MATHEURISTIC FOR MULTI-PERIOD HOME HEALTHCARE ROUTING AND SCHEDULING PROBLEM: A REAL LIFE CASE STUDY

A Thesis

by

Yağmur Selenay Selçuk

Submitted to the
Graduate School of Sciences and Engineering
in Partial Fulfillment of the Requirements for
the Degree of

Master of Science

in
Industrial Engineering

Özyeğin University
July 2023

MATHEURISTIC FOR MULTI-PERIOD HOME HEALTHCARE ROUTING AND SCHEDULING PROBLEM: A REAL LIFE CASE STUDY

Approved by:

Asst. Prof. Elvin Çoban Göktürk, Advisor
Dept. of Industrial Engineering
Özyeğin University

Prof. Burcu Balçık Koyuncu
Dept. of Industrial Engineering
Özyeğin University

Asst. Prof. Umman Mahir Yıldırım
Dept. of Industrial Engineering
Bilgi University

Date Approved: July 31, 2023

To my family...

ABSTRACT

The aging population's exponential growth has exerted considerable pressure on healthcare systems, necessitating the provision of enhanced healthcare services tailored to meet the unique needs of older adults, individuals with disabilities, and chronic patients. As a result, healthcare providers aim to offer varying healthcare services to patients in their homes, with the objective of improving the quality of care and optimizing the management of health systems. This thesis studies a *home healthcare routing and scheduling problem* (HHCRSP) over a multi-period planning horizon, considering caregivers' lunch breaks, prior service type, and patients' preferences. In this HHCRSP, some patients needing blood draw, and they have to be visited before noon, guaranteeing the corresponding caregiver's return to the hospital's lab before noon. Additionally, patients have preferred time windows for each day, corresponding to times due to reasons such as the need for someone to support them with the patients. In the thesis, the objective function is minimizing the total routing costs of caregivers' vehicles and the penalty costs incurred when patients cannot receive services within their preferred time windows. Our study is motivated by a real-life hospital that provides *home healthcare service* (HHCS). We develop a *mixed-integer linear programming* (MILP) model and propose a *simulated annealing-based matheuristic algorithm* (SAMA) that decomposes the original problem into two phases. Furthermore, we conduct a comparative analysis of the *k-nearest-neighbour* (KNN) algorithm, utilizing various k values to predict service times. The results of our study demonstrate significant improvements, up to 98.64% in cost-effectiveness achieved by the MILP in small-sized instances, and 98.58% by the SAMA in medium and large-sized instances, compared to the existing system in the motivational hospital. The numerical analysis provides insights for healthcare providers and policymakers in their efforts to optimize HHCS.

ÖZETÇE

Yaşlanan nüfusun hızla büyümesi; yaşlıların, engelli bireylerin ve kronik hastaların ihtiyaçlarını karşılamaya yönelik geliştirilmiş sağlık hizmetlerinin sağlanmasını zorunlu kılarak sağlık sistemlerine önemli bir baskı uygulamıştır. Sonuç olarak, sağlık hizmeti sağlayıcıları, hasta bakımında çeşitli sağlık hizmetleri sunmayı, bakım kalitesini artırmayı ve sağlık sistemlerinin yönetimini optimize etmeyi amaçlamaktadır. Bu tez, sağlık hizmeti sağlayıcılarının öğle aralarını dikkate alan çok dönemli bir planlama sürecindeki evde sağlık hizmeti rotalama ve çizelgeleme problemi üzerinde çalışmaktadır. Bu evde sağlık hizmeti rotalama ve çizelgeleme probleminde, bazı hastalardan kan alınması gerekmektedir ve bu hastalar öğleden önce ziyaret edilmeli ve ilgili hizmet sağlayıcıların öğleden önce hastanedeki laboratuvara dönmesi gerekmektedir. Ayrıca, hastaların zaman pencereleri, onlara hizmet sırasında destek olması için refakatçi ihtiyacı gibi nedenlerle her gün için tercih edilen zaman aralıklarını ifade eder. Tezde, amaç fonksiyonu, hizmet sağlayıcıların araçlarının toplam rotalama maliyetlerini ve hastaların belirlenmiş zaman pencerelerinde hizmet alamadıklarında ortaya çıkan ceza maliyetlerini minimize etmektir. Çalışmamız, evde sağlık hizmeti sunan gerçek bir hastaneden esinlenilmektedir. Problemi çözmek için, karışık tamsayılı lineer programlama modeli geliştirilir ve orijinal problemi iki aşamaya ayıran benzetilmiş tavlama tabanlı matematiksel sezgisel algoritma önerilir. Ayrıca, hizmet sürelerini tahmin etmek için çeşitli k değerlerini kullanan k-en yakın komşu algoritmasının karşılaştırmalı analizi yapılır. Çalışmamızın sonuçları, mevcut sisteme kıyasla, küçük boyutlu örneklerde karışık tamsayılı lineer programlama modeli tarafından %98.64'e ve orta ve büyük boyutlu örneklerde benzetilmiş tavlama tabanlı matematiksel sezgisel algoritma tarafından %98.58'e varan maliyet etkinliğinde önemli iyileştirmeler sağlandığını göstermektedir.

Sayısal analizler, sağlık hizmeti sağlayıcıları ve politika yapıcıları için evde sağlık hizmetini optimize etme çabalarında değerli içgörüler sağlar.

ACKNOWLEDGEMENTS

My thesis-writing experience was a lengthy voyage that encompassed both personal and academic growth. I was fortunate to have the unwavering support and guidance of numerous companions and mentors who accompanied me on this challenging path. Their invaluable fellowship was instrumental in my ability to bring this endeavor to a successful conclusion.

I am deeply grateful to express my heartfelt appreciation to my thesis advisor, Elvin Çoban Göktürk, for her unwavering support and invaluable guidance throughout this journey. Her profound academic expertise and extensive knowledge not only propelled the advancement of my research in both methodology and theory but also instilled in me a sense of confidence and belief in my own abilities and the significance of my research. As an advisor, mentor, and researcher, she has exemplified excellence, serving as an inspiration to me and leaving an indelible mark on my academic development.

I extend my deepest gratitude to all my professors, with special mention to Professor Burcu Balçık Koyuncu, Professor Okan Örsan Özener, and Professor Ali Ekici, for their unwavering and limitless contributions throughout this transformative journey. Their tireless efforts, guidance, and mentorship have played an instrumental role in shaping my academic growth and development. Their vast knowledge and expertise have enriched my understanding and enabled me to navigate the complexities of my research with greater depth and clarity.

I would also like to thank Zeynep Şişman, Nurseli Karaaytaç, Aybüke Ekşi, Ecem Yücesoy, and Berk Karasu who have accompanied me on this journey for two years. Without their friendship, everything would be much more boring and unbearable

during this process.

My family means everything to me, and they have always supported and encouraged me with their unconditional love and faith in me. I would like to thank my lovely mother, Rüveyda Selçuk, and my dear father, Mehmet Sadık Selçuk, for everything they have done for me.

I would like to extend my heartfelt gratitude to my beautiful sister, Irmak Eren, and my dear brother, Batuhan Eren, who have provided me with a remarkable sense of brotherhood despite not being my biological siblings. Throughout this process, they have consistently offered their unwavering support and served as attentive listeners during conferences, standing by me in the front row. Their presence and encouragement have been invaluable to me, and I am deeply appreciative of the unique bond we share.

Lastly, I would like to thank those who take to the streets peacefully for a free, equal, and just world. They are the source of inspiration for this thesis.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Advances in science and technology lead to an increase in life expectancy, also resulting in an increase in the density of the elderly population worldwide [1]. As birth rates decrease and the proportion of older individuals in the population increases, along with an increase in chronic health conditions, demand for healthcare services keeps growing [2]. The World Health Organization (2019) mentions that the global population aged 60 and above would witness a staggering growth of 56%, surging from 962 million in 2017 to 1.4 billion by 2030, and is anticipated to exceed 2.1 billion by 2050. By 2060, the proportion of individuals in Europe aged 60 years and above is projected to reach 54% [3]. To fulfill the increasing demand for healthcare services, home healthcare service (HHCS), during which patients are treated at their homes, becomes widespread. In the US, medicare spending accounted for approximately 41% of home health expenditures in 2009, and this percentage is expected to increase even more [4]. Since HHCS is an important channel to provide healthcare as HHCS increases the accessibility of health service that is accepted as the most basic right of people [1]. In essence, *home healthcare* (HHC) teams aim to provide comprehensive health services, including examinations, analysis, treatments, medical care, and wound care, to socially disadvantaged, sick, needy, and elderly individuals in the comfort of their homes [5, 6].

HHCS can significantly mitigate the issue of overcrowding in hospitals, which remains a critical concern, particularly during the ongoing COVID-19 pandemic [7]. By delivering healthcare services at home, HHCS reduces the need for hospital admissions, thereby minimizing the risk of infection transmission within healthcare facilities. This is especially vital during infectious disease outbreaks like COVID-19, where reducing the spread of the virus is of utmost importance. Moreover, HHCS not only safeguards patients from potential exposure to contagious diseases in hospital settings but also plays a crucial role in freeing up valuable hospital resources for more critical cases. By redirecting non-emergency cases to home-based care, HHCS contributes to the efficient allocation of limited hospital beds, medical equipment, and healthcare personnel. The combined impact of HHCS in reducing the burden on hospitals and providing care in a safe and personalized environment demonstrates its significance in addressing the evolving healthcare needs of the population. By facilitating access to healthcare services and optimizing resource utilization, HHCS emerges as a valuable approach to meet the demands imposed by aging populations and public health emergencies, ultimately enhancing the overall quality and effectiveness of healthcare delivery.

However, the successful implementation of HHCS relies on overcoming operational challenges, particularly in solving the scheduling and routing problems faced by caregivers. They must navigate the task of visiting each HHC patient within their time windows to prevent treatment delays that may lead to health deterioration [8]. Additionally, efficient scheduling and routing are critical within HHCS to optimize travel distances and minimize travel time between patients, enabling effective allocation of resources, such as medical equipment and staff. Through strategic route planning and patient visit sequencing, caregivers can minimize travel distances, ensuring the timely and high-quality delivery of care.

The thesis aims to solve a *home healthcare routing and scheduling problem* (HHCRSP) that is faced by a real-life hospital providing HHCS. The number of caregivers is limited, and these caregivers have to visit a set of patients with time windows defined for patients' availabilities with varying preferences and priorities over a multi-period planning horizon. Patients have preferred daily time intervals, and if caregivers visit outside this time interval, a penalty is realized. Moreover, patients needing blood draw, which demands the corresponding caregivers' vehicle to return to the hospital to deliver the blood tubes to the laboratory, preventing blood perishability and receiving the lab results on the same day to start treatment as soon as possible. The objective is to minimize the total routing costs of caregivers' vehicles and total penalty costs caused by early and late service time of patients while considering the delivery of blood tubes to the hospital before noon, caregivers' lunch breaks, and patients' preferences. We develop a *mixed integer linear programming model* (MILP) and propose a *simulated annealing-based matheuristic algorithm* (SAMA) to solve large instances in a reasonable computation time. Additionally, in the context of data generated from real-life HHCS, *k-nearest-neighbour* (KNN) algorithm is employed to predict the treatment time of patients. Our computational analysis reveals that the MILP model proves its effectiveness for small-sized instances, whereas the SAMA outperforms in medium and large-sized instances, resulting in significant cost improvements compared to the current system. Specifically, the key contributions of our research are as follows.

1. This thesis tackles a novel HHCRSP that considers diverse prior service types, including blood draw, caregivers' lunch breaks, and patients' preferences. Moreover, the traditional HHCRSP is extended by considering that patients are not only assigned to caregivers but also to daily periods. This extension adds an additional level of complexity to the problem as time windows are not defined as once, but they are defined as preferred time windows, requiring careful consideration and optimization of caregiver assignments within specified time intervals.

2. This thesis primarily examines two distinct cases: the main case, *Vehicle Dispatch for Lab Assignment without Meeting Point*, and a policy derived from it

which is called *Vehicle Dispatch for Lab Assignment with Meeting Point*, is also investigated.

3. To minimize both the overall routing costs and penalty costs, an MILP model is proposed. To address the computational challenges associated with medium and large-sized instances, we develop the SAMA as an effective solution method. By employing the SAMA, we can efficiently compute near-optimal solutions for HHCRSP instances of significant scale, further enhancing the optimization process and facilitating decision-making in the HHC context.

4. We employ the KNN algorithm with a range of k values to predict the treatment time of patients. By utilizing this algorithm, the thesis enhances the resource allocation process as more accurate treatment durations are computed based on the motivating hospital's past data by the KNN algorithm.

5. Numerical analysis is performed to assess the efficiency of the developed MILP and proposed SAMA in terms of computation time, objective function values, and the number of patients visited outside their preferred time windows. The analysis includes testing instances with varying coefficients of the objective function to observe the impact on the number of patients visiting early or late. This helps in understanding how changes in the number of patients requiring blood draw throughout the day affect the results. Additionally, the study examines the percentages of breakdowns in the objective function and their effects. The obtained results are compared to those of the current system to gain managerial insights.

The remaining chapters of this paper are organized as follows. In Chapter 2, we review the relevant papers from the literature in detail. In Chapter 3, the HHCRSP is introduced. In Chapter 4, our solution approach is explained. Then, the prediction of treatment durations is shown in Chapter 5. We summarize the computational results of the algorithms in Chapter 6. Finally, the conclusion is presented in Chapter 7.

CHAPTER II

LITERATURE REVIEW

This literature review provides a comprehensive analysis of the scheduling and routing aspects of the HHCRSP, taking into account their distinct objectives and constraints. The scheduling component aims to optimize the allocation of healthcare providers' time, considering factors such as patients' needs, caregivers' availability, and preferences. In the scheduling literature, different objectives have been explored, including minimizing overall costs, maximizing patient satisfaction, and balancing workloads. On the other hand, the routing aspect focuses on determining the most efficient routes for healthcare providers to visit patients' homes. The routing problem involves optimizing travel distances while considering time windows and accommodating various constraints such as caregivers' working hours and lunch breaks. Objectives in routing studies range from minimizing travel time and cost to ensuring timely and reliable service delivery. In the context of HHC literature, this problem combines scheduling and routing, integrating both aspects to address the specific challenges of HHCS.

Despite the extensive research conducted in the field, the significance of the HHCRSP has continued to increase in recent years. This is primarily attributed to the global rise in the aging population and the subsequent surge in chronic diseases, resulting in a growing demand for HHCS worldwide [9]. The recognition of the HHCRSP's importance stems from its potential to enhance patient outcomes, reduce healthcare costs, and improve patient satisfaction. By effectively scheduling and routing healthcare providers, HHCS can ensure that patients receive timely and appropriate care in the comfort of their own homes. This not only improves patient well-being and quality of life but also reduces the burden on hospital resources, leading to cost savings for healthcare systems [10].

The HHCRSP is an enhanced version of the *vehicle routing problem with time window* (VRPTW), incorporating several unique side constraints that are specific to the HHC context [11]. The HHCRSP is presented using varied terminologies in the existing literature such as, *home healthcare problem* (HHCP) [12], *home care worker scheduling and routing problem* (HCWSRP) [13], and *home healthcare routing and scheduling problem* (HHCRSP) [14] and in this thesis, home healthcare routing and scheduling term is used.

To the best of our knowledge, [15] and [16] propose the concept of the HHCRSP for the first time in the literature. As a result of the problem's inherent complexity, a variety of constraints and objective functions are formulated. The objective function has been formulated in collaboration with various HHCSs to optimize for a solution that achieves high quality, minimizes operational costs, and enhances satisfaction levels for both patients and caregivers. Objective functions commonly used in the literature are minimization of the traveling costs, time or distance [17], minimization of the unscheduled patient visits [18], a fair distribution of the workload among the caregivers [19], and maximizing the number of clients visited [20]. However, many real-life restrictions must be taken into account when trying to achieve these objectives, such as time window, maximum working hours, preferences, skill requirement, and lunch break [21, 22].

The time window constraints play a critical role in the HHCRSPs since it affects the overall efficiency and effectiveness of the caregivers' schedule. Considering the time window for each patient, the caregivers can optimize their route to minimize travel time and ensure that they are able to visit all patients within their specified time periods. Within the HHCRSP literature, two distinct types of time windows are commonly utilized, namely, *hard time windows* and *soft time windows* [21]. *Hard* *time windows* represent a strict constraint that must be adhered to without any flexibility. It means that caregivers must arrive at the patient's location and complete the service within the preferred time window [23]. For example, [24] study an HHCRSP and propose an ant colony system with cluster algorithm. They consider hard time windows for the patients, so the service of the patients must be between the defined time intervals. Similarly, [19] propose a set partitioning heuristic to solve an HHCRSP considering hard time windows. Their objective function corresponds to a weighted sum of costs associated with the routes, the weekly caregivers' overtime and idle time, and the unscheduled visits of patients.

On the other hand, *soft time windows* are a more flexible constraint that allows some degree of deviation [25]. Researchers such as [26] study an HHCRSP considering soft time windows. They consider that a penalty of one is incurred when the service starts outside the tight soft time window but within the loose time window. If the service starts outside the loose time window, a penalty of two is incurred. They aim to minimize the total costs and client inconvenience. To address the problem, they develop a mixed integer programming (MIP) and propose a multi-directional local search framework and large neighborhood search. Upon examining the existing comprehensive literature review, it becomes apparent that there has been a paucity of research conducted on the subject of soft time windows [21, 27, 22]. Thus, in this thesis, soft time windows are considered, as they are deemed more congruous with practical real-life settings and have received comparatively less scholarly attention.

The preference constraint is important for the HHCRSP since it allows for the consideration of the patient's individual preferences and needs based on gender or familiarity when creating schedules and routes for healthcare providers. [26] study a daily home care routing and scheduling problem to minimize total routing and overtime costs. They consider that patients may specify preferences regarding the nurses that perform the jobs. When a nurse is assigned to a job, a penalty is incurred based on their level of preference, which is indicated as preferred, moderately preferred, or not preferred, resulting in penalties of 0, 1, or 2, respectively. Moreover, for each job, the corresponding patient can prefer a time for the start of the service. Similarly, [28] study a home care service planning problem considering the preferences of both the caregiver and the patient. In addition, patients are given the opportunity to communicate their preferred visiting times so patients can prefer the service start time. To address the problem, they propose a two-phase method based on a set partitioning formulation. Furthermore, [14] propose to formulate the HHCRSP as a set partitioning problem that aims at planning the best routes for each caregiver. They consider patients' preferences based on gender or the language spoken by the assigned caregiver. As a result, the integration of the preference constraint within the HHCRSP is of paramount importance, as it enables healthcare providers to customize schedules and routes according to the unique preferences and requirements of individual patients. This consideration not only enhances patient satisfaction but also facilitates the delivery of personalized care, aligning healthcare services more closely with patients' specific needs and preferences.

When planning the route and schedule of HHC, it is important to carefully schedule the time and location of each caregiver's lunch break as it presents another challenging constraint. [29] study an HHCRSP considering synchronized visits and lunch breaks. They assume that lunch breaks are fictive visits and exist in each caregiver's route and that lunch break is mandatory. To address the problem, they develop an MIP and propose four hybrid metaheuristics, which are a memetic algorithm, a hybrid genetic general variable neighborhood search, a hybrid genetic, and a hybrid simulated annealing. According to numerical analysis, hybrid genetic general variable neighborhood search shows the best performance among four algorithms. [30] formulate an MILP for an HHCRSP with a daily planning horizon to minimize total operating costs. They consider flexible lunch break constraint that a lunch break is mandatory for a caregiver if his/her work timetable completely includes the given time window interval. Otherwise, the assignment of a lunch break is not necessary. In this study, since the lunch break is accepted as a need of the caregivers, a mandatory lunch break constraint is applied.

Due to the problem's nature and the diverse range of constraints involved, several solution methods have been utilized, including exact algorithms, heuristics, metaheuristics, and matheuristics [21]. In the context of solving the HHCRSP, the application of exact algorithms are highly valued due to their capability to deliver optimal solutions. These algorithms systematically examine the entire problem space, leaving no possibility for a better solution within the specified constraints and objectives. For example, [31] propose a branch-and-price algorithm for the HHCRSP considering stochastic service times and skill requirements. They aim to minimize total travel cost, fixed cost of caregivers, service cost, and penalty for late arrival at patients. [32] study an HHCRSP considering the synchronized services of multi-skilled caregivers necessitated by the simultaneous service requirements of the patients. To solve the problem, their approach involves formulating the problem as a set-partitioning problem and implementing a branch-and-price-and-cut solution algorithm. To obtain lower bounds for the problem, they develop a column generation scheme. These studies exemplify the diverse range of solution methods employed in tackling the HHCRSP, illustrating the need for innovative and customized approaches to effectively handle the problem's complexities and optimize scheduling and routing decisions. While exact algorithms may demand more computational resources and time compared to heuristic approaches, their significance lies in their ability to provide certainty and guarantee the most optimal solutions for the HHCRSP. This precision can greatly benefit healthcare providers and decision-makers by allowing them to make informed decisions and enhance the overall quality and efficiency of HHCSs.

Due to the complexity of the problem applying exact methods may not always be a straightforward solution method, especially with large-sized instances. Therefore, heuristics, metaheuristic, and matheuristic algorithms are frequently proposed in the literature for the solution of HHCRSP. For instance, [33] study an HHCRSP considering time windows and synchronization constraints. Both hard time windows and soft time windows of patients are considered. To tackle the problem, they develop an MIP and propose a memetic algorithm featuring two original crossover operators. To test the problem, they use real-life instances from an HHC provider in France. According to experimental analysis, the memetic algorithm provides efficient results. [34] study an HHCRSP with multiple days planning horizon. Caregivers assign both patients and days. An MIP is developed to solve the problem, and they propose a three-phase matheuristic approach based on the decomposition of the formulation to simplify the research, which is called the novel mixed problem. According to the analysis, the matheuristic algorithm is more efficient than the MIP model from the computational time and objective aspects. [35] solve an HHCRSP to maximize the total priority of the visited patients and to minimize the total traveling time by using real-life data. They propose assigning priorities to patients according to factors such as the last visit time. To address the problem, they develop the adaptive large neighborhood search algorithm and a matheuristic to generate near-optimal solutions. According to numerical results, both algorithms yield high-quality solutions, but the matheuristic outperforms the adaptive large neighborhood search algorithm in large-sized instances. In this thesis, matheuristics are applied, which are crucial in the HHCRSP, as they offer practical solutions within a reasonable computational time. These methods effectively address the complexity of the problem and facilitate efficient decision-making in the HHCRSP.

In this problem, we study an HHCRSP with a multi-period time horizon considering patients' preferences, caregivers' lunch breaks, and prior service (blood draw). A significant disparity between these investigations and our study is found in the requirement of visiting each patient on a given day in this study, whereas not every patient needs to be visited in the mentioned studies. In this study, unlike other studies, there is a preferred time interval for each day of the patients. Patients are assigned to days and caregivers, considering these time intervals. First, the MILP model is developed for small-sized instances to address this problem. Then, the SAMA is proposed to solve medium and large-sized instances. The KNN is applied to predict the treatment time of the patients. The problem properties and the performance of the MILP and the SAMA are analyzed, and the results of the solutions are compared with that of the current system. Then, sensitivity analysis is applied to test the objective function. Finally, managerial insights are provided considering the results of these analyses.

CHAPTER III

PROBLEM DEFINITION

In this thesis, we investigate an HHC setting where patients require a range of treatments, including regular examinations, throughout a specified time horizon. Caregivers, who are designated as the individuals responsible for delivering services, visit patients using healthcare vehicles. The main hospital is the origin for health caregivers, and their objective is to visit each patient within one of their preferred time windows. However, due to limitations in the number of caregivers and vehicles available, caregivers may need to arrive earlier or later than the preferred start time for a patient's treatment. The specific treatment days and caregiver assignments are not determined at the beginning of the week, with only the preferred time windows known for each day due to reasons such as the need for someone to support them with the patients. Therefore, patients are assigned to both days and caregivers.

Our model requires caregivers to visit each patient within the specified time horizon, even if it results in violating the patient's preferred time windows. Such violations are penalized due to their potential negative impact on the patient's health condition. Importantly, caregivers are not allowed to interrupt ongoing treatments, strictly prohibiting treatment preemption. The caregivers start their work at the hospital and return to the hospital after completing their assigned patient visits, following the operational procedures of the motivating hospital. The travel costs of caregivers, which are determined based on travel times between patients, capture the caregivers' travel to patients' locations.

The patients in this thesis are dispersed across a region and require examinations for various treatments. Figure 1 provides an exemplary representation of the locations of patient nodes and the hospital node. When a caregiver visits patients who need blood draw (indicated by red circles in Figure 1), it is crucial for the caregiver to return to the hospital before or at noon. Subsequently, the caregiver will continue visiting the remaining patients assigned to their route.

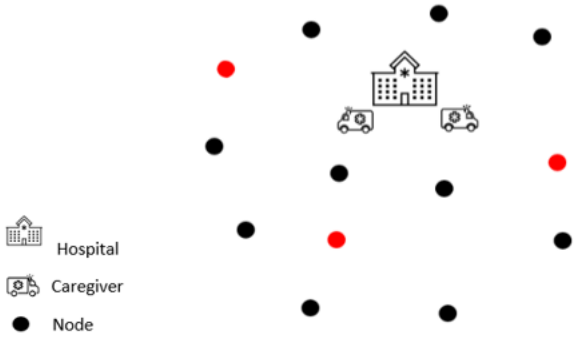

Figure 1: Representation of patient node distribution and hospital

Given the limited number of caregivers, each caregiver team possesses the same skill set and is assigned to a specific vehicle. Patients who need blood draw are given higher priority as their blood test tubes need to be sent to the hospital's lab before noon to get the test results on the same day. This priority ensures the timely prescription of necessary medications, if required. If a patient who needs blood draw finishes treatment early in the morning, the caregiver can still visit other patients who do not need blood draw and return to the hospital before noon. After delivering the blood test tubes to the hospital's laboratory, caregivers continue visiting the remaining patients in their assigned routes.

Some patients have preferences regarding caregivers based on factors such as gender and familiarity. Additionally, caregivers are required to take a 30-minute lunch break between 12:30 and 13:30. This break is scheduled either near the caregiver's current location after examining a patient or at the hospital if the vehicle returns to deliver blood test tubes. Caregivers must adhere to the maximum daily working hours and return to the hospital by that time for operational procedures and paper works.

The objective of this study is to develop a scheduling and routing plan that considers patients' preferences, prior service type, and lunch break of caregivers. The aim is to minimize the total routing cost of caregivers and penalty costs due to early and late start of service at preferred time windows of patients.

The proposed mathematical model encompasses all the routing and scheduling constraints inspired by the operational dynamics of the motivating hospital. The nomenclature used in our problem formulation is presented in Table 1, followed by the detailed description of the proposed mathematical model.

Table 1: Nomenclature

Sets	
$\mathcal{P}$	Patients ($i \in \{1, 2, \ldots, \overline{p}\}$)
$\mathcal{B}$	Patients who need blood draw ($i \in \{1, 2, \ldots, \overline{b}\}$)
$\mathcal{N}$	Nodes ($n \in \{0, 1, 2, \ldots, \overline{n}+1, \overline{n}+2\}$)
$\mathcal{V}$	Vehicles ($v \in \{1, 2, \ldots, \overline{v}\}$)
$\mathcal{D}$	Days ($d \in \{1, 2, \ldots, \overline{d}\}$)
$\mathcal{A}$	Patients who prefer the first caregiver ($a \in \{1, 2, \ldots, \overline{a}\}$)
$\mathcal{A}'$	Patients who prefer the second caregiver ($a' \in \{1, 2, \ldots, \overline{a'}\}$)
Parameters	
c_{ij}	Travel cost from node i to node j
$time_{ij}$	Travel time from node i to node j
dur_i	Treatment duration of patient i
$start_i^e$	Earliest start time of patient i's treatment
$start_i^l$	Latest start time of patient i's treatment
MW	Maximum working hour
B	Lunch break duration
$lunch^e$	Earliest start time of the lunch break
$lunch^l$	Latest start time of the lunch break
tp	Unit travel cost coefficient
$penalty^e$	Unit early arrival penalty if caregivers arrive earlier than the earliest start time of a patient's treatment
$penalty^l$	Unit late arrival penalty if caregivers arrive later than the latest start time of a patient's treatment
Decision variables	
x_{ijvd}	1 if the vehicle v travels from node i to node j on day d, 0 otherwise
y_{ivd}	1 if the vehicle v arrives at patient i that needs blood draw on day d, 0 otherwise
lb_{ivd}	1 if the caregivers at vehicle v takes the lunch break before visiting patient i on day d, 0 otherwise
t_{ivd}	Time when vehicle v arrives at node i, on day d
sl_{vd}	Lunch break start time for the caregivers at vehicle v on day d
w_{iv}	Time difference between the arrival of vehicle v and the earliest start time of patient i's treatment in case of an earlier arrival
l_{iv}	Time difference between the arrival of vehicle v and the latest start time of patient i's treatment in case of a late arrival

$$\min \quad tp \sum_{d\in D}\sum_{v\in V}\sum_{i\in N}\sum_{j\in N, i\neq j} c_{ij}x_{ijvd} + penalty^e \sum_{v\in V}\sum_{i\in P} w_{iv} + penalty^l \sum_{v\in V}\sum_{i\in P} l_{iv} \tag{1}$$

$$\text{s.t.} \quad \sum_{d\in\mathcal{D}}\sum_{v\in\mathcal{V}}\sum_{i\in\mathcal{N}\setminus\{\overline{n}+2\}} x_{ijvd} = 1 \qquad \forall j \in \mathcal{P} \tag{2}$$

$$\sum_{i\in\mathcal{N}\setminus\{\overline{n}+2\}, i\neq j} x_{ijvd} = \sum_{i'\in\mathcal{N}\setminus\{0\}, i'\neq j} x_{ji'vd} \qquad \forall j \in \mathcal{N}\setminus\{0, \overline{n}+2\}, v \in \mathcal{V}, d \in \mathcal{D} \tag{3}$$

$$\sum_{j\in\mathcal{N}\setminus\{0\}} x_{ijvd} \leq \sum_{j'\in\mathcal{N}\setminus\{\overline{n}+2\}} x_{j'ivd} \qquad \forall i \in \mathcal{N}\setminus\{0, \overline{n}+2\}, v \in \mathcal{V}, d \in \mathcal{D} \tag{4}$$

$$\sum_{j\in\mathcal{P}} x_{0jvd} = 1 \qquad \forall v \in \mathcal{V}, d \in \mathcal{D} \tag{5}$$

$$\sum_{i\in\mathcal{N}\setminus\{0, \overline{n}+2\}} x_{i(\overline{n}+2)vd} = 1 \qquad \forall v \in \mathcal{V}, d \in \mathcal{D} \tag{6}$$

$$\sum_{i\in\mathcal{N}\setminus\{\overline{n}+1, \overline{n}+2\}} x_{ijvd} \leq \sum_{i'\in\mathcal{P}} x_{i'(\overline{n}+1)vd} \qquad \forall j \in \mathcal{B}, v \in \mathcal{V}, d \in \mathcal{D} \tag{7}$$

$$x_{0jvd} = 0 \qquad \forall j \in \{\overline{n}+1, \overline{n}+2\}, v \in \mathcal{V}, d \in \mathcal{D} \tag{8}$$

$$x_{i0vd} = 0 \qquad \forall i \in \mathcal{N}\setminus\{0\}, v \in \mathcal{V}, d \in \mathcal{D} \tag{9}$$

$$x_{(\overline{n}+2)jvd} = 0 \qquad \forall j \in \mathcal{N}, v \in \mathcal{V}, d \in \mathcal{D} \tag{10}$$

$$\sum_{i \in \mathcal{P} \cup \{\overline{n}+1\}} lb_{ivd} = 1 \qquad \forall v \in \mathcal{V}, d \in \mathcal{D} \tag{11}$$

$$lb_{ivd} \leq \sum_{j \in \mathcal{N} \backslash \{0\}, i \neq j} x_{ijvd} \qquad \forall i \in \mathcal{P} \cup \{\overline{n}+1\}, v \in \mathcal{V}, d \in \mathcal{D} \tag{12}$$

$$t_{ivd} + (time_{ij} + dur_i) x_{ijvd} \leq \ t_{jvd} + (1 - x_{ijvd}) start_i^l \qquad \forall i \in \mathcal{N} \backslash \{\overline{n}+2\}, j \in \mathcal{N} \backslash \{0\}, v \in \mathcal{V}, d \in \mathcal{D} \tag{13}$$

$$sl_{vd} + (B lb_{jvd}) \leq \ t_{jvd} + (1 - lb_{jvd}) lunch^l \qquad \forall j \in \mathcal{P} \cup \{\overline{n}+1\}, v \in \mathcal{V}, d \in \mathcal{D} \tag{14}$$

$$t_{ivd} + (time_{ij} + dur_i)(x_{ijvd} + lb_{jvd} - 1) \leq \ sl_{vd} + (2 - x_{ijvd} - lb_{jvd}) start_i^l$$
$$\forall i \in \mathcal{P} \cup \{\overline{n}+1\}, j \in \mathcal{P} \cup \{\overline{n}+1\}, v \in \mathcal{V}, d \in \mathcal{D} \tag{15}$$

$$w_{iv} \geq \ start_i^e - t_{ivd} \qquad \forall i \in \mathcal{P} \cup \{\overline{n}+2\}, v \in \mathcal{V}, d \in \mathcal{D} \tag{16}$$

$$l_{iv} \geq \ t_{ivd} - start_i^l \qquad \forall i \in \mathcal{P} \cup \{\overline{n}+2\}, v \in \mathcal{V}, d \in \mathcal{D} \tag{17}$$

$$x_{ij1d} = 1 \qquad \forall i \in \mathcal{A}, j \in \mathcal{N} \backslash \{0\}, d \in \mathcal{D}, i \neq j \tag{18}$$

$$x_{ij2d} = 1 \qquad \forall i \in \mathcal{A}', j \in \mathcal{N} \backslash \{0\}, d \in \mathcal{D}, i \neq j \tag{19}$$

$$0 \leq \ t_{ivd} \leq \ MW \qquad \forall i \in \mathcal{P} \cup \{\overline{n}+1\}, v \in \mathcal{V}, d \in \mathcal{D} \tag{20}$$

$$lunch^e \leq \ sl_{vd} \leq \ lunch^l \qquad \forall v \in \mathcal{V}, d \in \mathcal{D} \tag{21}$$

$$t_{jvd} \leq lunch^l y_{ivd} \qquad \forall i \in \mathcal{B}, j \in \{\overline{n}+1\}, v \in \mathcal{V}, d \in \mathcal{D} \tag{22}$$

$$\sum_{i \in \mathcal{P} \cup \{0\}} x_{ijvd} \leq blood_{i'vd} \qquad \forall i' \in \mathcal{B}, j \in \{\overline{n}+1\}, v \in \mathcal{V}, d \in \mathcal{D} \tag{23}$$

$$x_{ijvd} \in \{0, 1\} \qquad \forall i, j \in \mathcal{N}, v \in \mathcal{V}, d \in \mathcal{D} \tag{24}$$

$$y_{ivd} \in \{0, 1\} \qquad \forall i \in \mathcal{B}, v \in \mathcal{V}, d \in \mathcal{D} \tag{25}$$

$$lb_{ivd} \in \{0, 1\} \qquad \forall i \in \mathcal{P} \cup \{\overline{n}+1\}, v \in \mathcal{V}, d \in \mathcal{D} \tag{26}$$

$$t_{ivd} \geq 0 \qquad \forall i \in \mathcal{N}, v \in \mathcal{V}, d \in \mathcal{D} \tag{27}$$

$$sl_{vd} \geq 0 \qquad \forall v \in \mathcal{V}, d \in \mathcal{D} \tag{28}$$

$$w_{iv}, l_{iv} \geq 0 \qquad \forall i \in \mathcal{P} \cup \{\overline{n}+2\}, v \in \mathcal{V} \tag{29}$$

The objective function (1) aims to minimize the total routing cost, and the total earliness and lateness arrival penalties. Constraints (2) ensure that each patient is visited exactly once, thereby avoiding duplicate visits. The flow balance constraints (3) are imposed to maintain the balance of caregivers' visits to patients. They ensure that a caregiver leaves a patient after providing the necessary services and that the total number of visits to and from each patient is equal. Constraints (4) to (6) specify the routing for the caregivers. They indicate that every caregiver starts their route from the hospital, can visit multiple patients during their route, and must return to the hospital after completing their assigned visits. To address the need for timely blood sample delivery, constraints (7) ensure that if a patient requiring blood draw is visited by a caregiver, the corresponding vehicle must return to the hospital to deliver the blood test tubes to the lab. This requirement facilitates prompt analysis of the blood samples, enabling timely treatment decisions. Constraints (8) to (10) impose specific restrictions on the vehicles' movements. They prevent the vehicles from going to the lab or returning to the hospital immediately after leaving in the morning, ensuring that the necessary services are completed before any additional movements. The constraints also restrict the vehicles from returning to the hospital after their evening visit, preventing further service activities for that day. Constraints (11) ensure that each vehicle takes a break during its route. These breaks, typically scheduled for lunchtime, allow caregivers to rest and attend to their personal needs. Constraints (12) establish the link between the lunch break variables and the vehicle routing variables. Constraints (13) ensure that the starting service time of a patient is greater than or equal to the starting time of the previous patient plus their respective service time. Constraints (14) and (15) determine the correct start times for the lunch breaks when caregivers rest at patients' locations before and after providing the required services. These constraints consider the travel time between patients and the time durations of the services. Constraints (16) and (17) define the early and late arrival times of the caregivers to patients' locations. These constraints capture the desired time window within which the caregivers should arrive, preventing both excessively early and late arrivals. Constraints (18) and (19) incorporate patients' preferences regarding caregivers, such as preferences based on gender or familiarity if there is any preference listed by the patients. Constraints (20) determine the maximum working hour for caregivers. Constraints (21) define the time windows for the start of the lunch breaks. Constraints (22) guarantee that all patients requiring blood draw are visited and attended to before noon, allowing sufficient time for the blood samples to be analyzed and necessary treatments to be prescribed promptly. Constraints (23) indicate that if blood is collected from a patient, the corresponding vehicle must have visited the blood node, ensuring that the collection of blood samples is accurately recorded and accounted for in the scheduling and routing plan. Finally, constraints (24) to (29) impose binary restrictions and nonnegativity on the decision variables.

CHAPTER IV

SIMULATED ANNEALING-BASED MATHEURISTIC ALGORITHM (SAMA)

Our proposed mathematical model is nontrivial to solve within a reasonable amount of time as the instances get larger since both routing and scheduling decisions are computed. Thus, we propose a matheuristic relying on decomposition and simulated annealing (SA). Since in terms of computational complexity, the suggested algorithm relies on SA, a widely recognized metaheuristic approach utilized for solving challenging problems [36]. Indeed, it is possible to enhance the results when the SA solutions are utilized as a starting point for an exact method [37]. Therefore, the proposed SA-based matheuristic uses the power of an exact method to obtain the best routes for the caregivers after the patients are assigned to the caregivers.

In our study, first, the algorithm decomposes the problem into two subproblems as *assignment* and *routing*. These two subproblems are solved iteratively through the SA algorithm framework. Firstly, as the initial solution for SA, we use random assignment to diversify the search space. During this assignment phase, patients' preferences and blood draws are considered. After assigning patients to vehicles and days, we plan the route for each vehicle using the MILP solver to find the best route. In each iteration of the SA, we randomly select patients and reassign them to other days and vehicles. Then, we determine the route again using MILP. These movements continue until the termination criteria of the SA are reached.

The assignment subproblem deals with assigning patients to vehicles. Since the schedules of the vehicles are independent for each day, a *dummy vehicle* is defined with an ID ($\hat{v} = \bar{d} \times \bar{v}$). At first, the patients are randomly assigned, and in order to be able to generate neighborhood moves, a chromosome-based answer representation is utilized. An exemplary chromosome is depicted in Figure 2 when there are ID set of 10 ($\bar{p}$) patients with 2 ($\bar{v}$) vehicles for a single day ($\bar{d}$=1). {0, $\bar{p}$+2} represent the hospital, and {$\bar{p}$+1} again represents the hospital but only when the hospital is visited in order to deliver the test tubes to the lab. The length of the chromosome equals the number of patients, and each patient has an assigned vehicle information in the chromosome. In phase 1, patients' preferences and prior service types are also considered. For example, in Figure 2, {0,11,12} represent the hospital and patient 8 prefers the first caregiver, and patient 2 needs a blood draw. Then, each patient is assigned to each dummy vehicle ($\hat{v}$). There are two dummy vehicles; let's assume the patients that must be visited by the first vehicle are {1,3,4,7,8,10}, whereas the patients that must be visited by the second vehicle are {2,5,6,9}. Needless to say, in phase 1, the algorithm only decides on the assignments of the patients to the vehicles and days. The routes that the vehicles will follow will be computed in the second phase of the algorithm.

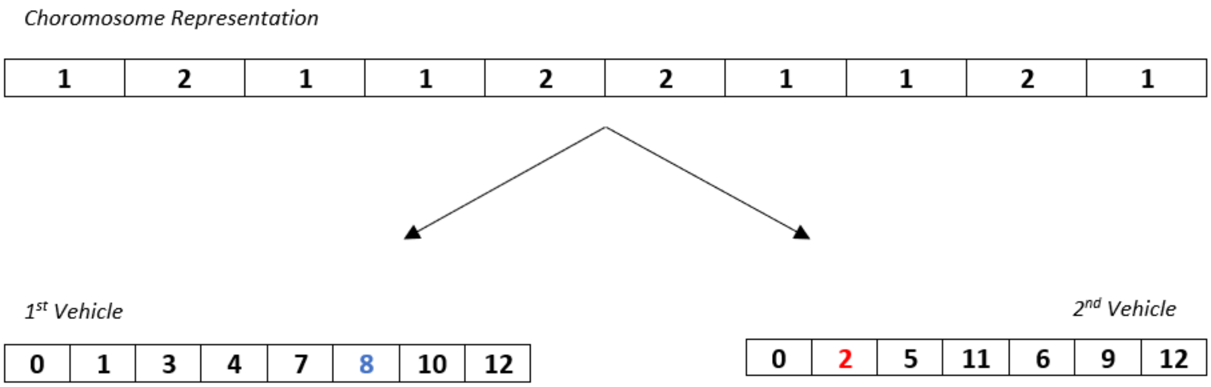

Figure 2: Chromosome representation

After assigning the patients to the vehicles and days, the routing schedule of the vehicles (dummy vehicles) is calculated by solving the MILP for each of them separately for the set of patients that are assigned to them.

We utilize the proposed decomposed phases within the SA framework to find the best assignment and routing accordingly. Pseudo-code of the proposed SAMA is presented in Algorithm 1. In the algorithm, T_0 represents the initial temperature which is the starting value of the temperature when the SA starts, and α represents the cooling rate, that is, the rate at which the temperature is reduced during the process in each iteration. First, the algorithm begins by initializing the SA parameters. Then, it iteratively applies *neighbourhood moves* to the current solution and evaluates the quality of the new solution by solving the MILP. If the new solution is an improvement, it becomes the current solution. If it is not an improvement, it may still be accepted as the current solution with a probability that it decreases as the temperature T decreases. The temperature is decreased at each iteration according to a cooling rate ($T_{New} = \alpha T_{Current}$). The algorithm terminates when the temperature reaches a certain threshold ($T > 1$). The best solution found during the course of the algorithm is returned as a result.

Algorithm 1 SAMA (α, T_0)

```
Initialize the values: $TC_{Best} = +\infty, TC_{Current} = +\infty, TC_{Temp} = 0, T = T_0, bestpop = \emptyset$
$pop$ := Generate a random string with the size of $\hat{v} + \bar{p} - 1$
$bestpop = pop$
for $k \in (1 : \hat{v})$ do
    C := Solve MILP with one vehicle
    $TC_{Temp} = TC_{Temp} + C$
end for
$TC_{Best} = TC_{Temp}, TC_{Current} = TC_{Temp}$
while $T > 1$ do
    $newpop$ := Choose a neighborhood move for $pop$
    $TC_{Temp} = 0$
    for $k \in (1 : \hat{v})$ do
        C := Solve MILP with one vehicle
        $TC_{Temp} = TC_{Temp} + C$
    end for
    if $TC_{Temp} < TC_{Current}$ then
        $TC_{Current} = TC_{Temp}, pop = newpop$
        if $TC_{Current} < TC_{Best}$ then
            $TC_{Best} = TC_{Current}$
            $bestpop = pop$
        end if
    else
        $\Delta = TC_{Temp} - TC_{Current}$
        $Rand$ := Random rational number $\in (0, 1)$
        if $Rand \leq e^{-\Delta/T}$ then
            $pop = newpop$
        end if
    end if
    $T_{New} = \alpha T_{Current}$
end while
```

CHAPTER V

PREDICTION OF TREATMENT TIMES

The healthcare industry has witnessed significant advancements with the integration of machine learning (ML), which has helped to manage large amounts of data and improve patient outcomes [38]. Improving prediction accuracy in treatment planning could greatly assist doctors in eliminating these difficulties [39]. For example, [40] study aims to explore the usage of a support vector machine in the prediction of dementia and validate its performance based on statistical analysis. [41] provide a survey of current techniques of knowledge discovery in databases using ML techniques which will be useful for medical practitioners to make effective decisions. They aim to predict more accurately the presence of heart disease with a reduced number of attributes. [42] study different computerized heart disease prediction systems with different classification techniques, which are naïve bayes, KNN, decision tree, and bagging to predict the disease more accurately and efficiently. In our previous study, we aim to predict the service time for HHC applications with the ultimate goal of enhancing patient care [43]. Our primary objective is to analyze the significant features that influence the determination of service time. To achieve this, we create a correlation matrix to examine the relationships between these features and their association with the service time.

In this problem, we use the KNN, which is one of the most widely used state-of-the-art ML technique for classification problems to predict patients' treatment durations [44]. The KNN, a supervised algorithm, predicts the classification of unlabeled data by considering the features and labels of the training data [45, 46]. It works by finding the k data points in the training set that is closest to the new data point and using these k data points to make a prediction. To use the KNN algorithm, the k value, the number of neighbors, is specified. The KNN algorithm does not require prior assumptions about the data distribution. Hence this makes it a suitable choice for our application, where data may not necessarily follow a particular distribution. Steps of the KNN algorithm are as follows [47].

1. *Calculating the distance of input records from all training records.*
2. *Sorting the training records based on their distances and selecting the KNNs.*
3. *Using the class which owns the majority among the KNNs.*

We apply the KNN algorithm to predict the treatment time required for a given patient based on features, such as their age, lab request, and treatment type, as presented in Table 2. The KNN algorithm is trained on a dataset of previous HHC visits, with the service time as the target variable and the patient features as the input variables. Once trained, the KNN algorithm is used to make predictions for new patients, allowing caregivers to schedule visits with appropriate service times.

Table 2: Summary of attributes utilized for the KNN algorithm

#	Attribute	Explanation
1	Patient ID number	Identity number for each patient
2	Age	Age of each patient
3	Lab	Is there a lab request for the patient or not?
4	Treatment type	Service types of patients
5	Treatment duration	Service duration of patients

CHAPTER VI

COMPUTATIONAL RESULTS

We evaluate the performance of the solutions obtained via solving the MILP and the SAMA based on computational times, objective function values, and the number of patients visited outside their preferred time windows and perform sensitivity analyses by varying the coefficients of the objective function, the number of patients requiring blood draw, and the number of vehicles. We first assume the motivating hospital's blood tube handling system as described in Chapter 3. If both vehicles visit patients requiring blood draw, both vehicles return to the hospital to deliver their blood tubes. We call this case *vehicle dispatch for lab assignment without meeting point.* Then, we study a slight modification of this case in which only one vehicle returns to the hospital to deliver the blood tubes after transferring all blood tubes to that vehicle at one of the predetermined meeting points. This case is called *vehicle dispatch for lab assignment with meeting point*, and we analyze the change in the performance by comparing the cases *with* and *without the meeting point.*

Additionally, we compare our proposed solutions' quality with the current system employed by our motivating hospital. In their current system, the whole service region is divided into smaller regions, and vehicles are assigned to visit the nearest patients within their respective regions. However, this system presents challenges, such as unbalanced patient and vehicle assignments, increased patient service delays, and inadequate comply patients' preferences.

The SAMA incorporates two user-defined parameters: an initial temperature (T_0) and a cooling rate (α). To determine appropriate parameter values, experiments are conducted on medium and large-sized instances using various parameter combinations. T_0 and α are set as 100 and 0.85, respectively.

The MILP model and SAMA are implemented using the Python programming language and the Gurobi optimization solver. To ensure computational efficiency, a running time limit of 4 hours (14,400 seconds) is set for the MILP. All computational experiments are performed on a computer system running Windows 11, equipped with an 11th Generation Intel(R) Core(TM) i7-1165G7 @ 2.80GHz 2.80 GHz processor.

First, we introduce our setting in Section 6.1. In Section 6.2, we analyze our main case study (i.e., without meeting points) and examine the results obtained by vehicle dispatch for lab assignment (i.e., with meeting points).

6.1 Our Setting

Hospital-Caregivers Information:

HHCSRP has been motivated by the existing challenges faced by a real hospital in Turkey. The hospital is situated in a city with a population exceeding 350 thousand. An exemplary graphical layout for our setting is represented in Figure 3, where grids represent patients' houses, and the triangle represents the hospital.

Figure 3: An examplary patients' distribution over the territory

The caregivers' working hours are from 8.30 to 18.00, but the last two hours of their day are dedicated to entering that day's data into the treatments' medical database and writing necessary prescriptions, especially after evaluating the lab results for patients requiring blood draw. In other words, caregivers provide HHCS between 8.30 and 16.00 (450 minutes). This additional time allocation ensures the completion of administrative tasks and allows caregivers to fulfill their responsibilities effectively within the designated working hours. Additionally, during the time period of 12:30 to 13:30, the caregivers engage in a scheduled lunch break lasting for a duration of 30 minutes.

Patient Information:

In the hospital's HHC patient registration system, the patients are categorized into five distinct groups based on disease types and treatment methods: (1) urethral catheterization treatment, (2) medical dressing, (3) blood draw, (4) medical checkup, and (5) initial examination and patient registration. Each group represents a specific type of treatment provided to the patients in their respective homes.

The first group, *urethral catheterization treatment*, involves the careful insertion of a urinary catheter by a caregiver. Due to the meticulous nature of this procedure and the need to avoid any complications such as leaking, the duration of this treatment is relatively longer compared to other types. *Medical dressing*, on the other hand, is a service provided to care for wounds at home following surgical procedures. This involves the application of dressings to promote healing and prevent infections. Patients in the *blood draw* group require treatments that involve the collection of blood samples. These samples are necessary for laboratory testing and further diagnosis of the patients' conditions. *Medical checkup* services encompass routine examinations and controls for elderly patients. These checkups are conducted to monitor their health status and detect any potential issues at an early stage. The final group consists of new patients who have not received any previous treatment at home. This service, known as *initial examination and patient registration*, involves conducting an initial assessment of the patient's condition and registering them into the HHC system. By categorizing patients into these distinct groups, the hospital's HHC system can effectively allocate caregivers and resources based on the specific treatment requirements of each patient.

The durations of different treatment types are not explicitly provided by the motivating hospital. Instead, we predict them via the KNN algorithm. The KNN algorithm utilizes historical data to estimate treatment durations for various patient groups. Specifically, a dataset containing 3028 records spanning a three-month period is used. This dataset consists of patient features, such as medical parameters, patient ID numbers, and the corresponding service times. To train and evaluate the algorithm, the dataset is divided into a training set (80% of the records) and a testing set (20% of the records). The KNN algorithm, as described in Chapter V, is applied to the dataset using different values of k from the set {2, 3, 5, 10, 15}. The accuracy of the predictions for each value of k is shown in Table 3. The results indicate that the highest prediction accuracy of 99.94% is achieved when k is set as 3, while the lowest accuracy of 93.96% is obtained when k equals 15. Based on these findings, a value of k equal to 3 is selected for predicting treatment durations, and the resulting predicted durations are represented in Table 3.

Table 3: Accuracy (%) computed via the KNN for varying k values

k value	2	3	5	10	15
Accuracy (%)	98.99	99.94	97.49	97.69	93.96

When applying the KNN algorithm with a value of k equal to 3, the predicted treatment durations for each patient group are obtained. These predicted durations provide estimates for the time required to complete the respective services. The results of the predicted durations for each patient group are presented in Table 4. This table displays the expected service times for the different patient groups, which

Table 4: Predicted treatment durations (minutes) of patient groups

Patient group	Service types	Treatment durations
1	Urinary catheterization treatment	40
2	Medical dressing	30
3	Blood draw	20
4	Medical checkup	20
5	Initial examination and patient registration	15

include urethral catheterization treatment, medical dressing, blood draw, medical checkup, and initial examination and patient registration. The predicted durations offer valuable insights into the expected time commitment for each type of service, aiding in the efficient scheduling and allocation of caregivers and resources.

The complete list of instances, including the number of patients, the presence of blood draw requirements, and patient preferences, is presented in Table 5. This table provides a comprehensive overview of the instances used in our study, allowing for a systematic comparison of the results obtained from each instance.

Table 5: Instance sets

Instances	Size	# of days	# of patients	# of patients who need blood draw	# of preference
1	Small	1	15	1	2
2	Small	1	15	4	1
3	Small	1	15	-	2
4	Small	1	15	5	-
5	Small	1	15	1	2
6	Medium	2	30	3	2
7	Medium	2	30	4	3
8	Medium	2	30	7	-
9	Medium	2	30	-	2
10	Medium	2	30	2	2
11	Large	4	45	-	2
12	Large	4	45	2	1
13	Large	4	45	1	-
14	Large	4	45	2	-
15	Large	4	45	5	-

The necessary data for our study, including the hospital's location, the set of caregivers, and the available procedures, are obtained from the motivating hospital. Subsequently, we generate 15 instances with varying sizes to evaluate the proposed approach. In the small-sized instances, there are 15 patients who need to be visited within a single day. Moving on to the medium-sized instances, there are 30 patients that should be visited within a span of two days. Lastly, the large-sized instances consist of 45 patients, and their visits are spread across a four-day period. In all instances, two vehicles are assumed in the motivating hospital.

Patient information, including their exact location and names, is handled with utmost care to ensure the confidentiality and privacy of individuals. To maintain confidentiality, patient location coordinates are generated randomly using a uniform distribution. The distances between patients and the hospital are assumed to be Euclidean, and travel durations are calculated proportionally based on these distances. The cost of travel is determined by considering travel time and fuel prices. Furthermore, the penalty for early arrival ($penalty^e$) is set to 5, while the penalty for late arrival ($penalty^l$) is set to 15. It is important to note that the penalty for late arrival is higher than the penalty for early arrival, emphasizing the importance of the possibility of degrading the health conditions of patients in case of late arrival.

6.2 Numerical Results

In the context of healthcare services, it is of utmost importance that a healthcare vehicle, when visiting a patient who needs blood draw, promptly returns to the hospital during midday hours. This enables the timely conduction of necessary laboratory tests and the retrieval of test results on the same day. In this case study, crucially, the vehicles need to prioritize visiting the laboratory after serving patients who need blood draw before noon. In cases where two vehicles are assigned to serve patients requiring blood draw, both vehicles should proceed to the laboratory. However, if only one vehicle is assigned to such patients, only that particular vehicle should make the trip to the laboratory.

For instance, Figure 4 provides a visual representation of exemplary routes generated by the MILP model for a specific instance (Instance 1). The routes depicted in the figure involve 15 patients and 2 vehicles, covering the duration of a single day. Notably, the patient requiring blood draw (P1) is highlighted in red, indicating the need for special attention and the requirement for the vehicle to return to the hospital.

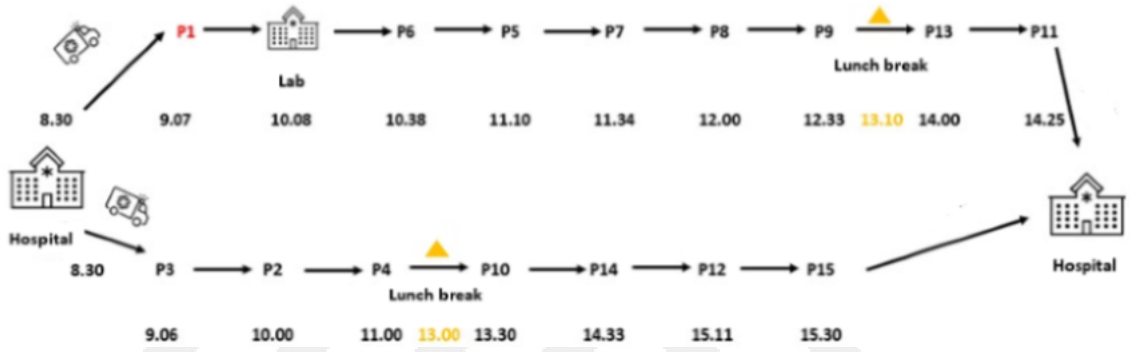

Figure 4: Routes result of the MILP for instance 1 with 2 vehicles and 1 day when patient needing blood draw is written in red

In Figure 4, the start time of each patient's service is indicated below their name. Additionally, the presence of a yellow triangle symbolizes the location where caregivers take their lunch break during their routes. Notably, in this specific instance, the second vehicle does not need to return to the hospital (lab). This is because the assigned caregiver does not visit any patients who need blood draw. Instead, the caregiver strategically plans their lunch break near the location of one of the patients, optimizing the overall efficiency of the route. In addition to the different patient groups based on treatment types, there can also be variations in patient preferences. Some patients may have specific preferences, such as preferring a caregiver of a particular gender. In such cases, the assignment of caregivers takes these preferences into account. For example, in Figure 4, both patient 1 (P1) and patient 9 (P9) have expressed a preference to receive services exclusively from the first caregiver.

Furthermore, each patient is assigned a specific appointment interval, such as 9.00 - 11.00 or 15.15 - 17.00, within which the caregivers are expected to start their services. In Figure 4, the time indicated below each patient represents the actual starting time of their treatment. For instance, patient 3 (P3) has been allocated an appointment interval of 9.00 - 11.00, and the treatment for P3 begins at 9.06. However, there is also a possibility for caregivers to arrive earlier than the earliest start time or later than the latest start time specified in the appointment interval. To discourage such deviations from the appointment intervals, early or late arrivals incur penalties in the objective function.

Considering the Home Health Care (HHC) environment, our objective function aims to minimize the total cost. A positive improvement indicates a reduction in costs compared to the current system. The improvement (%) of the MILP is calculated as follows.

$$\frac{Obj_{current} - Obj_{MILP}}{Obj_{current}} \times 100$$

Similarly, the improvement (%) of the SAMA approach is calculated using the same formulation.

$$\frac{Obj_{current} - Obj_{SAMA}}{Obj_{current}} \times 100$$

The improvement achieved by the MILP and SAMA approaches is presented in Table 6, providing a comparative analysis of their performance.

In the case of small-sized instances, significant cost improvements are observed when employing both the MILP and SAMA. Specifically, the MILP approach achieves the greatest improvement, reducing the total cost by 98.64% in instance 3, while the SAMA approach achieves a reduction of 98.58% in the same instance. To put it differently, both the MILP and SAMA methods result in a substantial 98.64% reduction in total cost for instance 3.

The CPU time (in seconds) for both algorithms varies across instances. The MILP algorithm demonstrates shorter CPU times for five instances compared to the SAMA

Table 6: Improvement (%) of the MILP model and SAMA based on current system

Instances	MILP (%)	CPU time	SAMA (%)	CPU time
1	98.02	6.54	97.96	128.07
2	97.60	13.26	97.60	1348.79
3	98.64	2.95	98.58	465.35
4	97.85	81.94	97.74	229.02
5	97.58	8.25	97.31	734.43
6	-	14,400	95.75	6,006.11
7	-	14,400	93.81	2,531.47
8	-	14,400	94.25	2,022.46
9	-	14,400	97.07	213.52
10	-	14,400	95.77	1,393.11
11	-	14,400	94.35	463.29
12	-	14,400	93.81	1,142.86
13	-	14,400	94.66	1,379.66
14	-	14,400	94.16	1,722.55
15	-	14,400	93.13	4,342.36
Average	97.94		95.57	
Best	98.64		98.58	
Worst	97.58		93.13	

method. However, the MILP algorithm experiences longer CPU time for instance 4 due to a higher number of patients requiring blood draw. Consequently, the MILP algorithm requires more time to identify the optimal solution, given the complexity introduced by a large number of patients requiring blood draw. Notably, according to the MILP results, instance 2 exhibits the second highest CPU time due to patients who need blood draw as well as patients with preferences.

The MILP approach does not yield results for medium and large-sized instances within the allocated time of 14,400 seconds. In medium-sized instances, the most significant improvement of 97.07% is observed in instance 9. Notably, this instance exhibits the lowest CPU time among the medium-sized instances, as there are no patients requiring blood draw. On the other hand, instance 6 experiences a considerably high CPU time due to the inherent characteristics of the SAMA approach. The SAMA encounters challenges in achieving better objective function values during the execution of neighborhood moves, leading to a higher number of iterations.

In the case of large-sized instances, the most substantial improvement of 94.66% is observed in instance 13. However, instance 15 has the highest CPU time among all instances, primarily due to the presence of five patients requiring blood draw. Once again, the problem-solving complexity increases as the number of patients needing blood draw rises.

As a result, in the case of the small-sized instance, the MILP result reveals that the best improvement achieved is 98.64%, while the worst improvement stands at 97.58%, with an average improvement of 97.94%. Conversely, based on the SAMA result, the best improvement obtained is 98.58%, with the worst improvement recorded at 93.13%. The average improvement, as per the SAMA result, amounts to 95.57%.

To gain insights into the impact of the MILP and SAMA on routing, as well as the disparities compared to the current system, we conduct testing on instance 1, which is comprehensively described in Chapter 6.2. The resulting routes from the MILP, SAMA, and the current system are presented in Figure 5. In the figure, the left side visually represents the route of vehicle 1, while the right side illustrates the route of vehicle 2. The patients requiring blood draw are highlighted in red (P1), and the green route shows the trajectory of the vehicle following the delivery of blood tubes to the hospital.

The application of the MILP algorithm results in the following routes for vehicle 1: {P1, lab, P6, P5, P7, P8, P9, P13, P11}, while the route for vehicle 2 is as follows: {P3, P2, P4, P10, P14, P12, P15}. Vehicle 2, which does not visit any patients requiring blood draw, does not return to the lab before the noon.

On the other hand, the SAMA generates the following routes: vehicle 1: {P1, P2, lab, P5, P7, P8, P9, P13, P11}, and vehicle 2: {P6, P3, P4, P10, P14, P12, P15}. Patients (P6 and P2) to different vehicles. According to the MILP, vehicle 1 visits P6, whereas the SAMA result indicates that vehicle 2 visits to P6. Similarly, the

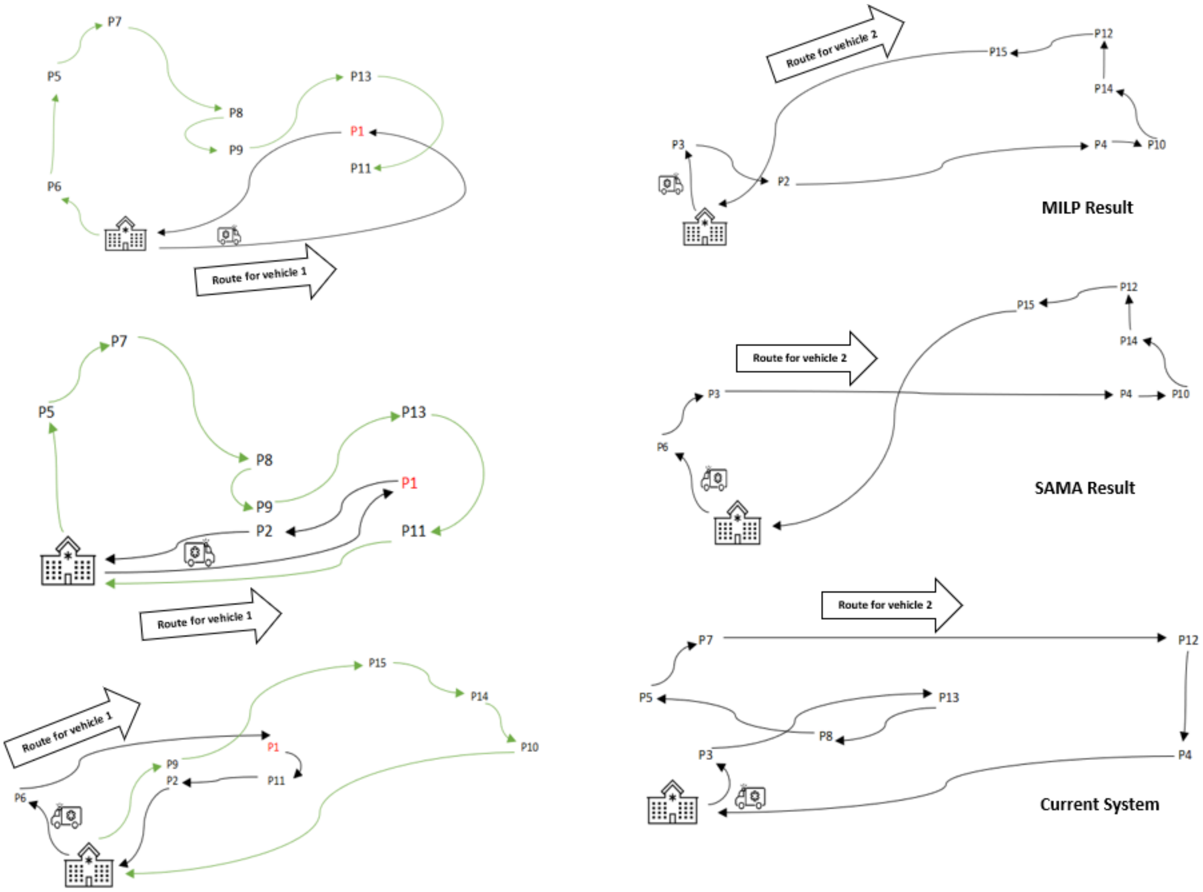

Figure 5: Comparison routing of the vehicles according to the MILP, SAMA, and current system

MILP assigns P2 to vehicle 2, whereas the SAMA result assigns P2 to vehicle 1.

In contrast to the previous results, the routing plan of the current system reveals the following routes: vehicle 1: {P6, P1, P11, P2, lab, P9, P15, P14, P10}, and vehicle 2: {P3, P13, P8, P5, P7, P12, P4}. In the current system, the routes of both vehicles differ significantly from the routes determined by the MILP and SAMA.

Table 7 illustrates the number of patients visited either early or late in each instance, as computed by both SAMA and the current system. It is important to note that the MILP can only compute results for 5 instances, and based on these results, there are no patients who are visited late or early within their preferred time window. As shown in Table 7, the current system incurs higher penalties compared to SAMA due to the occurrences of early or late visits.

Table 7: Comparison of the number of patients visited early (e) and late (l) based on results of the SAMA and the current system

Instances	# SAMA$_e$	# SAMA$_l$	# Current$_e$	# Current$_l$
1	-	-	10	2
2	-	-	10	3
3	-	-	12	2
4	-	-	10	2
5	1	-	10	3
6	1	-	17	2
7	3	-	18	4
8	3	-	18	6
9	-	1	19	6
10	-	-	17	6
11	-	-	34	-
12	-	-	32	3
13	1	-	35	-
14	-	-	35	1
15	-	1	34	-

The objective function is reported to have three components: *travel cost* (P_t), *late penalty* (P_l), and *early penalty* (P_e). Table 8 presents the values of these three components computed via the SAMA when the coefficients of routing cost, late penalty, and early penalty are set to 1, 5, and 15, respectively. As shown in Table 8, P_t is the dominant component in all instances except instances 8 and 15. This is because the number of patients needing a blood draw is higher in these instances compared to the other instances, resulting in more patients being outside their preferred visiting time window.

In addition, we have conducted additional sensitivity analyses on instance 5 to illustrate the impact of the number of patients requiring a blood draw. These analyses are presented in Section 6.2.1, as the number of patients needing a blood draw significantly influences the routing and scheduling plan.

Table 8: The objective function breakdown (%) for coefficients (1, 5, 15) is based on the results obtained from the SAMA

Instances	P_t	P_e	P_l
1	100	0	0
2	100	0	0
3	100	0	0
4	100	0	0
5	98.74	1.25	0
6	98.22	1.77	0
7	67.57	32.42	0
8	27.89	72.10	0
9	76.55	0	23.44
10	100	0	0
11	100	0	0
12	100	0	0
13	100	0	0
14	100	0	0
15	29.56	0	70.43

6.2.1 Sensitivity Analyses

In the sensitivity analysis, we examine the effects of varying coefficients, patient numbers, and vehicles on the performance metrics of computation times and objective function values. By systematically adjusting these parameters, we can observe how the system responds and identify potential areas for improvement of optimization.

Sensitivity Analysis of the Impact of Different Coefficient of Objective Function:

In this part, we refrain from analyzing the objective function results with different coefficients using the SAMA and MILP. This comparison would not be fair because multiplying different coefficients results in different magnitudes, making it challenging to draw meaningful conclusions.

The objective function breakdown is conducted in SAMA to analyze the impact of three cost components: total travel cost, early penalty cost, and late penalty cost. Table 9 presents the percentage of each cost's contribution to the objective function, where P_t represents the percentage of the travel cost's effect, P_e indicates the percentage of the early penalty cost's effect, and P_l denotes the effect of the late penalty cost on the objective function. Various breakdowns are observed in the objective function due to the distinct characteristics of each instance.

Table 9: Objective function breakdown (%) for instances in SAMA with different coefficients

	(1,5,15)			(1,5,5)			(1,1,1)		
Instances	P_t	P_e	P_l	P_t	P_e	P_l	P_t	P_e	P_l
1	100	0	0	100	0	0	100	0	0
2	100	0	0	100	0	0	40.56	0	59.43
3	100	0	0	100	0	0	100	0	0
4	100	0	0	100	0	0	99.07	0.22	0.7
5	98.74	1.25	0	100	0	0	97.49	0.71	1.79
6	98.22	1.77	0	100	0	0	87.84	0.25	11.9
7	67.57	32.42	0	87.92	0	12.08	100	0	0
8	27.89	72.10	0	98.82	0	1.17	99.43	0	0.56
9	76.55	0	23.44	88.75	0	11.24	97.71	0	2.28
10	100	0	0	100	0	0	98.68	0	1.31
11	100	0	0	100	0	0	99.52	0.47	0
12	100	0	0	99.6	0.4	0	100	0	0
13	100	0	0	99.55	0.44	0	100	0	0
14	100	0	0	100	0	0	100	0	0
15	29.56	0	70.43	99.60	0	0.39	99.06	0.7	0.22

Figure 6 displays the corresponding CPU time graph. In general, when the coefficient of late penalty increases, CPU time decreases. The reason behind the decrease in CPU time as the coefficient of the late penalty increases is likely due to the effect of the penalty's weight on the optimization process. When the coefficient of the late penalty is higher, the optimization algorithm may prioritize minimizing late visits more aggressively. Consequently, the search space may be limited, leading to faster convergence and reduced computation time.

The same analysis methodology is employed for the MILP in Table 10. In general,

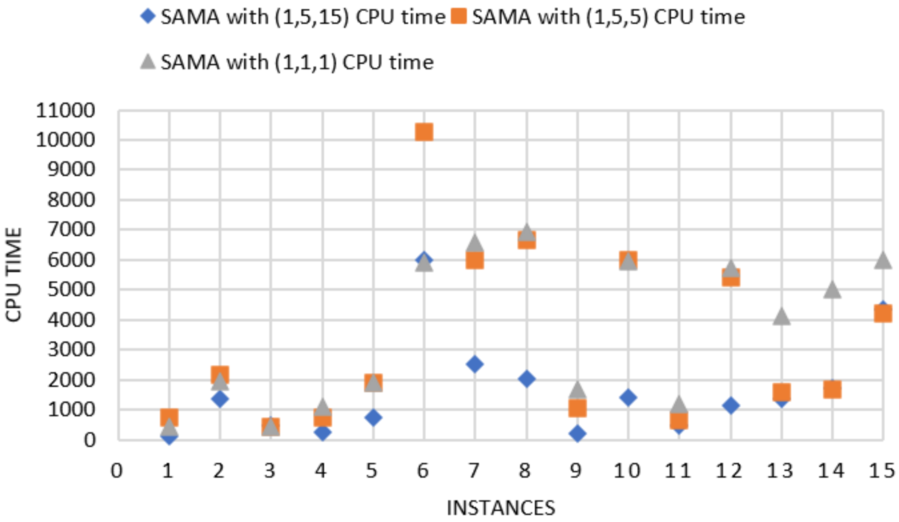

Figure 6: Comparison of CPU time (in seconds) in SAMA with different objective coefficients

we observe that a decrease in the coefficient of penalty costs results in a higher incidence of penalty costs. Nonetheless, it is essential to underscore that the total cost predominantly emanates from the total routing cost in the general context. It is crucial to note that this outcome is contingent upon the unique characteristics and specific attributes of each individual instance under consideration.

Table 10: Objective function breakdown (%) for instances in MILP with different coefficients

	(1,5,15)			(1,5,5)			(1,1,1)		
Instances	P_t	P_e	P_l	P_t	P_e	P_l	P_t	P_e	P_l
1	100	0	0	10	72	18	40	38	22
2	100	0	0	100	0	0	99	0	1
3	100	0	0	100	0	0	100	0	0
4	100	0	0	100	0	0	98	1	1
5	100	0	0	99	1	0	100	0	0

The analysis of the SAMA results with different coefficients in the objective function provides insights into the number of patients visited early and late. Table 11 reveals that when $penalty^l$ is set to 15, only two instances out of the total 15 show patients being visited late. In the case of the coefficient set (1,5,15), when there are patients visited earlier than their appointment time, no instances indicate patients being visited late. This observation can be attributed to the higher penalty cost associated with late visits compared to early visits.

Table 11: Comparison of patients visited early (e) and late (l) in MILP with different objective coefficients

Instances	Current Coefficients		Coefficients 1		Coefficients 2	
	# $(1,5,15)_e$	# $(1,5,15)_l$	# $(1,5,5)_e$	# $(1,5,5)_l$	# $(1,1,1)_e$	# $(1,1,1)_l$
1	-	-	-	-	-	-
2	-	-	-	-	1	-
3	-	-	-	-	-	-
4	-	-	-	-	1	1
5	1	-	-	-	1	1
6	1	-	-	-	1	-
7	3	-	2	1	-	-
8	3	-	-	1	-	1
9	-	1	-	1	-	1
10	-	-	-	-	-	1
11	-	-	-	-	2	-
12	-	-	1	-	-	-
13	1	-	2	-	-	-
14	-	-	-	-	-	-
15	-	1	-	1	2	1

Table 12 presents compelling insights regarding the impact of setting $penalty^e$ and $penalty^l$ to different coefficients. When employing the coefficient set (1,5,15), no instances indicate patients being visited late or early in the MILP. This distinction can be attributed to the higher penalty cost assigned to late visits in comparison to early visits, highlighting the significance of penalty cost differentials in influencing scheduling outcomes.

Considering the potential negative impact of late visits on the patients' health, it is recommended to assign a higher penalty for visiting patients late. Therefore, the coefficient set (1,5,15) yields the best results in terms of minimizing late visits and promoting timely healthcare delivery.

,

Table 12: Comparison of patients visited early (e) and late (l) in MILP with different objective coefficients

Instances	Current Coefficients		Coefficients 1		Coefficients 2	
	# $(1,5,15)_e$	# $(1,5,15)_l$	# $(1,5,5)_e$	# $(1,5,5)_l$	# $(1,1,1)_e$	# $(1,1,1)_l$
1	-	-	2	1	2	2
2	-	-	-	-	-	1
3	-	-	-	-	-	-
4	-	-	-	-	1	1
5	-	-	1	-	-	-

Sensitivity Analysis of the Impact of Increasing Number of Vehicle:

The comparison of the mathematical model for the problem with two and three vehicles is presented in Table 13. In this analysis, the number of vehicles is maintained at a constant of two, as specified in the original problem statement. However, the impact on the objective function is examined by increasing the number of vehicles to three. The objective function is evaluated for two scenarios based on the results obtained using the MILP approach.

Table 13: Objective function results comparison with different number of vehicles based on the MILP

Instances	Objective_{2_v}	Objective_{3_v}	μ_{2_v}	μ_{3_v}
1	154.25	154.06	77.12	51.35
2	252.92	239.69	126.46	79.89
3	138.55	171.09	69.27	57.03
4	234.38	206.72	117.19	68.90
5	225.08	215.23	112.43	71.74

The table focuses on the small-sized instances where the model can find feasible solutions. The first column represents the small-sized instances, while the second and third columns display the objective function values for the scenarios using two and three vehicles, respectively. In this context, the objective function solely considers the total travel costs, as no penalty costs are incurred during the resolution of the initial set of instances. Thus, the second and third columns directly reflect the total travel time. The fourth and fifth columns present the average travel cost per vehicle (μ) for the two scenarios.

The results indicate that, in most instances, using three vehicles leads to a lower total travel cost compared to using two vehicles. However, in instance 3, employing three vehicles results in a higher total travel cost. The reason for this is that the patients in this particular scenario are clustered in specific regions. As a result, when there are three vehicles available, the patients in these concentrated regions have to share the third vehicle, which leads to an extended travel distance.

We conduct a similar analysis using the SAMA method, and the findings are presented in Table 14. We observe consistent results with the MILP. When the number of vehicles is increased to three, the objective function decreases, except for instance 3. In instance 3, due to the specific distribution of patients' locations, the objective function value increases despite the additional vehicle.

Table 14: Objective function results comparison with different number of vehicles based on the SAMA

Instances	Objective_{2_v}	Objective_{3_v}	μ_{2_v}	μ_{3_v}
1	158.74	154.06	79.37	51.35
2	252.92	239.69	126.46	79.89
3	144.97	182.03	72.48	60.67
4	245.9	206.72	122.95	68.90
5	249.94	215.23	124.97	71.74

The key consideration here is that while employing three vehicles may result in lower routing costs, acquiring a new vehicle also comes with fixed costs and requires additional caregivers and resources. These added expenses should be taken into account when evaluating the overall cost-effectiveness of increasing the number of vehicles.

Sensitivity Analysis of the Impact of Increasing Number of Patients Needing Blood Draw:

Instance 5 is selected as the test case for performing a detailed analysis since it can be solved by both MILP and SAMA. Additionally, this instance includes patients who need blood draw and have preferences. Table 15 presents the results obtained from the MILP method. In the current configuration of instance 5, one patient needs a blood draw, while two patients have preferences. As shown in Table 15, augmenting the number of patients in need of blood draw leads to alterations in the routing plan. This adjustment necessitates visiting more patients in the morning and ensuring the timely delivery of blood tubes to the hospital before noon. Consequently, the objective function value increases as the number of patients requiring blood draw rises. Additionally, the increased number of patients with preferences causes a penalty cost for patients who are visited earlier, further contributing to the objective function value increment.

Table 15: Sensitivity analysis varying the number of patients who need blood draw using MILP

# Patients needing blood draw	# Preference	Objective	CPU time	# Patient$_e$	# Patient$_l$
-	2	176.27	3.65	-	-
1	2	224.86	8.25	-	-
2	2	235.53	46.89	-	-
3	2	319.6	247.94	1	-

The same analysis is applied to SAMA concerning patients needing a blood draw. Table 16 shows that as the number of patients needing a blood draw increases, both the objective function value and CPU time display an upward trend. This is because the model becomes more complex with an increasing number of patients needing blood draw.

Table 16: Sensitivity analysis varying the number of patients who need blood draw using SAMA

# Patients needing blood draw	# Preference	Objective	CPU time	# Patient$_e$	# Patient$_l$
-	2	183.63	728.47	-	-
1	2	249.94	734.43	-	-
2	2	338.06	2118.33	1	-
3	2	377.12	2112.72	-	2

6.2.2 Additional Analysis: What Happens If We Have Meeting Point for Delivering Blood Tubes Before Going to the Lab?

In our main study, which is mentioned in Chapter 6.2, the two vehicles are utilized for transporting blood samples from patients. They are dispatched to the lab before noon, where the blood tests are conducted, and subsequently, the samples are delivered. However, when two vehicles are dispatched to patients requiring a blood draw, both vehicles do not proceed directly to the lab. Instead, the blood tubes collected from both vehicles are combined and transferred to a single vehicle. Subsequently, the other vehicle changes its route to serve the other patients.We define two alternatives for meeting points and the choice of vehicles that will return to the hospital to deliver the blood tubes. In other words, we do not fix the meeting points; instead, we have two alternatives for them, and we also do not fix which vehicle will be used for the return to lab.

To facilitate these analyses, we update the existing model to incorporate the meeting point cases. We define a new set containing meeting point locations, and a binary decision variable is defined to determine which meeting point the vehicles should go to. However, we summarize the changes of mathematical model. In addition to our original setting, we enforce that vehicles should meet at one of the meeting point once if they both visit a patient needing blood draw. Then model also decide which vehicle return the hospital to deliver the blood tubes.

We analyze two different scenarios, namely *Scenario 1* and *Scenario 2*, and then compare the results with the *Current Case*. In the *Current Case*, both vehicles proceed directly to the lab without any meeting point. In *Scenario 1*, the vehicles meet at a designated meeting point before continuing their services. After the meeting, only one vehicle proceeds directly to the lab. In *Scenario 2*, the vehicles also meet at the designated meeting point, but there is a difference in their subsequent actions. One vehicle serves other patients if there is enough time, whereas if there is insufficient time, it goes directly to the lab.

Table 17: Comparative analysis of the MILP result: current case vs. meeting point scenarios

Instances	Patients needing blood draw	# Preferences	Current Case		Scenario 1		Scenario 2	
			Objective	CPU Time	Objective	CPU Time	Objective	CPU Time
2	4	1	252.92	6.97	232.93	13.26	221.74	35.46
4	5	0	238.17	55.1	226.49	81.94	226.49	77.56

As indicated in Table 17, we select instance 2 and 4 because they involve multiple numbers of patients needing a blood draw. The decision to analyze only these instances is based on the MILP model's outcome, which confirms that both vehicles are assigned to serve blood patients exclusively in these instances.

In instance 2, the *Scenario 2* yields the most favorable outcome. According to the numerical results, the vehicles prioritize serving the blood patients before converging at the designated meeting point. Following this, the first vehicle proceeds to attend to other patients, while the second vehicle proceeds to the lab. Notably, before heading to the lab, the second vehicle also serves two additional patients.

As shown in Figure 7, which is the result of instance 2 route, both vehicles leave the hospital in the morning; the first vehicle goes to the meeting point after going to the patient needing blood draw (P2). Meanwhile, the second vehicle visits other patients needing blood draw (P3, P1, and P4) and goes to the meeting point. At the meeting point, the blood tubes are collected in the second vehicle, and since the second vehicle has time to go to the lab, it serves two more patients (P5 and P6)

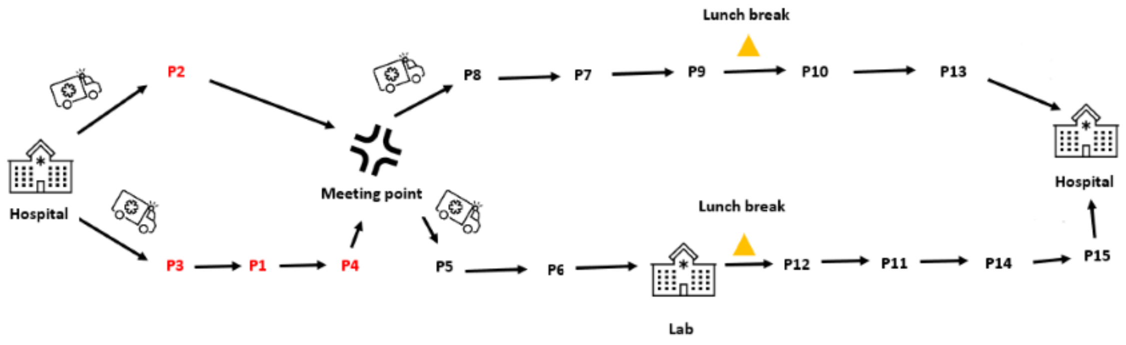

Figure 7: Scenario 2 result of the MILP for instance 2

and then goes to the lab and delivers the blood tubes. Vehicle 1 continues to serve other patients after the meeting point, and vehicle 2 continues to serve other patients after the lab. Furthermore, the first vehicle provides a designated lunch break before serving patient 10 (P10), whereas the second vehicle schedules its lunch break before serving patient 12 (P12).

In instance 4, both *Scenario 1* and *Scenario 2* produce the same and superior outcome compared to the current case. The results for *Scenario 1* and *Scenario 2* in this instance are identical, as the vehicle designated for lab transport does not have any additional time to serve other patients prior to proceeding to the lab.

The computational experiment results have convincingly demonstrated that a significant cost reduction can be achieved by implementing the MILP model and the proposed SAMA. By developing improved assignment and routing plans, our approach outperforms the current methods by a considerable margin. This finding sheds light on the potential of computational techniques to optimize resource allocation in complex systems.

CHAPTER VII

CONCLUSION

The increase in life expectancy due to advances in science and technology has led to a rise in the elderly population worldwide. As birth rates decline and the proportion of older individuals in the population increases, there is a growing demand for healthcare services. HHCS has emerged as a widespread approach to meet this increasing demand, as it provides healthcare services to patients in the comfort of their own homes. HHCS not only improves accessibility to healthcare but also helps alleviate the issue of overcrowding in hospitals, particularly during public health emergencies like the COVID-19 pandemic. By delivering care at home, HHCS reduces the need for hospital admissions, minimizing the risk of infection transmission and freeing up valuable hospital resources for more critical cases. This highlights the significance of HHCS in addressing the evolving healthcare needs of the population. However, the successful implementation of HHCS relies on overcoming operational challenges, particularly in solving the scheduling and routing problems faced by HHCS teams. Efficient scheduling and routing are crucial to optimize travel distances and ensure timely delivery of care. The HHCRSP is a complex problem that requires careful consideration of factors such as time windows and patient preferences.

In this context, this thesis aims to solve the HHCRSP faced by a real-life public hospital. The MILP and SAMA are proposed to solve the problem efficiently by analyzing different cases. Additionally, the study employs the KNN algorithm to predict the treatment time of patients, enhancing resource allocation and overall system efficiency. The research contributes to the field by addressing the complexity of the HHCRSP and providing accurate and effective solutions. It extends the traditional HHCRSP by considering diverse service types, caregiver assignments within specified time intervals, lunch breaks, patients' preferences, and prior service type (blood draw). The MILP model and SAMA offer systematic and efficient approaches to solve the problem, with the SAMA outperforming in medium and large-sized instances. Numerical analyses are conducted to assess the efficiency of the proposed solutions, considering computation time, objective function values, and the number of patients visited outside their preferred time windows. The results demonstrate the effectiveness of the MILP model and SAMA in optimizing operations in HHCS, leading to significant cost improvements compared to the current system used in the motivating hospital.

In future work, several promising areas can significantly advance decision-making in healthcare. One key focus is exploring the diversity of caregiver skills, availability, and resource allocation. Understanding caregivers' unique proficiencies allows strategic assignment to specific cases, enhancing patient outcomes and fostering a positive work environment. Effective resource allocation is critical as the industry faces evolving challenges. Leveraging advanced ML techniques enables continuous learning from patient data, leading to more accurate predictions and streamlined operations. This approach optimizes service times, patient care, and overall healthcare efficiency.

REFERENCES

[1] O. Guduk, O. Guduk, Y. Sertbas, and I. Satman, *Evde Sağlık Hizmetleri Raporu.* Türkiye Sağlık Enstitüleri Başkanlığı Türkiye Halk Sağlığı ve Kronik Hastalıklar Enstitüsü, 2021.

[2] P. Khodabandeh, V. Kayvanfar, M. Rafiee, and F. Werner, "A bi-objective home health care routing and scheduling model with considering nurse downgrading costs," *International Journal of Environmental Research and Public Health*, vol. 18, no. 3, p. 900, 2021.

[3] J. Decerle, O. Grunder, A. H. El Hassani, and O. Barakat, "A hybrid memetic-ant colony optimization algorithm for the home health care problem with time window, synchronization and working time balancing," *Swarm and Evolutionary Computation*, vol. 46, pp. 171–183, 2019.

[4] T. N. A. for Home Care and Hospice, "Basic statistics about home care," tech. rep., Washington, DC 20003, 2010.

[5] E. Lanzarone, A. Matta, and E. Sahin, "Operations management applied to home care services: the problem of assigning human resources to patients," *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, vol. 42, no. 6, pp. 1346–1363, 2012.

[6] "Home healthcare unit." Online, 2022. https://hasekieah.saglik.gov.tr/EN-522868/home-healthcare-unit.html.

[7] F. Ghiasvand Ghiasi, M. Yazdani, B. Vahdani, and A. Kazemi, "Multi-depot home health care routing and scheduling problem with multimodal transportation: Mathematical model and solution methods," *Scientia Iranica*, 2021.

[8] M. Di Mascolo, C. Martinez, and M.-L. Espinouse, "Routing and scheduling in home health care: A literature survey and bibliometric analysis," *Computers & Industrial Engineering*, vol. 158, p. 107255, 2021.

[9] A. M. Fathollahi-Fard, M. Hajiaghaei-Keshteli, R. Tavakkoli-Moghaddam, and N. R. Smith, "Bi-level programming for home health care supply chain considering outsourcing," *Journal of Industrial Information Integration*, vol. 25, p. 100246, 2022.

[10] S. Bahadori-Chinibelagh, A. M. Fathollahi-Fard, and M. Hajiaghaei-Keshteli, "Two constructive algorithms to address a multi-depot home healthcare routing problem," *IETE Journal of Research*, vol. 68, no. 2, pp. 1108–1114, 2022.

[11] M. Cissé, S. Yalçındağ, Y. Kergosien, E. Şahin, C. Lenté, and A. Matta, "Or problems related to home health care: A review of relevant routing and scheduling problems," *Operations Research for Health Care*, vol. 13-14, pp. 1–22, 2017.

[12] A. M. Fathollahi-Fard, M. Hajiaghaei-Keshteli, and S. Mirjalili, “A set of efficient heuristics for a home healthcare problem,” *Neural Computing and Applications*, vol. 32, pp. 6185–6205, 2020.

[13] R. Liu, B. Yuan, and Z. Jiang, “Mathematical model and exact algorithm for the home care worker scheduling and routing problem with lunch break requirements,” *International Journal of Production Research*, vol. 55, no. 2, pp. 558–575, 2017.

[14] F. Grenouilleau, A. Legrain, N. Lahrichi, and L.-M. Rousseau, “A set partitioning heuristic for the home health care routing and scheduling problem,” *European Journal of Operational Research*, vol. 275, no. 1, pp. 295–303, 2019.

[15] S. V. Begur, D. M. Miller, and J. R. Weaver, “An integrated spatial dss for scheduling and routing home-health-care nurses,” *Interfaces*, vol. 27, no. 4, pp. 35–48, 1997.

[16] E. Cheng and J. L. Rich, “A home health care routing and scheduling problem,” tech. rep., 1998.

[17] M. Bazirha, A. Kadrani, and R. Benmansour, “Stochastic home health care routing and scheduling problem with multiple synchronized services,” *Annals of Operations Research*, vol. 320, no. 2, pp. 573–601, 2023.

[18] M. Demirbilek, J. Branke, and A. K. Strauss, “Home healthcare routing and scheduling of multiple nurses in a dynamic environment,” *Flexible Services and Manufacturing Journal*, vol. 33, no. 1, pp. 253–280, 2021.

[19] M. I. Gomes and T. R. P. Ramos, “Modelling and (re-) planning periodic home social care services with loyalty and non-loyalty features,” *European Journal of Operational Research*, vol. 277, no. 1, pp. 284–299, 2019.

[20] F. Grenouilleau, N. Lahrichi, and L.-M. Rousseau, “New decomposition methods for home care scheduling with predefined visits,” *Computers & Operations Research*, vol. 115, p. 104855, 2020.

[21] C. Fikar and P. Hirsch, “Home health care routing and scheduling: A review,” *Computers & Operations Research*, vol. 77, pp. 86–95, 2017.

[22] J. Euchi, M. Masmoudi, and P. Siarry, “Home health care routing and scheduling problems: a literature review,” *4OR*, vol. 20, no. 3, pp. 351–389, 2022.

[23] M. Bazirha, A. Kadrani, and R. Benmansour, “Optimization of the stochastic home health care routing and scheduling problem with multiple hard time windows,” *International Journal of Supply and Operations Management*, vol. 9, no. 2, pp. 235–250, 2022.

[24] J. Euchi, S. Zidi, and L. Laouamer, "A hybrid approach to solve the vehicle routing problem with time windows and synchronized visits in-home health care," *Arabian Journal for Science & Engineering*, vol. 45, pp. 10637–10652, 2020.

[25] G. Du, Y. Tian, and X. Ouyang, "Multi-resources co-scheduling optimization for home healthcare services under the constraints of service time windows and green transportation," *Applied Soft Computing*, vol. 131, p. 109746, 2022.

[26] K. Braekers, R. F. Hartl, S. N. Parragh, and F. Tricoire, "A bi-objective home care scheduling problem: Analyzing the trade-off between costs and client inconvenience," *European Journal of Operational Research*, vol. 248, no. 2, pp. 428–443, 2016.

[27] M. Di Mascolo, C. Martinez, and M.-L. Espinouse, "Routing and scheduling in home health care: A literature survey and bibliometric analysis," *Computers & Industrial Engineering*, vol. 158, p. 107255, 2021.

[28] P. Maya Duque, M. Castro, K. Sörensen, and P. Goos, "Home care service planning. the case of landelijke thuiszorg," *European Journal of Operational Research*, vol. 243, no. 1, pp. 292–301, 2015.

[29] W. Liu, M. Dridi, H. Fei, and A. H. El Hassani, "Hybrid metaheuristics for solving a home health care routing and scheduling problem with time windows, synchronized visits and lunch breaks," *Expert Systems with Applications*, vol. 183, p. 115307, 2021.

[30] L. Xiao, M. Dridi, and A. H. El Hassani, "Mathematical model for the home health care scheduling and routing problem with flexible lunch break requirements," *IFAC-PapersOnLine*, vol. 51, no. 11, pp. 334–339, 2018.

[31] B. Yuan, R. Liu, and Z. Jiang, "A branch-and-price algorithm for the home health care scheduling and routing problem with stochastic service times and skill requirements," *International Journal of Production Research*, vol. 53, no. 24, pp. 7450–7464, 2015.

[32] H. Qiu, D. Wang, Y. Yin, T. C. E. Cheng, and Y. Wang, "An exact solution method for home health care scheduling with synchronized services," *Naval Research Logistics (NRL)*, vol. 69, no. 5, pp. 715–733, 2022.

[33] J. Decerle, O. Grunder, A. Hajjam El Hassani, and O. Barakat, "A memetic algorithm for a home health care routing and scheduling problem," *Operations Research for Health Care*, vol. 16, pp. 59–71, 2018.

[34] S. Moussavi, M. Mahdjoub, and O. Grunder, "A matheuristic approach to the integration of worker assignment and vehicle routing problems: Application to home healthcare scheduling," *Expert Systems with Applications*, vol. 125, pp. 317–332, 2019.

[35] A. Cinar, F. S. Salman, and B. Bozkaya, “Prioritized single nurse routing and scheduling for home healthcare services,” *European Journal of Operational Research*, vol. 289, no. 3, pp. 867–878, 2021.

[36] A. Blum, C. Dan, and S. Seddighin, “Learning complexity of simulated annealing,” in *Proceedings of The 24th International Conference on Artificial Intelligence and Statistics* (A. Banerjee and K. Fukumizu, eds.), vol. 130 of *Proceedings of Machine Learning Research*, pp. 1540–1548, PMLR, 13–15 Apr 2021.

[37] O. Ozkan, “Optimization of the distance-constrained multi-based multi-uav routing problem with simulated annealing and local search-based matheuristic to detect forest fires: The case of turkey,” *Applied Soft Computing*, vol. 113, p. 108015, 2021.

[38] G. Battineni, N. Chintalapudi, and F. Amenta, “Machine learning in medicine: Performance calculation of dementia prediction by support vector machines (svm),” *Informatics in Medicine Unlocked*, vol. 16, p. 100200, 2019.

[39] S. Huang, J. Yang, S. Fong, and Q. Zhao, “Artificial intelligence in cancer diagnosis and prognosis: Opportunities and challenges,” *Cancer letters*, vol. 471, pp. 61–71, 2020.

[40] V. A. Kumari and R. Chitra, “Classification of diabetes disease using support vector machine,” *International Journal of Engineering Research and Applications*, vol. 3, no. 2, pp. 1797–1801, 2013.

[41] V. Chaurasia and S. Pal, “Data mining approach to detect heart diseases,” *International Journal of Advanced Computer Science and Information Technology (IJACSIT)*, vol. 2, pp. 56–66, 2014.

[42] N. Khateeb and M. Usman, “Efficient heart disease prediction system using k-nearest neighbor classification technique,” in *Proceedings of the International Conference on Big Data and Internet of Thing*, p. 21–26, Association for Computing Machinery, 2017.

[43] Y. S. Selcuk and E. Coban, “Advancing home healthcare through machine learning: Predicting service time for enhanced patient care,” in *2023 5th International Congress on Human-Computer Interaction, Optimization & Robotic Applications (HORA)*, pp. 1–6, 2023.

[44] M. Shouman, T. Turner, and R. Stocker, “Applying k-nearest neighbour in diagnosing heart disease patients,” *International Journal of Information and Education Technology*, vol. 2, no. 3, pp. 220–223, 2012.

[45] D. Bzdok, M. Krzywinski, and N. Altman, “Machine learning: supervised methods,” *Nature Methods*, vol. 15, no. 1, p. 5, 2018.

[46] S. Uddin, I. Haque, H. Lu, M. A. Moni, and E. Gide, “Comparative performance analysis of k-nearest neighbour (knn) algorithm and its different variants for disease prediction,” *Scientific Reports*, vol. 12, no. 1, pp. 1–11, 2022.

[47] M. Kuhkan, “A method to improve the accuracy of k-nearest neighbor algorithm,” *International Journal of Computer Engineering and Information Technology*, vol. 8, no. 6, p. 90, 2016.

VITA

Yağmur Selenay Selçuk, after completing her B.Sc. degree at Özyeğin University in 2021, pursued her academic journey by enrolling in the M.Sc. program in Industrial Engineering, at Özyeğin University, under the supervision of Asst. Prof. Elvin Çoban Göktürk. Throughout her studies at Özyeğin University, she actively contributed to the academic community as a teaching assistant and research assistant in the Department of Industrial Engineering. Her academic pursuits and research focus revolve around the fields of healthcare management, transportation science, and machine learning.