Transportation Research Part C 64 (2016) 17–27

Contents lists available at ScienceDirect

Transportation Research Part C

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

Classifying the purpose of stopped truck events: An application of entropy to GPS data

Kevin Gingerich [a,b,*], Hanna Maoh [a,b,1], William Anderson [a,c,2]

[a] Cross-Border Institute, University of Windsor, 401 Sunset Ave., Windsor, ON N9B 3P4, Canada
[b] Department of Civil and Environmental Engineering, University of Windsor, 401 Sunset Ave., Windsor, ON N9B 3P4, Canada
[c] Department of Political Science, University of Windsor, 401 Sunset Ave., Windsor, ON N9B 3P4, Canada

ARTICLE INFO

Article history:
Received 20 February 2015
Received in revised form 11 December 2015
Accepted 4 January 2016
Available online 2 February 2016

Keywords:
Freight transportation
Stop purpose
Global positioning system (GPS) data
Entropy

ABSTRACT

This paper applies the concept of entropy to mine large volumes of global positioning system (GPS) data in order to determine the purpose of stopped truck events. Typical GPS data does not provide detailed activity information for a given stop or vehicle movement. We categorize stop events into two types: (1) primary stops where goods are transferred and (2) secondary stops where vehicle and driver needs are met, such as rest stations. The proposed entropy technique measures the diversity of truck carriers with trucks that dwell for 15 min or longer at a given location. Larger entropy arises from a greater variety of carriers and an even distribution of stop events among these carriers. An analysis confirms our initial hypothesis that the stop locations used for secondary purposes such as fuel refills and rest breaks tend to have higher entropy, reflecting the diversity of trucks and carriers that use these facilities. Conversely, primary shipping depots and other locations where goods are transferred tend to have lower entropy due to the lower variety of carriers that utilize such locations.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

Modern supply chain processes rely heavily on goods transferred between different manufacturers, distributors and retailers across various regions. While several modes of transportation can be used, the reliance on heavy trucks is very prominent in North America, particularly during the final segment of a supply chain. According to recent studies, truck freight in Canada has doubled in tonnage between 1992 and 2004 (HRSDC, 2013) and nearly doubled in value in the U.S. between 1993 and 2002 (USDOT, 2013). This increased economic and transportation dependency underlies the importance of trucks in North America. Increased freight mobility has prompted the need for detailed data to understand current processes and predict future freight travel demand. Previous data was often obtained through sampled surveys that can be expensive for the analyst and time consuming for the respondent. An emerging alternative to surveyed data is passive information collected through global positioning systems (GPS) technology.

* Corresponding author at: Cross-Border Institute, University of Windsor, 401 Sunset Ave., Windsor, ON N9B 3P4, Canada. Tel.: +1 519 253 3000x3737.
E-mail addresses: gingerik@uwindsor.ca (K. Gingerich), maohhf@uwindsor.ca (H. Maoh), bander@uwindsor.ca (W. Anderson).
[1] Tel.: +1 519 253 3000x4987.
[2] Tel.: +1 519 253 3000x2366.

http://dx.doi.org/10.1016/j.trc.2016.01.002
0968-090X/© 2016 Elsevier Ltd. All rights reserved.

Since the late 1990s, GPS transponders have been frequently used by freight carriers to track the current position of truck fleets. However, the usage of GPS truck data in freight analysis has only become more commonplace in recent years. Several studies have been published for the U.S. using truck GPS data from the American Transportation Research Institute (ATRI) (see for example: Liao, 2014; Kuppam et al., 2014; Bernardin et al., 2015). GPS data provides enormous potential for understanding the current patterns of freight based on the large volume of data alone. However, the volume of data also presents a substantial challenge since it was not originally intended as an input for transportation models and analysis. For example, the purpose of truck movements and stops observed with GPS data is initially unknown unless accompanied by travel diaries that provide ancillary information. Therefore a need exists for novel methods to mine this data in order to understand various patterns given that ancillary information is not typically collected.

This paper addresses the above challenge by applying the principle of entropy to illustrate how truck GPS data can be mined to identify the type of locations that are associated with an observed stopped truck event. The analysis is conducted using data collected by Shaw Tracking and provided to us by Transport Canada. The obtained dataset contains individual GPS pings for a large group of Canadian owned freight carriers with truck movements across Canada and the U.S. This paper uses the observed GPS records for the month of March, 2013. Here, a total of 40,650 individual trucks with approximately 101.6 million GPS pings (i.e. data points) are analyzed. Each GPS ping results in a data record containing the carrier ID, power (truck) ID, latitude, longitude, date, and time. The elapsed time between GPS pings and a dwell time that accumulates if the truck is stopped can be derived from the input GPS data. Pings associated with meaningful dwell times represent stop events that can be classified as (1) primary, which occurs when goods are transferred between the truck and location (or another truck), or (2) secondary, which occurs when a truck is stationary for other purposes such as driver breaks or fuel refills.

Although stops can be easily observed in the dataset, their purpose is unknown due to the lack of ancillary information on record. Primary stops are particularly important for models since they denote trip ends for a given truck. Likewise, secondary stops are useful since they complement primary stops by providing a complete picture on the nature of truck movements over space. However, it is important to make the distinction between the two types of stops since they correspond to different activities. Any transportation model that utilizes micro-data on truck movements (such as the GPS data) for its calibration will require an accurate representation of primary and secondary stops to properly represent travel activity patterns. Contemporary microsimulation models handle truck movements as the outcome of tours as oppose to two leg-trips. The latter is inspired by the available evidence from the literature which emphasizes the superiority of the tour based approach (Hunt and Stefan, 2007; Ferguson et al., 2012).

Typically, trips derived from GPS data should be delimited by stop events that constitute the trip ends (start and finish). These trip ends are used to determine the origin and destination of trips. Errors occur if secondary stops are misclassified as possible trip ends since goods are not actually transferred. Consequently, failing to classify truck stop purpose will lead to a miscalculation of the actual trip end locations and the industry pertaining to the trip. When analyzing the data, we expect secondary stop locations to attract a greater variety of trucks belonging to different carriers compared to primary stop destinations as they provide useful services to any truck passing nearby. This diversity of carriers dwelling at a stop location can be captured by associating a larger variety of carriers with a higher level of entropy. As such, we contend that entropy is an ideal concept for classifying the type of stops. To our knowledge, this is the first study to mine truck GPS data through the entropy technique in an attempt to identify the purpose of stopped truck events.

The remainder of this paper begins with a background section discussing the emerging trends in transportation modeling and the subsequent need for large, detailed data sources. Passive GPS data is capable of fulfilling these needs but requires further methods and pattern recognition to derive more detailed information. This is followed by a description of the GPS dataset used in this paper and the required pre-processing steps. Next, the formulation of an entropy method is introduced to help differentiate between primary and secondary stops. The validation results on a sample of stops for March, 2013 are then discussed before closing the paper.

2. Background

In 2008, the National Cooperation Highway Research Program classified current freight forecasting models into five categories (TRB, 2008). The first four classes consist of some combination of the four step UTMS method (trip generation, trip distribution, modal split, and network assignment) with the level of detail dependent on the project objectives. The fifth category is the economic input–output model made popular by Wassily Leontief (1986). All five categories generally represent aggregate models where the estimated total number of freight trips between each origin–destination pair is utilized. Several issues associated with the use of aggregate based models arise from their inability to effectively capture the behavior of multiple decisions from an individual. For example, traveling to several locations before returning to a shipping depot (trip chaining) would naturally require each stop to be conditional upon the previous choices. Moreover, aggregate models are generally not considered to be sensitive to policy changes (Miller et al., 2004).

Due to the above deficiencies, microscopic models based on individual vehicles are becoming more commonplace. Chow et al. (2010) append the five classes of freight forecasting models previously described by the TRB report (2008) with two extra categories based upon disaggregate, activity based methodologies. The first category is a logistics class of models that account for decision making throughout a supply chain process. Many of these models are capable of handling processes not available in the aggregate models such as empty trucks (Chow et al., 2010; Raothanachonkun et al., 2008; Holguín-Veras

Download English Version:

https://daneshyari.com/en/article/526269

Download Persian Version:

https://daneshyari.com/article/526269

Daneshyari.com