Exploring Activity Patterns of The Taipei Public Bikesharing System

Jinn-Tsai WONG a, Chun-Yen CHENG b

a,b Department of Transportation & Logistics Management, National Chiao Tung University, Hsinchu, 300, Taiwan, ROC

E-mail: jtwong@mail.nctu.edu.tw
E-mail: cyencheng@gmail.com

Abstract: This paper presents the insights of imbalanced public bicycle distributions, i.e. unable to pick up/return bikes due to empty/full stations through the analysis of spatio-temporal activity patterns of bike stations. YouBike, the public bikesharing system of Taipei City was examined. Taking advantage of Open Data policy, the changes of the number of available bikes across all stations were collected to identify station activity patterns. The relationship between spatial characteristics and station activity patterns were explored. The clustering results indicate that station activity patterns could be categorised into three groups and each reveals different activity patterns throughout the day. The visualisation of average temporal activity patterns and clustered groups are illustrated as well. Such results could provide better understanding of bikesharing usage and the underlying temporal and spatial dynamics of a city.

Keywords: Public Bikesharing System, Station Activity Pattern, Data Mining, Cluster Analysis

1. INTRODUCTION

Growing concerns about climate change, global warming, energy security, and unstable fuel prices have aroused a large number of policy makers and experts to explore sustainable travel solutions (Shaheen et al., 2010). As a result, a variety of non-motorised transport modes have been promoted by transport planners, professionals and policy makers in recent years, which are often considered as vital elements in sustainable transport (Rietveld and Daniel, 2004). Promoting active transport from motorised travel towards walking and cycling would expect to yield environmental benefits, such as limiting greenhouse gases emissions, reducing air pollution, noise, and alleviating traffic congestion (Rabl and de Nazelle, 2012; Woodcock et al., 2014).

Moreover, it is increasingly recognised that cycling and walking represent practical opportunities for people to integrate physical activity into daily life and yield positive impacts on public health (Cavill et al., 2008; Dill, 2009). Accordingly, public bikesharing system (PBS), which is economical, eco-friendly, healthy, ultra-low carbon emissions and more equitable, have increasingly received attention in the last decade and have rapidly emerged in many cities all over the world.

The PBS could be considered as an innovative scheme in the realm of urban transport (Parkes et al., 2013; Zaltz Austwick et al., 2013; Bührmann, 2007). It provides the missing link between existing points of public transportation and desired destinations (Midgley, 2009). It also bridges over the distances that are deemed too far to walk, but too close to justify a car/public transport trip (Casiello et al., 2013). In this role, bikesharing systems increase
transit accessibility. Accordingly, it could be an alternative and a complementary transport mode, as part of green and versatile public transport in cities (Borgnat et al., 2011; Faghih-Imani et al., 2014). Not only a PBS can be interpreted as an individual mode for short trips but also served as a vital segment of an intermodal route for long trips (Nair et al., 2013). If it serves for an “extension” of the existing transit system, PBS could be construed as a first-mile or last-mile connection (DeMaio, 2009; Liu et al., 2012; Casiello et al., 2013). In fact, it is suitable for its fast, convenient and flexible characteristics in short trips. As a result, commuters in urban area could use shared bicycles to connect to their desire destinations from their homes, working places, schools, transit stops or other places. It implies that bike-sharing systems play the role in increasing transit accessibility, which is strongly aligned with integrated transit systems that aim to increase the transit catchment area (Nair et al., 2013).

2. STATEMENT OF THE PROBLEM

Currently, there are almost 700 cities in the world have implemented bikesharing systems till the end of April 2014 and more than 200 cities are planning to install public bikesharing (DeMaio and Meddin, 2014). Many of them seem very successful. Nonetheless, there are problems resulting in user frustration. Research is certainly needed. A number of studies on bikesharing systems have mainly focused on (a) optimizing bike station locations (Lin and Yang, 2011; García-Palomares et al., 2012; Martinez et al., 2012; Sayarshad et al., 2012; Hu and Liu, 2013; Lin et al., 2013), (b) user perception and satisfaction of the systems (Bordagaray et al., 2012), and (c) exploring spatial and temporal characteristics from the station hire data (Froehlich et al., 2008; Froehlich et al., 2009; Kaltenbrunner et al., 2010; Borgnat et al., 2011; Vogel et al., 2011; Lathia et al., 2012; Borgnat et al., 2013; O’Brien et al., 2014).

According to the survey done by Taipei Friendly Environment Association (TFEA), publics reckoned that there were six major issues should be addressed in implementing the Taipei Youbike system: (a) bike stations required to integrate with the system of New Taipei City (46.2%), (b) bikes pick up and return problems due to imbalanced distribution (36.3%), (c) high broken rates of bikes (10.5%), (d) broken docks (9.3%), (e) uncomfortable saddle (6.6%) and (f) fare (6.6%). Similarly, there were also two biggest problems in Barcelona’s public bikesharing system: (1) unable to find a bike while trying to start their journeys and (2) unable to return the bike in the desired destinations due to stations occupied fully with bicycles (Kaltenbrunner et al., 2010). The problem of impossibility to pick up/from bikes due to imbalanced distribution seems to be a universal feature and is one of the main issues raised by many users in cities where have implemented bikesharing (Kaltenbrunner et al., 2010; Shaheen et al., 2010; Vogel et al., 2011; Shaheen and Guzman, 2011; Zhang et al., 2012; Liu et al., 2012; Nair et al., 2013). Normally, this issue could be addressed through design and management measures such as incentives provided for helping redistribution of bikes in the case of Vélib, Paris or informing the users about the closest stations via mobile apps.

The bike stations usage data reveals spatial and temporal bikesharing patterns; moreover, these patterns reflect the culture and the spatial layout of the city (Froehlich et al., 2009). For instance, station activity patterns may disclosure relations with certain type of customers using certain stations depending on the stations’ surroundings. Better understanding of station activity according to the specific city context may not only help transport planners and operators to evaluate the underlying temporal and spatial dynamics of a city but also facilitate daily operations. Temporal and spatiotemporal patterns among bike stations of Barcelona bikesharing system were explored by Froehlich et al. (2008, 2009). Clustering was applied to
illustrate the patterns associated with location, neighborhood, and time of day. Predictive models were implemented to predict the near-term availability of bicycles at each station. Results showed that neighboring stations are likely to share similar usage patterns. Numerous researches also used a hierarchical clustering method to generate clusters and investigate usage patterns geographically distributed in the city to understand the impact of the inhomogeneity of the city on the long-run activity of stations (Froehlich et al., 2009; Borgnat et al., 2011; Vogel et al., 2011; Lathia et al., 2012).

Kaltenbrunner et al. (2010) observed that different patterns appeared in different stations, depending on the station location and time of day. The results highlighted the importance of the dynamics of neighboring station for predicting bicycle availability at a given station. Borgnat et al. (2011) analyzed temporal and spatial patterns of station activity of Lyon's public bike-sharing system, Vélo’v. Temporal patterns of the bike-sharing system usage were examined. It revealed that there were three peaks on weekdays (i.e., 8am – 9am, 12am – 1pm, and 5pm – 7pm) and two peaks on weekends (1pm – 2pm and 4 pm – 6pm). Furthermore, spatial patterns were examined as well in terms of clustering flows of activity between stations, which clearly showed the dynamics on the network in space and time, indicating a preferred short distance use of the shared bicycles.

Vogel et al. (2011) also showed that there were five different activity patterns based on their temporal pickup and return activity at stations. Lathia et al. (2012) used hierarchical clustering algorithm to group the stations based on their usage patterns to investigate effects of the access policy change. Six clusters were generated and represented as three typical types of behavior among stations: Day-Time Origins, Day-Time Destinations, and Combined Origins/Destinations for both pre-and-post datasets. The results indicated neighborhood stations tend to have similar usage behaviors.

Addressing imbalanced bicycle distributions is still one of the research topics on public bike-sharing system. Better understanding of station activity can help operators to evaluate the underlying temporal and spatial dynamics of a bike-sharing system. Therefore, the primary aim of this study is to contribute to the solution of imbalanced distribution of bikes, i.e. unable to pick up/return bikes due to empty/full stations through the analysis of station spatio-temporal activity patterns.

3. DATA

YouBike deployed in Taiwan since 30 August 2012, aiming at providing 166 stations and 5350 bikes in the following seven years. It was the first large-scale public bike sharing system to be implemented in Taiwan. At the beginning of its launch, 41 stations and 1460 bikes were available. The business model of YouBike is publicly owned and privately operated for a given time by Giant, recognized as the world’s largest bicycle manufacturer. YouBikes are robust yet aesthetically pleasant and are convenient for users, with radio-frequency identification (RFID) embedded technology, low step-through design, adjustable seats, anti-theft and shirt-guard design, front and real lights as well as reflectors for enhancing safety and security (YouBike, 2014). This re-development bike-sharing scheme seems to be a success beyond city government and citizens’ expectations. By the end of April 2014, there are about 1.75 million registered users and almost 19 million total rides since operation. Besides, YouBike is proud of its high turnover rate, up to 11 times per bicycle.

This study covers the spatio-temporal activity data among stations in Taipei YouBike system. As the ubiquitous ICT embedded in transport systems, it not only enables users to
acquire real-time information about bicycle availability but also amenable to investigation for researchers. Taking the advantages of ICT and Open Data, this research explores the potential of retrieving and analyzing real-time bike station data. There are 165 stations of data collected. Notably this study focused on the temporal analysis and clustering of bike stations through the retrieved data of bicycle availability collected at a fixed timespan automatically rather than collecting flow data of bikes at stations. Therefore, instead of identifying riders travel behavior over space and time via OD matrix, it demonstrates another way of looking human mobility data in an urban area. The data was collected automatically from a dedicated web Application Programming Interface (API) in JSON (JavaScript Object Notation) format which is provided freely by Depart of Transportation, Taipei City Government. The official website also provides information service for users through the Google map API, illustrating a map of Taipei City overlaid with small smile markers to indicate YouBike station locations, the amount of available bicycles and free slots at a given time for each station.

In order to analyze the station activity patterns, we have collected data online since April 15th, 2014 to May 27th, 2014 every 5 minutes, parsing it and storing in a SQL database all the relevant information. Overall, the collected data in our study was from 166 stations with a total of 7,280 slots provided. Total of 1,966,228 observations were collected, with only 0.277% of all observations were eliminated afterwards due to unexpected values or invalid station capacity. The station size per station ranges from 26 to 180 slots.

It is important to note that the collected data does not incorporate trips, i.e., trip originations and trip destinations. As a result, flows of bicycle between stations cannot be examined, indicating that this data cannot be used to count journeys. It only simply observes that the available bicycles and parking spaces at a given time.

4. STATION ACTIVITY PATTERNS

Firstly, temporal activity patterns of some specific stations are illustrated by weekday and weekend data. Then average system-wide temporal trends aggregated by all stations are illustrated as well. Secondly, clustering algorithm is used to analyse how station activity patterns are geographically distributed in the city. Finally, the analysis results are discussed.

4.1 Temporal Patterns

It is expected that the temporal patterns of each bike station reflects the actual bikesharing usage patterns which are underpinned by the daily routines of citizens in the city. Figure 1 shows an example of the temporal activity pattern of a station close to MRT Zhongxiao Fuxing Station surrounded by number of shopping hotspots such as SOGO department stores and East metro shopping street, restaurants and offices.

The number of available bicycles basically fluctuates over time whereas the total slots mostly remain at the constant level (i.e., 54 slots). It can be found that some sudden increases of available bicycles occur around 5 am which may be caused by rebalancing trucks moving bicycles from other full stations to empty or almost empty stations. It seems that bicycle redistribution plays a vital role of daily operation as bicycle redistribution can be found in most of days during the 6 weeks period. Hence it is clearly that the amount of bicycles has a morning spike which is mainly to accommodate the morning commuting peak and starts to drop during the period of 5 am to 10 am. Similarly, it also has a spike which is normally in the evening then drops soon and stays at a low level of available bikes till the next day morning.
Figure 1. The number of available bicycles (blue line) and total slots (green line) at station No. 0111: MRT Zhongxiao Fuxing Station (Exit 2)

Average activity pattern is shown in Figure 2, providing the more clear view in terms of the mean activity pattern of working days and weekends. Note that the golden shaded area in Figure 2 represents one standard deviation above and below the mean activity patterns.

Overall, the standard deviation of available bikes illustrates the observed patterns are quite fluctuating and tends to be larger during the working days rather than weekends. It may be probably due to in weekend users more likely to have similar trip purposes (i.e., shopping). While in weekdays this station serves mixed groups of people including commuters and people going for shopping.
Figure 2. Average station activity patterns of station No. 0111: MRT Zhongxiao Fuxing Station (Exit 2)

It can be observed from Figure 2 that the activity pattern of working days differs from those in weekends. The weekday patterns show a peak in the number of available bikes in the early morning between 6:00 to 8:00, which might be typical for commuting. The amount of bikes drops sharply between 6 am and 8 am and remains at the low level of available bicycles. This corresponds to the facts that bicycles need to be refilled before the commuting hour.

By contrast, weekend patterns reveal that five-pronged spike in station activity which corresponds to the morning, lunch, afternoon and evening (containing 2 spikes) respectively. It is interesting that the morning commute is still shown in weekend patterns. The apparent lunch spike which appears across both Saturday and Sunday occurs between 12:00 and 13:00, reflecting the number of bikes checked in is far greater than checked-out. As one might expect, this station attracts people come to this station for launch and shopping. In addition, a small spike of available bikes in the afternoon may be caused by the other groups of people coming
for shopping in this area.

Normalised available bicycles (NAB) is used to normalised stations’ data by dividing each observation by the specified station size, aiming to adjust different values to a common scale; thus it allows to be able to compare the usage of station activity in different size. Note that NAB may be interpreted as the proportion of available bicycles occupied in parking slots for a given station. It can be seen as a key measure for each bike station and used by numbers of studies (Froehlich et al., 2008; Froehlich et al., 2009; Vogel et al., 2011; Lathia et al., 2012; O’Brien et al., 2014). NAB ranges from 0 (empty) to 1 (full). Figure 3 illustrates the average weekday and weekend normalised available bicycles (NAB) at station No. 0111. It can be seen that fluctuations across the week are averaged out and smoothed, allowing the overall number of available bikes throughout the day more intuitively. Generally speaking, both activity patterns in weekdays and weekends follow the similar pattern in terms of their “M” shaped activity tendencies. During the working days, it is more likely that people are hard to find a bicycle between 7 am and 11 am as average NAB is below 0.15 (less than 8 bicycles) and the average standard deviation is around 0.12. In contrast to the loose M shaped average activity patterns in weekdays, the pattern is tighter instead on weekends due to much earlier drop of available bikes from afternoon.

Due to pages limitation, here another two bike stations were selected to illustrate the heterogeneity of station activity patterns. Figure 4 illustrates the average weekday and weekend NAB at station No. 0049: Longmen Square where is only 500 metres far from the station No. 0111 discussed previously. Figure 5 demonstrates the patterns at station 0128: Chengong Public Housing where is in the proximity to the several financial offices and buildings as well as residential areas.

The weekday pattern in Figure 4 showed that the initial rising of available bicycles
starting from 9 am which may reflect the location characteristics of the bike station in proximity to the shopping area where shops open in the late morning. It is until 18:00 to drop continuously till midnight. It indicates that inbound flows dominate in the daytime while in the night-time outbound flows lead by people leaving for home. In weekends, the peak of available bike appears at 1 pm which is probably caused by the people coming this place looking forward their lunches. Afterwards, it decreases though with a clear raise around 6pm until 8 pm followed by the nearly zero level of bikes.

Figure 4. Average NAB during weekdays and weekends of station No. 0049

The pattern in Figure 5 shows that people start to leave the region in the earning morning with a little spike around 9:00 and people return the bikes later in the afternoon or late evening. The level of NAB value is averagely below 0.25 after 9:00 and until 21:00 it finally recovers to higher available bikes, which indicates that the bike station is popular almost throughout the day. The weekend’s pattern is generally similar to weekdays. The onset of weekend activity pattern occurs earlier than that of weekdays and has a spike at around 9 am and then fluctuates at the low level of available bikes afterwards. Until 6 pm it starts to recover, indicating that people are more likely to return to this region.

The activity patterns are somewhat different in comparison among stations. Unlike patterns of stations around the MRT station or shop areas, activity pattern of station located in residential area seems to have longer period of lower number of bicycles; in addition, it shows that YouBike is popular at this region, and as a first/last mile of transport mode to and from the station to their desired destinations.
Figure 5. Average NAB during weekdays and weekends of station No. 0128

Accordingly, different stations have different station activity patterns which may be affected by their locations and surroundings; in addition, weekend patterns are significantly distinctive from weekday’s patterns. It is also found that station with similar geographic characteristics and spatial layout would follow the similar patterns. Therefore, using clustering could help us to group the stations according to their temporal activities, examine whether these activities depend on spatial factors or not.

4.2. Cluster Analysis

In this section, clustering algorithm is used to analyse how station activity patterns are geographically distributed in the city based on their usage patterns and explore how these activity patterns relate to underlying cultural and spatial characteristics of Taipei City. This study selects normalised available bicycles (NAB) for each station as clustering inputs for analysis.

4.2.1 NAB clusters on weekdays

According to Silhouette coefficient, 3 clusters appear promising. The results of these three clusters are believed to represent the typical behaviour of various stations in weekdays as shown in Figure 8.
Figure 8 Usage patterns for average NAB of 3 clusters in weekdays
**Daytime origins nighttime destinations:** Cluster 1, on the top of Figure 8, is the group of locations where users rent bicycles from in the morning and flow to in the evening. It shows a precipitous drop in the number of available bicycles between 6 am and 9 am as people leave for work and still declines thus reaching the minimum around noon. It is followed by the fluctuations around 30% of available bicycles during the period from noon to 6 pm. And the available bicycles recover to the early morning levels by midnight.

**Daytime destinations nighttime origins:** Cluster 2 refers to the morning destinations and the average station activity pattern shows the inverse of that in daytime origins. People begin arriving in the morning between 7 am and 10 am approximately; and it shows a slight decrease of available bikes during lunch and return just after the lunch break at around 2 pm. People begin leaving from 4 pm in the evening and with a sharp decrease of available bicycles while approaching the typical commuting time in the evening. The decrease lasts until 9 pm.

**Mixed Origins/Destinations:** Cluster 3 represents stations that display the combination of behaviours: both morning and evening leave. With a delayed morning leave in comparison to cluster 1, thus it reaches the first minimum around 10 am. Followed by the fluctuations of available bicycles, it begins climbing up till the time before 6 pm and declines again till midnight.

Overall, with the help of clustering algorithms, stations could be grouped according to their temporal activity patterns. The geographical distribution of NAB clustering results which are based on temporal weekday patterns without any geographical knowledge is visualised in Figure 9.

Exploratory analysis of the cluster results and their surroundings brings about the following findings. It is observed that stations with the same clusters are likely to be located in neighbouring; in other words, neighbouring stations tend to share similar usage patterns. We also note that the clustering algorithm has separated stations spread around the edges of downtown and outer area of Taipei City. Incoming stations (i.e., daytime destinations, cluster 2) are primarily located in high density commercial areas and working places such as stations proximity to Taipei station and stations proximity to Taipei City Hall for example, which is according to Taipei urban planning and land use zoning map. This supports the objective of YouBike that serves for commuters. Stations within clusters 3 (dark red) are most likely to be spread around the MRT stations and rail station with working places and schools nearby. Therefore, it supports the vision of YouBike as a fist/last mile connection to public transport.
4.2.2 NAB clusters on weekends

We now repeat the same clustering procedures to look at NAB clusters on weekends and 4 clusters were derived as shown in Figure 10. Generally speaking, these four clusters seem to be categorised in two groups:

**Daytime origins nighttime destinations:** Cluster 1, 2 and 4 are belonged to this group according to Figure 10. They share the similar patterns where bicycles are flowing out in the morning and returned in the evening. They basically illustrate that the overall number of available bicycles reach the minimum just before the evening. Though these cluster groups illustrates the pattern of morning leaving, it should be note that their leaving times differ.

**Mixed daytime origins/destinations:** Cluster 3 represents this pattern, illustrating morning incoming, afternoon leaving and evening incoming of available number of bicycles. It illustrates a very early morning incoming flows of bicycles, continually climbing up before noon and then fluctuating. Until approximately 2 pm, the available bicycles start to decrease.
instead till evening.

Generally speaking, most of stations (82.2%) belong to the groups of daytime origins nighttime destinations though cluster 1 and cluster 2 have some significant bicycle incomings during the daytime.

Figure 10. Usage patterns for average NAB of 4 clusters in weekends

Figure 11 illustrates the clustering results in a geo-visualisation of Taipei and it is observed that neighbouring stations generally share similar activity patterns. It is interesting that most of stations in cluster 3 (mixed daytime origins and destinations) are located in Xinyi CBD and Taipei station surrounding area. Stations in cluster 4 tend to be located on the edge of city centre, outer area of Taipei City and in proximity to riverfront parks and public facilities. Regarding stations in cluster 1, they seem to be located in proximity to MRT red lines and in Songshan district. 63% of stations in cluster 2 are proximity to MRT and rail station.
5. CONCLUSION

As urban infrastructures are increasing digitalised through taking advantage of information and communication (ICT) technology, large-scale of human behaviour data will become ubiquitous and more easily accessible that can be used to measure the interrelationship between the policy, design, and usage of transport systems.

This study shows that how public bikesharing usage data which refers to the changes of the number of available bicycles across all stations can not only reveal the bicycle usage patterns but also demonstrate the underlying temporal and spatial dynamics of a city. Visualisation of the average daily variation in station activity allows us to observe the overall temporal tendency of activity patterns throughout the day. The clustering results indicate that station activity patterns during weekdays could be categorised into three groups: which are
daytime origins nighttime destinations, daytime destinations nighttime origins, and mixed origins and destinations. Each clustered groups reveal the different activity patterns throughout the day.

It is believed that the visualisation of average temporal activity patterns and the clustered results could lead to better understanding the bicycle availability information. In addition, it is expected to improve the Taipei YouBike service through an improved redistribution of bicycles via trucks. As a result, it would help to improve user satisfaction with the service and it is possible to attract more people to use YouBike as an enhanced green transport system.

Based on the research process and results as shown previously, there are still some work can be performed to enhance the research quality. In addition, as a growing number of bikesharing stations and bicycles are appearing in the near future, some of the key research opportunities and recommendations for further research are discussed as following.

**Collecting longer period of bikesharing usage data:** to collect longer period of bikesharing usage data would allow us to explore station activity patterns more comprehensively through investigating the longer effect and the possible changes which might be brought by system expansion or minor changes of number of bicycles and slot in certain stations.

**Conducting questionnaire to investigate why people opt to bikesharing:** it should be noted that the data we collected by passive sensors would only tell us how, when and what changes occur of the number of available bicycles. However, it fails to incorporate any information as to why people choose bikesharing at a certain station at a certain time. In other words, while we can investigate the variations of bikesharing activities through seeking hints by inspecting places such as transport depots, amenities, or residential area surrounded by the station, it still exist the gap between usage patterns and human behaviour. As a result, it may be useful to conduct a qualitative survey to help clarify and better understand why travellers use public shared-bikes.

**Incorporating location factors into clustering:** the research could be further extended by incorporating location factors such as docking station size, docking station area, or maximum load factor of station during the day. Therefore, it is expected to examine the relationship between shared-bicycles’ activity and station location and surrounded amenities.

**Developing effective redistribution mechanism and management:** Results of the station activity patterns of each station and clustering analyses offer insights on the location of bicycle redistribution hubs and provide the robust information of bicycle availability throughout the day. Therefore, the optimal redistribution routes can be achieved and it allows us to know where and when to redistributing the bicycles to and from the full/empty station at acceptable cost basis.

**REFERENCES**


New York, pp.267-284.
Casiello, B. et al. 2013. Increasing Mobility in Dubuque: Developing Alternative Mode-sharing Opportunities. The University of Iowa.
Martinez, L.M. et al. 2012. An Optimisation Algorithm to Establish the Location of Stations of a Mixed Fleet Biking System: An Application to the City of Lisbon. Procedia - Social and Behavioral Sciences. 54(0), pp.513-524.


