Research Article

The rise of big data on urban studies and planning practices in China: Review and open research issues

Jinwei Hao[a,*], Jin Zhu[b], Rui Zhong[c]

[a]*College of Architecture and Urban Planning, Tongji University, Rm C606, No. 1239, Siping Rd, Shanghai 200092, PR China*
[b]*Shanghai Urban Planning and Design Research Institute, No. 331, Tongren Rd, Shanghai 200040, PR China*
[c]*Urbanisation and Urban Rural Planning Research Center of Jiangsu, Floor 11, No. 88, Caochangmen Ave, Nanjing, Jiangsu 210036, PR China*

Abstract

The year 2014 marked the rapid expansion of big data in urban studies and planning practices in China. Big data has advantages of revealing individual characteristics rather than a general feature by traditional statistics, and it is consistent with the idea of people-oriented urbanization and urban–rural planning. The research progress of big data since 2000 in China is reviewed in this paper. Focusing on behavior big data mining and big data application in urban studies and planning practices, the review is proposed from four parts, involving behavior data acquisition & analysis, spatial analysis, plan making & management application and new methodologies with big data. Lastly, some open research issues such as potential challenges and possible directions of development are discussed.

1. Introduction

1.1. Big data in urban studies and planning practices

The year 2014 marked the rapid expansion of big data in urban studies and planning practices in China. The Annual National Planning Conference of China, as well as The Forum of China Urban Planning Studies, both involved big data as one of main subjects. Meantime, main academic journals in urban studies and planning practices in China like *Urban Planning International*, *Planners* and *Shanghai Urban Planning Review*, all introduced big data column to comply with the trend. Undoubtedly, big data is becoming a hot topic frequently discussed in urban studies and planning practices in China.

It should be noted, though, that this rise of big data is not only embodied in traditional academic achievements boom, but also in new characteristics covering academic community expansion and amalgamation, self-organized

research institutes development, and academic-industry-media integration. Specifically, academic community expansion is embodied in a deeper cooperation among geography, urban planning and information science, as well as a more frequent interaction between traditional research institutes and IT companies, such as Baidu, Alibaba, etc. Research institutes development is embodied in emergence of new self-organized research institutes, like Beijing City Lab (BCL), CITYIF, Xicheng & Tsinghua-Tongheng Urban Data Lab, Metro Data Team as well as a leading development of traditional research institutes represented by universities and urban planning institutes. Academic-industry-media integration is embodied in the emerging of new media represented by The Paper-Shi Zheng Ting (澎湃-市政厅), CAUP. NET (国匠城),Urban Data Party (城市数据派) and Fruitalk (果说). At the same time, Wechat group introduced a new approach in academic communication, cooperation platform construction, and public popularization of planning, contributing to expand related disciplines' social reputation and influence.

This paper reviews literatures of big data application in China in Chinese and English in the fields of urban studies and planning practices since 2000, as well as their references. Focusing on behavior big data mining and big data application in urban studies and planning practices, the review is proposed from four parts, involving behavior data acquisition & analysis, spatial analysis, plan making & management application and new methodologies with big data. Lastly, some open research issues such as potential challenges, possible directions of development are discussed as well.

1.2. Literatures selection and classification

According to sample size and update frequency of data, related literatures exploring regional & urban issues and planning issues through big data in China are selected in this review. This paper defines high-frequency with large-sample-size data and low-frequency with large-sample-size data as big data, which includes GPS Log Data from Handheld GPS Devices, Mobile Phone Data (MPD), Smart Card Data (SCD), GPS Data from Floating Cars (Taxis), Point of Interests (POIs), Volunteered Geographic Information (VGI), Search Engine Data, detailed digital landuse data, parcel data and road network data displayed by GIS, CAD, etc. Literatures published after the year of 2000 in Chinese and English both are reviewed in this paper. Literatures in Chinese mainly come from cnki.net, udparty.com, thepaper.cn, as well as Symposium on big data and spatio-temporal behavior planning held in Tongji Univ., Jan., 2015, while literatures in English mainly derive from Web of Science and Working Papers on beijingcitylab.com.

(1) Keywords searched in cnki.net include "big data AND cities (大数据AND城市)", "big data AND planning" (大数据AND规划), as well as (GPS/mobile phone (手机) /smart card(s) (公交刷卡) /floating car (浮动车) /taxi (出租车) /POI (兴趣点) /VGI (志愿地理信息) /weibo (微博) /checkin (签到) /search engine (搜索引擎)) AND (cities (城市) /spatial (空间) /spatiotemporal (时空)/ mobility (移动) /planning (规划)) in TITLE. Literatures selected must include at least one word in the first group AND at least one word in the second group. Then the related references of searched literatures also are selected.
(2) Keywords searched in Web of Science include (GPS/mobile phone/smart card(s)/floating car/taxi/POI/VGI/ checkin/search engine) AND (urban/spatial/spatiotemporal/mobility/planning) in TITLE. Then the related references of searched literatures also are selected.
(3) Related literatures on beijingcitylab.com, udparty.com, and Symposium on big data and spatio-temporal behavior planning hold in Tongji Univ., Jan., 2015.

Literatures selection mainly emphasizes journal articles, excluding monograph, academic dissertation, and some conference articles. Thus, we get 189 literatures in total, involving 183 core literatures and 6 representative literatures not directly relevant to the theme. Among 183 core literatures, 126 ones are in Chinese while 57 ones are in English.

1.3. Literatures analysis

These 183 core literatures have been analyzed from multi-aspects:

(1) Publication Quantity: Number of literatures has begun increasing since 2008, especially presenting a rapid increase trend after 2011. And literatures number has reached as high as 47 in 2014 and 41 by Apr., 2015 (see Fig. 1).