ARTICLE IN PRESS
Pervasive and Mobile Computing ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at ScienceDirect

Pervasive and Mobile Computing

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

Fast track article

Discovering and modeling meta-structures in human behavior from city-scale cellular data

Xiaming Chen [a,*], Haiyang Wang [a], Siwei Qiang [a], Yongkun Wang [b], Yaohui Jin [a,b]

[a] *State Key Lab of Advanced Optical Communication System and Network, Shanghai Jiao Tong University, China*
[b] *Network and Information Center, Shanghai Jiao Tong University, China*

ARTICLE INFO

Article history:
Received 22 March 2016
Received in revised form 30 September 2016
Accepted 2 February 2017
Available online xxxx

Keywords:
Meta-structure
Spatio-temporal patterns
Human mobility
Graph similarity
Behavior modeling

ABSTRACT

For a long time, researchers explore spatio-temporal properties in mobility to understand human behavior. They have discovered many statistical laws about human dynamics. Unfortunately, we still have limited knowledge about the spatio-temporal structure of individuals' movement at a large scale. In this paper, we studied the unified spatio-temporal structures (i.e., *meta-structures*) in human mobility. We hereby propose a meta-structure discovery algorithm by coupling both topology and spatio-temporal attributes of mobility graphs. With the construction of individual profiles from meta-structure analyses, we provided a novel mobility model from a process-driven perspective, which reduced the dependence of many existing models on the consistency between local and global mobility statistics. We gained some insights on the dominating meta-structures in human mobility by leveraging mobile data in a large city. The statistical distribution of meta-structures is found to be determined by the intrinsic heterogeneity of spatio-temporal properties in human behavior. Our model evaluation showed that a process with basic rules could demonstrate the key statistical properties in mobility meta-structures. We believe that these approaches and observations would be a good reference for management of human mobility in mobile networks and transportation systems.

1. Introduction

Understanding the properties of human mobility is very important for large-scale emergency detection [1], epidemic spread prediction [2], mobile network [3] and public transportation optimization [4]. The prevalence of smartphones with GPS helps to track the human mobile behavior at large scale. Due to the intrinsic complexity in human behavior [5–7], a basic method is to discover the spatio-temporal structure of human mobility [8,9], which requires robust algorithms [10,11].

Researchers have endeavored to discover human mobility in two ways. As in Fig. 1, one is to exploit the population statistics of human dynamics, as promoted by the pioneering work of Barabàsi et al. [5], which addressed the characteristics of population-level trip distributions; e.g., the visited location number follows a power-law distribution [12]. The other way is to mine the trajectory patterns from individual's movement history [13]. For example, with mobile phone data, a primary approach is transforming individual's records into a symbolic sequence [14] and searching for patterns with sequential pattern mining techniques.

* Corresponding author.
E-mail addresses: chen_xm@sjtu.edu.cn (X. Chen), jinyh@sjtu.edu.cn (Y. Jin).

http://dx.doi.org/10.1016/j.pmcj.2017.02.001
1574-1192/

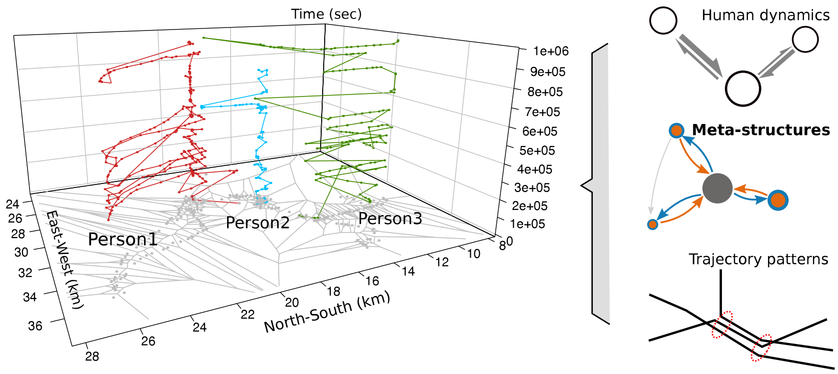

Fig. 1. Three user movement paths extracted from cellular trajectories. The Voronoi graph on the horizontal plain gives spatial distributions of frequently-visited cellular base stations of individuals.

Both ways have been widely accepted, but they may not be suitable for analyzing and modeling mobility with spatio-temporal structures. The human dynamics reports population-level characteristics with fitted parameterized distributions, which is convenient to compare mobility difference between two populations, yet less explainable and directive for realistic applications [1,4,2]. These methods might also suggest biased results about the structural difference between two small groups of people because of the high diversity in individual's mobility [6,10]. The trajectory pattern mining reserves the sequential information in movement rather than the mobility structures with spatial and temporal properties; e.g., people with the same topology structure might generate different trajectory patterns by considering their visit order and frequencies. For there reasons, we require a more robust representation of spatio-temporal patterns [9,8] in mobility.

In this paper, we revisit individual's mobility and model user mobility in cellular networks with a unified spatio-temporal structure, i.e., *meta-structure*. Studying the meta-structure in mobility not only complements existing discoveries about human behavior, but also paves a way towards understanding the connection between individual's mobility diversity and the change of macroscopic statistics.

In our model, individual's movement in a periodic interval is treated as a whole and reformatted into an *attributed graph* (or mobility graph) to unify both spatial and temporal information. A meta-structure is the common part shared by two or multiple mobility graphs. The benefits of this modeling are threefold: (1) **Spatial dependence**: As represented in a directed graph, the meta-structure encodes mobility similarity from not only the sequential displacements, but also the spatial dependence between locations. (2) **Temporal constraints**: The meta-structure contains an upper bound of the aligned periodic interval over mobility graphs and a lower bound of dwelling time for location preference. These bounds indicate energy limits for movement that are potentially connected with an ultraslow diffusion process [12]. (3) **Robustness**: Mobility mining usually suffers from both inaccuracy of data acquisition and intrinsic variability in mobility patterns. The meta-structure becomes robust to these difficulties by leveraging the maximal similarity in topology structure and attributes of mobility graphs.

To discover the meta-structures from cellular data, we first introduce a topology-attributes coupling similarity algorithm to derive the elementary (nodes and edges) similarities for two attributed graphs. The algorithm leverages the topology structure to broadcast similarity information to neighbors. A combined measure is then derived to solve the conflict between node- and edge-orient matching. The significant meta-structures are extracted via a clustering-based pruning technique and evaluated by the distance from proximal instances in the same cluster.

Mobility model is a vital proxy to involve empirical observations in an analytic manner for further usage. We propose a novel mobility model by leveraging spatio-temporal characteristics in mobility meta-structures within the intervening opportunity framework [15]. The modeled agent profiled by the local information in meta-structures employs an opportunity-driven process to plan his movement path in a simulation period. We evaluate the model performance by comparing its macroscopic statistics of modeled behaviors against empirical metrics derived from real movement data and previous observations.

Our model does not require the consistency between individual and population-level mobility statistics which is relatively necessary in previous models [16,12,9]. For example, the continuous-time random walk model [16] assumed that individual's dwelling time followed a scale-free distribution derived from the population-level statistics. However, we observe that most individuals' dwelling time could be well approximated by a two-mode mixture distribution. By involving the local profiles from meta-structures, our model could theoretically generate varying population-level statistics by changing the composition of simulated agent types. With this feature, the proposed model becomes a potential selection for future analysis of relationship between mobility diversity and macroscopic statistics.

Our contributions can be summarized as following:

1. *We provided an approach to discover significant meta-structures from multiple mobility graphs*. An iterative algorithm is proposed to extract the meta-structure by coupling the topology and attributes for both nodes and edges. We also

Please cite this article in press as: X. Chen, et al., Discovering and modeling meta-structures in human behavior from city-scale cellular data, Pervasive and Mobile Computing (2017), http://dx.doi.org/10.1016/j.pmcj.2017.02.001