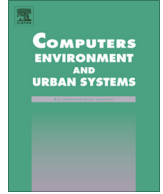




Contents lists available at ScienceDirect

## Computers, Environment and Urban Systems

journal homepage: [www.elsevier.com/locate/compenvurbsys](http://www.elsevier.com/locate/compenvurbsys)

## Urban sensing: Using smartphones for transportation mode classification

Dongyoun Shin<sup>a,\*</sup>, Daniel Aliaga<sup>b</sup>, Bige Tunçer<sup>c</sup>, Stefan Müller Arisona<sup>d</sup>, Sungah Kim<sup>e</sup>, Dani Zünd<sup>a</sup>, Gerhard Schmitt<sup>a</sup>

<sup>a</sup> Chair for Information Architecture, ETH Zurich, Wolfgang Pauli Strasse 27, 8093 Zurich, Switzerland

<sup>b</sup> Computer Science at Purdue University, 305N. University St., West Lafayette, United States

<sup>c</sup> Singapore University of Technology and Design and Delft University of Technology, 20 Dover Drive Singapore, 138682 Singapore, Singapore

<sup>d</sup> Future Cities Laboratory, Department of Architecture, ETH Zurich, 8092 Zurich, Switzerland

<sup>e</sup> Design Media and Technology Lab, Department of Architecture, Sungkyunkwan University, Suwon, South Korea

### ARTICLE INFO

#### Article history:

Available online xxxxx

#### Keywords:

Transportation mode classification  
Vehicle detection  
Social sensing  
Crowdsourcing  
Smartphone  
CITYing

### ABSTRACT

We present a prototype mobile phone application that implements a novel transportation mode detection algorithm. The application is designed to run in the background, and continuously collects data from built-in acceleration and network location sensors. The collected data is analyzed automatically and partitioned into activity segments. A key finding of our work is that walking activity can be robustly detected in the data stream, which, in turn, acts as a separator for partitioning the data stream into other activity segments. Each vehicle activity segment is then sub-classified according to the vehicle type. Our approach yields high accuracy despite the low sampling interval and does not require GPS data. As a result, device power consumption is effectively minimized. This is a very crucial point for large-scale real-world deployment. As part of an experiment, the application has been used by 495 samples, and our prototype provides 82% accuracy in transportation mode classification for an experiment performed in Zurich, Switzerland. Incorporating location type information with this activity classification technology has the potential to impact many phenomena driven by human mobility and to enhance awareness of behavior, urban planning, and agent-based modeling.

© 2014 Elsevier Ltd. All rights reserved.

### 1. Introduction

*“Much of our understanding of urban systems comes from traditional data collection methods such as survey by person or by phone. . . . They are hard to update and might limit results to a ‘snapshot in time’ Reades, Calabrese, et al., 2007”*

Urban planners may not satisfy with the data that is collected by the methods such as survey by phone or survey by Internet. This is because surveys usually collect data at a single point in time while modern cities become more complex and exhibit very dynamic conditions. Regarding the survey in transportation, people do not always make exact journey that they described on a survey. It cannot cover every trip, which is limited into collect average travel of the survey happened. It is also difficult to

measure changes unless two or more surveys are done at different temporal instances. People are easily tired of such frequent survey requests. Moreover, this repetition makes surveys expensive, time-consuming, and impractical.

In contrast, an emerging field of research uses mobile phones for “urban sensing” allowing us to collect scientific data in a new and innovative way (Cuff, Hansen, et al., 2008). Essentially, this is a type of “social sensing” with mobile devices enabling data collection from a large number of people in ways that were previously not possible. It provides the opportunity to track multiple data points in real time, and therefore to sample the dynamic behavior and inherent complexity of human activity within a city. The data points can be used to provide information and services not only to the individual user (e.g., location-based information), but also to urban planners who can use it to gain insight into the relationship between a city’s structure and its internal dynamics; for example, what are the patterns of people movement? how large is a person’s geographic living scope? how effective is a planner-proposed solution within the urban system? In addition, we can characterize the mobility pattern of a city including evaluating efficiency through

\* Corresponding author.

E-mail addresses: [shin@arch.ethz.ch](mailto:shin@arch.ethz.ch) (D. Shin), [aliaga@purdue.edu](mailto:aliaga@purdue.edu) (D. Aliaga), [bigetuncer@sutd.edu.sg](mailto:bigetuncer@sutd.edu.sg) (B. Tunçer), [arisona@arch.ethz.ch](mailto:arisona@arch.ethz.ch) (S.M. Arisona), [sakim@skku.edu](mailto:sakim@skku.edu) (S. Kim), [zuend@arch.ethz.ch](mailto:zuend@arch.ethz.ch) (D. Zünd), [schmitt@arch.ethz.ch](mailto:schmitt@arch.ethz.ch) (G. Schmitt).

<http://dx.doi.org/10.1016/j.compenvurbsys.2014.07.011>

0198-9715/© 2014 Elsevier Ltd. All rights reserved.

people's average moving distance, analyzing the change of transportation hot spots, estimating the favorite transportation mode, comparing with other cities, and monetizing the cost of socially-motivated inner-city travel.

Data capturing of human mobility become widely available in the real world by the advance of diverse location sensing technologies (Campbell, Eisenman, et al., 2006). However, despite its importance for urban planning and traffic forecasting, our understanding of the basic individual patterns of real-time human mobility remains limited due to the lack of tools to easily monitor and spatially-locate individuals. Moreover, it is hard to get insight into disaggregated travel styles since there is a lack of a medium or knowledge to automatically classify the different travel modes among the diverse transportation types.

Our belief is that automating the process of obtaining individualized and disaggregated human mobility will increase user participation in urban data collection and improve urban planning. For example, including transportation mode identification enables users to obtain an extended analysis of their travel patterns such as estimation of individualized CO<sub>2</sub> emission or measures of personal contribution to local transportation based on their trip diary. Urban planners will obtain more detailed and real-time observations of urban dynamics. Amongst many other possibilities, it will enable determining under which condition are vehicles or walking preferred and what kind of environmental characteristics affect the transportation choice.

In this context, we exploit an emerging field: mobile crowdsourcing.<sup>1</sup> We focus on automatically determining the travel mode used by an individual and analyze it to extract transport information. We have studied approximately 500 sensing samples from anonymous mobile phone users in Zurich, Switzerland whose acceleration measurements are logged during daily activities. For anonymity, all sensing data (e.g., acceleration and location) is collected by randomly generated ID numbers, which does not include any personal information.

Our prototype system and method provides answers to the following two key questions.

- **How to convert a mobile phone's sensor data into transportation information?** During movement, each travel mode (e.g., car, bus, tram, or train) has a different "signature" (e.g., rolling, waving, vibration, and/or acceleration pattern). Contrary to the general belief that a phone's acceleration measurements contain only information about when a vehicle leaves and stops, we found the cumulative set of acceleration measurements to be a very rich source of information that can be exploited to determine travel activities and transportation choices. Our method automatically classifies the travel mode by continuously inspecting readings from the mobile phone acceleration sensors as well as network-based approximate location data. While developing the transportation mode classification system, we found the heavy use of sensing and data processing to potentially cause fast battery drain. In fact, most previous research in mobility sensing (Hemminki, Nurmi, et al., 2013) recognizes the problem of high battery consumption. Thus, to minimize the impact on battery life and to permit long term sensing periods (e.g., weeks or months), our method focuses on providing a solution that does not use high power-consuming subsystems on a phone (e.g., its GPS system).

<sup>1</sup> Mobile crowdsourcing is a term that accounts crowdsourcing activities, which are supported by mobile devices. Thanks to the recent development in mobile devices such as diverse sensors and wireless network enables smartphone, the crowd-sourced information can be collected and shared without any further difficulties for a larger scale of analysis.

- **How to provide an overall sensing process without disturbing daily phone use?** The answer demands a fully-optimized design process (e.g., the solution must carefully orchestrate phone sensor types, sensing circles, processing procedure, detection algorithm, and server-user network usage). While most mobile sensing applications require user input, our entire sensing and data acquisition procedure operates in full automation. Users are not required to spend time manually typing out location information. The sensing data is collected and delivered to a central server, which subsequently measures and analyzes their urban activities. The computed results are transmitted back to the mobile phone and optionally displayed, thus providing a visual means to enhance awareness and environmental impact. Furthermore, we highlight two aspects of our approach, "network-based location sensing" and "low-frequency sampling", both of which help improve battery efficiency and method automation.

Altogether, the main objectives of this article include

- describing a mobile phone acceleration-sensor and network location-based transportation mode detection algorithm which collects information for human mobility,
- creating a prototype and preliminary deployment to collect field data in real urban environments, and
- providing an evaluation of the algorithm accuracy, battery efficiency, and a proposal for future uses and possibilities.

In the following, Section 2 will provide a summary of recent work addressing limitations and related problems. Section 3 will present the overall framework for the implementation, and Section 4 will describe our solution to automatic travel mode classification, and Section 5 will show statistical evaluations of our solution. Finally, we conclude with Section 6.

## 2. State of the art

Urban sensing (Campbell et al., 2006) enables us to collect scientific data in a new and innovative way. In fact, this is a very useful type of crowd-sourced sensing (Demirbas, Bayir, et al., 2010), which enables collecting urban data from a large number of people in ways that were previously not possible.

Miller (2013) stated that sensed transportation and geographic information can cultivate transportation systems where participants share information and resources to solve operation accessibility problems. Gould (2013) emphasized mobile phones will bring new opportunities enabling travel surveys as a medium of data collection.

Smartphones offer researchers the possibility to develop "crowdsourcing" applications that collect information from an extensive number of individuals and open the door to studying the dynamics of large populations (e.g., Joki, Burke, et al., 2007; Murty, Gosain, et al., 2008; Turner, White, et al., 2011). Moreover, smartphone-based data collection is becoming increasingly important to transportation planning (Barbeau, 2011; Berlingiero, Calabrese, et al., 2013) and to other planning agencies and private firms (e.g., Alt, Shirazi, et al., 2010).

Nevertheless, mobile crowdsourcing has not yet become a widespread medium for data collection because most of the existing applications require active user input. Since users typically do not receive any reward for their effort, they quickly lose interest and are not willing to collaborate (Santos, Cardoso, et al., 2010). A new research direction, called "social sensing", collects data without requiring any active user input (Adams, Phung, et al., 2008; Asakura & Hato, 2005; Kim, Shin, et al., 2012; Madan,