Analyzing the Common Features of Public Bicycle Share Data for Decision Support

Helen Tera
Institute of Computer Science
University of Tartu
Tartu, Estonia
helen.tera@ut.ee

Abstract—While public bicycle share (PBS) systems are becoming more popular around the world, the amount of public transportation data continuously increases. Analyzing this data can provide us with a better understanding of human mobility patterns and urban dynamics, as well as help us improve existing bicycle sharing solutions and the infrastructure for all bike users. This mapping study aims to give an overview of different features of the PBS data that have been used to visualize and analyze the usage of bicycle sharing solutions in order to provide decision support.

Keywords—public bicycle share, decision support, visualization, data analysis

I. INTRODUCTION

As sustainability and public health are major objectives of modern city planning, public bicycle share (PBS) programs play an increasingly important role in urban transportation system design. Reducing the number of personal motorized vehicles is one of the more common strategies for reducing carbon emissions and traffic congestion and introducing PBS systems can play an important role in increasing bicycle modality share in cities. Some programs have reported an increase in modal share of bicycles up to 1.5% owing to the introduction of bicycle sharing [1]. Besides demobilization and decarbonization, other common benefits of PBS systems include social interaction, accessibility and low cost [2].

Emerging PBS programs produce large amounts of data, which give insight to PBS usage and human mobility patterns. Analyzing this novel data can provide guidance for better bicycle scheduling, station planning and road network design [3].

PBS data has many different features that can provide value for city administrations and urban scientists and various visualization techniques can be used to analyze this data. The aim of this paper is to collect different features used for the analysis and visualization of public bicycle share data with the goal of providing decision support. Different PBS data metrics are introduced and their practicality for decision support purposes is discussed. This article aims to argue, that proper analysis and appropriate visualization of PBS data can lead to deep insight and impact in a policy making context.

In section II a brief overview of the history of PBS programs is presented along with the descriptions of two main types of PBS systems. Section III introduces the search methodology for the underlying mapping study, while section IV presents the results of the search along with ideas for decision support of features found.

II. BACKGROUND

This section presents background on the topic of public bicycle share and its history. Then, two main types of PBS systems, reoccurring in relevant literature, are described.

A. History of Public Bicycle Share

A public bicycle share (PBS) system is a short-term rental service in which bicycles are made available for shared use. The history of PBS systems dates back to 1965, Amsterdam, where the first PBS system named the White Bikes was introduced. [4] Ordinary bicycles, which were painted white for recognition, were left around the city for anyone to use. Due to customer anonymity it was easy to steal or damage the bicycles and thus the program collapsed within days. While the program ended in failure, it already included some of the components of PBS systems we know today, such as distinguishing public bikes by design or color [4].

According to Demaio [1], the first large-scale 2nd generation PBS system was launched in Copenhagen in 1991. The program introduced multiple improvements to 1st generation PBS systems, such as reinforced bicycles, fixed pick-up and drop-off locations and coin deposit system. While more organized than the previous programs, it similarly struggled with the problem on user anonymity and impelled the popularity of PBS systems [4].

As of mid-2021, there were at most 1999 simultaneously operating PBS systems around the world. Major portion of world PBS programs were employed in Asia (44%) and Europe (41%) [5].

B. Types of Public Bicycle Share

There are two types of public bicycle sharing systems: station-based and free-floating. In station-based systems bicycles have to be picked up and dropped off at designated...
stations, while in free-floating systems provides an option of picking up and leaving the used bicycle anywhere within the region of operation [6].

While free-floating PBS systems provide more comfort and increase usability, they also pose multiple difficulties. Efficient re-balancing strategies need to be considered for free-floating systems, to avoid congregation of shared bicycles in high-interest regions or scattering of the bicycles in remote areas. Station-based systems also provide easier solution for charging the bicycles and thus minimize maintenance activities and costs [7].

III. METHODOLOGY

This paper presents a systematic mapping study and is composed of the following steps: A) data search, B) study selection and C) data extraction. The results of the feature analysis follow in section IV.

A. Search strategy

The digital libraries chosen for this study were IEEE Xplore and EBSCO Discovery Service as they provided support for nesting and grouping of Boolean operators in the search string and provided a good set of initial results.

The first step of the search process was defining the search string. The search string consists of three main parts that are connected with AND operators. The first two sets of keywords include different naming conventions for PBS systems. The last part of the search string covered analysis, decision support and visualization.

The following search string was used to conduct the search:

\(((\text{public OR city}) \text{ (bike OR bicycle) (share OR sharing)}) \text{ AND} \text{ system OR program OR scheme}) \text{ AND} \text{ (analysis OR decision support OR visualization)}

Boolean operators were used in the search string for narrowing down the results. As PBS systems have various alternative namings, the search string included multiple possible variations of the terms.

B. Study selection

In addition to search string some supplementary limiters were used in the search. The following exclusion criteria was set:

- EC1. Exclude publications that are not in English.
- EC2. Exclude papers published before 2011.
- EC3. Exclude studies that are not part of academic journals or conference materials.

Additionally both digital libraries provided an option of filtering the results by subject. Thus, for narrowing down the results further, the recommended "bicycle sharing programs" tag was used in EBSCO Discovery Service and suggested "bicycles" tag was used in IEEE Xplore.

As a result, the search yielded 37 results from EBSCO Discovery Service and 69 results from IEEE Xplore. The rest of the study selection was conducted manually by inspecting the titles and abstracts of remaining articles. During that phase various articles connected to other PBS data related activities, such as demand prediction or bicycle reallocation, were discarded. Additionally, the publications concentrating on free-floating PBS systems were also excluded. Finally, a total of 14 relevant articles were selected: 9 from IEEE Xplore and 5 from EBSCO Discovery Service.

It is important to note that the chosen publications are different in their structure and goals - some concentrate more on visual analysis, while others tackle more specific problems like trip prediction or service planning. Regardless of the aim of a given paper, all the chosen publications feature different metrics and visualization techniques that support their goals and are relevant for in depth analysis of PBS data.

C. Data extraction

The selected 14 papers were worked through and relevant features were extracted into a data extraction table. The results can be found in the Table I below.

<table>
<thead>
<tr>
<th>Feature</th>
<th>Count</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Station level demand</td>
<td>9</td>
<td>[13] [18] [19] [110] [111] [112] [113]</td>
</tr>
<tr>
<td>Trips per time-frame</td>
<td>8</td>
<td>[13] [5] [16] [12] [10] [11] [18]</td>
</tr>
<tr>
<td>Weekday/weekend comparison</td>
<td>6</td>
<td>[13] [16] [12] [10] [19] [15]</td>
</tr>
<tr>
<td>Station spatial relations</td>
<td>5</td>
<td>[13] [11] [12] [14] [19]</td>
</tr>
<tr>
<td>Rent/return rate</td>
<td>2</td>
<td>[10] [11]</td>
</tr>
<tr>
<td>Trips by gender</td>
<td>1</td>
<td>[20]</td>
</tr>
</tbody>
</table>

The features are sorted by occurrence, starting with the most popular features. The metrics with different wording but similar idea have been assembled under a single description.

As can be seen in Table I there are multiple reoccurring features used in PBS data analysis. The most common feature used to describe PBS data is station level demand. There are also more uncommon features, that are used only in several of the studies.

IV. RESULT ANALYSIS

This section describes the extracted common features from the selected studies and details how analyzing each feature could be used for decision support.

A. Station level demand

Station level demand describes the number of bicycles rented from a given station, either per hour or over a span of a longer period. This feature is the most common feature used for analysis of PBS data. Out of 14 selected papers, 9 used some form of this metric.

There are numerous ways for illustrating station level demand: most of the articles combined it with a map view, either by utilizing a heat map [11] [13] or emphasizing the demand using different sizes or colours based on the station’s demand.
Another way of describing station level demand is using a heat map for rent/return frequency, which will be described in more details in section IV-E.

One of the most common problems with station-based PBS systems is the bicycle rebalancing issue - as bicycle demand differs from station to station, some stations can run out of bikes, while others can get overloaded. Being able to ensure that a user can rent a bike from a required station and return it at a station matching their destination, is of crucial importance to provide the expected flexibility and service quality of PBS. Understanding station level demand plays a major part in solving this problem. There are two main strategies for solving the rebalancing issue. One option is using designated cargo vehicles to reallocate the bicycles during the low-demand time (typically nighttime). Alternatively, a more static approach can be used, by incentivizing the PBS users to leave or pick-up bikes in specific stations.

B. Trips per time-frame

The other principal metric used in relevant literature is trips per hour/day/month. This feature represents the number of public share bicycle trips grouped by a period of time.

One of the most predominant ways of illustrating PBS data is visual distribution of trips by time of day, or presented the changes in trip counts as a line chart over a longer span of time, aggregating trips either daily or weekly.

Some of the articles also grouped PBS trips into a longer period of time, for example using monthly distribution, or presented the changes in trip counts as a line chart over a longer span of time, aggregating trips either daily or weekly.

This representation of PBS data gives its analyst insight into the public bicycle demand at different times of month or year. Analyzing changes in public bicycle demand over a longer period of time can reveal issues in PBS system or point to atypical factors affecting mobility. For example, figure illustrates, how PBS trips in Almaty followed a predictable distribution, with more trips happening during warmer months of spring, summer and autumn. Yet, it also highlights the exception of the months of April and May, when mobility patterns of Almaty were heavily influenced by the coronavirus pandemic.

Understanding public bicycle demand changes in span of one day can provide decision support for maintenance planning or bandwidth management, for example to avoid the evident disadvantages of scheduling station maintenance at a time when public bicycle demand is at its peak.
C. Weekday/weekend comparison

Another predominant metric used in various studies is comparison of number of trips on weekdays and weekends. For visualization, different techniques are used. Some opted to aggregate number of trips daily and display the total number per day on a line chart [17] or on a bar chart [18], while others also show differences in weekday and weekend bicycle usage by hour [8] [16] [15].

Significant differences in the number of trips for weekdays and weekends can point to different motivation for PBS use. If the total number of PBS trips is much higher during weekdays, it could indicate that public bicycles are primarily used for utilitarian purposes - commuting to work or school - while high number of PBS trips during weekdays could suggest that a given PBS program is also often used for recreational purposes. Similar idea was used in [10] for classifying clusters. Looking at figure 4, it can be assumed that cluster C5 belongs to the utilitarian usage pattern and thus could mostly consist of stations in the commercial areas.

![Fig. 4. Total number of trips by weekday. Source: [10]](image)

This metric can be a valuable complement to the hourly trips metric, as when averaging the number of trips by hour over a period longer than a week, the differences between weekday and weekend bicycle usage patterns get overlooked. Knowing the anticipated demand during different times can help transportation officials dispatch bikes in anticipation of high demand, in order to improve service quality.

D. Station spatial relations

Station spatial relations describe which origin-destination station pairs are more utilized. This metric illustrates the flow from a given station or to a specific station and helps to understand which are the most popular links between stations. Studying this metric can provide insight into city’s general mobility patterns and the flow of bikes between different areas.

Station spatial relations are most commonly displayed on a map with different color [11] [19] or link width [3] [12] [14] used to illustrate the number of trips on given route.

![Fig. 5. Station spatial relations for different days. Source: [11]](image)

Analyzing the actual routes taken by the PBS users can help to identify areas and road segments that are utilized by the PBS users the least. Having this information can be useful in road network planning for non-motorized vehicles, in order to further increase the modal share of green transportation modes.

E. Rent/return rate

Rent/return rate is the relation between the number of bikes rented at a station and number of bikes returned at the same station. It is another way of describing station level demand, but it adds a different level of information to it. While station level demand is mostly based on stations’ rental rates, it does not consider the return numbers. Rent/return rate helps to understand the actual number of bicycles available at a station or within a cluster.

Various heat maps are used to illustrate the rent/return rate in relevant literature. In [10] a heat map was used to illustrate the rent/return rate between different clusters. On the other hand, [14] used a temporal heat map to highlight the rent/return difference for different days (Fig. 6).

This simple matrix can give useful insight into understanding a station’s demand at a given time. The rent/return difference is color-coded: cells with red tones state that the rental number is greater than the return number, while in the blue cells the return number is greater. When the rent/return rate is balanced, the cell is colored white. Using this simple...
visualization technique, one can easily determine at which hours does the station experience outages and cannot provide enough bicycles for rent.

F. Trips by gender

Studying the bike-usage behaviors of different groups of people can provide policy makers insight into the commuting behavior of various demographic groups. This insight can help to encourage underrepresented groups to partake in active modes of transport. One common technique that can provide better understanding of PBS users’ mobility patterns is grouping PBS trips by users’ gender.

In [20] PBS users’ gender was one of the metrics used to predict the potential destination station and arrival time of PBS user at the start of their trip.

Another paper [23] analyzes the PBS trips by gender and discusses the reasons for differing traveling patterns among users of different gender. It was revealed, that the women were underrepresented among the PBS users in London and different travel patterns were discovered for men and women: male cyclists used the system predominantly for commuting, while the trip patterns of female PBS users showed a dominant leisure function.

V. CONCLUSION

In this paper, different metrics of PBS data have been presented along with examples from relevant papers. Ideas and instances of their uses in decision support have been highlighted. These metrics are station level demand, trips per time-frame, weekday/weekend comparison, station spatial relations, rent/return rate and trips by gender.

The author believes that the first step to policy making is thorough and deep understanding of the usage patterns of the PBS system. Having an understanding of when and for which purposes the PBS bikes are used, can lead to better strategies for encouraging people to engage with the PBS system. The more dimensions of PBS data is explored, the clearer picture can be obtained for decision support. Drawing on the features discussed in this paper, a thorough base for understanding the PBS data can be created and utilized, providing decision support for city administrators and transportation workers.

REFERENCES


