Università
degli Studi
di Padova

Dipartimento
di Ingegneria
dell'Informazione

Master Thesis in Control Systems Engineering

Reinforcement Learning approaches for fair mobility as a service

Master Candidate

Luca Vittorio Piron

Student ID 2056320

Supervisor

Prof. Gian Antonio Susto

University of Padova

Co-supervisors

Marina Ceccon,

Matteo Cederle,

Federico Chiariotti,

Marco Fabris

University of Padova

Alessandro Fabris

Max Planck Institute for Security and Privacy

Academic Year
2023/2024

Abstract

This thesis explores Reinforcement Learning (RL) approaches to enhance fairness in Shared Micromobility Services, with a case study based on bike-sharing systems. Traditional sharing models often grapple with the conflict between maximising profitability and providing equitable service coverage. This research aims to address this challenge by integrating RL to optimise the distribution and availability of bikes, ensuring accessibility across diverse urban areas, including underserved communities. The core of the thesis involves developing and testing a model that can adjust bike distribution in response to data on usage patterns, demands and geography. By leveraging RL techniques, the study not only predicts high-demand areas but also proactively manages resource allocation to balance profit and fair coverage. Empirical evaluations are conducted using a simulated environment based on real-world urban layouts and usage data. These experiments demonstrate the efficacy of RL in balancing the tradeoff between service fairness and profitability. The thesis contributes to the growing body of knowledge in applying RL aimed at fairness, particularly in urban planning and sustainable transportation. It offers a perspective on longstanding issues for more responsible urban mobility solutions.

Sommario

Questa tesi esplora gli approcci di Reinforcement Learning (RL) per migliorare l'equità nei Servizi di Micromobilità Condivisa, con caso studio basato sui sistemi di bike-sharing. I modelli di condivisione tradizionali spesso affrontano il conflitto tra la massimizzazione della redditività e la fornitura di una copertura equa del servizio. Questa ricerca mira a risolvere questa sfida integrando il RL per ottimizzare la distribuzione e la disponibilità delle biciclette, garantendo l'accessibilità in diverse aree urbane, inclusi i quartieri meno serviti. Il nucleo della tesi coinvolge lo sviluppo e il test di un modello che può regolare la distribuzione delle biciclette in risposta ai dati sui modelli di utilizzo e la domanda. Sfruttando le tecniche di RL, lo studio non solo predice le aree ad alta domanda, ma gestisce anche proattivamente l'allocazione delle risorse per bilanciare profitto e copertura equa. Le valutazioni empiriche sono condotte utilizzando un ambiente simulato basato su layout urbani reali e dati di utilizzo. Questi esperimenti dimostrano l'efficacia del RL nel bilanciare il compromesso tra equità del servizio e redditività. La tesi contribuisce al crescente corpo di conoscenze nell'applicazione del RL mirato all'equità, in particolare nella pianificazione urbana e nei trasporti sostenibili. Offre una prospettiva su questioni di lunga data per soluzioni di mobilità urbana più responsabili.

Contents

List of Figures

List of Tables

List of Algorithms

List of Acronyms

RL Reinforcement Learning

MSS Micromobility Sharing System

MDP Markov Decision Process

MaaS Mobility As a Service

MBDP Markovian Birth-Death Process

MMPP Markov Modulated Poisson Process

1 Introduction

1.1 Sustainable mobility, fairness, and artificial intelligence

With the world's latest advances and transformations, the necessity for a shift in the conversation towards how control systems can meet grand societal-scale challenges has grown [3, 50]. Particularly, over the past decade, Micromobility Sharing System (MSS) have become a key component of integrated urban transit [11], providing last mile service that complements with mass transit [40]. The quick development of micromobility has had a significant impact on pollution, with a reduction of tens of thousand tons of CO_2 emissions in New York City alone [9]. This boom has led to a significant interest in rebalancing techniques [17], i.e., schemes to move the shared vehicles (usually bicycles or scooters) from where the users leave them after taking a trip to where they are needed. Rebalancing represents the most significant cost for MSS operators, and needs to consider imbalances in demand patterns and traffic limitations for the trucks that physically transport the vehicles [13]. The latter method is also known in literature as *static rebalancing*. Additionally, the differences in the demand patterns for different vehicles need to be taken into account, as bike and scooter sharing services may cater to different needs and require specific solutions [54].

Despite the fast-paced growth of sharing services, recent research and advocates have highlighted a significant problem of most MSS, which can limit the potential benefits of the MaaS approach: like other forms of transportation, bikes and scooters are significantly more available in richer areas, excluding poorer communities [43]. This effect is compounded by the higher density of central areas, which makes them naturally more attractive for this kind of service, but also tends to starve suburban neighborhoods, and by the lower subscription rate from working class users, for whom the cost of the MSS is unaffordable [27]. A recent review of the MSS literature has also confirmed a significant gender dimension [19]: the population of bike sharing users tends to be disproportionately male, as well as skewing towards younger, well-educated professionals and university students. Accordingly, the most cost-efficient approach to designing a MSS will tend to cater to this demographic [21], leading to further iniquities due to the early adopter effect [20].

The rise of dockless MSS, in which bikes or scooters can be parked anywhere, has significantly reduced the iniquities inherent in dock-based systems [36] and increased availability, but there is still a significant issue in the distribution of bikes, with a recent study [38] finding a difference of 2 orders of magnitude in the density of bicycles available in different neighborhoods of the same city. The same type of difference was noted for scooters [45], which exhibit even higher levels of iniquity in their distribution. The difference between the planned and real service geographies of dockless systems, which was highlighted in [36], is generated by the combined effect of usage and service rebalancing policies that tend to cater to privileged users, often for the same cost-efficiency reasons.

We can then note a similarity between the equity issues in MSS operation and a well-known problem of ML solutions, which may exhibit bias in their performance due to issues with their training sets [4]. For this reason, the research on *fairness* in ML has focused on methods to identify and mitigate such bias, aiming to create fair models. The MSS equity problem is related to the concept of spatial fairness [21, 44], which relates to the idea of uniform allocation of resources in the considered area. As in most fairness scenarios, there is a fundamental trade-off. On one hand, optimizing the satisfaction of the expected demand usually implies concentrating the vehicles in popular areas with denser (and often richer) populations. On the other hand, a fair allocation of vehicles would require to equally distribute them in all areas, with a lower performance for the same cost [10].

While satisfying users in high-demand areas is required to attract as many users as possible, creating an unfair system has strong societal impacts, as it limits the accessibility of shared vehicles for people living in peripheral and disadvantaged areas. This can further enforce the discrimination of minority groups, since the area in which an individual lives often correlates with their ethnicity and economic status. Furthermore, individuals facing significant limitations in accessibility options may encounter challenges in engaging in everyday activities such as work, education, and leisure activities, thereby potentially exacerbating their exclusion from essential social functions [10].

For these reasons, our work focuses on analyzing this trade-off in dockless MSS, considering a RL scheme that can be tuned to consider the spatial fairness of the system. The contributions brought by this thesis can be summarized as follows.

- We propose a simplified fairness-aware MSS simulator, by clustering the areas into three categories: central, peripheral, and remote. Further details about this are given in Section 5;
- Through Monte Carlo simulations, we reveal the presence of an inherent trade-off between the MSS performance and the associated fairness level obtained by applying a parametric family of RL-based strategies; more precisely, at least a 7.5% improvement in terms of the Gini index can be achieved w.r.t. suboptimal strategies;
- We present a comprehensive analysis of the trade-off between fairness and performance in MSS operation, comparing the performance-oriented strategies that dominate the literature to a novel fairness-based approach that can be used to control the trade-off and define the relative importance of fairness, rebalancing cost, and disservice for the system users with simple parameters;
- While the abovementioned works deal with fairness in system planning, to the best of our knowledge, this is the first work on fairness in MSS operation and rebalancing.

The remainder of this manuscript unfolds as follows. Chapters 2, 3 and 4 cover the required preliminaries, addressing Micromobility Sharing Systems, Fairness and RL, respectively; Sections 5 and 6 delve into the proposed approach by providing a formulation of the problem under investigation and illustrating the RL-based solution. As a further support to the theoretical findings, Section 6 reports on a significant case study and examines the corresponding fairness achievements. Lastly, final remarks and future directions are briefly sketched in Section 7.

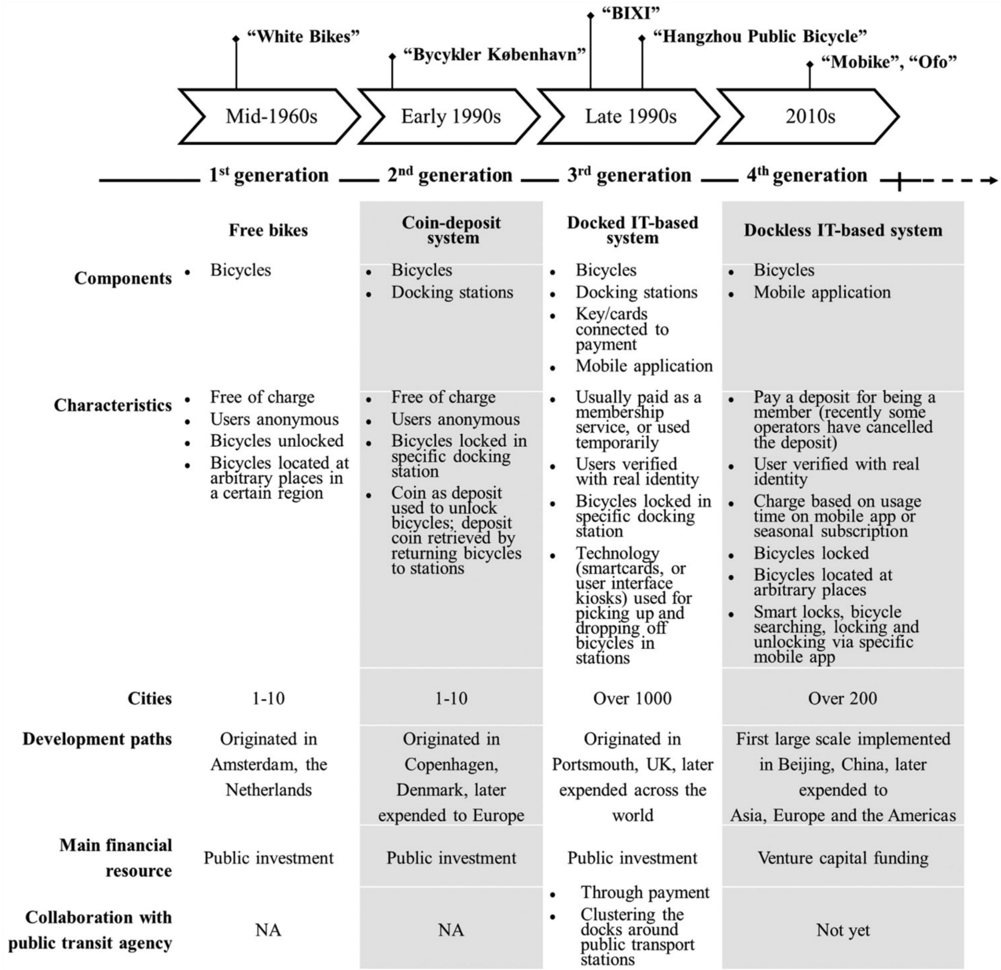
2

Micromobility Sharing Systems

2.1 The role of Micromobility Sharing Systems in urban transportation

The pervasive challenge of transitioning towards a more sustainable economy spans all sectors, from aiming to diminish fossil fuel dependency to controlling HVAC systems efficiently, urging all of society to consider their environmental footprint. This transition mandates a strategy that emphasizes not only technological advancements and the adoption of eco-friendly solutions but also regulatory frameworks and financial incentives. In this context, mobility plays a critical role due to its substantial contribution to emissions, urban pollution, and environmental degradation. Consequently, transforming transportation systems to embrace less polluting alternatives is crucial to mitigate mobility's environmental impact. The advent of digitization, Internet of Things (IoT), and smart city innovations have allowed policymakers and engineers to introduce new services designed to address commuting challenges. Among these solutions, bike-sharing and e-scooter sharing services constitute a modern, sustainable, and efficient transportation network. A typical motorized passenger vehicle emits several metric tons of carbon dioxide annually, thus cities worldwide have integrated these programs to cut emissions, reduce energy consumption, and protect the environment. Moreover, these programs facilitate the integration (first mile and last mile gap) with other transportation modes. Through its expansion over the past two decades, from Beijing to Buenos Aires, sharing systems have provided the option for a healthier, more economical, and environmentally friendly commuting alternative, significantly contrasted with traditional public transportation systems and the environmental costs of individual car usage. In the evolving landscape of urban transportation, the distinction between Mobility as a Service (MaaS) and Micromobility Sharing Systems is at the core of understanding the broader goal of redefining urban mobility for inclusivity and equal access. While both concepts aim to enhance urban transportation, they operate at different scales and with distinct objectives. MaaS represents an integrated platform that amalgamates various transportation services, including public transit, ride-hailing, car-sharing, and micromobility options like bike-sharing and scooter-sharing, into a cohesive digital ecosystem. This integration facilitates seamless travel for users across different modes of transportation, driven by the convenience of a single-access interface and unified payment system. On the other hand, Micromobility Sharing Systems specifically refer to the services that provide access to lightweight vehicles such as bicycles, e-bikes, and e-scooters for short-distance travel, often complementing the broader transportation network. The primary distinction lies in the scope and scale of the services they offer. MaaS is holistic, aiming to offer a comprehensive solution to urban mobility that encompasses a wide range of transportation options, thus addressing the diverse needs of urban dwellers. Micromobility Sharing Systems, while integral to the MaaS ecosystem, focus more narrowly on solving the challenges of last-mile connectivity and providing an eco-friendly alternative to personal vehicle use and public transit for shorter urban trips. By integrating micromobility services within the MaaS framework, cities can enhance the versatility and accessibility of their transportation networks, making urban mobility more adaptable to individual preferences and environmental sustainability.

First thought of in the mid-20th century, Micromobility Sharing Systems owe their more recent rapid expansion to a series of factors: technology, environmental awareness and a rising interest in healthy lifestyles [41]. An early attempt at implementation was made in Amsterdam with the "White Bikes" program in 1965, which faced challenges due to theft and vandalism [16]. Also, this first phase involved manual operations and limited scalability, but with the advent of later technology, the early 2000s allowed for the introduction of automated Bike-Sharing Systems. These used electronic docking stations and smart cards for access, significantly improving the user experience and system management. Such advancements facilitated the rapid expansion of BSSs across Europe and



	Free bikes	Coin-deposit system	Docked IT-based system	Dockless IT-based system
Components	• Bicycles	• Bicycles • Docking stations	• Bicycles • Docking stations • Key/cards connected to payment • Mobile application	• Bicycles • Mobile application
Characteristics	• Free of charge • Users anonymous • Bicycles unlocked • Bicycles located at arbitrary places in a certain region	• Free of charge • Users anonymous • Bicycles locked in specific docking station • Coin as deposit used to unlock bicycles; deposit coin retrieved by returning bicycles to stations	• Usually paid as a membership service, or used temporarily • Users verified with real identity • Bicycles locked in specific docking station • Technology (smartcards, or user interface kiosks) used for picking up and dropping off bicycles in stations	• Pay a deposit for being a member (recently some operators have cancelled the deposit) • User verified with real identity • Charge based on usage time on mobile app or seasonal subscription • Bicycles locked • Bicycles located at arbitrary places • Smart locks, bicycle searching, locking and unlocking via specific mobile app
Cities	1-10	1-10	Over 1000	Over 200
Development paths	Originated in Amsterdam, the Netherlands	Originated in Copenhagen, Denmark, later expended to Europe	Originated in Portsmouth, UK, later expended across the world	First large scale implemented in Beijing, China, later expended to Asia, Europe and the Americas
Main financial resource	Public investment	Public investment	Public investment	Venture capital funding
Collaboration with public transit agency	NA	NA	• Through payment • Clustering the docks around public transport stations	Not yet

Figure 2.1: Bike Sharing Systems Over Time. Chen, van Lierop & Ettema (2020) Dockless bike- sharing systems: what are the implications?, Transport Reviews

Asia, showcasing the potential for large-scale implementation. Indeed, in the 20 years from 2004 to 2024, there was an increase in the number of BSSs from 13 to nearly 3000 all over the world, differing in size and efficiency [23]. The global adoption of bike-sharing systems and more recently of e-scooter sharing systems was propelled by the integration of GPS technology and mobile applications. This period saw the rise of dockless sharing systems, which allowed users to pick up and drop off vehicles anywhere within a designated area, enhancing convenience and flexibility [37]. The dockless model significantly lowered the infrastructure costs associated with dock-based systems and enabled rapid expansion into new markets. Dockless sharing systems have dramatically transformed micromobility sharing markets by offering unprecedented convenience and ease of use. Their rapid growth highlights the significant shift towards more accessible and user-friendly modes of urban transportation. However, despite the evident benefits, dockless sharing systems do pose some challenges, including rebalancing and equitable access to the service. These will be discussed in detail throughout this document, as they are two of the core topics of this research.

More recently, research on the impact of the COVID-19 pandemic on bike-sharing usage patterns highlighted the system's adaptability and resilience [51]. Indeed, after the 2020 pandemic, which upended traffic patterns all over the globe, their use saw significant changes. The pandemic abruptly changed established mobility patterns, as the need for social distancing and lockdown orders drove citizens to reduce their movements and avoid crowded mass transit. Other research focused on New York City's bike-sharing system, one of the largest in the world, and revealed important trends during the lockdown, indicating that bike-sharing systems played a critical role in adapting to the new norms of urban mobility, demonstrating resilience and flexibility in response to the pandemic's challenges [15].

The impact of the COVID-19 pandemic on shared e-scooter systems, mirroring the resilience observed in bike-sharing schemes, showcases a significant adaptation in urban micromobility. [18]. A spatial panel model analysis from Austin, Texas, detailed how e-scooter usage changed during the pandemic, with a significant reduction in overall trips but an increase in the average distance per trip, highlighting how the pandemic reshaped mobility patterns and the role of e-scooters within them [34]. These findings are coherent with those shown for bike sharing systems and confirm the adaptability of sharing systems to changing urban mobility needs, positioning them as a resilient component of the transportation ecosystem.

The paper titled "A global comparison of bicycle sharing systems" provides a comprehensive analysis of bicycle sharing systems (BSS) across the globe, focusing on 322 schemes located on all continents [49]. It identifies five main types of BSS based on usage, contextual indicators, and user behavior: very large, high-use BSS; large BSS in major cities; medium BSS with extensive cycling infrastructure; small to medium efficient BSS; and small to medium inefficient BSS. This classification allows for a detailed comparison of BSS performance and offers a foundation for new schemes to identify similar established BSS as templates for anticipating user demand.

The paper highlights the importance of BSSs as a solution to urban congestion and as a mode of active transport that reduces emissions. The research emphasizes the need for global comparisons of BSS, especially since previous studies often exclude China, the largest BSS market.

Through a two-staged clustering analysis, the study creates a framework for comparing BSSs globally. It finds significant differences in BSS types, with varying levels of efficiency and usage patterns. The paper also notes the impact of socio-demographic factors, weather, climate, topography, and cycling infrastructure on BSS use. It suggests that cities planning to implement or improve BSS can learn from more efficient systems.

This research is notable for its use of a large and diverse dataset and contributes to a deeper understanding of how BSSs can be optimized to serve urban populations effectively and sustainably.

Similarly, a systematic study of ESS also reveals diverse operational models and user interactions across various urban settings. [28] This research categorizes ESS into distinct types based on scale, usage intensity, and integration with public transport networks, offering insights into their role in reducing urban congestion and emissions. Key findings emphasize the versatility of ESS in complementing existing transportation infrastructures, highlighting their potential to enhance first and last-mile connectivity. Moreover, the study investigates the influence of socio-demographic factors, urban layouts, and climate conditions on ESS adoption and usage patterns. By leveraging global data, the analysis provides a framework for cities considering the introduction or expansion of ESS, considering the importance of adapting strategies to local contexts for maximizing the benefits of this emerging mode of micromobility.

2.2 Technology used in Micromobility Sharing Systems

As previously mentioned, the recent success of micromobility sharing systems is largely driven by technology.

GPS (Global Positioning System) stands out as a cornerstone technology, offering multiple benefits. First, it enables efficient fleet management by allowing operators to monitor the location and status of each bike or e-scooter in the fleet. This capability facilitates the management of maintenance schedules, prediction of demand, and ensures vehicles are evenly distributed across the service area. Furthermore, GPS technology enhances recovery efforts. The ability to track bikes or e-scooters in real time significantly mitigates the risks associated with theft and loss, thereby ensuring they remain a reliable resource for the community. Lastly, the data-driven insights obtained from GPS tracking prove invaluable for urban planners and system operators. Analyzing travel patterns and bike usage yields informed decisions regarding service expansion or reduction, placement of new bike stations, and overall improvement in the design of the bike-sharing network to more effectively meet the community's mobility needs.

Mobile Apps also play a critical role in the user engagement and operational efficiency of sharing systems. They facilitate immediate access by enabling users to locate, unlock, and pay for rentals directly from their smartphones, significantly simplifying the rental process and lowering the barriers to usage. Beyond these basic functionalities, mobile apps introduce user-centric features that add substantial value. These can include route planning to identify the most bike-friendly paths, social sharing options for a communal experience, and fitness tracking that promotes healthy lifestyle choices by quantifying metrics like calories burned and distance traveled.

Moreover, these apps directly connects users and service providers. This allows users to effortlessly report issues, share feedback, or seek assistance, thereby boosting user satisfaction and enabling operators to swiftly address concerns, which in turn enhances the quality of service. Customization and personalization features are another significant benefit of apps. By analyzing user data, these apps can provide personalized recommendations and tailor notifications to fit individual preferences, thus enriching the user experience and encouraging repeated engagement with the service.

Another transformative force in the micromobility realm has been Big Data Analytics, offering profound insights that benefit both operators and users alike.

Among the key advantages is demand prediction. By analyzing usage patterns, BSS operators can pinpoint demand hotspots, ensuring bicycles are available precisely where and when needed. This is crucial for the success of dynamic rebalancing strategies, allowing for optimal resource allocation across the system.

Maintenance and operations optimization also benefits significantly from data analytics. Trends in bicycle wear and tear can be identified, enabling operators to prioritize maintenance tasks and reduce service downtime. Strategic optimization of bike distribution and rebalancing, informed by real-time data analysis, further enhances operational efficiencies.

Another area where big data analytics shines is in user experience. Insights gained from analyzing how users interact with the system lead to continuous improvements in BSS offerings. This can include developing more intuitive app interfaces, crafting targeted promotional activities, and boosting service reliability, all of which elevate the overall user experience. Additionally, the insights derived from BSS operations offer value for urban planning insights. Data on preferred cycling routes, peak service demands, and popular destinations provide city planners with the evidence needed to make informed decisions on infrastructure development.

A last technological development worth considering are batteries, which are central for e-bikes and e-scooters as the affect their efficiency, range, and overall usability. Recent advancements in battery technology have significantly improved the energy efficiency of these micromobility vehicles, enabling them to cover longer distances on a single charge, thus making them a reliable option for extended commutes and leisure activities. One of the standout features of modern batteries is their fast-charging capabilities, which minimize downtime and increase convenience for users, especially crucial for shared mobility systems where quick turnaround is key to meeting user demand. The durability and lifecycle of batteries have also seen considerable enhancements, leading to fewer replacements, lower maintenance costs, and a diminished environmental impact over the vehicle's operational life. As battery technology continues to evolve, it remains a central force in shaping the future trajectory of e-bikes and e-scooters, offering enhanced performance, user convenience, and a step forward towards sustainability in urban mobility.

2.3 Docked vs. Dockless Systems

The evolution of micromobility-sharing systems from docked to dockless formats represents a significant shift in urban mobility strategies. Each model presents distinct features, benefits, and challenges that influence the urban environment, user interaction, and the broader transportation ecosystem. This section will illustrate the main differences between the two systems, with a focus on Bike Sharing Systems, given its relevance in existing literature. However, with the exception of how batteries are integrated into the service, most of the considerations are applicable to both e-bikes and e-scooters too.

Docked Systems

Docked bike-sharing systems, characterized by their reliance on fixed stations for the pickup and return of bicycles, are strategically positioned across service areas to maximize accessibility and convenience. These systems are particularly prevalent near transit hubs and areas with high pedestrian traffic, smoothly integrating into urban infrastructure to augment public transportation networks. They offer a reliable and orderly service model, with predetermined locations for accessing and returning bicycles, which helps in maintaining the cleanliness and order of urban spaces by preventing the random dispersal of bikes that could obstruct sidewalks or public areas. The stationary nature of docked systems also facilitates streamlined fleet maintenance and rebalancing, as the specific locations of all bicycles are consistently monitored and managed. However, the establishment of docking stations involves considerable initial costs and spatial requirements, which can hinder the rapid deployment or expansion of docked BSSs in areas with limited resources or space. Furthermore, the necessity for users to return bikes to docking stations may reduce the system's convenience, especially if stations are not located near users' destinations. This limitation potentially detracts from the appeal of docked systems for some trips, indicating a trade-off between the orderliness and urban integration of docked systems and the flexibility desired by users.

Dockless Systems

Dockless bike-sharing systems have transformed the way we think about urban mobility, addressing many of the challenges posed by traditional station-based systems by offering increased convenience and accessibility. These systems rely on smartphone apps and GPS to locate bikes, thus eliminating the need for fixed docking stations. This innovation grants users the flexibility to start and end their trips practically anywhere, alleviating space constraints and drop-off restrictions around transit stations and operating more efficiently with the same resources. However, dockless systems face their own set of challenges, including regulation and distribution issues that can lead to irregular parking, oversupply, and, in some regions, competition among operating companies has led to problems with abandoned or damaged bikes. Despite these challenges, the dockless model's advantages, such as the ability to pick up and drop off bicycles anywhere in the service area, significantly enhance convenience, particularly for completing the last mile of a journey, making dockless systems particularly appealing to those seeking spontaneity and straightforward access.

By bypassing the need for physical docking infrastructure, dockless systems can launch and expand with minimal initial financial outlay, facilitating swift scalability and the agility to adapt to changing demand patterns. This is supported by the integration of advanced GPS and mobile technologies for tracking, access, and transactions, enabling comprehensive data collection that is invaluable for refining urban mobility strategies and enhancing system efficiency. Yet, the liberty to park bicycles indiscriminately can lead to congestion and obstructions in public areas, necessitating specific regulations and designated zones for parking dockless bikes to manage urban clutter and maintain pedestrian flow. Additionally, achieving an equitable distribution of bicycles throughout the service area remains a significant challenge, requiring sophisticated algorithms and proactive fleet regulation to prevent potential imbalances.

The openly accessible nature of dockless bicycles also heightens their vulnerability to vandalism and theft, necessitating diligent maintenance and surveillance to ensure continuous service availability and system dependability. Despite these challenges, the integration of dockless bike-sharing systems with public transit can improve travel time and convenience, enhancing overall user satisfaction. However, excessive bikes cluttering streets can deter non-users and hinder public transit accessibility. As dockless systems continue to evolve, better regulation and distribution management are essential to maximize their potential benefits and minimize their drawbacks, ensuring they remain a vital component of urban transportation ecosystems.

Ultimately, the choice between docked and dockless systems depends on a variety of factors including urban infrastructure, investment capabilities, and specific mobility needs of the city. While docked systems offer reliability and orderliness, dockless systems provide flexibility and ease of use. Cities around the world are experimenting with both models, and in some cases, integrating them to further benefit urban transportation networks.

2.4 Dynamic vs. Static Rebalancing

The success of MSS is significantly influenced by the efficiency of their rebalancing strategies, which ensure bikes or e-scooters are available across different service areas to meet user demand. Redistributing vehicles can involve moving bikes usually using trucks or vans operated by the service provider. The goal is to maintain a balanced distribution of bikes or e-scooters throughout the system to optimize availability and usability for users.These strategies are distinguished into dynamic and static rebalancing, each characterized by distinct methodologies, benefits, and challenges. Once again, we present an overview with a focus on BSS while stressing that most if not all considerations remain valid for more diverse MSS.

Static Rebalancing

Static rebalancing in bike-sharing systems relies on fixed schedules and predetermined routes for redistributing bikes throughout the service area, utilizing historical usage data instead of real-time analytics. This method offers predictability and simplicity, providing a clear and manageable system for fleet management and simplifying the planning and execution of redistribution tasks. Moreover, the absence of the need for continuous monitoring and immediate response renders static rebalancing more economical, as it minimizes the resources dedicated to oversight and intervention. Established schedules also ensure regular maintenance and consistent workloads for rebalancing teams, which facilitates routine system upkeep. However, static rebalancing faces challenges, including limited responsiveness to real-time demand fluctuations, which may lead to bike shortages or excesses in certain areas. This fixed approach might not always align with actual demand patterns, potentially leading to inefficiency in resource utilization and missed opportunities to optimize bike availability. Additionally, the inability to promptly adapt to shifting demand trends can result in reduced user satisfaction, as users may experience frustration due to the unavailability of bikes when and where they need them, potentially deterring future use of the system.

Static Rebalancing methods

The research landscape offers a broad spectrum of approaches to static rebalancing, demonstrating the potential of sophisticated computational techniques to address the challenges inherent in this process. Notable developments include the creation of exact algorithms aimed at minimizing the costs associated with rebalancing operations, highlighting the feasibility of achieving optimal solutions for systems encompassing up to 60 stations [22].

Furthermore, the utilization of graphical class diagrams has emerged as a novel approach to simplify the mathematical modeling involved in static rebalancing [31]. This methodology, inspired by the operational intricacies of a prominent Parisian bike-sharing system, showcases the utility of graphical tools in streamlining rebalancing efforts to enhance system efficiency.

Integrating maintenance requirements into rebalancing operations represents another significant advancement, addressing both the immediate availability of bikes and the long-term sustainability of bike-sharing infrastructure [53]. Moreover, the adoption of spatial cluster-based strategies underscores the evolving nature of bike-sharing system (BSS) management, revealing how data analytics can inform more effective rebalancing strategies by examining spatial and temporal patterns [30].

Additionally, the human dimension of BSS becomes evident through research focusing on user behavior in response to bike shortages at stations. Such studies underline the critical role of user engagement in the development of rebalancing strategies, linking system efficiency directly to user satisfaction [2].

Collectively, these contributions reflect the rich diversity of methodologies and perspectives shaping the discourse on static rebalancing within BSS. Spanning from computational algorithms and graphical modeling to the incorporation of maintenance needs and user behavior considerations, the field is characterized by a dynamic array of research efforts aimed at refining and enhancing the rebalancing process. This collective body of work ensures the continued appeal and vitality of bike-sharing systems as sustainable urban mobility solutions.

Dynamic Rebalancing

In contrast, dynamic rebalancing represents a more proactive approach to bike-sharing systems, adjusting bike distribution in real-time or near-real-time based on current demand. This strategy leverages advanced data analytics, including GPS tracking, usage patterns, and predictive modeling, to identify demand peaks and strategically relocate bikes. Such responsiveness to demand fluctuations allows for swift adaptation to changes, accommodating variations in user needs due to time of day, events, or seasons, thereby increasing user satisfaction and promoting system reliability. Furthermore, dynamic rebalancing optimizes the allocation of transportation and labor resources by focusing redistribution efforts where they are most needed, based on data-driven predictions. However, this approach comes with its set of challenges, including high operational complexity due to the extensive data collection and analysis required, necessitating sophisticated infrastructure for real-time bike tracking and logistics optimization. The need for ongoing monitoring and quick adjustments also leads to increased operational costs, driven by the deployment of specialized vehicles and personnel for bike relocation. Additionally, reliance on predictive algorithms might result in the risk of overcorrection, inadvertently moving bikes from areas where they will soon be in demand again, which can disrupt the balance between supply and demand.

Dynamic Rebalancing methods

The allure of dynamic rebalancing lies in its direct response to the fluid nature of urban mobility. The development of strategies based on historical data to predict network conditions ensures optimal bike and stall availability amidst demand fluctuations, highlighting the superiority of such rebalancing frameworks over static schedules and marking a notable advancement in the management of Bike Sharing System (BSS) operations [14].

Further exploration into user incentive strategies to improve the performance of BSS focused on integrating dynamic rebalancing efforts with gamification techniques to encourage user participation in the system, thereby enhancing both the efficiency and user satisfaction of bike-sharing services [12]. This approach demonstrates that a joint optimization strategy, which includes both manual rebalancing by operators and incentives for users, can lead to a higher service quality at a lower operational cost, proving the viability and effectiveness of the model in real-world scenarios.

Moreover, the challenges and opportunities presented by free-floating BSSs are addressed by proposing methods to mitigate intensive rebalancing needs during peak hours, suggesting that an increase in shared bikes at strategic locations can enhance the overall efficiency of BSS [52].

The intricate decision-making process behind dynamic rebalancing, which operators must navigate to maximize trips, profits, and adherence to service level agreements, showcases the tricky balance that must be struck [7].

Further expanding, the introduction of an evolutionary algorithm to optimize rebalancing operations focuses on minimizing both the total distance traveled by rebalancing vehicles and the total unmet demand, illustrating the potential of combining technological solutions and human-centric strategies to address the challenges faced by bike-sharing systems [46].

At the heart of dynamic rebalancing is the drive to achieve a delicate balance between operational feasibility and user satisfaction. As BSS continue to proliferate across urban landscapes, the insights garnered from these studies illuminate the path forward. By embracing data-driven strategies and predictive analytics, BSS operators can not only respond more effectively to the ebb and flow of urban demand but also ensure the sustainability and attractiveness of bike-sharing as a vital component of urban mobility.

3 Fairness

This chapter will aim to present the challenges involving fairness in the context of AI and sustainable mobility as a service. While these are not obviously overlapping fields, they share concerns related to equity and fair use and with the expanding number of applications in which AI is being deployed, it is worth discussing how this will affect mobility.

3.1 Fairness in Artificial Intelligence

The integration of AI systems into daily life—from healthcare diagnostics to financial lending decisions—has magnified the potential for algorithms to perpetuate or even exacerbate societal biases. These can often manifest as prejudices or discriminations based on attributes like race, gender, age, sexual orientation, religion, and socioeconomic status. These biases can influence behavior, decision-making, and policy development, contributing to unequal treatment or opportunities for individuals or groups.

In the context of AI, societal biases can be inadvertently incorporated into systems through biased data, algorithmic design, or interpretation of outputs. These can lead to AI systems that perpetuate or even exacerbate societal inequalities. The challenge of ensuring fairness in AI therefore requires serious research into how data is collected, how algorithms are built, and how applications are deployed.

For example, in exploring fairness and bias in algorithmic hiring, research examines the application of AI across the recruitment process, from sourcing to evaluation. This research highlights the significant influence of technologies on hiring decisions. It underscores the interaction between AI tools and social networks in job platforms, illustrating how these dynamics can affect the visibility of candidates and potentially introduce biases into the hiring process[33].

Other investigations delve into the implications of AI systems in healthcare, spotlighting the risk of perpetuating disparities across various healthcare domains, from diagnosis to treatment allocation. Their research is pivotal in guiding the ethical deployment of AI in healthcare, ensuring that advancements in medical technology contribute to reducing, rather than exacerbating, health disparities.[8]

The concept of fairness can be hard to grasp given that its meaning can vary significantly across cultures, legal frameworks, and individual perceptions.

In their cornerstone book, *Fairness and Machine Learning*, Barocas, Hardt, and Narayanan present a framework for discussing fairness from a computational perspective. They navigate through the complex field, articulating the nuances of bias and discrimination that can arise in automated decision-making systems. Their work emphasizes the importance of a principled approach to developing machine learning models that are not only technically sound but also ethically responsible.

They argue for the necessity of **transparency, accountability, and inclusiveness** in the design and implementation of algorithms, advocating for methods that ensure diverse and equitable outcomes. Through rigorous analysis and case studies, they demonstrate how machine learning models can inadvertently perpetuate existing inequalities, stressing the importance of interdisciplinary efforts to tackle these challenges.

One of the main challenges is indeed the translation of ethical principles into actionable criteria that can guide the development of fair AI systems. The collective efforts of researchers in this field prove why it is essential to embed ethical considerations into the lifecycle of AI development. Their work serves as a call for the AI community to prioritize fairness, not as an optional feature, but as a fundamental component of any system that seeks to make decisions affecting human lives. As these studies illustrate, fairness in AI is a dynamic and complex domain, that considers ethical, legal, and economic dimensions. Embedding fairness into the AI development process requires more than technical proficiency; it demands a commitment to ethical principles and social responsibility. Fairness should not be an afterthought but a core principle guiding our innovations.

3.2 Fairness in Mobility as a Service and Micromobility Sharing Systems

Incorporating fairness into Mobility as a Service (MaaS) and Micromobility Sharing Systems transcends the mere integration of diverse transportation options within a digital platform. It involves redefining urban mobility to foster inclusivity and ensure equal access to societal opportunities such as employment, education, and healthcare [32]. This shift signifies a movement towards urban environments where mobility solutions are designed with an equitable lens, guaranteeing that the benefits of both MaaS and micromobility extend to all segments of the population, including those traditionally marginalized.

The integration of MaaS with micromobility systems, such as bike-sharing and scooter-sharing, holds the potential to democratize urban mobility and bridge the existing mobility divide. By offering shared transportation modes, there is an opportunity to not only enhance urban accessibility but also to promote sustainability and reduce the environmental impact of private vehicle use [40]. However, realizing this potential requires overcoming barriers to adoption, which include technological challenges and socio-economic disparities [6]. Mobility justice advocates for a holistic approach in urban planning and policymaking, ensuring that transportation's broader societal impact supports marginalized communities [39].

Moreover, the adoption of MaaS and micromobility solutions conveys the importance of collaborative efforts between the public and private sectors to address challenges such as service coverage, public-private cooperation, and user acceptance [6]. This collaboration is crucial in making MaaS and micromobility solutions accessible and affordable, especially for communities that have historically been underserved by traditional transportation systems.

Addressing the unique mobility challenges faced by different population segments, particularly concerning safety, affordability, and accessibility, shows the significance of embedding fairness principles into transport policy and practice [25]. Recent studies have highlighted the gender dimension of mobility, revealing a skewed representation in the user base of bike-sharing and scooter-sharing services towards younger, well-educated male professionals [19]. This demographic imbalance necessitates MaaS and micromobility designs that cater to a broader user spectrum, thereby fostering a more equitable service distribution.

The journey towards equitable MaaS and micromobility systems is thus not solely a technological or infrastructural endeavor but a socio-political one too. It involves a commitment to addressing the underlying inequities within urban environments [43]. Through the collaborative and interdisciplinary efforts of policymakers, urban planners, and service providers, there lies a substantial opportunity to transform urban mobility landscapes into spaces that are more inclusive and equitable.

By embracing the insights provided by recent research, stakeholders in the urban mobility ecosystem are well-positioned to make impactful strides towards creating more accessible and inclusive cities. The path forward involves ensuring that these advancements are guided by a strong ethical framework centered on fairness and equity [21].

3.3 Metrics

In order to measure fairness, and in particular, fairness in the context of MSS, this section introduces some metrics which are widely used for such purposes.

Statistical Parity

Statistical Parity [26] is a simple but useful metric for evaluating fairness in algorithms, widely adopted in the assessment of MSS systems. It ensures that the decision-making process is equitable across different groups in a population. Statistical Parity is captured by the following equation:

$$P(\hat{Y} = 1|G = g_1) = P(\hat{Y} = 1|G = g_2) = \ldots = P(\hat{Y} = 1|G = g_n), \tag{3.1}$$

where the terms of Equation 3.1 are defined as follows:

- $P(\hat{Y} = 1|G = g_i)$: Represents the probability of a favorable outcome or decision $\hat{Y} = 1$ given the membership in a particular group $G = g_i$. This probability is measured for each group defined by a protected attribute (e.g., gender, race).

- $G = g_i$: Denotes the group membership based on protected attributes. Each g_i represents a distinct group within the population, for which fairness is to be assessed.
- $\hat{Y} = 1$: Indicates a favorable decision or outcome, such as being selected for a job interview or receiving a loan approval.
- The equation asserts that the probability of receiving a positive outcome should be equal across all groups $G = g_1, G = g_2, \ldots, G = g_n$, to ensure nondiscrimination. This equality in probabilities across different groups signifies the achievement of Statistical Parity.

Achieving Statistical Parity is desirable for fairness in MSS systems, particularly in scenarios where historical biases could influence the distribution of services or resources. It embodies the principle that all individuals, irrespective of group affiliation, should have equal access to benefits provided by the system.

GINI Coefficient

Another metric used to quantify fairness within the MSS context it the Gini index [48] , defined as:

$$g(x) = \frac{1}{2n^2\bar{x}} \sum_{j=1}^{n} \sum_{k=1}^{n} |x_j - x_k|, \tag{3.2}$$

where the components of the equation are as follows:

- $g(x)$: The Gini index for the distribution x, quantifying the fairness in the system with values ranging from 0 to 1, where 0 denotes a perfectly fair system and 1 indicates maximal unfairness.
- $n = 3$: The number of area categories, indicating the scope of the assessment across different geographic or service areas.
- x_j and x_k: Represent the metrics of interest for the j^{th} and k^{th} categories, respectively. These metrics could include the number of rebalancing operations performed or the probability of service failure at finding an available vehicle within a given category.
- $\bar{x}$: The mean of the quantities x_k for $k = 1, \ldots, n$, used as a normalization factor to account for variations in the scale of the metrics of interest.
- The expression $\sum_{j=1}^{n} \sum_{k=1}^{n} |x_j - x_k|$: Aggregates the absolute differences between all pairs of categories, which is central to assessing the dispersion across the distribution.

This formulation of the Gini index allows for a detailed examination of the distributional fairness across categories, providing a quantitative foundation to evaluate and address disparities in service access.

JAIN Fairness Index

Jain's Fairness Index, defined below, serves as another measure for evaluating the fairness of a system [29] , providing an useful complement to the Gini index within the MSS context:

$$J(x) = \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n \sum_{i=1}^{n} x_i^2}, \tag{3.3}$$

The elements of Equation 3.3 are detailed as follows:

- $J(x)$: Jain's Fairness Index for the distribution x, offering a measure of equity in resource allocation or service levels. It quantifies how uniformly resources or services are distributed across a population.
- n: The total number of entities or users within the system, affecting the range of the index from $\frac{1}{n}$ (indicating complete unfairness) to 1 (denoting perfect fairness).
- x_i: Represents the specific metric being evaluated for fairness for the i^{th} entity. This could be, for instance, the number of services each entity has accessed or the quality of service each has experienced.
- $\sum_{i=1}^{n} x_i$: The sum of the fairness metrics across all entities, used in both the numerator and denominator, reflects the aggregate level of resource or service allocation.
- $n \sum_{i=1}^{n} x_i^2$: The denominator scales the squared sum of individual metrics by the number of entities, helping to normalize the index and accentuate disparities in allocation.

Implementing these metrics in the evaluation of MaaS systems allows policymakers and engineers to identify disparities in service provision and address them effectively. By analyzing data on service usage and availability with these metrics, stakeholders can understand areas where interventions are needed to improve equity, such as increasing the number of vehicles in underserved areas or adjusting pricing models to make services more affordable for lower-income users.

Ultimately, the goal is to create a MaaS ecosystem that provides reliable, convenient, and affordable access to all, ensuring that mobility services contribute positively to social equity and the reduction of socioeconomic disparities. By regularly measuring and adjusting for fairness, MaaS can become an example of sustainable and inclusive urban development.

4

Elements of Reinforcement Learning

As outlined in the previous chapters, the aim of this research is to address the trade-off between cost optimization and fairness in MaaS, through the use of Reinforcement Learning. To this end, this chapter will first introduce the core concepts in the RL framework in order to then discuss how RL was applied to the problem in question in the next chapter.

4.1 Reinforcement Learning

Reinforcement Learning is a type of machine learning where an agent learns to make decisions by performing actions in an environment to achieve some goal [47]. The agent learns from the outcomes of its actions, rather than from being preemptively programmed to explicitly know what to do. This learning process is driven by the feedback the agent receives from the environment in the form of rewards or penalties, which guide the agent towards achieving the optimal behavior or policy for maximizing cumulative rewards over time.

In recent years it has found numerous successful application, in a diverse number of fields, such as Robotics, Finance, Healthcare, Gaming, Recommendation Systems, Supply Chain Management and Energy Management.

Reinforcement Learning framework

The fundamental concepts in RL include the agent, environment, state, action, reward, policy, and value function:

- **Agent**: The learner or decision maker.
- **Environment**: Everything the agent interacts with.
- **State (S)**: A representation of the current situation of the agent in the environment.
- **Action (A)**: Any decision or move the agent makes.
- **Reward (R)**: An immediate return given to the agent for performing an action in a particular state. It serves as feedback to the agent.
- **Policy (π)**: A strategy used by the agent, mapping states to actions.
- **Value Function**: A function that estimates how good it is for the agent to be in a given state (or how good it is to perform a certain action in a given state). The "goodness" here is measured in terms of expected future rewards.

Exploration vs Exploitation

One of the early decisions to make when using Reinforcement Learning (RL) concerns the trade-off between exploration and exploitation. Exploration involves the agent trying out different actions to discover new knowledge about the environment. It is necessary for learning the value of actions in different states, especially those not yet visited or less understood. On the other hand, exploitation refers to the agent using its current knowledge to make the best decision based on what it already knows. This means choosing actions that are known to yield the highest reward based on the current policy and value function.

Balancing exploration and exploitation is critical. Too much exploration can lead to inefficiency and missed opportunities for maximizing rewards, while too much exploitation can prevent the agent from discovering potentially better strategies. Effective reinforcement learning algorithms design mechanisms to balance these two aspects, often through strategies like ϵ-greedy, where the agent explores randomly with probability ϵ and exploits its current knowledge with probability $1 - \epsilon$.

4.1.1 Markov Decision Process

Markov Decision Processes (MDPs) provide the mathematical framework for modeling decision making in situations where outcomes are partly random and partly under the control of a decision maker [47]. MDPs are an extension of Markov Reward Processes (MRPs), which are in turn an extension of Markov Processes (MPs). A Markov Process is characterized by a set of states S and a transition matrix P, which describes the probability of transitioning to state s' from state s, denoted as:

$$P_{ss'} = P(St + 1 = s' \mid St = s). \tag{4.1}$$

MRPs introduce the concept of rewards and discounting. The discount factor $\gamma \in [0, 1]$ is introduced, and the reward R, which indicates how aligned a transition is with the end goal, is defined as:

$$R_s = E[R_{t+1} \mid St]. \tag{4.2}$$

With MRPs, we can define the concept of returns as the sum of discounted rewards from time t of a trajectory under the transition matrix P, formulated as:

$$G_t = R_{t+1} + \gamma R_{t+2} + \gamma^2 R_{t+3} + \ldots = \sum_{i=0}^{\infty} \gamma^i R_{t+i+1}. \tag{4.3}$$

Since G_t describes the return of a single trajectory, we can define the value function V that describes the expected return as:

$$V(s) = E[G_t \mid St = s]. \tag{4.4}$$

The return G_t can be recursively defined, which leads to:

$$G_t = R_{t+1} + \gamma(R_{t+2} + \gamma R_{t+3} + \ldots) = R_{t+1} + \gamma G_{t+1}. \tag{4.5}$$

MDPs introduce actions into the framework, allowing the transition probability and reward function to be redefined to account for actions. The transition probability given an action is:

$$P^a_{ss'} = P(St + 1 = s' \mid St = s, At = a), \tag{4.6}$$

and the reward function is:

$$R_{as} = E[R_{t+1} \mid St, At = a]. \tag{4.7}$$

Actions allow an agent to make decisions at each step, necessitating the definition of a policy π, which specifies the agent's action preference in a given state:

$$\pi(a \mid s) = P(At = a \mid St = s). \tag{4.8}$$

A policy fully determines the agent's behavior and defines how it learns over time. Since the value function V depends only on the state, it does not need to be redefined; however, the introduction of actions leads to the definition of the Q function to take advantage of the agent's ability to act, formulated as:

$$Q(s, a) = E[G_{t+1} \mid St = s, At = a]. \tag{4.9}$$

By leveraging the definition of the return G_t, as previously done for the value function V, we can redefine the Q function recursively, thus establishing a relationship between V and Q functions and introducing them into their definitions. Learning these functions is a complex task that has given rise to multiple algorithms, all of which need to navigate the exploration-exploitation dilemma inherent in training agents.

4.1.2 Introduction to Q-Learning

Learning in the context of Markov Decision Processes (MDPs) involves determining an optimal policy that maximizes the expected return from any given state. An important aspect of learning in MDPs is balancing the trade-off between exploration, or trying new actions to discover their effects, and exploitation, or using known actions that yield high rewards. This section introduces Q-Learning, a fundamental approach for solving reinforcement learning problems.

Initially, learning methods such as Monte Carlo approaches and Temporal Difference (TD) learning were developed to estimate value functions and policies. While Monte Carlo methods wait until the end of an episode to update value estimates based on the returns received, TD learning updates estimates based on other, already estimated values, allowing for learning to occur from incomplete episodes. Both methods, however, focus on learning the value of states or state-action pairs without directly addressing how to learn an optimal policy.

Q-Learning, a form of Temporal Difference learning, emerges as a solution to

Algorithm 1 Q-Learning Algorithm

Initialise $Q(s, a)$ arbitrarily for all $(s, a) \in \mathcal{S} \times \mathcal{A}$
Choose α (learning rate), γ (discount factor), and policy π (e.g., ε-greedy)
for each episode **do**
  Initialise state s
  while s is not terminal **do**
    Choose action a from s using policy derived from Q (e.g., ε-greedy)
    Take action a, observe reward r and next state s'
    $Q(s, a) \leftarrow Q(s, a) + \alpha[r + \gamma \max_{a'} Q(s', a') - Q(s, a)]$
    $s \leftarrow s'$
  end while
end for

this challenge by learning the value of taking a particular action in a particular state directly. It updates its estimates of state-action values (Q-values) towards target values that combine immediate rewards with the discounted value of the subsequent state, as estimated under the current policy. The Q-Learning algorithm is model-independent and can learn optimal policies even in deterministic environments.

In this algorithm, $Q(S, A) \leftarrow Q(S, A) + \alpha[R + \gamma \max_a Q(S', a) - Q(S, A)]$ represents the core of Q-Learning, where:

- $Q(S, A)$ is the current estimate of the state-action value.
- α is the learning rate, controlling the extent to which new information affects existing estimates.
- R is the reward received for taking action A in state S.
- γ is the discount factor, determining the importance of future rewards.
- $\max_a Q(S', a)$ is the estimate of optimal future value, maximized over all possible actions at the next state S'.

Q-Learning algorithmically formalizes the process of iteratively improving the policy by updating the Q-values based on the equation provided. Through repeated interactions with the environment and updates to the Q-values, the algorithm converges to the optimal policy that maximizes the cumulative reward.

The Q-Learning algorithm is guaranteed to converge to the optimal solution as long as some conditions are satisfied, as stated by the following theorem.

Theorem 1 (Convergence of Q-Learning [35]). *Given a finite MDP, the Q-Learning algorithm given by the update rule*

$$\begin{aligned} Q_{t+1}(s_t, a_t) \leftarrow & Q_t(s_t, a_t) + \alpha_t(s_t, a_t) \cdot \\ & (R_{t+1} + \gamma \max_{a'} Q_t(s_{t+1}, a') - Q_t(s_t, a_t)), \end{aligned} \tag{4.10}$$

converges with probability 1 to the optimal state-action value function as long as

$$\sum_t \alpha_t(s,a) = \infty \quad \wedge \quad \sum_t \alpha_t^2(s,a) < \infty \tag{4.11}$$

for all $(s,a) \in \mathcal{S} \times \mathcal{A}$.

Q-Learning stands out as a powerful and flexible algorithm in the realm of reinforcement learning. The algorithm's simplicity and model-free nature make it applicable to a wide range of problems, from simple tasks to complex decision-making challenges. As research in reinforcement learning progresses, Q-Learning continues to be a foundational technique. It encapsulates the essence of learning from interaction, progressively improving decision-making to maximize rewards. Its continued relevance and adaptability to new challenges affirm its importance in the toolkit of methods for tackling the intricacies of autonomous decision-making and learning in uncertain environments.

5

System Model

This section is dedicated to the preliminary notions needed for the modelling of a MSS, considering a Bike Sharing System as reference.

5.1 Network

A dock-based MSS is naturally defined as a fully connected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where a node in $\mathcal{V}$ represents a station and $\mathcal{E} = \mathcal{V} \times \mathcal{V}$ denotes the set of connections between each pair of stations. Each node $i \in \mathcal{V}$ is characterized by its current occupancy, i.e., the number of bikes present at the i-th station at time t.

In constructing our model for a Micromobility Sharing System, we considered system types and the structuring of the network within the urban environment. The contents of Chapter 2 and the observed trends in micro-mobility, led us to favor a dockless system over a dock-based one. This decision is grounded in the increased flexibility offered to users by dockless systems, a feature that is becoming increasingly prevalent in urban micro-mobility landscapes. By transitioning the focus from fixed stations to service areas, we redefine the set of nodes $\mathcal{V}$ as a partition of the city map. In this context, each node represents a specific area where the count of vehicles is maintained, thereby allowing the extensive body of literature on docked systems to be applicable in a dockless framework through this spatial reconfiguration.

Regarding the scale and structure of the network, our model is designed with adaptability in mind, capable of accommodating any number of nodes. However, for experimental purposes and to facilitate learning, we established a network consisting of 100 nodes. This choice reflects a scale that is representative of a medium-large Bike Sharing System (BSS). Within this network, nodes are categorized into three groups based on their demand patterns. This classification forms the basis for further investigation into how different areas within a city contribute to and interact with the overall micro-mobility ecosystem.

A limitation in the network was the absence of considerations linked to distance between nodes. These were characterized by their demand patterns and not by how far they are from each other or for how costly it is to move vehicles from one node to another. While this is a limitation, it will be addressed in chapter 7.

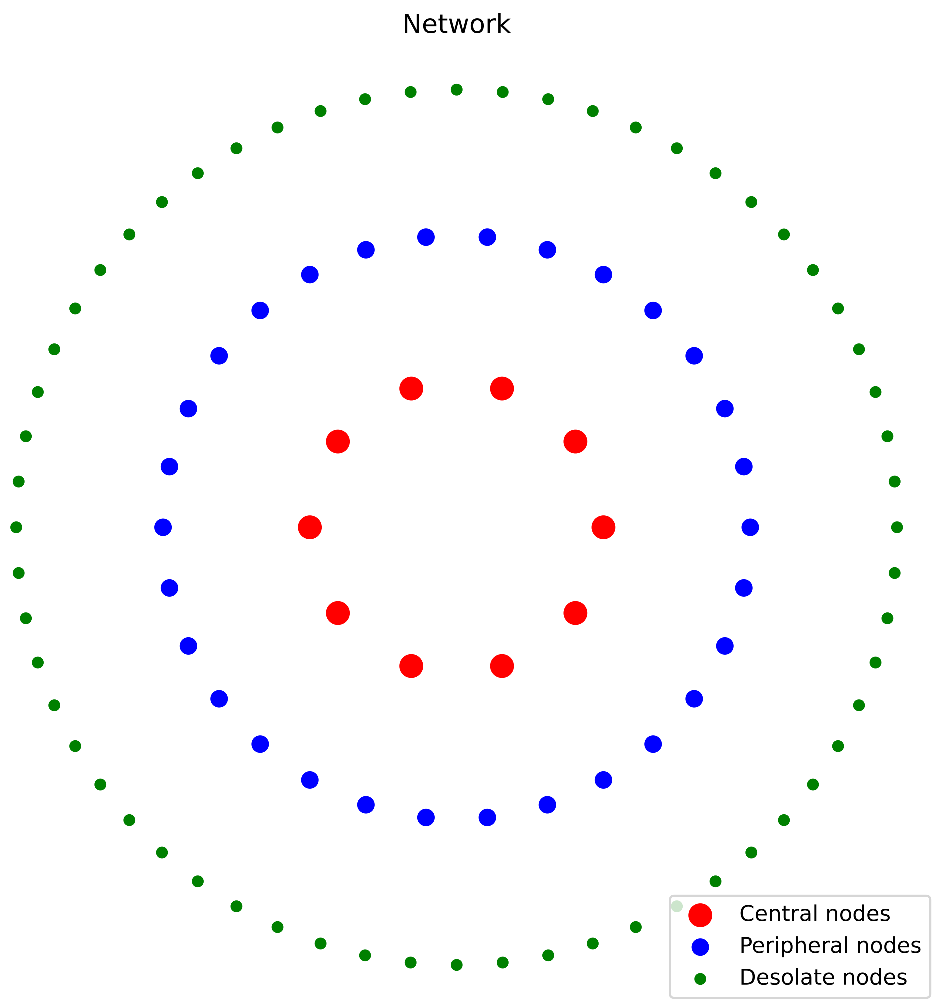


Figure 5.1: Network visualization

5.2 Demand

Demand is a crucial aspect in working with MaaS systems. In order to simulate the environment in which to deploy the RL strategies, there were several considerations to take into account when thinking about it. A first approach would be to use an existing dataset containing all the information for all bikes in a network.

There are plenty of such datasets in literature, including:

- The NYC Citi Bike dataset, which offers comprehensive trip records and station information in New York City.
- The London Santander Cycles dataset, offering detailed ride data in London.

Typically, these datasets have the form/include information like:

Field	**Description**
Trip Duration	The length of time the trip took from start to finish.
Start Time and Date	The time and date when the trip began.
Stop Time and Date	The time and date when the trip ended.
Start Station ID	A unique identifier for the station where the trip originated.
Start Station Name	The name of the station where the trip began.
Start Station Latitude	The latitude of the start station.
Start Station Longitude	The longitude of the start station.
End Station ID	A unique identifier for the station where the trip concluded.
End Station Name	The name of the station where the trip ended.
End Station Latitude	The latitude of the end station.
End Station Longitude	The longitude of the end station.
Bike ID	A unique identifier for the bicycle used in the trip.
User Type	The type of user (e.g., "Subscriber" or "Customer").

Table 5.1: Description of Fields in Bike Sharing Dataset

However, censoring is a major issue with this approach. The dataset may not fully capture unfulfilled demand. For instance, potential riders might decide not to use the service if no bikes are available at their nearest station or if all docking points are occupied at their destination. This type of demand is not directly observable in the data and could require estimation techniques to infer the true demand levels.

With this in mind, we decided to proceed with synthetic data. It is important to observe that accurately modeling and predicting the dynamics of such

Table 5.2: Example dataset of bike sharing in Pandas DataFrame format.

Trip Duration	**Start Time**	**Stop Time**	**Start Station ID**
670	2023-03-01 00:01:00	2023-03-01 00:12:00	3186
754	2023-03-01 00:04:00	2023-03-01 00:16:00	3183
1223	2023-03-01 00:06:00	2023-03-01 00:26:00	3183

Start Station Name	**Start Lat.**	**Start Long.**	**End Station ID**	**End Station Name**
Grove St PATH	40.719586	-74.043117	3199	Newport Pkwy
Exchange Place	40.716247	-74.033459	3203	Hamilton Park
Exchange Place	40.716247	-74.033459	3202	Jersey & 3rd

networks in their entirety is not a computationally tractable problem for large MSS services, like the ones that we are interested in. We therefore focused on a stochastic model of an individual service area, considering an independent MMPP for the arrivals and departures, which is consistent with experimental results on large sharing systems [14]. The demand rates vary according to daily, weekly, and seasonal cycles, and are affected by geographic factors as well. The vehicle occupancy of the area then follows a left-censored continuous-time MBDP, i.e., a stochastic process in which Poisson events represent either an increase or a decrease of the state by 1, and in which the rate of these events is the outcome of a Markov process with discrete time steps. The left censoring limits the state to positive values: while new arrivals are always possible (unlike in dock-based systems, in which stations have a maximum capacity), a new departure from the area is impossible if there are no vehicles to take.

The transition probability from state m to state n over time t is then approximated by

$$P_{m,n}(t) \simeq \begin{cases} \sum_{\ell=m}^{\infty} p_{\text{Sk}}(-\ell; t, \lambda, \mu), & \text{if } n = 0; \\ p_{\text{Sk}}(n - m; t, \lambda, \mu), & \text{if } n > 0; \end{cases} \tag{5.1}$$

where $p_{\text{Sk}}(n; t, \lambda, \mu)$ is the Skellam distribution [42], which corresponds to the difference of two Poisson r.v.s:

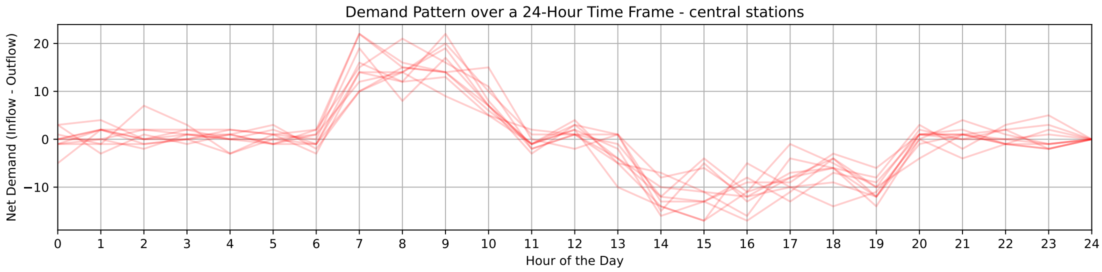
$$p_{\text{Sk}}(k; t, \lambda, \mu) = e^{-t(\lambda+\mu)} \left(\frac{\lambda}{\mu}\right)^{\frac{k}{2}} I_k \left(2t\sqrt{\lambda\mu}\right), \tag{5.2}$$

where λ and μ represent the arrival and departure rates, respectively, and $I_k(\cdot)$ is the modified Bessel function of the first kind [1]. The approximation follows the work in [14] and is necessary due to the left censoring.

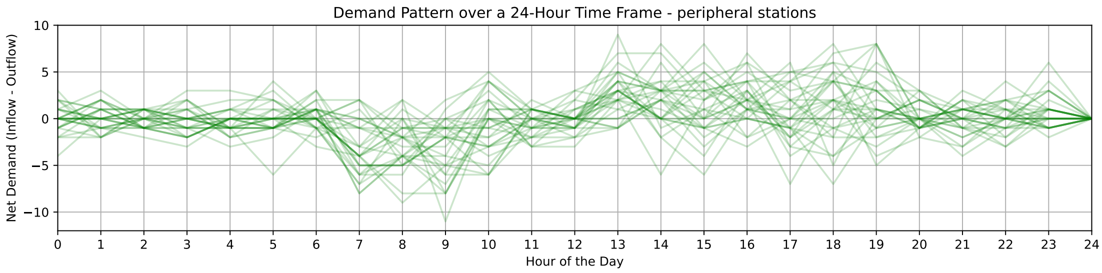
As mentioned above, will consider a system with V service areas, which we divide in three categories according to common spatial patterns in US and European cities:

- *Central* areas in the city center, where large businesses and attractions are clustered, typically have a high traffic volume, with more arrivals than departures in the morning and the opposite in the evening due to commuter traffic;
- *Peripheral* areas are typically residential areas close to the center, which typically have a lower density, and thus less traffic, but present an inverted pattern with respect to daily activities, i.e., more departures in the morning and arrivals in the evening;
- *Remote* areas are typically underserved suburbs and lower-income communities. The traffic patterns are similar to peripheral areas, but with an even lower traffic volume due to the factors we highlighted above.

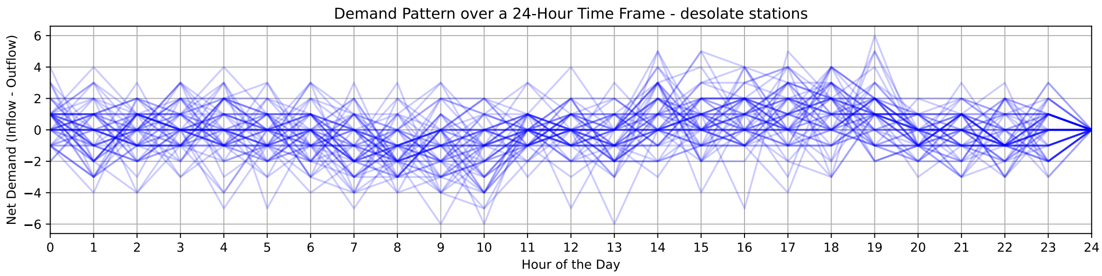
As for modelling each arrival and departure accurately, given the overall traffic level for each station, once again these were designed using a Possion distribution. Finally, in order to ensure a degree of robustness, the bike fleet was distributed randomly across the network's nodes.



(a) Demand pattern over a 24-hour time frame for central stations.



(b) Demand pattern over a 24-hour time frame for peripheral stations.



(c) Demand pattern over a 24-hour time frame for desolate stations.

Figure 5.2: Demand patterns over a 24-hour time frame for different types of stations.

6

A novel Reinforcement Learning approach for fairness-oriented MSSs

This chapter makes use of the concepts introduced in the previous ones to show how RL was applied to the specific context of static rebalancing in MSS.

6.1 Learning Framework

We now illustrate the control approach of this study by modeling the problem as a multi-agent MDP and defining the solution. We start from the multi-agent Reinforcement Learning (RL) approach for the operation and control of MSS networks. The discussion continues with the presentation of the model for individual agents and the adopted reward mechanism, which stands at the core of the proposed fairness-oriented strategy. However, we note that the main contribution of the paper does not lie in the RL solution itself, which follows a relatively common separability approach, but in the application of fairness principles to the control of rebalancing operations in an MSS network.

Our statistical model relies on an independence assumption: the MMPP representing arrivals and departures in each area are assumed to be independent both from each other and from the processes in other areas. Naturally, this assumption is not verified in real systems, as trips usually begin in an area and end in another a few minutes later, but the approximation error is surprisingly low in large-scale systems [13]: any individual area makes up such a small fraction of the total traffic that local events have negligible effects elsewhere.

This independence property makes it possible to consider individual rebalancing actions in different areas as separate problems, modeling the system as a transition- and reward- independent MDP: actions from one agent have no effect on the state transitions of others, although the overall reward function might be a non-linear function of the individual reward. The overall state of the multi-agent problem can then be factored into individual state components for each area. Even this case may not be solvable in polynomial time [5], so we will need to structure the reward function in such a way that we can further simplify the problem.

The agents use an exponentially-annealed ε-greedy policy, which guarantees convergence thanks to Theorem 1. In view of that, we design the learning rate in this way:

$$\alpha_t(s_t, a_t) = \begin{cases} c & \text{if } t < T; \\ \frac{1}{t} & \text{otherwise;} \end{cases} \tag{6.1}$$

where $c > 0$ is a constant value and T is a pre-specified threshold. In this way both the requirements in (4.11) are satisfied, thus ensuring convergence.

6.1.1 Factorized MDP representation

As proposed in Section , we cluster the nodes of the MSS network into three different categories: central, peripheral and remote. These are distinguished by their demand patterns, as well as by the overall traffic volume, which is high for central areas and gradually decreasing for peripheral and remote ones. This spatial categorization is necessary for understanding how to measure a fair allocation of vehicles across the network.

The elements that constitute the state of each individual agent are the following:

- The time of the day, either morning or evening;
- The area type, i.e. *central, peripheral or remote*;
- The number of vehicles currently available in the area.

Naturally, the time is a shared component among different agents, but its deterministic transition allows for factorization by simply replicating it N times. The action space for each agent is designed to be granular enough to offer meaningful choices without overwhelming the agent with too many options. Actions include adding up to 30 vehicles or removing up to 20 vehicles, by increments of 5.

6.1.2 Reward design and fairness considerations

We propose an innovative rebalancing agent designed to perform *fair* static rebalancing within a MSS environment. We partition $\mathcal{V}$ into $\mathcal{V}_c$, $\mathcal{V}_p$, and $\mathcal{V}_r$: the three subsets contain the central, peripheral, and remote areas, respectively. The global reward function measures how effective the system was at meeting demand in the preceding 12 hours and it also takes into account the number of vehicles that were added or removed from the network:

$$R_t = -\sum_{i \in \mathcal{V}} (f_{t_i} + 0.05 a_{t,i}) + \sum_{j \in \mathcal{V}_c} (\beta f_{t_j}) - \sum_{j \in \mathcal{V}_r} (\beta f_{t_j}), \tag{6.2}$$

where $f_{t,i}$ represents the number of failures in area i during the considered interval, i.e., the number of users who fail to find a shared vehicle in that area, and β is a hyperparameter used as a *temperature*, to measure the degree of importance[1] that is given to remote areas with respect to central and peripheral ones.

Given the global reward function in (6.2), the optimal solution to the problem is given by the combination of the individual solutions to the agent problems, with the following reward function:

$$R_{t,i} = \begin{cases} -(1-\beta)f - 0.05a_{t,i} & \text{if } i \in \mathcal{V}_c; \\ -f - 0.05a_{t,i} & \text{if } i \in \mathcal{V}_p; \\ -(1+\beta)f - 0.05a_{t,i} & \text{if } i \in \mathcal{V}_r. \end{cases} \tag{6.3}$$

The resulting solution then enjoys the convergence property given in Theorem 1.

Proof. We know that the state is separable, as the state transition probability of agent i is only affected by its own action $a_{t,i}$: two components of the state (the time of day and the area type) evolve deterministically, while the third

[1] In general, the adjustment of the temperature plays a pivotal role in controlling the delicate balance between optimizing performance metrics [24], such as accuracy, and ensuring equity in sociotechnical systems.

follows an independent process in each area. We can also trivially prove that the global reward function in (6.2) is the sum of the individual reward functions in (6.3). The distributed Q-Learning algorithm is then optimal for the global problem. □

The final control pipeline is summarised in Algorithm 2.

Algorithm 2 MSS Q-Learning

Initialise $Q(s, a) \in \mathbb{R}$ for all $(s, a) \in \mathcal{S} \times \mathcal{A}$
$\varepsilon \leftarrow 1$ (exploration rate)
$\delta \leftarrow 0.999$ (decay factor)
for day = 1 **to** $D \to \infty$ **do**
  Setup initial state $s_0 = \bigcup_{i \in \mathcal{V}} s_{i,0}$
  for $t = 0$ **to** 23 **do**
    Observe transition process from s_t to s_{t+1}
    if $t == 11$ or $t == 23$ **then**
      for all areas $i \in \mathcal{V}$ **do**
        Choose $a_{i,t}$ from $s_{i,t}$ using ε-greedy policy derived from Q:

$$a_{i,t} = \begin{cases} \arg\max_a Q(s_{i,t}, a) & \text{w/ prob. } 1 - \varepsilon \\ \text{random action } a \in \mathcal{A} & \text{w/ prob. } \varepsilon \end{cases} \tag{6.4}$$

        Take action $a_{i,t}$, observe $R_{i,t+1}$
        Update Q according to (4.10)
      end for
    end if
  end for
  $\varepsilon \leftarrow \varepsilon \cdot \delta$
end for

7 Simulations and Results

We provide an extensive investigation of different rebalancing strategies, to demonstrate the trade-off between performance and equity measured by indexes defined in Section 3 with the goal of finding a viable compromise.

As an example of a dockless MSS, we consider a medium-sized bike sharing system. Its spatial representation is modeled as a network with 100 nodes in total; of these, 10 are set to be *central* (labeled by $k = 1$), 30 are *peripheral* (labeled by $k = 2$), and 60 are *remote* (labeled by $k = 3$). The simulation starts with each area being subject to the demand introduced in Section 5, which is synthetically generated. On averae, the network is subject to over 7000 daily trips. At every hour $t = 0, \dots, 23$, the number of vehicles present at each area is updated based on arrivals and departures. If at a certain moment a node is unable to satisfy the demand, i.e. there are no vehicles left at one particular area and there is request for a departure, this is registered as a single failure for that node. The agent performs its control actions at 11a.m. and at 11p.m. every day through static rebalancing, as described in Section 6. The training phase for the proposed strategy is run through 1000 days and evaluated over the last 100 steps in the following analysis. According to (6.1), the learning rate α_t is initially fixed at $c = 0.01$ and gradually decreased after $T = 800 \cdot 24$ hours.

With the above setup, we analyse the Pareto-efficiency of the provided approach. Also, we point out that certain RL strategies should be preferred over others, as they lead to fewer service failures and higher fairness scores.

Firsly, we highlight the presence of an inherent trade-off between the control performance and fairness degree of the proposed approach applied to this

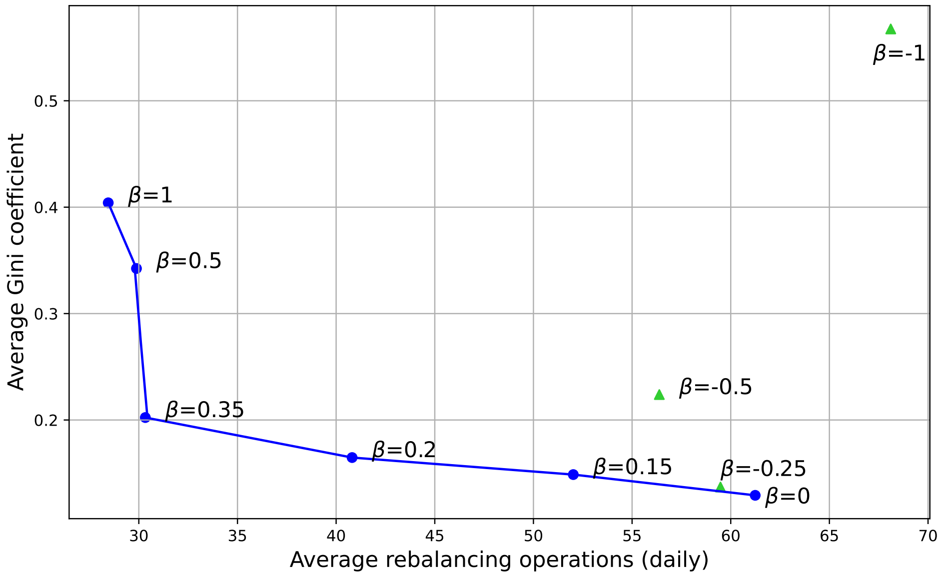


Figure 7.1: Pareto front: Relationship between Gini scores and number of rebalancing operations obtained by averaging the corresponding outcomes of a 100-run Monte Carlo simulation.

setup. As a cost index, the overall number of rebalancing operations x_k executed by Algorithm 2 for each category k is considered. On the other hand, the Gini index defined in (3.2) is computed over the the rebalancing operations x_k, with $k \in \{1, 2, 3\}$, to assess the fairness degree. To this aim, we run a Monte Carlo simulation examining the role of different local reward values parameterized by $\beta \in [-1, 1]$ in (6.3). More in detail, we consider such an evaluation over the grid $\beta \in \{-1, -0.5, -0.25, 0, 0.15, 0.2, 0.35, 0.5, 1\}$. Fig. 7.1 shows the aforementioned trade-off, illustrating the intrinsic complexity of selecting an "optimal" rebalancing strategy among those parameterized by β. Nonetheless, it is also evident that certain strategies (i.e., those characterized by $\beta < 0$) should be discarded, since they do not exhibit Pareto-efficiency (see Fig. 7.1). As a further support to the previous claim, Fig. 7.3a and Fig. 7.3b illustrate the distributions of the cost scores and fairness scores, respectively. The boxplots highlighted in light blue refer to Pareto-efficient strategies, while those depicted in green represent the suboptimal choices. In particular, the depicted distribution behaviours allow to determine that the case $\beta = -0.25$ is just nearly optimal but does not lie on the Pareto-front. Similar conclusions can be drawn by observing the same plots using the Jain Fairness index instead, as shown in Fig. 7.2 and Fig. 7.4

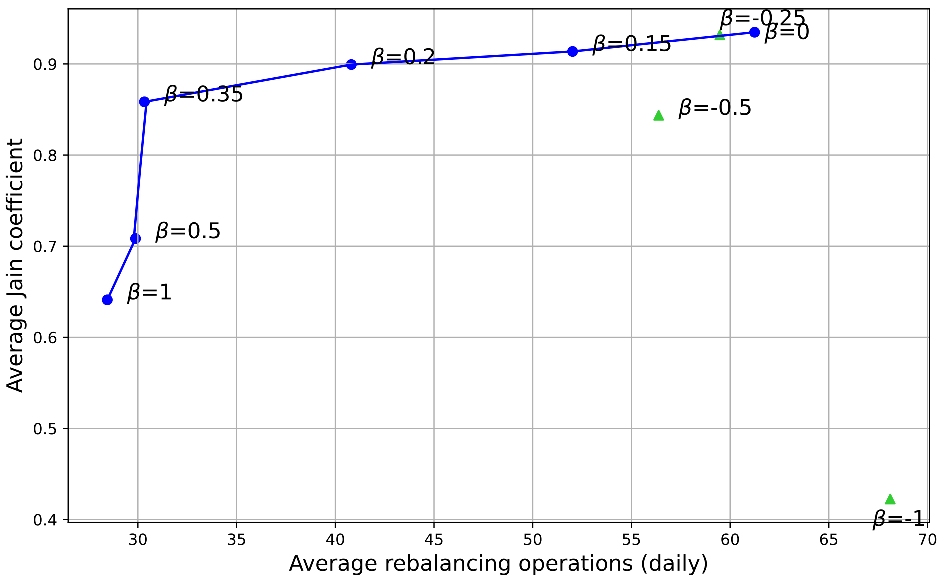


Figure 7.2: Pareto front: Relationship between Jain indeces and number of rebalancing operations obtained by averaging the corresponding outcomes of a 100-run Monte Carlo simulation.

We finally compare two of the strategies depicted in Fig 7.5, namely those with $\beta = -0.5$ (unfair) and $\beta = 0.2$ (fair), to better illustrate certain aspects related to fairness and cost. Fig conveys how both strategies lead to a similar value in terms of absolute failures, however there are more details to consider. Firstly, the strategy obtained using $\beta = 0.2$ is more expensive in terms of daily rebalancing operations and in terms of fleet size. Indeed, by observing Fig, one can notice how on average this strategy requires more daily rebalancing operations, in particular at remote areas. Given the high reward the RL agent receives for avoiding such failures, this is in line with expectations. Another worthwhile observations is that for $\beta = -0.5$, the reward for avoiding failures in remote areas is so low that it discourages the agent from rebalancing in these areas altogether. As for the fleet size, this strategy converges to needing a fleet of 1285 bikes, compared to the 944 for the on obtained using $\beta = -0.5$. So, for the reasons stated above, the reward function using $\beta = 0.2$ clearly leads to a more costly strategy.

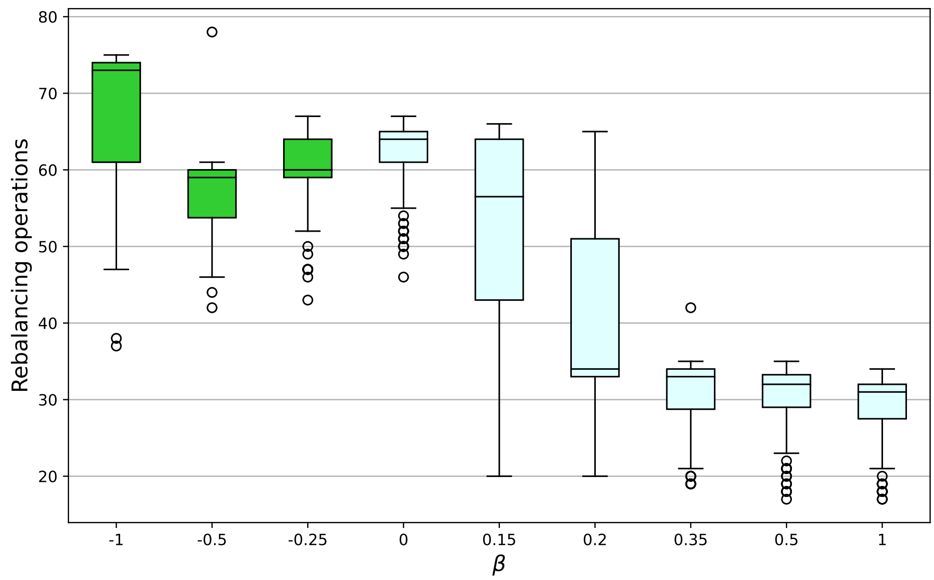
By comparison, if instead of analysing absolute failures we turn to relative failure rates, we better understand how the two strategies differ in terms of fairness. This measure is calculated for each area type as the percentage of failures for each area type given the total potential demand which could be satisfied for each area type. By comparing the two strategies obtained using the two different RFs, it clearly emerges that setting $\beta = -0.5$ penalises remote areas heavily when compared with other area types whereas as by using $\beta = 0.2$. This is, as expected, reflected in both Gini and Jain indexes. For the reasons stated above we can conclude that unfair solutions that do not belong to the Pareto-front can be avoided in general. Conversely, fair Pareto-optimal strategies are worth considering despite their higher cost.

	RF 1	RF 2
gini	0.13	0.31
jain	0.95	0.73

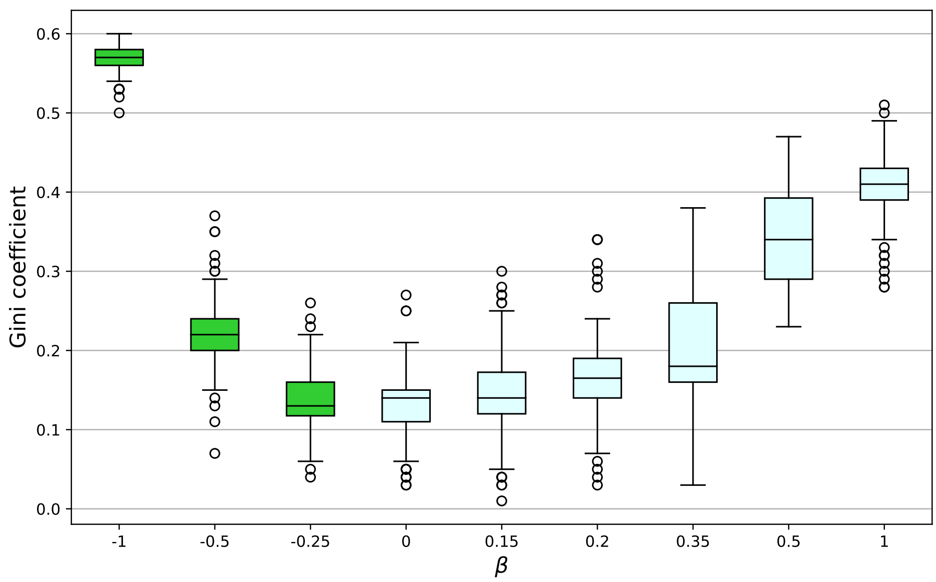
Table 7.1: fairness coefficients for different Reward Functions

By reflecting on these findings, system designers and policy makers should interrogate themselves on several issues. How much is it worth investing in fairer systems, given their cost? Considering a longer term than the one adopter in this model, is there a return on the investment for making a system fairer? Is it ultimately possible to reduce the concept of fairness to a scalar number or are there intangible factors and consequences which require deeper insight?

Ultimately, given the multi-objective nature of the problem, and the multiple possible definitions of spatial fairness and equity for MSS in complex urban environments, policymakers and engineers will have to work together to decide how to build mobility systems in the future, considering the tradeoffs illustrated above.



(a) Number of rebalancing operations: This graph illustrates the frequency of rebalancing operations across different scenarios.



(b) Gini scores: A visual representation of fairness in the distribution, highlighting the effectiveness of various strategies in maintaining balance.

Figure 7.3: Relationship between Gini scores and number of rebalancing operations obtained by averaging the corresponding outcomes of a 100-run Monte Carlo simulation. This comparison shows a trade-off indicating that only a subset of the proposed family of strategies can be considered Pareto-efficient in terms of both fairness and control performance.

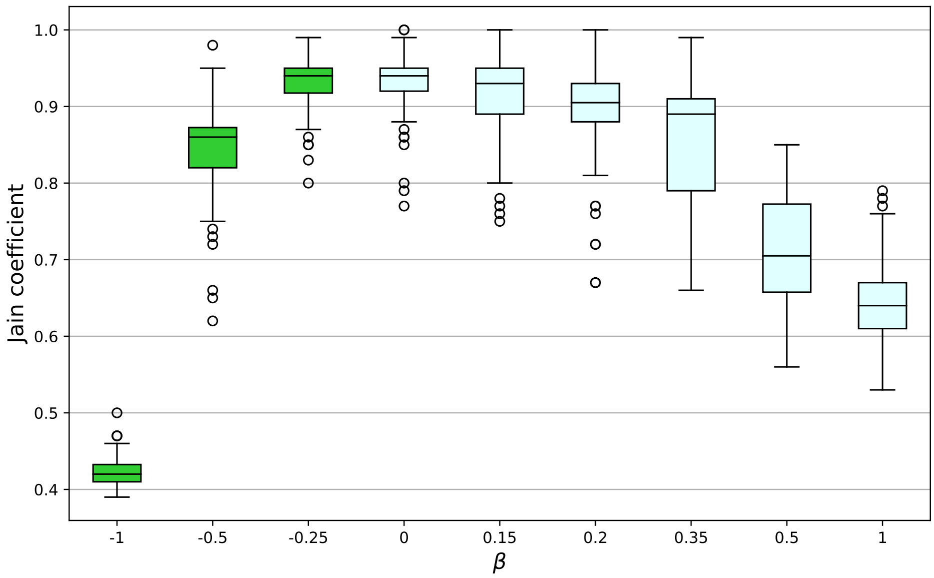
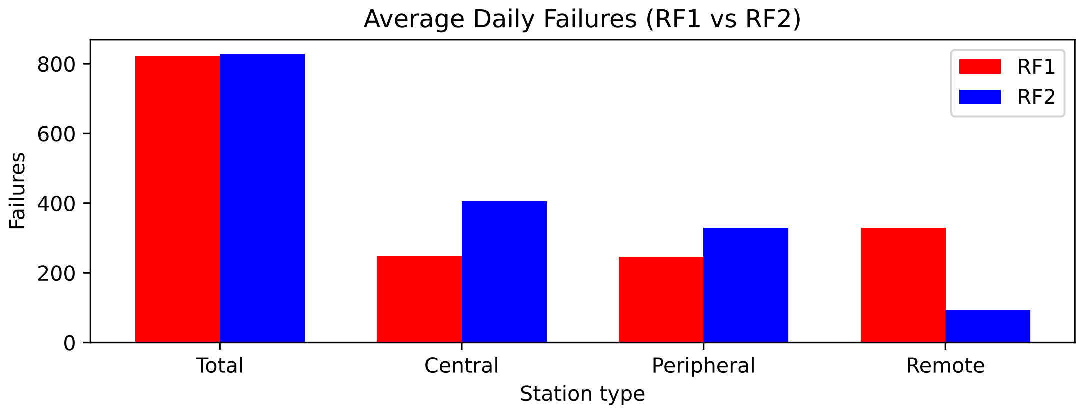
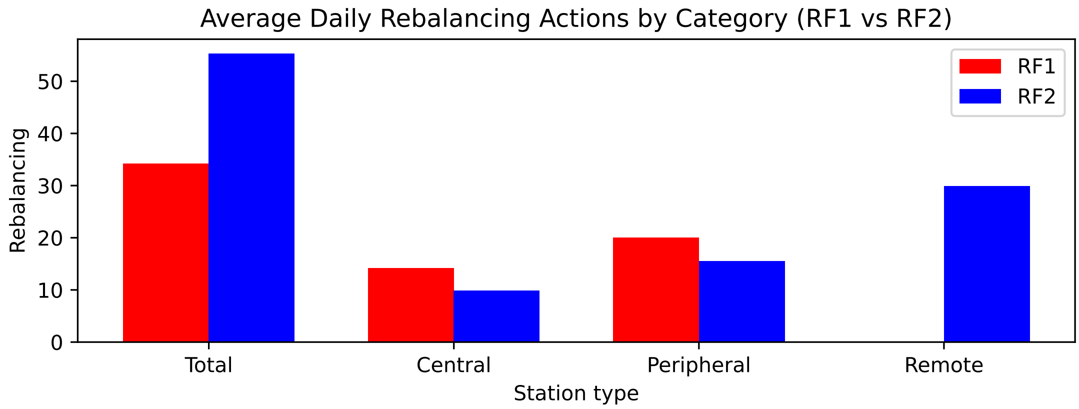


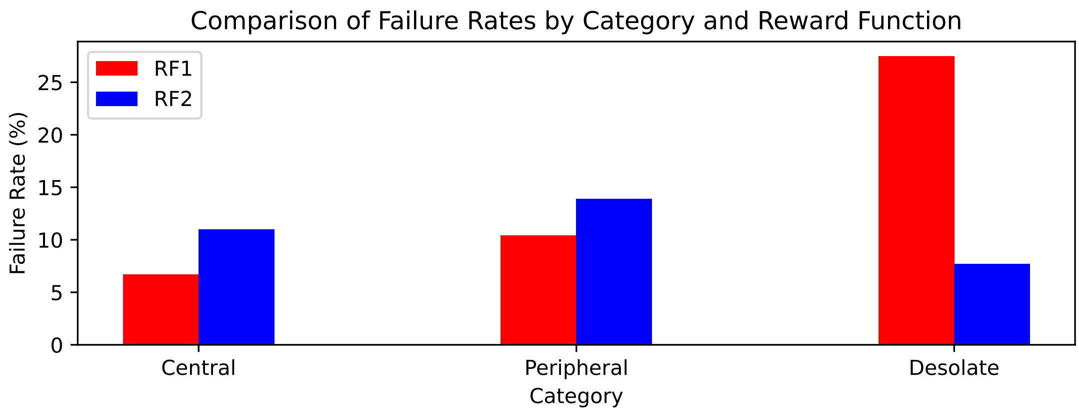
Figure 7.4: Jain indeces: A visual representation of fairness in the distribution, highlighting the effectiveness of various strategies in maintaining balance.



(a) Average daily failures across the network



(b) Average Daily Rebalancing actions



(c) Comparing the relative failure rates for different area types in the network

Figure 7.5: A comparison of RFs with $\beta = -0.5$ (unfair) and $\beta = 0.2$ (fair)

8 Conclusions and Future Works

In this study, we consider the problem of rebalancing an MSS, with a focus on spatial fairness. In particular, an MSS simulator is designed based on the definition of three categories of city areas, and tested considering several RL approaches, distinguished by their reward function definitions. Each reward function assigns a distinct weight to equalizing the distribution of rebalancing operations and failures across different area types.

Our findings, based on the cost of rebalancing operations for each strategy and the Gini index reflecting vehicle accessibility across categories, indicate that a balanced trade-off between efficiency and fairness is achievable. Indeed, a Pareto front of solutions is found: among these RL methods, any can be selected for each application, depending on the cost and fairness requirements of the specific situation. However, it is recommended to prioritize solutions with a low Gini score, since the negligence of non-central areas could have strong societal impacts, exacerbating the discrimination against residents in these regions.

In future studies, we will aim to refine the defined framework by relaxing some assumptions. In particular, we aim to define a rebalancing cost function that depends on the type of urban area vehicles are moved to, highlighting the difference in distance to cover between different areas. Furthermore, it is crucial to account for time-varying distributions of the demand while modeling, which take into consideration the effect of a failure in a given zone and a consequential lower demand characterizing that area in the future. Lastly, it would be interesting to loosen the independence assumption, hence to cope with the correlation between arrival and departure processes at different stations.

References

[1] Milton Abramowitz and Irene A Stegun. *Handbook of mathematical functions with formulas, graphs, and mathematical tables*. Vol. 55. US Government printing office, 1968.

[2] Carla Affonso, Olivier Couffin, and Philippe Leclaire. "Modeling user behavior in response to bike shortages at stations". In: *Transportation Research Part A: Policy and Practice* 145 (2021), pp. 45–60.

[3] Andrew Alleyne et al. "Control for societal-scale challenges: Road map 2030". In: *2022 IEEE CSS Workshop on Control for Societal-Scale Challenges*. IEEE Control Systems Society. 2023.

[4] Solon Barocas, Moritz Hardt, and Arvind Narayanan. *Fairness and Machine Learning: Limitations and Opportunities*. `http://www.fairmlbook.org`. fairmlbook.org, 2019.

[5] Raphen Becker et al. "Solving transition independent decentralized Markov decision processes". In: *Journal of Artificial Intelligence Research* 22 (2004), pp. 423–455.

[6] Luke Butler, Tan Yigitcanlar, and Alexander Paz. "Barriers and risks of Mobility-as-a-Service (MaaS) adoption in cities: A systematic review of the literature". In: *Cities* 103 (2020), p. 102726. DOI: `10.1016/j.cities.2020.102726`.

[7] Annelise Chardon et al. "The tricky balance of dynamic rebalancing in Bike Sharing Systems". In: *Journal of Public Transportation* 19.2 (2016), pp. 111–132.

[8] Richard J. Chen et al. "Algorithm Fairness in AI for Medicine and Healthcare". In: *ArXiv* abs/2110.00603 (2021).

[9] Yan Chen et al. "An environmental benefit analysis of bike sharing in New York City". In: *Cities* 121 (2022), p. 103475.

[10] Zheyan Chen, Dea Van Lierop, and Dick Ettema. "Dockless bike-sharing systems: what are the implications?" In: *Transport reviews* 40.3 (2020), pp. 333–353.

[11] Lu Cheng et al. "The role of bike sharing in promoting transport resilience". In: *Networks and spatial economics* (2021), pp. 1–19.

[12] Federico Chiariotti et al. "Integrating dynamic rebalancing and user incentives in Bike Sharing Systems". In: *Transportation Research Part A* 116 (2018), pp. 327–344.

[13] Federico Chiariotti et al. "A bike-sharing optimization framework combining dynamic rebalancing and user incentives". In: *ACM Transactions on Autonomous and Adaptive Systems (TAAS)* 14.3 (2020), pp. 1–30.

[14] Federico Chiariotti et al. "A dynamic approach to rebalancing bike-sharing systems". In: *Sensors* 18.2 (2018), p. 512.

[15] Marco Chiarotti, Tomaso Villa, and Xiang Yan. "Adaptability and resilience of bike-sharing services to the COVID-19 pandemic in New York City". In: *Sustainable Cities and Society* 74 (2022), p. 103138.

[16] Adri A De la Bruhèze and Frank Veraart. *The White Bikes*. Foundation for the History of Technology, 1965.

[17] Mauro Dell'Amico et al. "The bike sharing rebalancing problem: Mathematical formulations and benchmark instances". In: *Omega* 45 (2014), pp. 7–19.

[18] G. Dias, E. Arsenio, and Paulo J. G. Ribeiro. "The Role of Shared E-Scooter Systems in Urban Sustainability and Resilience during the Covid-19 Mobility Restrictions". In: *Sustainability* 13.137084 (2021). DOI: `10.3390/su13137084`.

[19] Jennifer Dill and Nathan McNeil. "Are shared vehicles shared by all? A review of equity and vehicle sharing". In: *Journal of Planning Literature* 36.1 (2021), pp. 5–30.

[20] David Duran-Rodas et al. "Demand And/oR Equity (DARE) method for planning bike-sharing". In: *Transportation Research Part D: Transport and Environment* 97 (2021), p. 102914.

[21] David Duran-Rodas et al. "How fair is the allocation of bike-sharing infrastructure? Framework for a qualitative and quantitative spatial fairness assessment". In: *Transportation Research Part A: Policy and Practice* 140 (2020), pp. 299–319.

[22] Gunes Erdogan et al. "An exact algorithm for minimizing rebalancing costs in bike sharing systems". In: *Journal of Urban Transportation* 23.4 (2015), pp. 277–292.

[23] Elliot Fishman, Simon Washington, and Narelle Haworth. "Bike share: A synthesis of the literature". In: *Transport Reviews* 33.2 (2013), pp. 148–165.

[24] Sorelle A Friedler et al. "A comparative study of fairness-enhancing interventions in machine learning". In: *Proceedings of the conference on fairness, accountability, and transparency*. 2019, pp. 329–338.

[25] Yvonne Hail and Ronald W. McQuaid. "The Concept of Fairness in Relation to Women Transport Users". In: *Sustainability* 13.5 (2021), p. 2919. DOI: 10.3390/su13052919.

[26] Corinna Hertweck, Christoph Heitz, and Michele Loi. "On the Moral Justification of Statistical Parity". In: *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (2020).

[27] Kate Hosford and Meghan Winters. "Who are public bicycle share programs serving? An evaluation of the equity of spatial access to bicycle share service areas in Canadian cities". In: *Transportation Research Record* 2672.36 (2018), pp. 42–50.

[28] A. Hosseinzadeh et al. "Spatial analysis of shared e-scooter trips". In: *Journal of Transport Geography* 92 (2021), p. 103016. DOI: 10.1016/J.JTRANGEO.2021.103016.

[29] Mario Köppen, Kei Ohnishi, and Motoyuki Tsuru. "Multi-Jain Fairness Index of Per-Entity Allocation Features for Fair and Efficient Allocation of Network Resources". In: *2013 5th International Conference on Intelligent Networking and Collaborative Systems*. IEEE. 2013, pp. 841–846.

[30] Benyamin Lahoorpoor et al. "A spatial cluster-based approach for efficient static rebalancing of bike-sharing systems". In: *Journal of Transport Geography* 77 (2019), pp. 182–193.

[31] Philippe Leclaire and Olivier Couffin. "Using graphical class diagrams to simplify mathematical model creation for static rebalancing in bike-sharing systems". In: *International Conference on Urban Transportation*. IEEE. 2018, pp. 145–159.

[32] Melinda Matyas and Maria Kamargianni. "The potential of mobility as a service bundles as a mobility management tool". In: *Transportation* (2019). DOI: 10.1007/s11116-018-9913-4.

[33] Ninareh Mehrabi et al. "A Survey on Bias and Fairness in Machine Learning". In: *ACM Computing Surveys (CSUR)* 54 (2019), pp. 1–35. DOI: 10.1145/3457607.

[34] Farzana Mehzabin Tuli, Arna Nishita Nithila, and S. Mitra. "Examining the Impact of the Covid-19 Pandemic on Shared E-Scooter Usage: A Spatial Panel Model". In: *SSRN Electronic Journal* (2022). DOI: 10.2139/ssrn.4275398.

[35] Francisco S Melo. "Convergence of Q-learning: A simple proof". In: *Institute Of Systems and Robotics, Tech. Rep* (2001), pp. 1–4.

[36] Si'an Meng and Anne Brown. "Docked vs. dockless equity: Comparing three micromobility service geographies". In: *Journal of Transport Geography* 96 (2021), p. 103185.

[37] Peter Mooney, Adam Winstanley, and Padraig Corcoran. "Dockless bike sharing: Data sharing and privacy". In: *Data Science Journal* 18.1 (2019), pp. 1–12.

[38] Stephen J Mooney et al. "Freedom from the station: Spatial equity in access to dockless bike share". In: *Journal of Transport Geography* 74 (2019), pp. 91–96.

[39] Carol Mullen and Greg Marsden. "Mobility and aging: New directions for public health action". In: *American Journal of Public Health* 102.8 (2012), pp. 1508–1515. DOI: 10.2105/AJPH.2011.300631.

[40] Susan Shaheen and Nelson Chan. "Mobility and the sharing economy: Potential to facilitate the first-and last-mile public transit connections". In: *Built Environment* 42.4 (2016), pp. 573–588.

[41] Susan A Shaheen, Stacey Guzman, and Hua Zhang. "Bicycle-sharing schemes: Enhancing sustainable mobility in urban areas". In: *United Nations Department of Economic and Social Affairs* (2010).

[42] John G Skellam. "The frequency distribution of the difference between two Poisson variates belonging to different populations". In: *Journal of the Royal Statistical Society Series A: Statistics in Society* 109.3 (1946), pp. 296–296.

[43] C Scott Smith, Jun-Seok Oh, and Cheyenne Lei. *Exploring the equity dimensions of US bicycle sharing systems.* Tech. rep. Western Michigan University. Transportation Research Center for Livable . . ., 2015.

[44] Edward Soja et al. "The city and spatial justice". In: *Justice spatiale/Spatial justice* 1.1 (2009), pp. 1–5.

[45] Lin Su, Xiang Yan, and Xilei Zhao. "Spatial equity of micromobility systems: A comparison of shared E-scooters and docked bikeshare in Washington DC". In: *Transport Policy* 145 (2024), pp. 25–36.

[46] Qin Su et al. "An evolutionary algorithm for optimizing Bike Sharing rebalancing operations". In: *Transportation Science* 56.2 (2022), pp. 459–478.

[47] Richard S. Sutton and Andrew G. Barto. *Reinforcement Learning: An Introduction*. MIT Press, 2018.

[48] *The Gini Coefficient*. Basic Concepts in Statistics and Epidemiology, 2018.

[49] Daniel Todd et al. "A global comparison of bicycle sharing systems". In: *Transportation Research Part D: Transport and Environment* 96 (2021), p. 102867.

[50] Eugenia Villa, Valentina Breschi, and Mara Tanelli. "Fair-MPC: a control-oriented framework for socially just decision-making". In: *arXiv preprint arXiv:2312.05554* (2023).

[51] Feng Xin, Zhiwei Wang, and Xiao Liu. "Impact of the COVID-19 pandemic on bike-sharing usage patterns: A multi-city study". In: *Journal of Transport Geography* 95 (2023), p. 103115.

[52] Li Zhang et al. "Challenges and strategies for free-floating Bike Sharing Systems". In: *Journal of Sustainable Mobility* 9.3 (2022), pp. 58–72.

[53] Ming Zhang, Wei Li, and Fang Wang. "Integration of bike maintenance for optimal rebalancing in bike-sharing systems". In: *Journal of Sustainable Mobility* 5.2 (2018), pp. 103–117.

[54] Rui Zhu et al. "Understanding spatio-temporal heterogeneity of bike-sharing and scooter-sharing mobility". In: *Computers, Environment and Urban Systems* 81 (2020), p. 101483.