>
<td>Two fluid model strategy (Hong S., Lee, Bong and Kho, 2005)</td>
</tr>
<tr>
<td></td>
<td>Travel time reliability</td>
<td>Connecting travel time reliability and the Network Fundamental Diagram of Traffic Flow (NFD) (Mahmassani, Hou and Saberi, 2013)</td>
</tr>
</tbody>
</table>

Two main strategies in the literature have been proposed, emphasizing on the evaluation of network condition and network reliability. Evaluating transport network conditions focuses on the macroscopic characteristics of a transportation network and the travel demand of its users. The method is essentially designed for detecting congestion areas within the network. As this strategy can be assessed with different measures, the detection process will vary depending on these applied measures. In the literature three studies were identified which contained three different assessments of this strategy.

The other strategy, evaluating transport network reliability, mostly targets to the application in large cities where space is becoming increasingly intense (Oliveira, Portugal and Junior, 2016). The network reliability here represents an aspect of congestion that is a consequence of unpredictable travel conditions, rather than everyday delays. The main performance measure which can be used to assess this strategy includes the travel time reliability. Although the two strategies use completely different measures and investigate different aspects of the transport network, ultimately both approaches are to solve congestion issues, detecting congestion areas and congestion road segments on a transportation network. In this sense, the two strategies are perfect ingredients for developing a suitable methodology for this paper.

Regarding the use of real data observation, Lu and Li (2014) used temporal travel efficiency measures to build a travel demand model directly based on an enormous amount of GPS data. Authors used temporal efficiency measures to build a time dependent Origin-Destination matrix and matched it into statistical indicators. Although the objective is interesting, they consider only one type of performance measures. Combining or adding more performance measures seems promising to make the approach more accurate and more reliable. In a recent work conducted by (Cui, et al., 2016-A; 2016-B) who used taxi data, network performance were measured by identifying the mismatches between travel demand and transportation network services. Two performance measures, including temporal efficiency and spatial travel efficiency, were used to assess network performance.

From the perspective of accessibility evaluation, (Oliveira, Portugal, & Junior, 2016) conducted a study on two road network performance attributes, network
reliability and vulnerability. The objective of this study is to compare the network reliability and vulnerability in complex networks to ultimately get more insight on the way they are related to each other. Within the area of transport network reliability the author used travel time reliability measures. The results revealed no correlation between reliability and congestion. Instead it was discovered that reliability is related to increasing road saturation or level of congestion. In the latter, the study revealed the nature of travel time reliability. In this way more insight was obtained on how travel time reliability could be applied as a performance measure for the assessment of the transport network performance.

Furthermore, (Mahmassani, Hou and Saberi, 2013) used the Network Fundamental Diagram (NFD) of traffic flow to link NFD and travel time reliability performance, and establish a bridge between network traffic flow theory and travel time reliability. This bridge extend travel time variability models from the link and path levels to the network level. Results show that travel time variability increases with network density and flow rate. Although the work provides a macro level analysis for analyzing transport network reliability, the conclusion was drawn based on the data from a single day.

Hon, Lee, Bong and Kho (2005) analyzed the reliability of transportation networks using the two-fluid model. As most reliability studies, this model used travel time reliability measures which included travel time and stop and running time per unit distance. To obtain values of these measures researchers proposed to build the input data for the two-fluid model using GPS devices. The results of this study proved that GPS data could indeed be used to utilize the two-fluid model and provide reliable results. To filter the best fitting approaches, previous methods are assessed on two criteria, applicability and feasibility. The applicability of each method highly depends on its relevance to the research objective. Additionally, the success of a specific method depends on the available data and how it can be utilized. Based on the assessment of the approaches using the following criteria, 1) a methodology based on a combination of network reliability and network conditions strategies, 2) the performance measures include temporal, spatial and travel time measures, and 3) the data which will be used is individual inferred GPS data. In summary, in the context that policy makers can analyse performance of a transportation network easily and at low cost, this paper extends the methodology in the aspects of higher accuracy using more accurate data and incorporating a travel time reliability analysis.

3. METHODOLOGY

Various types of data can be used to generate the measures including travel time, travel date and type of day, timestamp information. For spatial efficiency measures, location information is important including the place of origin (start of the trip) and the place of destination (end of the trip). With this information, the route directness can be derived. The travel time reliability measures are much similar to the temporal efficiency measures. However the difference is that travel time reliability measures focus on the variation of travel time rather than travel time itself.

Although GPS data collected using new technologies are full of rich information, the necessary information for the performance measures are not directly available. Imputation of GPS data is generally needed.
To analyze the network performance using GPS data, the methodological process is divided into several analytical stages. It includes processing of GPS data, modelling region-wide travel patterns, generating performance measures, detecting regions with serious mismatch problem and method evaluation. A flowchart of the analysis process is shown in Figure 1.

### 3.1 Processing of GPS data

Using GPS data, the driving speed at each point, represented by \( \text{speed}_i \), can be calculated by dividing the direct distance two following points with the time duration of moving between these two points. Calculating this speed is important as it provides the first data filter. Data (cars) with a speed higher than 130 km/h are filtered out to exclude invalid or erroneous data.

\[
\text{speed}_i = \frac{HD_{(i+1,i)}}{t_{i+1} - t_i} \quad (1)
\]

where, \( HD_{(i+1,i)} \) is the direct distance of two following points, \( t_{i+1} - t_i \) is the time it took to move from the first point to the following one.

The direct distance of two following points can be calculated in different ways. Existing literature normally use the Euclidean distance because of its easy implementation and straightforwardness. However, this distance represents the distance of the straight line between two coordinates also known as the “shortest path” distance, and not the real distance. For this reason, the real distance is used in this study, which is extracted from a Haversine equation and will be represented as the Haversine Distance (HD). A Haversine formula is journey seq. This formula was extracted from (Feng et al., 2011; Feng and Timmermans, 2014; Feng and Timmermans, 2016) and is illustrated in Equation 2.
\[
HD(l_{i+1}, l_i) = 2 \times R \times 
\left( \sqrt{\sin^2 \left( \frac{y_{i+1} - y_i}{2} \right) + \cos(y_i) \times \cos(y_{i+1}) \times \sin^2 \left( \frac{x_{i+1} - x_i}{2} \right)} \right)
\]

where, \( R \) is the radius of the sphere, \( l_i = \{x_i, y_i\} \) is the coordinates of first point, \( l_{\text{origin}} = \{\text{lonfrom, latfrom}\} \), \( l_{i+1} = \{x_{i+1}, y_{i+1}\} \) is the coordinates of following point, \( l_{\text{destination}} = \{\text{lonto, latto}\} \).

Apart from calculating “HD”, “speedi” and removing data with a higher speed, than 120 km/h, it is also important to clean the GPS data in this first step of the analysis. Raw GPS data usually contains random errors. In the current paper, the data is cleaned by removing data with latitude and/ or longitude zero.

### 3.2 Modeling region wide travel patterns

For the construction of the travel pattern analysis, the trip characteristics of travel speed and route directness are needed. For the purpose of spatial and temporal measurements, the data needs to be divided into smaller study cells and different time.

#### Trip indicators

In order to analyze passenger travel patterns, two indicators/ characteristics are presented. The first one describes the actual travel speed which is calculated by dividing the actual distance of the journey by travel time.

\[
v_k = \frac{\text{Link}(l_o, l_d)}{(t_d - t_o)}
\]

where, \( v_k \) is the travel speed each trip \( k \), \((t_d - t_o)\) is the travel time from origin to destination. \( \text{Link}(l_o, l_d) = \sum HD(l_{i+1}, l_i) \) represents the sum of the actual travel distance of each trip, and it is estimated with the sum of the Haversine distance between each two following points along the path.

The second indicator represents the route directness or circuity. The circuity or route directness of travel paths, describe the ratio between the actual travel distance between a pair of origin and destination in the network and the Haversine distance of two locations.

\[
r_k = \frac{\text{Link}(l_o, l_d)}{\text{ED}(l_o, l_d)} = \frac{\sum ED(l_{i+1}, l_i) }{ED(l_o, l_d)}
\]

where \( r_i \) is the route directness each trip \( k \), \( \text{ED}(l_o, l_d) \) represents the Euclidean distance between the origin and destination, as show in Equation (5).

\[
ED(l_{i+1}, l_i) = \sqrt{(y_{i+1} - y_i)^2 + (x_{i+1} - x_i)^2}
\]

where \( l_i = \{x_i, y_i\} \) is the coordinates of point \( i \), then \( l_{\text{origin}} = \{\text{lonfrom, latfrom}\} \). \( l_{i+1} = \{x_{i+1}, y_{i+1}\} \) represents the coordinates the following point \( i+1 \). \( l_{\text{destination}} = \{\text{lonto, latto}\} \).
Table 2. Performance indicators

<table>
<thead>
<tr>
<th>Measure</th>
<th>Poor performance</th>
<th>Good performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speed (V)</td>
<td>Low value</td>
<td>High value</td>
</tr>
<tr>
<td>Route directness (R)</td>
<td>High value</td>
<td>Low value</td>
</tr>
<tr>
<td>Travel time (TT)</td>
<td>High value</td>
<td>Low value</td>
</tr>
</tbody>
</table>

Table 3. Network performance based on combinations V and R

<table>
<thead>
<tr>
<th>V</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>Low</td>
<td>High</td>
</tr>
</tbody>
</table>

These indicators are important as they are used in the last step of travel pattern modelling to generate two performance measures. These include the average speed per trip, represented by ‘V’, and the average route directness per trip, represented by “R”. These measures will be described in the following sections.

**Spatial partition and temporal partition**

Using a grid-based method, the study area is divided into smaller regions which is in general represented with a cell ID or region ID. The grid is generated as a layer and each region is set at 1 km² in size. Using spatial analysis function, GPS are connected with grids. Thus, the path followed by each individual can be matched with regions. This is fundamental for detecting the regions with serious problems because the regions where user experience certain problems can be then identified at the grid level.

Apart from the spatial partition, each trip is also divided into different time periods of a day. This is represented by TimeP. In this analysis, a certain time period for TimeP is chosen in which traffic congestion commonly occurs. Additionally, as suggested by the literature (Cui, et al., 2016-A), the type of day is also distinguished. Previous studies revealed that travel behavior and travel patterns generally differ across various time periods of a day and also across different type of days. The different types of day are represented by DayT.

### 3.3 Generating performance measures

A commuter travel pattern is described as: OD (Reg (x₀, y₀), Reg (xₐ, yₐ), TimeP, Day, DayT). Each model represents all the trips that originate from region Reg (x₀, y₀) and end in region Reg (xₐ, yₐ). With this matrix two transport performance measures are derived, including the average travel speed per trip, represented by V and the average route directness per trip, represented by R.

**Generating measures**
V and R are performance measures for spatial and temporal travel efficiencies. Additionally, the ratio of both, V/R, could be used as a third measure. This measure assesses transport performance by taking both temporal and spatial travel efficiency measures at the same time. Within this context, a low value means low travel speed and poor connection between two locations. However, because it is not used to detect problem regions, it is left out of the analysis.

The following formulas are used for generating these measures:

\[
V = V(\text{Reg}(x_o, y_o), \text{Reg}(x_d, y_d), \text{TimeP}, \text{Day}, \text{DayT}) = \frac{\sum_{k=1}^{M} v_k}{M} \tag{6}
\]

where \( V \) is the average travel speed over all trips, \( k \) means for every single trip separately. \( M \) is the total number of trips to and from a certain region. \( \text{Reg}(x_o, y_o), \text{Reg}(x_d, y_d) \) represents the origin and destination within the region. \( \text{TimeP} \) is the time of day, \( \text{Day} \) is the date and \( \text{DayT} \) is the type of day (e.g. weekend, weekday or holiday).

\[
R = R(\text{Reg}(x_o, y_o), \text{Reg}(x_d, y_d), \text{TimeP}, \text{Day}, \text{DayT}) = \frac{\sum_{k=1}^{M} r_k}{M} \tag{7}
\]

where \( R \) is the average route directness over all trips, \( r_k \) is the route directness.

**Travel time reliability**

The last performance measure is the travel time reliability. Ideally, the distribution function of travel time should be estimated using historical data. However, due to the fact that GPS data was used, the travel times were easily extracted. Travel time is measured in (travel) time per unit travelled distance. The travel time is represented by \( TT \) and is calculated as below.

\[
TT = \frac{t_d-t_o}{HD(t_o, t_d)} \tag{8}
\]

where, \( TT \) is the travel time per unit travelled distance between origin and destination.

### 3.4 Detecting problem regions

To detect the daily transport problems, the performance measures \( V, R \) and \( TT \) are tested against threshold values. The first step in doing so includes identifying regions with higher volumes of traffic demand. In this sense, the amount of trips \( M \) to and from a certain region need to be higher than a certain amount. The threshold is represented by \( T_M \). A higher volume of travel demand in the morning in residential or employment areas make it more likely the candidates for traffic congestion. Additionally, the large number of passenger trips increases the accuracy of the estimated parameters \( V, R \) and \( TT \), which can better represent the general traffic conditions.

After that the regions with high travel demand are detected, \( V \) and \( R \) are used to measure the temporal and spatial travel efficiencies within these regions. Here, regions with \( R \) higher than the threshold \( (T_R) \) or \( V \) lower than the threshold \( (T_V) \), are defined as problematic regions with temporal or spatial travel efficiency problems. Additionally \( TT \) is also implemented to measure travel time reliability within the regions with high volume travel demand.
The regions with a higher travel time per unit distance than a threshold $T_{TT}$ are also defined as problematic regions. The regions with $V$ lower than $T_V$, $R$ higher than $T_R$ and $TT$ higher than $T_{TT}$, represent regions experiencing the worst problems. Within these regions commuters are forced to take detours for travel, while still moving at slow speed due to heavy traffic along these detours and taking a significantly longer time to reach their destination (Cui, et al., 2016-A).

In summary, the measures can be specified as below:

1) $V < T_V$ and $M > T_M$ (temporal performance measure)
2) $R > T_R$ and $M > T_M$ (spatial performance measure)
3) $TT < T_{TT}$ and $M > T_M$ (travel time reliability performance measure)

The worst regions are described as regions experiencing all three problems at once. Within these regions commuters are forced to take detours, where they also experience some level of traffic congestion. These regions are identified with:

1) $V < T_V$ and $R > T_R$ and $M > T_M$ (temporal and spatial performance measures)
2) $V < T_V$ and $R > T_R$ and $TT < T_{TT}$ and $M > T_M$ (temporal, spatial and travel time reliability performance measures)

### 3.5 Method validation

The proposed approach is validated comparing the results with results derived using an approach from the literature. As explained at the end of the literature review, existing works mainly used taxi GPS data. For this reason, the most effective validation is to compare the improved and extended approach proposed in this paper to the original one from literature. The method validation consists of four steps, which include generating synthetic taxi data, using different strategies, making statistical analysis and comparing results. For the purpose of method validation, only a fraction of the entire data set is used. Here a random week in one of the three months was chosen.

![Figure 2. Map of the study area with GRID](image)
To recreate the method from literature, first taxi trip data needs to be generated from the original raw GPS data. From literature it was revealed that taxi trip data includes large time gaps between recorded locations. For this reason the large portions of the raw data were randomly removed, creating a synthetic taxi GPS dataset.

After creating the taxi data, both approaches from the literature and the proposed approach were applied. After all measures are generated using both approaches, the problematic regions are detected. For the validation not only an empirical value will be used to detect the regions, but a variation of threshold values. By doing so, the quality of each method can be estimated using a statistical analysis on the results.

The variation of the threshold values is based on their empirical values found in the previous literature, $T_V = 20$ km/h, $T_R = 1.5$, $T_M = 20$ trips. In case of the threshold of $T_{TT}$, this value was not directly available in literature, but it is estimated using a statistical method. The literature suggest using the 95th percentile method, which estimates the 95% probability of the worst travel time that commuters may experience once per month, namely one of twenty week days. The value for this threshold is then estimated using existing data.

For the purpose of comparison, we extract the standard deviation and the mean. With these variables, the coefficient of variation can be determined. The coefficient of variation ($C_V$) is used to measure the consistency of the results across all validation experiments. A high $C_V$ value reflects inconsistency among the results within the group of a specific validation method. In turn, a lower $C_V$ value reflects a higher consistency in the results and thus indicate the most reliable method which yields the most accurate results. In other words, by determining the coefficient of variation for each method, the one producing the best quality result is revealed.

4. DATA

The data used in this study is part of a massive data collection, which was held during the entire year of 2014, in Eindhoven region. The survey was conducted in four waves. In each wave, around 1200 respondents were recruited to join the survey for three consecutive months. A web survey is used to collected the real activity-travel data. Figure 2 shows the study area and the grid created for performance measurements.

![Figure 3. A hypothetical journey from home to the supermarket](image-url)
The inferred GPS data are available in three parts, journey data, stage data and GPS trace data. The trace data is the raw GPS data of each commuter, containing all location data from a commuter from when it starts recording until it stops recording. In this paper, the GPS raw data represented according to the diary data by the origin and destination for each individual day.

A GPS trace is one bundle of recorded information, which contains timestamp, location coordinates, etc. Such a bundle of information is recorded almost every second. A journey is comprised of multiple journey sequences and each journey sequence contains one or more stages/trips (see Figure 3). Unfortunately, the journey data only describes origin and destination of each journey sequence. The trace data on the other hand split in separate user files. Each user file contained the GPS trace of that person during a months’ time.

5. RESULT

5.1 Description of raw data

Because the focus of this study is on the transport network performance, only the data related to car trips are used for the analysis. The information of transportation mode that is available in the stage files are used for the data selection in our analysis.

The grid was created for the Netherlands using MapInfo Professional and generating a total of 90,117 cells with a size of 1 km². After that, the coordinates were imported and each GPS point was linked to a region. This was done by joining the grid table with the data tables for March, April and May 2014. The travel patterns of the
commuters located in the Noord-Brabant province (9,432 regions) in May, 2014, as an example, are presented in Figures 4.

The temporal partition is based on dividing the trips according to the time of day (TimeP) and the type of day (DayT). The type of day was divided into weekends, weekdays and holidays. For this study, the focus is on weekdays to capture the peak use of road networks.

After the temporal and spatial partition, the performance measures were generated for all three months. These include the average speed per trip \( V \), the average route directness per trip \( R \), the average amount of trips and the average travel time per kilometer per trip \( TT \).

Results of average speed are shown in Figure 5-1. It can be seen the average speed is very low during the morning peak. In April and May, the average speed between 7.00 AM and 8.00 AM is close to zero. This could be caused by traffic jams occurring in some parts of the study area.

Figure 5-2 presents the average route directness \( R \) per trip for the study period. It shows that the route starts increasing until approximately 7:30AM and then stays constant or decreases towards the end of the morning peak 9.00AM.

Figure 5-3 presents the average number of trips during the study period. Here it can be seen that, similar to the route directness, the amount of trips increase until 7:30 and then decrease towards the end. Between 7.00AM and 7:30 AM, the number of trips reaches the peak level.

Figure 5-4 presents the average travel time per km per trip. It shows that during the morning peak, average travel times increase until 7:45 AM- 8:05 AM and then start decreasing towards the end of the morning peak. Within this time period, commuters experience worst network performance.

5.2 Detection of problematic regions

Regarding the threshold, the empirical values derived from existing literature are used. \( TT \) is estimated using the percentile statistical method which estimates the probability of the worst travel time that commuters may experience once per month in percentages.
In the literature 95 percentile is common choice, but in this paper the 75% and 90% were also estimated. The reason for this is that the 95th percentile describes a kind of worst case as it only shows the 95% probability of commuters experiencing the worst travel time once per month. Although the 95% has been applied empirically, the data output shows that the 90% is much closer than the mean value. For this reason the threshold $T_T$ will be set at 500 seconds per. This translates into an approximate $T_T$ of 8 min/km. This means that during this time commuters have driving speed of 7.5 km/h, indicating that a $T_T$ of 500 seconds per kilometer a good threshold value.

Figures 6 present the number of detected regions that experience the predefined problem indicators in values and percentages. Here the total number of regions is represented by “Regions total”. Five indicators are used. The paired combination of indicators include:

- $V < T_V$ and $M > T_M$, represented with VM
- $R > T_R$ and $M > T_M$, represented with RM
- $TT > T_{TT}$ and $M > T_M$, represented with TTM

The worst regions are described as regions experiencing all three problems at once. Within these regions, commuters are forced to take detours, where they also experience some level of traffic congestion. These regions are identified based on:

- $V < T_V$ and $R > T_R$ and $M > T_M$, represented with VRM
- $V < T_V$ and $R > T_R$ and $TT > T_{TT}$ and $M > T_M$, represented with VRTTM

It can be seen that 1261 regions were used for traveling. Among the 1261 regions almost half experiences mismatch between travel demand and the ability to actually travel (transport network services). In fact a total of 592 regions experience all problems. This is 47% of all the regions commuters travelled from and to. Combined with the travel time reliability only 240 regions were detected, which counts for 19% of all the regions. The latter suggests that although 49% experience serious mismatch problems, only 19% of the total or 41% of the mismatch regions experience the mismatch problems and have a poor network reliability. Important to mention is that some regions experience multiple problems, and are therefore detected separately for each problem.

Figures 7-9 illustrate the thematic maps of the detected regions based on the first three problem indicators. It is understandable that the last two indicators can be visualized by combining these maps. These thematic maps present a visualization of the problem regions in the study area. Separate thematic maps for each of the three months of 2014 can be found by following the same procedure. By combining the thematic maps the areas with serious mismatch problems (areas detecting using the last two indicators) can be extracted. These include the areas with more than one problem.

![Detected Regions March 2014](image)

![Detected Regions March 2014 (%)](image)

Figure 6. Detected problematic regions (March 2014)
Figure 7 presents the thematic map of areas where the average speed is smaller than 20km/h. The yellow parts show areas were commuters drive at an average speed between 7 and 20km/h. These areas are most likely the congested areas. The blue and grey parts show areas where cars move at very slow pace. For these areas further investigation is recommended. As these slow speeds could be caused by commuters having to decrease speed for traffic lights. However the grey parts could also be areas where severe congestion occurs on weekdays between 7.00 AM – 9.00 AM.

Figure 8 shows the thematic map of areas where the route directness is higher than 1.5. As mentioned before, this indicates that commuters are forced to take detours due to the congestion. From this map it can be concluded that the blue and grey areas suffer from congestion on weekdays between 7.00 AM - 9.00 AM. For this thematic map it is recommended to not take the yellow parts into consideration parts. The cause for the extremely high values for route directness could also be traffic light stops related. However in this case the values seem too high for it to be caused by severe congestion.

Figure 9 present the thematic map of areas where it takes commuters more than 500 second to move 1 km (8 min/km). From this map it can be concluded that areas with a higher travel time per unit distance than 900 sec/km should not be considered. These areas need further investigation. However, the blue and dark blue parts should be the focus. These areas are most likely caused by traffic congestion. In these areas it takes commuters more than 8 minutes to move 1 km. The blue areas are most likely areas with severe traffic congestion and the dark blue areas normal congestion on weekdays between 7.00 AM – 9.00 AM.

5.3 Method validation

For the method validation relative to the taxi data used in previous studies, only a small sample of the data is sufficient. Data in a random week during the morning peak, between 7:00 AM and 9:00 AM was selected.
The validation method consists of detecting regions with the base approach using the combined criteria, VRM. In order to examine the impacts of threshold setting on the result stability, each time the values were varied for each indicator while keeping other two constant (empirical values). This was performed three times where, each time, one threshold varies and the two others remain constant (empirical values). The threshold variation chart is presented in Table 5. Each value is basically a variation of the empirical value (either smaller or larger).

When applying the approach proposed in this paper (VRTTM), similar to the previous validation, here only $T_{TT}$ is varied. Other threshold are constant (empirical values). $T_{TT}$ is varied over (200, 400, 450, 500, 550, 600, 1000), with 500 being the empirical/estimated value for $T_{TT}$.

Results of the amount of identified regions are shown in Table 6. In case of the speed threshold, both approaches have the same amount of regions identified for each variation. However the proposed approach does identify three times the amount of regions compared to the base approach. This is perhaps because of the high-resolution of used GPS data.

Results by varying the threshold of route directness show that the proposed approach identifies more regions than the base approach. However, the proposed approach results into a much higher standard deviation. This may because the base approach measures the Euclidean distance that is the shortest distance or straight-line distance between two locations (and not the real distance based on the travelled path of the commuter). The route directness for the first approach will yield smaller values than the route directness for the proposed approach. However, this does not mean that the base approach produces better quality results, because for estimating route directness the actual travelled routes/paths should be used instead of the straight lines.

Following the same procedure by varying the trip threshold, the proposed approach identifies more regions than the base approach and a more constant amount of regions for each variation of the trip threshold. Moreover, the proposed approach

<table>
<thead>
<tr>
<th>Measures</th>
<th>Value variation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T_V$</td>
<td>(15, 18, 19, 20, 21, 22, 25)</td>
</tr>
<tr>
<td>$T_R$</td>
<td>(0.5, 0.7, 1.1, 1.5, 1.7, 1.9, 2.5)</td>
</tr>
<tr>
<td>$T_M$</td>
<td>(15, 18, 19, 20, 21, 22, 25)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 5. Validation variation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Measures</td>
</tr>
<tr>
<td>$T_V$</td>
</tr>
<tr>
<td>$T_R$</td>
</tr>
<tr>
<td>$T_M$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 6. Validation: Variation of $TV$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Min</td>
</tr>
<tr>
<td>-------------------------------</td>
</tr>
<tr>
<td>VariateTV proposed approach</td>
</tr>
<tr>
<td>VariateTV base approach</td>
</tr>
<tr>
<td>VariateTR proposed approach</td>
</tr>
<tr>
<td>VariateTR base approach</td>
</tr>
<tr>
<td>VariateTM proposed approach</td>
</tr>
<tr>
<td>VariateTM base approach</td>
</tr>
<tr>
<td>VariateTTTM proposed approach</td>
</tr>
<tr>
<td>VariateTTTM base approach</td>
</tr>
<tr>
<td>VariateTTT VRM_proposed_approach</td>
</tr>
<tr>
<td>VariateTTT VRM_base_approach</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 6. Validation: Variation of $TV$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Min</td>
</tr>
<tr>
<td>-------------------------------</td>
</tr>
<tr>
<td>VariateTV proposed approach</td>
</tr>
<tr>
<td>VariateTV base approach</td>
</tr>
<tr>
<td>VariateTR proposed approach</td>
</tr>
<tr>
<td>VariateTR base approach</td>
</tr>
<tr>
<td>VariateTM proposed approach</td>
</tr>
<tr>
<td>VariateTM base approach</td>
</tr>
<tr>
<td>VariateTTTM proposed approach</td>
</tr>
<tr>
<td>VariateTTTM base approach</td>
</tr>
<tr>
<td>VariateTTT VRM_proposed_approach</td>
</tr>
<tr>
<td>VariateTTT VRM_base_approach</td>
</tr>
</tbody>
</table>
leads to a smaller standard deviation than the base approach, indicating that the approach produces better quality results than the base approach, when the trip threshold values are varied.

The validation based on travel time reliability was done by identifying regions according to the travel time reliability, where all three problems occur. Results show that almost an even amount of regions are detected with both approaches. It also proves the validity of the proposed indicator as a performance measure. Both approaches have almost the same standard deviation. However, the proposed approach yields a slightly smaller value, indicating that the method proposed in this paper is better for detecting regions based on the travel time reliability. Also important to note is that both methods produced results that are very close to each other, proving the significance and good quality of the travel time reliability in performance measurement.

By varying the value of the threshold of travel time reliability, we found that similar amounts of regions were detected with the previous problem indicators. As this indicator identifies the worst problematic regions, it can be argued that regions with long $TT$ generally also experience other problems ($V < T_v$, $R > T_R$ and $M > T_M$). This has important meaning for traffic managers and policy makers as it proves that travel time reliability is one of the most effective performance measures. These findings is important as it reveals the regions with the worst network performance.

6. SUMMARY AND CONCLUSION

Measuring the road network performance and reliability has been important issue for planners to detect problematic regions. Although different measures have been proposed in the literature, the possibility of using big mobility data has been not well reported. This paper therefore presented such an approach to measure transport network performance using GPS data. In comparison with previous studies, the paper evaluates the network performance based on the defined performance measures and examines the sensitivity and variability of different threshold settings. Different measures including temporal and temporal performance and travel time reliability are incorporated into the existing measuring framework.

As demonstrated using the GPS data, the proposed measures can be applied to measure network performance. The travel time reliability is measured with the travel time per unit distance. Ideally this measure should be derived from historic data. However as this data is scarce and hard to obtain, the 95th percentile method as proposed in the literature was adopted and extended considering the 75th and 90th percentiles, which include statistically estimating the travel time reliability rather than calculating it or deriving it from historic data. Results showed that the 90th percentile value around 500 seconds per km was the closest to the mean value.

The combined indicator related to temporal, spatial and travel time performance measures and the one using travel time reliability both can detect regions with the most serious mismatch problem between travel demand and transport network services. However, each indicator identifies around the same amount of regions. This, along with the validation of the approach, proves that travel time reliability is an excellent overall performance measure. Moreover, from the results of the generated measures, high values of travel times occur around the same time as higher route directness values and higher trip counts. This is in line with the literature that travel time reliability is not necessarily related to congestion itself but to the level of congestion.

The validation procedure showed that the proposed approach produces better and more accurate results than the approach from the literature. This is proved by the fact
that the statistical analysis of standard deviation and the coefficient variation of the identified regions using the variation of average speed, travel time, and the amount of trips yield lower values relative to the ones obtained using the proposed method.

The proposed approach can be used for policy makers to identify regions with specific problems and/or experiencing multiple problems, and is potentially beneficial for individual commuters as an indication of dynamic network performance. However, further research on this topic is still necessary to find tailored sustainable solutions for specific problems. In this regard, the seriousness of the mismatch problem needs to be identified by calculating the occurrence of a specific problem in a certain area with probabilities. By determining the probability of a problem, traffic managers and policy makers can strategically target these areas a priority. It is also recommended to study the side effects of targeting these areas first, as this could cause the probability of certain problems occurring to decrease or even disappear in other areas.

REFERENCES


