Transport mode and cluster separation in road traffic network travel times

Industry Representatives:
Wayne Dale, Queensland Department of Transport and Main Roads, Brisbane, QLD

Problem Moderators:
James McGree, Mathematical Science School, Queensland University of Technology, Brisbane, QLD
Gentry White, Mathematical Science School, Queensland University of Technology, Brisbane, QLD

1 Problem Summary

Management of current roadways infrastructure and planning for maintenance and future growth are important projects for local and statement governments. The maintenance and construction of new roads are large undertakings, requiring considerable dedication of resources and potential disruption for many users. Making decisions about these activities requires a good understanding of the patterns of usage for existing roadways and a means of planning and estimating the likely impact of future additions or modification of the existing network.

The Queensland Department of Transport and Main Roads (TMR) collects usage data by recording the movement of media access control (MAC) addresses from bluetooth enabled devices through sensors placed at major intersections throughout Queensland. These data represent a sample of the actual traffic and can be used to make inference about travel times throughout the network. Based on these data, TMR presented with three questions:

1. Is the current method used by TMR to estimate typical (car) travel times using the MAC address data accurate in comparison with real-world data collected by guide cars?
2. Is it possible to discern multiple transport modes from the MAC address data, i.e. discerning buses, cars, and bicycles and pedestrians?
3. Is there any way to implement change point detection methods to identify significant changes in traffic patterns?

2 Approaches and Methodology

In order to address these questions, individual groups used a variety of methods and approaches including applying spline based methods to, for example, estimate average travel time and identify change points in traffic patterns, cluster analysis to identify modes of travel and a global network analysis of travel speeds for a given MAC address for travel mode identification.

2.1 The viability of current TMR methodology

Currently TMR uses the collected MAC address data to estimate traffic speed. The current practice of TMR is as follows:

1. Clean the data by removing addresses travelling less than 5 km/h from the data set;
2. Pooling data into five minute bins; bins with less than 5 valid addresses are not analyses;
3. The removal of outliers using the median absolute deviation (MAD) method; and
4. Calculating the sample mean for the remaining addresses an estimator for the travel speed in that interval.

The initial request from TMR is that this methodology be evaluated against other methods, and actual observed data from "spotter" cars. The current methodology employed by TMR showed the best fidelity with the spotter car data at slower speeds (peak hour) but tended to under predict travel time at other times of day.

As an alternative to the TMR approach, traditional spline based methods were compared with median based smoothers. The results indicate that as expected the spline based methods tended to over-smooth periods of rapid change around peak hours. The spotter car data showed significant and sharp increases in travel time approaching the peak hour, especially in the late afternoon. The median based smoother was able to detect the change points without over smoothing, but tended to over-estimate travel speeds earlier in the day, and under-estimate travel speeds during peak hours.

The global network analysis used information gathered about a MAC address from the whole network to filter out modes of transport not of interest when calculating typical travel times. For example, filtering out pedestrains, cyclists, buses and multiple MAC addresses. This enabled typical travel times to be calculated based on cars only. Flux on a segment of road was also considered in determining average travel times. The approach showed good agreement with "spotter" data, even at higher speeds.

2.2 Identification of multiple transport modes

MAC addresses are identified as they travel past stationary detectors at intersections. Since vehicles can quite obviously have multiple MAC addresses, it can be deceptive in computing average speeds. For instance an individual may have a mobile phone, laptop computer, and tablet device, which all have visible MAC addresses to be collected by the detectors. A bus may cross through an intersection carrying as many as 62 people, each recording multiple MAC addresses. In addition pedestrian and bicycle traffic is also recorded through intersections. It is desirable for TMR to be able to identify and monitor each of these three (or four) different modes of transport. In order to do this, methods are needed to identify each of these modes.

The global network approach was considered for this identification. If a MAC address reached an average speed of over 40km/h over a 1km stretch of road, this was classified as a vehicle. Further, similarities in travel times and locations throughout the network were explored as a means of identifying a vehicle with multiple MAC addresses.

Clustering methods were also used to determine mode of travel. This was only undertaken on a road segment basis but clear clusters were present in the data. One would need to track cluster grouping per MAC address throughout the network to gain insight into travel mode.

2.3 Change-point detection

Travel times vary over given sections of roadway over a 24 hour cycle, in particular peak hour traffic can drastically slow travel speeds. In designing and planning for roadway construction and management TMR would like to be able to analyse real-time data to detect significant changes in travel speed. This would need to be calculated on a near real-time basis. Most approaches to calculating travel times were able to capture significant changes in travel time, say around peak hour. More investigation needs to be undertaken to determine if any method can capture this on a near real-time basis.

3 Future Work

These data and these questions present an excellent opportunity for applying and developing methodologies for mathematical/statistical modelling techniques. The challenges here include topics such as robust smoothing and change-point detection, network analysis, cluster identification and classification models, as well as the need to implement a working solution in a real-world environment. Several of the workshop participants are continuing the research they began and are pursuing future collaborative research opportunities with TMR.