

# Bike Sharing as a Key Smart City Service: State of the Art and Future Developments

Federico Chiariotti, Chiara Pielli, Angelo Cenedese, Andrea Zanella, Michele Zorzi

Department of Information Engineering (DEI)

University of Padova, via G. Gradenigo 6/B, Padova, Italy

Email: {chiariott, piellich, cenedese, zanella, zorzi}@dei.unipd.it

**Abstract**—Bike sharing has outgrown its first failures in the '60s and '70s and has become ubiquitous around the world. This rapid growth is strongly intertwined with the rise of Smart Cities: the use of connected bikes makes the service more practical for users, avoids thefts and provides a large amount of data for system planners. Over the past few years, research on bike sharing has bloomed, providing several innovative solutions to improve the service and to encourage citizens to use environmentally friendly modes of transportation, reducing both traffic and commuting times. In this work, we present the most promising developments towards a tighter integration between Smart City data and techniques and the operation and planning of bike-sharing systems, focusing on two model use-cases: New York City's CitiBike service, a large system with hundreds of stations, and Padova's GoodBike system, which has just 28 stations.

## I. INTRODUCTION

Smart Cities are quickly becoming a key application for the Internet of Things [1]; the integration of data from a network of pervasive sensors can be used to make existing city services better and enable entirely new ones. Bike sharing is one of these services: from its modest beginnings in Europe in the '60s [2], which were plagued by theft and vandalism, it has quickly become ubiquitous in cities all around the world when communication technology allowed automatic unlocking and billing. Users can be identified by using electronic keys, providing a strong deterrent on theft, while the possibility to access real-time information about the availability of bikes can make the service more usable and convenient. Mass bike sharing is one of the first real Smart City-enabled services, and the benefits in terms of traffic reduction [3] and public health [4] are already being enjoyed by citizens in Europe, China and North America. With proper planning, bike sharing can become the missing “last mile” connection [5] to mass transit systems, allowing citizens to quickly reach any point of the city without the massive infrastructure and maintenance costs that a capillary mass transit network would require.

Besides the service offered to citizens, bike sharing usage data are an invaluable source of information about mobility and transport modes in modern cities: over the past 10 years, the availability of demand datasets has pushed the research community towards several attempts at understanding bike sharing patterns [6], both to shed light on citizens' habits and to improve the service and mobility in general. In this work, we will first use data from New York City's CitiBike service<sup>1</sup> as a significant use-case to present the possible challenges and opportunities that can arise from a further integration of bike-sharing systems into the Smart City paradigm. We will then

compare it to the data from the GoodBike Padova<sup>2</sup> system, discussing the differences between large-scale and small-sized systems.

The rest of the paper is organized as follows. In Section II, we present some of the open possibilities to improve bike-sharing systems and integrated public transport, while Section III focuses on the analysis of the two datasets and on the possible models of the behavior of the system. Finally, in Section IV we offer our concluding remarks and describe some possible avenues of future research.

## II. CHALLENGES AND OPPORTUNITIES

Although bike-sharing services appeared in Europe as early as in the 1960s, they became a smart service only in recent years, and many issues still need to be addressed. Embedded sensors allow system administrators to obtain real-time data that can be used to improve the bike-sharing systems and adapt them to the needs of the users. Data is in fact fundamental to characterize the request for bikes over time and space, plan the location and capacity of the stations accordingly, and derive models to understand when intervention is needed, in terms of both bike maintenance and redistribution. In the following, we present the main opportunities to study and upgrade bike-sharing systems.

*a) Improving the system:* historical data can be exploited to gain insights into the complex bike activity patterns at stations, inferring the bike demand and identifying the most crowded areas in a city. A recent study [7] presents a method to detect broken bikes and lockers, which reduce the stations' capacity and negatively affect the service; such information is necessary to guarantee a prompt intervention. Historical data can also be analyzed to identify where bikes need to be at a given moment in order to meet the expected demand. In fact, external intervention (e.g., the deployment of a fleet of rebalancing trucks) is often needed to relocate bikes from overcrowded stations to those with a shortage of bikes, and designing an efficient rebalancing scheme requires to model the expected bike usage and availability over time. Real data can be used to validate and feed analytical models: several studies model the bike demand at each station as a queuing system with finite capacity and derive the target inventory level that limits the probability that stations are either completely full or empty [8]. Machine learning can be a powerful tool to extrapolate valuable information from real datasets and incorporate additional information such as the weather conditions [9], which are often troublesome to consider in an

<sup>1</sup><https://www.citibikenyc.com/system-data>

<sup>2</sup><http://www.goodbikepadova.it/>

analytical model. Rebalancing techniques can be divided into two macro-categories, namely, static and dynamic, where bike reallocation is performed at predetermined times or adapts to the changing network conditions, respectively. Common approaches to solve static rebalancing problems are mixed integer programming techniques [10], [11] and clusterization, where stations are grouped according to their demand patterns and rebalancing can be performed at a cluster level [8], [12]. Static approaches, however, may not be sufficient to avoid network failures during the day; dynamic rebalancing continuously monitors the current state of the network in order to redistribute bikes when and where needed. This strategy includes a scheduling component based on the users' activity during the rebalancing operation, and a routing problem. The network is typically modeled as a Markov Decision Process [13] and rebalancing is performed when the cost of moving the trucks does not exceed the gain it provides, which can be expressed in terms of, e.g., unfulfilled user demand [14] or expected due date violations [13]. A promising approach is given by hybrid schemes, e.g., in [15] static rebalancing overnight moves the network to an optimal configuration that minimizes the probability of stations becoming either empty or full in the following day, and daily clustering optimization handles rush-hour usage and ensures that users are never too far from an available bike or dock.

*b) Pricing:* pricing strategies that incentivize users to return bikes to the least loaded among the closest stations can limit the use of trucks and personnel to rebalance the system, thereby reducing costs. For example, [16] gauges the impact of assigning a price to each possible route, so as to optimize the total number of trips by discouraging routes that unbalance the system. In [17], the authors investigate the impact of incentives given to users so that they are pushed to choose less used stations. An interesting approach to improve the existing bike-sharing systems by taking advantage of the users themselves is gamification, i.e., the use of game design elements in non-game contexts, a powerful tool that can be used to tackle real-world problems. For example, in the Smart Cities context, citizen's behaviors and practices that are synergic with city policies could be promoted through virtual and real incentives [18]. By designing an effective reward system, bike-sharing users may be encouraged to take bikes from overcrowded stations and drop them at stations that are in shortage of bikes.

*c) Planning the bike-sharing system:* similarly to the usage of historical data to improve bike-sharing systems, Geographic Information Systems (GIS) can help plan additions to existing services. GIS methodology can be used to infer optimal locations for new routes, links between existing ones, and upgrades in the bike-sharing networks [19], [20]. Interestingly, [21] shows that combining GIS and multi-criteria evaluation analysis to plan for optimal bicycle facilities is much more effective than using only either supply- or demand-based planning criteria.

*d) Bike lanes and urban planning:* the potential benefits of bike-sharing data are not limited to the bike-sharing system itself: bike rides are a useful way to analyze mobility in general, and correlations in the data can help clarify citizens'

interaction with the urban environment. The most obvious example is the presence of bike lanes: studies based on the data from Montreal [22] and Washington DC [23] show that roads and neighborhoods with bike lanes highly incentivize using bikes over other modes of transportation. A study using the Brisbane bike-sharing data [24] considers urban topography, land use and zoning as well as bike-specific road infrastructure.

A recent work reverses the perspective [25]: instead of planning the bike-sharing system based on the existing infrastructure, the authors use GPS data from public bikes to propose possible new routes for bike lanes. The same could be done for zoning and mass transit: instead of building a bike-sharing system for a Smart City, the integration of data from multiple sources will allow city planners to evolve the city and its public transport system to support each other. This approach is better able to support the strain on the transport infrastructure due to the growing density of cities, and fits with the overall shift towards full integration of different domains that is being proposed for Smart Cities in other sectors [26].

### III. USE-CASES: THE CITIBIKE AND GOODBIKE SYSTEMS

New York City's CitiBike was launched in May 2013, and it is now the largest bike-sharing service in North America, with more than 700 stations in Manhattan, Brooklyn and Jersey City. It is also one of the most widely investigated systems in research on bike-sharing, since its usage data are released monthly since its inception. The other use case considered in this work is Padova's GoodBike system, launched in July 2013 and currently consisting of 25 stations (note that the analyzed data refer to a former displacement with 28 stations). The analyzed dataset is related to a period of one year, from August 2013 to August 2014.

#### A. Analyzing the data

Fig. 1(a) shows the stations in the Manhattan and Brooklyn areas and the demand in May 2017: the dots representing the stations are color-coded from green to red according to their relative total traffic. The figure also depicts the most common pairs of departure and arrival stations, using a darker shade of teal to indicate the most popular ones. These geographical patterns are very useful both for rebalancing and for planning future improvements to the system and to public transport in general.

Geographical data need to be complemented by temporal patterns and other considerations: for example, some of the most common routes go directly across the Brooklyn Bridge. This heavy traffic consists mostly of tourists and is less affected by the weekday rush hours than other routes between residential and commercial zones, which are commonly taken by commuters. The most used station, close to the top right corner of the map, is right next to Grand Central Station.

Fig. 2 shows the average demand for the month of July 2015 for a typical commuter station: the rush hours in the morning and evening are clearly noticeable during weekdays; since the station is probably used by many students and workers, there are more arrivals than departures in the morning and more departures than arrivals in the evening. On Fridays, there is a more homogeneous pattern in the afternoon, likely because

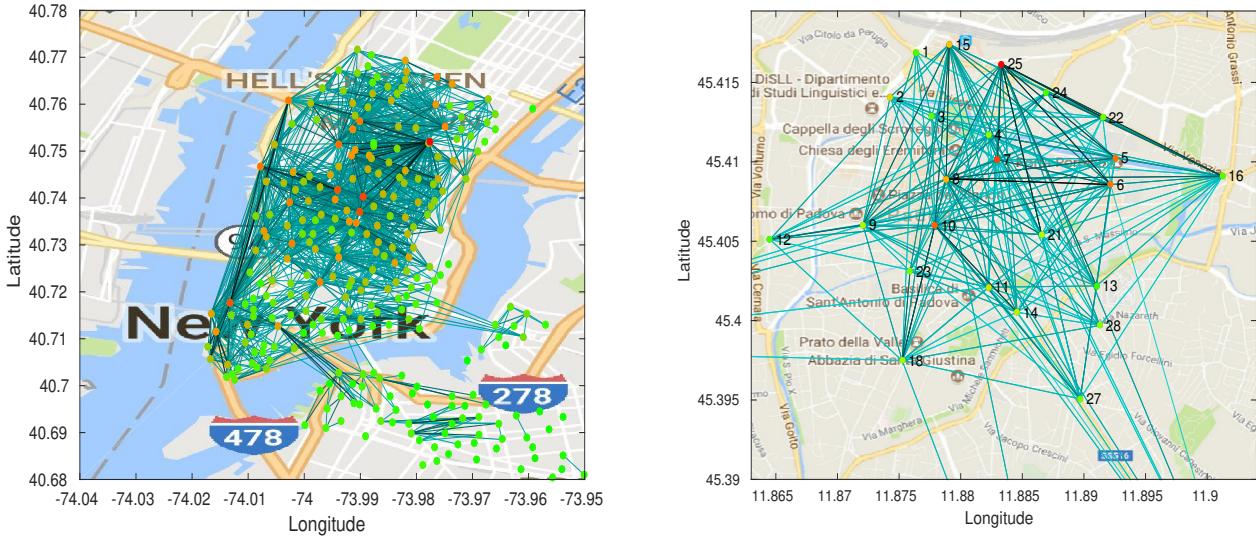


Fig. 1. Maps of the demand in a) the Lower Manhattan and Brooklyn area (May 2017) and in b) Padova (March 2014), color-coded from green (low demand) to red (high demand). The most common routes are plotted in teal, with a darker shade representing a more frequently taken route. In b), station numbers are shown alongside their location.

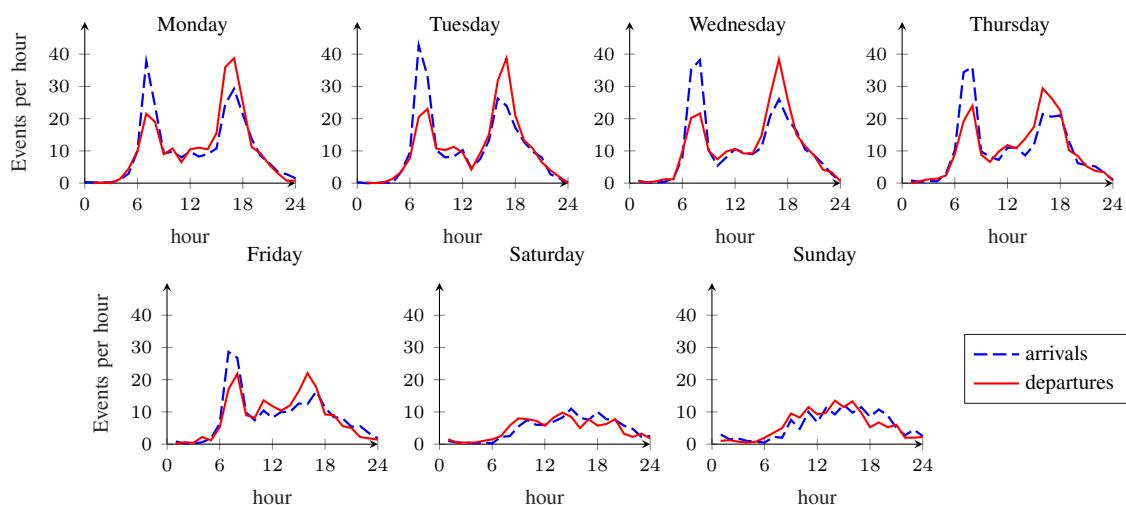


Fig. 2. NY CitiBike. Average traffic patterns for the month of July 2015 at station 537 (Lexington Ave. and East 24th St.)

many people leave work early, and the demand is much lower during the weekend. A complete analysis of the demand for each station, taking into account both geographical and time-based patterns, can provide more meaningful insights on the behavior of the system and of the citizens that use it, allowing city planners to further improve the service.

Similar considerations can be drawn for the Padova system (Fig. 1(b)), whose smaller size allows a more detailed analysis of the flows: in this case, the busiest stations are those in the city center and the University areas (#6-Marzolo, #7-Morgagni, #10-Antenore) and the train and bus station (#15-StationA, #25-StationB); the three most traveled routes are #25→#6, #8→#6, and #6→#8. The usage of the system increases from August to October, then remains constant until

June, and then decreases again in July and August; two breaks occur around Christmas (for two weeks) and Easter (for one week). Moreover, the average number of transactions considering all the stations during weekdays is 542 (502, 572, 571, 566, and 497 from Monday to Friday) while during weekends it drops to 190 (226 on Saturday and 154 on Sunday). This analysis allows us to state that University students are the main user base of the GoodBike system (Padova is, indeed, a college town with over 60 thousand University students over a total population of 220 thousand citizens). This has a strong impact on the policy employed to rebalance the system, by adopting a strategy that is based on the predictable behavior of University students along specific routes. In fact, it can be observed (Fig. 3) that the service demand dynamics during

daily hours is characterized at 8-9 AM by high pick-ups at the train stations and high drop-offs in the University area while in the afternoon the behavior is exactly the opposite.

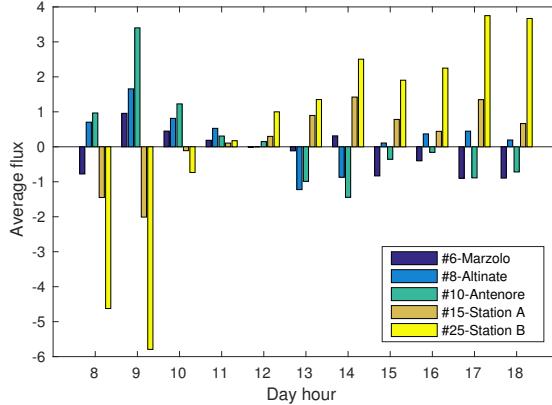


Fig. 3. PD GoodBike. Balance between pick-ups and drop-offs at the most employed bike stations of the train station and the University areas.

Another meaningful aspect to consider is the characteristics of the bike-sharing system travels with respect to time duration and spatial distance, in relation to the pricing policy and the integration with other transportation systems. In particular, the distribution of the duration of trips in both systems follows a consistent pattern: Fig. 4 shows that most trips take less than 10 minutes. The Padova system has slightly longer trips, with a significantly lower fraction of trips under 5 minutes; this could be due to the denser distribution of stations in New York City. Most of the trips in both bike-sharing systems cover less than 2 km, which is consistent with the duration distribution and with the “last mile” model of bike sharing usage.

Indeed, one important aspect that has often been overlooked by the research is multi-modal transportation: as bike sharing can complement mass transit, we expect the location of bus stops and subway stations and their timetables to affect the bikes usage. Fig. 5 shows the relation between the distance from bike-sharing stations to the closest subway entrance and the average hourly demand at those stations for the CitiBike system. The extreme outlier in the upper left of the plot is the bike-sharing station right outside Grand Central Station: it is the station with the highest average demand, which fits the pattern of usage of bike sharing as a complementary service to mass public transport.

Weather also has an impact on the demand of bikes: predictions of future bike traffic can become more accurate by including weather patterns [27]. Table I lists the effect of temperature, rain, and snow on the daily bike traffic city-wide. The cross-correlation between temperature and traffic, expressed using the  $R^2$  score [28] is particularly strong, as users tend to bike more during the warmer months. Rain does not affect the system too heavily, as the demand only drops by 10% on rainy days; snow has a stronger deterrent effect on bikers, as it makes it more difficult to ride. In order to account for the combined effect of snow and low temperature, we only considered the winter months (from December to March), and the demand still dropped by almost half on snowy days and

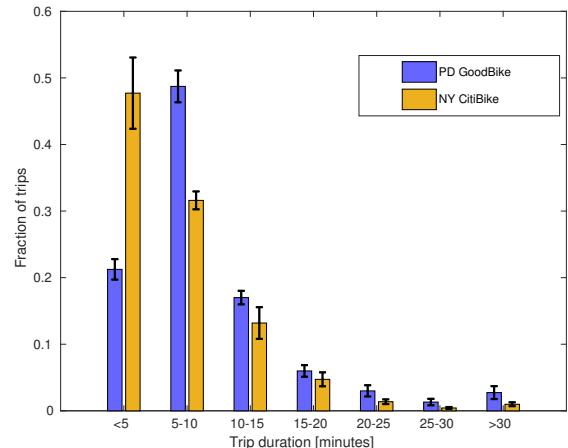


Fig. 4. Distribution of the average time duration of the trips for both systems.

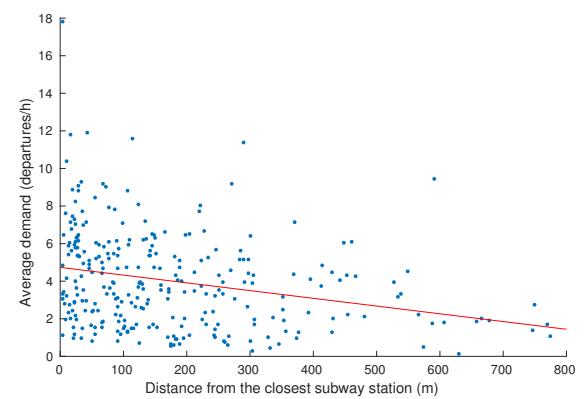


Fig. 5. NY CitiBike. Scatter plot of the average demand by station, plotted against the distance from the closest subway entrance.

by more than a third when there was snow on the ground.

Interestingly, the Padova data show a completely different seasonal pattern: as we remarked above, the service is mostly used by students, resulting in a huge drop in the demand during the summer months. For this reason, the correlation between temperature and bike sharing demand is extremely weak. While the response to rain is slightly stronger than in New York, the numbers are essentially similar, and since snow is extremely rare in Padova, we did not consider it as a factor due to the lack of meaningful data.

As a final remark, it is important to highlight that the available data does not always represent the true demand patterns, because of the *censoring problem*: stations that are empty or full prevent customers from taking or returning bikes, respectively, and this unsatisfied demand is not included in the historical data. The optimization of the system is an ongoing process, as improving the service and reducing unsatisfied demand mitigates the censoring problem for future data.

#### B. Building models from the data

Clearly, as the results discussed in Section III-A show, an extensive analysis of these data allows for a better understanding of the users’ behavior and the optimization of the bike-

TABLE I  
EFFECTS OF WEATHER VARIABLES ON THE DAILY BIKE DEMAND

WEATHER PATTERN	$R^2$		DEMAND DROP	
	NY	PD	NY	PD
Temperature	0.45	0.01	—	—
Rainfall	0.06	0.08	10%	15%
Snowfall*	0.13	—	49%	—
Snow on the ground*	0.10	—	36%	—

\* In the case of snow and snowfall, the drop in the demand is calculated only on the winter months (from December to March)

sharing service, but is not the only way to attain an increase in the system performance. Indeed, an improvement of the Quality of Service (QoS) can also be reached through the synthesis of a model able to provide predictive information.

A discrete time mathematical model of the service can be synthesized starting from a set of observed transitions in a chosen interval, from which a transition estimation matrix can be inferred as  $\mathbf{N}(t)$ , where the entry  $N_{i,j}(t)$  refers to the number of bicycles that move from station  $j$  to station  $i$  in the time unit. Hence, the number of bicycles being picked up from and dropped off at the different stations is given by  $n_i(t) = \sum_j N_{i,j}(t)$ . Let  $x_i(t)$  be the number of bicycles available at station  $i$  at time  $t$ . A dynamical model can be written as

$$x_i(t+1) = x_i(t) + n_i(t). \quad (1)$$

By introducing a state vector  $\mathbf{x}$  for the whole system as the set of all  $x_i$  and defining all variables accordingly, the model takes the following form

$$\mathbf{x}(t+1) = \mathbf{x}(t) + \mathbf{n}(t) + \mathbf{u}(t), \quad (2)$$

where in addition to the dynamics induced by the users behavior  $\mathbf{n}$ , a term  $\mathbf{u}$  is inserted as a generic control input that can be seen as the action of the system rebalancing policy. Obtaining a representative model induces two distinct problems: the *identification* problem consists in inferring a suitable time-varying  $\mathbf{N}(t)$  matrix that describes the time evolution of the system with no data overfitting, whilst the *design* problem requires to choose an appropriate control strategy  $\mathbf{u}$  to reach suitable QoS performance and user satisfaction. An evolution example of such model is given in Fig. 6: a predicted availability at the station is shown by considering a model based on an identification horizon of 8 days and a dynamics constrained between the empty and the full stall conditions.

Interestingly, this model serves several purposes: to produce scenario simulations able to describe a predictive evolution of the system and optimize the initial distribution of bicycles among all the stations; to simulate different rebalancing strategies within an if-then approach; to dynamically adapt the rebalancing strategy according to the prediction results.

However, for large bike-sharing systems like New York's CitiBike, it is impractical to account for the relations among all stations. In that case, a powerful and validated approach consists in modeling each station separately as a Markov Modulated Poisson Process. The occupancy of station  $i$  follows a finite Markov Birth-Death Process (BDP)  $m_i(t)$ , where

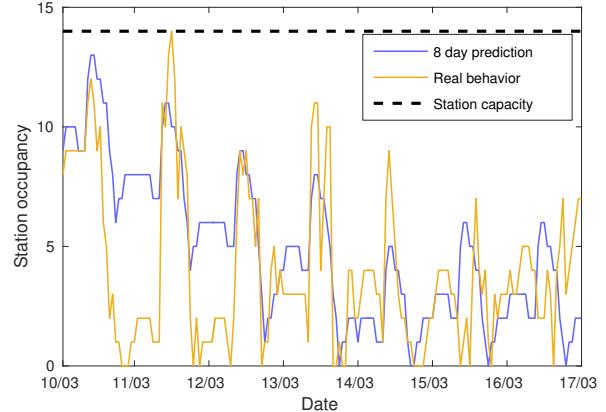


Fig. 6. PD GoodBike. Prediction of bicycle availability at the station based on a dynamical model identified over an 8-days horizon (orange solid line) compared with the true values (blue solid line). The station saturation limit is shown as a black dashed line.

the birth and death processes are Markov-modulated Poisson processes whose time-varying rates can be inferred from the historical data. The BDP is limited by the station's capacity. This model has been validated against real datasets and the approximation of independence among stations is not restrictive [29]. Moreover, it is a useful tool in the planning of rebalancing operations (see Sec. IIc), that allows to characterize the behavior of each station's state and estimate when external intervention is needed and to which extent [8].

#### IV. FINAL REMARKS

In this work, we presented some of the current trends in the development of bike-sharing systems in Smart Cities, showing an ever tighter integration between operations, planning and data analysis. We presented some of the directions the research in the field is taking, and introduced New York's CitiBike and Padova's GoodBike as two significant examples of bike-sharing systems of different sizes and with different user bases.

Although CitiBike has already started implementing some of the techniques suggested by the research community, particularly on rebalancing [15], users' satisfaction is still not ideal, with more than a third of negative reviews on websites such as TripAdvisor. The service could be further improved by including other types of data, such as weather and subway schedules, in the bike sharing optimization, making more accurate estimates of future demands and rebalancing the system more effectively. The integration of the other techniques we introduced in Sec. II, such as gamification and pricing, in the bike sharing management platform would also make the service more efficient, fitting the Smart City paradigm. Padova's approach to data and optimization is still more rudimentary, and the quality of the service is lower: according to the system's own website, only 3 stations have a rating above 3 stars out of 5, and several stations are below 2. In this case, opening the system to Smart City approaches would be even more beneficial.

Over the next few years, we expect research to focus on exploiting multiple sources of data and working on multiple

issues at once, consolidating bike sharing's position as one of the most successful Smart City-enabled services and integrating different modes of transportation to reduce traffic and make citizens' commutes faster and eco-friendlier.

## REFERENCES

- [1] A. Zanella, N. Bui, A. Castellani, L. Vangelista, and M. Zorzi, "Internet of Things for smart cities," *IEEE Internet of Things journal*, vol. 1, pp. 22–32, Feb. 2014.
- [2] P. Midgley, "The role of smart bike-sharing systems in urban mobility," *Journeys*, vol. 2, pp. 23–31, 2009.
- [3] E. Fishman, S. Washington, and N. Haworth, "Bike shares impact on car use: evidence from the United States, Great Britain, and Australia," *Transportation Research Part D: Transport and Environment*, vol. 31, pp. 13–20, Aug. 2014.
- [4] D. Rojas-Rueda, A. de Nazelle, M. Tainio, and M. J. Nieuwenhuijsen, "The health risks and benefits of cycling in urban environments compared with car use: health impact assessment study," *British Medical Journal*, vol. 343, p. d4521, 2011.
- [5] S. Jäppinen, T. Toivonen, and M. Salonen, "Modelling the potential effect of shared bicycles on public transport travel times in Greater Helsinki: An open data approach," *Applied Geography*, vol. 43, pp. 13–24, Sep. 2013.
- [6] P. Vogel, T. Greiser, and D. C. Mattfeld, "Understanding bike-sharing systems using data mining: Exploring activity patterns," *Procedia-Social and Behavioral Sciences*, vol. 20, pp. 514–523, 2011.
- [7] M. Kaspi, T. Raviv, and M. Tzur, "Detection of unusable bicycles in bike-sharing systems," *Omega*, vol. 65, pp. 10–16, Dec. 2016.
- [8] J. Schuijbroek, R. C. Hampshire, and W.-J. Van Hoeve, "Inventory rebalancing and vehicle routing in bike sharing systems," *European Journal of Operational Research*, vol. 257, pp. 992–1004, Mar. 2017.
- [9] L. Caggiani and M. Ottomanelli, "A dynamic simulation based model for optimal fleet repositioning in bike-sharing systems," *Procedia-Social and Behavioral Sciences*, vol. 87, pp. 203–210, Oct. 2013.
- [10] D. Chemla, F. Meunier, and R. W. Calvo, "Bike sharing systems: Solving the static rebalancing problem," *Discrete Optimization*, vol. 10, pp. 120–146, May 2013.
- [11] 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*, vol. 2, pp. 187–229, Aug. 2013.
- [12] 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*, vol. 71, pp. 230–247, Jan. 2015.
- [13] J. Brinkmann, M. W. Ulmer, and D. C. Mattfeld, "Short-term strategies for stochastic inventory routing in bike sharing systems," *Transportation Research Procedia*, vol. 10, pp. 364–373, Jan. 2015.
- [14] C. Kloimüller, P. Papazek, B. Hu, and G. R. Raidl, "Balancing bicycle sharing systems: an approach for the dynamic case," in *European Conference on Evolutionary Computation in Combinatorial Optimization*. Springer, Apr. 2014, pp. 73–84.
- [15] E. O'Mahony and D. B. Shmoys, "Data Analysis and Optimization for (Citi) Bike Sharing," in *29th AAAI Conference on Artificial Intelligence*, Jan. 2015, pp. 687–694.
- [16] A. Waserhole and V. Jost, "Pricing in vehicle sharing systems: Optimization in queuing networks with product forms," *Euro Journal on Transportation and Logistics*, vol. 5, pp. 293–320, Aug. 2016.
- [17] C. Fricker and N. Gast, "Incentives and redistribution in homogeneous bike-sharing systems with stations of finite capacity," *Euro Journal on Transportation and logistics*, vol. 5, pp. 261–291, Aug. 2016.
- [18] R. Kazhamiakin, A. Marconi, M. Perillo, M. Pistore, G. Valetto, L. Piras, F. Avesani, and N. Perri, "Using gamification to incentivize sustainable urban mobility," in *IEEE First International Smart Cities Conference (ISC2)*, Oct. 2015, pp. 1–6.
- [19] J. Larsen, Z. Patterson, and A. El-Geneidy, "Build it, but where? the use of geographic information systems in identifying locations for new cycling infrastructure," *International Journal of Sustainable Transportation*, vol. 7, no. 4, pp. 299–317, Jun. 2013.
- [20] J. C. García-Palomares, J. Gutiérrez, and M. Latorre, "Optimizing the location of stations in bike-sharing programs: a GIS approach," *Applied Geography*, vol. 35, no. 1, pp. 235–246, Nov. 2012.
- [21] G. Rybarczyk and C. Wu, "Bicycle facility planning using GIS and multi-criteria decision analysis," *Applied Geography*, vol. 30, no. 2, pp. 282–293, Apr. 2010.
- [22] A. Faghhih-Imani, N. Eluru, A. M. El-Geneidy, M. Rabbat, and U. Haq, "How land-use and urban form impact bicycle flows: evidence from the bicycle-sharing system (BIXI) in Montreal," *Journal of Transport Geography*, vol. 41, pp. 306–314, Dec. 2014.
- [23] D. Buck and R. Buehler, "Bike lanes and other determinants of Capital bikeshare trips," in *91st Transportation research board annual meeting*, vol. 12, Jan. 2012, pp. 3539–3550.
- [24] I. Mateo-Babiano, R. Bean, J. Corcoran, and D. Pojani, "How does our natural and built environment affect the use of bicycle sharing?" *Transportation Research Part A: Policy and Practice*, vol. 94, pp. 295–307, 2016.
- [25] J. Bao, T. He, S. Ruan, Y. Li, and Y. Zheng, "Planning bike lanes based on sharing-bikes' trajectories," in *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 2017, pp. 1377–1386.
- [26] F. Chiariotti, M. Condoluci, T. Mahmoodi, and A. Zanella, "SymbioCity: Smart cities for smarter networks," *Transactions on Emerging Telecommunications Technologies*, vol. 29, no. 1, pp. e3206–n/a, Jan. 2018.
- [27] T. H. T. Nguyen, T. T. Le, T. P. D. Chu, L. T. Nguyen, and V. H. Le, "Multi-source data analysis for bike sharing systems," in *International Conference on Advanced Technologies for Communications (ATC)*, Oct. 2017.
- [28] H. E. Brodgen, "On the interpretation of the correlation coefficient as a measure of predictive efficiency," *Journal of educational psychology*, vol. 37, no. 2, p. 65, Feb. 1946.
- [29] M. Dell'Amico, E. Hadjicostantinou, M. Iori, and S. Novellani, "The bike sharing rebalancing problem: Mathematical formulations and benchmark instances," *Omega*, vol. 45, pp. 7–19, Jun. 2014.