Smart Bike-Sharing Systems for Smart Cities

Gus Lipkin

4 December 2020

Image source:
https://en.wikipedia.org/wiki/Bicycle-sharing_system#/media/File:Melbourne_City_Bikes.JPG

Introduction

As cities grow and more people are pushed to public transit, the so-called “last-mile problem”, transportation from the home and workplace to and from public transportation, needs to be solved. One solution is a bike-share program. Bike shares allow people to pick up a bike near them and take it to their destination where someone else can then do the same. Smart bike share systems use data analysis to monitor stations, collect usage data, and predict future bike use and availability so that bikes can be better distributed throughout the city to match needs.

Literature Review

Bike sharing is a self-service transportation option where users can check out a bike from a docking station then ride to their destination and park the bike in a nearby docking station (Bryce, 2016). Bike share bikes are usually built in the cruiser style which are more comfortable and upright as well as more sturdy than other styles (Bryce, 2016). Bike shares are mostly used by non-cyclists and can have a positive impact on nearby commercial areas (Bryce, 2016).

As the technology changes to meet consumer needs, dockless bikes have been introduced. Companies such as Lime allow users to simply pick up the nearest bike, located through a smartphone app, and ride it to their destination where they can leave it wherever they want (Vuocolo, 2019). These dockless systems are frequently criticized for leading to the cluttering of sidewalks with discarded bikes leading many cities to effectively ban dockless bike sharing systems (Vuocolo, 2019).

In an effort to better understand docked bike sharing system trends, many people have researched and studied bike availability prediction models, clustering to explore data trends, and bike rebalancing routes (Alavi & Buttlar, 2018).

By predicting where bikes are available and how many there will be at each station, redistribution efforts can be better planned and executed (Alavi & Buttlar, 2018). Past models have concentrated efforts on measuring time, weather, the urban environment, nearby infrastructure, and more in predicting availability (Alavi & Buttlar, 2018). Alavi et al suggested that these models could be improved by introducing machine learning algorithms, improving large network predictions, and clustering stations.

Clustering stations is important to docked bike share system research because it shows which areas are in high demand and which docking stations may serve to meet surge demand in an area. Geography and time of day so far have been the most heavily measured data points when attempting to use clustering (Alavi & Buttlar, 2018).

Alavi et al propose using data mining and increase adaptability of the models to improve clustering.

Rebalancing is the most important part of any bike share system. If there are not enough bikes to meet demand at every station, then the system will lose users. Rebalancing also has the most research already as path finding algorithms can be adapted to include criteria on the capacity of the bike distribution truck and the number of bikes at each station (Alavi & Buttlar, 2018). By splitting the rebalancing problem into two phases, repositioning and routing, Alavi et al developed a novel approach to routing which allowed for improved bike rebalancing routes.

Data

In developing their model, Alavi et al gathered data from a San Francisco bike share program between August 2013 and August 2015. Data was collected every minute and included the station ID, the number of docks and bikes available at each station, and a date-time stamp of the data recording time. In order to reduce the dataset size, only data from every fifteen minutes was used. In addition, if there was no data for a station at that time, that entry was eliminated. They also collected information on the weather which included, mean temperature, mean humidity, mean visibility, mean wind speed, precipitation, and the general weather condition for the day.

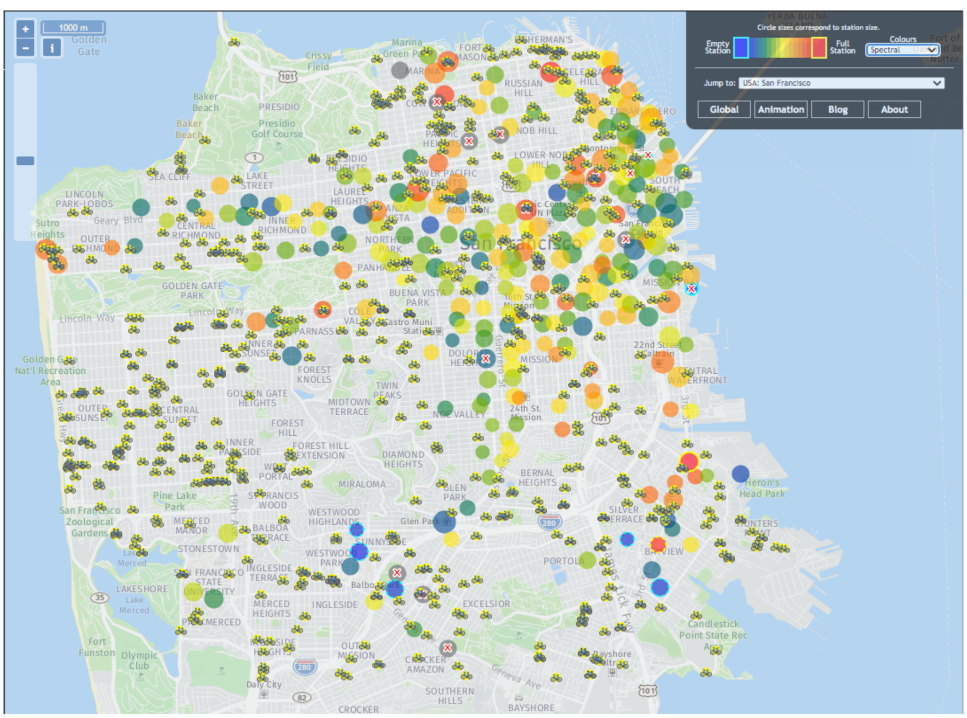

Figure 1: A map of Bay Area Bike Share Stations and Availability as of 4 December 2020 at 13:47 Local Time (*Bike Share Map: San Francisco*, n.d.)

Analysis and Results

When developing their bike prediction models, Alavi et al split the models into univariate and multivariate models, each with their own methods. Univariate models were developed with random forest and least-square boosting algorithms to predict the bikes available at each of the seventy stations.

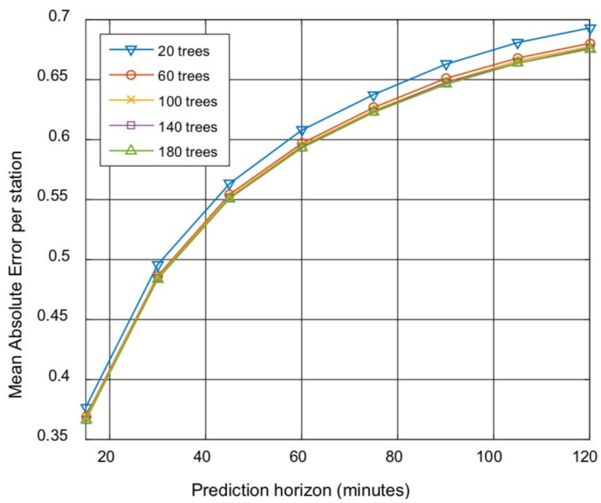

Figure 2: “RF MAE at different prediction horizons and number of trees” (Alavi & Buttlar, 2018)

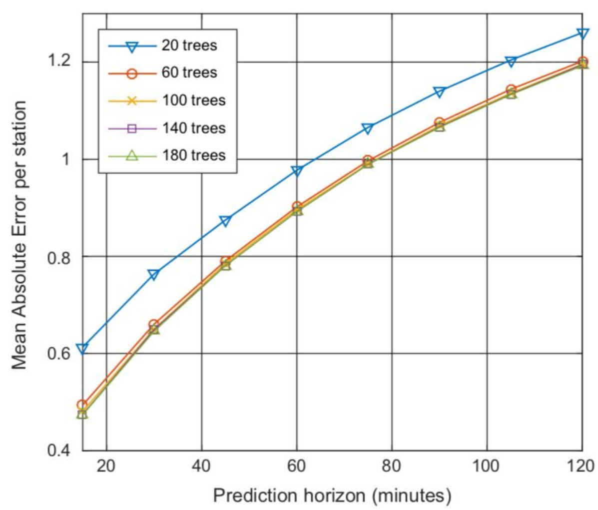

Figure 3: “LSBoost MAE at different prediction horizons and number of trees” (Alavi & Buttlar, 2018)

Figures 2 and 3 show the mean average errors in each the RF and LSBoost models. The RF models had a smaller error at each prediction horizon and as the prediction horizon increased for both models, the error also increased.

For the multivariate models, partial least square regression was used. They found that the network could generally be split into five distinct regions and that trips usually did not leave the region that they originated in. The MAE of the PLSR model was roughly between that of the RF and LSBoost models but increased at a linear rate instead of a logarithmic rate. However, the PLSR is still better for large bike share networks as it takes into account more nuances in the system than the univariate models could.

When clustering, Alavi et al wanted to increase the purity and similarity of the cluster while only considering the number of clusters in the model. They ignored the weather data and clustered based on day of the week and time of day across the entire year.

When rebalancing the bikes, it is important to appropriately address those so-called “NotSpots,” those spots which are either empty or full so that people cannot take a bike or leave a bike, respectively (Goodyear, 2014). Certain threshold values were used in determining Alavi et al's NotSpots as they are effectively the same as a station that is empty or full. The bike rebalancing tour calculation was split into two phases: the first constructed the tour and the second improved and selected the optimal tour from phase one. In phase one, the main goal was to optimize so that each NotSpot was included once and only once. In phase two, the best tours from phase one were optimized and the best route was selected. The final result was an algorithm that performed, on average, .824% worse than best-known routes.

Next Steps

As cities continue to grow and bike sharing increases in demand, it is important to keep the future in mind so that the bike share system of tomorrow can be optimized today. Two such future endeavors may include using existing data to analyze user movement patterns to increase or decrease the number of bays available at each docking station or looking into ways to incentivize travel between at capacity NotSpots and empty NotSpots. Both of these would optimize the bike rebalancing tour that happens at the end of the day and has a chance of boosting business in areas that are frequently empty NotSpots.

Conclusion

Bike shares have three main areas of interest to data scientists: bike prediction models, supervised clustering to determine station similarities, and bike rebalancing tour optimization. Alavi et al expanded the knowledge available in each area and contributed by improving on existing research with machine learning techniques. Their models were either simpler or produced better results than existing methods so that in the future, smart bike share systems will be more adequately able to serve the people in the communities in which they operate.

Bibliography

Alavi, A. H., & Buttlar, W. G. (2018). *Data Analytics for Smart Cities*. https://doi.org/10.1201/9780429434983

Bike Share Map: San Francisco. (n.d.). Retrieved December 4, 2020, from https://bikesharemap.com/sanfrancisco/

Bryce, K. (2016, November 17). *The Many Benefits of Bike Sharing Programs*. https://www.commuteoptions.org/the-many-benefits-of-bike-sharing-programs/

Goodyear, S. (2014, March). *Mapping the Imbalances of New York's Popular, Troubled Bike-Share*. Citylab.com. www.citylab.com/commute/2014/03/mapping-imbalances-new-yorks-popular-troubled-bike-share/8699/

Vuocolo, A. (2019, March 4). *Why Are Bike Share Companies Pivoting to E-Scooters?* Bicycling. https://www.bicycling.com/news/a26623484/lime-bikes-scooters/