International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

COPY RIGHT

IJIEMR Transactions, online available on 6th Mar 2018. Link :http://www.ijiemr.org/downloads.php?vol=Volume-08&issue=ISSUE-03

Title: **ANALYZING BIKE SHARING SYSTEM BASED ON USER BEHAVIOR**

Volume 08, Issue 03, Pages: 7–12.

Paper Authors

MR.R.VEERABABU, CH.SWETHA HARSHINI

Vignan's Lara Institute of Technology & Science

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

ANALYZING BIKE SHARING SYSTEM BASED ON USER BEHAVIOR

MR.R.VEERABABU[1], CH.SWETHA HARSHINI[2]
Assistant Professor[1], Department of M.C.A ,Vignan's Lara Institute of Technology & Science
M.C.A Student[2],Department of M.C.A ,Vignan's Lara Institute of Technology & Science

Abstract:
The rapid development of bike-sharing systems has brought people enormous convenience during the past decade. On the other hand, high transport flexibility gives rise to problems for both users and operators. For users, dynamic distribution of shared bikes caused by uneven user demand often leads to the check in or check out service unavailable at some stations. For operators, unbalanced bike usage comes with more bike broken and growing maintenance cost. In this paper, we consider to enhance user experiences and rebalance bicycle utilization by directing users to different stations with a higher success rate of rental and return. For the first time, we devise a trip advisor that recommends bike check-in and check-out stations with joint consideration of service quality and bicycle utilization. To ensure service quality, we firstly predict the user demand of each station to obtain the success rate of rental and return in the future. Experiments indicate that the precision of our method is as much as 0.826, which has raised by 25.9% as compared with that of the historical average method. To rebalance bike usage, from historical data, we identify that biased bike usage is rooted from circumscribed bicycle circulation among few active stations. Therefore, with defined station activeness, we optimize the bike circulation by leading users to shift bikes between highly active stations and inactive ones. We extensively evaluate the performance of our design through real-world datasets. Evaluation results show that the percentage of frequently used bikes decreases by 33:6% on usage number and 28:6% on usage time.

Bike sharing systems

It is important to distinguish between three generations of services. According with several authors there are three generations of services of bike-sharing: free bike system, coin-deposit systems and information technologybased systems (Shaheen, Guzman, & Zhang, 2010), (Wang, Liu, Zhang, & Duan, 2008) and (DeMaio, 2009). The free bike-sharing system is characterized by a set of bicycles (with unusual colors and/or shapes) that are available without costs to the user. Typically the stations are located near public facilities that have their own staff which are responsible for the users' identification, reducing the needs of human resources of the system. The use of the bicycle is, in the most cases, free to the user. The first bike sharing system was emerged in Amsterdam, the Netherlands in 1965. A set of fifty free bicycles was seen as the solution for traffic problems. However the Witte Fietsen (white bikes) Plan failed after its launch due to the bicycle damages and thefts. In the coin-deposit systems the bicycles are not freely available, once the users have to use a coin to unlock the bicycle from the docking stations. At the same time, some concerns about the location of the stations are introduced to ensure the efficiency of the operation. Although some significant changes on the motorized transportation patterns in some cities the coin-deposit system did not solved the thefts problem. To overcome this problem, the third generation of bike-sharing emerged based on automatic services. This generation uses smart technology (mobile phones, mag-stripe cards, smartcards or codes) to unlock the bicycles from the stations allowing the automatic identification of the users (with a code for instance). The casual users pay a security deposit to ensure the return of the bicycle, and the use of the bicycles is paid depending on the time interval of the usage. Typically the service is free in the first specified time interval and the price gradually increases after the interval depletion. This system is simpler to manage in terms of human resources, but requires a higher investment in technology. Some of the great advantages of the technology introduction are the possibility of 24h service, the easier location of stations in the city and the data collection about the usage of the service. Shaheen, Guzman, & Zhang (2010) identified also the fourth generation of bike-sharing systems. Fourth generation bike-sharing systems are multimodal systems. Their main concern is an improvement of the service to the user needs, in other words it is demand-responsive. It includes an improvement in technological mechanisms in the stations and bicycles that facilitate their use and share, electric bicycles, bicycle relocations and the integration of the several transport services in the same access card (public transportation or car-sharing). In Portugal it was implemented a free bicycle sharing system in Aveiro, called Bugas, that was launched on April 2000. It stated with a stock of 350 bicycles spread over 33 parks all over the city. However, after the pilot period some of the bicycles were vandalized or stolen. Currently the system works as a less ambitious service with only one station and some degraded bicycles. The successful of the bike-sharing programs depends on how the demand is satisfied. However, the definition of bike-sharing demand is not yet a popular subject in the literature. Next section provides a literature review about general bicycle demand models and a focus on existing bike-sharing demand definition strategies

Demand studies for cycling and bike-sharing

One of the biggest concerns of the urban transportation planners is to provide the most adequate response to traveller's needs, estimating transportation demand and its variation. Planners are also aware of the strong relation between transportation and land use, and as this relation should be incorporated in demand studies. It is complex and risky to predict the number of bicycle trips, especially in cities where the bicycle is not yet widely used. There are various studies on the prediction of non-motorized travel demand. Turner, Hottenstein, & Shunk (1997) and (Schwartz et al., 1999) present an overview of different approaches to determine the bicycle travel

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

demand. One of the methods more frequently referred is the Latent Demand Score Method (Landis, 1996) and it is specially adapted in cases where bicycles are not yet a popular choice. The methodology provides a coefficient of potential demand for bicycle trips throughout a transportation network (in each arc of the network), based on the influence of generator/attractors points in the city on the number of bicycle trips for all road segments. One of the advantages of this model is that it acts as a geographic information system. However the trips estimated are not directional (the method considers the total number of the trips that were generated and attracted), meaning that the method compromises an Origin-Destiny evaluation. An adaptation of this method was used in a demand study for the city of Tomar (Portugal) where the main difference of this adaptation to the Latent Demand Score Method is that it considers the number of trips in each origin-destination point, and the choice of shortest path between origin and destination (Ribeiro, Frade, & Correia, 2012). The current scientific studies or real world applications use 'revealed' or 'stated' preference surveys as methods for bike sharing systems demand estimation (dell'Olio, Ibeas, & Moura, 2011) (ConBici, 2007) (PROBICI team, 2010). In the cases of bike-sharing systems expansion, the revealed surveys can be very useful; however in some cases the responses to the stated preference surveys can be strategic and may not reflect the real intentions of the interviewee. Surveys results must be used with care, mainly in the cases where similar services were not yet implemented. In order to avoid the constrains caused by the surveys, the demand modelling approach will study different bike-sharing systems around the world defining the profile of the users and potential users, the factors that can influence the demand (as the geographical conditions, the variation of demand during the day or over the seasons, and the travellers characteristics age; sex, and/or job, etc.) and how they affect it. The demand of New York City bike-sharing system was designed using the user group patterns of successful bike-share programs: Velib'in Paris, Velo'v in Lyon and Bicing in Barcelona; from which three typical user groups were identified: commuters, recreational/errand riders and tourists. The authors estimated the number of people in each potential user category in New York and applied to them different uptake rates (3%, 6% and 9%) to quantify the users of bike-share program. The uptake rates are defined based on London and Paris surveys (NYCDCP, 2009). Krykewycz, Puchalsky, Rocks, Bonnette, and Jaskiewicz (2010) use a methodology to estimate the demand for a new bicycle-sharing program in Philadelphia (Pennsylvania). The authors' defined two market areas using raster based geographic information system analysis and applied three bike share trip diversion rates determined through surveys in Lyon, Paris (France) and Barcelona (Spain) in order estimate the modal shift from other modes to bike-sharing, establishing different demand scenarios (low, middle and high). In the Seattle case, the demand study was based in the Philadelphia study. However, the market areas were defined considering a GIS raster dataset of weighted sum indicators that influence bike-share use (population density, non-institutionalized group quarter population density, job density, retail job density, commute trip reduction companies, tourist attractions, parks/recreation areas, topography, regional transit stations, bicycle friendly streets, streets with bicycle lanes and local transit stops). Rates observed in Lyon, Paris and Barcelona, to the defined market areas, were also applied (Gregerson, Hepp-buchanan, Rowe, Sluis, Vander, Wygonik, et al., 2010). Daddio (2012) presents a regression approach to relate the surrounding characteristics with the station demand. The dependent variable is the number of trip departures per station, using the data provided by Capital Bikeshare (bike-sharing system of Washington Metropolitan Area). The independent variables are measure within 400 meter walk distance from each station. The variables considered are divided in three sets of characteristics: trip generation, trip attraction and transportation network. In the District of Columbia, the variables statically significant are the population between the ages of 20 and 39, the proportion of population that belongs to a race other than "white alone", the number of retail establishments selling alcohol, the number of metro stations and the distance from weighted mean (ridership) from the center of full DC and CA Capital Bikeshare system. The use of public bicycles increases potentially when they are complemented with other transportation modes (intermodality), or when parking problems exists in the origin or destination of the trip. In The Netherlands for instance a growth in bicycle use for non-recurrent trips, besides a reduction in car use and a growth in train trips, was observed after the introduction of a public bicycle sharing service, (Martens, 2007). Krizek & Stonebraker (2010) presented a methodology - developed for Puget Sound Regional Council in Washington in 2002 - that determines the total number of potential users of a bicycle station (in different scenarios) depending of the respective user groups, defined as: bicycle commuters who work within a quarter mile of the bicycle station; bicycle users who park their bicycles at transit stations and bicycle users who travel with their bicycles. The methodology relates the number of the users with the employment data, the number of transit trips, the bicycle share within 3 miles of a proposed bicycle station, and the number of bicycle commuters to within a quarter mile of the bicycle station. The validation of this method was done considering the data of two existing bicycle stations and the methodology was considered reasonably accurate.

Existing system:

For stations, the user demand is ever-changing and unbalanced, which often leads to the check in or check out service unavailable at some stations and has a negative impact on user experience. For bikes, the usage frequency of each bike is unevenly distributed, posing a problem for both riders and system operators. On the one hand, due to the high flexibility of bike sharing system, the system typically ends up with an uneven distribution of bikes across the different stations (due to the uncontrolled, uneven demand), often rendering the check in or check out service unavailable at some stations where bicycle docks are either fully occupied or empty. During peak periods, user demand characteristics differ among stations in certain areas. For example, rental demand usually gets larger in workday morning near residential areas, whereas return demand gets larger near commercial districts.

Disadvantages:

At present, operators perform bike redistribution based on monitor video and user complaints. However, this method has exposed the serious lag. It is usually when service unavailable events occur that operators start to give some scheduling instructions. When the vehicle arrives, service unavailable events may have passed for some time, which makes it difficult to meet the needs of users at rush hour.

Proposed system:

The proposed system will give you some recommendations to the system admin. We are going to predict in which perspectives most of bike rides happening. To do this we are using machine learning algorithms and finding the accuracy scores. Linear regression, decision tree and random forest to compare the accuracy scores. This help to the system manager by placing more bikes in the peak time and they can maintain their business in efficient way

Advantages:

- Accurate results
- Gives future results and then can establish wealthy system
- It can reduce loss to the system admin

Modules:

Numpy:

NumPy enriches the programming language Python with powerful data structures, implementing multi-dimensional arrays and matrices. These data structures guarantee efficient calculations with matrices and arrays. The implementation is even aiming at huge matrices and arrays, better know under the heading of "big data". Besides that the module supplies a large library of high-level mathematical functions to operate on these matrices and arrays.

pandas

Pandas is a high-level data manipulation tool developed by Wes McKinney. It is built on the Numpy package and its key data structure is called the DataFrame. DataFrames allow you to store and manipulate tabular data in rows of observations and columns of variables.

Sklearn

interfaceScikit-learn provides a range of supervised and unsupervised learning algorithms via a consistent in Python.

It is licensed under a permissive simplified BSD license and is distributed under many Linux distributions, encouraging academic and commercial use

matplotlib

It is a plotting library used for 2D graphics in python programming language. It can be used in python scripts, shell, web application servers and other graphical user interface toolkits.

There are several toolkits which are available that extend python matplotlib functionality. Some of them are separate downloads, others can be shipped with the matplotlib source code but have external dependencies.

Seaborn:

Seaborn is a Python data visualization library based on matplotlib. It provides a high-level interface for drawing attractive and informative statistical graphics.

Conclusion:

Therefore, with defined station activeness, we optimize the bike circulation by leading users to shift bikes between highly active stations and inactive ones. We extensively evaluate the performance of our design through real-world datasets. Evaluation results show that the percentage of frequently used bikes decreases by 33:6% on usage number and 28:6% on usage time.

References

[1] P. DeMaio, "Bike-sharing: History, Impacts, Models of Provision, and Future," Journal of Public Transportation, vol. 12, no. DeMaio 2004, pp. 41–56, 2009. [Online]. Available: http://www.transitinformatics.org/test/nctr/wp-content/ uploads/2010/03/JPT12-4DeMaio.pdf

[2] P. Midgley, "Bicycle-sharing schemes: enhancing sustainable mobility in urban areas," United Nations, Department of Economic and Social Affairs, pp. 1–12, 2011.

[3] L. MetroBike, "2016 Year-end wrap-up will appear at the end of January," http://bike-sharing.blogspot.com/2017/01/ 2016-year-end-wrap-up-will-appear-at.html.

[4] J. Froehlich, J. Neumann, and N. Oliver, "Sensing and Predicting the Pulse of the City through Shared Bicycling," in IJCAI, 2009.

[5] A. Kaltenbrunner, R. Meza, J. Grivolla, J. Codina, and R. Banchs, "Urban Cycles and Mobility Patterns: Exploring and Predicting Trends in a Bicycle-based Public Transport System," Pervasive and Mobile Computing, vol. 6, no. 4, pp. 455–466, 2010