Data-Driven Bike Sharing System Optimization: State of the Art and Future Opportunities

Longbiao Chen[1],Zhihan Jiang[1], Jiangtao Wang[2], Yasha Wang[2]
[1]Fujian Key Laboratory of Sensing and Computing for Smart Cities, School of Information Science and Engineering, Xiamen University, China
[2]Institute of Software, School of Electronics Engineering and Computer Science, Peking University, China
longbiaochen@xmu.edu.cn

Abstract

The prevalence of public bike sharing systems (BSS) in recent years provides a last-mile-trip supplement to other modes of public transit systems in a convenient and environmentally-friendly way. However, due to the inherent imbalance between the renting and return rates, it is very challenging for the operators to optimize the resources (e.g., stations and bikes) to meet the fluctuating demand appropriately. Traditionally, urban planners use surveys to guide the BSS optimization, which is costly and inefficiently. With the availability of multi-source data in the city, it is now an emerging trend to leverage multi-source urban data for optimizing BSS. In this article, we specifically focus on the data-driven BSS optimization problem and provide a comprehensive literature review with future research opportunities. We first identify the main challenges and present a general technical framework for BSS optimization, and then present the representative studies in each sub-problem. Finally, we charter the research directions and opportunities in the future.

Keywords

Bike Sharing System Optimization; Demand Prediction; Station Placement; Bike Balancing

1 Introduction

Recent years have witnessed prevalence of public bike sharing systems (BSS) [9], which have been widely adopted in many large cities worldwide (e.g., Chicago (Divvy), New York (Citi), Washington, D.C. (Capital Bikeshare), Paris (Vélib), Beijing (Ofo Bike), etc.). In BSS, people can rent and return bikes from nearby stations, which can be used as a first-and-last mile trip supplement to other modes of public transportation systems (e.g., metro trains and buses). As the bike usage is green and low-carbon, BSS indeed provides an environmentally friendly and sustainable solution for people's short-distance trip demand. Besides, due to a large number of rental stations scattered in the city, people can pick up and drop off very conveniently, which avoids the upset situation while searching the available parking lot with private vehicles and saves the waiting time needed for the public transportation.

Although BSS brings significant benefits, it is rather challenging for the operators (e.g., the government or the commercial operators) to manage and maintain these systems effectively and efficiently. Particularly, one primary challenge is how to dynamically adjust the resources according to the fluctuating rent-and-return demand. Due to inherent imbalances between the renting rates and return rates at the various stations, the bike resources may be inadequate at some stations while too redundant at others. The insufficiency leads to the unavailability of bikes, while the redundancy may bring the concern of parking spot or vacant lockers[11]. Traditionally, urban planners use surveys to collect information on bike demand to guide the station deployment and bike placement [1]. However, these approaches have two main disadvantages. First, it incurs a great amount of time, labor, and money so that the method's scalability is relatively poor. Second, as the survey data collection frequency is relatively low, it cannot respond to the dynamic change of BSS systems (e.g., the number of bikes at each station).

In recent years, with the availability of multi-source data in the city (e.g., bike usage records, human check-in/mobility traces, meteorology, weather condition, and point of interests), it is an emerging trend that researchers from various communities (such as Data Mining, Urban Computing, and Ubiquitous Computing) begin to study BSS optimization problem from the urban data perspective [2, 14, 8]. In this paper, we call this paradigm as urban-data-driven BSS optimization. Specifically, in this article, we divide the optimization objectives into the following two categories, according to the frequency with which the optimization operation is performed.

- **Long-term Optimization** (Station Placement Optimization). This problem aims at maximizing the utility when designing the locations of bike stations to best meet users' trip demand [10, 13, 6]. Generally speaking, as the update of station placement locations is not very frequent (e.g., days or weeks), we classify it as the long-term optimization in this article.

EWSN '19, February 25 - 27, 2019, Beijing, China.

- **Short-term Optimization** (Bike Balancing Optimization). For a fixed set of stations, one main problem is how to reposition the bikes (usually by vehicles and staffs) among each station according to users' fluctuating trip demand [15, 12, 7]. As the balancing operation is conducted more frequently (e.g., hours or a day) compared to the station placement, we classify it as the short-term optimization in this article.

In this article, we specifically focus on the above two kinds of urban-data-driven BSS optimization problems and provide a comprehensive literature review with future research opportunities. We first identify the main challenges for BSS optimization in Section 2, and then present the general technical framework in section 3. Finally, we charter the research directions and opportunities in Section 4, and conclude the entire article in Section 5.

2 BSS Optimization: Challenges

Before the prevalence of bike sharing system, the optimal placement of public facilities (e.g., gas and hydrogen filling stations) is already widely studied in urban planning field. Compared to these traditional problems, the BSS optimization is more challenging due to the following reasons.

- **Complicated Optimization Operations.** Traditional public facility placement merely focuses on adding new stations. In contrast, the BSS optimization involves with multiple operations. In terms of long-term optimization (i.e., station placement), both new stations and bike lockers can be added, while the existing ones may need to be removed. In terms of short-term optimization (i.e., bike balancing), the operations such as inventory decision and routing decisions are involved. Moreover, these complicated operations in BSS are correlated and should be jointly optimized.
- **Dynamic Demand Changing.** The traditional public facility placement is usually adjusted in a long-term manner (e.g., weeks or months). However, the BBS has to optimize its resources in a much more real-time way. For example, multiple vehicles and staffs have to be scheduled continuously (e.g., hours) in certain urban regions to reposition the bikes in each station, which are based on the accurate prediction of dynamically changing station-level bike usage demand.
- **Multiple Influencing Factors.** The BSS optimization relies on the accurate station-level prediction of bike demand, which is non-trivial due to many influencing factors (e.g., historical trip pattern, weather condition, POI, and surrounding public transit network). Moreover, the importance of each factor is not fixed for each station. For example, the demand change and imbalance may follow the temporal pattern at stations within the CBD areas and residence areas. However, the imbalance sometimes may be persistent (e.g., low return rate in stations located on the top of hills). Thus, it is quite challenging to build a prediction model by jointly considering multiple factors in different scenarios.

3 BSS Optimization: A Framework

Although the state-of-the-art research work for BSS optimization differs in problem formulation, optimization algorithm and adopted datasets, a general BSS optimization conceptual framework (called the BikeOptimizer in this article) can be summarized by identifying their common components and procedures, as illustrated in Figure 1.

Specifically, the BSS optimization framework mainly consists of the following modules:

- **Bike Demand Prediction.** For either short-term (station placement) or long-term (bike balancing) BSS optimization, one fundamental issue is the prediction of bike demand (pick-up and drop-off) in each rental station [7, 4]. As the bike demand is influenced by multiple factors (e.g., historical records, weather condition, POI information, events, and public transportation networks), existing studies use either data mining techniques or classical statistical method to build a multi-factor prediction model based on multi-source urban data.
- **Station Placement Optimization.** One of the key factors for maximizing the overall utility of a BSS is placing bike stations at locations that can best meet users' trip demand. Traditionally, urban planners rely on dedicated surveys to understand the local bike trip demand [4], which is costly in time and labor, especially when they need to compare many possible places. Although existing literature has identified a large set of factors that may affect bike trip demand in general, each city has its own environmental, social, and cultural characteristics, resulting in different adoption patterns of bike sharing programs. Therefore, it is necessary for BSS operators to identify the relevant factors to BSS demands and make optimal station placement plans.
- **Bike Balancing Optimization.** Based on the predicted station-level bike usage demand, the bike balancing problem (also called bike repositioning/redistribution problem) involves two major issues: inventory decisions and routing decisions. The inventory decision is to determine the number of bikes should be placed in each station within each time interval, while the routing decision involves determining the number of bikes should be removed or placed in each station at each visit of an operation vehicle, with the objective of maximizing the users' satisfaction while minimizing the balancing cost.

4 Research Directions and Opportunities

Various techniques have been proposed to optimize the resource allocation in BSS. We next highlight the research gaps with several future research directions.

- **Station-Centric Demand Predictor.** As mentioned above, one fundamental issue of the BSS optimization is the ability to accurately predict the bike trip demand. Most of the existing work builds a unified predictive mode for all stations. However, the features influencing the trip demand may be diverse across different stations [5, 3]. For example, many stations may be highly influenced by temporal features (e.g., the number of bikes

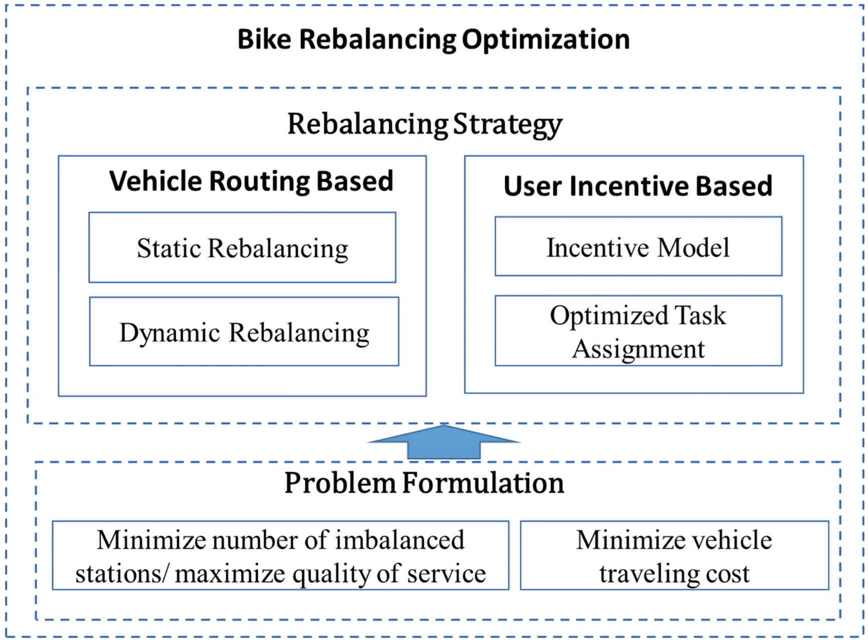

Figure 1. The BSS optimization framework.

at stations within the CBD areas usually varies with the morning and evening commuting peak), while some may not follow this pattern (e.g., constant low return rate in stations located on the top of hills). Therefore, it is necessary to build a station-centric trip demand predictive model for each station. However, learning and choosing the relevant features for each station is not easy, especially for those stations without sufficient historical training data. In this case, we may exploit data of other similar stations with sufficient data to collaboratively address this issue, while measuring the similarity would be a key problem.

- **Hybrid BSS Optimization.** Existing studies focus on station placement and bike balancing separately. However, in real-world application scenarios, the operator of BSS should optimize BSS by jointly considering these two strategies. For example, if a certain rental station S is always entirely full, then the vehicles will be frequently routed to remove the bikes only by adopting the balancing strategy. In this case, the traveling cost of vehicles and operation cost of staffs would be high with a long time period, so that it may be an alternative to deploying a new station near S to split the flow. Therefore, a new research proposal may be to design a hybrid plan of short-term and long-term optimization with the objective of minimizing the total cost (traveling cost of vehicle routing and deployment cost of station placement).
- **Multi-Operator BSS Optimization.** State-of-the-art research work only aims at optimizing BSS from the perspective of a single operator. In fact, there are usually multiple operators in a city with their own BSS. It is necessary to optimize one city's bike sharing ecosystem by considering multiple BSSs managed by different operators. However, this task is quite challenging due to the following reasons. First, different operators run their own systems, and they are not willing to share their back-end data (e.g., the number of bikes at each station) for the commercial purposes. Thus, a fair and all-win data sharing mechanism should be developed by fully considering the commercial benefit of different operators. Second, even if an appropriate data sharing mechanism is established, it is non-trivial to jointly optimize the overall utility of multiple systems, because different operators may have different budget constraints and service level optimization requirements.

5 Conclusions

In this paper, we present a survey of bike sharing system optimization problem based on urban data. Specifically, we first identify three key issues in bike sharing system optimization (i.e., bike demand prediction, station placement and bike balancing). We then introduce a general BSS optimization framework, BikeOptimizer, by summarizing the state-of-the-art research works, and elaborating the key components of the framework. Finally, we point out some future research directions and opportunities, which may help further optimize the bike sharing systems with more practical concerns.

6 Acknowledgments

This research was supported by the China Fundamental Research Funds for the Central Universities No. 0630/ZK1074, Natural Science Foundation of Fujian Province No. 2018J01105, and Natural Science Foundation of China No. 61802325.

7 References

[1] A. M. Burden, R. Barth, and others. Bike-share opportunities in new york city. *New York: Department of City Planning*, 2009.

[2] D. Chemla, F. Meunier, and R. W. Calvo. Bike sharing systems: Solving the static rebalancing problem. *Discrete Optimization*, 10(2):120–146, 2013.

[3] L. Chen, J. Jakubowicz, D. Yang, D. Zhang, and G. Pan. Fine-Grained Urban Event Detection and Characterization Based on Tensor Cofactorization. *IEEE Transactions on Human-Machine Systems*, 47(3):380–391, 2017.

[4] L. Chen, X. Ma, T.-M.-T. Nguyen, G. Pan, and J. Jakubowicz. Understanding bike trip patterns leveraging bike sharing system open data. *Frontiers of Computer Science*, pages 1–11, 2016.

[5] L. Chen, D. Yang, J. Jakubowicz, G. Pan, D. Zhang, and S. Li. Sensing the Pulse of Urban Activity Centers Leveraging Bike Sharing Open Data. In *2015 IEEE 12th Intl Conf on Ubiquitous Intelligence and Computing and 2015 IEEE 12th Intl Conf on Autonomic and Trusted Computing and 2015 IEEE 15th Intl Conf on Scalable Computing and Communications and Its Associated Workshops (UIC-ATC-ScalCom)*, pages 135–142, 2015.

[6] L. Chen, D. Zhang, G. Pan, X. Ma, D. Yang, K. Kushlev, W. Zhang, and S. Li. Bike Sharing Station Placement Leveraging Heterogeneous Urban Open Data. In *Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing*, volume UbiComp'15, pages 571–575. ACM, 2015.

[7] L. Chen, D. Zhang, L. Wang, D. Yang, X. Ma, S. Li, Z. Wu, G. Pan, T.-M.-T. Nguyen, and J. Jakubowicz. Dynamic Cluster-based Over-demand Prediction in Bike Sharing Systems. In *Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing*, UbiComp'16, pages 841–852. ACM, 2016.

[8] M. Dell'Amico, E. Hadjicostantinou, M. Iori, and S. Novellani. The bike sharing rebalancing problem: Mathematical formulations and benchmark instances. *Omega*, 45:7–19, 2014.

[9] P. DeMaio. Bike-sharing: History, impacts, models of provision, and future. *Journal of public transportation*, 12(4):3, 2009.

[10] I. A. Forma, T. Raviv, and M. Tzur. A 3-step math heuristic for the static repositioning problem in bike-sharing systems. *Transportation research part B: methodological*, 71:230–247, 2015.

[11] J. R. Lin and T. H. Yang. Strategic design of public bicycle sharing systems with service level constraints. *Transportation research part E: logistics and transportation review*, 47(2):284–294, 2011.

[12] J. Liu, L. Sun, W. Chen, and H. Xiong. Rebalancing bike sharing systems: A multi-source data smart optimization. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 1005–1014. ACM, 2016.

[13] M. Rainer-Harbach, P. Papazek, B. Hu, and G. R. Raidl. Balancing bicycle sharing systems: A variable neighborhood search approach. In *European Conference on Evolutionary Computation in Combinatorial Optimization*, pages 121–132. Springer, 2013.

[14] T. Raviv, M. Tzur, and I. A. Forma. Static repositioning in a bike-sharing system: Models and solution approaches. *EURO Journal on Transportation and Logistics*, 2(3):187–229, 2013.

[15] J. Schuijbroek, R. C. Hampshire, and W.-J. Van Hoeve. Inventory rebalancing and vehicle routing in bike sharing systems. *European Journal of Operational Research*, 257(3):992–1004, 2017.