Chapter 14
Using Call Detail Records of Mobile Network Operators for Transportation Studies

Erki Saluveer
Positium LBS, Estonia

Rein Ahas
University of Tartu, Estonia

ABSTRACT
Contemporary transportation research is seeing a steady incline in the use of new, digital tracking data. The necessity for such new types of data has risen from the continuous growth of society’s mobility, the attractiveness of big data, and the need for data that could be collected automatically to be used in developing various monitoring systems. However, introducing new types of data requires the developing and thorough testing of new methodologies. In the current chapter, the authors introduce some methodological issues related to using passive mobile positioning data in transportation research. The Call Detail Records (CDR) and Data Detail Records (DDR) of Mobile Network Operators are a set of data that are automatically recorded and used with increasing frequency by scientists, including transportation researchers and developers of monitoring systems. The authors go on to introduce the Estonian experience in managing such data, spatial and temporal interpolation, the determining of anchor points, and activities responsible for movement. They assess both positive and negative aspects of using passive mobile positioning data and briefly consider legislation and privacy issues in regards to such data.

INTRODUCTION
The current chapter shall introduce the methodology for using passive mobile positioning data in transportation studies which was developed by Positium LBS and the Department of Geography of the University of Tartu in collaboration with Estonia’s largest mobile network operator EMT (Telia & Sonera Group). This methodology has been used in scientific research (Ahas et al., 2010; Järv et al., 2012; Järv et al., 2014), and in urban planning and transportation planning projects (Järv et al., 2007; Positium, 2009, 2010, 2011, 2013) in Estonia since 2007. Mobile positioning

DOI: 10.4018/978-1-4666-6170-7.ch014
can be divided into active and passive positioning (Ahas, 2010). Active mobile positioning is used for mobile tracking in which the location of the mobile phone of the respondent is determined with an intended query within mobile network operator (MNO). Active positioning is normally conducted using a clearly defined and recruited sample of respondents and combined with a questionnaire. Passive mobile positioning data are generated as a part of the MNO operations and are automatically recorded in memory files as secondary data (log files) of phone or network use. The more commonly available are Call Detail Records (CDR) and Data Detail Records (DDR). Different operating systems (Ericsson, Nokia-Siemens, Motorola, etc.) have different technical solutions and terminology for such data outlets, but most operating systems have a similar structure under 2G, 3G and 4G standards (Dahlman et al., 2011; Smoreda et al., 2013).

In this chapter, we shall focus on the methodology, which uses call detail records (CDR), for transportation studies as CDR is the most commonly used passive positioning data (Asakura & Hato, 2007; Bar-Gera, 2007; Phithakkitnukoon et al., 2012; Calabrese et al., 2013). The advantages of CDR data over traditional methods used in transportation studies such as census, surveys and GPS are the higher spatial and temporal resolution of data; a longer observation period; and the digital and automatic data collection. The digital data collection and high quantities can be also problem in transportation research because of extra work in data processing and extra costs (Moiseeva et al., 2010). CDR data is collected by every MNO for billing purposes and therefore it is technically very easy to obtain. Also, in the European Union MNOs have to comply with the Council Directive 2006/24/EC for data retention for the period of six months to two years as from the date of communication.

Nevertheless, it must be stressed that passive mobile data are by no means a “miracle cure” for transportation research and modelling. CDR data presents us with the movement of anonymous ID’s over a certain period of time, but we have very little information regarding those moving about and their reasons for doing so. CDR data are also less precise both in time and space than travel behaviour data obtained through surveys and transportation census. Yet, CDR data make it easier to obtain the traffic flow of roads or areas in different time frames from the memory files of MNO-s without the need to conduct a specific survey or census. The broad spatio-temporal coverage of the data enables us to perform trip generation and trip distribution and to generate “observation based” origin-destination matrices for selected time periods and locations. Completing the mode choice and route assignment is, however, more complicated and CDR data is also problematic when trying to determine the activities related to trips. However, it is possible, with analytical algorithms and complimentary databases (land usage, census data), to discern additional information regarding the respondents, their activities and mode of transport.

The most important advantage of using mobile positioning data is the opportunity to survey traffic in near real-time and over a large area, and, based on this, to develop monitoring systems and decision support systems (Jarv et al., 2014). Such benefits of using mobile data are the reason why scientists and civil engineers are testing, step by step, the possibilities of using mobile data in travel behaviour research and transportation modelling. However, the largest challenge when working with passive mobile positioning data is obtaining the data from MNO’s, as there are problems with legislation at both national and international level; privacy and surveillance concerns with clients and public, and business secrets of MNO-s. Nevertheless, the usage of mobile data in research and practice has seen a vast increase over recent years, and will probably continue to see further increase. Once mobile data has been obtained, the following methodological problems must be considered:
Related Content

Predictive Modeling of Surgical Site Infections Using Sparse Laboratory Data

Clustering Techniques Within Service Sector
brahim Yazici, Ömer Faruk Beyca and Selim Zaim (2017). *Applying Predictive Analytics Within the Service Sector* (pp. 74-87).
[www.igi-global.com/chapter/clustering-techniques-within-service-sector/177317?camid=4v1a](www.igi-global.com/chapter/clustering-techniques-within-service-sector/177317?camid=4v1a)

[www.igi-global.com/chapter/point-density-estimation-of-changes-in-income-polarization-in-tanzania-19922001/79691?camid=4v1a](www.igi-global.com/chapter/point-density-estimation-of-changes-in-income-polarization-in-tanzania-19922001/79691?camid=4v1a)

Measuring the Quality of Healthcare Services in Bangladesh
[www.igi-global.com/article/measuring-the-quality-of-healthcare-services-in-bangladesh/232323?camid=4v1a](www.igi-global.com/article/measuring-the-quality-of-healthcare-services-in-bangladesh/232323?camid=4v1a)