Inference of road traffic congestion from sensor events

Nicholas Gould and James Cheng, Manchester Metropolitan University, Manchester, UK
William Mackaness, University of Edinburgh
n.gould@mmu.ac.uk

Abstract

We describe a current project that attempts to identify high level road traffic congestion events from low level events determined by sensors.

The UK Department for Transport (DfT) defines road congestion as having both a physical dimension and a relative dimension (Department for Transport, 2015). The former, relates to the impact vehicles have on each other’s presence, impeding progress and causing congestion, whereas the latter attempts to account for the driver’s perception of congestion; a queue that might seem noteworthy in a rural setting may seem trivial to the urban motorist. The importance of the driver’s goal is another aspect that is not captured simply by metrics. Two drivers in the same queue might have different perceptions of congestion if one is on route to an important business meeting and the other is on a shopping trip.

Many different sources are used to identify traffic congestion in addition to that data collected from sensors installed by municipalities: Google use data from GPS enabled smart phones running Android; fleet vehicles or high-end cars fitted with GPS can provide historical journey time data; Uber has started to make its GPS data available to city planners (Morozov, 2015).

However, these sources may not persist. For example, data services may suffer temporary outages or be permanently withdrawn; sources that were once free may start extracting a charge or change their terms and conditions. It is therefore necessary to ensure that any model can embrace multiple data sources.

As an example of the heterogeneous nature of the data, journey time data can be derived from smart phones carried by vehicle occupants using both GPS tracking and by using stationary, passive sensors to identify Bluetooth signals. However, each method has a different approach to calculating journey time (Figure 1).
We suggest therefore that a numerical model is not sufficient to capture the complexities of road congestion, in particular the relative dimension. In order to understand congestion we require an open model that is not reliant on opaque data sources, not limited to road network sensors, but can be expanded to incorporate other data sources such as weather forecasts, air quality measurements and social media (Pan et al., 2013).

We now describe a current research project funded by the DfT. This research aims to build on the method of Rude and Beard (2012) of identifying congestion by regarding it as a high-level event that is comprised of a number of low-level or primitive events. The primitive events are identified using sensors in the road network, and these consist of measures such as increasing journey times and increasing traffic flows. The spatial-temporal pattern of these events will then be used to identify congestion and predict its behaviour. The eventual aim is to develop an early warning system for congestion, similar to those developed for storms.

For example, consider Figure 2. The increasing journey time and decreasing flow may be an indication of traffic congestion spreading from West to East. The use case applied by Rude and Beard (2012) used simulated data along a single section of road, whereas we aim to study real sensor data in a 2D plane. We have two data sources, journey time data over road links derived from Bluetooth sensors and traffic volume sensors. In this use case the simple model of a sensor network providing point data such as temperature or air pressure does not apply. The road count data is point based but the journey time data applies to a road link or set of road links. The solution in the first instance is to abstract the journey time data to a point in the centre of a road link. The model is further complicated by bidirectionality of the phenomena; a journey time or volume in one direction in the road network may be very different in the opposing direction (note, Figure 2 does not represent this).

---

1 Provided by Transport for Greater Manchester.
2 Is this the equivalent of wind speed and direction in a storm?
In an effort to represent one aspect of the relative dimension of road congestion, we compare the study day data with the data for a set of “typical” days at the same location. The aim is to be able to identify the cause of a congestion event; for example, a large sports event such as a Premier League football match will have a peak either side of a two hour time period.

Even if we use this approach to identify congestion, problems will occur when we try to apply our model outside of its current case study region, the city of Manchester, UK. We have devised the model to consider two types of sensors, one used for measuring journey times and one for measuring traffic volumes but to apply the model to other cities assumes that the same sensor types are available. To make the model general we need to incorporate other sensors types. This is where a semantic approach might be applicable, particular when investigating the relative dimension of congestion. There is a need for a semantic description of congestion that can be shared between not only applications but also agencies.


www.theguardian.com/commentisfree/2015/feb/01/cities-need-to-fight-uber-transport-choice-evgeny-morozov