

Macro Analysis of Urban Structure Based on Point of Interest

Weili Huang

Information Engineering Institute, Jiangxi University of Technology, Nanchang 330098, China

huangweili@jxut.edu.cn

Keywords: Point of Interest; Macro analysis; Urban structure; Regional GDP

Abstract. Point of interest (POI) is the geographic entity to implement urban function. With the development of the city, the number of POI is exploding. We analyzed the POI's macro characteristic of 31 major cities in China. The results show that POI is strongly associated to urban macro economy. Many macroeconomic indicators are significantly positive correlated with the number of POI, especially regional GDP. Further comparative analysis on the proportions of different POI kinds reveals the development of urban structure is unbalanced and unreasonable.

Introduction

Researches on urban structure involve many fields such as population [1], land [2], economy [3], society [4], etc. Usually, the traditional data is achieved from questionnaire survey, population census, economic census or land-use data. With the development of information technique, more novel data sources are available, such as GSM [5], GPS [6], social networking services [7], Wi-Fi hotspot, smart card [8], and nighttime satellite image [9]. Model driven analysis methods have been turned into data driven [10].

Point of interest (Point of Interest, POI) is the geographic entity, such as shopping mall, railway station and school. POI is an important part of navigation electronic map. It is possible to mine the massive POI data for the function, structure and form of urban space. POI data has rich information and wide coverage. These advantages make POI more popular. In recent years, many POI-based researches on urban structure come out. In this paper, we analyzed macro urban structure based on POI. The relationship between POI and macroeconomic indicators is discussed. Furthermore, the proportions of different POI kinds are used to estimate whether the urban structure is healthy.

Data

Comprehensive and real-time map data has been available from many famous map providers such as Google, Baidu, Open Street Map, AMAP, etc. Powerful APIs and map services are provided for free in these platforms. In addition to geographical position, POI data has much extra information. Baidu POI includes address, telephone number, label, price, business hours, scores in taste, service, environment, health and technical, picture, comment, discount, etc. AMAP POI data is divided into basic, detailed and depth information. The depth information provides the descriptions of restaurants, hotels, cinemas and scenic spots.

Our map data is the China's electronic map data from MapInfo in 2012. For keeping consistency, the other relevant statistical data is also in 2012. The original map data has no POI layer. It is subdivided into government, park, building, hotel, health care, school, bank, pharmacy, restaurant, bus station, entrance, toll station, park, supermarket etc. Therefore, it needs to be aggregated to form a complete view of the POI data. After aggregation, we got about 5 million 200 thousand POI data. Moreover, the dynamic and timeliness of POI data need to be paid enough attention. Failure or duplicate data is inevitable in mass data. Data preprocessing is necessary. The failure data and duplicate data have a common characteristic, that is, there are more than one POI data in the same position. We mainly focus on the geographical position with multi-POI. By filtering kinds and names, one is reserved. For example, the kind "station" always contains both "bus station" with code 808B and "bus station entrance" with code 808F in the same geographical coordinate. This is a typical duplication data, only one can be reserved.

Macro Analysis of Urban Structure

POI Regional Density. POI regional density is defined as a ratio of the number of POIs in the region and regional area. Assuming the area of A is S_A , N_A is the number of POI, let ρ_A be POI region density.

$$\rho_A = \frac{N_A}{S_A} \quad (1)$$

POI density of a city can be defined as following.

$$\rho_{City} = \frac{N_{City}}{S_{City}} \quad (2)$$

where N_{City} is the number of POI in the urban area, S_{City} is the area of the city. ρ_{City} is mainly used for urban analysis at the macro scale.

Correlations between POI and Macroeconomic Indicators. As the urban cells, POI is everywhere. It must be associated with urban macroeconomic indicators. The major urban macroeconomic indicators include total population(year-end), regional GDP, passenger traffic, freight traffic, budgetary revenue and expenditure of local governments, total investment in fixed assets, total retail sales of consumer goods and total value of import and export etc. The regional GDP is a comprehensive indicator, which directly reflects the degree of urban economic development.

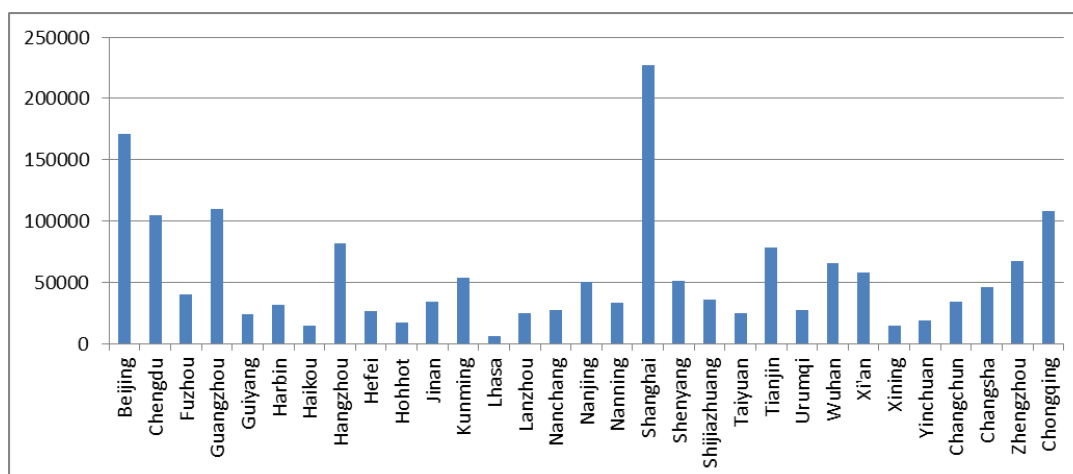


Figure 1. The number of POI for China's major urban areas

POI represents urban vitality. Fig.1 shows the number of POI for China's major urban areas. Correlation analysis has been done between the number of POI (N_{City}), the construction land area of municipal district (S_{City}), POI regional density (ρ_{City}), population, population density and regional GDP. The correlation matrix for these factors is given in Table 1.

Table 1 The Correlation matrix for POI and major factors

	S_{City}	N_{City}	ρ_{City}	Population	Population Density	Regional GDP
S_{City}	1.000	.913	-.834	.494	.590	.866
N_{City}		1.000	.842	.615	.636	.946
ρ_{City}			1.000	.259	.833	.793
Population				1.000	.333	.633
Population Density					1.000	.646
Regional GDP						1.000

It can be found that:

- (1) There was a significant positive correlation between the number of POI N_{City} and the construction land area of municipal district S_{City} with correlation coefficient 0.913;
- (2) There was a positive correlation between the number of POI N_{City} and POI regional density ρ_{City} with correlation coefficient 0.842;
- (3) The positive correlation between POI regional density and population density is obvious;
- (4) It was noticeable that there is a significant positive correlation between the number of POI and regional GDP with the maximum correlation coefficient 0.946, Fig. 2 shows the detail.

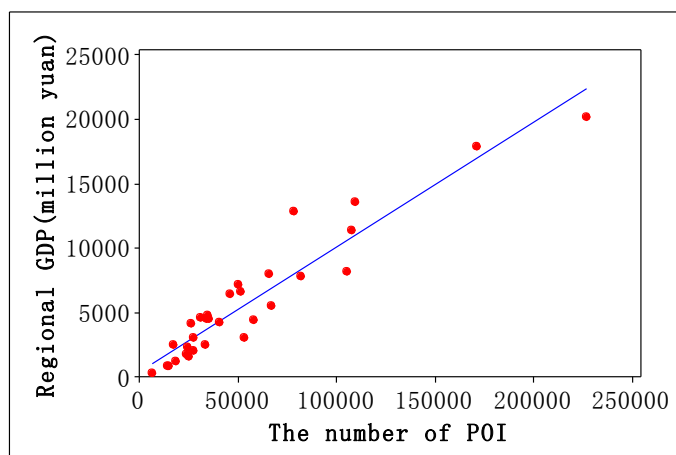


Figure 2. Correlation diagram of the number of POI and regional GDP

On the whole, there were higher positive correlations between the number of POI N_{City} , POI regional density ρ_{City} and regional GDP. So both N_{City} and ρ_{City} can be used as the effective indicators to characterize the macro economy. For N_{City} is more relevant to regional GDP, it reflects the degree of the urban economic development more accurately. The development of urban function is independent to the area of the city. The larger S_{City} is not always with the larger N_{City} , but usually the larger N_{City} might correspond to the larger ρ_{City} and the larger population density might correspond to the larger ρ_{City} .

POI Structure Analysis. The number of POI in urban area has a significant positive correlation with regional GDP, which indicates that POI can be regarded as a measure of urban general function. Different functional areas make up urban structure. Each kind of POI corresponds to a specific urban function. The urban macro structure could be observed from analyzing the proportion of different kinds of POI.

We list the results for some representative kinds of POI in Table 2. The correlation coefficients between the proportion of major kinds of POI and regional GDP are given. All of them with average proportions over 1% accounted for 79% of the total number of POI. In order to analyze the relationship between urban functions and macroeconomic indicators, we calculated the proportion of each kind of POI in each city. We found the kind of company has the highest correlation coefficient (0.707) with regional GDP, followed by the station, park, toilet, supermarket etc. There is undoubted that company is the most important kind for the development of urban economy. On the other hand, the city with higher regional GDP will have higher fiscal expenditure for urban development. So the proportions of facilities as station, park and toilet also correlate with the regional GDP positively. Another interesting result is that the kinds of education and health care showed obvious negative correlation with regional GDP. In the city with higher regional GDP, the proportion of education and medical is lower. The difficulty of getting medical treatment and receiving an education become more common in the big city. It also represents the development of urban function structure is unbalanced and unreasonable currently.

Table 2 The Correlations for the percentage of major kinds of POI and regional GDP

Kind	Proportion	Correlations	Kind	Proportion	Correlations
Company	9.9%	0.707	Building Material	1.6%	-0.271
Station	5.0%	0.532	Bank	5.4%	-0.278
Park	1.2%	0.524	Restaurant	8.1%	-0.358
Toilet	1.3%	0.364	Non-Staple Food	2.1%	-0.362
Supermarket	1.9%	0.342	Hotel	2.2%	-0.380
Bath	1.1%	0.241	Education	4.9%	-0.430
Fast Food	1.2%	0.192	Cybercafe	1.4%	-0.456
Barber	2.7%	0.178	Hospital	1.3%	-0.476
Neighborhood	4.2%	0.178	Telecom	1.4%	-0.524
Pharmacy	3.4%	-0.096	Auto Maintenance	1.6%	-0.539
Cafe	1.1%	-0.104	Government	5.5%	-0.553
Convenience Store	2.2%	-0.190	Auto Service	1.8%	-0.588
Clothing And Shoes	4.5%	-0.262	Community Medical	2.1%	-0.593

Conclusions

The POI data is used to study the urban structure in this paper. First, we defined the POI regional density. Then an analysis of POI and the macroscopic indicators is given. The study of 31 China's major urban areas shows there is a significant positive correlation between the number of POI and regional GDP. POI can be used to characterize urban macro economy. Further proportion analysis of different POI kinds reveals the main frame of urban structure. Company is the core power of urban development. On the other hand, the drawback of urban structure is observed. Due to the development of current urban function structure is unbalanced and unreasonable; the medical and education service especially in the big city is insufficient.

Acknowledgements

This work is supported by Science and Technology Project of Jiangxi University of Technology (Grant No. xtcx201319), Science and Technology Project of Education Department of Jiangxi Province ("Research on urban heterogeneous community structure based on Co-location model").

References

- [1] F. Wang, Y. Zhou, Modeling urban population densities in Beijing 1982-90: Suburbanization and its causes, *Urban Studies* 36 (1999) 271-287.
- [2] M. Batty, P. A. Longley, The morphology of urban land use, *Environment and Planning B: Planning and Design* 15 (1998) 461-488.
- [3] X. Zhang, E. S. Zhong, X. J. Zhang and S. Wang, Spatial distribution and clustering of commercial network in Beijing during 2004-2008, *Progress in Geography* 32 (2013) 1207-1215.
- [4] B. W. Jones, B. Spigel and E. J. Malecki, Blog links as pipelines to buzz elsewhere: the case of New York theater blogs, *Environment and Planning B: Planning and Design* 37 (2010) 99-111.
- [5] F. Calabrese, C. Ratti, Real time Rome, *Networks and Communication Studies* 20 (2006) 247-258.
- [6] B. Jiang, J. Yin and S. Zhao, Characterizing Human Mobility Patterns in a Large Street Network, *Physical Review E* 80 (2009) 1711-1715.
- [7] B. Jiang, Head/tail breaks for visualization of city structure and dynamics, *Cities* 43 (2015) 69-77.
- [8] Y. Long, Y. Zhang and C. Y. Cui, Identifying Commuting Pattern of Beijing Using Bus Smart Card Data, *Acta Geographica Sinica* 67 (2012) 1339-1352.

- [9] J. Bennie, T. W. Davies, J. P. Duffy, R. Inger and K. J. Gaston, Contrasting trends in light pollution across Europe based on satellite observed night time lights, *Scientific Reports* 4 (2014).
- [10] Q. Q. Li, D. R. Li, *Big data GIS*, Geomatics and Information Science of Wuhan University, 39 (2014) 641-644.