M.Sobana *et al*, International Journal of Computer Science and Mobile Applications,
Vol.3 Issue. 8, August- 2015, pg. 80-86 **ISSN: 2321-8363**

FUNCTIONALITY BASED REGIONS DISCOVERY IN A CITY BASED ON HPC FRAME WORK

R.Gayathri[1] M.C.A., M.Phil.
M.Sobana[2]

1. Assistant Professor, Department Of Computer science, Kaamadhenu Arts & Science College.
2. M.Phil Scholar, Kaamadhenu Arts & Science College.

Abstract:
The city development is gradually fosters by different functional regions, such as educational areas and business districts. we propose a HPC framework that Finds Regions of different Functions in a city using human mobility , points of interests (POIs) located in a region and check-ins. A city segmented into disjointed regions. The functions of each region are inferred by using a topic-based inference model. By this a region is represented by a distribution of functions, and a function is featured by a distribution of mobility patterns. The intensity of each function is identified for different locations. The results generated by this framework benefits' a variety of applications. This method is evaluated using large-scale and real-world datasets. The results justify the advantages of our approach over baseline methods solely using POIs or human mobility.

Keywords: human mobility, point of interest, functional regions, check-in.

1. INTRODUCTION

The urbanization and civilization leads to various functional regions in a city, e.g., residential areas, business districts, and educational areas, which support different needs of people's urban lives and serve as a valuable organizing technique for framing detailed cognition of a metropolitan. These regions may be artificially designed by urban planners, or naturally formulated according to people's actual lifestyle, and would change functions and territories with the development of a city.

Here, it is aimed to discover regions of various functions in urban areas using Human Mobility, Point of Interest, and Check-in (HPC) framework. A city is partitioned into individual regions. This partition is achieved by major roads, like high way and ring roads (refer to Figure 1(a)). Human mobility is traced by people's movement trajectories, which can be cell-tower delineates in a cellular network, or trajectories of driving routes, or a sequence of posts (like geo-tweets, geo-tagged photos, or check-ins) in location-based services. POI is associated with a coordinate and a category like restaurants and shopping malls. Check-ins constitutes different locations sharing the same temporal distributions.

Figure 1 shows the identity of the functionality intensity of each function in different locations. This is motivated by the following observations. On certain occasions, only a part of a region contributes to a function. But then, a function could be formulated across several individual regions (e.g., a shopping street). Each function is titled with some tags in a semi-manual way based on the output.

M.Sobana *et al*, International Journal of Computer Science and Mobile Applications,
Vol.3 Issue. 8, August- 2015, pg. 80-86 **ISSN: 2321-8363**

Fig 1 Functional Regions

Discovering regions based on the variety of functions can enable different worth full applications. 1st, it can provide people with an immediate understanding of a composite city (like New York City, Tokyo, and Paris) and social recommendations. For example, tourists can easily differentiate some scenic regions from business concern given these functional regions, thereby reducing effort for trip planning. Local people can also flourish their knowledge about a city by determination of regions that have exchangeable functions (e.g., entertainment areas). It is very common that local people will not understand each part of a metropolitan even if they have been in the metropolitan for a several years. The urban planning of a city is calibrated by these functional regions which contribute to the future planning. A city does not evolve as it is in original plan, given the complexity of urban planning itself and there is a difficulty in predicting the development of a city. Third, these functional regions will benefit location preferring for a business and advertisement. For instance, while building a supermarket there is a need to consider the distance to the residential areas. Shop planners would benefit from this information, when making a selection of possible sites for daytime dependent business.

2. Related Work

Urban computing with taxicabs: Urban computing is emerging as a concept where every sensor, device, person, vehicle, building, and street in urban areas can be used as a component to probe city dynamics and further enable a city-wide computing for serving people and their cities. The increasing availability of GPS-embedded taxicabs provides us with an unprecedented wealth to understand
human mobility in a city, thereby enabling a variety of novel urban computing research recently. For example, Gee al. [7][6]and Yuan et al. [20] respectively study the strategies for improving taxi drivers' income by analysing the pick-up and drop-off behaviour of taxicabs in different locations. [19] aims to find the practically fastest driving route to a destination according to a large number of
taxi trajectories, and Zheng et al. [21] glean the problematic urban planning in a city using the corresponding taxi trajectories. Based on the traffic flow represented by taxi trajectories, the technology for detecting anomalies in urban areas has been reported in [10].

The work presented in this paper is also a step towards urban computing. Different from the above-mentioned research, however, we focus on the discovery of functional regions in a city, which we have never seen before in this research theme .*Discovery of functional regions*: Functional regions [1] have been studied in traditional fields of GIS and urban planning for years, as the discovery of them can benefit policy making, resource allocation, and the related research. [8] gives a good survey on the related literatures which are mainly based on clustering algorithms. Some algorithms classify regions in urban area based on remote-sensor data, as thoroughly compared in [16]. Other network-based clustering algorithms (e.g., spectral clustering), however, employ interaction
data, such as the economic transactions and people's movement between regions.

Recently, a brunch of work aims to study the geographic distribution of some topic in terms of user-generated social media. For example, Yin et al. [18] study the distributions of some geographical topics (like beach, hiking, and sunset) in USA using geo-tagged photos acquired from Flickr. Poz noukhov et al. [12] explore
the space-time structure of topical content from a large number of geo tweets.

The social media generated in a geo-region is still used as static features to feature a region. On the other hand, a few literatures have reported that human mobility can describe the functions of regions. For instance, Qi et al. [13] observe that the getting on/off amount of taxi passengers in a region can depict the social activity dynamics in the region. Our work is different from the research mentioned above in the following aspects. First, to the best of our knowledge, our method is the first one that simultaneously considers static features (POIs) of a region and interactions (human mobility) between regions when identifying functional regions. Second, rather than directly using some clustering algorithm, we propose a topic-model-based solution which represents a region with a distribution of functions. The function distribution is more practical than a single function for a region. Moreover, it reduces the data sparseness problem before clustering regions. We justify the advantage of our method over just using clustering approach in the experiments.

3. HSC FRAMEWORK

HSC Framework implementation involves the following steps.

3.1 Road Network Construction

A road network is comprised of major roads like highways and ring roads. This will partition the city into regions. For example, as shown in Figure 1, the red segments denote freeways and city expressways in Beijing, and blue segments represent urban arterial roads. The three kinds of roads are associated with route level 0, 1, and 2 respectively (in a road network database), forming a natural segmentation of the urban area of Beijing. These segmented regions are named as a "formal region". The road network is represented by raster-based model and map segmentation is represented by morphological image processing techniques.

3.2 Analogy Building

A topic model-based method is used to identify the functions of individual regions. Individual regions are obtained by morphological image segmentation approach. This method considers a region as a document, a function as a topic, uses human mobility related to the region as words, and treats POIs located in a region as metadata (like titles, authors, affiliations, and key words). Which result in manner where distribution of topics (functions) represents region and a distribution of words represents each topic. As shown in Table 1, an analogy is drawn between discovering functions of a region and the topic discovery of a document.

Transition cuboids → vocabulary
Formal regions → documents
Function of a region → topic of a document
Mobility patterns → words
POI feature vector → metadata of a document

Table 1: Analogy from region-functions to document-topics

A POI is commemorated with a tuple (in a POI database) consisting of a POI category (as listed in Table 2), name and a geo-position (latitude, longitude). For each formal region nr, the number of POIs in each POI category can be counted.
The *frequency density* vi of the i th POI category in r is calculated by:

$$vi = \frac{\text{Number of the POIs of the ith POI}}{\text{Area of region r}}$$

and the POI *feature vector* of r is denoted by x r = (v1, v2, . . . , v F , 1)where F is the number of POI categories and the last "1" is a default feature to account for the mean value of each topic. The POI feature vector is esteemed the metadata of each region, which is an analogue of the observed features such as author/email/institution of a document. Basic LDA model along with the mobility patterns are used to identify region topics.

Table 2: POI category taxonomy
code POI categor

POI category taxonomy

code	POI category
1	car service service
2	car sales
3	car repair

4	motorcycle service
5	Café/Tea Bar
6	sports/stationery shop
7	living service
8	sports
9	hospital
10	hotel
11	scenic spot store
12	residence
13	governmental agencies and public organizations
14	science and education
15	transportation facilities
16	banking and insurance
17	corporate business
18	street furniture
19	entrance/bridge
20	public utilities
21	Chinese restaurant
22	foreign restaurant
23	fast food restaurant
24	shopping mall
25	convenience store
26	electronic products
27	supermarket
28	furniture building materials market

3.3. Cluster Implementation For Region Identification

This step combines similar formal regions in terms of region topic distributions with a help of clustering algorithm. Regions from the like cluster have like functions, and unlike clusters re present unlike functions.

The silhouette value of a point i in the dataset, denoted by s(i) is in the range of [−1, 1], where s(i) close to 1 means that the point is appropriately clustered and very distant from its neighbouring clusters; s(i) close to 0 indicates that the point is not distinctly in one cluster or another; s(i) close to -1 means the point is probably assigned to the wrong cluster. The average silhouette value of a cluster measures how tightly the data in this cluster are grouped, and the mean silhouette of the entire dataset reflects how appropriately all the data has been clustered.

3.4 Interpretation of Region

In this step each cluster of regions are annotate with some semantic terms, which reflects its real functions. A functional region is annotated by considering the following 4 aspects:

1) The POI configuration in a functional region. An average POI feature is computed by taking the regions a cross in functional region. According to the calculated POI feature vector, rank each POI category in a functional region (termed as internal ranking) and rank each POI category all the functional regions.

2) The most dominant mobility patterns.

3) The functionality intensity. the representative POIs located in each functionality kernel is studied, e.g., a function region will an educational area if it is replete of universities and schools.

4) The human-labelled regions. People may know the functions of a few well-known regions, e.g., the region contains the For bidden City is an area of historic interests. After clustering, the human labelled regions will help us understand other regions in a cluster.

3.5 Results & discussion

Fig2. Shows the combined functional regions disclosed by different methods, where different colours indicate different functions.

Diplomatic and embassy areas[c0]. The feature POI category in this functional region is the governmental agencies and public organizations, with a substantial higher density than other functional regions.

M.Sobana *et al*, International Journal of Computer Science and Mobile Applications,
Vol.3 Issue. 8, August- 2015, pg. 80-86 ISSN: 2321-8363

Fig. 2. Segmented regions after connected component labelling.

Embassies are located mostly in these areas, which are well configured for the diplomatic function, e.g., they have the highest external rank of Pub/Bars and transportation facilities, and the second highest rank of residential buildings, hospitals, and hotels, among all the clusters.

Education and science areas [c2]. This cluster region contains the utmost number of science and education POIs (e.g., T sinu a university and Beijing university).

Developed residential area s[c6]. This cluster region is clearly a mature residential area with the majority of residential building, living services, hospitals, hotels. In these areas, there are enough number of services supports' the people's living, such as the restaurants , shopping malls, banking services, schools, sports centres.

Emerging residential area s[c8]. This area is glossed as the emerging residential area since it has a balanced POI configuration, such as living services, residential buildings, sports centres, hospitals and some companies etc.

Developed commercial / entertainment areas[c5]. They are distinctive entertainment areas with the highest external rank of theatres, foreign restaurants and café/tea bars. Moreover, there are a great many shopping malls, Chinese restaurant and convenience stores. Functionality intensity and frequent mobility patterns derived for each functional region in addition to the POI configurations.

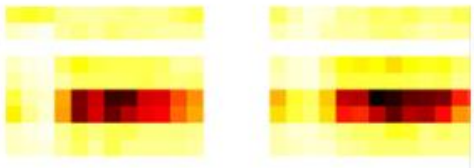
Developing commercial/business/ Entertainment areas [c1].
The POI contour of this cluster is similar to cluster c5 and c7, but in conditions of the absolute quantity, c1 is less than c5 while more thanc7. A sealed number of shopping malls, restaurants and banking services lineament this cluster as a developing commercial/ business/ Entertainment functional region (either of them is possible).

Regions under construction
This region will potentially become regions 1 or 8 since the POI configuration brings out a rudiment of the commercial/residential area with related supporting services. So it will belongs to c1 or c7.

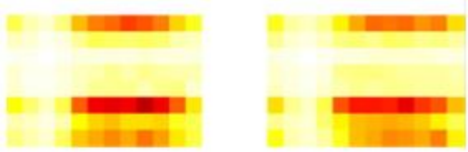
(a) functional region c1 (b) functional region c4

Fig 3: Functionality intensity of functional regions

By considering the functionality intensity estimated by the mobility patterns, it is easy to find that they are places of historic interests in Beijing. This has been shown in Figure 3(b), the famous historical places like Forbidden City and Temple of Heaven are located in these areas.[c2]



a) arriving, c6 b) leaving,c6



a) arriving, c8 b) leaving,c8

Fig 4. Arriving & Leaving of c6 & c8

Fig 4 shows the comparisons of arriving/leaving transitions of (region clusters) c6 and c8 with that of other clusters respectively, where the x-axes are time of day (by hour) and y-axes are the functional regions that come from (left subfigures) and leave for (right subfigures).Both c6 and c8 adopt the trend that most leaving transition sin the morning (8-9am, go to work) and most arriving transitions in the early evening (5-6pm, go back home), which is a typical pattern for the residential area. Nevertheless, in terms of the absolute quantity,c8 is much lower than c6, which shows that there are more people living in c6.

From the view of mobility graph & POI, region cluster c0, c2, c5, c6, c8 are matured and developed areas to greater extend as compared to other clusters .Compare to other previous methodology, HPC methodology outperforms in region identification and labelling.

4. Conclusion

Framework for discovering functional zones (e.g., educational areas, entertainment areas, and regions of historic interests) in a city using human trajectories, which imply socio-economic activities performed by citizens at different times and in various places. It have been have evaluated that, this framework with large-scale datasets including POIs, road networks, taxi trajectories and public transit data. It is founded that public transit data can be used as an accompaniment to the taxi trips in representing urban mobility, so as to achieve a better performance for discovering functional zones. The results yielded by this framework benefits' an assortment of applications including urban planning, location choosing for a business, and social recommendations.

References

[1] J. Antikainen. The concept of functional urban area. Findings of the Espon project, 1(1), 2005.

[2] S. Bednarz et al. Geography for Life: National Geography Standards. 1994.

[3] D. Blei. Introduction to probabilistic topic models. Communications of the ACM, 2011.

[4] D. Blei, A. Ng, and M. Jordan. Latent dirichlet allocation. TheJournal of Machine Learning Research, 3:993–1022, 2003.

[5] R. Estkowski. No Steiner point subdivision simplification isNP-Complete. In Proc. 10th Canadian Conf. Computational Geometry. Citeseer, 1998.

[6] Y. Ge, C. Liu, H. Xiong, and J. Chen. A taxi business intelligencesystem. In Proc. KDD '11, pages 735–738, 2011.

[7] Y. Ge, H. Xiong, A. Tuzhilin, K. Xiao, M. Gruteser, and M. Pazzani. An energy-efficient mobile recommender system. In Proc. KDD '10, pages 899–908, 2010.

[8] C. Karlsson. Clusters, functional regions and cluster policies. JIBS and CESIS Electronic Working Paper Series (84), 2007.

[9] L. Lam, S. Lee, and C. Suen. Thinning methodologies-acomprehensive survey. IEEE Transactions on pattern analysis and machine intelligence, 14(9):869–885, 1992.

[10] W. Liu, Y. Zheng, S. Chawla, J. Yuan, and X. Xing. Discovering spatio-temporal causal interactions in traffic data streams. In Proc.KDD '11, pages 1010–1018, 2011.

[11] D. Mimno and A. McCallum. Topic models conditioned on arbitraryfeatures with dirichlet-multinomial regression. In Uncertainty inArtificial Intelligence, pages 411–418, 2008.

[12] A. Pozdnoukhov and C. Kaiser. Space-time dynamics of topics in streaming text. In Proc. LBSN '11, pages 8:1–8:8, 2011.

[13] G. Qi, X. Li, S. Li, G. Pan, Z. Wang, and D. Zhang. Measuring social functions of city regions from large-scale taxi behaviors. In IEEE PERCOM Workshops, pages 384–388, 2011.

[14] P. Rousseeuw. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. Journal of computational and appliedmathematics, 20:53–65, 1987.

[15] L. Shapiro and G. Stockman. Computer Vision. Prentice Hall, 2001.

[16] R. R. Vatsavai, E. Bright, C. Varun, B. Budhendra, A. Cheriyadat, and J. Grasser. Machine learning approaches for high-resolution urbanland cover classification: a comparative study. In Proc COM.Geo '11,pages 11:1–11:10, 2011.

[17] M. Wand and M. Jones. Kernel smoothing, volume 60. Chapman & Hall/CRC, 1995.

[18] Z. Yin, L. Cao, J. Han, C. Zhai, and T. Huang. Geographical topic discovery and comparison. In Proc. WWW '11, pages 247–256, 2011.

[19] J. Yuan, Y. Zheng, X. Xie, and G. Sun. Driving with knowledge from the physical world. In Proc. KDD '11, pages 316–324, 2011.

[20] J. Yuan, Y. Zheng, L. Zhang, X. Xie, and G. Sun. Where to find my next passenger. In Proc. Ubicomp '11, pages 109–118, 2011.
[21] Y. Zheng, Y. Liu, J. Yuan, and X. Xie. Urban computing with taxicabs. In Proc. Ubicomp '11, pages 89–98, 2011.
[22] Y. Zheng and X. Zhou. Computing with spatial trajectories. Springer-Verlag New York Inc, 2011.