Comparing the use of actual space and virtual space: A case study on Beijing’s Wangfujing area

Qiang Sheng
School of Architecture and Design, Beijing Jiao Tong University, China
shengqiang1234@hotmail.com

Ning Liu
Space Syntax Ltd. Beijing Office
ning.liu@spacesyntax.com

Abstract

Recent development of user assessment website such as Dazhongdianping (a very popular website for choosing and reviewing restaurants in China) has great impact on the way people use urban functions. This paper analyzes the relationship between the configuration of actual space, number of customers observed in fieldwork and the visibility of shops on website in Beijing Wangfujing area. The result suggests that the use of informational technology is strengthening the spatial distribution logic of these restaurants instead of weakening the role of actual space. By comparing the customer number of shops inside one of the case shopping mall with the number of reviews posted on web, this study also explores the potential and limitation of web data being used in spatial analysis.

Keywords

Big data, shopping mall, Wangfujing, Dazhongdianping.

1. Introduction

Recent development in the informational technologies has great impact on people’s everyday life in big cities. From the 1990s, one popular issue in the field of urbanism and architecture is how informational technologies can affect the locational logic of regional economic structure and customer behavior in local places. With the rise of ‘global market’ and further intensification of the economical networks supported by these new technologies, some scholars predicted that the proximity of places will eventually become irrelevant (Manuel, 1996).

However, many other scholars are against this pessimistic viewpoint. At the regional or global scale, they believe that in the information age, the centrifugal and centripetal force co-existed as two sides of a coin (Liu, 2002; Gillespie et al., 2000). On the one hand, the advance of informational technologies lead to dispersion of manufactory activities towards marginal places. On the other hand, management and advanced services are becoming more and more concentrated in few global cities (Sassen, 2001).

At the urban scale, the booming online shopping has some influences on the way people choosing, purchasing and reviewing the commodities in last decade. For some experience-based commercial functions such as restaurants, the choosing and reviewing on line also provide new possibilities for
both customers and restaurants. On some of these very popular websites, such as ‘Dazhongdianping’ in China, the user can easily compare and select the restaurants based on several filters such as the location, cost, popularities and ranks. These data are summarized based on the reviews posted by previous customers. Together with the use of GPS system in cars and smart phones, it seems the location of restaurant should not matter any longer, at least not as much as before.

On the other hand, some scholars hold alternative viewpoints. Read pointed out that when the informational accessibility is no longer a problem, the role played by actual spatial accessibility will rather be enhanced (Read, 2009). In this line of thinking, in the age of information the value of location could be even more important than before.

By studying on the data of these restaurants, especially on the correlation between these data with their geo-locations, this paper will start the analysis on three scales: First, at metropolitan scale, over 1000 restaurants on Dazhongdianping will be sampled and analyzed to examine if their distribution can reveal certain spatial logic. Second, at urban block scale, all restaurants in wangfujing/dongsi area will be surveyed and mapped. The pedestrian movement on 73 street segment will also be surveyed in fieldwork. The data on Dazhongdianping website in the same area will also be visualized in map. Then we will use depthmap to calculate the correlations between . For the space syntax research, it can reduce the field work and help to choosing the appropriate spatial parameters.

2. Data and research method

2.1 About Dazhongdianping

Dazhongdianping Company is founded in 2003 in Shanghai. Its website (www.dazhongdianping.com) is one of the first Chinese web forum and service platform which provides information based on users’ feedback on the quality of the restaurants. In the last 10 years Dazhongdianping developed into a leading website reviewing the catering service. To a large extent it also changed the way Chinese people use catering functions in big cities. A user can make his choice based on many characters such as the style of flavor, quality of environment and service. It also shows the reviews from other customers and the average rank of certain restaurant based on their feedbacks.

The data collection of this research started from May 2013. The research area covers 5 main administrative districts, dongcheng, xicheng, chaoyang, haidian and fengtai. Within each district, there are 10 to 20 smaller sub-area listed on the website. Figure 2 shows the distribution density of restaurant within each of this sub-area listed on the website.

For each sub-area, we select top 15 restaurants in the default ranking filter. After delete some restaurants which are listed by multiple sub-areas, there are in total 1123 restaurants selected as samples. In the default setting of the website, each page can show up to 15 restaurants in one page, practically the restaurant listed in the first page are the most visible ones. For each restaurant, this research records three types of data, the average cost per meal per person (in RMB), the number of reviews from previous customers and the average rank based on their reviews (0-5 stars). Among these data, the number of reviews could be understood as comparable with the number of visits in a comparative study between these restaurants. As mentioned before, although food delivery is an emerging new service provided by many modern restaurants, face to face service is still a dominant sector in their business. The average cost per meal per person can reveal not only the economic status, but also the scale of customer of each restaurant. Normally, expansive restaurants are visited by customers from larger region. Therefore, these two sets of data are related with the actual movement in urban space. All these data are illustrated in a map in Figure 2. The regression analysis shows very low correlation between these three sets of data (R-square value lower than 0.078). This research will focus on if the spatial location of these sampled restaurants reveals certain underlying logic.
**Figure 1**: Number of restaurants and density in each sub-area listed in Dazhongdianping.

**Figure 2**: Number of reviews, average cost per meal per person and ranks of the sampled restaurants.
2.2 Integration and Nach value in space syntax

The area of space syntax analysis covers the whole metropolitan region including the satellite towns such as Changing, Shunyi, Tongzhou and Mentougou. The analysis is mainly focusing on two types of parameters in the angular analysis mode: integration and Nach value. The Nach value (normalized angular choice) used in this research is a new development of space syntax in 2012 [Hillier et al., 2012].

Figure 3: Integration and Nach value of Beijing in radius 1km and 10km.

Figure 3 shows integration and Nach value in 10km and 1km radius. Integration can reveals different scales of central areas. Nach can show different scales of movement networks. In terms of shopping behavior, integration and Nach value could be understood as revealing two different kinds of spatial potentials: integrations shows the potential of a place (or a shop inside a shopping mall) being visited as a target of the trip (To-movement potential); Nach shows the potential of a place being passing-by (Through-movement potential). Using these two parameters, this research will explore how they can influence the use of catering functions in different scales.

3. Metropolitan-scale analysis based on Dazhongdianping data

This research starts with analyzing the relationship of 3 sets of data (reviewing number, average cost and rank) with the spatial parameters. Table 1 lists their locations’ integration and Nach value in different radius. The darker the background color means the higher dependency of certain group of data on certain type of spatial parameters. For instance, in the list of reviewing number, the 77 restaurants with top 20% number of reviews tend to be influenced by 50km radius choice value most. Their location’s choice value in 50km radius is above the average rate by 45%.
Table 1: The relationship between reviewing number, average cost and rank with spatial parameters.

<table>
<thead>
<tr>
<th>Review</th>
<th>NodeCount</th>
<th>NodePerm</th>
<th>NodePermStd</th>
<th>MCHain</th>
<th>MCHainStd</th>
<th>MCHainPer</th>
<th>MCHainPerStd</th>
<th>SpatialCost</th>
<th>SpatialCostPer</th>
<th>SpatialCostPerStd</th>
</tr>
</thead>
<tbody>
<tr>
<td>top20</td>
<td>77</td>
<td>1.294</td>
<td>1.2173</td>
<td>22963</td>
<td>1.2773</td>
<td>1.2136</td>
<td>1.2047</td>
<td>124.23</td>
<td>124.23</td>
<td>124.23</td>
</tr>
<tr>
<td>top20-40</td>
<td>239</td>
<td>1.242</td>
<td>1.2173</td>
<td>22963</td>
<td>1.2773</td>
<td>1.2136</td>
<td>1.2047</td>
<td>124.23</td>
<td>124.23</td>
<td>124.23</td>
</tr>
<tr>
<td>top40-60</td>
<td>316</td>
<td>1.225</td>
<td>1.2173</td>
<td>22963</td>
<td>1.2773</td>
<td>1.2136</td>
<td>1.2047</td>
<td>124.23</td>
<td>124.23</td>
<td>124.23</td>
</tr>
<tr>
<td>top60-80</td>
<td>147</td>
<td>1.225</td>
<td>1.2173</td>
<td>22963</td>
<td>1.2773</td>
<td>1.2136</td>
<td>1.2047</td>
<td>124.23</td>
<td>124.23</td>
<td>124.23</td>
</tr>
<tr>
<td>bottom20</td>
<td>53</td>
<td>1.225</td>
<td>1.2173</td>
<td>22963</td>
<td>1.2773</td>
<td>1.2136</td>
<td>1.2047</td>
<td>124.23</td>
<td>124.23</td>
<td>124.23</td>
</tr>
</tbody>
</table>

As the table shows, the reviewing number is influenced by spatial parameters than rank and average cost. The top 20% reviewed restaurants tends to located on the street segments with high 50km radius Nach value and high 2km radius integration value. In fact, the table shows that other levels of reviewing numbers reveal the similar tendency. However, there are also some differences in the analysis on average cost and rank of those restaurant. On the average cost, table 1 shows cheap restaurants depends more on the spatial condition than expensive ones. We speculate that cheap restaurants offer standard food, so if they have higher number of reviews, it is very possible that they located in very accessible spaces. Therefore their visibility in virtual space might depends more on the possibility of being visited by chance in actual space. On the ranks of the restaurants, the result is even more interesting. Principally the rank of the restaurant should only depend on the customer's subjective feelings on the quality of food, but the result still shows a clear spatial logic which at least stronger than that in the analysis of average cost. In fact, the results could be seen as a combination of the analysis on the number of reviews and average cost. Very highly ranked (5 star) restaurants do not have strongest dependence on spatial condition, because they are mostly quite expensive. Rank 4.5-4.9 restaurants are dependent on the spatial conditions most. Rank 3.5-3.9 restaurants least. While very low ranked restaurants tend to depend more on spatial condition probably because they are mostly very cheap. There is a Chinese poem saying that: If the wine is good, it does not matter where the bar is located. This result shows a different reality even in the information age. Although very expensive restaurant depends on spatial condition less. But majority of restaurant still depends on the location. Good location can bring in more customers. More customers give higher profit which could be used to hire better chef, to serve better food, then to get more reviews on web, so on and so forth.

Based on the previous finding which indicates the spatial condition of 2km integration and large scale Nach (r=50km), this part of research focuses on the total number of restaurants listed in each small district in Dazhongdianping and the spatial condition of that particular area.
Using the tool of regression analysis with two factors (2km integration and 50km Nach), the R-square value of distribution density of the restaurants on Dazhongdianping and integrated spatial parameter can reach 0.51 (see Figure 4). The weight between Nach and 2km integration is 4.613:1.

So far these findings can lead to two preliminary conclusions: First, it indicates that in the information age the spatial condition still plays an important role in the distribution of catering functions. Second, this spatial condition could be understood as an integration of two main aspects: how local area is accessible in the whole metropolis (50km Nach) and how it is centralized in local street fabric itself (2km integration). Furthermore, the connection in city as a whole is more important than its local connections. This finding also could be understood as related with the way people are using these restaurants through the help of information and transportation technologies. The restaurants which is well accessible in the whole metropolis are easier to reach by vehicles through highway and motor way system. But when we reach the local area and get off the vehicles, it is more possible we could find them in a street which is well centralized in pedestrian scale. Furthermore, restaurants with this spatial advantages have better possibility to be visited and reviewed, so they can also benefit more from the information technologies. To put this process in a simple way, the information technologies are not functioning independently as separate technical networks. It cannot create and sustain a virtual world. They have to be imbedded with other technologies that is already there. They are intensifying the spatial conditions rather cancel it out.

Can these preliminary findings be tested in smaller scale? Is there a relationship between the number of customers and the reviewing numbers on web? These questions will be explored in a detail case study in Wangfujing area in Beijing.
4. Urban Block scale and architecture scale analysis: case study in Wangfujing area and the Malls at New Oriental Plaza

4.1 data mapping and fieldwork

Wangfujing/dongsi area on Dazhongdianping has highest number of restaurants. Previous study shows that the number of reviews has strong correlation with the spatial accessibility of two scales. In this part of the research will focus on Wangfujing area by visualizing the data from Dazhongdianping and mapping the actual location, the pedestrian flows in this area through survey. Figure 5 shows the visualization of three types of data in Wangfujing area: the number of reviews, the average cost per person per meal and the rank given by the users. There are 423 restaurants listed in Dazhongdianping in this area. Among these 423 restaurants, 30 of them have no data on the average cost or reviews, 61 of them have no average cost data. So in total 21.5% of them do not have sufficient data for detail analysis. These restaurants are mostly small restaurants located on small alleys.

Figure 5: Visualization of restaurant data in wangfujing area.

After comparing the number of restaurants mapped in field trip (fig.6) with the number of restaurants listed in Dazhongdianping on the street (excluding those inside shopping malls or hotels), the result shows that only 54 restaurants (11.3%) are not listed on that website. They are also mostly located in small alleys. In fact, Dazhongdianping’s survey has covered almost everywhere in Beijing’s
metropolitan area, including those small alleys. The data can give a whole picture regarding where most restaurants are located and how they are functioning. But some small restaurants also change quite often so their information are not updated in time or they have few reviews.

As mentioned before, this paper intends to compare the uses of those restaurants in actual and virtual space. A field work is organized to collect detail data regarding the pedestrian flow intensity in Wangfujing area and inside some case shopping malls. The number of customer in those individual restaurants inside case shopping mall is also mapped to compare with the number of reviews on web. Wangfujing area and shopping malls in this area have very detailed information on Dazhongdianping, even each individual shop inside shopping mall also have reviewing data. This can give us an ideal empirical data base to study the relationship between virtual and actual use of urban space.

A fieldwork is organized to compare the web data with the actual use of space on the streets and inside these shopping malls. 73 street segments are selected for gate count study. Pedestrian flows are measured four times in a weekday and a weekend day. Furthermore, there are 346 corridors inside the three shopping malls are selected for gate count study using similar method with the study on street segments. The number of customers inside 893 shops or restaurants inside those three case shopping malls were noted down three times a day (morning, lunch time and afternoon).

Figure 7 visualizes the flow intensity on the streets and at the entrances of three shopping malls, it also shows the number of customers inside shops in each floor in weekday and weekend. The flow at the entrances shows the Malls at New Oriental Plaza has highest number of people coming in and out of the building, Beijing Department Store has lowest number. Comparing the number in weekday and weekend, the New Oriental Plaza tends to run a very stable business while the New Dong’an APM tends to benefit most from the flow in weekend.

The following part of the research will usedepthmap to analyze the spatial parameters of the Wangfujing area and the case shopping malls. It will compare the actual spatial condition and usage with the virtual visibility.
4.2 Spatial analysis on Wangfujing area and the New Oriental Plaza

At the urban block scale, a detail pedestrian network including the ground level of major shopping malls are mapped. Figure 7 shows the observed pedestrian flow intensity with the space syntax analysis result of 3km integration. The R-square value can reach up to 0.612. Furthermore, the average 3km integration and choice value of three major shopping malls are also listed.

![Figure 7: The flow intensity in Wangfujing area and number of customers in three shopping malls (weekday and weekend day).](image)

![Figure 8: Pedestrian flow intensity in Wangfujing area and the spatial analysis of 3km integration value.](image)
The result shows the New Oriental Plaza (Case C) has the highest Nach value, New Dong An APM (Case B) has the highest integration value, the Beijing Department Store ranks the last in both values. Comparing this result with the data presented in Figure 7, Case C has the highest number of people entering and going out of the building, while case B has the highest number of customers observed inside the shops per square meter. Case A is the most unsuccessful case. So this ranking fits with the spatial analysis quite well. The New Oriental Plaza (case C) benefits most from the through-movement, while New Dong An APM (case B) benefits most from the to-movement.

### Table 1

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>r=500m</td>
<td>0.414</td>
<td>0.001</td>
<td>0.022</td>
<td>0.388</td>
<td>0.038</td>
<td>0.021</td>
<td></td>
</tr>
<tr>
<td>r=1km</td>
<td>0.543</td>
<td>0.003</td>
<td>0.000</td>
<td>r=1km</td>
<td>0.446</td>
<td>0.0001</td>
<td>0.004</td>
</tr>
<tr>
<td>r=3km</td>
<td>0.641</td>
<td>0.0002</td>
<td>0.000</td>
<td>r=3km</td>
<td>0.46</td>
<td>0.001</td>
<td>0.004</td>
</tr>
</tbody>
</table>

*Figure 9: The flow intensity and 3d spatial analysis of the Malls at New Oriental Plaza.*

Three sets of data (number of reviews, average cost and rank) are input into the segment map. The scatter plot shows the number of review has clear spatial logic and the integration value r=3km has the top R-square value 0.641. The analysis on the distribution of pedestrian flow shows the same result. This finding indicates that at the urban block scale, the number of reviews on Dazhongdianping web could be considered as related positively with the actual number of visits. On the other hand, the average cost per person per meal turns out to be not related with spatial analysis. As we can see from the map in Figure 9, the most expansive restaurants locates on the street segment deep inside the fabric. This phenomena confirmed our previous analysis at metropolitan scale.

What about those good restaurant then? The left map in Figure 9 shows the restaurants ranked 4.5 and above are mostly located on other main streets except Wangfujing. For decades Wangfujing has developed into a very busy pedestrian street mainly for tourism and the street front are mostly occupied by souvenir shops and other retails. Good restaurants do not necessarily join this competition of high valued land, but they still need rather well-connected spaces.

As building scale, a detail study on the case of New Oriental Plaza (case C) starts with the analysis of observed flow on the corridors (see Figure 10). Using the ‘unlink’ tool in space syntax, a 3D model of the mall is established. 1km radius Nach value is strongly correlated with the observed flow inside the building (R-square=0.623).
When focusing on the catering functions inside the building, most of them are located on the least visible and least accessible part of corridors (Macdonald’s is the only exception). They locate on the underground floor, at the east and west end of the building and a sideway in the south part which is parallel with the main corridor. This distribution is both a design strategy and a natural logic for the catering function because they are only visited in two times of a day. Similarly, at urban block scale there are also very few restaurants located on the Wangfujing street and open its gate towards main street directly.

Figure 10: The flow intensity and 3D spatial analysis of the Malls at New Oriental Plaza.

Figure 11: The comparison of observed number of customers and reviewing numbers on web in the restaurants in the New Oriental Plaza.
Furthermore, when comparing the actual use of these restaurants with the reviewing numbers, there is almost no correlation (R-square=0.03). There are two possible reasons: First, this might due to the way we do this survey: only noting down the number of customers three times a day will underestimate the fact that smaller yet popular restaurants may have even more customers in the whole day instead of have high number of customers right on the moment we are counting them. Second, some restaurants or cafes like Macdonald’s and Starbucks have high number of customers in most time of a day, but they have relatively very low number of reviews on web. Obviously the standard food and beverage they serve make them less interesting to be reviewed. Therefore, unlike the finding at urban block scale, on this stage the number of review data still cannot be used to substitute the actual number of customers observed through fieldwork. Further research still needed to explore the relationship between actual visits and number of reviews by considering other types of data.

5. Conclusions and discussions: The impact of information technology on experience-based commercial functions

It seems that any kind of claims underestimated the impact of the information technologies on our contemporary urban life is so powerless. Indeed, without Dazhongdianping, it is very hard to collect the similar set of data in this research. Especially at urban scale it will be a mission impossible. However, it is not fair to claim that the role of actual space will be irrelevant. This research suggests catering function still depends on the physical experience in actual space. For this type of experience-based functions, the location still matters. A good location can bring in more customers, and tend to be reviewed more often in virtual space.

In detail, this research can give us some preliminary empirical findings: at urban scale most popular restaurants on web are located in the area which is well-connected in the metropolis as a whole and also well centralized in the local street fabrics. At urban block scale the well-connected shopping malls tend to have more popular restaurants, but at the architecture scale there is a clear difference between the observed number of customers and the reviewing number of each restaurants inside the mall. This result suggests that a target consumption based on web depends less on the spatial configuration inside the building, but more on the location of the building in its urban context.

The web-based big data is definitely a gift for spatial research in the information age. It allow us to see more things in such a big amount that could never be possible before. Instead of analyzing the locational logic of urban centers, now we could have the possibility to develop a more targeted approach to analyze the spatial condition for specific functions and social groups (the social elites go to expensive restaurants, for instance). We could also analyze the potential of a site without being there and counting the number of people on site, but to achieve that we also need more fundamental research to establish and double check those data-linkages. The findings of this paper could help us integrate the big data and the traditional way of survey. User assessment based, open sourced web data proved to be very promising. It already can have some good correlation with the spatial analysis in urban block scale and above. But in finer scale it still need to be double-checked. Further research will focus on including more data from Dazhongdianping and trans-scale analysis to analyze the customer behavior in shopping area and inside malls.

Acknowledgements

This paper is supported by the National Natural Science Foundation of China (51208343).
References


Gillespie, A., Richardson R, Cornford J. (2000), ‘Regional Development and the New Economy’. In Research paper of the Center for Urban and Regional Development Studies, University of Newcastle Upon Tyne, UK,

