Transportation Research Part B 76 (2015) 81–92

Contents lists available at ScienceDirect

Transportation Research Part B

journal homepage: www.elsevier.com/locate/trb

Repeatability & reproducibility: Implications of using GPS data for freight activity chains

Johan W. Joubert *, Sumarie Meintjes [1]

Centre of Transport Development, Department of Industrial & Systems Engineering, University of Pretoria, South Africa

A R T I C L E I N F O

Article history:
Received 9 December 2014
Received in revised form 10 March 2015
Accepted 11 March 2015
Available online 27 March 2015

Keywords:
Repeatability
Reproducibility
Freight
GPS data

A B S T R A C T

As transport modellers we are interested in capturing the behaviour of freight vehicles that includes the locations at which vehicles perform their activities, the duration of activities, how often these locations are visited, and the sequence in which they are visited. With disaggregated freight behaviour data being scarce, transport modellers have identified vehicle tracking and fleet management companies as ideal third party sources for GPS travel data. GPS data does not provide us with behavioural information, but allows us to infer and extract behavioural knowledge using a variety of processing techniques. Many researchers remain sceptical as specific human intervention, referred to as 'expert knowledge', is often required during the processing phase: each GPS data set has unique characteristics and requires unique processing techniques and validation to extract the necessary behavioural information. Although much of the GPS data processing is automated through algorithms, human scrutiny is required to decide what algorithmic parameters as considered 'best', or at least 'good'. In this paper we investigate the repeatability and reproducibility (R&R) of a method that entails variable human intervention in processing GPS data. More specifically, the judgement made by an observer with domain expertise on what clustering parameters applied to GPS data best identify the facilities where commercial vehicles perform their activities. By studying *repeatability* we want to answer the question *'if the same expert analyses the GPS data more than once, how similar are the outcomes?'*, and with *reproducibility* we want to answer the question *'if different experts analyse the same GPS data, how similar are their outcomes?'* We follow two approaches to quantify the R&R and conclude in both cases that the measurement system is accurate. The use of GPS data and the associated expert judgements can hence be applied with confidence in freight transport models.

1. Introduction

The economic vitality of a country greatly depends on its ability to effectively move freight (Cherrett et al., 2012; Hensher and Figliozzi, 2007), yet commercial road traffic is still not fully understood at a macro and system-wide level. It is critical to understand the behaviour of freight vehicles so that representative freight models can be developed for transport planning purposes (Joubert and Axhausen, 2011). With disaggregated freight behaviour data being scarce (McCormack and Hallenbeck, 2006), decision-makers have to make use of a variety of data collection methods.

* Corresponding author. Tel.: +27 12 420 2843; fax: +27 12 362 5103.
E-mail addresses: johan.joubert@up.ac.za (J.W. Joubert), sumarie.meintjes@gmail.com (S. Meintjes).
[1] Tel.: +27 12 420 5411; fax: +27 12 362 5103.

http://dx.doi.org/10.1016/j.trb.2015.03.007
0191-2615/

Survey-based techniques are predominantly used to collect travel behaviour data. There are many limitations to this method. Companies are usually hesitant to allow surveys to be conducted due to privacy concerns (Ma et al., 2011); surveys are expensive and time-consuming (Sharman and Roorda, 2011), restricting the size of the data set collected; and the accuracy of the data collected is heavily dependent on the respondent (Bohte and Maat, 2009). Under-reporting of trips also occur, as explained by Wolf et al. (2001): drivers might not log trips immediately, and are subject to memory decay, which may result in parts of trips, or even complete trips, being forgotten and not logged; the drivers may fail to understand or follow survey instructions; the full details of trips might not be logged, since drivers often see some parts of their trips as insignificant or unnecessary; drivers are also subject to simple carelessness, reducing the accuracy and completeness of the travel data collected. Consequently, global positioning system (GPS) devices are increasingly being used either to enhance travel data collected from surveys, or as the sole method of collecting travel data (Bohte and Maat, 2009).

GPS data can be obtained by inviting firms to participate in a data collection effort, as was done by Greaves and Figliozzi (2008). The effort required to identify firms willing to share sensitive information, and the cost and time associated with the installation of GPS devices creates the need for an alternative source for GPS data (Sharman and Roorda, 2011). Vehicle tracking and fleet management companies have been identified as ideal third party sources for GPS travel data. These companies have already incurred the cost of installing GPS devices in freight vehicles, and their data have been collected from a range of different companies, for many vehicles. Depending on the age of the company, they can provide transport modellers with ample historical travel data. We acknowledge that users of such data sets should be cautious about the inherent selection bias: the vehicles in the data set may only reflect a (potentially) bias sample of users who subscribe to the particular vehicle tracking or fleet management company's service offering.

The observation period is increased greatly with the use of GPS devices (Bohte and Maat, 2009; Sharman and Roorda, 2011), such as a data collection period of three months used by Sharman and Roorda (2011), six months used by Joubert and Axhausen (2011), and more than a year used by Ma et al. (2011). These large data sets are more accurate and complete than survey-based data, because it is not affected by the driver. Data is collected passively and automatically by the GPS device, with little to no interaction required by the driver. This has been shown in a number of studies that compare the accuracy and completeness of driver-reported trips to GPS-reported trips: Bohte and Maat (2009) found that more trips per tour were recorded with GPS devices in a large-scale study conducted in the Netherlands; Stopher et al. (2007) found that the Sydney Household Travel Survey under-reports 7% of trips, compared to the much higher rate of 20–25% found in the standard computer-assisted telephone interview (CATI) survey conducted in the United States; and Wolf et al. (2001) studied the CATI method used in the California Statewide Household Survey, and found that GPS devices identified 29% more trips.

Despite its advantages over survey-based methods, collecting travel data using GPS devices is not perfect. There are external factors that may influence the device's ability to collect travel data, as explained by Czerniak (2002), Du and Aultman-Hall (2007) and Greaves and Figliozzi (2008).

GPS devices typically record the time that the vehicle is at a specific location. The locations are recorded as longitude and latitude values, and the accuracy of the locations depends on the available number of satellites that the GPS device is locked to. A GPS device needs a signal lock on at least four satellites in order to determine its location. Large infrastructure can block or interfere with the signal, such as when driving between tall buildings, through a tunnel, or under a bridge. While the GPS device searches for satellites, incorrect locations can be reported, or a complete loss of data points can occur.

Some freight vehicles do not switch off their engines when performing their activities, and idle at the location where the activity is being performed. This may result in signal jiggle (Czerniak, 2002; Ma et al., 2011): a false report of movement that occurs because the GPS points fluctuate around a stationary position.

As transport modellers, we are interested in capturing the behaviour of freight vehicles, such as the locations at which vehicles perform their activities, the duration of activities, how often these locations are visited, and the sequence in which they are visited. GPS data do not provide us with behavioural information, but allow us to infer and extract behavioural knowledge using a variety of processing techniques (Du and Aultman-Hall, 2007; Ma et al., 2011; Sharman and Roorda, 2011). Expert knowledge is often required during this processing phase, since each GPS data set has unique characteristics and requires unique processing techniques to extract the necessary behavioural information. Examples of expert opinion include visually inspecting and splitting a GPS trace to account for a commuter's likely mode change; route choice identification from infrequent GPS records in a trace; and validating a clustering algorithm to ensure the identified activity location is in a reasonable and likely place, for example inside the boundary of a property (Sharman and Roorda, 2011). Many of these examples depend on automated algorithms of pattern recognition, but rely on human intuition and judgement for validation.

The behavioural knowledge we gain is often used as inputs into and embedded in transport planning models, which, in turn, is used to support decision making for large and expensive infrastructure and policy projects. The quality of decision support that our models provide may therefore be very sensitive to the quality of expert knowledge that went into the source data. For example, Joubert and Axhausen (2013) used expert knowledge to build their freight model by applying a *gold standard* approach. Activity chains are extracted from GPS data, and from the activity chains a complex network is built to study the connectivity between facilities. The conversion from activity chains to a complex network is highly dependent on the expert approach as it influences which points (freight activities/trip ends) to include in their complex network, and which to ignore. If their approach is not repeatable or reproducible, one cannot make confident decisions from such a complex network, and it calls the use of GPS data into question.