RESEARCH ANALYSIS OF URBAN AND SOCIAL PATTERNS IN THE CITY

Shared bicycles and their influence on urban fabric

Araújo, Gonçalo[1]; Correia – Guedes, Manuel[1]; Ercument Gorgul[2]

[1]Instituto Superior Técnico, Universidade de Lisboa, Lisboa, Portugal
[2]CAUP, Tongji University, Shanghai, China

ABSTRACT: Shared bicycles have been around for a while and growing steadily in China. Recently, concept and volume of this new form of shared transportation vehicles captured a widespread attention and usage. This study is focused in two areas known as former French concession in Shanghai and in Xintiandi. Using one of the popular bike sharing app "Mobike" location and number of available bicycles is captured during a period of one week, three times a week, and every eight hours. Furthermore, this data is correlated with the existent urban framework by analysing certain aspects such as proximity to building services and daily life of locals. Allowing a quantitative and comparative evaluation with other sites regarding predictors of urban development, cyclers safety and urban quality. Then a proximity factor is introduced measuring the distance to key services, such as supermarkets, restaurants or office buildings, that impact life in the area. Finally, it will be possible to determine the comparative quality of these areas and take conclusions regarding future area studies and comparisons.
KEYWORDS: Transportation, Urban quality, Shared Bikes, Social Patterns, Urban Planning

1. INTRODUCTION

The main goal of this research is to map and analyze the dispersion of available shared bicycles in Xintiandi, Shanghai, and apply the results to a macroscale event, establishing correlations between the position of these bicycles, their availability, their location and proximity to different services and various infrastructures of the city. Throughout this study, one will try to understand and answer questions that are of crucial importance to urban development of a city and try to identify cores of economic and urban development. It is important to mention that this is a hypothesis-oriented study and will follow accordingly, establishing questions and trying to answer them through data gathering and analysis.

First, start with identifying and understanding what is a social phenomenon and the inherent questions that relate it to shared bicycle usage. Secondly, tracing what is called key services. In this study, key services are designation given to services and infrastructure typologies that are considered relevant for shared bicycle usage and play a major role in their dispersion. Following identification, explanation and calculation of what is called "proximity factor", which translates the relevance of each service and the role it plays in the positioning of *Mobike*, one of the most popular shared bicycle brands, further allowing to make assumptions over the relevance of these services and understanding the scope of urban development with new and innovative transportation models, which are recently booming all over China.

2. METHODOLOGY

The data gathering was done simply by using the *Mobike* app, through screening positions of available bikes in aforementioned area every morning, afternoon and night, over a week period. Screening of selected key services was made using an altered variation using map services of *Baidu* that allows to identify and locate all type of infrastructure typology of an area.

After these predefined settings (choosing key services and brand of bikes to study), the only thing left to do would be to map these services, and cross reference this data with the Mobike data. To do that, it would be necessary to scale both areas of the application map and the *Baidu* Map variation into same scale, and then overlap them [Fig.1,2].

By mapping with points in an xyz-coordinate system, both bicycles and existing services in the perimeter would allow us to have a common quantifiable vector system that may, or may not, soon prove accurate.

All this data, after being gathered, will need a significance test, to see if any error implied by this cross referencing would be of a large importance that would influence and turn the future results obsolete. Consequently, the most suitable test would be a comparison using heatmap between available *Mobike* density over time, and position of these defined key services [Fig.3,4]. This will lead to ideas about whether if this study is worth conducting and take primary conclusions over the importance of it.

Understandably, there are some quantifiable relations between these parameters in several images where the heat pattern of bike distribution relates similarly to the position of residences and office buildings in the area.

Figure 1 Mobike on one morning and office buildings

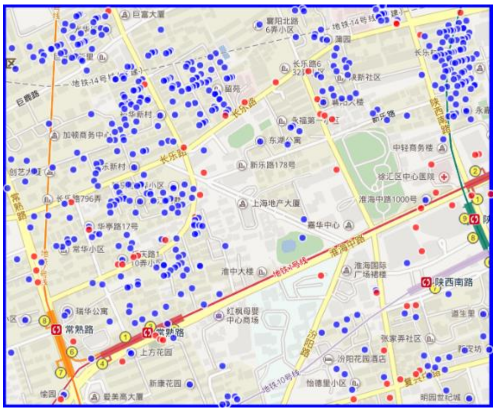

Figure 2 Residences (Blue) and Supermarkets (Red)

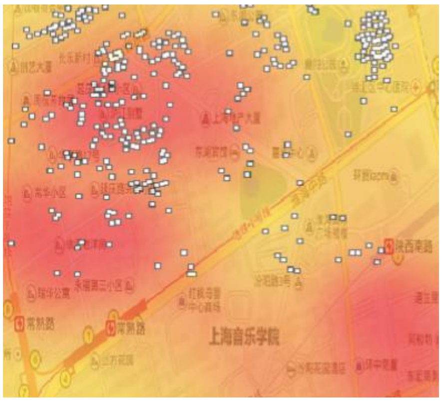

Figure 3 - Heatmap of bikes and residences (White) in one night

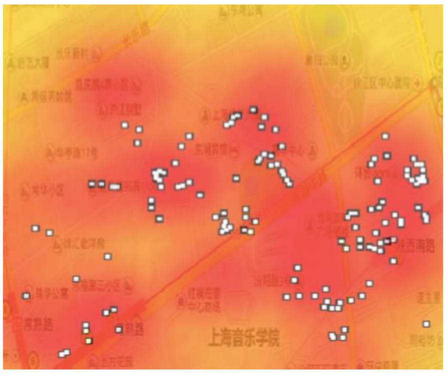

Figure 4 - Heatmap of bikes and office buildings in one morning

3. FIRST HYPOTHESIS AND BASIC SOCIAL ANALYSIS

With these gathered insights, definitions of proximity factor and key services, as well as the raw analysis of the data gathered, one can already draw some basic information and hypotheses, that must be proven in the next steps.

At this stage, one can logically assume by sociological experience that, for example, there are more bikes with a lower proximity factor, meaning that, there will be more bikes close to key services, than further away, counting for the total of Mobikes registered in the study. This might look like a basic assumption, but if true, it can have several different influences in urban fabric, arguably directing city's development in a way that implies a cordiality and symbiosis between urban fabric, various layers of infrastructure, influencing the flow of shared bicycles all over the area of study and, therefore, the flow of people, life and image, in the city.

Another assumption or hypothesis that one can make is that there is a higher number of available bikes during weekdays, though, this hypothesis is exclusive to the area of interest, since analysis of the urban framework at that scale, describes and shows the area of Xintiandi as a large and famous touristic, retail and office area. This will imply that more people will move from their houses to the area, leaving a trace of bikes. Consequently, exists a larger offer of available bikes in that area during the weekdays, since weekends consist of non-working days. Thus, by observing data, it should be noticed a lower number of bikes during weekends.

Finally, during these weekends, it is probable that the user's commute to office and work, transfer to other areas, namely residential areas of Shanghai, which retail and general office workers can afford. These residential areas obviously have a minor density of services (such as retail, office buildings, restaurants or malls) than the same service fabric of the perimeter of Xintiandi. Thus, by analysing this information, it is implied that the calculated proximity factors for all the available Mobikes during the weekend will be, overall, bigger than the one calculated during weekdays. Meaning that there will be an inferior overall number of available bicycles and that these will be more scattered through areas, further away from some critical key services, such as shopping centres, subway stations, office buildings and so on.

However, this may seem like a logical and unimportant hypothesis, yet it is of major importance to understand the implication and the significance of the singular content of the key services, and in the end, will allow to study, and filter the critical key services that have a general influence in the city and not in specific districts. Thus, the next stage would be to verify and determine the value of significance of each key service, by testing how many bikes have each service as their closest one.

4.PROXIMITY FACTOR

The proximity factor described during this paper, works to evaluate the quality of versatility in building typologies that surround the area of study, meaning that a constant and averagely lower proximity factor for most of the sessions and bikes accounted would indicate that most available bikes are dropped in places: a) With good urban quality and versatility; b) Very passible of being picked up, since services that promote mobility and interactions would be constantly nearby.

More information we gather from the analysis of the proximity factor for each bicycle are patterns of bike distribution during the week. We can analyse and compare the weekend record to the week and with enough resources, one can compare areas of development in between zones of city and lower the average proximity factor in the area which would eventually need improvement. Also, by reading high constant values on a specific site, it is indicative of a necessity for an *urban requalification*, meaning there are not enough key service infrastructures around the area.

4.1 Accuracy tests, closest key services and overview

Analysis of such data and tests might allow a deeper study of urban framework and lifestyle, optimizing the streets, facilities and services in regard of user experience. Instead of conditioning user experience to urban framework, it is possible to adapt the urban framework to user experience. Thus, proportionating a better symbiosis between systems. Subsequently, to narrow down the most significant key services, correlation with the number of closest key services is necessary.

As observed in Fig. 5, there is a natural and sociological correlation between restaurants and office buildings in urban fabric and daily lifestyle of people. It is visible in mapping graph that, there are usually more available bicycles closer to restaurants. But most interesting part is the chronological proportionality between these two, as the variation of bikes close to each service develop almost identically, at a given time. This is helpful as a condition for future development of areas, as it arguably can show strategies for planning.

It is interesting to compare the results obtained in the figures 5 and 6 as both pairs of key services present a hand in hand variation over time. It is uttermost important to understand the influence and significance that each key service prints in daily lifestyle of locals, so it can be assumed that services can have different values of significance in the quantification of proximity factor. Furthermore, by looking at the graphs, we can already assort natural relation existing between some services, providing a reliable pattern studying over time. The same can be pictured in Fig. 6, related to supermarkets and residences, but in this situation, the difference between the number of bicycles that are closer to residences and supermarkets is much higher than the one observed in Fig. 5. This is probably due to the massive domain of housing and residence-oriented areas, that exist in specific areas on sight.

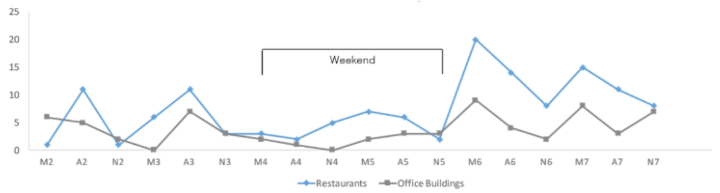

Figure 5 - Correlation between closest restaurants (blue) and office buildings (grey) to each bike over time.

Figure 6 - Correlation between closest supermarkets (green) and residences (orange) to each bike over time.

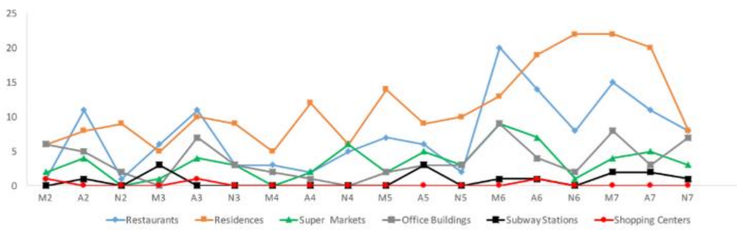

Figure 7 - Correlation of all closest key services (red – shopping centres, black - subway stations) over time.

Finally, by reading and overlapping closest key services to each Mobike over time [Fig. 7], it is easily understandable how to filter key services and select the most significant ones in the area being studied. It was decided to exclude the embedded key services within the residential areas, due to the overwhelming concentration of this typology in area of research as well as difficulties to extract this embedded information within the fabric of residential lanes. Also shopping centres usually do not have nearby bicycle parking lanes, since most of them have a big pedestrian plaza, landscaping, sculptures or monumental entrances. These steps lead to a data deviation in favour of other services nearby, indicating the shopping centres as elements of intensification on proximity factor calculation. Subway stations are obvious drop off points for Mobike. This also means that the data gathered from these might be corrupted since they had no influence in the studied social phenomenon. Finally, for reasons stated above it is down to three most significant services: Restaurants, office buildings and supermarkets. They represent the most balanced distribution among all the services.

Consequently, after the filtering of these significant services, it is important to understand if this data is accurate and therefore ready to be used. This results in the need of proceeding to a hypothesis null test against these services. In other words, this means that a way to see in a large scale if parking users are actually going to the service registered in the session and not to other services around must be found. Fortunately, there is a simple test to be done to verify if there is any major error in the methodology applied. Simply by calculating the number of second closest key services to all the bikes, and determine their building typology, and then, correlate it with the number of closest services of the same building type [Fig 8, 9, 10, 11].

	A	B	C	D	E	F	G
1	Closest S.	Restaurants	Office	Super	Shopping	Subway	
2	M2	1	11	2	1	1	
3	A2	14	7	7	0	1	
4	N2	1	7	6	0	0	
5	M3	8	4	2	0	3	
6	A3	14	12	7	1	0	
7	N3	4	10	4	0	0	
8	M4	2	4	2	1	0	
9	A4	9	2	5	0	0	
10	N4	8	2	7	0	0	
11	M5	12	4	9	0	0	
12	A5	9	7	7	0	3	
13	N5	3	5	9	0	0	
14	M6	26	11	13	0	1	
15	A6	19	13	12	1	1	
16	N6	15	11	9	0	0	
17	M7	20	19	10	0	2	
18	A7	19	8	12	0	2	
19	N7	13	8	3	0	3	
20		197	145	126	4	17	489
21	2nd closest	Restaurants	Office	Super	Shopping	Subway	
22	M2	1	10	2	1	1	
23	A2	12	6	7	2	2	
24	N2	4	7	6	0	0	
25	M3	4	4	3	0	4	
26	A3	11	10	10	2	0	
27	N3	4	8	6	0	0	
28	M4	3	5	2	0	0	
29	A4	7	4	5	0	1	
30	N4	2	5	9	0	1	
31	M5	9	4	12	0	1	
32	A5	8	7	7	0	4	
33	N5	5	6	6	0	0	
34	M6	16	17	14	0	3	
35	A6	18	14	10	1	3	
36	N6	9	11	12	2	1	
37	M7	11	24	14	0	2	
38	A7	11	13	14	1	2	
39	N7	8	10	5	0	3	
40		143	165	144	9	28	489

Figure 8 –Table of closest services registered for each bike

Figure 9 – Correlation between closest and 2nd closest restaurants over time

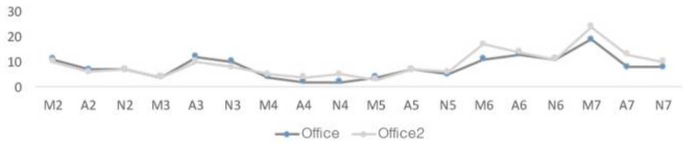

Figure 10 – Correlation between closest and 2nd closest office buildings over time

Figure 11 – Correlation between closest and 2nd closest supermarkets over time

For a simpler explanation, it is interesting to look at the phenomenon of picking or dropping a bike of one individual, and then apply this logic to a large-scale event. Imagining that someone would park the bicycle in a specific location, it does not mean that this person would obligatorily move to the closest key service in the area. This happens because of several conditions, either for parking spots available, or just individual commodity, just the thought of a person always going to the closest service to where they park their bicycle sounds illogical. What happens is that this person would probably go to the second or third closest key service, and with our current data it would be unclear what the actual destination is. This implies that in one recording session at an individual scale, the unity representing this person's Mobike would disappear from the closest key service and instead be added in the list of the second closest key service. If we multiply this event to a group of people, what we would see in the data would be a different value between the closest and second closest key service, meaning that people who were originally parking closer to a restaurant, maybe would go to the nearest supermarket, and the number corresponding to these people, would shift to another key service on the second closest service list [Fig. 9]. This implies a change in the slope of the graphs of the closest and second closest key services. Consequently, to verify if this large-scale shift of the numbers occurs, the only thing to do is to correlate the same graphs with the closest service and the second closest, and if there is a massive deviation of the pattern of the graph, it means that this test model is not accurate for a large-scale study of this phenomenon, while if the graphs have a similar slope over time, it means that the distribution is even and at a large scale. These errors either end up compensating and cancelling each other out or the number of people that are not going to the closest key service does not represent a significant percentage of the studied population. [Fig. 9,10,11] Either way, it ends up safe proving the reliability of the test model that is ideal for this study.

4.2 Proximity factor supported calculation and methodology

Now that the accuracy of the test model is proven, there is only one step missing to completion and final reading of the studied data. The calculation of the Proximity factor tuned according to area specifics, and in this case, Xintiandi.

Our data: the total amount of bikes accounted for closest key services; and the total amount of closest services by category, will allow us to establish three constants in the calculation of our proximity factor.

These appear from the percentage model [Fig. 12], by calculating the number of restaurants, supermarkets and office buildings that are closest to Mobikes and dividing it by the total number registered, we now have a percentage constant, attributed to each significant key service. This constant is area specific and must be calculated again, in order to proceed to the calculation of the proximity factor in another site, since, as explained before, other areas have further or different preponderance to infrastructure typologies and therefore, the results or the significant services might change. In this case it was office buildings, supermarkets and restaurants, in other cases it might be industry, residences, green public spaces or every other significant building typology that is present in the study.

Simple things that we can read from the graph below are, for example, that office buildings have a higher percentage of closest services during weekdays while restaurants appear to have a higher number during afternoon sessions. It is important to notice the different percentages during morning, afternoon and night, according to the service and lifestyle of the users.

Finally, the proximity factor is represented by equation (1), which represents the sum of the distance of one available *Mobike* to the closest service multiplied with the respective significance factor, with the distance of the same bicycle to the closest supermarket multiplied with the significance factor of supermarkets in the area and so on, for every significant building typology on the site. This results in an accurate split of the overall distance to the selected key services with the respective significance factor applied in a percentage model.

It is also important to mention that weather and other conditions that influence bicycle usage in the city should be accounted for. In this case, it was planned to be used as a binary exclusion factor, when accounted for rain or other events in the area that would incapacitate or influence the use of bikes, but since the data was only recorded for one week, and no such event or weather-related condition occurred during the week of data collection, it was not necessary. This will finally lead to our tuned data graphs and will allow us to further analyse the results and draw the necessary information about the specific area's social activity.

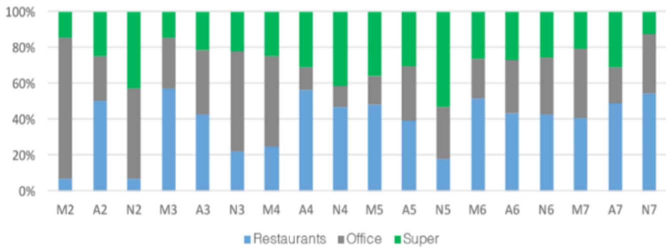

Figure 12 Percentage model of the number of bikes with closest services and their typology

$$P = \sum_{i=1}^{n} D_i \times F_i \quad (1)$$

Where:

P = Proximity factor for one registered bicycle
D_i = Distance of bicycle to closest indexed service
F_i = Significance of indexed service
n = Number of correspondent key services

4.3 Final results analysis

With the proximity factor calculated for all the bicycles recorded, it is possible to visualize the data in a heat graph and verify the previously stated hypothesis [Figure 13]. As seen previously, the weekend starts at Morning 4 (M4 in the x axis) and a large decrease of available bikes from the night before is visible. Xintiandi is a very well-known night-life area in Shanghai and this observation supports that statement, seeing as almost half of the available bicycles in N3 (Friday night) were taken from outside the area of study, to other areas.

On the other hand, M5 (Monday morning) presents almost 50 available bikes in the area with very low average proximity factor (below 0.4 average). This can be read as a successful distribution of the bicycles and that they are reasonably close and accessible to any city user in the area, indicative of a good urban efficiency and quality.

Overall, in the Xintiandi area, bicycle proximity factors seem generally low, which is good, because considering overall range to services and infrastructures one can assume an urban quality that regulates accordingly to the city life, except maybe, at night where some records tend to demonstrate less available bikes and higher proximity factors.

One of the next interesting steps to take would be to compare it with other Shanghai areas and other cities similar zones and determine which areas might need urban requalification or improvement in service and building quality, by estimating an adaptive model that falls between the standards of city regulation and classification.

Another optional interesting step is to automate this analysis with a computer and use a machine learning approach to be constantly evaluating and analysing the data, allowing a more accurate result with less intervals between periods of screenings and new evaluation, intervention and analysis methodologies.

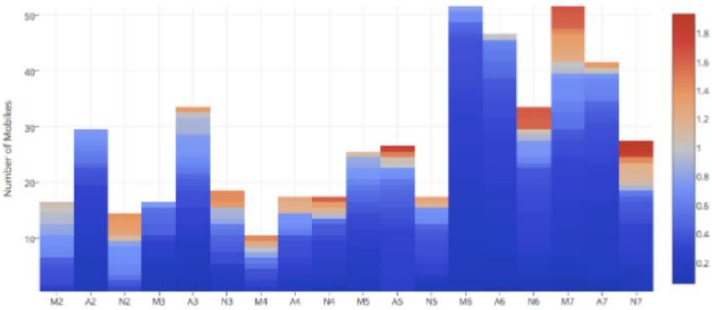

Figure 13 – Heatmap of proximity factor for registered bicycles over time.

5. CONCLUSION

After this thorough analysis and test model tuning and development, some basic but interesting facts that support the hypothesis can be assumed. In the night, records tend to register less available bicycles in the Xintiandi area. This is important, since Xintiandi is also a famous night-life area, with several night facilities and the fact that less bikes are available during night time, means that people using nocturnal facilities will not have access to such a big offer of available bicycles. Although, this does not happen in late morning records and afternoons. This occurs due to the registered increased tendency of available bicycles in the residential area. Also, offices and restaurants register their highest peaks on mornings and in afternoons. This translates the normal people working routine and strengthens the bond between urban fabric and daily life. It can be deduced that the influence of the urban fabric in people's lives is enormous and implies a direct **one-way influence of the urban fabric.** The final outcome should not be this one, but one where user experience and building typologies have a reciprocating influence on the co-existence of people and buildings.

This will allow a more efficient urban planning process, over the master plan of the area, and its respective function and services. Furthermore, with this new accurate test model, there are several things that can be improved in the scope of urban planning, among them are:

- The providing of a better insight under which areas should suffer future redevelopments. Ultimately, would increase the area life-style efficiency;
- The improvement of cyclers safety, by adding or re-structuring the bike lane frameworks in the existent streets;
- Allows to establish common grounds and predictors, between city mapping, Urban mobility and future development in a large scale.

ACKNOWLEDGEMENTS

This research was partially supported by Tongji University, Tecnico Lisboa and Disney Research Lab in Shanghai. We thank our colleagues from Tongji university who provided insight and expertise that greatly assisted the research, although they may not agree with all the interpretations/conclusions of this paper.

My friends Jonas Herzer, MBS, Phillips-Universitat Marburg; Hejia Jin, Student, TU Berlin and Andrea Grossi, Urban Planner, Politecnico di Milano for comments that greatly improved the manuscript.

We would also like to show our gratitude to the “anonymous” reviewers for their so-called insights. We are also immensely grateful to Carmo Cardoso, Architect; Sheryl Shen, Student and Alexandra Erfort, Student for their comments on versions of the manuscript, although any errors are our own and should not tarnish the reputations of these persons.

REFERENCES

1. Brennan, Matthew (May 2017). “Bike Sharing and the City, 2017 White Paper”. China Channel. Retrieved May 17, 2017.
2. Andrews-Speed, Philip (November 2014). “China's Energy Policymaking Processes and Their Consequences”. The National Bureau of Asian Research Energy Security Report. Retrieved May 20, 2017.
3. Russel, John, (January 2017) “China's Mobike raises $215M led by Tencent for its bicycle sharing service”. Techcrunch.com. Retrieved May 20, 2017.
4. Hersey, Frank, (May 2017) “China's bike rental startups are learning a lot about how people spend their free time”. Techcrunch.com. Retrieved May 21, 2017.
5. “China's electric vehicles to hit 1 million by 2020: report”. Reuters. 2010-10-16. Retrieved 201705-10.
6. The High Road; “If China steers its auto industry toward hybrids and perhaps hydrogen cars, the world may have no choice but to follow.”(Cover Story), Newsweek International. Retrieved May 10, 2017.