Customer Profiling Based on Mobile Apps GPS Data

- A Case Study on Westfield Shopping Malls

Fangzhou Zhou, Xianghui Zhang, Xinglei Wang, Tao Cheng
SpaceTimeLab for Big Data Analytics
Department of Civil, Environmental and Geomatic Engineering
University College London
London, UK
{fangzhou.zhou.21; xianghui.zhang.20; xinglei.wang.21; tao.cheng; @ucl.ac.uk}

***Abstract*—In order to provide a personalised experience to customers, it's essential for shopping centers to understand its customer base and their shopping behaviors. Building a well-developed customer profile is critical for improving marketing efficiency, expanding market share, and building long-term, stable business ties with trading partners. Currently most shopping malls or retail business use footfall or customer surveys to grasp the customer behaviors, which are insufficient to obtain accurate and representative information about the customers. This study aims to provide a detailed customer profile for shopping centers using GPS datasets. We choose the two Westfield shopping malls in London as the case study area. In order to uncover additional customer information, this study focuses four research questions: (1) Origin places of customers; (2) Their transportation mode to the mall; (3) The average dwell time of customers; (4) The pattern of return visitors. According to the results, malls can develop a range of marketing initiatives to provide a better shopping experience for customers and attract more of them.**

Keywords-Customer Profile; Mobile Apps GPS Data; Origin of place; Transportation mode; Dwell time; Return visitor patterns

I. INTRODUCTION

In order to provide a personalised experience to customers, it's essential for shopping centers to understand its customer base and their shopping behaviors. Building a well-developed customer profile is critical for improving marketing efficiency, expanding market share, and building long-term, stable business ties with trading partners. At present, most shopping malls or retail business use footfall to grasp the customer behaviors [1], which is defined as the count of the number of people, or traffic, entering a store. Customer surveys are conducted to obtain detailed information, such as personal information of gender, age, home address, preferences for brands and events, frequency of visits and the money spend [2]. This information can help shopping centers to customise staffing schedules based on traffic patterns, evaluate marketing programs, and improve sales efficiency; as well as gain deeper insights into customers' shopping habits and experiences to increase their loyalty.

However, both the footfall count and the customer survey have their limitations. To conduct footfall analysis, it is usually for mall staffs to use manual clicker to keep track of the number of store visitors, or to use video cameras, thermal imaging, etc. to count the number of visitors. The former method is labor-intensive, and the latter two methods also require the purchase of appropriate equipment and analysis. Moreover, the footfall index can only indicate the flow of visitors and contains less information. Customer surveys can reveal more personal information, but the willingness of customers to fill out the questionnaire is usually low, and the return rate of valid questionnaires is generally not high. The low sample rate of survey limits shopping malls to obtain accurate and representative information about the customers.

To overcome the shortcomings of the two methods, Mobile GPS data has becoming popular to derive footfall of shop visitors. For example, Hoomans [3] used the footfall generated by GPS dataset provided by Colliers Spots to study the relationship between attractiveness indicators of a retail area and its performance. Mobile Apps GPS data collect geolocation and time information when the Apps are used. Compared to traditional GPS datasets, whose time interval of recording location is typically 1s, 5s, 30s, or 1min, the time interval of mobile Apps GPS dataset depends on user usage. When people open these apps, their Device IDs, coordinates, time, speed, the precision of location, etc. are recorded. With this information, we can make geospatial analysis related to people's activities besides footfall. Based on the above advantages, we can use the GPS Data to carry out customer profile analysis for shopping malls, which not only can eliminate the labor and material consumption in the footfall collection stage, but also can help malls to understand more about their customers' details, so that they can formulate sales policies accordingly.

The aim of this study is to provide a detailed customer profile for shopping centers using Mobile GPS datasets. In order to uncover additional customer information, this study focuses the following four research questions:

- Where are they from?
- How do they come to the mall?
- How long do they stay in the mall?
- The pattern of return visitors;

By understanding where the customers come from and how long they stay in the mall, the mall can plan its regional marketing policy, such as advertising; knowing what kind of transportation the customers take to the mall can help to plan the parking lots and place advertisements along the tube lines where the customers take more rides; by understanding the composition of the return visitors (are they local residents, mall workers or customers?) and the frequency of visits and footfall

information of these return visitors, the mall can keep track of the operations, obtain effective information for staffing and targeted marketing activities.

II. Case Study Area And Data

We choose the two Westfield shopping malls in London as the case study area. Westfield London and Westfield Stratford City are two big shopping malls in London, which are also two of the largest shopping malls in Europe, both belong to Westfield Group, respectively located in Borough of Hammersmith and Fulham and Borough of Newham. Westfield London has 293 stores, including an area of Luxury Retailers (The Village). Westfield Stratford City has 227 stores, including three hotels and one casino. The comparison of the analysis results of the East and West malls can reflect the difference in the needs of different customer groups in different regions; due to the huge customer base of these two malls and their important position in London retailer industry, the malls and the retail industry can understand the needs of their huge customer base and thus improve their sales performance.

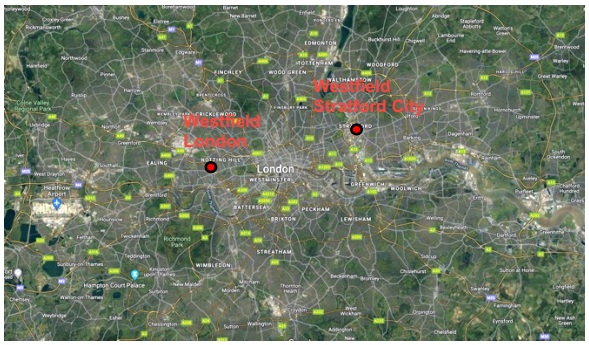

Figure 1. Location of Westfield London and Westfield Stratford City

The MSOA and LSOA boundary data used for the spatial analysis were obtained from https://data.cambridgeshireinsight.org.uk/dataset/output-areas/resource/0e5ac3b8-de71-4123-a334-0d1506a50288, and we extracted the data of London city for later analysis. There are 983 MSOA and 4835 LSOA in Greater London. The map data for the Westfield London and Westfield Stratford City malls were manually marked in the ArcGIS platform, and a 20 meters buffer zone was added to the exterior of each mall boundary to contain the activity space and pedestrian paths outside the malls.

We chose two months of data for February 2020 for analysis. In February 2020, GPS activity data were monitored for 81,860 customers in Westfield Stratford City and 30,355 customers in Westfield London.

III. Workflow

Here we define the workflow to derive answers to the four questions defined in Section 2.

In order to analyse the origin information of customers, firstly we used the data for each day of February 2020 and 2022, and used spatial join function to connect it with MSOA and LSOA data to extract the GPS point data that appeared inside the two Westfield malls on that day; then we merged all the data of that day with the extracted data within the mall area based on user IDs to identify customers who came to the mall on that day; and then we used groupby function to extract the first data of that day for customers who have visited the mall, the region where the data is located as their origin of place; finally, we used spatial join function to connect the origin information with MSOA and LSOA data to calculate the number of customers who came to the mall from each MSOA and LSOA, merged the data for one month, and visualized the results in the London map.

For the traffic mode analysis of customers, we developed a set of rules: first we extracted the data of customers who came to the mall that day for 20 minutes before they entered the mall, which contains their travel mode information at each record point (there are 7 types of travel modes: walk, car, cycle, bus, tube, train, stationary). We then determined how the customers came to the mall based on the extracted travel mode information.

For the dwell time of customers, since each log of GPS data has its time difference from the next log of data, we first counted the dwell time of each customer in the mall using the groupby function for the time difference of each log of GPS data; then we used the merge function to connect the dwell time information of customers with their origin of place information; finally, we used the groupby function to calculate the average dwell time of customers from each MSOA and LSOA in the mall, and visualised the results in the London map.

For the pattern of return customers, we extracted the patterns of local residents, workers and visitors (normal customers) based on a set of rules. We defined the local residents as the users whose home is in Newham 013G or Hammersmith and Fulham 004A (LSOA where Westfield Stratford City and Westfield London are located respectively). And we defined workers as if one person comes to the shopping mall at least 3 times one week, and the average dwell time in one day of one week is over 7 hours, then s/he is the worker.

After obtaining the lists of residents and workers, we began to analyze the time pattern of return visitors. We calculated the frequency pattern and time pattern(footfall) of return visitors separately. For the frequency pattern, firstly we took the concatenation of the lists of residents and workers to facilitate the subsequent calculations related to the visitors, and then we used value_counts function to count how many times each customer appears in these 28 days and calculate the percentage to count the frequency of 1-28's proportion. Finally we drew a bar chart based on the frequency and number of different type of visitors (total customers/local residents/workers/normal visitors). For the time pattern, we first extracted the visit hour information of every user in 28 days and used the groupby function to compute the time information data based on the date and hour, made a statistic on the visits of every hour in every day by 4 types of customers. Finally we generated 4 heatmaps based on the 4 time information data.

IV. Results

Figure 2 shows the results of the analysis of where customers come from. People tend to shop at malls that are closer to them, Westfield Stratford City has more customers from east London, customers of Westfield London are almost from west London. The largest number of customers are from MSOA where the mall is located.

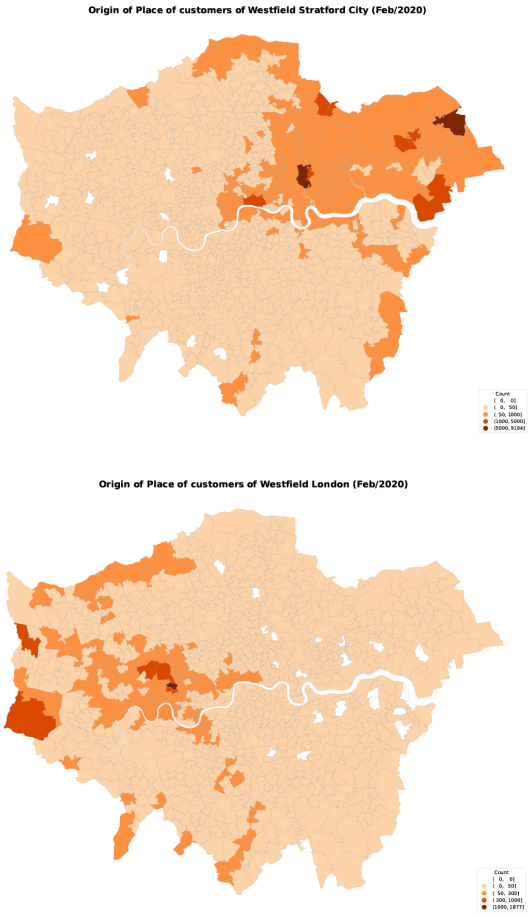

Figure 2. Origin of place of customers

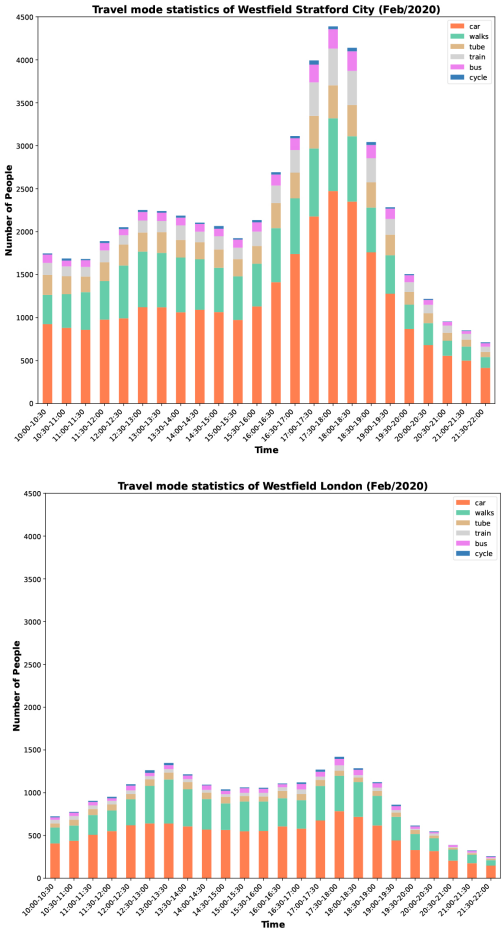

Figure 3. Travel mode statistics

Figure 3 shows the travel mode statistics of the mall where the customer came. The figures show that the largest proportion of customers came by car. The number of people coming in Westfield Stratford is much greater than that in Westfield London, partly because of the difference of the mall size. And there is a significant spike during lunch and dinner hours. In Westfield London, this peak occurs between 13:30 and 14:00 at lunch time and 17:30 and 18:00 at dinner time; in Westfield Stratford City, this peak only occurs between 17:30 and 18:00 for dinner.

Figure 4 shows the average dwell time of customers. The average dwell time of customers in Westfield London is higher than in Westfield Stratford City. With the newer facilities and clearer partitioning of the mall, people may have a better shopping experience and thus tend to hang out and rest for longer periods of time in Westfield London. It is also clear that some customers from more distant areas will also tend to stay for a longer period.

Figure 5 shows the frequency of visits of customers of these two malls. Except for the workers groups, the frequency of visits of other groups follows a decreasing pattern by frequency: the resident and visitor groups have the highest number of one-time visits in a month, both above 50%. Specifically, the percentage of residents visiting once is higher in Westfield London, at over 70%, and in Westfield Stratford City, at over 50%, indicating that the residents of Westfield Stratford City visit the mall more frequently; the same is true for the visitor group. This indicates that Westfield Stratford City has a higher number of repeat customers. For the workers group, the frequency of their visits shows a segmented pattern, with Westfield London's workers concentrating on 10-20 visits and Westfield Stratford City's workers concentrating on 3-12 visits, with the difference in frequency representing the different schedule of workers in two malls.

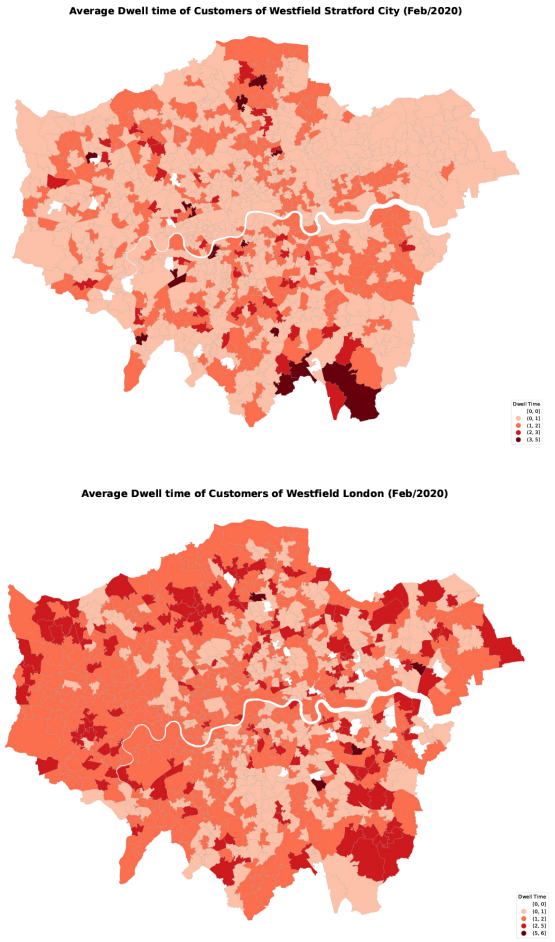

Figure 4. Average dwell time of customer

Figure 6 shows the hourly footfall of customers of these two malls. Each figure represents the visit counts of all customers, residents, workers, and visitors respectively. Each square indicates the number of customers to the mall for that date and one-hour period. Westfield Stratford City shows a clear pattern. Each type of customers visits the mall most on weekdays from 6am-9am and 4pm-7pm (with some missing data for workers). It's also the rush hour in the morning and evening, the data may record the commuting activity of people at stations near the mall. In Westfield London, people also visit more in the morning and evening peak, which is more pronounced in the visitor group, less pronounced in the resident group. For the worker group, it's more pronounced in the morning peak. Unlike Westfield Stratford City, Westfield London has more visitors from 9am to 4pm on weekends, even more than the weekday morning and evening peak, reflecting that Westfield London has given full

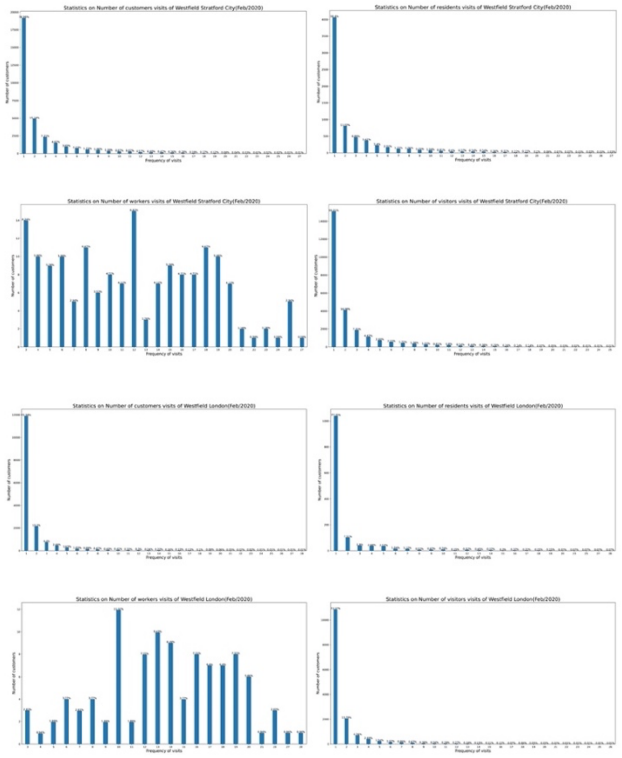

Figure 5. Frequency of visitis of customers

play to its function as a shopping mall, providing a good shopping space for people on weekends and attracting more visitors.

V. DISCUSSION

This study attempted to create a customer profile for Westfield shopping mall by analysing four aspects of customer behavior patterns: the origin of place, the mode of transportation to the mall, the average dwell time in the mall, and the pattern of return visitors. The results show the number of people who came to the mall from each MSOA in London and the average dwell time of customers from each MSOA; the number of customers by different modes of transportation and their arrival time; and the pattern of arrival time of the three groups of residents, workers, and visitors to the mall. The results show that the largest number of customers arrive by car, with two peak periods during lunch and dinner time. Westfield London attracts more customers and fully utilises its function as a mall; this is also indicated by the fact that the average dwell time of customers in Westfield London is higher than that of Westfield Stratford City.

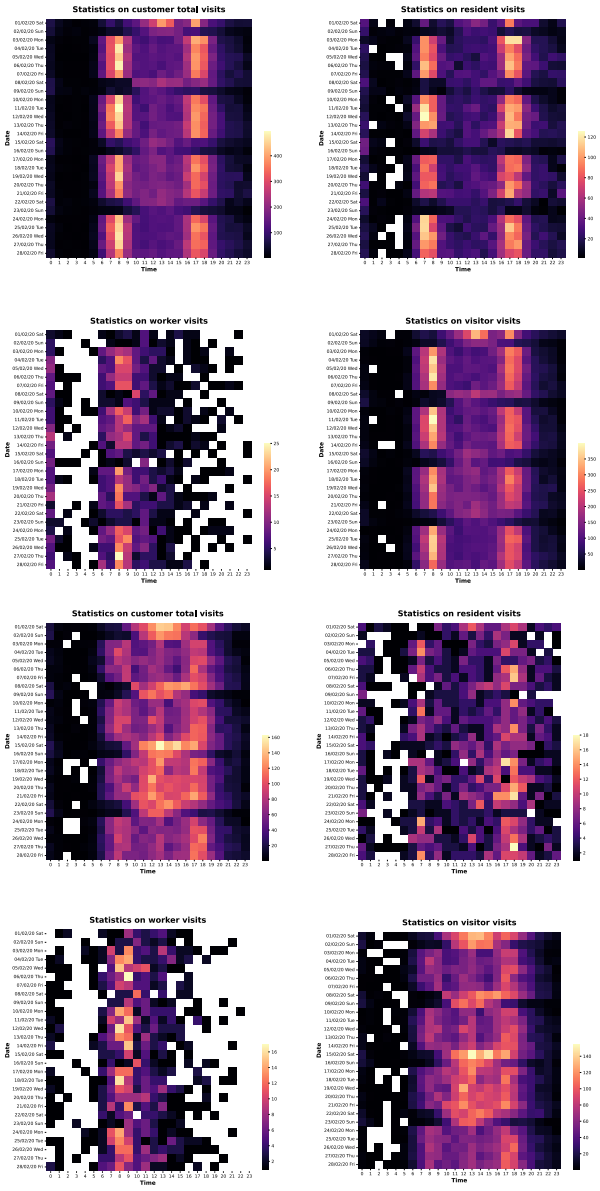

Figure 6. Hourly footfall of customers

According to the results, malls can develop a range of marketing initiatives to provide a better shopping experience for customers and attract more of them. For example, by determining the number of customers arriving by car and bicycle, the number of parking spaces can be adjusted accordingly; during lunch and dinner hours, the mall could have more staff to serve more customers than usual. In Westfield Stratford City, for commuting residents and visitors, a food market can be hosted on the outdoor during the morning and evening rush hours so that people passing by can quickly buy food; and for weekends when people come less frequently and do not stay as long, the layout can be re-planned and new marketing strategies can be developed to attract more visitors.

Since our data is updated in real time, we can adjust our marketing strategy in the future by combining the current real-time data; we can also analyse the seasonal pattern of customer visits based on the previous data and plan the marketing strategy by season. Based on the information of customers' origin and dwell time, the socio-demographic data can be combined to analyze the characteristics of each type of customers, such as which age group or which income group of customers visit the mall most and stay the longest, so as to develop targeted marketing strategies.

ACKNOWLEDGEMENT

This work is part of the Consumer Data Research Centre (CDRC) project supported by the UK Economic and Social Research Council (ES/L011840/1).

REFERENCES

[1] Mumford, C., Parker, C., Ntounis, N., & Dargan, E., "Footfall signatures and volumes: Towards a classification of UK centres," Environment and Planning B: Urban Analytics and City Science, vol. 48, pp. 1495-1510, 2021

[2] Makgopa, S. S., "Determining shopping malls customers' satisfaction and loyalty," Journal of Business and Retail Management Research, vol. 48, 2018

[3] Hoomans, N., "The relationship between attractiveness of retail areas and their performance", 2022