Chapter 1: Introduction

Draft version

1.1 Context

`*NOTE: First, talk about the advantages of sustainable mobility in cities. Then, explain what bike sharing systems are, and describe the different generations of bike sharing systems. Then, cover in more detail how dockless bike sharing systems function. Mention that in the past year, they have rapidly expanded over the whole world. Mention advantages, but also problems.*`

In order to offer a serious alternative for motorized transport, bike sharing systems need to be reliable. For station-based systems, this means that situations in which a user either can not start a trip at the desired time and location because of an empty docking station, or can not end a trip at the desired time and location because of a full docking station, should be avoided. In dockless bike sharing systems, the latter is not an issue, since bikes can be left anywhere and anytime. The first, however, remains a challenge. A user does not have the certainty of finding an available bike near the desired trip origin, at the desired time of departure.

Accurate, short-term forecasts of the distribution of available bikes both over space and time, enable system operators to anticipate on imbalances between supply and demand, such that the occurrence of situations as described above is limited to a minimum. Furthermore, users can take advantage of these forecasts also in a direct way, to plan their trips effectively, and reduce the time spend on searching for a bike. Hence, forecasting bike availability is of great importance when turning the shared bike into a reliable, pleasant and uncomplicated mode of transport.

1.2 Objective

Several approaches have been developed to forecast the bike availability in station-based bike sharing systems. However, dockless bike sharing systems remain fairly unexplored in that sense, since their rapid expansion only started recently, and in contradiction to the station-based system, large historical data sets are not widely and openly available.

Furthermore, the spatio-temporal nature of dockless bike sharing data brings additional challenges to the forecasting task. In station-based systems, forecasts are only required at fixed spatial locations, and although some works include spatio-temporal relationships between stations, they can mostly be treated as distinct entities, with different forecasting models for each station. In dockless bike sharing systems, however, available bikes can be at any spatial location inside the system area. Besides, the bike availability not only has a spatial dependence, with more bikes being available in certain areas, and a temporal dependence, with more bikes being available at certain times, but those dependencies are also linked to each other. That is, temporal patterns may differ over space, and vice versa.

The objective of this thesis is to deal with those challenges, and develop a generally applicable methodology for bike availability forecasting in dockless bike sharing systems.

1.3 Related work

`*NOTE: Literature review is done, but needs to be summarized and structured in text. Cover the following topics: forecasts in station based systems, forecasts in dockless systems, forecasts in other transportation areas, like car sharing, taxis, public transport,`

```
et cetera, influences on the usage of bike sharing systems (weather, etc), descriptions
of the spatio temporal dynamics of dockless bike sharing system, studies on user profiles
of bike sharing system users*
```

1.4 Approach

The few attempts made to design forecasting systems for dockless vehicle sharing, either with bikes, scooters or cars, can be divided into two groups of approaches, which are labeled here as the *grid based approach* ad the *distance based approach*. The grid based approach (Yi et al. 2018; Caggiani et al. 2017) starts with laying a spatial grid over the system area, and treating each grid cell as being a docking station. That is, from the historical data on locations of available vehicles, the number of vehicles in each grid cell is counted at several timestamps in the past, creating a time series of counts. Hence, for a given geographical point, the forecasted value will be the expected number of vehicles inside the grid cell that contains the point.

The distance based approach (Formentin, Bianchessi, and Savaresi 2015) uses the historical data to calculate the distance from a given geographical location to the nearest available vehicle for several timestamps in the past, creating a time series of real-valued distances. Hence, the forecasted value will be the expected distance from the given point to the nearest available vehicle, at a given timestamp in the future.

The bike availability forecasting system as proposed in this thesis uses the distance based approach, because it has the following advantages over the grid based approach.

- Forecasts will not be dependent on the chosen spatial resolution of the grid.
- Forecasts will not be made with count data, which would have limited the choice of suitable forecasting models.
- Forecasts can be interpreted in the same way at every location in space. This in contradiction to the grid based approach, where for a location at the edge of a grid cell, there might as well be closer bikes within the neighboring cell.
- Forecasts give more freedom to the user, who can decide for himself if he considers the forecasted distance acceptable or not. In the grid based approach, the cell size is fixed, and does not take into account that some people are willing to walk further than others.

Throughout the thesis, the proposed forecasting system will be referred to as the **D**ockless **B**ike **A**vailability **F**orecasting **S**ystem (DBAFS). The focus will lay on the design of a general methodology for it, rather than on the creation of a fully configured, practical application. A forecast can be requested for any location within the system area of a dockless bike sharing system, and, theoretically, any timestamp in the future. However, short-term forecasts, up to one day ahead, form the target, and longer forecasting horizons will not be handled.

1.5 Outline

The rest of the document is structured as follows. Chapter 2 provides a theoretical background of the concepts used in this thesis, which are all linked to the field of time series analysis. In Chapter 3, the methodology of the proposed forecasting system is described into detail. Chapter 4 presents the data on which the system is tested, and describes the experimental setup. In Chapter 5, the results of the experiments are shown, interpreted, and discussed. Finally, Chapter 6 lists the conclusions of this thesis.

References

Caggiani, Leonardo, Michele Ottomanelli, Rosalia Camporeale, and Mario Binetti. 2017. "Spatio-temporal clustering and forecasting method for free-floating bike sharing systems." In *Advances in Intelligent Systems and Computing*, 539:244–54. Springer Verlag. http://dx.doi.org/10.1007/978-3-319-48944-5{_}23.

Formentin, Simone, Andrea G. Bianchessi, and Sergio M. Savaresi. 2015. "On the prediction of future vehicle locations in free-floating car sharing systems." In *2015 Ieee Intelligent Vehicles Symposium (Iv)*, 1006–11. doi:10.1109/IVS.2015.7225816.

Yi, Ai, Zongping Li, Mi Gan, Yunpeng Zhang, Daben Yu, Wei Chen, and Yanni Ju. 2018. "A deep learning approach on short-term spatiotemporal distribution forecasting of dockless bike-sharing system." *Neural Computing and Applications*, 1–13. doi:10.1007/s00521-018-3470-9.