Sensing and Predicting the Pulse of the City through Shared Bicycling

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

[1]Computer Science and Engineering
University of Washington
jfroehli@cs.washington.edu

[2]Telefonica Research, Barcelona, Spain
{joachim, nuriao}@tid.es

Abstract

City-wide urban infrastructures are increasingly reliant on network 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 focus on the *digital footprints* from one type of emerging urban infrastructure: shared bicycling systems. We provide a spatiotemporal analysis of 13 weeks of bicycle station usage from Barcelona's shared bicycling system, called Bicing. We apply clustering techniques to identify shared behaviors across stations and show how these behaviors relate to location, neighborhood, and time of day. We then compare experimental results from four predictive models of near-term station usage. Finally, we analyze the impact of factors such as time of day and station activity in the prediction capabilities of the algorithms.

1 Introduction

Observing and modeling human movement in urban environments is central to traffic forecasting, understanding the spread of biological viruses, designing location-based services, and improving urban infrastructure. However, little has changed since Whyte (1980) observed in his "Street Life Project" that the *actual* usage of New York's streets and squares clashed with the original ideas of architects and city planners. A key difficulty faced by urban planners, virologists, and social scientists is that obtaining large, real-world observational data of human movement is challenging and costly (Brockman et al., 2006).

As websites have evolved to offer geo-located services, new sources of real-world behavioral data have begun to emerge. For example, Rattenbury *et al.* (2007) and Girardin *et al.* (2008) used geo-tagging patterns of photographs in Flickr to automatically detect interesting real-world events and draw conclusions about the flow of tourists in a city. In addition, as city-wide urban infrastructures such as buses, subways, public utilities, and roads become digitized, other sources of real-world datasets that can be implicitly sensed are emerging. Ratti *et al.* (2006) and González *et al.* (2008) used cellular network data to study city dynamics and human mobility. McNamara *et al.* (2008) used data collected from an RFID-enabled subway system to predict co-location patterns amongst mass transit users. Such sources of data are ever-expanding and offer large, under-explored datasets of physically-based interactions with the real world.

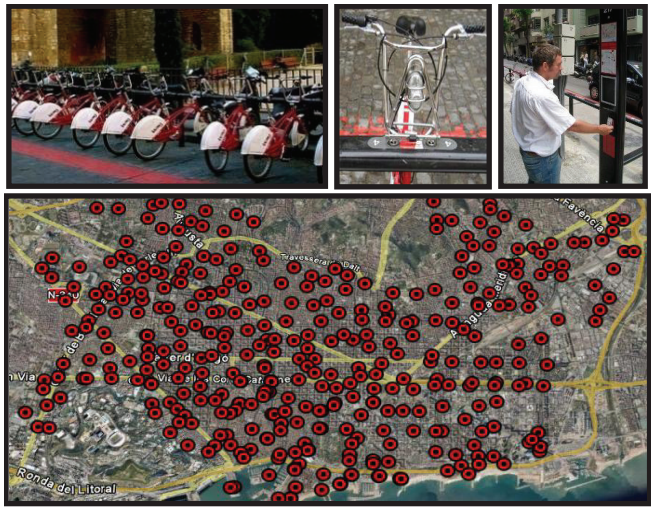

Figure 1. (top) A Bicing station; a close-up of a bicycle and parking slot; and a user at a station kiosk using RFID to check-out a bicycle; (bottom) The 390 Bicing stations distributed across the city of Barcelona, Spain.

In this paper, we introduce a novel source of real-world human behavioral data from a new type of urban infrastructure: shared bicycling systems. We show how station usage data from Barcelona's Bicing system (Figure 1) can be used to infer cultural and geographic aspects of the city and predict future bicycling station usage behavior, which corresponds to human movement in the city.

In particular, the main contributions of this paper are: (1) demonstrating the potential of using shared bicycling as a data source to gain insights into city dynamics and aggregated human behavior; (2) exploring the relationship between spatiotemporal patterns of bicycle usage and underlying city behavior and geography; and (3) studying patterns in bicycle station usage, including the prediction of usage patterns and an analysis of how factors such as the time of the day affect this prediction. In our analysis, we emphasize not just what the bicycling station usage data

reveals about shared bicycling patterns, but also how these patterns reflect the culture and the spatial layout of the city.

We believe this work not only has direct implications for the design and operation of future shared bicycling systems (*e.g.*, more precise load balancing and station dimensioning, improved online services for Bicing users), but also for urban planning (*e.g.*, for blueprinting new bicycle paths and roads), traffic forecasting (*e.g.,* understanding where and when people are in the city), the social sciences (*e.g.,* studying *how* people move about a city, see Latour, 2007) and the development of novel context-based mobile services. In addition, we expect that similar types of analyses can be applied to other sources of urban digital traces, such as those provided by subways (*e.g.,* London's Oyster), buses (*e.g.*, Seattle's Orca program), parking management (*e.g.*, San Francisco's SFpark), and cellular networks (González *et al.*, 2008). Our work thus emphasizes the increasing role that machine learning (ML) and pattern recognition techniques will play to assist the aforementioned fields in analyzing traces of human behavior.

2 The Bicing Dataset

Community shared bicycling programs offer an environmentally friendly, healthy, and inexpensive alternative to automobile transportation. Recent technological advances have led to a third generation shared bicycling system whose real-time usage data can be collected, archived, and analyzed. Currently, there are over forty such programs in the world including SmartBikeDC in Washington D.C. and Vélib' in Paris, which has 20,000 bicycles and 1,450 stations (approximately 1 station every 300 meters). Barcelona's shared bicycle program, Bicing, was launched in March of 2007. It currently has 390 stations with 6,000 bicycles and over 150,000 yearly subscribers.

Bicycles are checked out by swiping an RFID membership card at a Bicing station kiosk (Figure 1), which then unlocks a bicycle and displays its rack location on an LCD screen. Check-out information is uploaded to a web server that provides real-time information about the number of available bicycles and vacant slots at each station. A check-out provides 30 min of free ride time, and every 30 min 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 12AM on Sunday through Thursday and 24 hours during the weekend. The Bicing website reports the status of all bicycle stations via a Google Maps visualization[1]. We scrape this webpage every two 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. Our dataset contains 13 weeks of contiguous observations from Aug. 27, 2008 to Dec.1, 2008. In total, we have collected 26.1 million observations from 390 stations.

In this paper, we focus primarily on weekday data (Monday to Friday, excluding holidays), because weekdays typically correspond to more regular shared bicycling usage patterns.

2.1 Definitions and Notation

In this subsection, we define the terms, notation and intermediary processes used in our analysis.

Station size: Station sizes are not directly reported by the Bicing website, but can be calculated as the sum of available bicycles B_t and free parking slots S_t at time t at each station. Note that this sum fluctuates over time, likely due to a combination of temporarily broken bicycles or parking slots, station growth as the Bicing system evolves, and/or stations reporting invalid numbers while being serviced. Consequently, we infer station size as the 95th percentile of all observed B_t+S_t values for a specified time window for each station. The average inferred station size is 25.2 slots.

Observation normalization: Stations range in size from 15 to 36 slots. Therefore, a station's data is normalized by dividing each observation by the inferred station size. In particular, analyses presented in this paper will be carried out using normalized available bicycles (*NAB*), unless specified otherwise. Note that *NAB* is analogous to the station's percentage full.

DayView: A DayView is calculated by averaging station data that matches certain criteria into a 24 hour window, discretized into five-minute bins (288 bins/day). For example, a station's "*NAB* Weekday DayView" is created by computing the station's average *NAB* values in each of the 288 bins for each weekday over the 13-week observation period. Example weekday and weekend DayViews are shown in Figure 2 for a single station.

Activity and Event Score: The Activity Score (AS) is a measure of how active a station is at a given time: $AS(t)=|B_t - B_{t-1}|$, where B_t is the number of bicycles at time t. The Event Score (ES) is a binary version of AS: $ES(t)=1$ if $AS(t)>0$, otherwise $ES(t)=0$. We typically compute average Activity (AS_{AVG}) and Event Scores (ES_{AVG}) over a specified period of time to measure station activity.

Distance metric: In order to compute the distance between two DayViews or between a single day's data and a station's DayView, we use a Dynamic Time Warping (DTW) based metric with a one-hour Sakoe-Chiba band (Sakoe, 1978). DTW was used in preference to a Euclidean-distance measure because we were interested in comparing overall temporal patterns and wanted to allow for up to one hour of temporal shifts in the data.

2.2 Data Cleansing

The data scraped from the Bicing website is noisy as a result of temporary station closures, technical issues at the stations caused by maintenance work, internet connectivity failures, server-side SQL time-outs, and broken bicycles and parking slots. We employed a three step process to detect and eliminate these faulty observations. Data cleansing was critical to ensure that the data used to train our prediction models was valid.

[1] http://www.bicing.com/localizaciones/localizaciones.php

	Raw Dataset	Cleaned Dataset
Stations	390	370
Days	25K	22.7K
Observations	26.1M	20.2M
Parking Slots	9831	9315

Table 1. Size of the raw and cleaned dataset

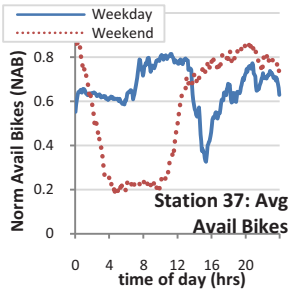

Figure 2. Normalized weekday and weekend DayViews of available bicycles at Station 37.

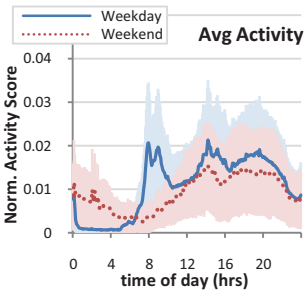

Figure 3. Normalized weekday and weekend DayViews of activity scores for all stations.

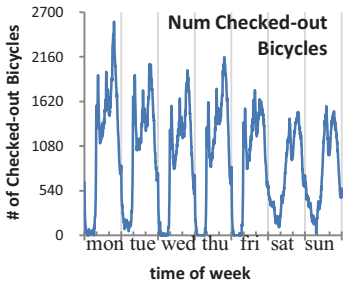

Figure 4. The total number of bicycles checked-out from all stations over a week.

Observation Removal. We eliminated approximately 10% of the 26.1 million individual observations collected from the Bicing website due to: (a) a fluctuating $B_t + S_t$; (b) unattainable values, where $B_t + S_t$ exceeded the station's size; and (c) unusually low station activity, where $ES < 0.1 * ES_{AVG}$ within a twelve hour sliding time window.

Day Removal. We then eliminated 13.7% of all observed days because: (a) there were too few observations for the day (less than 70% of the 720 possible observations/day); or (b) the day exhibited extremely anomalous behavior, *i.e.,* the DTW distance between the day and the station's DayView was larger than two standard deviations from the station's average DTW distance computed across all days.

Station Removal. Finally, we removed all stations that had less than 30 weekdays (out of a total of 64) of *good* data, according to the previously described criteria. This step reduced the number of stations from 390 to 370 (5.1%).

At the end of this process, a total of 20.2 million observations remained, as depicted in Table 1.

3 Temporal and Spatiotemporal Patterns

Before exploring the predictability of individual station usage, we discuss temporal and spatiotemporal patterns and highlight how these patterns reflect the underlying cultural and spatial characteristics of Barcelona.

3.1 Temporal Patterns

Figure 3 compares DayViews of normalized weekday and weekend Activity Scores (AS) for all stations. The shaded background in Figure 3 plots the standard deviation of the scores within each bin. Figure 4 shows the total number of bicycles checked-out from all stations over a week, averaged across the 13-week observation period. Both figures reveal a repeating three-pronged spike in station activity during the weekday, which corresponds to the morning, lunch, and evening commutes. The "lunch spike," which appears across all days, occurs at 2PM, reflecting the late lunch culture of Spain. As one might expect, the morning commute is absent in the two weekend days, resulting in a two-pronged spike. Fridays are the least active weekday, most likely due to reduced working hours in many offices in Barcelona, Mondays are the most active, which is also the most congested vehicle traffic day. The fact that people tend to use Bicing more during the work week reflects results of an online survey[2] we conducted earlier this year with 212 respondents, in which 75% of Bicing users stated commuting as a motivation to sign up for membership.

3.2 Spatiotemporal Patterns

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 and topographic features. To investigate how Bicing usage patterns are shared across stations and geographically distributed in the city, we used a hierarchical clustering technique called *dendrogram clustering* (Duda, 2000) over each station's DayViews.

We built two sets of clusters: one based on weekday Activity Score DayViews ("Activity Clusters," Figure 5) and the other on weekday Available Bicycle DayViews ("Bicycle Clusters," Figure 6). In both cases, a normalized weekday DayView representation was created for each station and a similarity matrix constructed to store the DTW distance between each cluster. The clustering algorithm began with 390 clusters (one for each station). Cluster-to-cluster similarity was calculated as the average of all DTW distances between the weekday DayViews of each station within the cluster. At each clustering iteration, the two most similar clusters were grouped together. Clustering terminated when the average intercluster-to-intracluster distance was maximized with a weight applied to decrease the total number of clusters. Note that the clustering algorithm had no knowledge about the station's geo-location.

Figures 5 and 6 (top) illustrate the clustering results in a geo-visualization of Barcelona and show that neighboring stations generally share similar usage patterns. Each node represents a station location colored according to cluster membership. The line graphs shown in Figure 5 and 6 (bottom) were created by averaging normalized activity/bicycle data across all the stations in each of the clusters. The shading around the lines depicts the standard deviation across stations in each of the 288 DayView bins. The line colors are consistent with the cluster colors used in the map visualization.

[2] https://catalysttools.washington.edu/webq/survey/jfroehli/56481

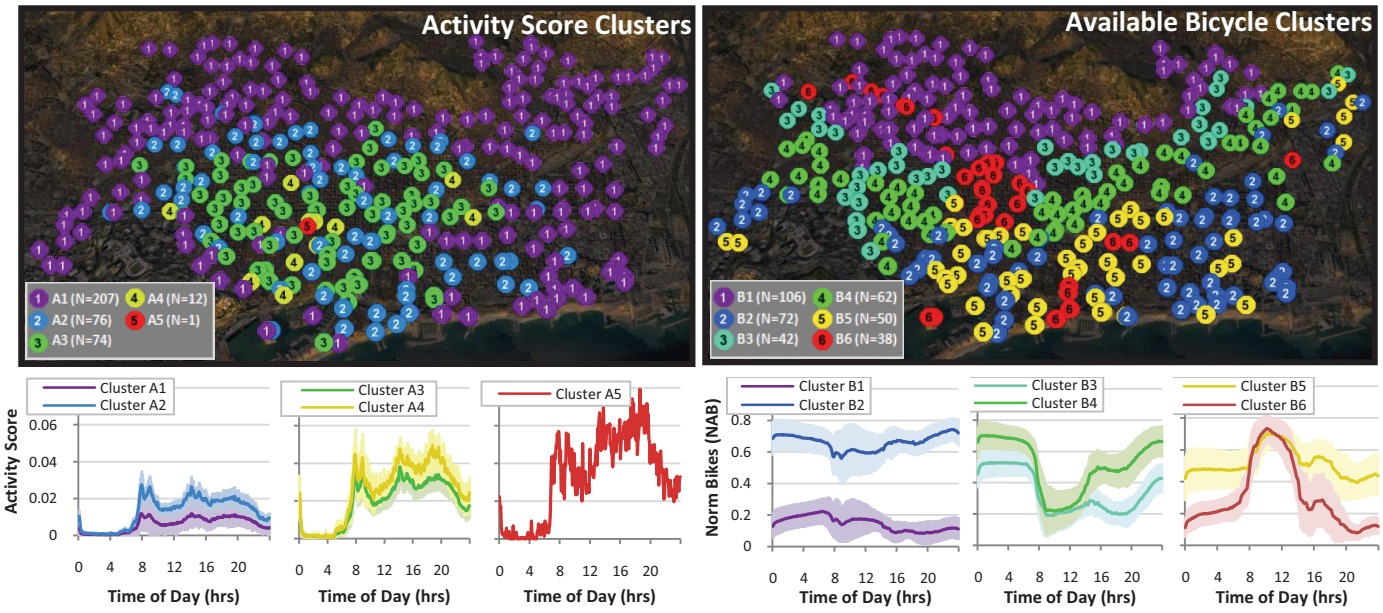

Figure 5. The five Activity Clusters created by progressively combining similar Activity Score station DayViews via dendrogram clustering.

Figure 6. The six Bicycle Clusters created by progressively combining similar available bicycle DayView clusters via dendrogram clustering.

Activity Clusters. The clustering algorithm returned five activity clusters (Figure 5), each with a similar three-pronged spike shape (see also Figure 3). The clusters generally become more active (from A1 to A5) as one moves from the outward edges of the city into downtown. The least active cluster, Cluster A1 (N=207), surrounds nearly the entire perimeter of Barcelona. The more active clusters (A3, A4, and A5) become noticeably more active as the day advances.

Bicycle Clusters. Our algorithm identified six bicycle clusters (depicted in Figure 6) with three classes of behavior: *outgoing* (Clusters B3 and B4), *incoming* (Clusters B5 and B6) and *flat* (Clusters B1 and B2). The *outgoing* clusters show a precipitous drop in available bicycles around 7-8AM as people leave for work, a slight rise at 2-3PM during lunch and a return to early morning levels by 10-11PM. These stations are spread around the edges of the downtown and midtown sections of the city. The *incoming* stations are located in high density commercial areas and along two major arterial routes: Rambla de Catalunya and Avinguda Diagonal. The incoming station shape is nearly inverse of the outgoing stations: people begin arriving around 7-8AM and begin leaving around 1-2PM. Many financial businesses in Barcelona open at 9AM and close around 2:30PM, which aligns well with the temporal patterns of these two clusters. Finally, clusters B1 and B2 have relatively *flat* usage patterns. Cluster B2 tends to have a high degree of available bicycles (on average, it is 66% full) whereas cluster B1 is just the opposite (15% availability). One reason for this discrepancy is likely due to Barcelona's topography: the city itself is built on a long incline. Stations located at the top of Figure 6 are between 80-110 meters above sea level versus those at the bottom, which are at 0-10 meters above sea level. People tend not to bicycle up to the higher altitudes, thus leaving those stations in Cluster B1 starving for bicycles.

4 Station Behavior Prediction

We focus next on the prediction of station usage. In particular, we are interested in predicting the number of available bicycles at each station at a given time in the future. This work is related to traffic forecasting. Most approaches in traffic engineering rely on flow theory and incorporate queue-theoretic models (Vandaele *et al.*, 2000). Horvitz *et al.* (2005) took an alternative approach that is closely related to ours: they successfully modeled key traffic bottleneck areas using a Bayesian network and ignored the underlying flows. Similarly, we do not attempt to model individual bicycle movements in the city but rather focus on modeling Bicing station usage directly.

Predictive station usage models would (a) allow for more accurate load balancing of the stations; (b) assist urban planners and city officials by providing them with information about *expected* activity in the city; and (c) open the way to new mobile services for Bicing users. In the previously mentioned online survey, Bicing users identified *finding an available bicycle and parking slot* as the two most important problems in their Bicing experience (76% and 66% of respondents, respectively). Therefore, we are also interested in predicting the probability of finding a free bicycle and slot in a station at a given time in the future. Furthermore, our models shed light on some of the factors that influence the *predictability* of station usage behavior.

4.1 Models of Station Behavior

We have implemented four simple predictive models, including a Bayesian network (BN) to predict the availability of bicycles at each station. All models have three input parameters: (1) the current time t_0; (2) the last

known number of bicycles at time t_0 (B_{t0}); (3) and a prediction window (PW) that specifies how far into the future to predict. In this paper, we focus on PW values ranging from 10 min to 120 min, as they correspond to the most common Bicing usage scenario. Models that use historic data have a fourth input parameter: all previous observations up to time t_0.

Last Value (LV). This model predicts that the current number of available bicycles remains constant at its last measured value of B_{t0} for all PW:

$$\boldsymbol{Pred_{LV}(t_0, B_{t_0}, PW) = B_{t_0}}$$

Historic Mean (HM). This model dynamically constructs a DayView based on all observations up to t_0 , and returns the value at the corresponding DayView's time bin ($t_0 + PW$). More formally: Let the average number of bicycles in time bin TB_{t0} within a DayView be denoted as $\bar{B}_{TBt0}$. The HM predictor forecasts the number of bicycles at time $t_0 + PW$:

$$\boldsymbol{Pred_{HM}(t_0, B_{t_0}, PW) = \bar{B}_{TB_{t_0+PW}}}$$

In other words, this predictor returns the average number of bicycles in the time bin $t_0 + PW$.

Historic Trend (HT). This model builds upon the LV and HM predictors. The historic trend is extracted by computing the difference in the average number of bicycles at t_0 and $t_0 + PW$, and adding it to the current number of bicycles, B_{t0}:

$$\boldsymbol{Pred_{HT}(t_0, B_{t_0}, PW) = B_{t_0} + \bar{B}_{TB_{t_0+PW}} - \bar{B}_{TB_{t_0}}}$$

Bayesian Network (BN). This model consists of a simple BN per station with three observed (input) nodes and one hidden (output) node, where all observed nodes are the parents of the hidden or output node (naïve Bayes), as depicted in Figure 7.

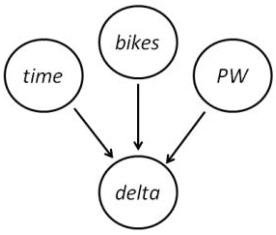

Figure 7. Graph of the BN

The nodes of the network are: (1) *time*, a discrete observed node with 24 possible values corresponding to each of the hours of the day; (2) *bikes*, the value of *NAB* at time *t*, discretized into five bins, where a value of one corresponds to 0-20% of bicycles, two corresponds to 20-40%, etc; (3) *PW,* the size of the prediction window, with six possible values corresponding to 10, 20, 30, 60, 90 and 120 minutes into the future; and (4) *delta*, a continuous Gaussian variable that contains the change in the number of bikes at time *t* + *PW* with respect to its value at time *t*. Predictions in the number of bicycles are made by adding the value of the *delta* node to the most recent observation:

$$\boldsymbol{Pred_{BN}(t_0, B_{t_0}, PW) = B_{t_0} + delta}$$

The dimensionality of the values in the three observed discrete nodes will have an effect on the accuracy of the BN. We carried out extensive experiments with two pilot stations to determine the optimal dimensionality of the discrete nodes. The above configuration had the smallest prediction error and therefore has been used for all stations.

The three history-based predictors used three weeks of historic data, from Nov. 2, 2008 to Nov. 23, 2008 (see Figure 10) to build their models. The HM and HT predictors used the corresponding DayViews built from the same three-week period. Each BN was trained by computing a posterior over the parameters given the fully observed data from *time*, *bikes*, *PW* and *delta,* with $time$ covering the three-week period of training data, in five minute increments. We used the Matlab BN toolkit by Kevin Murphy[3] to learn the parameters of the Bayesian network ("learn_params") and to compute the ML estimation of the unobserved delta note (in the toolbox labeled as "marginal_nodes" of a "jtree_inf_engine").

4.2 Prediction Evaluation

We used the *next* five weekdays following our training data (*i.e.,* from Nov. 24 to Nov. 28, 2008) to evaluate the models. To simulate real-world conditions, we did not remove unusual days from the test data. Starting at midnight on Nov 24th, the prediction models were fed the current time (in five minute increments), the current number of available bicycles, and each of the six previously mentioned *PW* values. For each *PW,* our models returned the predicted number of available bicycles.

The *prediction error* was computed as the absolute difference between the predicted number of bicycles and the ground truth observation at time $t_0 + PW$. The error was normalized by the station's size (thus, given in *NAB* units). For completeness, we also compared our results to a *Random predictor* (Rand), which returns a random value drawn from a uniform distribution between zero and the station's size.

Table 2 lists the prediction error for each of the models, averaged over all stations, all days and all values of *PW*. The BN predictor had the smallest average error of 0.08 NAB. At a station with 25 slots, this corresponds to an average error of two bicycles (8%).

Interestingly, the HM predictor performed worst of all - excluding the random model, implying that a station's daily activity is quite varied when compared to its historic mean. This highlights the importance of giving more weight to the most recent observations.

Full-Empty Station State Prediction. From the perspective of both Bicing users and managers, a more useful measure is the accuracy in predicting if the station is *empty* (*i.e.*, relevant when someone wants to pick up a bicycle) or if the station is *full* (*i.e.*, someone wants to drop off a bicycle). In both cases, the station *state* can easily be derived from the number of available bicycles as predicted by our models.

[3] http://www.cs.ubc.ca/~murphyk/Software/bnsoft.html

Model	Avg Error	Stdev of Error
Rand	0.37	0.27
HM	0.17	0.16
LV	0.09	0.14
HT	0.09	0.13
BN	**0.08**	**0.12**

Table 2. Average prediction error of the four models and the random predictor.

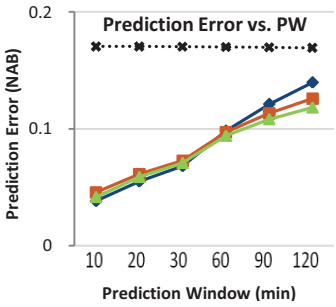

Figure 8. Prediction error as a function of the size of the prediction window

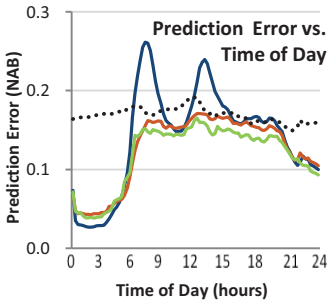

Figure 9. Prediction error as a function of the time-of-day using a two-hour prediction window

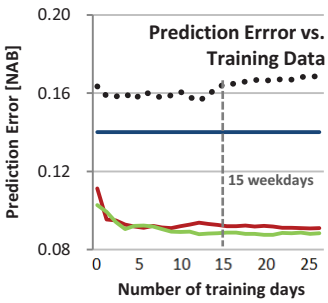

Figure 10. Prediction error as a function of days of training data from one stations per cluster at *PW=120*

In Table 3 below, we compare the resulting classification errors applying the above models with a PW=*120min*, which is the most challenging scenario. Note that in the analysis below, we: (1) exclude the *Rand*, HM and LV models, as they perform poorly for *PW=120 min*; (2) include the classification error of two additional binary classifiers that have been trained to predict the station states: a decision tree classifier (ID3) and a support vector machine classifier (SVM) with a dot product kernel. Both classifiers have been trained and evaluated with observation triplets that consist of the time of the day (in minutes), the *percentage full* of the station (*i.e.* a number between 0% and 100% specifying how full a station is with respect to its capacity) and the state of the station (full/not full; and empty/not empty) at *PW=120 min*. Note that we train one classifier per binary decision: full/not full and empty/not empty.

Method	Prediction error for "station is empty"	Prediction error for "station is full"
BN	**13.3%**	**12.0%**
HT	13.7%	12.0%
ID3	13.3%	12.5%
SVM	14.7%	13.8%

Table 3. Comparison of the error of four classifiers at *PW=120 min*. The predicted number of bicycles of the HT and BN models were used to obtain the predicted station state. In contrast, the ID3 and SVM classifiers are directly trained and evaluated with the station state.

Interestingly, the four classifiers listed in Table 3 show very similar classification errors, despite using different input variables and optimizing different cost functions (generative *vs.* discriminative approaches). The similarity between prediction errors may indicate that the lower bound of station state predictability is ~12%. Data from additional sources (*e.g.,* weather, local event information, train schedules) might be needed to overcome this barrier.

4.2 Factors that Impact Predictability

Next, we discuss factors that impact the *predictability* of the stations' usage.

Prediction Window. The average *prediction error* as a function of *PW* is shown in Figure 8. The HM predictor resulted in an average error of 0.17 regardless of the *PW*. If station usage was perfectly consistent across weekdays, this predictor would perform flawlessly. The LV predictor did nearly as well as the best performing predictors for *PW <= 60 min*. Thus, for small *PWs* the number of available bicycles at t_0 is a strong indicator of the number of available bicycles at t_0 + *PW*. For *PWs > 60 min*, however, the HT and BN predictors begin to perform significantly better than the rest: an unpaired t-test showed that the difference in prediction error at *PW=120min* is significant between all four predictors: $t(1002229) = 34\text{-}158$, $p < 0.001$ (Bonferroni's adjustment for comparing five models applied). A mobile Bicing user interested in predicted station vacancies would likely be most interested in *PW <= 60 min,* where the LV, HT, and BN models are able to predict the number of available bicycles to within a single bicycle.

Time of the Day. Figure 9 shows the accuracy of the four predictors as a function of time of day. As expected, all predictors perform well during the night, when there is little activity, and become less accurate during the day. Although all predictors performed worse during the most active times of day, the BN was the most resilient.

Amount of Historic Data. Three of the station usage models required previous observations to build their predictions. Given that our models are data driven, an open question is how much data is needed to build accurate models. Figure 10 shows the average prediction error for six stations (one per Bicycle cluster) at *PW=120min*. The more observations used, the lower the prediction error of the HT and BN predictors. After 15 weekdays of training data, the average prediction error plateaus, showing a 22% (HT) and 26% (BN) improvement over predictors trained with only a single day. Note the error of the HM predictor actually *increases* with more observations, corroborating the intuition that more recent data should be weighted higher.

Station Clusters. To better understand how station usage patterns affect predictability, we grouped our prediction results by the Activity and Bicycle Clusters presented in Section 3. In general, the more active stations are more difficult to predict (Figure 11). The BN is *significantly* better in the more active clusters than other predictors. The *flat* clusters (B1 and B2) result in the lowest overall prediction error as there is minimal variation in bicycle availability over time (Figure 12). There is no significant difference between *incoming* and *outgoing* stations, probably due to having similar, yet reversed, dynamics in their behavior.

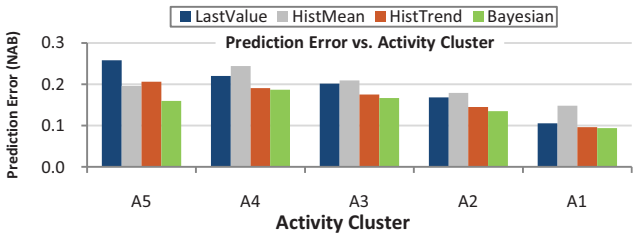

Figure 11. Average prediction errors broken down by activity cluster. Note how active stations have higher prediction errors.

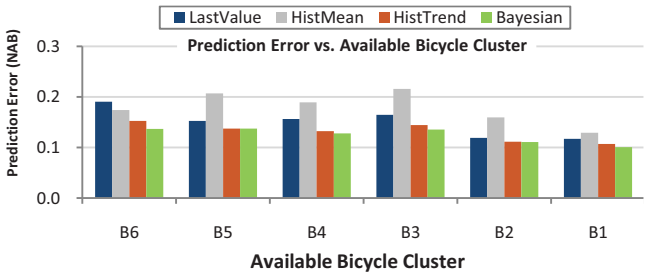

Figure 12. Average prediction errors broken down by bicycle cluster. Note how the BN error is the least sensitive to the type of cluster.

5 Conclusions and Future Work

As urban infrastructures are increasingly digitized, human behavior data will become ubiquitous. Pattern recognition and ML techniques will be necessary to sort through and analyze the large amounts of real-world behavioral data being produced. We have shown how shared bicycling station usage data can reveal not just patterns of bicycling usage, but also the underlying temporal and spatial dynamics of a city. We have also shown that fairly simple predictive models are able to predict station usage with an average error of only two bicycles and can classify station state (full, empty, or in-between) with 80% accuracy up to two hours into the future. Our experiments indicate that 10 to 15 weekdays of historic data are enough to build station models.

Although some of the results presented in this paper align with our intuition of human behavior patterns in Barcelona, other details were surprising. For example, shared temporal trends in station usage can allow us to infer attributes about neighborhoods (*e.g.*, residential vs. commercial, proximity to downtown). Station usage is also far less regular than initially suspected, as revealed by the poor performance of the HM predictor. Although shared bicycling station usage is only one source of human movement in the city, it has the potential to uncover underlying temporal and spatial dynamics of the city, as shown in this paper. We believe similar analyses can be expanded to and combined with other digitized urban infrastructures such as roadways, buses, and subways.

In the future, we would like to incorporate contextual features into our BN, such as weather, season, special events (*e.g.,* concerts or soccer matches), public transportation schedules and locations, and data from additional urban infrastructure (*e.g.*, cellular networks). We also plan to experiment with dynamic BNs and continuous dynamic models. We are currently working to leverage our prediction results as a backend for mobile and web applications that would provide real-time information about the probability of finding a bicycle or a free parking slot at a station. Finally, we are logging 18 additional shared bicycling programs including Paris and Washington D.C. and are planning a large-scale analysis that compares behavioral patterns across cities.

Acknowledgements

We thank Xavier Anguera, Mauro Cherubini, Leah Findlater, James Fogarty, Aaron Hertzman, Raphael Hoffman, John Krumm, Rodrigo Oliveira, Shwetak Patel, Scott Saponas, Mike Toomim for their helpful comments on early drafts of this paper. This work is sponsored by Telefonica Research and Jon Froehlich's Microsoft Research Fellowship.

References

[Brockmann *et al.,* 2006] Brockmann, D. D., Hufnagel, L. & Geisel, T. (2006). The scaling laws of human travel. Nature 439, pp. 462–465

[Duda *et al.,* 2000] Duda, R., Hart, P., Stork, D. (2000) *Pattern Classification (2nd Edition)*. Wiley, New York.

[Girardin *et al.,* 2008] Girardin, F., Calabrese, F., Dal Fiore, F. , Ratti, C., and Blat, J. (2008). Digital footprinting: Uncovering tourists with user-generated content. *IEEE Pervasive Computing*, 7(4)

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

[Horvitz *et al.,* 2005] Horvitz, E., Apacible, J., Sarin, R. and Liao, L. Prediction, expectation, and surprise: Methods, designs, and study of a deployed traffic forecasting service. *Proceed. IUI '05*, July 2005.

[Latour, 2007] Latour, B. Beware, your imagination leaves digital traces, *Column for Times Higher Education Supplement*, 6th of April 2007. http://www.brunolatour.fr/poparticles/poparticle/P-129-THES-GB.doc

[McNamara *et al.,* 2008] McNamara, L., Mascolo, C., & Capra, L. Media sharing based on colocation prediction in urban transport. *Proceed. ACM MobiCom '08.* San Francisco, CA, Sept. 14 - 19, 2008, pp. 58-69.

[Rattenbury *et al.,* 2007] Rattenbury, T., Good, N., & Naaman, M. Towards automatic extraction of event and place semantics from flickr tags. *Proceed. ACM SIGIR '07.* Amsterdam, July 23-27, 2007, pp. 103-110.

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

[Sakoe & Chiba, 1978] Sakoe, H. & Chiba, S. (1978). Dynamic programming algorithm optimization for spoken word recognition. *IEEE Trans. on Acoustics, Speech & Sig Proc,* V. ASSP-26. pp. 43-49.

[Vandaele *et al.,* 2000] Vandaele, N., Van Woensel, T., & Verbruggen, N. (2000). A Queuing-Based Traffic Flow Model. *Transportation Research-D: Transportation and Environment*, 5(2), pp. 121-135.

[Whyte, 1980] Whyte, W. H. The social life of small urban spaces. *Washington, D.C.: Conservation Foundation*, 1980.