Cities 55 (2016) 9–21

Contents lists available at ScienceDirect

Cities

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

Urban function connectivity: Characterisation of functional urban streets with social media check-in data

Yao Shen *, Kayvan Karimi

Space Syntax Laboratory, Bartlett School of Architecture, University College London, England, United Kingdom

ARTICLE INFO

Article history:
Received 9 November 2015
Received in revised form 15 March 2016
Accepted 25 March 2016
Available online xxxx

Keywords:
Network accessibility
Connectivity
Social media check-in data
Land use
Street network
Urban design

ABSTRACT

Social media check-in data, one type of crowdsourcing open data about individual activity-related choices, provides a new perspective to sense people's spatial and temporal preference in urban places. In this paper, through the analysis of the interaction between these scored places on streets, we aim to advance our knowledge of network accessibility with social media check-ins to portray urban structure and related socioeconomic performance more explicitly. By conceptualising an interface graph to reflect the interplay between land-use points and the co-visual paths, we propose a novel framework to characterise the urban streets with land-use connectivity indices that are measured with a new type of place-function signature. A "3-Ds" model is introduced to package three principal dimensions of urban function network, including accessible density, accessible diversity and delivery efficiency, as one integrated index that works towards a comprehensive understanding of function connectivity from each street's midpoints to all reachable land-use points. Streets are further partitioned to the annotated function regions based on function connectivity in different types of active land-use. The results of preliminary studies in the city of Tianjin, China show that the proposed metrics can explicitly describe the inherent function structure and the regions' typology across scales. Compared with space syntax measurements at the same radius for describing the variation of empirically observed house price, the integrated metric can improve the predictability of statistic models sufficiently, and each specified index is confirmed to be statistically significant by controlling other factors. Overall, this research shows that the usage of ubiquitous big social media data can enrich the current description of the urban network system and enhance the predictability of network accessibility on socioeconomic performance.

1. Introduction

The recent growth in the provision of location-based services has inspired people to share their location preferences in social media networks, through which the ubiquitous user-generated information of location choices has produced a new generation of 'human knowledge' regarding urban spaces at a fine-grained resolution (Wyly, 2014). Although such a provision offers new opportunities for research, the use of social media data has its limitations, including sampling problems, context-related uncertainty, and lack of theoretical composition (Boyd & Crawford, 2012). However, the finer resolution of these datasets has the potential to enable people to ask different questions than those based on conventional data that is oftentimes aggregated and out of date (Shelton et al., 2015). As a new type of fine-scaled datasets that contains detailed information about urban land-use, Points-of-Interest (POIs) and Check-in data in modern social media are a new focus in urban studies, resulting in research that identifies urban regions with POIs and taxi trajectories (Yuan et al., 2012), that determines the characteristics of urban parcels using vector cellular automata (Liu & Long, 2015), that maps urban areas of cities (Long et al., 2015), etc. Simultaneously, geo-tagged data have been increasingly discussed with reference to modelling human mobility patterns. Studies have applied individual-based check-in datasets to map mobility patterns (Hasan et al., 2013), to calibrate the parameters of distance decay (Wu et al., 2014), to infer daily activity clusters (Hasan & Ukkusuri, 2014; Jiang et al., 2012), to validate retail store replacement (Karamshuk et al., 2013) and to analyse urban structure (Long & Thill, 2015; Ratti et al., 2010; Zhong et al., 2015). All these efforts imply that now the open 'big data' can represent the variations of place in people's minds. However, no prior study has examined the function connections among various land-uses based on social media check-in data, other than Liu et al. (2016), who classified land-use clusters using spatial interaction patterns between parcels. The aforementioned studies have instead focused more on origin–destination patterns and have not considered the impact of physical layout on the detailed spatial interaction between real urban public spaces in their models that limit the contribution of location-based 'big data' to urban morphology, land-use planning and urban design studies.

* Corresponding author at: Bartlett School of Architecture, University College London, 140 Hampstead Road, NW1 2BX London, United Kingdom.
E-mail address: y.shen.12@ucl.ac.uk (Y. Shen).

http://dx.doi.org/10.1016/j.cities.2016.03.013
0264-2751/

An important issue that should be noted in any urban design and planning task is that streets are the fundamental spatial elements for movements; they interlink urban functions physically and cognitively (Gehl, 2011; Jacob, 1993). Urban function locations, as the destinations of urban activities, interact with one another through a network of urban public spaces and formulate a network-based land-use system. Consequently, the distance-based and cognitive proximity between urban land-uses influence not only the movement but also the choices of land-use locations (Geurs et al., 2015). A vibrant urban location can revitalise the urban context, in which it is embedded; additionally, the proximity of active urban locations can increase the popularity of the places that are connected to them. Nevertheless, few examples have focused on the importance of the role of streets and the physical layouts in the traditional models of accessibility, which are based on assumptions that do not recognise spatial heterogeneity in the urban space (Batty, 2009). This gap has been addressed with the configurational studies, such as the work undertaken by researchers from the space syntax community. These studies have demonstrated that the network properties of urban grids can adequately capture the influence of cognitive efforts on pedestrian movement patterns (Hillier et al., 1993), car volume distributions (Hillier & Iida, 2005), land-use distributions (Penn & Turner, 2004; Scoppa & Peponis, 2015; Shen et al., 2013) and other socioeconomic issues (e.g., Hillier, 2007; Karimi, 2012; Vaughan, 2007).

There has been some criticism of configurational studies for not considering the impacts of land-use distribution or other attractors on the spatial network analysis (Ratti, 2004), but these criticisms seem to lack a certain depth in understanding of theoretical and methodological propositions of space syntax (Hillier & Penn, 2004). In fact, the space syntax model is capable of offering significant potential for further development precisely because it links cognitive costs to navigational energy expenditures in spatial analysis (Kim & Penn, 2004). However, the topological/geometrical interaction between the land-uses through streets is an important dimension in which to scrutinise the underlying structures of functional streets that are typically neglected in conventional studies on land-use distribution and accessibility (Geurs et al., 2012). Recently, by taking into account the reachable densities of activities distribution in the space syntax model, Ståhle et al. (2005) developed a toolbox called 'place syntax' to calculate accumulated opportunities within the buffers defined by the metric/topological/geometrical radius. With the emphasis on the value of perceived density, Marcus and his team suggested that the space syntax model could be extended to a more general concept, the 'spatial capital', with the possibility of translating the urban form to other social, economic and cultural capitals (Berghauser Pont & Marcus, 2014; Marcus, 2010). Simultaneously, other areas of focus include modelling the interplay between reachable metric distance and directional distance to enhance standard space syntax in predicting human pedestrian patterns (Ozbil et al., 2011; Peponis et al., 2008), analysing transit riderships (Ozbil et al., 2009), and modelling the pattern of commercial frontage (Scoppa & Peponis, 2015). These studies implicitly considered the detailed possibility of improving the existing space syntax model, but did not propose a systematic perspective.

In this paper, we propose an original method for computing urban function connectivity by considering the spatial interaction between the scored urban spaces and partitioning the urban streets based on the composition of the defined spatial interactions. Social media check-ins are used to infer the significance of a place for a specific type of active urban function, to weight respectively the *accessible density* and *accessible diversity*, and to measure the *delivery efficiency* and the so-called *urban function connectivity*. A statistical data mining approach is adopted to characterise urban streets with the similar composition of function accessibilities for different types of land-uses. The proposed method is applied in a case study in Tianjin and its feasibility is verified by confirming the enhancements of the predictability of the statistic models that capture explicitly the variation of the house prices.

2. The method

2.1. Preliminary definition

In this study, *urban function connectivity* (*UFC*) is defined as the relatedness information between land-uses through the street networks, representing the sense of function potentials from every street's midpoint to all the reachable land-use points. This particular form of connectivity, therefore, is constructed on the basis of the street network where urban land-uses are assigned spatially. An *urban function region* (*UFR*) is identified as a group of places where the properties of function connectivity for different active land-uses are similar. Apart from the conventional definition of the functional region for comparing economic development in regional studies (Antikainen, 2005; Williams et al., 2010), we use this format of UFRs to refer to the clusters of streets within which urban functions operate similarly. Given this definition, we introduce an alternative approach to partition urban space from the bottom up by considering the spatio-functional relationships in a specified land-use system.

The land-use system in this study is conceptualised as a *path-point model* (PPM), or as a '*network interface model*' (NIM) to abstract the co-existential relationship between urban function points and the visual paths as graphs. In such a model, scored urban function locations (points) are assigned to the nearest paths based on their spatial interlinkage which is identified as the interface between buildings and public spaces (Alexander et al., 1977; Hillier & Hanson, 1984). By converting the spatial relationship between the main elements in PPM/NIM to edges and nodes, the land-use system can be transformed to an interface graph/network. The land-use locations and the directly visible street segments are defined as 'function nodes' and 'segment nodes', respectively, whereas the interfaces (the directly physical relationship) between nodes – including the entrances from the street to the locations and the intersections between the roads – are identified as 'entrance edges' and 'intersection edges'.

Fig. 1 illustrates the basic conceptual method used to construct an interface graph step-by-step. We first prepare the necessary data maps including the road network and land-use pattern so that these entities can be transformed into function and segment nodes with the entrance edges in an interface map on the basis of their interface connections. In the following stages, the dual graph of the interface network is created by converting the street junctions to the intersection edges that connect the segment nodes and assigning the cognitive cost at every junction to the graph as the weights of those intersection edges. The cognitive cost for the intersection edges is specified as the angular change at each junction according to space syntax theory (Dalton, 2000, 2001; Haq, 2003; Hillier & Iida, 2005; Kim & Penn, 2004; Turner, 2001a) and earlier evidence in the field of cognitive neuroscience and way-finding (e.g., Bailenson et al., 1998, 2000; Crowe et al., 2000, Montello, 1991). Using angular-weighted adjusted graphs in a simple land-use system, we represent the manner in which the angle change through a journey along the shortest path is calculated (Fig. 2). As the current evidence suggests that humans are not sensitive to very slight directional change (Figueiredo, 2009), a cut-off angle is used to filter the imperceptible angular deviations (α) from straight lines to enable a more appropriate approximation of the real movement decision making. Urban streets are the basic spatial units for the function connectivity model, as they are the real conduits for human movement.

Notably, scores or any other information can be used to weight the function nodes to capture the various levels of the significance of urban functions. In this research, check-ins and POIs derived from social media service providers are adopted to present the diverse types of urban activity locations and the proxies for the relative preferences of people in urban destinations. By adding weights for the function nodes, many aspects of function connectivity can be addressed to develop a comprehensive and robust methodology.