 USING GPS DATA TO MEASURE THE PERFORMANCE AND RELIABILITY OF TRANSPORTATION SYSTEMS

1Eric Scharnhorst, 2Edward McCormack, 2Wenjuan Zhao

1Institute of Planning, Royal Danish Academy of Architecture
Philip de Lange Alése 10, Copenhagen, Denmark
1435 Copenhagen K
E-mail: eric.scharnhorst@karch.dk. Ph.:+45 50 15 96 50

2Washington State Transportation Center (TRAC)
University of Washington, Box 354802
University District Building
1107 NE 45th Street, Suite 535
Seattle, Washington 98105-4631

Global positioning system (GPS) devices are embedded in everything from wristwatches to freight trucks. The persistent stream of data generated by GPS devices can be tapped and then used to measure how different systems are performing.

This paper develops a new systematic methodology for identifying and ranking the bottlenecks based on data collected by commercial GPS devices mounted on trucks. This data is initially processed in Geographic Information System (GIS) software where it is assigned to a roadway network to provide performance measures for individual roadway segments. Then, the truck speed distributions of roadway segments are represented by either a unimodal or bimodal probability density function, thus proposing a new reliability measure. This new reliability measure is used to evaluate the roadway performance and to classify the performance into three categories: unreliable, reliably fast, and reliably slow. The bottlenecks are ranked basing on both the reliability and congestion measurements, and this method is applied to evaluate the performance of the entire Washington State roadway network, where it proved efficient in identifying and ranking truck bottlenecks.

The new methods of using GPS and GIS to measure and rank truck roadway performance in Washington State can be custom tailored and then reapplied to finer-grained networks, such as the pedestrian and cycling environments of cities.

Keywords: GPS, GIS, Data, Transportation, Systems, Reliability

1. Introduction

The bulk of this paper describes one approach to using global positioning system (GPS) data in order to describe a system. Specifically, a continuous stream of GPS data from devices attached to trucks was applied to the problem of identifying, ranking, and then describing the statistical reliability of freight bottlenecks on the roadways in Washington State, USA. Because GPS devices are now commonly embedded in phones, and phones are carried around by people, the truck bottleneck analysis described in this paper should be understood as only one of many potential ways that GPS data can be used for systems analysis.

GPS data infrastructure is diffuse, persistently updated, and inexpensive compared to other data collecting hardware. These characteristics result in ‘big data,’ a phenomenon that creates both challenges and opportunities for researchers. GPS vendors often discard much of this data shortly after it is collected, despite the fact that historical data can be leveraged to reveal patterns across large extents, and a very fine grain. Future research can apply similar methods to those described below to analyze many systems used by people. In the conclusion section of this paper, one of such future researches – project in Copenhagen, Denmark – is briefly discussed.

2. Description of the Research: Using GPS to Identify and Rank Truck Bottlenecks

The Washington State Department of Transportation (WSDOT), Transportation Northwest at the University of Washington, and the Washington Trucking Associations have partnered on a research effort to collect and analyze GPS truck data from commercial, in-vehicle, truck fleet management systems. This project is collecting commercially available GPS data and evaluating the feasibility of using them to support a state truck freight network performance-monitoring program. An important part of this effort consists of using these data to locate and rank truck bottlenecks on Washington road network. This section of the paper discusses the bottleneck identification process.
Data

The project team currently receives daily GPS data from about 6,000 trucks traveling on roads throughout Washington State. A contract negotiated with a large GPS vendor has resulted in a near real-time raw data feed for all its client trucks when they travel in the state (Fig. 1). The commercial in-vehicle GPS devices report, via cellular technology, both at preset intervals (every 10 to 15 minutes) and when the trucks stop. The resulting GPS data set includes reads for individual truck longitude and latitude, the truck ID (scrambled for privacy), spot (instantaneous) speeds, and a date and time stamp. Other variables in the data set include GPS signal strength, travel heading, and the status of a truck stop (parked with engine on or engine off). More details about the data collection effort and the GPS-based performance measures program can be found in McCormack et al. (2010) [1].

Bottleneck Identification Process

The bottleneck identification process developed for Washington State is designed to find sections of Washington roadways that perform poorly for trucks and then to develop quantitative measures that allow these bottlenecks to be ranked and compared. The process is designed to be automated, repeatable, and statistically valid while also producing results that are usable and readily comprehended by transportation professionals and decision makers. Because the results are oriented toward WSDOT mission, the measures also have to reasonably mesh with measures that WSDOT has already developed for evaluating congestion and roadway performance for all types of vehicles. The bottleneck process must also account the network reliability, since travel reliability (travel time consistency) is as important as travel speed to the trucking community. For example, the USDOT office of operation noted, “Shippers and freight carriers require predictable travel times to remain competitive” [2]. This view is supported by input from local truckers’ focus groups, as well as by input from WSDOT Freight Systems Division staff.

The bottleneck identification and ranking process developed for Washington State involved the following tasks:

1. Segment the Roadway. Separate the state entire roadway network into analysis segments based on the locations of ramps /major intersections and, in some cases, roadway length.
2. Add Attribute Information to the Segments. Assign to each analysis segment the appropriate roadway attributes (speed limits, classification, etc.) along with heading information to determine travel direction.
3. Geo-locate the Trucks. Assign each probe truck GPS location reads to the appropriate segments. Account for the truck travel direction on the segment.
4. Locate Bottlenecks. For segments with enough truck data, use the GPS trucks travel speeds averaged over time to quantify the reliability and overall performance of each segment and identify as bottlenecks locations where trucks are performing unreliably or slowly.
5. Rank the Bottlenecks. Rank the truck bottlenecks on the basis of a range of metrics, including averaged segment travel speeds, geographic location, and the segment Freight Goods Transportation System (FGTS) category.
Each task is explained in more details below.

**Task 1: Roadway Segmentation**

WSDOT entire road network, with the associated roadway attributes is readily obtainable in a geographic information systems (GIS) compatible format. GIS software was used to divide this network into analysis segments according to the locations of junctions, ramps, or signalized intersections. Any segment longer than 3 miles was further divided into shorter segments.

Because most roadways involve two-way travel, the increasing and decreasing mileposts of the GIS linework (i.e., the roadway on a GIS map) were used to determine the travel direction of each roadway segment. In essence, except for a few one-way roads, each roadway segment was processed as two segments, one for each travel direction.

The following rules were applied to the GIS databases to segment Washington state routes. Fig.1 shows the segmentation steps graphically.

1. The roadway linework was split at every junction.
2. For interstates and other major roads with ramps, the linework was further split at access ramps. Where both on- and off-ramps were situated in close proximity, the midpoint between ramps was calculated and then used as a splitting point.
3. If the route was not an interstate, it was also split at signalized intersections. If, after these steps were followed, a resulting segment was longer than 3 miles, it was subdivided into several 2-mile segments plus one remainder segment between 1 and 2 miles long.

The following roadway data were used (downloaded from the WSDOT GeoData Distribution Catalog, www.wsdot.wa.gov/mapsdata/geodatacatalog/default.htm):

1. State Highway Linework at 24K
   1.1 Increasing roadways
   1.2 Decreasing roadways
   1.3 Ramps
   1.4 FGTS classifications
2. Intersections
   2.1 Signalized intersections
3. Highway Urban and Urbanized Areas

![Figure 2. Steps used to segment the roadway network](image)

**Task 2: Add Attribute Information to the Segments**

Within the GIS, a 50-foot buffer was added around the analysis segments developed in Task 1. The resulting area was given identifying attributes from the state roadway network. Some of these attributes already existed in the state highway GIS files from Task 1. These attributes are listed below:
1. RID: State route ID, or numerical name of the route (for example, 099 refers to highway 99).
2. ROAD_SPEED: Posted speed limit.

Additional attributes were added to the segments with both custom and stock GIS processing tools:

3. SECTION_ID: 6 character unique numerical identifier for each segment. For example, “099010” refers to the tenth (010) segment of SR 99 (099). This can be used in future studies to identify adjacent or related bottlenecks and assist in tracing the cause of bottlenecks. For example, 099010 and 099011 represent adjacent segments.

4. FGTS2009: 2009 Washington State Freight and Goods Transportation System (FGTS) is a classification of the state highways, county roads, and city streets according to the average annual gross truck tonnage they carry [3].
   T-1: more than 10 million tons per year
   T-2: 4 million to 10 million tons per year
   T-3: 300,000 to 4 million tons per year
   T-4: 100,000 to 300,000 tons per year
   T-5: at least 20,000 tons in 60 days

5. ROAD_DIRECTION: The compass heading (0-360) of the roadway. Because state route segments are often sinuous, each small portion of a segment was dynamically encoded with a range of location-specific directional headings (Fig. 3). This ensured that the road direction and truck travel direction could be accurately matched.

6. FMEAS: The lowest milepost measure of a segment. This attribute did not have to be a whole number and often had three or more decimal places.

7. TMEAS: The highest milepost measure of a segment. This attribute did not have to be a whole number and often had three or more decimal places.

8. LENGTH_MILES – Segment length in miles.

9. PUGET: Identified whether a segment was in the “Central Puget Sound” analysis extent

10. URBAN: Identified whether a segment passed through an urban area.

[Figure 3. Adding directional headings to the roadway segments]

Task 3: Geo-Locate Trucks

On a daily basis, 24 hours of new truck GPS location reads are fed into the project database. On a typical day, this data stream includes about 350,000 GPS location records. Using a GIS process, each of these locations needs to be assigned or geocoded by latitude and longitude to a point on the map. After being geo-located to the map, the location and heading of each point are compared to the segmented
linework from Task 2, and any GPS reads taken from trucks that were not traveling along a WSDOT route are filtered out. After this process, about 10 percent of all GPS reads are retained.

The GPS points are filtered in a two-step process. First, the location of each point is compared to the state route segments created in Task 2. Points that fall outside of a zone or buffer created around each segment (roughly 50 feet from the roadway center) are eliminated.

Next, the heading of each point is compared to the closest heading of a short section of the analysis segment. Points with a difference in heading of greater than 15 degrees are eliminated. Points with a difference in heading of 15 degrees or less are retained and tagged with a value indicating whether it was traveling in an increasing or decreasing direction. This process filters out trucks traveling along intersecting or non-state route roadways, and it also identifies which direction (such as north- or southbound) on a roadway segment a truck was traveling (Fig. 4). Finally, each truck GPS records was assigned the ten segment attributes listed above in Task 2.

![Figure 4. Steps used to geo-locate the individual trucks](image)

**Task 4: Locate Bottlenecks**

This task evaluates the performance of the GPS trucks as they traveled on each analysis segment. This task can only be completed if there are enough GPS truck data reads for a valid analysis. Because the GPS data for the central Puget Sound region have been collected since October of 2008, most major truck routes have hundreds or, in many cases, thousands of reads, so there are typically adequate data for most roads commonly used as freight routes. Data for roads serving the non-central Puget Sound part of the state have only been collected since August 2010, so a number of lower-level roads do not yet have enough probe truck reads for analysis to be valid. This situation will improve as the data continue to accumulate over time. Fortunately, important truck corridors, by definition, have higher truck volumes and are therefore more likely to have been used by more of our probe trucks and can be included in the bottleneck analysis.

After the roadway segments with enough reads (a minimum of a 200 trucks per segment) to fit a mixture of two Gaussian distributions have been identified, the bottleneck identification process determines the travel reliability of that segment. Because each GPS report by a device in a truck includes a spot or instantaneous speed, the distribution of speeds for the different trucks that travel on that segment can be calculated, plotted, and fitted (Fig. 5). On the basis of this speed distribution, the reliability of each roadway segment can be placed into one of three categories: reliably slow, reliably fast, and unreliable. For the urban areas in the central Puget Sound region, the truck speed data are analyzed by determining speed distributions during four different time periods:

1. AM (6:00 AM to 9:00 AM)
2. Midday (9:00 AM to 3:00 PM)
3. PM (3:00 PM to 7:00 PM)
4. Night (12:00 AM to 6:00 AM and 7:00 PM to 12:00 AM)
Each time period is fit with a statistical distribution based on the trucks spot speed data. Because the other areas of the state have fewer data points and less time-related congestion, their analysis period is a full day. The reliability evaluation criteria include two levels:

1. Evaluate whether the travel condition during a given time period (the central Puget Sound region) or a given day (the non-Puget Sound areas of the state) is reliable given the speed distribution and a statistical fitting process. Generally speaking, if the speed distribution has two speed “humps” and is bimodal (such as in Fig. 3a), then the travel condition is considered unreliable. Otherwise, the travel condition is unimodal and is considered reliable with one average speed (as in Fig. 3b) and goes to level B for further evaluation of whether the travel speed is reliably slow or reliably fast. For the Puget Sound roadways, this measure can identify how the reliability condition changes over different times of the day.

2. Evaluate whether the travel condition during a given time period is reliably slow or fast. If the average speed is higher than 70 percent to 85 percent of the posted speed limit, then the travel condition is considered reliably fast; otherwise, the travel condition is considered reliably slow. This measure has been adopted by WSDOT as the speed threshold for evaluating the duration of a congested period [3]. For example, on a freeway with a 60 mph speed, the threshold value is 45 mph because this is between 70 percent and 85 percent of the posted speed.

Task 5: Roadway Performance Evaluation

Once the bottlenecks have been identified on the basis of their level of (un)reliability, three different measures are used to further rank and sort the truck bottlenecks severity. They are as follows:

1. Average Speed
   For the central Puget Sound region, this method first calculates the mean speed for each freeway segment during four different time periods. Four mean values are combined to get an average speed measurement. This measure weighs different time periods equally and reflects the overall performance of the freeway segment. For the non-Puget Sound roadways the ranking of average speed for the entire day is used.

2. Frequency of Truck Speed Falling Below 60 Percent of thePosted Speed Limit
   This measure reflects the severity of congestion on the freeway segments. For the central Puget Sound region, this method calculates the percentage of truck speed falling below 60 percent of the posted speed limit during the different time periods and then averages those four percentage values to evaluate the frequency of truck speed falling below 60 percent. WSDOT uses the 60 percent value as the speed threshold for evaluating whether the freeways are experiencing severe congestion [2]. This threshold is adjusted for roadways with different posted speed limits. For the central Puget Sound region, this method first calculates the mean speed for each freeway segment during the four different time periods. The four mean values are combined to get an average speed measurement. This measure weighs different time periods equally and reflects the overall performance of the freeway segment. For the non-Puget Sound roadways the ranking of average speed for the entire day is used.

3. Geographic Areas and Freight and Goods Transportation Systems (FGTS) Categories
   This evaluation step reflects policy decisions made by WSDOT. The bottleneck process as currently developed separates the state into the central Puget Sound area and the rest of the state. This step acknowledges that the Puget Sound area freight issues and truck volumes are notably different than those in other areas of the state. An additional sorting of the bottlenecks based on the Freight Goods Transportation System (or FGTS) is possible. Use of the FGTS allows freight
routes with inherently more importance (a higher FGTS category) to be given the greater weight. This measure captures a segment strategic importance to freight mobility.

Task 6: Bottleneck Analysis Reports

The results of this bottleneck process are placed in tables sorted by severity. These tables can be based on a range of measures, as it is identified in Task 5. However, for easier use by transportation professionals and decision makers, the results can also be summarized into a one page per bottleneck report. An example for a bottleneck report is shown in Fig. 6. The bottleneck report includes a map and description of the bottleneck location, daily truck volume if available, and other relevant information.

Severe truck bottleneck in Central Puget Sound:
1-5 southbound

- Location: 1-5 southbound between NE 63rd St and NE Pacific Ave.
- Length: 1.3 mile
- Daily Truck Volume: 11,000
- Average truck travel speed: 38 mph
- Percentage of travel speed below 35 mph: 48%

Travel Reliability:
- Time Period | Reliability
- All | Unreliable
- Midday | Unreliable
- PM | Unreliable
- Night | Reliably Fast

Figure 6. Example of a bottleneck report

3. Conclusion

After developing a system for pre-processing a continuous stream of GPS data in a GIS, new truck bottleneck reliability measures were developed and applied to evaluate the performance of a truck transportation network. This was done by fitting the collected truck speed data with a mixture of two Gaussian distributions, and then using a set of rules to determine whether the travel condition was: (1) unreliable, (2) reliably slow, or (3) reliably fast basing on the estimated distribution parameters. The methodology presented has proved efficient in identifying the worst truck bottlenecks within Washington State road network. This research provides a useful tool for decision makers to systematically locate the worst bottlenecks, and pinpoint the locations where bottleneck alleviation may provide the greatest benefit.

This paper methods can be applied to other systems that create streams of GPS data. Future research may consider pedestrian and cycling networks and other fine-grained urban human systems. Using GPS data to identify and then describe the urban attributes in cities that attract pedestrians and cyclists will be valuable for: 1) urban and transportation planners in the municipality of study; 2) architecture and planning students working to promote healthy cities; and 3) other cities interested in sustainable transportation and inviting urban environments.

References