ARTICLE IN PRESS

Pervasive and Mobile Computing ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at SciVerse ScienceDirect

Pervasive and Mobile Computing

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

Review

Mining GPS data for mobility patterns: A survey

Miao Lin *, Wen-Jing Hsu

School of Computer Engineering, Nanyang Technological University, Singapore

ARTICLE INFO

Article history:
Received 26 November 2012
Received in revised form 17 June 2013
Accepted 26 June 2013
Available online xxxx

Keywords:
Mobility pattern
GPS data
Ubiquitous computing
Mobile computing

ABSTRACT

With the help of various positioning tools, individuals' mobility behaviors are being continuously captured from mobile phones, wireless networking devices and GPS appliances. These mobility data serve as an important foundation for understanding individuals' mobility behaviors. For instance, recent studies show that, despite the dissimilarity in the mobility areas covered by individuals, there is high regularity in the human mobility behaviors, suggesting that most individuals follow a simple and reproducible pattern. This survey paper reviews relevant results on uncovering mobility patterns from GPS datasets. Specially, it covers the results about inferring locations of significance for prediction of future moves, detecting modes of transport, mining trajectory patterns and recognizing location-based activities. The survey provides a general perspective for studies on the issues of individuals' mobility by reviewing the methods and algorithms in detail and comparing the existing results on the same issues. Several new and emergent issues concerning individuals' mobility are proposed for further research.

Contents

* Corresponding author. Tel.: +65 97270920.
E-mail addresses: linm0018@e.ntu.edu.sg (M. Lin), hsu@ntu.edu.sg (W.-J. Hsu).

1574-1192/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.pmcj.2013.06.005

Table 1
A comparison of different types of data repositories for the study of individuals' mobility. Velocity, pause time, inter-contact time (ict for short) are the basic features needed in describing individuals' mobility. Resolution indicates the degree of spatial accuracy of the data. Scale indicates the area covered by an individual's data. GSM cell-tower data contain the ict information, whereas both the wireless network traces and GPS data need manual labels to indicate the ict.

Data type	GSM cell-tower data	Wireless network traces	GPS data
Resolution	km	m	m
Scale	Metropolitan area	Campus/work place	Global
Velocity	No	No	Yes
Pause time	Approximate	Approximate	Exact
Ict	Exact	Labels needed	Labels needed
Path info	Rough	Rough	Exact

1. Introduction

Mining individuals' mobility patterns is an emergent research area. Owing to the prevalence of mobile phone, a wide variety of mobility-related data are now being captured, including, e.g., the usage of WiFi, and communications between individuals via sending messages or making phone calls. The availability of such data on personal devices will inevitably engender the study of individuals' mobility patterns at an unprecedented scale both in terms of the areas covered by the trajectories and also the number of individuals involved in the study. Many interesting results have been obtained [1–4]. For instance, despite great dissimilarities in individuals' mobility areas, ranging from a few square kilometers to more than a few thousands square kilometers, individuals' mobility is found to be highly regular [1]. Most individuals return to a handful of locations and share a rather universal probability density function for places visited [2]. After removing individuals' idiosyncrasies, the rules that govern the individuals in exploring new locations or returning to the previously visited locations are found to be similar [3]. Also, members of a same social group, such as researchers from the same labs, present similar mobility behaviors [4].

The study on mining individuals' mobility patterns have found applications in a wide range of researches, such as personal positioning [5,6], complex system [1,7], computational sociology [4,8], wireless network analysis [9–11], etc. Both individuals and the society as a whole can benefit from the study of individuals' mobility. For instance, understanding individuals' mobility can help recognize regular activities or abnormal events [12–15], predict future movements or destinations [2,5,16,17], identify similarities among individuals [18,19]. The study may also be applied in traffic forecast [20], city planning [21,22], mobile virus control [7], epidemic prevention [23–25], evaluating mobile network protocols [9–11], logistics management [26], energy consumption and carbon footprint [27,28].

Due to the coarse description of individuals' locations provided by the GSM cell-tower data, however, many existing studies of mobility patterns, are incomplete and less than convincing. In GSM cell-tower data, a location is indicated by the cell tower that provides the connecting services. Hence the resolution for the location is constrained by the service area of a cell tower, whose size is in the order of hundreds of square meters to several square kilometers. The paths, often indicated by a series of discontinuous sudden jumps, are hardly observable in fine details between the destinations.

Besides the GSM cell-tower data [29], some other data repositories, such as wireless network traces [30–32] and GPS data [33,34], have recently become available.[1] A comparison among the data repositories in presenting the basic mobility features, is summarized in Table 1. Specifically, GPS data allow to track the movements of individuals in latitude and longitude along with timestamp. GSM cell-tower data are constrained by the service area of cell towers, whereas wireless network is mostly available within even smaller regions, such as campuses or work places, although their accessibility is being improved gradually. By comparison, GPS devices are able to provide positioning data globally. GPS data, however, suffer from three major limitations, (1) GPS signals are usually blocked indoor or underground places, (2) GPS devices may get interferences near tall buildings, and (3) continuously collecting GPS data may consume devices' energy quickly. These limitations have prevented collection of continuous high quality GPS records over long durations.

This survey will focus on the mobility studies based on high resolution positioning data, GPS data in particular, with a target on researches that aim to uncover the patterns in individuals' movements, e.g., [37–39]. The study of individuals' mobility may be broadly classified into two subareas, (1) mining mobility patterns, and (2) constructing mobility models. Typically, the following mobility patterns are of great interest: locations of significance, modes of transport, trajectory patterns and location-based activities. The mobility models of great interest are the ones that emphasize on the statistical patterns of the individuals learned from trajectories [3,9–11], or i.e., the trace-based mobility models. Since there is an excellent survey on this topic [40], we will not include this topic in this paper.

The remainder of this paper is organized as follows. Section 2 to Section 5 focus on issues and results in mining various mobility patterns. Section 6 highlights a few new and emergent research issues. In Section 7, we briefly conclude this review.

[1] A spatial–temporal datasets generator is given in [35]. Since this survey targets on mining of real-world mobility patterns, the issue of the generation of synthetic dataset will not be addressed here. Nevertheless, one of the studies based on synthetic data is included in Section 4.1 [36].

Please cite this article in press as: M. Lin, W.-J. Hsu, Mining GPS data for mobility patterns: A survey, Pervasive and Mobile Computing (2013), http://dx.doi.org/10.1016/j.pmcj.2013.06.005