Measuring the Pulse of the City through Shared Bicycle Programs

Jon Froehlich[1], Joachim Neumann[2], and Nuria Oliver[2]

[1]Department of Computer Science and Engineering, University of Washington, Seattle, WA 98195, USA
Email: jfroehli@cs.washington.edu
[2]Telefonica Research, Barcelona, Spain
Email:{joachim, nuria}@tid.es

Abstract

City-wide urban infrastructures are increasingly reliant on networked technology to improve and expand their services. As a side effect of this digitalization, large amounts of data can be sensed and analyzed to uncover patterns of human behavior. In this paper, we sense and analyze data from a new type of urban infrastructure called shared bicycling. We provide a spatio-temporal analysis of six weeks of usage data from Barcelona's shared bicycling system called Bicing. We show how these digital traces can be used to uncover daily routines, cultural influences and the role of time and space in city dynamics.

I. INTRODUCTION

We are nearing a time when even the most mundane objects and services will be digitized. As a result of this computational pervasiveness, our interactions in the physical world are increasingly leaving behind *digital footprints*. Recent work has shown the value in sensing these footprints to uncover new insights into human behavior [2], urban dynamics [3], and tourist movements [1]. In this paper, we explore the underlying "pulse of the city" of Barcelona through the lens of a new type of urban infrastructure: a 3rd generation shared bicycling program called *Bicing*. We emphasize not just what the data reveals about the patterns of human movement but also how these patterns reflect culture and the overall spatial context of the city. Our objective is twofold: (1) to highlight the potential of using shared bicycling as a new data source to gain insights into city dynamics and human behavior and (2) to introduce preliminary analysis techniques that we are developing to analyze the ever-expanding amounts of spatio-temporal data produced by urban infrastructures.

II. BICING: BARCELONA'S SHARED BICYCLING PROGRAM

Community shared bicycle programs offer an environmentally friendly, healthy and inexpensive alternative to automobile transportation. Recent technological advances have lead to a new generation of systems, which utilize technology such as RFID, mobile data services, and automated micro-payments to increase operational efficiency and reduce bicycle theft and vandalism. Barcelona's shared bicycle program, Bicing, was launched in March of 2007 (Figure 1). It is one of a series of extremely successful 3rd generation shared bicycle programs recently deployed in Europe. In late Summer 2008, Bicing grew to 373 stations with an average of 25.5 bicycle parking slots, 6,000 bicycles, and over 150,000 subscribers.

To check-out a bicycle, users swipe their RFID membership card at a Bicing station's kiosk, which then displays, via an LCD screen, the bicycle on the rack that has been unlocked. This information is uploaded to a server that keeps track of who has checked out what bicycle and updates the Bicing website with real-time information about the number of available bicycles and vacant slots at each station. A check-out provides 30 minutes of free ride time, every half-hour beyond that costs €0.30 for up to two hours. Bicycles can be returned to any station, where they are placed in an auto-locking rack. Warnings, monetary penalties (€3/hr), and eventually suspension of membership are possible if a user consistently returns a bicycle beyond the two hour limit. Bicing is open from 5AM to midnight on Sunday through Thursday and 24 hours during the weekend. To maintain even distribution of the bicycles (load balancing), a small number of trucks equipped with trailers move bicycles around the city.

Figure 1. (a) A nearly full Bicing station; (b) A station kiosk; (c) A close-up of a locked bicycle; (d) A map of Barcelona showing the location of the 373 Bicing stations. The five highlighted stations are discussed below.

A. Our Bicing Dataset

The Bicing website reports the status of all bicycle stations via a Google Maps visualization[1]. We scrape this webpage every five minutes and extract three data elements per station: the station's *geo-location*, the number of *available bicycles* and the number of *vacant parking slots*. We do not obtain personally identifiable information. Our data logger automatically accounts for new Bicing stations as they appear online. About 1% of the raw data extracted from the Bicing website suffers from one of the following problems: (1) the numbers are unreasonably large; (2) the numbers jump by more than five bicycles and then return to their previous values at the next observation; (3) sometimes all stations simultaneously report zero available bicycles and/or zero vacant parking slots. We clean the data simply by replacing the erroneous value with the most recent valid value. Our cleaned dataset includes six weeks[2] of station observations starting on May 23rd and ending on July 3rd, 2008 for a total of over 4.3 million data points. Note that 2 of the 373 stations consistently reported invalid data and were thus disregarded from our analysis.

In the following section, we analyze the temporal and spatial patterns of the Bicing dataset, in order to explore the underlying human behaviors and movement dynamics in the city of Barcelona.

---

[1] http://www.bicing.com/localizaciones/localizaciones.php
[2] Although our web scraper has been logging continuously since May 23rd, the Bicing website was down from July 3rd - 30th.
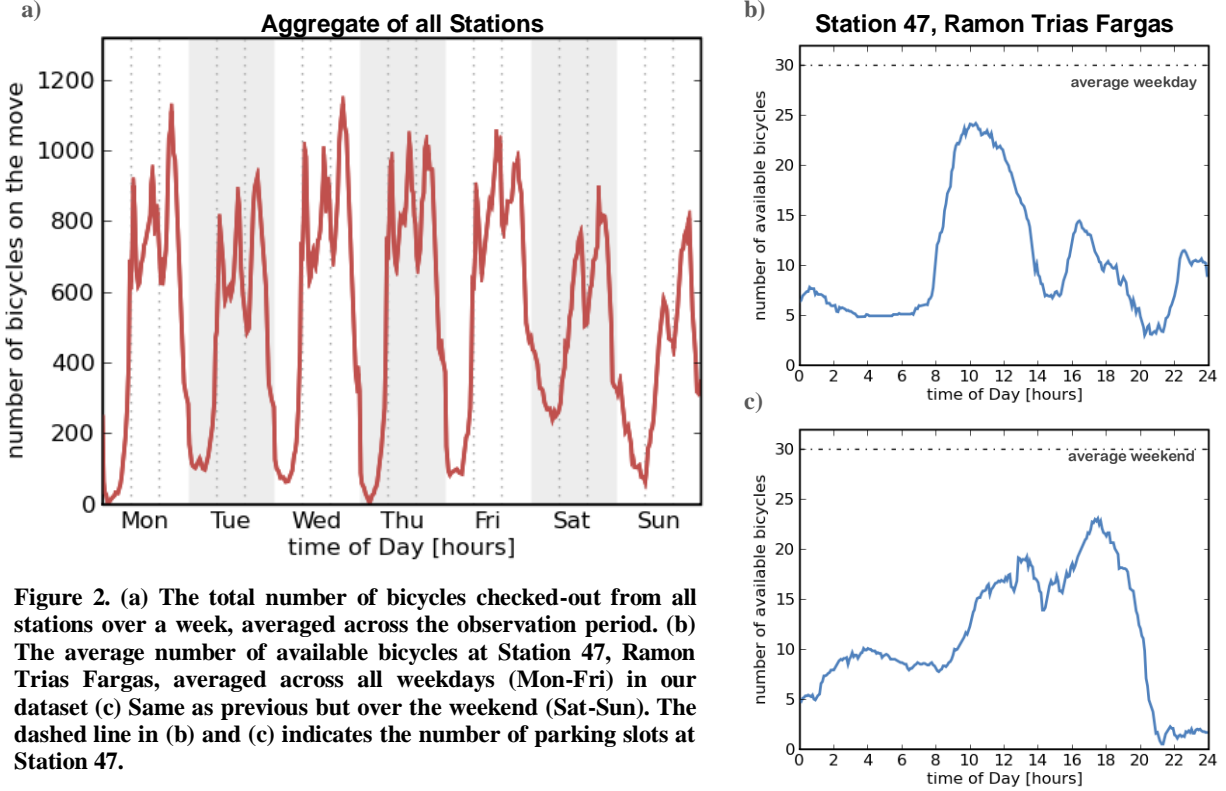
Figure 2. (a) The total number of bicycles checked-out from all stations over a week, averaged across the observation period. (b) The average number of available bicycles at Station 47, Ramon Trias Fargas, averaged across all weekdays (Mon-Fri) in our dataset (c) Same as previous but over the weekend (Sat-Sun). The dashed line in (b) and (c) indicates the number of parking slots at Station 47.

III. ANALYZING THE "PULSE OF THE CITY"

A. Temporal Patterns: Sensing Culture and Daily Routines

The temporal patterns of a city are a reflection of the daily routines of its citizens. Figure 2a shows the average number of bicycles "on the move" during a given point along the week. Perhaps the most salient feature of this graph is the repeating three-pronged spike, which corresponds to the morning, lunch, and evening commutes. As one might expect, the morning commute is absent in the two weekend days, resulting in a two-pronged spike. In addition, observe that the "lunch spike" occurs at 2PM, reflecting that Spaniards tend to eat a late lunch. In addition, the two most popular Bicing periods—*i.e.,* the periods with the largest number of bicycles on the move— are Monday and Wednesday night at around 10PM and that, on average, people tend to use Bicing more during the work week than they do during the weekend.[3]

To further highlight the contrast between weekday and weekend activity, Figure 2b and c portray the number of available bicycles at station 47, Ramon Trias Fargas, which is situated next to the University of Pompeu Fabra (see Figure 1d). Early on weekday mornings there are relatively few bicycles at the station. Then, at around 8AM, students, staff and faculty begin arriving on campus and the number of available bicycles increases quickly as people begin dropping them off. A local minimum occurs at 2PM as people leave for lunch and a second dip occurs around 7-8PM as people seem to be leaving for the night. In contrast to the weekday activity, on the weekends there is no sign of the 8AM commute. Instead, bicycles slowly trickle in throughout the day. Interestingly, however, both the weekday local minima and the weekend local minima seem to temporally align (at around 2PM and 8PM, respectively). We are not certain if this is a reflection of the lunch and dinner routines of these Bicing users or perhaps an artifact of the load-balancing implemented in Bicing via trucks. We are currently exploring methods to automatically detect the presence of the trucks in our data. However, our intuition is that they do not significantly bias our analysis.

---

[3] A keen reader may observe that the number of bicycles on the move does not drop to zero on weekday nights when the system is closed. We believe this an artifact of Bicing's operational logistics (bicycles are taken off racks during this time for maintenance) and not a reflection of human behavior.
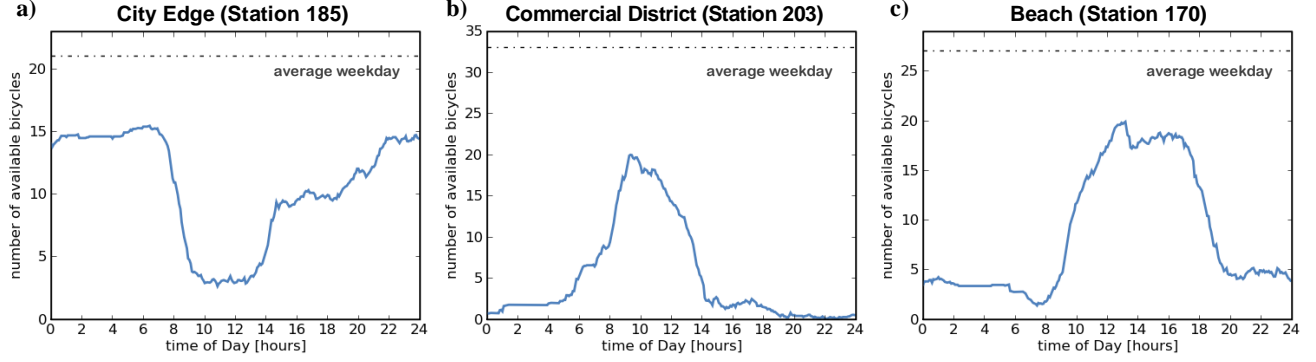
Figure 4. Average number of available bicycles at stations (a) 185 Olzinelles, (b) 203 Diagonal, and (c) 170 Litoral.

B. Spatial Patterns: Sensing the Flow of the City

The spatial layout of a city has an obvious influence on the movement patterns and social behaviors found therein. Barcelona has a mixture of residential, commercial, and recreational areas connected via narrow streets, one-way avenues and a multitude of public transportation options. In Figure 4, we explore the interrelation between a station's location and its underlying temporal usage pattern. What is interesting here is not just that the temporal patterns differ from place to place but what these differences seem to reveal about the *type* of place. For example, station 185 (Figure 4a) is on the city edge and station 203 (Figure 4b) is in a commercial district along Avinguda Diagonal, a major arterial road (these stations are marked in Figure 1d). Their weekday patterns are near opposites: on the city edge, bicycles are checked out starting around 7AM just as bicycles begin arriving in the commercial district. These usage patterns are an indication of the surrounding locale, e.g., the city edge is more residential: people take bicycles in the morning and return them in the evening as they commute to and from work.

Although not as pronounced, the subtle differences between station 203 and station 170 (Figure 4c) reflect a fundamental difference in the reasons why people travel. Whereas stations 185 and 203 are "commuter stations," station 170 is located at the beach. Thus, the 7AM rush of activity does not occur and the rise in incoming bicycles is less pronounced than in the commuter stations, as people casually arrive without the pressures of an explicit work schedule. Although the weekend figures are not shown above, we also observed that while the city edge and commercial stations have considerably different weekday patterns from weekend patterns, the beach station's usage pattern remains relatively constant throughout the whole week.

The "one hump" pattern depicted in both Figure 4b and Figure 4c is one of approximately eight common temporal patterns in our dataset. It is the simplest representation of an "incoming flow" station, often positioned in an area that attracts people during the day. In the next section we report a preliminary analysis that begins to tease out how different types of usage patterns relate across stations.

C. Understanding the Role of Time and Space

In previous sections, we observed the influence of daily routines, culture and location on a station's usage pattern. The usage of Bicing involves a multitude of underlying motivating factors such as commuting, shopping and going to eat. Clustering allows us to measure the relative importance of these effects. By placing the clustering results on a map, we can begin to see the interrelationship between activity and space. This allows us to explore questions like "Do co-located stations share the same usage pattern?" and "How are these usage patterns distributed in the city?"

For clustering, we used the Expectation-Maximization (EM) algorithm found in the Weka toolkit [4]. To compute our clustering features, we split each weekday into seven time bins: early morning (7-9AM), mid-morning (9AM-1PM), lunch (1-3PM), afternoon (3PM-5PM), early evening (5PM-8PM), late evening (8PM-12AM) and night (12AM-5AM). In addition, we considered the entire 24 hour period and only the "open" period (from 5AM to 12AM). For each time bin, we calculated the average number of available bicycles, the difference between the number of bicycles at the beginning and end of the bin's edge, and a measure of the station's activity defined as the percentage of change in the features with respect to the features observed 5 minutes earlier (i.e. the
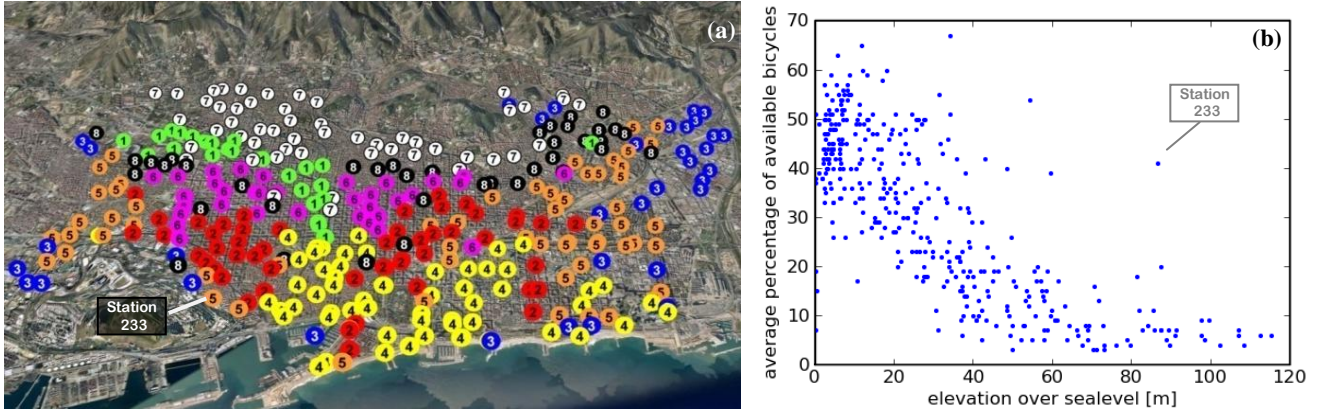
Figure 5. (a) The clustering results; (b) A scatter plot of station elevation vs. average number of available bicycles.

previous observation). Note that EM requires an expected number of clusters as its input. We implemented a visual analysis tool to inspect the temporal patterns of each station within their geo-spatial context. With this tool, we determined eight as the ideal number of clusters.

Figure 5a displays the clusters that contain stations with similar temporal patterns. Close stations tend to behave similarly. However, this is not *always* the case. Cluster 3 (the blue cluster) is clearly the "city edge" cluster; note the number of blue nodes that surround the east, west and southern edges of the Bicing system. Cluster 1 (green) is a "commercial district" cluster, tracing the outline of the northern part of Avinguda Diagonal. Cluster 7 (white) is scattered in the upper part of the map covering nearly the full extent of the city. This cluster represents the highest elevation in the Bicing system and is likely a consequence of a preference to bicycle downhill as Barcelona rests on a slight, ever-increasing slope. Figure 5b reinforces this dynamic by showing how stations with high elevation very rarely have available bicycles. Station 233, marked above, is an interesting exception. Although it is located 86 meters above sea level, its usage pattern is distinct from Cluster 7. This shows that although elevation is a dominant factor affecting usage, other factors may overcome this dominance. In this case, Station 233 is close to both a major park called Montjuïc and residential housing.

IV. CONCLUSION AND FUTURE WORK

In this paper, we have introduced the notion of using shared bicycling's digital footprints to gain an understanding of human behavior and city dynamics. Although our dataset does not contain details about individual movement, our results show how Bicing footprints expose the underlying daily routines and patterns of Barcelona's citizens. We are currently working on incorporating data from other sources of urban infrastructure including cellular networks and automobile parking sensors to investigate how they might augment our analysis and provide different insights into human behavior. We are also interested in building internet- and mobile-based applications that suggest nearby stations based on predictive models. In addition, we have begun logging eighteen additional shared bicycling programs including Paris, Auckland and Washington D.C. We are planning a large-scale analysis that compares behavioral patterns across cities. Some of our results have obvious implications for the design and operation of the shared bicycling system itself. For Bicing, we have met with the Barcelona city government to discuss our findings and continue to correspond with them about our progress.

REFERENCES

[1] Girardin, F., Calabrese, F., Dal Fiore, F., Ratti, C., Blat, J. (2008). Digital footprinting: uncovering the presence and movements of tourists from user-generated content. *IEEE Pervasive Computing*.

[2] González, M. C., Hidalgo, C. A., and Barabási, A. L. (2008). Understanding individual human mobility patterns. Nature 453, 779-782.

[3] Ratti, C., Pulselli, R. M., Williams, S., and Frenchman, D. (2006). Mobile landscapes: Using location data from cell-phones for urban analysis. Environment and Planning B: Planning and Design, 33(5). 727 – 748.

[4] Witten, I. H. and Frank, E. (2005). Data Mining: Practical machine learning tools and techniques. 2nd Edition, Morgan Kaufmann, San Francisco.