

10th International Conference on Transport Survey Methods

## Workshop synthesis: System based passive data streams systems; smart cards, phone data, GPS

Martin Trépanier <sup>a,\*</sup>, Toshiyuki Yamamoto <sup>b</sup>

<sup>a</sup> Polytechnique Montréal, CIRRELT (Interuniversity Research Centre on Enterprise Networks, Logistics and Transportation) P.O. box 6079,  
station Centre-ville, Montréal, Québec H3C 3A7, Canada

<sup>b</sup> Nagoya University, Furo-cho, Chikusa-ku, Nagoya 464-8603, Japan

---

### Abstract

Each day, huge quantities of passive data streams are collected by smart card, GPS, Bluetooth and mobile phone systems all over the world. These data happen to be very useful to transport planners, because of the valuable spatial and temporal information they contain. This paper is a synthesis of a workshop on passive data streams held in Australia in November 2014. Many issues are discussed here: definitions, data collection and processing, privacy, how to use these data for transport planning, how to integrate these data with traditional and more “active” data sources, such as household travel surveys that respondents are explicitly asked to participate in, and how to cope with the absence of socio-demographic attributes in passive data streams.

© 2015 Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Peer-review under responsibility of International Steering Committee for Transport Survey Conferences ISCTSC

*Keywords:* Passive data stream; smart card; smart phone; GPS; Bluetooth; privacy

---

### 1. Introduction

The use of passive data streams has become increasingly popular among transport planners. Streams are said to be "passive" because there is no acknowledgement of the user that "his/her" data are used for planning. Hence, there is no additional input that could be used to qualify, for example, the purpose of the trips or the socio-demographic attributes of the users. The huge quantities of data collected continuously by passive stream technologies, such as

---

\* Corresponding author. Tel.: +1-514-340-4711 ext. 4911; fax: +1-514-340-4173.

E-mail address: [martin.trepanier@polymtl.ca](mailto:martin.trepanier@polymtl.ca)

mobile phones, GPS devices, Bluetooth sensors and smart card fare collection systems in public transport, can surely be used to analyze the behavior of transport users or vehicles.

This document synthesizes the discussion of the W8 workshop for the International Steering Committee on Travel Survey Conference that was held in Leura, Australia, in November 2014. The workshop can be regarded as the extension of the workshop at the previous conference of the same series in 2011. The previous workshop focused on passive public transportation data streams (Morency, 2013), and several other types of passive data, such as GPS, Bluetooth and mobile phone systems, are added to the focus of the current workshop. It presents the topics as they were discussed during the 6-hour workshop:

- privacy issues related to the use of passive data streams, may it be primary or secondary data as defined by the participants;
- challenges related to the use of passive data (GPS, mobile phone, Bluetooth, smart card);
- the possibility to add socio-demographic information to collected data;
- the role of passive data streams in transport planning, for a wider perspective.

For greater convenience, the different types of data streams are discussed in separate sections because each type has its own issues and challenges. The paper finishes with a concluding section that discusses the open questions that emerged during the workshop. Further research directions are also suggested.

## **2. Context**

This section presents contextual elements about the definition, privacy and control of passive data streams.

### *2.1. Definition of passive data streams*

There are two types of passive data streams. A traveler, through the use of a smart phone application or a portable GPS device, produces a primary data stream voluntarily. However, in all cases there are no interactions between the device and the user during data collection, but post-validation can occur in specific research studies. A secondary data stream is the most common source: data are collected without any knowledge from the user, by a system that sometimes is not intended to collect mobility data. Smart card fare collection system data, mobile phone location collection by phone operators and roadside Bluetooth detection are examples of secondary data.

### *2.2. Privacy*

Passive data streams can be used to trace the mobility behavior of a single transport user. This raises privacy questions because the confidentiality of data must be assured by people collecting and analyzing data. For primary data streams, the issue is the same as for active surveys: a consent agreement can be established between the respondents and the survey organization. A secondary data stream is more complicated because there is usually no consent agreement. In this case, the data stream must be kept anonymous according to the legislations in place. To avoid reverse identification, an irreversible encryption algorithm can be used to encrypt the identification number of the card, phone or Bluetooth device. In this way, the mobility pattern associated to each device can be analyzed, but the link to the transport user cannot be performed. Privacy questions about the usage of data can be addressed through ethical reviews completed by institutional review boards. During the workshop, the privacy issue was presented by Yamamoto (2014). In his study, a driver's willingness to provide vehicle travel information using a GPS device is investigated, and the results suggest that frequent vehicle users are more willing to provide the most information, including travel routes, origins and destinations, but less willing to provide data on speed.

### *2.3. Data control of secondary data*

Although researchers or the transport authority usually collect primary data, this is not the case for secondary data, which causes some issues. First, there can be a cost of data acquisition, and agreements must be signed

Download English Version:

<https://daneshyari.com/en/article/1106790>

Download Persian Version:

<https://daneshyari.com/article/1106790>

[Daneshyari.com](https://daneshyari.com)