MINISTRY OF EDUCATION

TECHNICAL UNIVERSITY
OF CLUJ-NAPOCA, ROMANIA

Field: Systems Engineering

# PhD THESIS
## – ABSTRACT –

## Optimization strategies of the operations from the logistics systems

**PhD Student:**
**Eng. Horațiu Florian**

**PhD Supervisor:**
**Prof. Eng. Adina Mariana Aștilean, PhD**

**Examination committee:**
Chair: Prof. Dr. Ing. **Petru Dobra**, PhD – Technical University of Cluj-Napoca;
PhD Supervisor: Prof. Eng. **Adina Mariana Aștilean**, PhD – Technical University of Cluj-Napoca;
Members:
- Prof. Eng. **Ioana Făgărășan**, PhD – Politehnica University of Bucharest;
- Prof. Eng. **Gheorghe Dorin Șendrescu**, PhD – University of Craiova;
- Prof. Eng. **Tiberiu Ștefan Leția**, PhD – Technical University of Cluj-Napoca.

**– Cluj-Napoca –**
**2024**

# Introduction

This PhD thesis provides an overview of the current problems and approaches regarding the optimization of operations in logistics systems. The study was oriented towards identifying and solving the unbalanced operation situations within the bike-sharing systems. From this perspective, the research presents new solutions developed for the purpose of predicting the demand for resources needed for rebalancing in bike-sharing systems and methods that can be applied to correct the imbalances.

For this purpose, several methods and models have been proposed, the most important of which are: the model of the temporal variations of the number of bikes in the stations of bike-sharing systems, the model of the behavior of the groups of stations, the grouping method based on the measurement of the degree of similarity of stations, the method of generating priority lists regarding the visiting order of the bike stations in order to rebalance them, the method of assigning stations to the agents that carry out the rebalancing and the method of integrating routing algorithms within the proposed relocation strategy.

## Problem Analysis

In the last decades, the field of logistics has had a permanent upward evolution, materialized by the diversification of the approaches and the increase in the number and usefulness of applications. Optimizing the operations of a logistics system determines its competitiveness.

The identification of the classes of operations and the analysis of the characteristic processes are the basis of the subsequent optimization strategies. In the general context of the studies aiming the logistics systems management, the following are regarded:

1. Ensuring the correct and safe delivery of goods and services. This involves carrying out loading and unloading operations to the required standards, avoiding blockages caused by moving goods, packaging while avoiding the damage of goods, and avoiding confusion when labeling packages.

2. Optimizing the stocks, so that there is no excess or shortage of goods, achieved by: providing sufficient storage space, avoiding over-sizing the warehouse, creating an operational framework, correct stacking that avoids damage, ensuring security, and having a precise inventory system.

3. Consolidating the freight, which involves: avoiding confusion between similar products, easy tracking of products, full loading of the transport vehicles, ensuring synchronization of the arrival times.

4. Preparing the commercial assortment, having the following problems: the correct planning of stocks to avoid cases of oversupply or undersupply, the creation of an operational framework between departments (purchases, sales, storage etc.), and the correct estimation of the demand.

5. Transporting, having the following problems: avoiding traffic jams, ensuring security, avoiding inefficient routes, and choosing routes in a flexible manner, which takes into account different risks (fluctuating product demand, adverse weather conditions etc.).

6. Increasing operational efficiency by identifying and eliminating time and other cost related waste throughout the supply chain.

7. Satisfying customers' needs by providing a high level of service and meeting their requirements.

The logistics systems include a variety of categories, the most important being:

- Supply logistics systems
- Production logistics systems
- Distribution logistics systems
- Reverse logistics systems
- Military logistics systems
- Medical logistics systems
- Transport and infrastucture logistics systems
- Supply chain management logistics systems

In a world that is marked by fast urbanization and global trends in sustainable development, the field of logistics plays a vital role within the complex systems. A distinctive aspect of the logistics systems is the emphasis on the integration of the functionalities of the various components, the result materializing in coherent and well-coordinated entities. The efficiency of a logistics system is determined by its ability to ensure the smooth movement of goods, from the point of origin to the destination, in a way that minimizes costs and optimizes the use of resources. Moreover, the ability to react quickly to environmental changes and to anticipate future demands are essential characteristics of successful logistics systems.

Beyond costs, business logistics is increasingly important, being influenced by certain factors, such as:

- **The market globalization**. The world trade is in continuous development, and the trend of globalization of the markets is also growing. Moving products from the point of manufacture to the point of consumption on a global scale has obvious logistical challenges.
- **The deregulation**. The transport has changed from a highly regulated industry to an increasingly free market industry. The effects have been the diversification of options, increased complexity in logistics services and costs, and more opportunities to improve business operations.
- **The quality of services provided to customers**. The deregulation, global markets and other factors create a more competitive business environment, directly reflected in the need to ensure supply chains that can deliver products on time and that can adapt to fast market changes.
- **The technological development**. Accelerating technological advances are significantly changing and improving logistics operations. For example, there are: automated product barcode tracking, transportation asset management via satellite communications, e-commerce, and computerized decision support.
- **The environment**. The current and anticipated environmental regulations have significant implications for logistics and can have a fundamental impact on the locations of utilities, factories, storage facilities and recycling centers.

In a broader context, the increased interest in sustainable urban mobility has led to the development and implementation of bike-sharing systems. They were introduced to the market with the aim of providing an environmentally friendly and cost-effective alternative to urban transport, especially in congested areas. Optimizing resource management and ensuring an efficient flow are just two of the most difficult problems these systems face.

In this system, the bikes are made available to the public for short-term shared use for a reduced fee or free of charge. Generally, they allow the user to borrow a bike from one parking space and return it to another parking space belonging to the same system. A key property of these logistics systems is that the resuource base remains within the system, i.e. it is not passed on to the customer or any other entity. Therefore, unlike other categories of logistics systems, a large part of the problems encountered are focused on the management of the resource base.

Some of the specific problems that may arise within bike-sharing systems are:

- Many bikes can end up being parked outside areas of interest where the level of demand is very high, after only a few uses; the causes may be related to the polarization of the traffic. Moreover, there can be areas where the number of the parking spots is below the required level.

  In the dynamic case, there are some practical solutions to ensure a proper number of bikes and accessible parking spots in a reasonable time frame. Some of the current solutions are based on user-driven rebalancing and focus on finding ways to convince users to park borrowed bikes at other destinations than what they would find as the most convenient [1,2]. Operator rebalancing is another suitable solution to ensure the availability of the resources, being particularly effective in cases requiring urgent relocation. For these cases, different rebalancing methods have been proposed, with the aim of minimizing the number of stations where customers cannot find bikes or parking spots [3, 4]. There are also proposals to combine the two approaches [5].

- The phenomenon of bike station emptying is non-deterministic, which can lead to situations for which the rebalancing vehicles are not prepared.

  The prediction of bike demand within a certain acceptable margin can be performed both at the station level and at the level of a group of stations with similar behavior. In the article [6], a similarity measure method is proposed for clustering stations, in which the station loads are used as the basic feature used by the model. Even though it involves more computational effort, station-level load prediction can provide superior results in certain situations. The authors of [7] use a method that is based on bike numbers to predict hourly demand for any station. Previous journeys and external factors such as weather, time of the day and the available infrastructure are taken into account for the prediction model.

- The system usage patterns are often unclear, and the impact of the external factors such as weather or infrastructure is difficult to measure.

  The article [8] analyzes the historical information to extract and use critical behavior patterns for dynamic rebalancing operations. The disfunctional state of the bikes is tracked to optimize the operations in bike-sharing systems, with article [9] proposing a model based on integer programming algorithm and article [10] proposing a greedy heuristic method.

- The batch of bikes is large and consequently the number of stations, which makes it difficult to track them and control operations for the entire system.

  An in-depth analysis of the detailed behavior of large-scale systems is often impractical. Thus, reductionist approaches are needed, that allow the study of simpler subsystems, or other elements that provide an understanding of the system's behavior. In this sense, dividing the system into groups plays a crucial role in solving the problem. By applying clustering techniques using attributes such as location [11, 12], usage patterns [13, 14] or usage preferences [15, 16], valuable information can be obtained from the system in order to to optimize the allocation of bikes [17], the location of stations [18] or the planning of routes [19].

- The need to build case-specific frameworks for making bike relocation decisions, assessing the need for stations to be rebalanced, determining the required number of bikes to transport to the stations, or deciding the number of vehicles needed to serve a set of stations.

  In this context, a travel consistency indicator is used in the paper [20] as a feature descriptor for a machine learning method, with the aim of controlling the static rebalancing operation, increasing the degree of demand satisfaction and decreasing at the same time the number of station visits by rebalancing operators. The authors of the paper [21] developed a simulation framework to evaluate different rebalancing strategies.

- The efficient relocation of bikes by choosing routes that satisfy several objectives (fuel consumption of vehicles involved in relocation, reduced time for transport, avoiding the loss of customers, reducing the time stations operate in critical state).

  In order to increase the performance of the system, many strategies have been proposed, using exact, hybrid, heuristic, metaheuristic and hyper-heuristic algorithms [22, 23]. In order to find new routing strategies, some studies adapt and develop typical formulations for the capacity-constrained vehicle routing problem. This considers the generalization of the traveling salesman problem that has been intensively studied in the literature, in deterministic and stochastic variants. Many applications, using different assumptions and constraints, have been defined in this context, with solutions proposed for both deterministic and stochastic approaches. The stochastic approaches to capacity-limited vehicle routing problems and the methods developed to solve them, over two decades, are presented and briefly analyzed in [24, 25] papers. In [25], several categories of representative solutions were analyzed, while in [24], the authors focused on presenting problem types and their approaches.

The bike-sharing systems have a number of important characteristics that place them in the category of complex logistics systems. The groups of operations and activities carried out within them cover the following aspects:

- **Resource management**: In bike-sharing systems, the resources are the bikes and the docking stations, the intermediate storage and the fleets of vehicles used in rebalancing operations.
- **Route optimization**: Finding the most suitable routes for customers, but also the optimal routes for the vehicles involved in station rebalancing operations, are essential to minimize costs.
- **Spatial planning**: Ensuring the necessary spaces for the purpose of permanent and temporary storage of all resources, as well as establishing their locations, are elements that influence the long-term efficiency of the entire system's operation.
- **Continuous monitoring of system status**: Gathering and storing information about the location and status of all existing resources, using current techniques and technologies, are important factors that can contribute both to a better system operation and to the development of new optimization strategies.
- **Management of uncertainties**: User requirements may vary by time of day, season or special events. Adapting to fluctuations in demand is difficult.
- **The need for maintenance**: Includes checking and repairing bikes or docks.
- **Inventory planning**: Inventory planning and management, to ensure resources are available when needed, is essential.
- **Interaction with customers**: Includes information on bike availability and rental process.
- **Regulations and legislation**: The regulations and legislation governing the operations and safety of resources or users must be applied.
- **Capacity planning**: It is important to plan and adjust capacity to meet the fluctuations in demand and ensure adequate service to users.

- **Collaboration with third parties**: Logistics and bike-sharing systems may involve collaboration with third parties, such as suppliers or external partners, to secure supplies or extend services.
- **User experience**: In both types of systems, the aim is to improve the user experience related to the rental processes and ensure high quality services.

## The actuality of the topic. The need to optimize the operations in bike-sharing systems

The last decades have been marked by modernization and urbanization, processes that have led to new levels of traffic congestion, endangering the viable development of urban areas. The congested traffic is associated with high levels of chemical and noise pollution, accidents and economic losses. In this context, a new field has emerged, called ”Smart Mobility”. Emerging technologies such as autonomous driving have the potential to transform the way mobility systems work [7]. However, many solutions proposed to reduce the negative effects of congestion and pollution also have disadvantages and unwanted side effects [26]. Even though it looks promising, the large-scale implementation of existing solutions in this field remains a challenge, requiring time, complex analysis and incurring considerable costs [27, 28].

A palette of available measures was considered as an alternative approach, which involves restricting the use of existing vehicles and replacing them with other means of transport [29]. Also, in recent years, transport systems have been adapted to meet the transport needs of certain categories of users [30]. In this sense, cycling is seen as an environment-friendly [31] solution that can shorten travel time, if appropriate conditions are provided. Consequently, there are currently a large number of bike-sharing systems, involving a total of millions of bikes [31].

For carrying out in good conditions the activity of renting bikes from a certain lot, from which customers can take bikes, in order to move to different parts of the city, several categories of problems related to capacity, transfer, reliability and integration, must be solved in bike-sharing systems, which are also found for other modes of transport [32].

Although some progress has been made, many articles from literature mention the fact that, the rebalancing of the stations in bike-sharing systems remains an open problem. The studies carried out and the alternatives for solving the specific problems, elaborated and implemented in this thesis, were motivated by the need to find solutions that would ensure superior performance for the mentioned categories of applications and for those with similar characteristics.

## The thesis objectives

The research undertaken in this PhD thesis was focused on the development and implementation of new strategies and methods, in order to solve some of the current optimization problems from the logistics field.

In this context, the proposed objectives were:

- **The first objective:** modeling the behavior of the bike-sharing systems, taking into account:
  - the moments of the day and the durations of the periods when the parking lots of the stations are either completely occupied or completely free (the stations are in critical state);
  - the characteristics of the fluctuation over time of the number of bikes in the stations;
  - the evolution of the number of stations in critical state within some subsystems.
- **The second objective:** developing of methods for analyzing the behavior of bike-sharing systems taking into account:

  - the system load;
  - the probability of the transition of subsystems from one state to another;
  - the number of resources required for rebalancing and planning their use.

- **The third objective:** developing, implementing and testing the methods of rebalancing bike stations, that integrate

  - statistical models;
  - algorithms belonging to artificial intelligence.

- **The fourth objective:** modeling the variations over time of the number of bikes in the stations of bike-sharing systems, using statistical methods.

- **The fifth objective:** evaluating the system performances resulting from rebalancing operations, with respect to:

  - the lost customer rate;
  - total operating times under critical state of the stations;
  - the length of the rebalancing routes;
  - the number of bikes that were relocated.

- **The sixth objective:** validating the proposed solutions, based on a comparative analysis between the results obtained using reference algorithms in the scope of the method, respectively the proposed algorithm.

- **The seventh objective:** extracting the features of a large-scale bike-sharing system using the recorded data of the system Citi Bike New York.

## The content of the thesis

The PhD thesis consists of 2 main sections and 9 chapters, corresponding to distinct parts of the undertaken research, which had as its main goal the optimization of logistics systems' operations. Below is a summary for each section and chapter, respectively.

The chapters of the first section have complementary content, each contributing to the formation of an overview of the categories of applications, their specific requirements, existing solutions and proposed approaches, in order to find new high-performance solutions.

In order to identify the research directions that are of interest to the logistics system, **Chapter 1** presents an overview of the problems of the logistics field, in direct connection with the main categories of logistics systems and the classes of operations related to them. The economic and social factors that led to the evolution of concepts and the need to diversify existing approaches are highlighted. In this context, the categories of bike-sharing systems, being complex logistic systems, are mentioned, in which case it is necessary to continue the research efforts to transform them into a valid way for urban transport, by reaching the expected level. Based on the evaluation of the published results, the problems for which the existing solutions are not satisfactory are identified, the rebalancing of the bike stations being one of them. Based on the mentioned aspects, below, the thesis objectives are listed and, at the end, the summary of its chapters is presented.

**Chapter 2** details the presentation of the main aspects that directly influence the way decisions are made in logistics systems. The connection between the grouping of optimization strategies and the specificity of the categories of activities to which they are intended is highlighted. The role of modeling the behavior of logistics systems in the generation of operating alternatives is emphasized. The common elements found within the models are mentioned, regardless of the category for which they are built. The difficulties and limitations of the modeling and development of activity optimization methods are highlighted, both depending on the categories of applications and taking into account different approaches.

**Chapter 3** presents a study that mainly analyzes recently proposed solutions to solve the main categories of problems in the logistics field. The mentioned articles are of interest both from the point of view of orientation towards new theoretical approaches for the logistics field, and from the point of view of the emphasis they place on the necessity and possibility of integrating recent techniques and technologies in order to find solutions with a high degree of practical applicability. A wide range of solutions and optimization strategies that have been applied in different logistics applications are examined both in terms of their demonstrated usefulness and their limitations. An aspect of interest is related to the possibilities offered by certain research directions, in order to develop innovative, useful and widely applicable solutions. In this sense, a subchapter is dedicated to the bike-sharing systems and to the needs to rebalance its stations.

The second section focuses on the personal contribution, by establishing the common basic concepts for all the case studies that were carried out, the presentation of the representative applications for the related probles, the proposed solutions, the results obtained and the conclusions.

In **Chapter 4**, the hypotheses are formulated and features and structures are presented that are used in several models, methods, algorithms and subsequent applications presented in the current work. .

In **Chapter 5**, the problem of the variation of the number of bikes in stations at different times of the day is resolved, and a model has been proposed to reflect the variation of the demand during peak hours. The model was created using probabilistic Petri nets and validated using the available data in the public database of the Citi Bike New York system. The obtained results led to the conclusion that, for different time intervals, the proposed model falls within the accepted tolerance interval. The usefulness of the model lies in the fact that it is able to provide information in advance related to possible critical states and offers a suggestive way of tracking the variations in the behavior of the resources.

The research undertaken in **Chapter 6** pursued a similar objective to that addressed in the previous chapter. The major difference is how the rebalancing is planned, targeting groups of stations with common characteristics. Using the existing information about the loads of bike stations, the behavior of the stations was studied. In this sense, a new method for determining the similarity of behavior for grouping the stations was developed, and, by using the obtained groups, it was possible to identify a model for the behavior of groups of stations. By observing the moments of the day when the stations are in a critical state was the basis of the method for determining the similarity of their behavior, modeling the behavior of the group of stations being carried out using Markov chains. Based on the obtained results, the location of the subsystems with specific features can be identified. In this case, the subsystems that require rebalancing, and the corresponding resource requirements, can be determined.

**Chapter 7** presents a study that is motivated by the need to increase the performance of the rebalancing strategies and the methods applied in bike-sharing systems. Using recorded information from the Citi Bike system regarding the number, location, and loads of stations and trucks, a high-performance rebalancing routing method was developed, implemented, and tested. The basis of the method is a modified genetic algorithm, with the role of determining a rebalancing order for the subsystems' stations, and an inference mechanism, with the role of assigning the rebalancing tasks to the trucks. The results, obtained by the applications in which the method was implemented, highlighted reduced costs of the rebalancing routes, satisfying the conditions for avoiding the loss of customers and ensuring a uniform distribution of tasks to trucks.

In **Chapter 8**, a comparative study is presented, which aims to evaluate the performances of the rebalancing method, proposed in chapter 7, in the variants in which the modified genetic algorithm is replaced by other reference algorithms: ACO (Ant Colony Optimization), HHO ( Harris Hawks Optimization), the Tabu Search Algorithm (TSA) and the standard genetic algorithm. For this purpose, histograms were created corresponding to the data describing the results of applying each algorithm within the method, for several cases, in order to cover several existing situations that characterize the behavior of real bike-sharing systems. The use of histograms allows determining the values of the performance indicators for the obtained routes, these indicators reflecting the cost reduction and scalability of the proposed methods. The outcome of the analysis is that the proposed modified genetic algorithm has superior performance compared to the reference algorithms in all cases, except for the Harris Hawks optimization, compared to which it shows superior performance only in the case of small subsystems.

**Chapter 9** includes the general conclusions and the main contributions of the thesis, with an emphasis on innovative elements and including aspects related to implementation.

The thesis also contains the author's lists of abbreviations, references, tables, figures and publications.

Next, the proposed methods and models are presented, targeting the optimization of the operations within the logistics systems, discussing the results obtained and drawing conclusions to which they led.

# Modeling the temporal variations of the number of bikes in the stations of bike-sharing systems, using probabilistic Petri nets

The demand prediction is one of the main research topics of bike-sharing systems, with the main goal of preventing situations where certain bike stations become unusable, due to the presence of an inappropriate number of bikes at the station. In this sense, a model capable of reflecting the evolution over time of the number of bikes and docks available in different bike stations, spread over an urban area, is presented. Using the probabilistic Petri nets defined in [33] and historical data, the model describes the changes in bike demand at the station level at different time intervals, with the stations being part of groups with different characteristics. These characteristics differ depending on the positioning of the areas in the city. The peripheral or central position of the bike station, the presence of various institutions, transport stations or tourist attractions in its immediate vicinity could have an important influence on the estimated value of requests over a shorter or longer period of time. Designed to reflect the dynamics of demand during peak periods, the proposed model can also describe the likely evolutions of the number of bikes available at other time intervals during a day in different seasons.

For modeling purposes, a probabilistic Petri net was used, defined in [33] as follows:

$$PPN = \{N, F, \lambda, M\}, \text{ unde} \tag{1}$$

$N = (P, T, Pre, Post)$ corresponds to the structure of the Petri net consisting of: the set of locations, $P$; the set of transitions, $T$; the functions $Pre$ and $Post$, defined on the set of oriented arcs from locations to transitions, respectively, on the set of oriented arcs from transitions to locations;

$F$ is a function defined on the set of locations $P$, which associates to each location $p$ a probability distribution, $fx(t)$ defining the probability for the postset of $P$ ;

$\lambda$ is a function defined on the set $T$, which imposes constraints on transitions;

$M : P \to N$ represents the token.

The initial probability value of each token is equal to 1 and changes after each execution of a transition, the previous value of the probability of the token being multiplied by the value of the probability associated with the arc entering the transition. An example of how transitions are executed is shown in Figure 1:

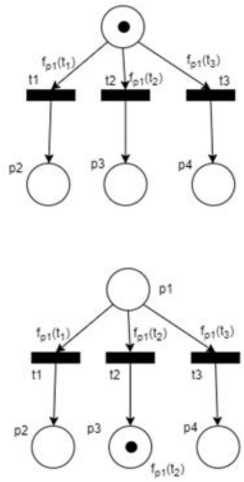


Figure 1: The evolution of the token after the execution of transition t2 [34].

The data were collected over a time period of specified duration, at equal time intervals. It is assumed that initially the initial number of bikes is known and the station load is balanced. From this starting point there are several possible developments, over successive time intervals.

The general structure of the model is illustrated in Figure 2:

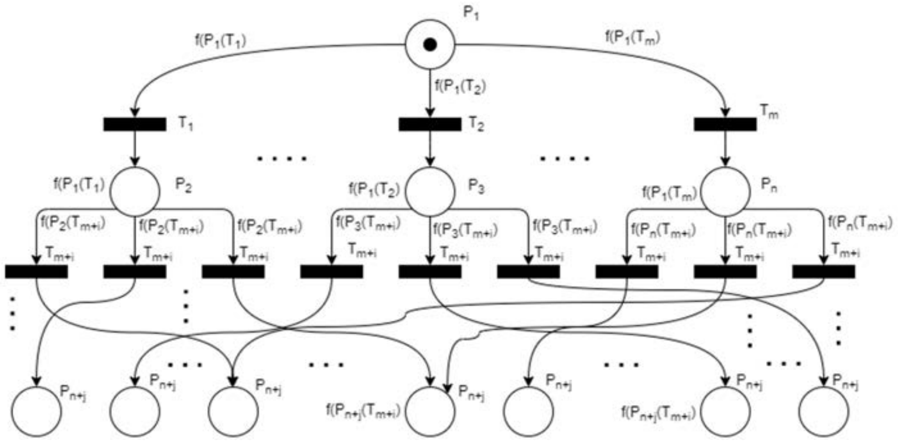


Figure 2: The general structure of the model [34].

Each marked location will correspond to a likely system state. The probable number of bikes in the station at a certain moment can be determined by taking into account the probability values associated with the previous segments belonging to the followed route. Transitions represent events corresponding to the passing of a certain time interval.

To determine the probability of following a route among the possible ones, two categories of data were considered: the number of bikes registered at the station at a given time (corresponding in this case to the end of the set time intervals, the same for all days ) and the total number of previous records leading to approximately the same number of bikes. The probability associated with a certain route segment, i.e. an arc connecting a location and a transition, will be calculated as the ratio between the number of past days in which, in the same time interval, the same number of bikes was registered, and the number total records (days of the period considered). For practical reasons, to avoid a large but irrelevant number of possible routes, small differences between the number of bikes at given times were not taken into account.

The main advantages of this model are related to the high degree of confidence regarding the estimated values and the suggestive way of tracking resource variations (available number of bikes, respectively free seats). Moreover, the influences of different factors on the estimated final result can be highlighted, including the time intervals that make up the entire monitoring period.

The values obtained at the outputs of the model can serve rebalancing purposes, providing a priori information on possible critical states of bike stations.

## Similarity measure for station clustering in bike-sharing systems

In the context of bike-sharing systems, the unbalanced temporal flows associated with commuting cause one of the most common problems associated with such systems. Bike sharing systems tend to drift from an equilibrium point towards an unbalanced state where there are stations with no bikes available and stations that are full. Maintaining bike and dock availability is one of the challenges faced by operators of bike sharing systems.

Most of the bikes sharing systems have a strategy in place for rebalancing the number of bikes available at each station to a predetermined level. The rebalancing problem is looked at as an optimization problem in which routes are planned out based on predicted bike use. Doing this at station level is a computationally intensive task. To reduce prediction complexity, most studies employ a grouping or clustering solution to predict bike use for a given time period.

In this context, it is presented a new similarity measure [6] that can be used to cluster stations based on time periods when bike availability is below a certain threshold, thus outlining the clear need for rebalancing efforts to be directed at a group of stations at a certain time period.

In order to group the stations of the system, the criteria that determine the membership of each individual to each group must be specified. In this sense, the following elements and characteristics can be taken into account:

- the bike flow;
- the geographical position;
- proximity to centers of interest (e.g. tourist areas, parks, public institutions);
- density of nearby stations.

Another option may be to take into account the load of the station [6]. Expressed either as a real number, having values in the range [0;1], or as a percentage, (0-100%), this is the normalized number of the bikes inside the stations.

Given the mentioned definition, the load, $L$, has the following formula:

$$L = \frac{nb}{nb + nd} \tag{2}$$

where $nb$ represents the number of bikes that are parked to the station and $nd$ represent the number of available docks of the same station, which indeed excludes the docks that are not functional anymore.

In Figure 3, the load, $L$, over time evolution of an arbitrary chosen station from Citi Bike is presented.

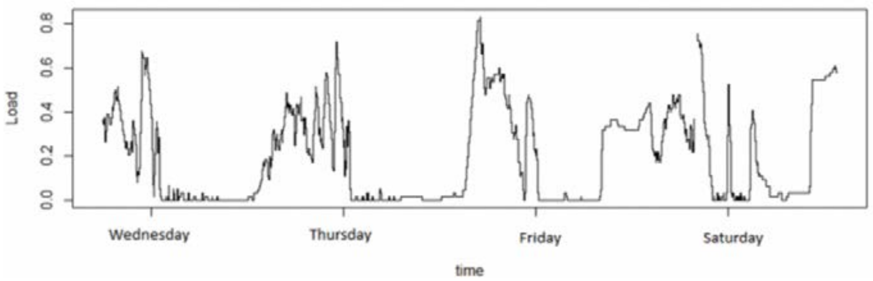


Figure 3: The evolution over time of the load.

An important observation about this representation, which can be extended to the rest of the stations in the collected data set, is that the value of the function less often fluctuates around the minimum value.

This observation regarding the evolution over time of the load, and which allows the elimination from the evaluation of some extended parts of the representation, allows abstraction from a large part of the represented values.

Thus, after obtaining this characteristic, a single numerical indicator will result that describes the station in a particular way, and which is based on the recorded load values.

To obtain the value corresponding to the numerical indicator, the data remaining after the elimination process are retrieved and stored. For each individual entry and for the set of values obtained, the average time at which the data was collected is calculated.

The threshold taken into account in order to eliminate the minimum value is calculated according to the relation:

$$k_\eta(x) = \begin{cases} 0, & x > \eta \\ 1, & x \leq \eta \end{cases} \tag{3}$$

where $k_\eta : [0, 1] \rightarrow \{0, 1\}$ is the characteristic function of the set of possible values of the load; $x$ represents the load value; $\eta$ is the constant value set for the threshold.

*The average hour corresponding to the empty state*, $\overline{h}$, is calculated with the following expression and is the essential feature used to determine the degree of similarity.

$$\overline{h} = \frac{\sum_i h_i \cdot k_\eta(L_i)}{\sum_j k_\eta(L_j)} \tag{4}$$

where $h_i$ represents the moment when the load information, $L_i$, was collected.

Based on the hourly average value, there is the possibility to classify the bike stations.

For example, stations in residential areas tend to deplete in the morning during the working week due to the fact that the majority of urban dwellers have similar working schedule, while for the business centers in which the bike riders have the activity, the corresponding stations tend to deplete around evening, marking the return from the work of the employees.

It is important to note that all of these trends superposed and with noise applied is affected by seasonal change and weather conditions.

How classification was done in this paper is by partitioning the full hour interval of each day in smaller subintervals and assigning each subinterval a class. Ideally, it is desirable to obtain, for the stations of the same class, geographical clusters in which stations of the same class stay in proximity. Obtaining a uniform geographic distribution of all stations implies low performance of the algorithm chosen.

The set of steps for classification is described as follows:

1. For each station, the load value is calculated.

2. Based on the load value, the average hour corresponding to the empty state for each bike station is calculated.

3. For the purpose of graphic representation, each station is assigned a color corresponding to the time partition that corresponds to the average hour corresponding to the empty state.

4. All stations are represented by colored points in a system of coordinate axes, having the longitude on the X axis and the latitude on the Y axis.

The input data on which the load over time is computed represents bike sharing system's information, collected once per 5 minutes, throughout the whole day, about the following:

- The location, via geographical coordinates (latitude, longitude);
- The number of available docks;
- The number of available bikes;
- The number of disabled docks and bikes.

The results obtained by applying the proposed method on Citi Bike system data, partitioned in several intervals of around three and a half days, using the parameter set described in the chapter before, are showed in Figure 4. The stations are grouped in class-specific colors.

It can be observed that there are six different clusters that denominate each a specific hour interval, with the following correspondence: RED – [0; 4), ORANGE – [4; 8), YELLOW – [8; 12), GREEN – [12; 16), BLUE – [16, 20) and PURPLE – [20; 24).

The X axis represents the Longitude axis while the Y axis represents the Latitude axis. Each colored point corresponds to one bike sharing station.

In this case, the majority of stations are of color blue and yellow, meaning the majority of stations are considered as empty either in the morning part of the day, between 8 and 12, or in the evening part of the day, between 16 and 20.

The transition from the blue to the yellow groups is gradual, with green intermediate stations representing the interval [12; 16).


Figure 4: Data grouping based on the average hour corresponding to the empty state.

The traffic is congested in the morning and evening in most metropolitan areas, which matches the measurements made.

From the point of view of the derived comparative results, it can be seen that the biggest areas, colored blue and yellow, tend to have the same location, and outside of that, the color of the stations fluctuates pronouncedly.

In the validation method, the defined disparity at the station level was determined and shown to be within the established limits, taking into account a specified deviation. The resulting groupings can be used for both rebalancing vehicles routing and resources demand prediction.

# Modeling the behavior of the station groups

Docked bike-sharing systems often experience load imbalances among bike stations, leading to uneven distribution of bikes and to challenges in meeting users' demand. To address the load imbalances, many docked Bike Sharing Systems employ rebalancing vehicles that actively redistribute bikes across stations, ensuring a more equitable distribution and enhancing the availability of bikes for users. The determination of the number of rebalancing vehicles in docked Bike Sharing Systems is typically based on various criteria, such as the size of the system, the density of stations, the expected demand patterns, and the desired level of service quality. This is a determining factor, in order to increase the efficiency of customer service at a reasonable cost. For this purpose, a Markov chain-based model was proposed, offering support for determining the optimal fleet size, when the rebalancing operations are performed for highly dynamic bike stations, within a subystem.

To derive a behavioral model of the subsystem, a state vector $X(t)$ ($t \rightarrow$ time dependent) is assigned to the subsystem:

$$X(t) = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix}, \tag{5}$$

where $m$ is the number of bike stations contained inside the subsystem and $x_i$ is a function attached to $B$, defined as follows:

$$x_i = x(B_i) = \begin{cases} 1 & B_i \text{ has CRITICAL STATE} \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

*The overall criticality* of the subsystem, $\overline{x}$, is then extracted from the state vector, to serve as input for the states of the model proposed. The overall criticality is defined as the percentage of bike stations from the subsystem that are in a critical state, formulated as below:

$$\overline{x} = \frac{\sum_{i=1}^{m} x_i}{m} \quad (\cdot 100\%) \tag{7}$$

Figure 5 summarizes the method of computing the overall criticality ($\overline{x}$), using the information of the available bikes ($k_i$) and the capacity ($C_i$) of all the bike stations ($B_i$) from a subsystem comprised of $m$ bike stations, $1 \leq i \leq m$.

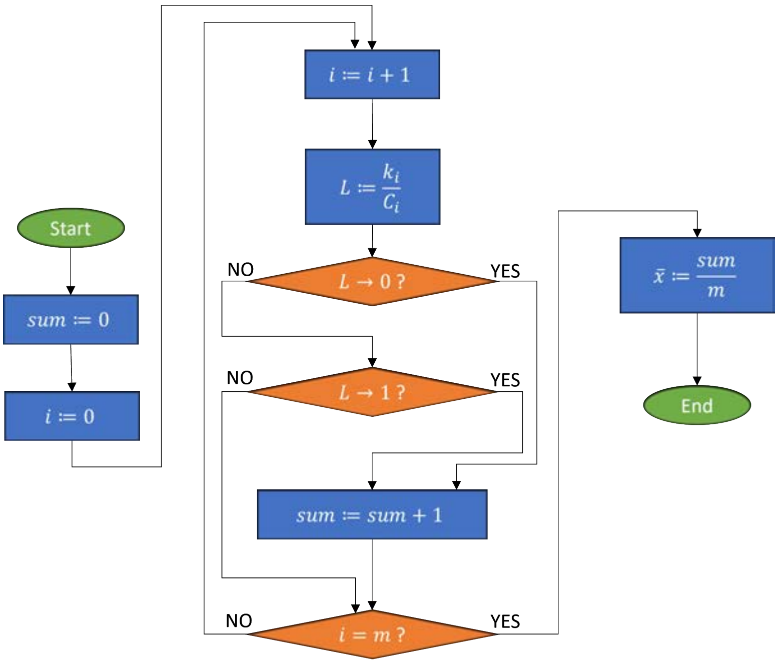


Figure 5: The diagram of the method that determines the general critical state of the subsystem.

When using the Markov chain-based model for describing the behavior of the system, two issues must be addressed: the state explosion problem and the interdependence between states and the number of bike stations. Both issues are resolved by a quantization strategy, described as follows.

Until now, we have associated a specific value from the domain $[0\%, 100\%]$ to the subsystem state vector. The resolution of the measurement is dependent on $m$ (defined in Equation (7) as the number of bike stations). The states of the MC must be independent of the number of bike stations, otherwise subsystem scalability is not ensured (i.e., if $m$ is increased, the number of states should remain constant; otherwise, the state explosion problem appears).

Therefore, a quantization is performed on the value range $[0\%, 100\%]$ of $\bar{x}$, described as follows:

$$\begin{array}{ccccccccc} [0\%, 100\%] = & [0\%, \upsilon_1) & \cup & [\upsilon_1, \upsilon_2) & \cup & \dots & \cup & [\upsilon_{N-1}, 100\%], \\ & \Uparrow & & \Uparrow & & & & \Uparrow \\ & S_1 & & S_2 & & & & S_N \end{array} \tag{8}$$

where $S_1, S_2, ..., S_N$ represent the Markov chain states; $\upsilon_1, \upsilon_2, ..., \upsilon_{N-1}$ are constants ($0\% < \upsilon_1 < \upsilon_2 < ... < \upsilon_{N-1} < 100\%$).

At any given state ($S_i$), a subsystem has the theoretical opportunity to transition to any other state ($S_j$), as can be seen in the state evolution example from Figure 6 or the general state transition diagram from Figure 7.

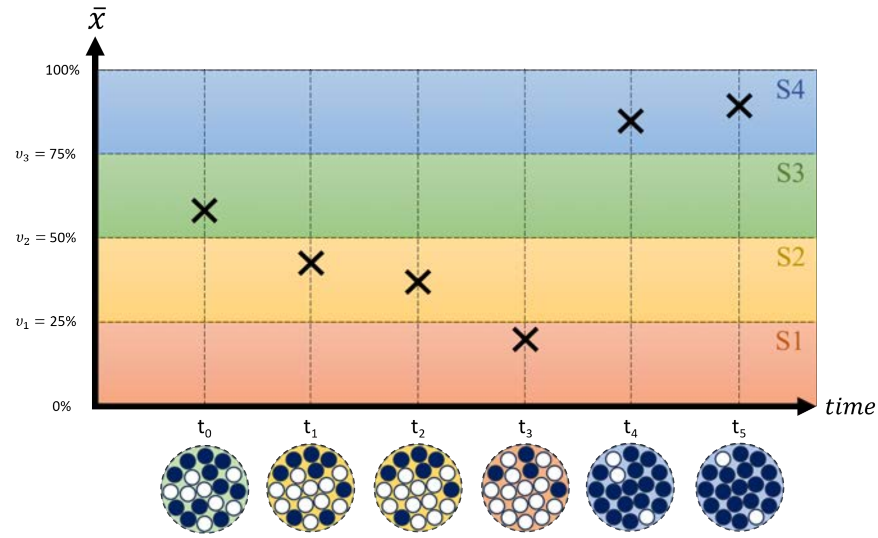


Figure 6: Evolution of subsystem state.

In this example, from Figure 6, the subsystem representation is featured by a circle which:

- has a dashed outline,
- has a color filling identical to the strip that is determined by the percentage values interval correspond to an individual state,
- over which, there is a set of smaller scirls,
  - some of them having dark filling.

For example, the left-most representation of the subsystem contains eleven dark and eight white small circles, surrounded by a green filled circle. This means that, in the context of this example, the subsystem contains nineteen bike stations. Above the subsystems' representations, there is a graph with a horizontal axis denoting time and a vertical axis corresponding to overall criticality ($\overline{x}$). All variables from Equation (8) and their relationship are reflected in the graph. Each measurement from the graph, symbolized by $\times$, means that the overall criticality $\overline{x}$ of the subsystem at time $t_i$ is of value $\overline{x}_i$, which translates to the affiliation to one of the states $S_1$, $S_2$, $S_3$, or $S_4$. Each value $\overline{x}_i$ is computed by dividing the number of small dark circles over the number of all small circles inside the bigger circle.

The state $S_i$ of the MC is defined based on the ratio of bike stations in critical load. The state changes when the ratio of bike stations in critical load goes outside of a determined interval. Therefore, the state can be characterized by the specific range $[\upsilon_{i-1}, \upsilon_i) \rightarrow S_i$ of the ratio of bike stations in critical load.

The timing associated with the state transition is expressed by a random variable $X$ with an exponential distribution function:

$$F_X(x) = P(X \leq x) = 1 - e^{-\lambda x}, \tag{9}$$

If $t_{ij}$ represents the time interval for which a subsystem remains in state $S_i$ until it reaches state $S_j$, expressed by a random variable of exponential distribution with mean value $\gamma_{ij}$, then the execution frequency of the transition will be $\lambda_{ij} = \frac{1}{\gamma_{ij}}$, which, from now on, will be called transition rate.

A general representation of the Markov chain which is associated to the subsystem is given in Figure 7.

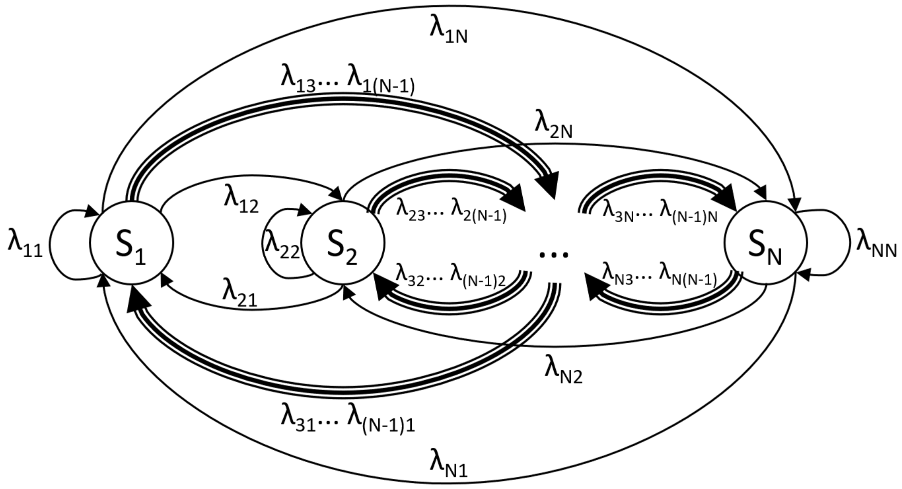


Figure 7: The general state transition diagram for a subsystem.

The Markov Chain was constructed with the following assumptions:

- Once a truck has left a bike station, it cannot return to the same station unless it has visited other bike stations in between;
- The bikes are not exchanged in between trucks that meet in their journeys;
- When observing real subsystems for the defined time interval of model identification, the transitions in between some states will never happen;
- The transition probability between states that correspond to neighboring ranges of critical load ratios is much higher compared to distant ranges.

Using the proposed Markov chain, the steady state can be determined for the subsystem, which matters for deciding the number of trucks for rebalancing purposes. The method used for determining the steady staty is described as follows.

The degree of subsystem demand (for rebalancing) is defined in the equation below:

$$D = \overline{x} \cdot m \tag{10}$$

The relationship between optimal fleet size and different subsystem demands can also be visualized in Figure 8.

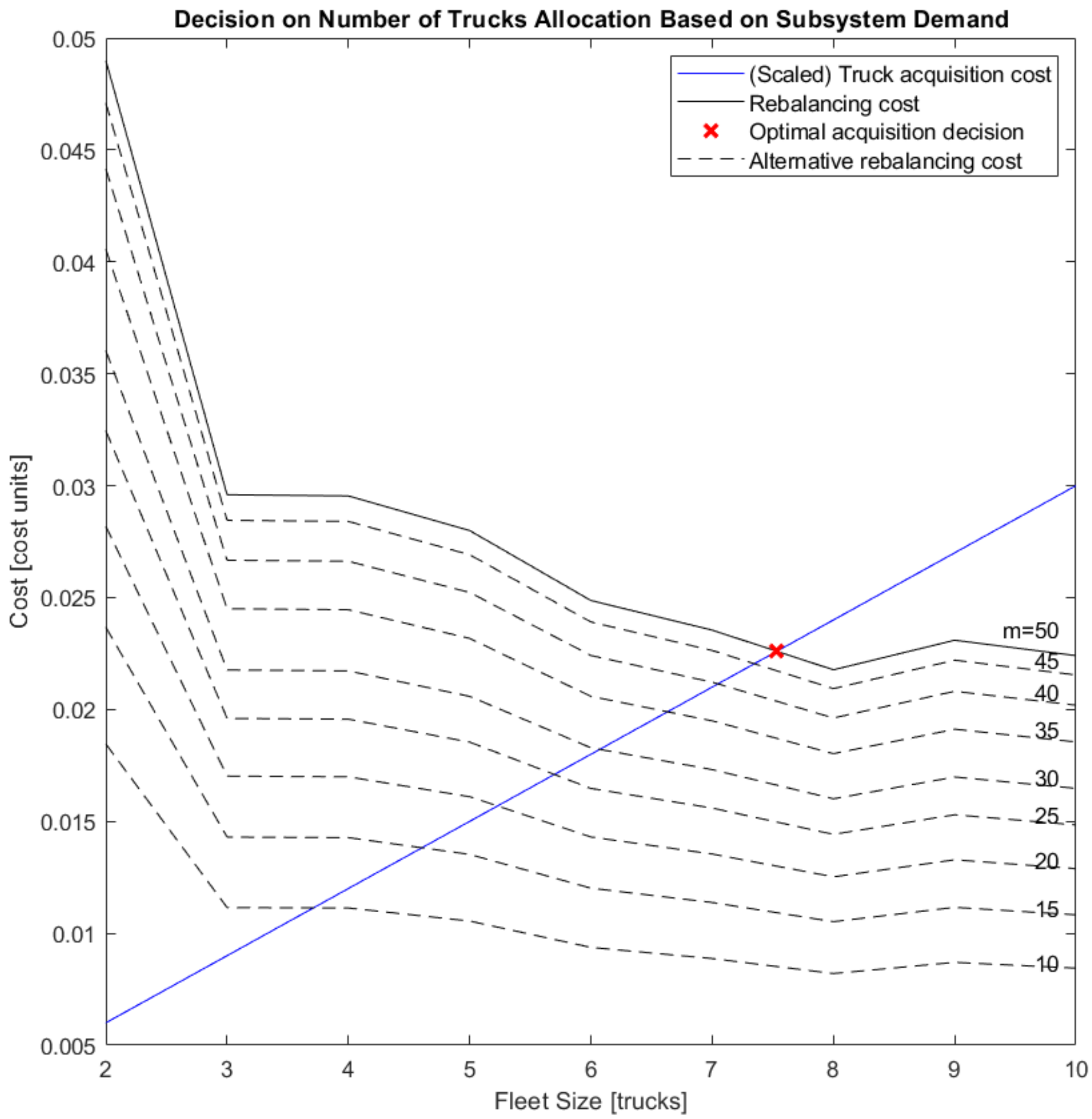


Figure 8: The graphical representation of the fleet acquisition and operational costs.

For exemplification purposes, inside Figure 8, based on the available results from [35], the actual rebalancing cost is drawn, in the form of the continuous black line for which the subsystem demand was maximum ($D = m = 50$—marked on top of line); dashed lines are used for other different demands ($D = 10, 15, 20, \ldots, 45$—marked on top of lines). The red point corresponds to the optimal number of trucks, as a trade-off between the initial fleet acquisition cost and the rebalancing cost; the blue line is the scaled truck acquisition cost, which grew linearly, based on the number of trucks involved, considering that each additional truck represented a fixed cost. The scaling of the truck acquisition cost is formulated below:

$$\mathrm{FK}(\mathrm{Tr}) = \alpha \cdot K_t \cdot \mathrm{Tr}, \tag{11}$$

where $\mathrm{FK} : \mathbb{N} \to \mathbb{R}^+$ is the scaled fleet cost; $\mathrm{Tr} \in \mathbb{N}$ is the number of trucks allocated for the subsystem; $\alpha \in \mathbb{R}^+$ is a constant scaling factor; $K_t \in \mathbb{R}^+$ is the fixed cost of the acquisition of one truck.

The scaling factor $\alpha$ is selected so that the optimal criteria from Equation (12) are met:

$$\hat{\mathrm{Tr}} \text{ este OPTIM} \iff \mathrm{FK}(\hat{\mathrm{Tr}}) \approx \mathrm{CK}(\hat{\mathrm{Tr}}), \tag{12}$$

where $\hat{\mathrm{Tr}}$ is the optimal number of trucks allocated to the subsystem; $\mathrm{FK} : \mathbb{N} \to \mathbb{R}^+$ is the scaled fleet cost; $\mathrm{CK} : \mathbb{N} \to \mathbb{R}^+$ is the rebalancing cost. Approximation ($\approx$) instead of equality ($=$) is used because it is impossible to allocate fractions of trucks, so the closest natural number is used instead (for example, in Figure 8, $\hat{\mathrm{Tr}} = 7$ or 8).

Equation (7) shows that the overall criticality $\overline{x}$ changes over time and, based on Equation (10), the subsystem demand $D$, therefore, also changes over time. The optimal number $\hat{\mathrm{Tr}}$ of trucks is also dependent on $D$. Because the subsystem to be modeled belongs to a real-world BSS, another constraint is that $\hat{\mathrm{Tr}}$ is ideally selected once in the beginning phase of the subsystem. In order to determine a time invariant demand $\hat{D}$ of the subsystem, the steady state probabilities for the subsystem are used. The steady state probabilities vector $P^*$ is defined in the equation below:

$$P^* = [P_1^* \quad P_2^* \quad \cdots \quad P_N^*], \tag{13}$$

where $N$ is the number of individual states of the subsystem; $P_i^*$ is the probability for the subsystem to be in the state $S_i$, and it can be determined by solving the following system of equations:

$$\begin{cases} P^* Q = 0 \\ \sum_{i=1}^{N} P_i^* = 1 \end{cases} \tag{14}$$

Considering the quantization Equation (8), which states that $S_i$ corresponds to an overall criticality as an average of the interval boundaries ($\frac{1}{2}(\upsilon_{i-1} + \upsilon_i)$), the steady demand $\hat{D}$ of the subsystem will be calculated as follows:

$$\hat{D} = P^* \Upsilon m, \tag{15}$$

where $\Upsilon$ is the associated average overall criticality vector to the subsystem states ($\Upsilon \mapsto \{S_1, S_2, \ldots, S_N\}$), defined as follows:

$$\Upsilon = \begin{bmatrix} \overline{\upsilon_{0,1}} \\ \overline{\upsilon_{1,2}} \\ \overline{\upsilon_{2,3}} \\ \vdots \\ \overline{\upsilon_{N-2,N-1}} \\ \overline{\upsilon_{N-1,N}} \end{bmatrix} = \frac{1}{2} \begin{bmatrix} \upsilon_1 \\ \upsilon_1 + \upsilon_2 \\ \upsilon_2 + \upsilon_3 \\ \vdots \\ \upsilon_{N-2} + \upsilon_{N-1} \\ \upsilon_{N-1} + 1 \end{bmatrix} \tag{16}$$

The model usage can be applied in other bike-sharing systems, because it has the capacity to describe the behavior of the system with a high enough resolution so that the rebalancing vehicles can comply with the imposed time constraints. Establishing of the optimal rebalancing fleet size ensures an increased operational efficiency. The obtained results represent an argument for extending the model usage, for describing the behavior of other systems.

## Support strategies based on modified genetic algorithms for resources relocation in bike-sharing systems

Even if the quality of services of bike sharing service systems was permanently improved, there are some issues which still need new and more efficient solutions. In order to find new adequate solutions, ensuring high performance, a new rebalancing strategy of the highly dynanmic stations is proposed [35]. The core elements of the method are a fuzzy logic-controlled genetic algorithm (FLCGA) [36] for bike station prioritization and an inference mechanism aiming to do the assignment between the stations and trucks.

Figure 9 is a schematic representation of the method in which: a static rules list specifying rebalancing constraints along with dynamic traffic data constitute the input of a deci-sion algorithm; the genetic algorithm is used to determine the order of the services. Truck assignment rules are then applied to obtain the truck scheduling list.

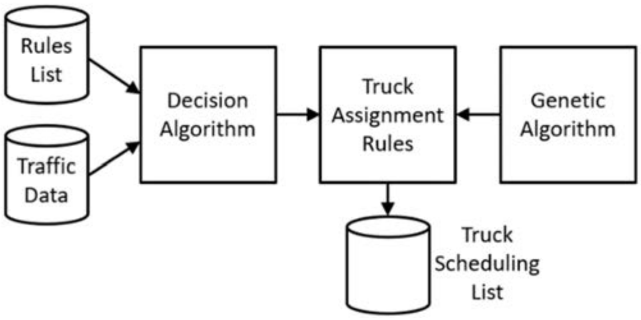


Figure 9: The schematic representation of the proposed method.

The block diagram of the FLGCA algorithm, proposed for the bike-sharing rebalancing problem, is given in the next figure:

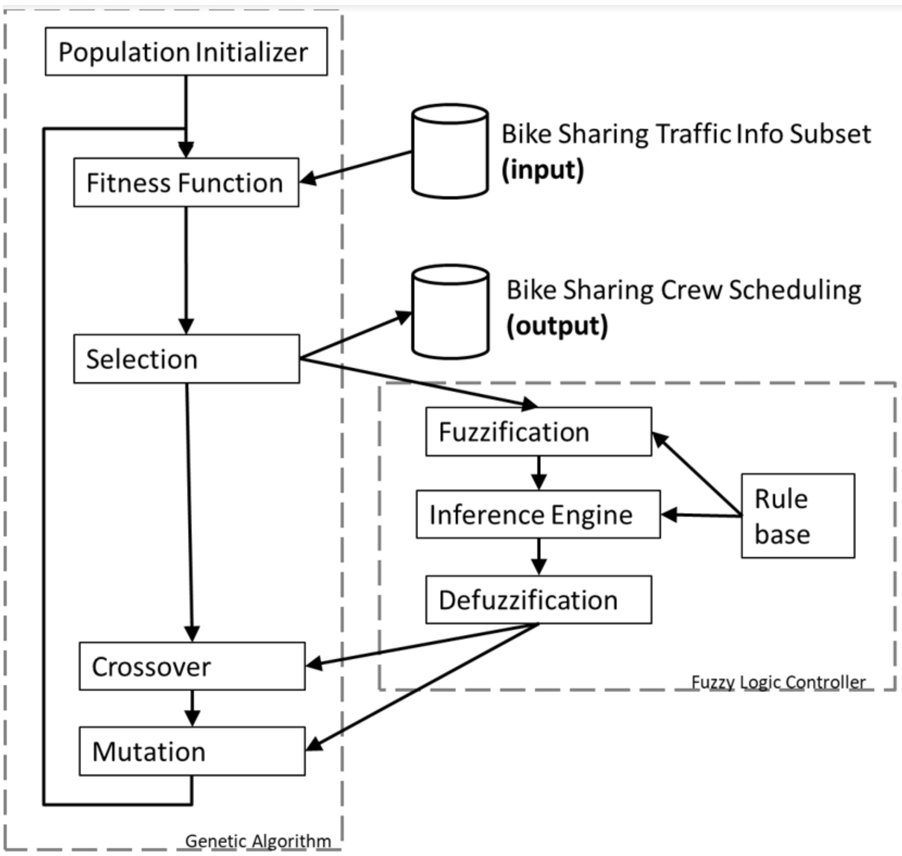


Figure 10: The Block Diagram of the FLCGA.

The diagram contains two main parts corresponding to the Genetic Algorithm (GA) and respectively to the Fuzzy-logic controller (FLC).

The controller dynamically adjusts the crossover and mutation probabilities for the next generation based on the statistics of the current candidate solutions.

The result of the FLCGA is a cost-effective order of serving bike stations that need rebalancing.

The steps of the decision algorithm for the truck assignment are:

**Algorithm 1** Decision algorithmul for association the lists of priorities to the trucks.

```
INPUTS: S, T
for station in S do
    truckAssigned ← FALSE
    truck ← getFirstElement(T)
    repeat
        journey ← getJourneyHistory(truck)
        if length(journey) = MINIMUM then
            pair(truck, station)
            if ableToFulfillDemand(truck) then
                route ← distance(truck,station)
            else
                route ← distance(truck, depot) + distance(depot, station)
            end if
            journey ← journey + route
            updateJourneyHistory(truck, journey)
            truckAssigned ← TRUE
        else
            truck ← getNextElement(T)
        end if
    until truckAssigned = TRUE
end for
OUTPUT: journey
```

The algorithm considers the special event when the truck misses the capacity to directly serve the next bike station. The additional visiting of the depot will lead to the increase of the total cost of the trip.

An example of routes assigned to trucks, resulting from running the application for a subsystem comprised of 5 trucks and 20 stations, is given in Table 1.

Table 1: Exemple of obtained routes.

| Truck ID | Route | Route length | Load [%] |
|---|---|---|---|
| 1 | 0→7→16→4→20→0 | 14.34 | 72 |
| 2 | 0→17→23→12→0→15→9→14→0→3→2→0 | 13.80 | 86 |
| 3 | 0→6→10→8→0 | 16.96 | 53 |
| 4 | 0→11→19→13→18→0 | 15.35 | 35 |
| 5 | 0→1→0→5→0 | 12.55 | 64 |

The performance indicators, with the values computed based on the routes obtained for the rebalancing trucks, reflect relevant aspects regarding the correct functioning of the system: reduction of the total time of operation in a critical state of the system; reducing the rate of lost customers; obtaining a uniform distribution of the loads assigned to the rebalancing trucks. The efficiency of the method was tested in various representative case studies using data recorded from the Citi Bike system.

# Comparative analysis of the results obtained by integrating reference algorithms into the relocation method

The bike relocation method in bike-sharing systems, presented previously, was tested on traffic data from the Citi Bike New York system and the results were compared to those obtained by replacing the FLCGA algorithm with each of the following algorithms: the standard genetic algorithm (SGA), Ant Colony Optimization (ACO) [37], Harris Hawks Optimization (HHO) [38], Tabu Search Algorithm (TSA) [39]. The applicability of all methods is demonstrated by performing numerical experiments using real historical traffic data on 1000 data sets, each data set corresponding to a group of stations (subsystem) and a scenario, each scenario being a unique combination of numbers of trucks and stations involved in the experiment.

A histogram was used for representing the results of each method. On the horizontal axis of the histogram the possible values of the objective function corresponding to the routing of the rebalancing trucks are represented, and on the vertical axis the number of subsystems for which the respective value was determined. Each distinct subsystem is identified based on the number and locations of the stations it contains, as well as the number of rebalancing trucks involved. Each subsystem is assigned a set of routes corresponding to the visits of the rebalancing trucks at the stations. The performance of the set of the determined routes is expressed by the objective function value. Each value of the objective function corresponds to an element that is added to the histogram. An example of a histogram is shown in Figure 11:

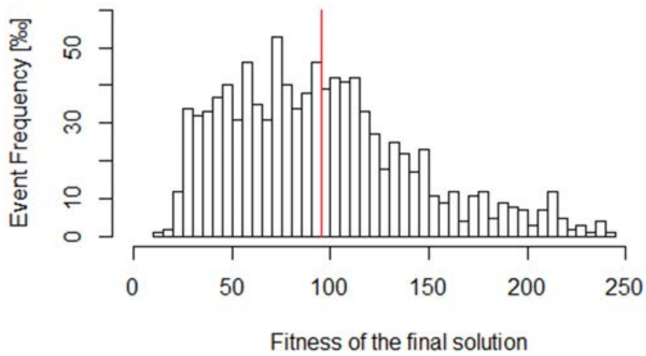


Figure 11: Histogram of the performances of the bike relocation methods.

Each subsystem was generated by employing a window search through the data of the geographical location (latitude, longitude) of bike stations. In this context, the win-dow search in this context represents a random positioning of a fixed-size window on the map of bike stations from the BSS. The window size has been chosen to capture constant numbers of stations. If a window covers more stations than desired, a random selection of residual stations to be removed from the sample is performed. The proce-dure is exemplified in Figure 12.

Figure 12: Example of window search performed on bike stations of Citi Bike, New York

Several scenarios, which included a variable number of trucks and bike stations, were defined in order to evaluate the performance and scalability of the method. The comparative analysis of the results was based on the following properties, determined by analyzing the histograms: the median value of the objective function, the standard deviation of the values, the value of the objective function corresponding to the 10th percentile of the histogram and the value of the objective function corresponding to the percentile of the 90th percentile.

An example of the results, which were obtained by integrating the proposed modified genetic algorithm and the ACO algorithm in the rebalancing method, is presented in Table 2:

Table 2: The results obtained after applying the FLCGA and ACO algorithms, respectively, within the relocation method.

| **Trucks nr.** | **Stations nr.** | **Median fitness [bikes/km]** | **Standard deviation [bikes/km]** | **Fitness (10th percentile) [bikes/km]** | **Fitness (90th percentile) [bikes/km]** | **Total distance traveled [km]** |
|---|---|---|---|---|---|---|
| FLCGA - DIFFERENT NUMBERS OF STATIONS | | | | | | |
| 2 | 10 | 9.52 | 2.33 | 4.40 | 16.89 | 39.72 |
| 2 | 15 | 11.90 | 3.01 | 4.74 | 20.49 | 50.94 |
| 2 | 20 | 13.85 | 4.28 | 5.16 | 23.63 | 60.64 |
| 2 | 25 | 15.52 | 5.82 | 5.43 | 28.07 | 69.83 |
| 2 | 30 | 17.347 | 6.77 | 5.21 | 30.12 | 77.53 |
| 2 | 35 | 18.15 | 8.41 | 5.84 | 33.02 | 87.28 |
| 2 | 40 | 19.03 | 7.22 | 6.19 | 34.14 | 95.01 |
| 2 | 45 | 19.78 | 8.95 | 6.71 | 32.45 | 101.36 |
| 2 | 50 | 20.41 | 8.36 | 6.55 | 37.05 | 105.43 |
| ACO - DIFFERENT NUMBERS OF STATIONS | | | | | | |
| 2 | 10 | 5.91 | 1.06 | 4.18 | 10.17 | 52.27 |
| 2 | 15 | 7.38 | 1.59 | 4.21 | 12.36 | 78.21 |
| 2 | 20 | 9.79 | 1.67 | 4.71 | 12.78 | 78.39 |
| 2 | 25 | 10.29 | 3.99 | 3.78 | 15.74 | 104.47 |
| 2 | 30 | 12.58 | 3.35 | 4.57 | 16.91 | 109.50 |
| 2 | 35 | 13.07 | 3.69 | 4.58 | 17.00 | 134.73 |
| 2 | 40 | 12.69 | 3.22 | 5.24 | 17.83 | 114.38 |
| 2 | 45 | 11.96 | 2.83 | 5.20 | 18.66 | 139.90 |
| 2 | 50 | 14.34 | 4.09 | 5.82 | 18.91 | 134.86 |
| FLCGA - DIFFERENT NUMBERS OF TRUCKS | | | | | | |
| 2 | 50 | 20.41 | 8.36 | 6.55 | 37.05 | 105.43 |
| 3 | 50 | 33.79 | 15.97 | 10.29 | 62.75 | 114.75 |
| 4 | 50 | 33.85 | 17.94 | 9.42 | 67.19 | 135.73 |
| 5 | 50 | 35.72 | 18.36 | 9.13 | 69.37 | 147.51 |
| 6 | 50 | 40.21 | 17.53 | 12.90 | 71.59 | 166.84 |
| 7 | 50 | 42.47 | 19.61 | 11.53 | 73.71 | 185.97 |
| 8 | 50 | 45.93 | 20.03 | 13.87 | 78.26 | 203.48 |
| 9 | 50 | 43.30 | 20.84 | 12.58 | 75.28 | 213.39 |
| 10 | 50 | 44.63 | 22.56 | 14.39 | 78.04 | 231.62 |
| ACO - - DIFFERENT NUMBERS OF TRUCKS | | | | | | |
| 2 | 50 | 14.34 | 4.09 | 5.82 | 18.91 | 134.86 |
| 3 | 50 | 25.30 | 14.86 | 8.26 | 49.98 | 144.22 |
| 4 | 50 | 28.74 | 16.11 | 7.93 | 51.52 | 163.18 |
| 5 | 50 | 31.30 | 15.52 | 7.49 | 63.74 | 177.76 |
| 6 | 50 | 34.34 | 13.70 | 11.19 | 53.98 | 213.99 |
| 7 | 50 | 38.08 | 18.38 | 8.96 | 58.22 | 251.56 |
| 8 | 50 | 35.21 | 20.71 | 13.03 | 65.58 | 253.67 |
| 9 | 50 | 39.80 | 19.75 | 10.10 | 63.77 | 251.59 |
| 10 | 50 | 38.68 | 18.29 | 12.90 | 66.23 | 292.42 |

The results of the comparative analysis led to the conclusion that from the point of view of the average performance of the rebalancing operation, the results of the proposed algorithm are superior to those of the ACO, TSA and standard genetic algorithms, and for small size subsystems they are superior to those of the HHO algorithm. The proposed modified genetic algorithm is also shown to have a shorter convergence time than the standard genetic algorithm. The results recommend the method to be used in other similar applications.

# Thesis' contributions

The main contributions are presented below, listed according to the order of the thesis chapters:

**Chapters 1, 2 and 3**:

1. Performing a literature review regarding the optimization of operations in logistics systems: manipulation, storage, fractioning/grouping, preparation of commercial assortment, transport and bike-sharing related operations. This contribution is aimed at achieving the first, second, third and fifth objective.

**Chapter 4**:

2. Elaborating of methods for selecting stations from the system, which are highly dynamic. This contribution is aimed at achieving the seventh objective.

3. Formulating the hypotheses and determining the characteristics considered in the description of the bike stations. This contribution is aimed at achieving the second objective.

4. Formulating the hypotheses and determining the characteristics considered in the description of the rebalancing agents. This contribution is aimed at achieving the second objective.

5. Formulating the hypotheses regarding the operational cases of the warehouses. This contribution is aimed at achieving the second objective.

6. Establishing the selection criteria of the stations which are highly dynamic, from the bike-sharing systems. This contribution is aimed at achieving the second objective.

7. Implementing the extraction and preprocessing functions of the generated data, taken from the system database of the Citi Bike New York system. This contribution is aimed at achieving the seventh objective.

**Chapter 5**:

8. Using the probabilistic Petri nets for modeling temporal variations of the number of bikes. This contribution is aimed at achieving the first, second and fourth objectives.

9. Implementing the algorithms for linearizing the behavior of bike stations reflected by the variation of the number of bikes over time. This contribution is aimed at achieving the seventh objective.

10. Implementing the models based on probabilistic Petri nets that describe the evolution of the load of the bike stations. This contribution is aimed at achieving the third and sixth objectives.

**Chapter 6**:

11. Defining the similarity of the behavior of stations from the bike-sharing system. This contribution is aimed at achieving the first and second objectives.

12. Using Markov chains to model the behavior of the groups of bike stations. This contribution is aimed at achieving the first, second and fourth objectives.

13. Elaborating of a method to identify and evaluate the behavior of the groups of stations, based on the selection of time intervals, corresponding to behavior templates. This contribution is aimed at achieving the first, second and fourth objectives.

14. Developing a method to determine the required number of trucks for the purpose of rebalancing stations where the total demand for bikes changes over time. This contribution is aimed at achieving the first, second and third objectives.

15. Designing and implementing the software application used to extract the features of subsystems, in order to model the behavior of the groups of stations. This contribution is aimed at achieving the second and seventh objectives.

16. Implementing the method for determining the similarity of the stations' behavior. This contribution is aimed at achieving the second and seventh objectives.

17. Designing and implementing the Markov chain-based modeling used to describe the behavior of the groups of stations. This contribution is aimed at achieving the third and fifth objectives.

**Chapter 7**:

18. Elaborating an inference mechanism for the purpose of assigning the lists containing the visiting order of the stations to the agents performing the rebalancing. This contribution is aimed at achieving the third objective.

19. Elaborating an algorithm in order to determine the routes of the rebalancing agents. This contribution is aimed at achieving the third objective.

20. Integrating the fuzzy logic controlled genetic algorithms in the bike relocation method. This contribution is aimed at achieving the third objective.

21. Defining, in a stepwise manner, of an objective function in order to determine the performance of the rebalancing operations. This contribution is aimed at achieving the third, fifth and sixth objectives.

22. Parametrizing of a fuzzy controller in order to obtain superior performance of the rebalancing operations. This contribution is aimed at achieving the third and sixth objectives.

23. Designing of the software application architecture implementing the lists containing the visiting order of the stations to the agents performing the rebalancing. This contribution is aimed at achieving the third objective.

**Chapter 8**:

24. Carrying out a comparative analysis of the performances obtained by integrating some representative routing algorithms within the bike relocation method. This contribution is aimed at achieving the fifth and sixth objectives.

25. Using search windows with randomly selected locations to generate the datasets used to validate the rebalancing methods. This contribution is aimed at achieving the fifth and sixth objectives.

**The following contribution applies to Chapters 6, 7 and 8**:

26. Implementing all methods of identification and validation of the models proposed in the thesis. This contribution is aimed at achieving the third, sixth and seventh objectives.

# Dissemination

The contributions of this thesis have been published in prestigious journals, in the fields of *Information Theory Applied in Scientific Computing* and *Intelligent Transportation System Technologies and Applications*, and in the volumes of international conferences in the field of *Automation, Quality and Testing, Robotics*. The list of the published articles is presented as follows:

**Papers published in Web of Science indexed journals**

- Papers published during the development of the thesis
  - **Florian H.**, Avram C., Radu D., Aștilean A. *Decision system based on Markov chains for sizing the rebalancing fleet of bike-sharing stations*, Applied Sciences, vol. 14, 2024
  - **Florian H.**, Avram C., Pop M., Radu D., Aștilean A. *Resources Relocation Support Strategy Based on a Modified Genetic Algorithm for Bike-Sharing Systems*, Mathematics, vol. 11, 2023
- Papers published before starting the thesis
  - **Florian H.**, Mocanu A., Vlasin C., Machado J., Carvalho V., Soares F., Aștilean A., Avram C. *Deaf people feeling music rhythm by using a sensing and actuating device*, Sensors and Actuators A: Physical, 2017

**Papers published in the volumes of international conferences**

- IEEE Proceedings
  - Tudoroiu R., Santa M., **Florian H.**, Zaheeruddin M., Radu S., Tudoroiu N. *Nonlinear Neural Control Strategies versus Conventional Control — Case Study and Performance Comparison*, 2024 IEEE International Conference on Automation, Quality and Testing, Robotics (AQTR), Cluj-Napoca, Romania, 2024
- Web of Science indexed conference volumes
  - **Florian H.**, Avram C., Pop M., Mocanu A., Radu D., Aștilean A. *Probabilistic Petri Nets Model for Assessing Temporal Variations of the Number of Bikes in Bike-Sharing Stations*, 2022 IEEE International Conference on Automation, Quality and Testing, Robotics (AQTR), Cluj-Napoca, Romania, 2022.
  - **Florian H.**, Pop M., Avram C., Aştilean A. *Similarity measure for station clustering in bike sharing systems*, 2020 IEEE International Conference on Automation, Quality and Testing, Robotics (AQTR), Cluj-Napoca, Romania, 2020.

# Bibliography


[1] X. Guoxun, L. Yanfeng, J. Daxiang, and L. Jun, "A user-based method for the static bike repositioning problem," *Future Gener. Comput. Syst*, vol. 40, pp. 426–436, 2020. [Online]. Available: https://doi.org/10.12011/1000-6788-2018-0884-11

[2] R. E. Sibai, K. Challita, J. B. Abdo, and J. Demerjian, "A new user-based incentive strategy for improving bike sharing systems' performance," *Sustainability*, vol. 13, no. 5, p. 2780, Mar. 2021. [Online]. Available: https://doi.org/10.3390/su13052780

[3] B. Legros, "Dynamic repositioning strategy in a bike-sharing system; how to prioritize and how to rebalance a bike station," *European Journal of Operational Research*, vol. 272, no. 2, pp. 740–753, Jan. 2019. [Online]. Available: https://doi.org/10.1016/j.ejor.2018.06.051

[4] Y.-C. Lin, "A demand-centric repositioning strategy for bike-sharing systems," *Sensors*, vol. 22, no. 15, p. 5580, Jul. 2022. [Online]. Available: https://doi.org/10.3390/s22155580

[5] S. Reiss and K. Bogenberger, "A relocation strategy for munich's bike sharing system: Combining an operator-based and a user-based scheme," *Transportation Research Procedia*, vol. 22, pp. 105–114, 2017. [Online]. Available: https://doi.org/10.1016/j.trpro.2017.03.016

[6] H. Florian, M. Pop, C. Avram, and A. Aştilean, "Similarity measure for station clustering in bike sharing systems," in *2020 IEEE International Conference on Automation, Quality and Testing, Robotics (AQTR)*, 2020, pp. 1–5.

[7] S. Sohrabi, R. Paleti, L. Balan, and M. Cetin, "Real-time prediction of public bike sharing system demand using generalized extreme value count model," *Transportation Research Part A: Policy and Practice*, vol. 133, pp. 325–336, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0965856417316105

[8] M. Cipriano, L. Colomba, and P. Garza, "A data-driven based dynamic rebalancing methodology for bike sharing systems," *Applied Sciences*, vol. 11, no. 15, p. 6967, Jul. 2021. [Online]. Available: https://doi.org/10.3390/app11156967

[9] M. Kruszyna, "NOAH as an innovative tool for modeling the use of suburban railways," *Sustainability*, vol. 15, no. 1, p. 193, Dec. 2022. [Online]. Available: https://doi.org/10.3390/su15010193

[10] F. Lu, W. Feng, M. Gao, H. Bi, and S. Wang, "The fourth-party logistics routing problem using ant colony system-improved grey wolf optimization," *Journal of Advanced Transportation*, vol. 2020, pp. 1–15, Oct. 2020. [Online]. Available: https://doi.org/10.1155/2020/8831746

[11] F. Chiariotti, C. Pielli, A. Zanella, and M. Zorzi, "A dynamic approach to rebalancing bike-sharing systems," *Sensors*, vol. 18, no. 2, 2018. [Online]. Available: https://www.mdpi.com/1424-8220/18/2/512

[12] E. Eren and V. E. Uz, “A review on bike-sharing: The factors affecting bike-sharing demand,” *Sustainable Cities and Society*, vol. 54, p. 101882, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2210670719312387

[13] C. S. João Filipe Teixeira and F. M. e Sá, “Empirical evidence on the impacts of bikesharing: a literature review,” *Transport Reviews*, vol. 41, no. 3, pp. 329–351, 2021. [Online]. Available: https://doi.org/10.1080/01441647.2020.1841328

[14] C. Sun and J. Lu, “The reliability model for bike-sharing dispatch based on hotspot detection and hypothesis test: A case study in beijing,” *Discrete Dynamics in Nature and Society*, vol. 2022, p. 1–13, Jan. 2022. [Online]. Available: http://dx.doi.org/10.1155/2022/7049765

[15] ——, “Modeling spatial riding characteristics of bike-sharing users using hotspot areas-based association rule mining,” *Journal of Advanced Transportation*, vol. 2022, p. 1–14, Apr. 2022. [Online]. Available: http://dx.doi.org/10.1155/2022/5705080

[16] C. M. Vallez, M. Castro, and D. Contreras, “Challenges and opportunities in dock-based bike-sharing rebalancing: A systematic review,” *Sustainability*, vol. 13, no. 4, 2021. [Online]. Available: https://www.mdpi.com/2071-1050/13/4/1829

[17] J. Zhou, Y. Guo, J. Sun, E. Yu, and R. Wang, “Review of bike-sharing system studies using bibliometrics method,” *Journal of Traffic and Transportation Engineering (English Edition)*, vol. 9, no. 4, pp. 608–630, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2095756422000174

[18] J. Zhao, W. Fan, and X. Zhai, “Identification of land-use characteristics using bicycle sharing data: A deep learning approach,” *Journal of Transport Geography*, vol. 82, p. 102562, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0966692318309232

[19] M. E. Mohammed H. Almannaa and H. A. Rakha, “Dynamic linear models to predict bike availability in a bike sharing system,” *International Journal of Sustainable Transportation*, vol. 14, no. 3, pp. 232–242, 2020. [Online]. Available: https://doi.org/10.1080/15568318.2019.1611976

[20] V. Mahmoodian, Y. Zhang, and H. Charkhgard, “Hybrid rebalancing with dynamic hubbing for free-floating bike sharing systems,” *International Journal of Transportation Science and Technology*, vol. 11, no. 3, pp. 636–652, Sep. 2022. [Online]. Available: https://doi.org/10.1016/j.ijtst.2021.08.002

[21] Y. Jin, C. Ruiz, and H. Liao, “A simulation framework for optimizing bike rebalancing and maintenance in large-scale bike-sharing systems,” *Simulation Modelling Practice and Theory*, vol. 115, p. 102422, Feb. 2022. [Online]. Available: https://doi.org/10.1016/j.simpat.2021.102422

[22] Y. Fan, G. Wang, X. Lu, and G. Wang, “Distributed forecasting and ant colony optimization for the bike-sharing rebalancing problem with unserved demands,” *PLOS ONE*, vol. 14, no. 12, p. e0226204, Dec. 2019. [Online]. Available: https://doi.org/10.1371/journal.pone.0226204

[23] Q. Tang, Z. Fu, D. Zhang, M. Qiu, and M. Li, “An improved iterated local search algorithm for the static partial repositioning problem in bike-sharing system,” *Journal of Advanced Transportation*, vol. 2020, pp. 1–15, Nov. 2020. [Online]. Available: https://doi.org/10.1155/2020/3040567

[24] J. Oyola, H. Arntzen, and D. L. Woodruff, "The stochastic vehicle routing problem, a literature review, part i: models," *EURO Journal on Transportation and Logistics*, vol. 7, no. 3, pp. 193–221, Sep. 2018. [Online]. Available: https://doi.org/10.1007/s13676-016-0100-5

[25] ——, "The stochastic vehicle routing problem, a literature review, part II: solution methods," *EURO Journal on Transportation and Logistics*, vol. 6, no. 4, pp. 349–388, Dec. 2017. [Online]. Available: https://doi.org/10.1007/s13676-016-0099-7

[26] Y. Zhu, X. Ye, J. Chen, X. Yan, and T. Wang, "Impact of cruising for parking on travel time of traffic flow," *Sustainability*, vol. 12, no. 8, p. 3079, Apr. 2020. [Online]. Available: http://dx.doi.org/10.3390/su12083079

[27] C. Bıyık, A. Abareshi, A. Paz, R. A. Ruiz, R. Battarra, C. D. Rogers, and C. Lizarraga, "Smart mobility adoption: A review of the literature," *Journal of Open Innovation: Technology, Market, and Complexity*, vol. 7, no. 2, p. 146, Jun. 2021. [Online]. Available: http://dx.doi.org/10.3390/joitmc7020146

[28] J. Sonnleitner, M. Friedrich, and E. Richter, "Impacts of highly automated vehicles on travel demand: macroscopic modeling methods and some results," *Transportation*, vol. 49, no. 3, p. 927–950, Jun. 2021. [Online]. Available: http://dx.doi.org/10.1007/s11116-021-10199-z

[29] J. Pucher and R. Buehler, "Cycling towards a more sustainable transport future," *Transport Reviews*, vol. 37, no. 6, pp. 689–694, Jun. 2017. [Online]. Available: https://doi.org/10.1080/01441647.2017.1340234

[30] S. Shaheen, "Mobility and the sharing economy: Potential to overcome first- and last-mile public transit connections," *Transport Reviews*, 2016. [Online]. Available: https://escholarship.org/uc/item/8042k3d7

[31] F. Kon, É. C. Ferreira, H. A. de Souza, F. Duarte, P. Santi, and C. Ratti, "Abstracting mobility flows from bike-sharing systems," *Public Transport*, vol. 14, no. 3, pp. 545–581, Mar. 2021. [Online]. Available: https://doi.org/10.1007/s12469-020-00259-5

[32] The Geography of Transport Systems. [Online]. Available: https://transportgeography.org/contents/chapter1/what-is-transport-geography/challenges-transport-systems/

[33] M. Albanese, R. Chellappa, V. Moscato, A. Picariello, V. Subrahmanian, P. Turaga, and O. Udrea, "A constrained probabilistic petri net framework for human activity detection in video," *IEEE Transactions on Multimedia*, vol. 10, no. 8, pp. 1429–1443, 2008.

[34] H. Florian, C. Avram, M. Pop, A. Mocanu, D. Radu, and A. Astilean, "Probabilistic petri nets model for assessing temporal variations of the number of bikes in bike-sharing stations," in *IEEE International Conference on Automation, Quality and Testing, Robotics, AQTR 2022, Cluj-Napoca, Romania, May 19-21, 2022*. IEEE, 2022, pp. 1–5. [Online]. Available: https://doi.org/10.1109/AQTR55203.2022.9802065

[35] H. Florian, C. Avram, M. Pop, D. Radu, and A. Aștilean, "Resources relocation support strategy based on a modified genetic algorithm for bike-sharing systems," *Mathematics*, vol. 11, no. 8, 2023. [Online]. Available: https://www.mdpi.com/2227-7390/11/8/1816

[36] E. Khmeleva, A. A. Hopgood, L. Tipi, and M. Shahidan, "Fuzzy-logic controlled genetic algorithm for the rail-freight crew-scheduling problem," *KI - Künstliche Intelligenz*, vol. 32, no. 1, pp. 61–75, Oct. 2017. [Online]. Available: https://doi.org/10.1007/s13218-017-0516-6

[37] M. Dorigo and C. Blum, "Ant colony optimization theory: A survey," *Theoretical Computer Science*, vol. 344, no. 2-3, pp. 243–278, Nov. 2005. [Online]. Available: https://doi.org/10.1016/j.tcs.2005.05.020

[38] A. A. Heidari, S. Mirjalili, H. Faris, I. Aljarah, M. Mafarja, and H. Chen, "Harris hawks optimization: Algorithm and applications," *Future Generation Computer Systems*, vol. 97, pp. 849–872, Aug. 2019. [Online]. Available: https://doi.org/10.1016/j.future.2019.02.028

[39] F. Glover, "Tabu search—part i," *ORSA Journal on Computing*, vol. 1, no. 3, pp. 190–206, Aug. 1989. [Online]. Available: https://doi.org/10.1287/ijoc.1.3.190