Pervasive and Mobile Computing 9 (2013) 629–642

Contents lists available at ScienceDirect

Pervasive and Mobile Computing

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

Combination of spatio-temporal correction methods using traffic survey data for reconstruction of people flow

Yoshihide Sekimoto*, Atsuto Watanabe, Toshikazu Nakamura, Hiroshi Kanasugi, Tomotaka Usui

Center for Spatial Information Science, University of Tokyo, 5-1-5, Kashiwanoha, Kashiwa-shi, Chiba, Japan

ARTICLE INFO

Article history:
Available online 2 November 2012

Keywords:
Reconstruction of people flow
Traffic survey data
Large-scale spatio-temporal analysis

ABSTRACT

Data on people flow has become increasingly important in various fields, including marketing and public services. Although mobile phones enable the user's position to be located with a certain degree of accuracy from a large number of people and become one of the most promising devise, unwillingness to share related with privacy issues still remain. Therefore, it is also important to establish a practical method for reconstructing people flow from various kinds of existing fragmentary spatio-temporal data, such as public traffic survey data, from a view of complementariness with mobile phone data. In this study, we propose a combination of spatio-temporal correction processes to a previously published method, to generate continuous spatio-temporal people flow data sets at chosen intervals in selected cities. The correction methods include temporal smoothing of departure time using kernel density estimation, network data correction in OpenStreetMap data, and spatial smoothing in geocoding with MODIS data. We also compare the reconstruction accuracy by deriving correlation coefficients for different combinations of correction methods. Such reconstructed people flow data can potentially be used as infrastructure data in various fields, including emergency planning and related events in areas where data collection and real-time awareness are weak.

1. Introduction

Recently, the monitoring of dynamic changes in people flow has become necessary in order to mitigate secondary disasters following earthquakes, fires, or other major events, as well as to relieve congestion at nodes in terminal stations. For example, more than 12,000 people were killed while trying to escape from the tsunami of the Great East Japan Disaster on March 11, 2011. As another example, consider the Shinjuku Station in Tokyo; about 4 million people on average ride trains to and from this station on a daily basis, making it the most crowded station in the world. It is necessary for public facility managers to have comprehensive information about the flow of people, if they are to design safe and comfortable spaces, and devise and implement appropriate urban transport policies. Such information is also valuable in the commercial fields of outdoor advertising; price systems governing effective advertising, depend on the traffic volume of each location.

A number of previous studies have looked at population distribution data. For example, the National Center for Geographic Information Analysis (NCGIA) produced the Gridded Population of the World (GPW) in 1995 as the first raster global data set of population [1]. A second version (GPW2) was produced in 2000 by the Center for International Earth Science

* Corresponding author. Tel.: +81 4 7136 4308.
E-mail addresses: sekimoto@csis.u-tokyo.ac.jp (Y. Sekimoto), atsuto@csis.u-tokyo.ac.jp (A. Watanabe), ki_ki_gu@csis.u-tokyo.ac.jp (T. Nakamura), yok@csis.u-tokyo.ac.jp (H. Kanasugi), usui@csis.u-tokyo.ac.jp (T. Usui).
URL: http://shiba.iis.u-tokyo.ac.jp/member/sekimoto/index.html (Y. Sekimoto).

1574-1192/$ – see front matter
doi:10.1016/j.pmcj.2012.10.005

Information Network (CIESIN) at Columbia University, which used a higher resolution.[1] Data are gridded by country on the basis of original census units at the highest spatial resolution they are available at. A smoothing method, which assumes that grids closer to units with high density are higher, is used. Similar information is provided by LandScan, developed by the Oak Ridge National Laboratory (ORNL). LandScan models are tailored to match the data conditions and geographical nature of each individual country and region. At a resolution of approximately 1 km (30″ × 30″), they provide the finest resolution of global population distribution currently available [2]. Whereas GPW provides the nighttime population, LandScan represents the ambient population (the average over 24 h).

However, recent developments in sensing technology have introduced various ways to measure people flow more dynamically. Especially, mobile phone sensing can be widely applied for taking dynamic measurements, and some research exists in this area [3–6]. Moreover, acquisition/process costs of spatio-temporal positions can be reduced, and mobile phone will be one of the most promising devise. However, some privacy issues related to the mobile phone still remain, from a view of infrastructure data set that can give an overview of the mass flow by integrating the various acquired data mentioned in the previous paragraph. In this sense, utilization of public survey data or other participatory data are also very important, not depending mobile phone data only. This means that combination or complementariness of various data sources will lead more resilient infrastructure data set.

Hence, our group proposed a data process for the reconstruction of spatio-temporal positions of large numbers of people using existing person trip survey (hereinafter referred to as ''PTS'') data, collected by various transportation planning agencies or commissions for several Japanese cities [7,8]. Some results are summarized on the ''People Flow Project'' website.[2] However, some interpolations and corrections will be necessary when applying our proposed core method [7] to traffic survey data from different regions of the globe, because each city's zones and network data vary in size.

In this study, we apply the proposed process to PTS data of developing cities provided by the Japan International Cooperation Agency (JICA), and in doing so, resolve difficulties in preparing infrastructure data, such as road network data. Section 2 provides a problem statement of our existing method, an outline of our proposed method, and a brief discussion of the JICA-PTS data. Section 3 describes the proposed interpolation and correction methods, and gives the results for each. Section 4 provides a comparison of reconstruction accuracies for different combinations of methods, and Section 5 concludes.

2. Preliminaries

This section describes current problems based on our existing methods, applying to JICA-PTS data. Actually, our existing method is explained in 2.1, and JICA-PTS data is briefly described. Finally, we state some problems for the application and provide standpoint for our approach of this study in the next section.

2.1. Existing method: a basic reconstruction framework using a transportation survey [7]

To date, we have applied the core concept to Japanese PTS data. In doing so, we performed spatio-temporal reconstruction based on fragmentary spatio-temporal location information, such as person trip surveys, to reproduce the spatio-temporal positions of a large number of people at high-resolution time intervals of a minute. The reconstruction process can be summarized in three steps. First, place information must be converted to latitude and longitude (hereinafter, ''lat/lon'') from each PTS zone through an address matching process. Second, each route must be selected according to the origin and destination (hereinafter, ''OD'') positions of sub-trip information on the basis of road and railway topology. Third, the spatio-temporal position needs to be interpolated according to the form of the people flow data set, on the basis of detailed road and railway geometry.

Fig. 1 is a reconstruction result cited from [7], and illustrates each step of the process for the spatio-temporal interpolation of positions from OD data using one trip in Central Tokyo. Fig. 1(a) shows data for a trip consisting of three sub-trips for one person after the address matching process. Fig. 1(b) shows route selection from the OD data of each sub-trip using road and railway topology (the latter, in particular, is linked with the timetable). Fig. 1(c) illustrates spatio-temporal interpolation at 1 min intervals along detailed road and railway geometry.

We have two key points about static infrastructure data, which strongly support the feasibility of this reconstruction process. One is to use railway timetables as the topology in the route selection process. Choosing railway routes is also inevitable in order to reconstruct the total people flow, which is not limited to vehicle flow as dynamic data. We consulted the railway timetables for time data between any two stations throughout Japan. These are available at a relatively low cost from an API (Application Program Interface) provided by Val Laboratory Co., Ltd.[3] The other key point is to interpolate along detailed network geometry data. For road and railway networks, we used DRM (Digital Road Map) data,[4] including 4.67 million road network links throughout Japan, provided by Sumitomo Electric System Solutions Co., Ltd.

1 http://sedac.ciesin.columbia.edu/gpw/wps.jsp.

2 http://pflow.csis.u-tokyo.ac.jp.

3 http://www.ekiworld.co.jp.

4 http://www.drm.jp.