Available online at www.sciencedirect.com
ScienceDirect
Tourism Management 29 (2008) 469–486
TOURISM MANAGEMENT
www.elsevier.com/locate/tourman

Evaluating passive mobile positioning data for tourism surveys: An Estonian case study

Rein Ahas[a,*], Anto Aasa[a], Antti Roose[a], Ülar Mark[b], Siiri Silm[a]

[a]*Institute of Geography, University of Tartu, Vanemuise 46, Tartu 51014, Estonia*
[b]*Positium ICT, Niine 11, Tallinn 10414, Estonia*

Received 9 June 2006; received in revised form 19 April 2007; accepted 16 May 2007

Abstract

This paper introduces the applicability of passive mobile positioning data in studying tourism. Passive mobile positioning data is automatically stored in the memory files of mobile operators for call activities or movements of handsets in the network. For tourism studies we use database of the locations of roaming (foreign phones) call activities in network cells: the location, time, random ID and country of origin of the called phone. We describe the peculiarities of data, data gathering, sampling, the handling of the spatial database and some analysis methods, using examples from Estonia. The results proved that mobile positioning data has valuable applications for geographical studies. Correlations with conventional accommodation statistics in Estonia were up to 0.99 in the most commonly visited tourist regions. Correlations of positioning data with accommodation statistics were lower in regions with a high number of transit tourists and less tourism infrastructure. The results show that positioning data has advantages: data can be collected for larger spatial units and in less visited areas; spatial and temporal preciseness is higher than for regular tourism statistics. Random IDs allow one to study tourists' movements, for example to study typical routes of tourists of certain nationalities. The weaknesses of data are related to problems with accessing data, as operators do not wish to share data and because of privacy and surveillance concerns. Problem is also that positioning data is another quantitative dataset with limited features.

Keywords: Mobile positioning; Tourism; Geography; Space–time behaviour; Social positioning method; Estonia; Surveillance

1. Introduction

Tourism is of growing importance for the global economy, since mobility has risen rapidly during recent decades. Every tourism activity has a specific geography and temporal sequence, and contemporary tourism geography has an increasing field of study in this matter. In operationalising tourism flows, most of the research emphasises the supply–demand balance and visitors' behaviour, as well as their impacts on the physical environment in recent decades (Hall & Page, 2006). For the recording of tourism flows there are different data sources and methods such as border, transportation and accommodation statistics, questionnaire surveys and a plethora of modelling exercises. These quantitative datasets and methodologies are standardised nationally and internationally by organisations such as the World Travel and Tourism Council (WTTC), World Trade Organisation (WTO) and Eurostat as visitor monitoring and management tools, and do not raise many methodological questions. Nevertheless, conventional quantitative methods are too limited and restricted to answer complicated questions about international tourism flows in a globalising world. For example, the traditional statistics on tourist flows, such as border and accommodation statistics, do not provide researchers information concerning the choice of destination or the evaluation of objects of interest and the infrastructure visited. Also, in many European Union (EU) member states as in Estonia, border statistics are no longer recorded. Accommodation statistics often have problems with tax violations in Eastern European and other countries, and overnight stays do not show the daily

*Corresponding author. Tel.: +372 5035914; fax: +372 7375825.
E-mail addresses: rein.ahas@ut.ee (R. Ahas), ular.mark@positium.ee (U. Mark).

0261-5177/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.tourman.2007.05.014

geographical movement of persons. Tourism analyses play a more and more important role in forming tourism strategies, plans and marketing tools, in which money from public administration is very often used. Because quantitative tourism databases remain too simple, more and more tourism studies tend to follow qualitative approaches or focus on very narrow applied studies. This is stimulating a controversial debate on the role and potential of academic research and about methodological approaches in tourism research and geography (Casino & Hanna, 2000; Hall & Page, 2006).

Despite the comprehensiveness of qualitative approaches, tourism studies still need quantitative datasets in order to monitor tourism flows and to perform academic analysis of trends, perspectives and geography. Recent developments in information and communication technologies (ICT) such as geographical information systems (GIS) and digital databases are advancing surveying methods in geography and tourism studies. The GIS and relevant visualisation methods have many applications in tourism studies (Buhalis & Licata, 2002; Frihida, Marceau, & Thériault, 2004; Lew & McKercher, 2006; Wang & Cheng, 2001). One of the emerging subjects in geographical studies is connected with mobile (cellular) phone positioning datasets and location-based services (LBS) (Ahas, Aasa, Silm, & Tiru, 2007; Ahas & Mark, 2005; Ratti, Frenchman, & Pulselli, 2006; Spinney, 2003). Mobile positioning data has great potential for applications in space–time behaviour studies addressed in studying tourism geography, though there are various restriction and pre-conditions in ICT applications. Nevertheless, mobile positioning data tends increasingly to complement conventional data sources and destination management, as it is becoming crucial and more important how a destination utilises its resources.

The objective of this paper is to introduce and to evaluate the applicability of the passive mobile positioning data in studying tourism. Passive mobile positioning data is stored automatically in the memory files of mobile operators as billing information, technical references of hand-over or other relevant logs with the precision of network cells (Ahas, Aasa, Silm, et al., 2007). There is a huge amount of geographical information available on the hard drives of mobile operators, which can be used in geographical and other relevant studies. The paper assesses the strengths and weaknesses of the method, such as how to access those databases, issues of privacy, sampling and spatial resolution. We also illustrate our discussion with some examples from Estonian study projects with such data.

2. Data and methods

2.1. Specific terminology

Mobile positioning—tracing location coordinates of handsets via the cellular network. There are different methods for location, such as cell ID, triangulation direction or/and distance from antenna, determined using radio waves, A-GPS. Different positioning methods are used because of different network standards (GSM, CDMA, 3G) and for different purposes. The rapid development of mobile positioning began with regulations to locate emergency calls (US 911 bill).

Active mobile positioning data—mobile tracing data in which the location of the mobile phone is determined (asked) with a special query using a radio wave. In order to ask the location of certain phones, a special environment and permit from the phone holder is required. Active mobile positioning is used in emergency calls, "friend finder" and many other LBS applications.

Passive mobile positioning data—automatically stored in the log files (billing memory; hand-over between network cells, Home Location Register, etc.) of mobile operators (Ahas, Aasa, Silm, et al., 2007). The easiest–billing memory is recorded when a person uses a mobile phone (call activity). Passive mobile positioning data is normally collected with the precision of network cells. For the collection of passive mobile positioning data, mobile operators can aggregate anonymous geographical data from log files, ultimately not violating personal identity and privacy, and researchers can use it in surveys for scientific purposes. Issues of privacy and surveillance are very important aspects of this data, and are discussed later in this paper. Passive mobile positioning data has been used in many transport and urban studies. The information on the crowdedness of network cells is used for research, planning or traffic management. Several applications display visualised network information in the handset screen.

Mobile tracing—the location of a mobile phone is determined with special queries with determined time sequence using a radio signal.

Social positioning method (*SPM*)—analyses of the location coordinates of mobile phones (tracing or passive data) and the personal (social) features of phone holders for space–time movement analyses (Ahas & Mark, 2005).

Base station—the cellular network is based on base stations. A base station usually has one tower and several directed antennae in the tower. The radio coverage of many antennae forms a cellular network.

Cell ID—every network cell in a mobile phone network has a unique ID, and the location of a phone in the cell can easily be determined for every call activity. The location of the cell is normally determined with the base station which has one antenna (omni-round radio coverage) or several directed antennae (radio coverage sector, determined with bearing degrees).

The size of a network cell is not fixed; the phone normally switches to the closest antenna or that with the strongest radio coverage. If the network is crowded, the phones can be switched not to the nearest station but to any other in the neighbourhood. The optimal distance from handset to antenna in the GSM network is less than 60 km. The