A DATA DRIVEN SOLUTION APPROACH FOR THE HOME HEALTH CARE PROBLEM

by
Aleyna Gürsoy

Submitted to Graduate School of Natural and Applied Sciences
in Partial Fulfillment of the Requirements
for the Degree of Master of Science in
Industrial and Systems Engineering

Yeditepe University
2023

A DATA DRIVEN SOLUTION APPROACH FOR THE HOME HEALTH CARE PROBLEM

APPROVED BY:

Prof. Dr. Dilek Tüzün Aksu ....
(Thesis Supervisor)
(Yeditepe University)

Assist. Prof. Dr. Eylül Damla Gönül Sezer ....
(Yeditepe University)

Assist. Prof. Dr. Umman Mahir Yıldırım ....
(İstanbul Bilgi University)

DATE OF APPROVAL: ..../..../2023

I hereby declare that this thesis is my own work and that all information in this thesis has been obtained and presented in accordance with academic rules and ethical conduct. I have fully cited and referenced all material and results as required by these rules and conduct, and this thesis study does not contain any plagiarism. If any material used in the thesis requires copyright, the necessary permissions have been obtained. No material from this thesis has been used for the award of another degree.

I accept all kinds of legal liability that may arise in case contrary to these situations.

Name, Last name Aleyna Gürsoy

Signature …..…………………….……………………

ACKNOWLEDGEMENTS

I would like to express my thanks to my supervisors, Dilek Tüzün Aksu and Semih Yalçındağ for their invaluable time and guidance throughout my thesis work. I would also like to express my gratitude for their efforts and understanding. Many thanks to my friends and distinguished professors of Industrial and Systems Engineering department for their kind support throughout my graduate study at Yeditepe University. Finally, I would like to thank my family for their continuous encouragement, support, and love.

ABSTRACT

A DATA DRIVEN SOLUTION APPROACH FOR THE HOME HEALTH CARE PROBLEM

Development in technology and healthcare has extended the average life expectancy globally and has led to an increase in the elderly population. Due to the change in population and increase in health awareness, the demand and need of individuals for healthcare services that prevent and treat diseases have increased as well. Especially during the COVID-19 pandemic, solutions have been developed for patients and elderly people who should not leave their houses to receive treatment at home. These services are called Home Health Care (HHC) systems. In the HHC system, there are patients who need to receive treatment at home and health care personnel who provide various health services to patients at their homes. In this context, human resource planning in the HHC system is crucial. The decisions that need to be made in human resources planning for HHC system consist of assigning the health care personnel to patients, scheduling the visits of health care personnel to patients, and constructing routes for the visits. The problem that integrates these three decisions is called the Home Health Care Problem (HHCP). Since HHCP is of NP-Hard complexity, solving it in an integrated manner within a reasonable time is not always possible. In this thesis, a HHCP variant that takes continuity of care, time windows of patients, and workload of operators into consideration is solved using a two-phase solution approach where the assignment and scheduling problems are handled in the first stage and the routing problem is handled in the second stage. In order to deal with the possible infeasibility in the routing stage and achieve high-quality results, the solution obtained from the two-phase approach is improved using a matheuristic framework that employs a Variable Neighborhood Search metaheuristic. Within the framework, the resulting routes are evaluated by a data driven method that estimates the route durations using the geographical locations of the patients from the historical routing information. In this context, the main topic of this thesis is to solve the HHCP using a matheuristic framework and to develop a data driven method to improve the solution obtained by the two-phase approach.

ÖZET

EVDE SAĞLIK HİZMETLERİ SİSTEMİ İÇİN VERİYE DAYALI BİR ÇÖZÜM YÖNTEMİ

Gelişen teknoloji ve sağlık imkanları, tüm dünyada ortalama yaşam süresinin uzamasını sağlamış ve yaşlı nüfusun artmasına sebep olmuştur. Değişen nüfus yapısı ve toplumda artan sağlık bilinci ile beraber, bireylerin hem hastalıkları önleyici, hem de hastalıkların tedavi edilmesine yardımcı olan sağlık hizmetlerine talebi ve ihtiyacı artmıştır. Özellikle COVID-19 salgını döneminde evden çıkmaması gereken hastaların ve yaşlıların evde tedavi görebilmesine yönelik çözümler geliştirilmeye çalışılmıştır. Bu talepler ve ihtiyaçlar doğrultusunda, Evde Sağlık Hizmetleri (ESH) sistemi ortaya çıkmıştır. ESH sisteminde hastalar ve hastaların evlerinde onlara çeşitli sağlık hizmetleri sağlaması gereken personeller mevcuttur ve bu bağlamda, sistem içindeki insan kaynakları planlaması önemli bir rol oynamaktadır. İnsan kaynakları planlaması kapsamında personellerin hastalara atanması, personellerin hastaları ne zaman ziyaret edeceğine dair çizelgeleme yapılması ve hasta ziyaretlerinde kullanılacak rotalamaların oluşturulması gerekmektedir. Bahsedilen adımlardan oluşan problem, Evde Sağlık Hizmetleri Problemi (ESHP) olarak adlandırılmaktadır. ESHP, NP-Zor karmaşıklıkta bir problem olduğundan dolayı, bütünleşik olarak makul sürede çözülmesi her zaman mümkün olmamaktadır. Bu tezin ana konusunu ise bakım sürekliliğini, hastaların zaman pencerelerini ve operatörlerin iş yükünü dikkate alan ESHP; ilk aşamada atama ve çizelgeme problemlerinin, ikinci aşamada ise rotalama probleminin ele alındığı bir çözüm yaklaşımı kullanılarak çözülmüştür. İki aşamalı yaklaşımdan elde edilen çözüm, rotalama aşamasında ortaya çıkabilecek olası rotalama sorunlarını gidermek veya yüksek kaliteli çözümler elde etmek için Değişken Komşuluk Arama metasezgiseli kullanan bir matsezgisel çerçeve içinde iyileştirilmiştir. Kullanılan matsezgisel çerçevede; çözümden elde edilen rotalar, geçmiş rotalama verilerinde bulunan hastaların coğrafi konumları kullanarak rota sürelerini tahmin eden veriye dayalı bir yöntem kullanılarak değerlendirilir. Bu bağlamda, ESHP'yi matsezgisel bir çözüm yöntemi kullanarak çözmek ve iki aşamalı yaklaşımla elde edilen çözümü iyileştirmek için veriye dayalı bir yöntem geliştirmek bu tezin ana konusudur.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS/ABBREVIATIONS

A	all possible paths
a_i	daily capacity of operator $i \in O$
D_{id}	workload of each operator
EST_j	earliest start time of service for patient $j \in N$
EST_{dj}	earliest start time of service for patient $j \in N$ on day $d \in H$
H	planning horizon
k_{max}	number of neighborhood structures
LFT_j	latest finish time of service for patient $j \in N$
LFT_{dj}	latest finish time of service for patient $j \in N$ on day $d \in H$
M	a very large number
m_u	variable for maximum workload
N	set of patients
n	number of patients on a route
NR_d	group of patients in RL for $d \in H$
N_k	set of neighborhood structures
n_{max}	maximum number of patients on a route
O	set of operators
O_d	set of operators available on day $d \in H$
P	set of patterns
r_j	care plan of patient $j \in N$
s_j	service time of patient $j \in N$
$t_{jj'}$	distance between patients j and j', $(j, j') \in A$
u_{ij}	binary variable for patient-operator assignments
u_{ij}^d	binary variable for daily patients visits of operators
$x_{jj'}$	binary variable for the arcs operators travel on
$x_{jj'}^{id}$	binary variable for the arcs operators travel on each day
y_j	start time of visit to patient $j \in N\backslash\{0\}$
y_{ijd}	start time of visit by operator $i \in O_d$ to patient $j \in N\backslash\{0\}$ on day $d \in H$
z_{jp}	binary variable for patient-pattern assignments

τ_j	average travel time to reach patient $j \in N$

ANN	Artificial neural network
B&P	Branch-and-price
D	Decomposition approach
GA	Genetic Algorithm
GVNS	General Variable Neighborhood Search
HHC	Home Health Care
HHCP	Home Health Care Problem
I	Integrated approach
KR	Kernel Regression
LNS	Large Neighborhood Search
LT	Long-term
MILP	Mixed integer linear programming
NN	Neural network
P	Partial
RL	Restriction List
S	Soft
ST	Short-term
TS	Tabu Search
TSP	Traveling Salesman Problem
TSPTW	Traveling Salesman Problem with Time Windows
VNS	Variable Neighborhood Search

1. INTRODUCTION

The elderly population has increased in recent years as worldwide average life expectancy has gotten longer. The rise in the elderly population resulted in increasing health problems such as chronic illnesses, cancer cases, and the number of physically disabled people. Along with the increasing health problems due to the aging population, the demand for preventative, protective, and supporting health care services has also increased which has induced the development of the Home Health Care (HHC) system. The need for the HHC system has drastically increased, especially during the COVID-19 pandemic as the elderly people with various health problems required to receive health care services at their homes [1].

In countries such as Türkiye, where the population is aging, the significance of HHC services is growing day by day. As it is stated in the 582nd objective of the Eleventh Development Plan prepared for Türkiye, the coverage of the HHC services is aimed to be expanded in the most efficient way [2].

In the HHC system, the primary objective is to deliver various medical and social services to patients at their home where they feel most comfortable, without putting themselves or their caretakers at risk of infection. This system also results in lower hospitalization expenses and better capacity as beds will be spared for the patients that require more intense care.

Since there are various patients and operators in the HHC system, the key problem here is to determine the most efficient way to utilize each operator based on their availabilities and patient requirements. Such a problem includes capacity planning, assignment of patients to operators, constructing schedules for operators, and routing of patient visits. In this thesis, the main focus is the assignment, scheduling, and routing steps of the aforementioned resource planning, and the problem handling these steps in an integrated manner will be referred to as the Home Health Care Problem (HHCP). In this regard, the integrated solution approach of HHCP deals with the assignment, scheduling, and routing problems simultaneously.

In HHCP, there is a set of patients who need to be visited and a set of operators who will serve those patients within a specified planning horizon, e.g., a week. The objective of HHCP can be evaluated under three decisions: assigning each patient to an operator considering their availability, scheduling the patient visits along the planning horizon, and constructing routes the operators will use to visit the patients on each day of the planning horizon considering the time windows of the patients. All of these decisions are considered under the assumption of full continuity of care, which indicates that the patient-operator assignments will remain unchanged during the planning horizon. HHCP is classified as a NP-Hard problem. Therefore, obtaining its optimal solution in a reasonable time may not always be possible, especially for real-life problems. As more criteria are included in the problem such as time windows of the patients, the workload capacity of operators, and continuity of care, reaching a feasible solution within a reasonable time becomes increasingly challenging. Thus, in the literature, two-phase methods are preferred over the integrated solution approach [4–7]. In the literature, the two-phase methods are developed by decomposing the integrated model and solving the assignment, scheduling, and/or routing problems at different levels [6, 8]. One of those variations is solving the assignment-scheduling problem (the problem where the assignment and scheduling problems are solved simultaneously will be called the assignment-scheduling problem from now on) first and then solving the routing problem with the obtained assignment and scheduling information. However, this strategy may lead to infeasibility in the routing phase as the visiting sequence and the time window information are not available when solving the assignment-scheduling problem. Similarly, due to the lack of routing information in the first phase, low-quality solutions regarding operator routes can be observed. In order to deal with possible routing issues and low-quality solutions, data-driven route duration estimation methods are used in the literature [7].

The primary purpose of the data-driven route duration estimation is to evaluate the routes each operator must use to visit their patients by utilizing historical data of past routes. In most cases, the operators may choose to follow a visiting sequence based on the characteristics of patients and their locations. The routes each operator must follow to visit their patients impact the assignment or scheduling decisions since the visiting sequences may be different based on the patient characteristics. In this thesis, only the geographical locations of the patients are considered in the route duration estimation. Through the historical data, the routing choices of the operators can be captured and used to obtain high-quality solutions by incorporating them into the assignment and scheduling decisions. Therefore, the advantage of using the two-phase solution approach rather than the integrated solution approach lies in the incorporation of data-driven route duration estimation which helps achieving high-quality solutions benefiting from the historical routing data.

In this thesis, a data-driven estimation method is developed using a neural network (NN) model which is a machine learning process operated by a collection of neurons. In order to make realistic estimations regarding the routing preferences of the operators, the developed method utilizes geographical locations of the patients visited in past routes. Thus, the data-driven estimation method provides more insight regarding the newly established routes by comparing them to past routes based on the locations of the patients. Additionally, the developed method is compared to a previously developed data-driven estimation method that uses Kernel Regression (KR) [7]. Two methods are compared based on their performances on estimating the route durations with geographical locations belonging to different regions than the ones used in historical data. Based on the results, the NN model is employed as the estimation method.

Moreover, a matheuristic framework is proposed for the solution of the HHCP by improving the initial solution obtained from the two-phase solution approach, in which the assignment-scheduling problem and routing problem are solved consecutively. Matheuristic algorithms combine metaheuristics and mathematical modeling techniques to provide a feasible solution. The proposed method comprises a Variable Neighborhood Search (VNS) for exploring the solution space, and it also modifies the mathematical model of the assignment-scheduling problem by adding constraints to prevent assigning specific patients at the same time to an operator if they cause an infeasibility on a route. Additionally, the framework uses the NN model estimations for evaluating the route durations to overcome the infeasibility resulting from the assignment-scheduling problem and obtain high-quality solutions by capturing the historical routing preferences of operators.

1.1. OBJECTIVE OF THE THESIS

In this thesis, the main objective is to enhance the solution obtained from the assignment-scheduling problem by targeting potential infeasibilities that may arise in the routing stage. This is accomplished through the use of a matheuristic framework that employs the VNS metaheuristic and a data-driven approach for estimating the route durations of operators. It is intended to offer high-quality and feasible routes that take the time windows of the patients into account as an output of the designed matheuristic framework. The developed solution method works by investigating the obtained solution iteratively to target the potential routing issues that are caused by the lack of routing information in solving the assignment-scheduling problem. Thus, the presented solution approach is aimed to produce feasible solutions for HHCP that incorporate time window restrictions and satisfy the continuity of care assumption for a time horizon of one week and a set of operators with identical skills.

Hence, the objectives of this thesis can be listed as follows:

- Solving the HHCP using two-phase solution approach with data-driven method
    - Selection of the best data-driven estimation method (i.e., neural network model)
- Developing a matheuristic framework with a feedback mechanism to improve the solution quality

1.2. OUTLINE OF THE THESIS

The organization of this thesis is given as follows: Chapter 2 includes a thorough literature review on HHCP. In Chapter 3, the HHCP is presented. Next, Chapter 4 is dedicated to the in-depth presentation of the data driven matheuristic framework developed for the HHCP. Chapter 5 presents the computational experiments. Lastly, the conclusions are drawn regarding the outcomes of this thesis in Chapter 6.

2. LITERATURE REVIEW

The literature shows that the interest in HHCP has grown in recent years [9–12], and the key characteristics of different studies regarding HHC are identified by their modeling approaches and the duration of the planning period. Different modeling approaches consist of integrated and two-phase modeling approaches, whereas the planning period can be short or long-term. These studies can also be further classified under different constraints such as continuity of care, time windows, and maximum operator workload. Additionally, various aspects of objectives are considered [11] such as lowering the operational costs, minimizing the traveling time between patients, maximizing patient and operator satisfaction, and minimizing the operator workload. In this section, a detailed review of the literature regarding HHC services is provided. The studies are mainly investigated based on modeling approaches and the planning period, and the details concerning other constraints are also explained.

As an example of a recent study focuses on the integrated solution approach with a short planning period, Liu et al. [13] model HHCP in a time horizon of one day while considering time windows, operator skills, and workload with an objective of minimizing the patients that can’t be visited. The developed model is solved using the branch-and-price (B&P) algorithm for up to 100 patients and 12 operators. In another study conducted by Frifita and Masmoudi [14], a daily model with an objective of minimizing the travel cost is proposed. The proposed model includes constraints on time windows and operator skills and is solved by three different Variable Neighborhood Search (VNS) metaheuristic algorithms for at most 40 operators and 200 patients. Nasir and Kuo [15] model HHCP for a daily planning period under the constraints of patients’ time windows, operator skills, and maximum workload, and presented a hybrid Genetic Algorithm (GA) metaheuristic algorithm in order to solve the model. The proposed method is able to produce feasible solutions for a dataset consisting of up to 12 operators and 60 patients. In the study of Lin et al. [16], multiple objectives such as minimizing the total workload, minimizing the total travel time, and balancing the workload of operators are considered along with constraints on time windows, operator skills, and maximum workload. The suggested model is then solved with B&P algorithm for at most seven operators, 13 patients. In the most recent study involving an integrated modeling approach and a short planning period, Akbari et al. [17] propose an integer programming model and a General Variable Neighborhood Search (GVNS) metaheuristic algorithm. The model incorporates time windows of patients and aims to minimize the waiting time of patients until receiving a service while prioritizing the patients in relation to their severity. Large problem instances which contain up to 500 patients are used to test the proposed model.

Integrated models that address a long-term planning period (e.g., one week) are recently favored in order to provide a solution for more realistic data. For example, Nickel et al. [18] use an integrated modeling approach to solve problems with a long-term planning period while minimizing travel and labor costs as well as maximizing the number of patients served. They also consider time windows and operator skills as additional constraints and used Tabu Search (TS) and Large Neighborhood Search (LNS) to solve the model for up to 12 operators and 361 patients. In a similar study conducted by Trautsamwieser and Hirsch [19], the same constraints are imposed while aiming to minimize the total workload. Therefore, an algorithm composed of B&P and VNS is presented and it is tested on a large dataset with nine operators and 203 patients. Similarly, Grenouilleau et al. [20] propose a model with an objective of minimizing the inconsistency in the continuity of care under the same constraints but by using the set partitioning heuristic for at most 20 operators and 150 patients. An example that consists of multiple objectives including minimizing total travel time and the deviation from the patients' designated time windows is the study proposed by Mosquera et al. [21]. The model is solved using a two-staged heuristic framework with the largest dataset containing 28 operators and 127 patients. Another comprehensive research by Grenouilleau et al. [22] takes into account the continuity of care, time windows, operator skills, and the maximum workload as constraints and aims to maximize the number of new patients accepted in the system. In this study, a matheuristic framework is presented and tested on a small dataset with six operators and 60 patients. Gong et al. [23] minimize traveling and overtime labor costs together with the penalty incurred by not meeting patients' preferences and not ensuring continuity of care. The proposed method is built under multiple operator skills and maximum workload constraints and solved using a matheuristic methodology. The presented methodology is tested on datasets consisting of up to 100 operators and 40 patients. In another related work proposed by Çınar et al. [24] consider the time window and maximum workload constraints; however, operator skills and continuity of care are not incorporated into the model. Furthermore, the objective is to maximize the prizes acquired from the visited patients, and the prizes are updated in a dynamic structure throughout the planning horizon. The matheuristic and heuristic algorithms are developed and tested on sample datasets with 180 patients. Finally, Bhattarai et al. [25] suggest an integrated model to tackle an HHC problem with the objectives of balancing operators' workloads, ensuring continuity of care, and meeting patients' time windows along with maximizing the revenue obtained by patient visits. They also define the operator satisfaction levels as a function based on the workload and incorporate them into the model. Instances with at most four operators and 20 patients are used to assess the performance of the proposed model.

There are also studies conducted that utilize a two-phase approach for HHCP with a short-term planning period (e.g., one day). For instance, Yalçındağ et al. [7] focus on daily assignment and routing problems with maximum workload restriction. In this context, they solve both problems consecutively with an objective of balancing the workloads and minimizing the travel time. They also concentrate on estimating travel times using Kernel Regression (KR) which is a data-driven method since the targeted assignment problem is handled without any information regarding routes, so they suggest an estimation method based on Kernel Regression (KR) technique. The suggested method is utilized to predict each operator's travel time by considering the geographic locations of previously visited patients. They employ a Genetic Algorithm (GA) based solution method in order to obtain solutions for the assignment and routing problem, and this method is tested on instances up to 15 operators and 150 patients. Moreover, the study of Chaieb et al. [4] suggests a decomposition-based model that handles patient clustering, operator assignments to clusters, and route construction for operators. The model takes into account a variety of objectives, including minimizing operator and travel costs, balancing operator workloads, and visiting the maximum number of patients, and is solved using several heuristic methods. The largest dataset used to test the model includes 28 operators and 154 jobs.

Aside from the ones cited above, some of the studies employing the two-phase modeling approach take into account a long-term planning period. As an example, Yalçındağ et al. [6] develop decomposition-based solution approaches for assignment, scheduling, and routing problems while aiming to balance workload and minimizing the traveling costs. Continuity of care, operator skills, and maximum workload constraints are all included in all of the provided solution approaches. The largest dataset which is used in testing consists of 16 operators and 300 patients. As it was previously stated, the routing information is not available while solving the assignment and scheduling problems; thus, routing constraints need to be included in the model using estimation methods. Regarding this, Yalçındağ and Matta [26] suggest a probabilistic estimation method in order to incorporate the time windows in the assignment model; however, the presented method is only tested on small-sized problem instances with two operators of the same skill and 15 patients. Different from these studies, Gomez and Ramos [5] emphasize dynamic planning and use a weekly rolling horizon. Although the continuity of care assumption is satisfied on a weekly basis, Gomez and Ramos allow the patients to be assigned to different operators between weeks within the scope of dynamic planning. The model they built minimizes the travel time and the changes in operator-patient assignments while considering workload balancing. Similar to the decomposition approach, the assignment problem and the scheduling and routing problems are solved consecutively. Scheduling and routing problems are modeled based on teams instead of individual operators, whereas the assignment problem decides which operators are included in each team. The routing and scheduling problem is solved separately for each team, assuming each team member is of the same skill while considering the time windows and maximum workload constraints. The solution method is tested on several problem instances, the largest of which consists of 12 operators and 60 patients.

This thesis proposes a matheuristic framework for the solution of an HHCP with a one-week planning period (i.e., long-term) and single-skilled operators. The addressed problem ensures continuity of care and incorporates maximum workload and time window constraints. For the developed framework, an initial solution is obtained using the two-phase approach where the assignment, scheduling, and routing problems are solved on different levels, and a data-driven machine learning method is developed to estimate route durations in order to identify the possible routing issues. The initial solution is systematically updated using a VNS metaheuristic algorithm. The structure of the VNS algorithm evaluates the solution in terms of the maximum operator workload and workload balance and aims to improve the solution in both aspects.

In obtaining the initial solution with the two-phase approach, the assignment-scheduling problem and the routing problem are solved consecutively as proposed by Yalçındağ et al. [6], and a data-driven route duration estimation method is presented similar to the study of Yalçındağ et al. [7] considering geographical locations of patients. However, the former study lacks the incorporation of time windows, whereas the latter only solves the HHCP with short-term planning period and maximum workload constraint. Therefore, this thesis extends those studies by developing a matheuristic framework to solve a HHCP with long-term period that considers time windows of patients and maximum workload restriction for operators. In addition to that, a data-driven route duration estimation method with higher accuracy is developed and incorporated into the matheuristic framework that is used to solve the HHCP.

In Table 2.1, comparable studies on HHCP are presented, and the contribution of this thesis is highlighted.

Table 2.1. A summary of studies that are comparable to this thesis. (I: Integrated approach, D: Decomposition approach, ST: Short-term, LT: Long-term, S: Soft, P: Partial, B&P: Branch-and-price, MILP: Mixed integer linear programming)

	Modeling Approach	**Time Horizon**	**Continuity of Care**	**Time Window**	**Maximum Workload**	**Data-Driven**	**Solution Strategy**
Liu et al. [27]	I	ST		✓	✓		B&P
Frifita and Masmoudi [14]	I	ST		✓			Heuristic
Nasir and Kuo [15]	I	ST		✓	✓		Metaheuristic
Lin et al. [16]	I	ST		✓	✓		B&P
Tanoumand and Ünlüyurt [28]	I	LT		✓	✓		B&P
Nickel et al. [18]	I	LT		✓	✓		Heuristic
Trautsamwieser and Hirsch [19]	I	LT		✓	✓		B&P, Heuristic
Grenouilleau et al. [20]	I	LT	S	✓	✓		Heuristic
Grenouilleau et al. [22]	I	LT	✓	✓	✓		Matheuristic
Mosquera et al. [21]	I	LT		✓			Heuristic
Gong et al. [23]	I	LT	S		✓		Matheuristic
Çınar et al. [24]	I	LT		✓	✓		Matheuristic, Heuristic
Bhattarai et al. [25]	I	LT	✓	✓	✓		MILP
Akbari et al. [17]	I	ST					Metaheuristic
Yalçındağ et al. [7]	D	ST			✓	✓	Metaheuristic
Chaieb et al. [4]	D	ST		✓	✓		Metaheuristic
Yalçındağ et al. [6]	D	LT	✓		✓		MILP
Yalçındağ and Matta [26]	D	LT	✓		✓		MILP
Gomes and Ramos [29]	D	LT	P	✓	✓		MILP, Heuristic
This thesis	D	LT	✓	✓	✓	✓	Matheuristic

3. PROBLEM DEFINITION

In this thesis, a framework for solving the Home Health Care Problem (HHCP) is presented. In the addressed problem, there is a set of patients who require operators to visit them at home in order to receive health care services. Patients are assigned to operators taking operators availability and capacity into consideration as well as the balance between their workloads, and the sequence of the visits is determined for each day of the planning horizon. In addition to that, there is a certain time window for each patient, and this constraint is also taken into account in assignment decisions.

In this context, the problem is defined by a completed direct network $G = (N, A)$ with n nodes belonging to the set $N = \{1, \dots, n\}$ where each node j represents a patient. An additional node is also presumed to be included in the set to define the node which the operators visit at the start and end of their daily tour. The set A is defined for all possible paths between nodes in the network, and there exists an arc $(j, j') \in A$ with distance denoted as $t_{jj'}$ corresponding to each node pair (j, j') in set N. The set O is defined for the operators and the subset $O_d \subseteq O$ consists of operators that are available on day d throughout the planning horizon denoted with H where $d \in H = \{1, \dots, H\}$. The requirement of patients is denoted by the vector r_j which represents the number of visits required by the patient j during the time horizon. In a time horizon of one week, this vector indicates the total number of days each patient has to be visited, and in the HHC literature, this is called the care plan. The care plan of each patient is determined by using patterns. In this context, the patterns offer a combination of one or more days of visits to form a care plan for the patients. In order to demonstrate, let us assume that a patient requires two visits from any operator in a week. Some of the candidate patterns suitable for that patient could be the patterns with visits scheduled on Tuesday-Wednesday or Wednesday-Friday, or Monday-Thursday. Each pattern p is defined in a set P and it is defined as $p(d) = 0$ if a pattern p suggests that no visit will be operated on day d. It is also assumed that only a single visit can take place on a given day. In addition to that, there is a time window during which each patient can be visited each day. Regarding the time windows, EST_{dj} and LFT_{dj} denote the earliest start time and latest finish time of service for patient j on day d respectively. Lastly, there is a specific duration of care for each patient j denoted as s_j, and a daily capacity for each operator i that is defined by a_i.

With all the information provided regarding the HHCP, there are three major decisions to make in solving the problem. First, each patient must be assigned to an operator which will remain the same throughout the time horizon as the continuity of care constraint implies. Second, a pattern must be assigned to each patient that is consistent with their care plans. Consequently, routes for each operator must be formed for each day a visit is scheduled. In making these decisions, the capacity of each operator must be considered as well. These three decisions correspond to the assignment, scheduling, and routing decisions respectively. Hence, the HHCP is defined as a problem which handles assignment, scheduling, and routing problems simultaneously in an integrated manner, as illustrated in Figure 3.1.

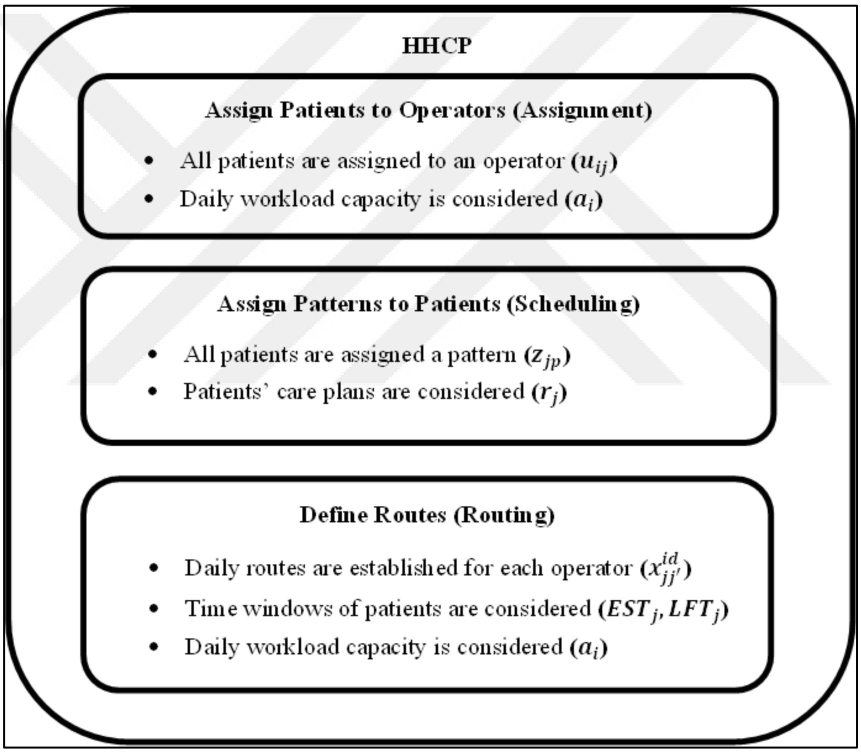

Figure 3.1. Schema of assignment, scheduling, and routing problems for HHCP

The addressed problem is reduced to the widely recognized Traveling Salesman Problem (TSP) if modeled for a single day for a single operator. Since TSP is of NP-Hard complexity [3], HHCP is also of NP-Hard complexity, therefore real-life sized problems cannot be solved within a reasonable time. Thus, the integrated problem is often solved by decomposing it into two levels or by using heuristic algorithms. As summarized in Figure 3.2, there are several decomposition approaches that can be used [6]. In the first approach shown on the left most side of the figure, the assignment problem is solved first and then the assignment information obtained is used in solving the scheduling and routing problems simultaneously. The second approach displayed in the figure is when the assignment and scheduling problems are solved simultaneously, and the routing problem is handled in the second phase with the obtained assignment and scheduling information. A third option is to incorporate the scheduling problem into both phases, which may result in better coordination between two phases. In this procedure, first, the assignment and scheduling problems are handled together. Later in the second phase, the scheduling and routing problems are solved by fixing the values of the assignment variables that were previously obtained in the first phase. Finally, the integrated solution approach where each of the assignment, scheduling and routing problems are handled simultaneously is shown on the right most side of the figure. Although the integrated method yields the best solution in terms of quality and feasibility, the integrated model is of NP-Hard complexity, hence it cannot be used to solve realistic problems.

Among the described two-phase solution methodologies, the first method with the assignment problem in the first level and the scheduling and routing problems on the second level did not achieve any feasible solutions with any of the problem instances in the experiments conducted by Yalçındağ et al. [6]. Furthermore, it is observed in the referred study that the second and third decomposition approaches provide near optimal and good quality solutions in small, medium, and large problem instances. Between the two methodologies, it is seen that incorporating the scheduling decisions in both levels (the third option explained above) resulted in higher quality solutions compared to the alternative in which the scheduling decisions are only made in the first level (the second one). However, the computational time required to solve the third decomposition method is higher than that of the second method. Since each one of three main decisions are made without considering one another in the two-phase solution approaches, these approaches may generate infeasibility in the second level or low-quality solutions, such as having inconsistent workloads between the operators. Therefore, in order to address these issues, a matheuristic framework with a feedback mechanism is developed. The second decomposition approach is employed within the matheuristic framework as it requires lower computational time, and it shows no significant difference in solution quality compared to the third alternative. Thus, an initial solution is obtained with the selected decomposition approach, and the possible infeasibility arising from the two-phase solution is treated inside the framework.

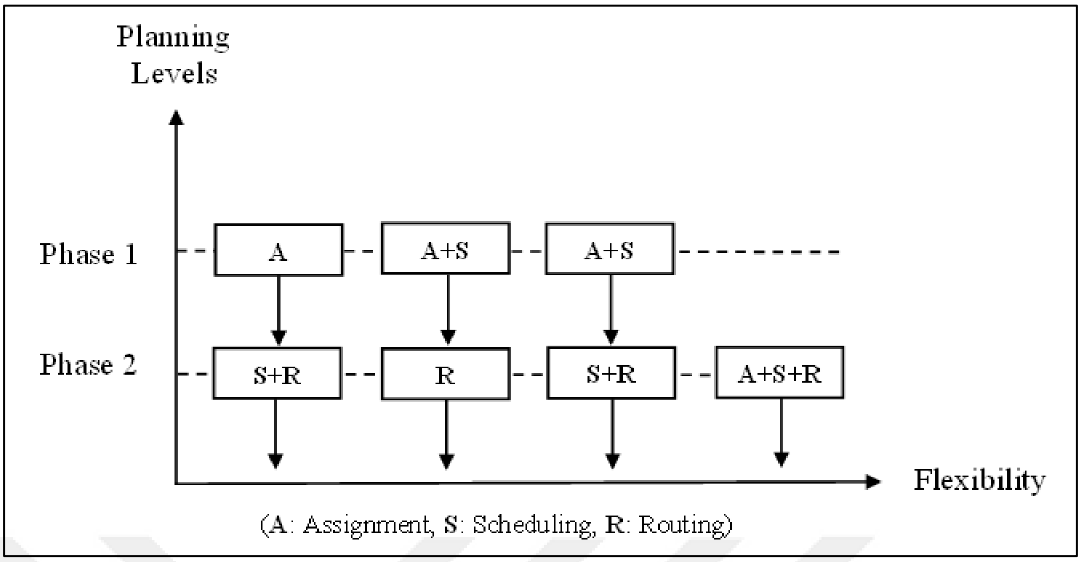

Figure 3.2. Different decomposition approaches (adapted from [6])

As mentioned earlier, the decomposition approach adopted in this thesis involves solving the assignment and scheduling problems synchronously in the first level, and the routing problem in the second level. The problem in which the assignment and scheduling decisions are made together is called the assignment-scheduling problem (A+S). The rest of the chapter is dedicated to detailed explanation of the two-phase and integrated solution strategies. Below is given the necessary information regarding the notation used in the mathematical models.

Sets:

A	:	all possible paths, $(j, j') \in A$
H	:	planning horizon, $d \in H$
N	:	set of patients, $j \in N$
O	:	set of operators, $i \in O$
O_d	:	set of operators available on day $d \in H$
P	:	set of patterns, $p \in P$

Parameters:

a_i	:	daily capacity of operator $i \in O$

EST_j : earliest start time of service for patient $j \in N$

EST_{dj} : earliest start time of service for patient $j \in N$ on day $d \in H$

LFT_j : latest finish time of service for patient $j \in N$

LFT_{dj} : latest finish time of service for patient $j \in N$ on day $d \in H$

r_j : care plan of patient $j \in N$

s_j : service time of patient $j \in N$

$t_{jj'}$: Distance between patients j and j', $(j, j') \in A$

τ_j : average travel time to reach patient $j \in N$

Decision Variables:

$$u_{ij} = \begin{cases} 1, & \text{if operator } i \in O \text{ is assigned to patient } j \in N\backslash\{0\} \\ 0, & \text{otherwise} \end{cases}$$

$$u_{ij}^{d} = \begin{cases} 1, & \text{if operator } i \in O \text{ visits patient } j \in N\backslash\{0\} \text{ on day } d \in H \\ 0, & \text{otherwise} \end{cases}$$

$$x_{jj'}^{id} = \begin{cases} 1, & \text{if operator } i \in O_d \text{ travels on arc } (j, j') \in A \text{ on day } d \in H \\ 0, & \text{otherwise} \end{cases}$$

$$x_{jj'} = \begin{cases} 1, & \text{if the operator travels on arc } (j, j') \in A \\ 0, & \text{otherwise} \end{cases}$$

y_{ijd} = start time of visit by operator $i \in O_d$ to patient $j \in N\backslash\{0\}$ on day $d \in H$

y_j = start time of visit to patient $j \in N\backslash\{0\}$

$$z_{jp} = \begin{cases} 1, & \text{if pattern } p \in P \text{ is assigned to patient } j \in N\backslash\{0\} \\ 0, & \text{otherwise} \end{cases}$$

3.1. TWO-PHASE APPROACH

In the two-phase approach, the assignment-scheduling problem is solved first, and it is followed by solving the routing problem in which the information obtained from the assignment-scheduling problem is incorporated. In this section, the necessary mathematical models regarding the two-phase strategy are presented.

3.1.1. Mathematical Model of The Assignment-Scheduling Problem (A+S)

In the assignment-scheduling problem, operators are assigned to patients and the visiting schedules are determined through the assignment of patterns to each patient. These decisions are made considering the workload of the operators. The workload can be defined as the total time spent by the operator traveling between the patients and serving each patient, and there exists a daily workload capacity for each operator. Since the travel times between patients are not yet available at this stage, an estimation regarding the route durations is made, and this estimation is utilized in computation of the workload for each operator. Such an estimation can be made by simply calculating the average time to reach each patient from all other nodes including the depot [6] as displayed in Equation (3.1) below.

$$\tau_j = \frac{1}{n} \sum_{j \neq j'} t_{jj'} \tag{3.1}$$

Following is the mathematical model for the assignment-scheduling problem.

$$\text{Min. } m_u \tag{3.2}$$

$$\text{s.t. } \sum_{i \in O} u_{ij} = 1 \qquad \forall j \in N \tag{3.3}$$

$$\sum_{p \in P} z_{jp} = 1 \qquad \forall j \in N \tag{3.4}$$

$$D_{id} = \sum_{j \in N} (s_j + \tau_j).u_{ij}^d \leq a_i \qquad \forall d \in H, \forall i \in O_d \tag{3.5}$$

$$\sum_{j \in N} \sum_{i \in O} x_{jj'}^{id} \leq \sum_{p \in P: p(d)=1} z_{j'p} \qquad \forall j' \in N \backslash \{0\}, \forall d \in H \tag{3.6}$$

$$\sum_{j \in N} \sum_{i \in J} x_{jj'}^{id} \geq \sum_{p \in P} z_{j'p} \qquad \forall j' \in N \backslash \{0\}, \forall d \in H \tag{3.7}$$

$$x_{jj'}^{id} \leq u_{ij} \qquad \forall (j, j') \in A, \forall d \in H, \forall i \in O_d \tag{3.8}$$

$$\frac{\sum_{d\in H} D_{id}}{a_i.H} \leq m_u \qquad \forall i \in O \qquad (3.9)$$

$$u_{ij} \in \{0,1\} \qquad \forall j \in N\backslash\{0\}, \forall i \in O \qquad (3.10)$$

$$u_{ij}^{d} \in \{0,1\} \qquad \forall (j,j') \in A, \forall d \in H, \forall i \in O_d \qquad (3.11)$$

$$z_{jp} \in \{0,1\} \qquad \forall j \in N\backslash\{0\}, p \in P \qquad (3.12)$$

In the presented model, the objective function (3.2) aims to minimize the workload of the operator with the maximum workload (m_u) calculated in (3.9). Constraints (3.3) ensure that each patient is assigned to exactly one operator. Similarly, constraints (3.4) guarantee that each patient is assigned exactly one pattern. Constraint set (3.5) computes the daily workload of each operator utilizing the average travel time estimation provided in (3.1) and ensures that the daily workload capacity of the operators is not exceeded. Constraints (3.6) and (3.7) connect the assignment and scheduling variables. In detail, constraints (3.6) ensure that a patient can only be visited by a particular operator on a certain day if a visit is scheduled for that patient on that day. Constraints (3.7) state that an operator must visit a patient on a particular day if a visit has been scheduled for that day. Constraints (3.8) ensure that in order for an operator to visit a patient, they must be assigned to each other. Lastly, constraints (3.10-3.12) state the domain restrictions.

3.1.2. Mathematical Model of The Routing Problem (R)

Upon solving the assignment-scheduling problem, the visits each operator must make every day throughout the planning horizon are determined. This information is used in establishing routes that each operator must follow for each scheduled patient visit. Therefore, a Traveling Salesman Problem with Time Windows (TSPTW) is solved at the routing phase for each operator and for each day on the planning horizon.

$$\text{Min.} \quad \sum_{\forall j\in N} \sum_{\forall j'\in N} t_{jj'} \cdot x_{jj'} \qquad (3.13)$$

$$\text{s.t.} \quad y_j - t_{0j} \cdot x_{0j} \geq 0 \qquad \forall j \in N\backslash\{0\} \qquad (3.14)$$

$$y_j \geq EST_j \qquad \forall j \in N\backslash\{0\} \qquad (3.15)$$

$$y_j - y_{j'} + \left(LFT_j - EST_{j'} + t_{jj'}\right) \cdot x_{jj'} \leq LFT_j - EST_{j'} \qquad \forall j \in N\backslash\{0\}, \forall j' \in N\backslash\{0\} \qquad (3.16)$$

$$\sum_{j \in N} x_{jj'} = 1 \qquad \forall j' \in N\backslash\{0\} \qquad (3.17)$$

$$\sum_{j' \in N\backslash\{0\}} x_{jj'} = 1 \qquad \forall j \in N\backslash\{0\} \qquad (3.18)$$

$$y_j + s_j \leq LFT_j \qquad \forall j \in N\backslash\{0\} \qquad (3.19)$$

$$y_j + t_{j0} \leq y_0 \qquad \forall j \in N\backslash\{0\} \qquad (3.20)$$

$$y_j \geq 0 \qquad \forall j \in N \qquad (3.21)$$

$$x_{jj'} \in \{0,1\} \qquad \forall (j,j') \in A \qquad (3.22)$$

The objective function (3.13) minimizes the total distance traveled. Constraints (3.14) compute the time at which the patient is reached once the operator leaves the depot. Constraints (3.15) ensure that the service does not start before the earliest start time. Constraints (3.16) is the subtour elimination constraint. Constraint sets (3.17) and (3.18) guarantee that each patient is visited exactly once. Constraints (3.19) guarantee that the service must be completed before the latest finish time. Constraints (3.20) compute the finish time of the tour. Lastly, (3.21) and (3.22) are non-negativity and domain restriction constraints respectively.

3.2. INTEGRATED APPROACH

In solving the integrated model, all decisions regarding the assignment of patients to operators, scheduling the patient visits through pattern assignments and establishing routes for patient visits are made concurrently.

3.2.1. Mathematical Model of The Integrated Problem

Min. $$m_u \qquad (3.23)$$

s.t. $$\sum_{i \in O} u_{ij} = 1 \qquad \forall j \in N \qquad (3.24)$$

$$\sum_{p \in P} z_{jp} = 1 \qquad \forall j \in N \tag{3.25}$$

$$\sum_{j \in N} x_{jj'}^{id} = \sum_{j \in N} x_{j'j}^{id} \qquad \forall j' \in N \backslash \{0\}, \forall d \in H, \forall i \in O_d \tag{3.26}$$

$$D_{id} = \sum_{(j,j') \in A} \left(s_j + t_{jj'}\right) \cdot x_{jj'}^{id} \leq a_i \qquad \forall d \in H, \forall i \in O_d \tag{3.27}$$

$$y_{ijd} + s_j + t_{jj'} - M \cdot \left(1 - x_{jj'}^{id}\right) \leq y_{ij'd} \qquad \forall j \in N, \forall j' \in N \backslash \{0\}, \forall d \in H, \forall i \in O_d \tag{3.28}$$

$$EST_{jd} - M \cdot \left(1 - u_{ij}\right) \leq y_{ijd} \qquad \forall j \in N \backslash \{0\}, \forall d \in H, \forall i \in O_d \tag{3.29}$$

$$y_{ijd} + s_j - M \cdot \left(1 - u_{ij}\right) \geq LFT_{jd} \qquad \forall j \in N \backslash \{0\}, \forall d \in H, \forall i \in O_d \tag{3.30}$$

$$\sum_{j \in N} \sum_{i \in O} x_{jj'}^{id} \leq \sum_{p \in P : p(d) = 1} z_{j'p} \qquad \forall j' \in N \backslash \{0\}, \forall d \in H \tag{3.31}$$

$$\sum_{j \in N} \sum_{i \in J} x_{jj'}^{id} \geq \sum_{p \in P} z_{j'p} \qquad \forall j' \in N \backslash \{0\}, \forall d \in H \tag{3.32}$$

$$x_{jj'}^{id} \leq u_{ij} \qquad \forall (j, j') \in A, \forall d \in H, \forall i \in O_d \tag{3.33}$$

$$u_{ij'} \leq \sum_{j \in N} \sum_{d \in H} x_{jj'}^{id} \qquad \forall j' \in N \backslash \{0\}, \forall i \in O \tag{3.34}$$

$$\frac{\sum_{d \in H} D_{id}}{a_i \cdot H} \leq m_u \qquad \forall i \in O \tag{3.35}$$

$$u_{ij} \in \{0,1\} \qquad \forall j \in N \backslash \{0\}, \forall i \in O \tag{3.36}$$

$$x_{jj'}^{id} \in \{0,1\} \qquad \forall (j, j') \in A, \forall d \in H, \forall i \in O_d \tag{3.37}$$

$$y_{ijd} \geq 0 \qquad \forall j \in N, \forall d \in H, \forall i \in O_d \tag{3.38}$$

$$z_{jp} \in \{0,1\} \qquad \forall j \in N \backslash \{0\}, p \in P \tag{3.39}$$

In the provided model, the objective function (3.23) minimizes the utilization rate of the operator who has the maximum utilization rate (m_u) among the operators in the resulting assignments. With this objective function, it is aimed to balance the workload of the operators. Constraint set (3.24) enables that each patient is assigned to exactly one operator over the time horizon. It is crucial for continuity of care that the operator-patient assignment remains the same throughout the planning horizon. Constraints (3.25) ensure that each patient is assigned to a pattern. Constraints (3.26) are the flow conservation constraints regarding the routing variables. Constraint set (3.27) computes the workload of each operator (D_{id}) based on the total time spent on the route and the time spent serving the patient on the specified day. These constraints also ensure that the daily capacity of the operators is not exceeded. (3.28) are subtour elimination constraints that also compute the start time of service on patients for each day. It should also be noted for those constraints that the values of y_{0d} and s_0 are assumed zero. Constraint sets (3.29) and (3.30) ensure that each patient is visited during their designated time windows where (3.31) and (3.32) connect the scheduling and routing variables. Constraints (3.31) impose that a patient must be scheduled a suitable pattern if they are to be visited on a particular day, whereas constraints (3.32) guarantee that a patient is visited on that day by exactly one operator if they are scheduled a visit for that day. Constraint sets (3.33) and (3.34) link the assignment and routing variables. Constraints (3.33) ensure that an operator can visit a patient only if they are assigned to each other. (3.34) prevents an operator from being assigned to a patient if the operator never visits that patient over the planning horizon. Constraints (3.35) compute the workload of the operator with the maximum workload (m_u) among all operators. Finally, (3.36) - (3.39) are domain restriction constraints for variables and M is a very large number.

4. DATA DRIVEN MATHEURISTIC FRAMEWORK FOR THE HOME HEALTH CARE PROBLEM

As it was previously stated in the earlier chapters, we attempt to solve the integrated problem of HHCP by following the two-stage solution approach due to its NP-Hard complexity. The two-stage solution approach consists of the assignment-scheduling problem and the routing problem which are solved consecutively. However, the assignment-scheduling problem does not include any decision variables for routing and takes average travel times into consideration for route duration estimation. Hence, the time windows of the patients are neglected. As a result, this method may often yield infeasible routes later in the routing stage. This chapter is dedicated to the matheuristic framework we developed to produce a feasible and high-quality solution by improving the solution of the two-stage method.

4.1. MATHEURISTIC FRAMEWORK

As displayed in Figure 4.1, the proposed matheuristic framework mainly consists of three stages.

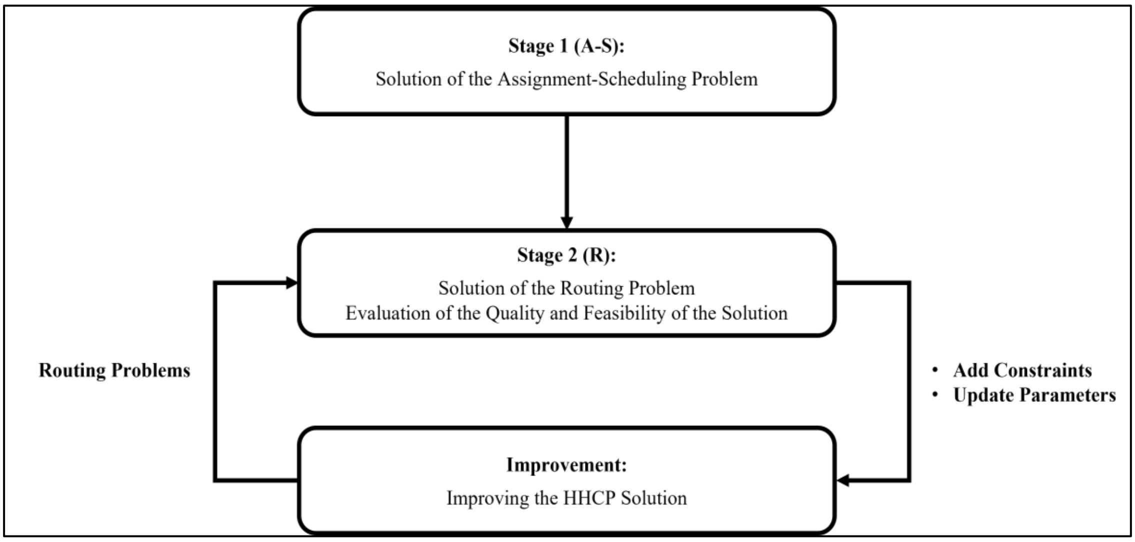

Figure 4.1. The proposed matheuristic framework for HHCP

First, the assignment-scheduling problem is solved, and the initial operator-patient and pattern-patient assignments are obtained. In solving the assignment-scheduling problem, the routing constraints are neglected and average travel times of all patients to be visited in a route are used to estimate the route duration parameter existing in the mathematical model.

In the second stage, the routing problem is solved in order to establish the routes that each operator must utilize on each of the days that they are available. In solving the routing problem, the assignment and scheduling decisions obtained from the previous stage are used. At this stage, the solution may turn out to be infeasible due to negligence of routing constraints in the previous stage. If the infeasibility is due to violation of time windows, it is essential for the upcoming improvement stage to detect the source of infeasibility. For this purpose, a set of deviational variables are incorporated into the routing problem. When a patient’s time window is violated, the mentioned variables take a non-zero value. For each route, the information of the patients for which the deviational variables take a non-zero value is stored and utilized in the algorithm in order to decide on the modification that will be made in the current solution during the improvement stage.

In the third stage, the initial solution found in the first stage is enhanced in an iterative manner by utilizing mathematical modeling techniques and Variable Neighborhood Search (VNS) metaheuristic. In the improvement stage, the solution space is meticulously explored with the goal of arriving at a feasible solution while avoiding low-quality solutions. This is achieved by iteratively updating the solution and solving the routing problem, therefore creating a feedback mechanism between the second and third stages. In the adopted VNS algorithm, the route duration must be estimated in order to assess if the change to be implemented has a favorable influence on the solution or not. For this purpose, a neural network model is employed in predicting the route durations. The incorporation of VNS and neural network model is further explained in the upcoming sections.

4.1.1. Machine Learning Method for Route Duration Estimation

Following the solution of the assignment-scheduling problem, two major problems can arise. One of them is having a large difference in the length of the routes each operator must follow. Another potential problem is failing to construct routes that satisfy the time window constraints of the patients. In both cases, the problem needs to be recognized in order to make the necessary adjustments to the solution. In addition to that, in some cases the operators’ routing preferences may change due to different geographical features, and the estimation method enables embedding the past preferences to the solution by utilizing historical routing data. Therefore, a route duration estimation method is developed to be used inside the metaheuristic algorithm within the matheuristic framework. According to this method, an estimation is obtained for each day an operator must operate along the planning horizon. In the literature, a time travel estimation method has been proposed by Yalçındağ et al. [7], based on the Kernel Regression (KR) technique, which is a non-parametric regression method that does not rely on a specific assumption regarding the distribution [30]. This method uses patients' geographical locations and the route durations observed in the historical routing data. Let $J = \{1, \dots, J\}$ be the set of patients an operator must visit on a particular day, where the $2J$-dimensional vector $x = \{x_i, i = 2(j-1) + l, j \in J, l = 1, 2\}$ denotes the geographical locations of patients in set J. The set $S_J = \{1, \dots, S\}$ contains the historical routes where J patients are visited where $y^0 = \{y_s^0, s \in S_J\}$ is a vector that holds the durations of historical routes whereas the vector $x^0 = \{x_{si}^0, i = 2(j-1) + l, j \in J, l = 1{,}2, s \in S_J\}$ with dimension $S \times 2J$ that contains the geographical locations of the patients visited in the historical data.

The KR technique computes the value of the random variable Y according to X, namely as $\mathbb{E}(Y|X)$, and the relationship between X and Y is defined as $Y = \tau(X) + \varepsilon$. In this context, τ denotes an unknown function where ε is defined for an independent and identically distributed error term with a mean of zero and variance of $\sigma^2(X)$. In this case, the multivariate KR method which estimates the value of variable Y conditionally on the value of the exogenous variable X also calculates the following conditional expectation value: $\mathbb{E}(Y|X) = \mathbb{E}\left(Y \middle| x_1, \dots, x_{2J}\right) = \tau(x)$. In this formulation, x represents the patients in set J that will be visited. Nadaraya-Watson estimator can be used to estimate the unknown function of τ which is defined in the following equation:

$$\hat{\tau}(x) = \frac{\sum_{s \in S_J} K\left(\frac{x_s^0 - x}{h}\right) y_s^0}{\sum_{s \in S_J} K\left(\frac{x_s^0 - x}{h}\right)} \tag{4.1}$$

In Equation (4.1), $K(.)$ denotes the $2J$-dimension kernel function, where h denotes the bandwidth vector, and finally $x_s^0 = (x_{s,1}^0, \dots, x_{s,2J}^0)$ represents the geographical location of each patient in the set J on route s. The Gaussian kernel function is one of the most popular Kernel functions, and it is described as follows:

$$K(z_s) = \prod_{i \in J} \frac{1}{\sqrt{2\pi}} e^{-z_{si}^2} \tag{4.2}$$

It should be noted that in the given equation, $z_{si} = \left(\frac{x_{si}^0 - x_i}{h_i}\right)$ and $J = \{1, \dots, 2P\}$.

It is observed that the KR method does not perform well if the patient coordinates to be used for estimation are far away from the coordinates on the past routes. Furthermore, the method can only be used if the number of patients on the routes provided in the historical data and the data for which the estimation is sought are the same. The proposed estimation method employs a neural network model, and it aims to improve the mentioned shortcomings of the KR method.

Neural networks (NN), also known as artificial neural networks (ANN), can be defined as a collection of algorithms that perform tasks involving training and optimization inspired by the human brain [31]. Each unit in the NN is called a neuron, and a set of neurons combined together forms layers. In a basic NN structure, there is an input and an output layer. Additionally, most of the structures include one or more hidden layers. Every neuron or unit in the network is connected to other neurons by weighted connections. Through the input layer, the information is fed into the network and the data is transmitted between the units all the way to the output layer where the final output is produced. The neurons, or units, are fired or activated according to an activation function. The weights in the NN are updated after each new data is fed through the network which enables it to "learn" patterns in the data. As a result of this well-designed structure, NN can explain complex functions and therefore are widely used in the machine learning field.

The NN model employed in route duration estimation is trained by utilizing patients' coordinates and historical routing data similar to the previously mentioned KR approach. Different from the KR method, the trained NN model can handle the difference in the locations of the patients and can work with routes in which there are different numbers of patients. Therefore, a model is trained once at the beginning of the matheuristic algorithm using a large set of historical data, and estimations are obtained using the trained model throughout the algorithm. The historical data used in training the NN model consist of information about the past routes of different lengths. In order to deal with varying route lengths, the maximum number of patients in a route needs to be determined first. Since the routes consist of patients an operator visits on a specific day, this decision is made according to the maximum number of patients that can be visited on any day in relation to the size of the problem. This decision defines the dimension of the historical data. Let us define the maximum number of patients in any route with n_{max}. Then, any row in the historical data with n_{max} patients on it would consist of $2(n_{max} + 1)$ columns where $2n_{max}$ columns correspond to the x and y coordinates of the patients on the route and 2 columns correspond to the x and y coordinates of the depot at which each route starts and ends. Moreover, any route with n patients where $n < n_{max}$, x and y coordinates of the depot are used to fill in the $2(n_{max} + 1) - 2n$ columns remaining from those used for the location information of the patients on the route. Finally, a set of historical data for all possible route lengths up to n_{max} patients is provided to the NN model in the form described in detail for training purposes.

4.1.2. Variable Neighborhood Search for The Home Health Care Problem

Recently, the Variable Neighborhood Search (VNS) metaheuristic has become prominent due to its efficacy, particularly in routing and scheduling problems [32–34]. Therefore, we decided to utilize a VNS algorithm within the proposed matheuristic framework.

In the VNS algorithm proposed by Hansen and Mladenovic [35], the search is performed over k_{max} predetermined neighborhoods within a set N_k, where $k = \{1,2, \dots, k_{max}\}$. The main rationale behind the algorithm is to improve the initial solution x by modifying it using these neighborhood structures. During the search for a better solution, the neighborhood that is used to generate other solutions is systematically changed. Below, the algorithm of the basic VNS structure is provided.

Algorithm 4.1. Basic VNS algorithm

```
Function VNS (x, k_max, t_max):
x = InitialSolution ()
t = CurrentTime ()
while t ≤ t_max:
    k = 1
    improvement = 0
    number_of_iterations = 0
    while k ≤ k_max:
        number_of_iterations = number_of_iterations + 1
        x' = LocalSearch (x, N_k)
        if x' is better than x:
            x = x'
            improvement = 1
        else:
            k = k + 1
        end
        if number_of_iterations = 10 and improvement = 0:
            x = Shake (x)
            k = 1
            improvement = 0
            number_of_iterations = 0
        end
    end
end
return x
```

Given a set of neighborhood structures, the basic VNS algorithm initiates by generating an initial solution x and iterates until a defined terminating condition (e.g., a time limit) is satisfied. Starting from the first neighborhood, a local search procedure is applied on solution x, and a local optimal solution x' is obtained. The new solution x' is evaluated in two aspects: the objective function value and the feasibility. If solution x is feasible, then it is replaced with x' when it yields a better objective function value. Otherwise, the replacement is made only if a feasible solution x' is obtained. In all remaining scenarios, the search continues from the next neighborhood. If no improvement is made in the current solution for 10 consecutive iterations, a random solution is generated in the shaking step using one of the shaking procedures and the current solution x is replaced with that random solution. Further along this section, the VNS algorithm is discussed in depth.

The algorithm works with a set of four neighborhoods and two kinds of shaking procedures. The VNS algorithm employs the following neighborhoods in the given order: patient group restriction, pattern change, patient swap, and operator change. The initial solution x is provided to the algorithm by solving the assignment-scheduling problem with a limited set of patterns, and the patterns employed in the solution may change throughout the algorithm.

The algorithm begins with a local search procedure, which is the backbone of the method. In the local search, the motive is to find a local optimal solution by performing diverse changes on the current solution. The following is a detailed explanation of how the mentioned neighborhood structures work.

I. Patient Group Restriction

This neighborhood is mainly dedicated to eliminating the infeasibility in the current solution by imposing constraints on a subset of patients who cannot be visited by the operators without violating time windows. Therefore, the first step of this neighborhood is to detect such patients. For this purpose, the routing problem containing a set of deviational variables is solved. In case of a time window violation, the deviational variables for the corresponding route take a non-zero value indicating that there is an infeasibility for the route. The neighborhood stores the groups of patients in infeasible routes in a list called the Restriction List (RL). This list is used to add the necessary restriction constraints whenever the assignment-scheduling model is solved throughout the algorithm. Depending on the result of the move attempted by the neighborhood, the list is updated by adding a new set of patients. Concisely, each element in the RL consists of a group of patients that can cause infeasibility on the route whenever they are visited all together.

The steps performed in this neighborhood are described in detail below:

- For each daily route, RL is created such that it contains patients on the route whose deviational variables have a non-zero value. If there is more than one such patient on a route, they are all added to RL. If there exists only one such patient, then all of the patients scheduled for that day are added to RL. Lastly, no patients are added to the list if there is no time window violation in the route.
- A set of constraints are added to the assignment-scheduling problem for each group of patients in RL. The following equation displays the extra constraints to be added.

$$\sum_{j \in NR_d} u_{ij}^d \leq |NR_d| - 1 \qquad \forall i \in O,\ \forall d \in H \qquad (4.3)$$

The set NR_d represents the group of patients in RL for day d in time horizon H where $NR_d \subseteq N$. The constraints aim to prevent all operators from visiting the particular set of patients on any day during the time horizon.

- The assignment-scheduling problem is solved by adding the new set of constraints while keeping the operator assignment variables fixed for all other patients on routes with no time window violation.
- If the assignment-scheduling problem returns a feasible solution, then the RL is updated to be used later throughout the algorithm. Otherwise, the algorithm proceeds with the next neighborhood after deleting the previously added patients from the RL.
- Consequently, the routing problem is solved to check whether the solution has improved in terms of the objective function value or feasibility. If it has improved, then the current solution is updated, and the algorithm continues with the same neighborhood until the solution cannot be further improved. Otherwise, the pattern change neighborhood is executed.

II. Pattern Change

This neighborhood aims to reduce the workload of the day with the highest workload by assigning different patterns to patients scheduled for that day. The alternative patterns are drawn from a predetermined set of patterns defined in the algorithm during the initialization step. To ensure a fair comparison with the solution of the integrated model, the matheuristic algorithm is restricted to utilizing only the predetermined patterns, because the integrated model cannot handle a large set of patterns.

In the pattern change neighborhood and the next two neighborhoods, changes that we perform relate to the patients visited on a given day by a certain operator. Therefore, starting from this neighborhood, the term “day/operator combination” to define each day each operator visits their assigned patients, and this term is used throughout the remainder of the section.

The steps performed in this neighborhood are described in detail below:

- The day/operator combination with the highest workload is determined.
- A patient who is visited in the day/operator combination is randomly selected.
- Considering the number of days of service required by the selected patient, a new pattern is randomly chosen from the predetermined set of patterns such that no patient visit is needed on that specific day within the selected pattern.
- If there exists no such pattern for the selected patient, then the last two steps are repeated until a patient is found for whom a pattern change can be made. The algorithm proceeds to patient swap neighborhood if no pattern change is possible for the patients scheduled on that day.
- Once the move is identified, the pattern-related variables in the assignment-scheduling model solution are updated based on the change.
- The workload for each day/operator combination is estimated using the NN model. If the maximum workload is estimated to be reduced by the move, then the new solution is saved as a candidate solution. If not, the obtained solution is still accepted as a candidate solution with a 25% probability in order to account for potential errors in the NN model estimation.
- Finally, the routing model is solved for the returned candidate solution and the objective function value is evaluated to determine if any improvement has been achieved. If the resulting objective function value is better than that of the current one or a feasible solution is obtained from an infeasible one, then the current solution is updated, and the algorithm continues with the pattern change neighborhood repeatedly so long as the solution is improved. Otherwise, the patient swap neighborhood is initiated.

III. Patient Swap

The main objective of this neighborhood structure is to balance the days with the highest and the lowest workload. This is achieved by swapping the assigned operators of the patients who are visited on the identified days if it is possible. The swapping operation does not affect the patterns the patients are visited in the current solution; the swapped patient is visited on the same days by another operator.

The steps performed in this neighborhood are described in detail below:

- The first step of this neighborhood is to determine the operators who will exchange patients. In order to make this decision, the day/operator combinations with the highest and lowest workload are determined first. If both days are operated by the same operator, then another day is selected instead of the day initially chosen for the minimum workload, however the day with the maximum workload remains the same. This is achieved by first selecting another operator who is different from the one that operates on the day with the highest workload. Then, the day that the selected operator has the least workload is identified.
- Next, a patient is randomly chosen for the swap operation from each of the day/operator combinations identified in the previous step. Necessary updates are made in the patient assignment variables in the assignment-scheduling model solution to reflect the change to the current solution so that the selected patients are swapped between the selected operators. This update is done through the patient assignment variables in the model.
- The workload for each day/operator combination is calculated using the estimations obtained by NN. If the maximum workload is estimated to be reduced by the move, then the new solution is maintained as a candidate solution. Otherwise, the obtained solution is still accepted as a candidate solution with a 25% probability in order to account for potential errors in the NN model estimation.
- Finally, the routing model is solved for the returned candidate solution, and the resulting solution is assessed in terms of the objective function value or feasibility to check whether any improvement has been accomplished. If so, the current solution is modified, and the algorithm repeats the patient swap neighborhood until the solution is no longer improved. If no improvement is made, then the operator change neighborhood is activated.

IV. Operator Change

In order to reduce the workload, this neighborhood concentrates on the day/operator combination with the highest workload. This is accomplished by removing one of the patients scheduled for that day from that operator and assigning that patient to another operator. Similar to the patient swap neighborhood, the patient is visited on the same days by the newly assigned operator.

The steps performed in this neighborhood are described in detail below:

- The neighborhood initiates by identifying the day/operator combination with the highest and lowest workload. If both days are operated by the same operator, then another operator is randomly chosen among other operators in order to determine the operator to which the patient will be transferred.
- A patient is randomly picked among the ones who are visited by the operator in the day/operator combination with the highest workload. Afterwards, the patient assignment variables in the assignment-scheduling model solution are updated so that the selected patient is transferred from the operator who operates on the highest utilized day/operator combination to the one who operates the lowest utilized day/operator combination determined in the previous step.
- The workload for each day/operator combination is estimated using the NN model. If the maximum workload is estimated to be reduced by the move, then the new solution is kept as a candidate solution. Otherwise, the obtained solution is still accepted as a candidate solution with a 25% probability in order to account for potential errors in the NN model estimation.
- Finally, the routing model is solved for the returned candidate solution, and the objective function value is assessed to check whether any improvement has been accomplished. If the resulting objective function value is less than the current one or a feasible solution is obtained from an infeasible one, the current solution is modified, and the algorithm repeats the operator change neighborhood until the solution is no longer improved. If no improvement is made, then there are two options. The shaking procedure is executed if the necessary conditions are satisfied. The conditions and the shaking procedure are explained in detail later in this section. Otherwise, the algorithm repeats the aforementioned neighborhood structures in the given order starting from the first one, namely the patient group restriction neighborhood.

The VNS algorithm scans the four neighborhoods as described above as long as the current solution can be improved. If no improvement can be achieved in 10 consecutive neighborhood scans, then a shaking procedure is applied to restart the search from a new region of the solution space. The proposed framework involves two types of shaking procedures: swap shake and pattern shake. These two procedures are different in terms of the magnitude of the change they induce in the solution. The decision of which shake procedure will be used in each iteration is determined probabilistically for each type. Swap shake is applied to the solution with 20% probability, whereas pattern shake mechanism is utilized with 80% probability.

I. Swap Shake

This shake procedure aims to change the current solution by swapping a number of patients between two operators. The steps performed in this procedure are described in detail below:

- A pair of random operators are picked from the set of operators.
- From each of the previously selected operators, one-third of the patients assigned to them are randomly picked.
- The selected patients are then swapped between the two operators. This change is reflected in the solution by updating the patient assignment variables in the assignment-scheduling model solution.
- Sequentially, the assignment-scheduling problem is solved by fixing the values of the patient assignment variables that were updated according to the swap operation.
- If the obtained solution is feasible, it replaces the current solution. Otherwise, the current solution does not change.

II. Pattern Shake

This shake procedure attempts to change the current solution by introducing new patterns to the model from the list of predetermined patterns. The steps performed in this procedure are described in detail below:

- First, a pattern that is not used in the current solution is selected from the pattern pool.
- Next, the selected pattern is added to the set of patterns of the current solution and the assignment-scheduling problem is solved with the updated set of patterns.
- If the obtained solution is feasible, it is accepted as the current solution. Otherwise, no change is made in the current solution.

Following the shaking procedure, the algorithm continues with the neighborhood scanning procedure and the loop is repeated until the terminating criterion is reached.

5. COMPUTATIONAL RESULTS

In this chapter, results of the computational experiments for evaluating the performances of the proposed route duration estimation method and matheuristic framework are presented. The matheuristic algorithm, the NN model, and all mentioned mathematical models are coded in Python 3.8, and the models are solved using CPLEX 20.1.

In Section 5.1, we provide the results of an experiment that compares the performance of the two route duration estimation methods along with the description of the data generation method used for creating the TSP instances utilized for this comparison. Next, Section 5.2, we introduce the procedure for constructing the HHCP instances used for evaluating the performance of the proposed matheuristic framework followed by the computational results on the matheuristic performance.

5.1. PERFORMANCE COMPARISONS OF THE ROUTE DURATION ESTIMATION METHODS

Within the matheuristic framework, the solution is iteratively updated using the VNS metaheuristic. As explained in the previous section, an NN model is employed in the VNS algorithm to obtain the route duration estimates after making a certain update in the solution in order to identify the workload of each operator. Depending on the result, the new solution can be accepted as a candidate solution. Hence, the NN model is utilized in the evaluation of a certain move within the VNS algorithm.

In the study conducted by Yalçındağ et al. [7], the KR method is used to estimate the route durations of operators. The developed NN model is compared with the previously presented KR method. In this section, the results regarding this comparison are provided. A set of TSP instances is generated to test the performance of both methods. The following subsection describes the method used to generate these instances.

5.1.1. Generation of the TSP Instances

Yalçındağ et al. [7] observed that the KR technique does not perform well if the locations of patients used in the estimation are far away from those in the historical data used in construction of the Kernel estimator. Therefore, comparisons are made considering this deficiency of the KR method. In this context, three groups of data are generated as follows:

- Historical patient locations, or the training data
- Test data consisting of patient locations that are far away from the training data
- Test data consisting of patient locations that are nearby the training data

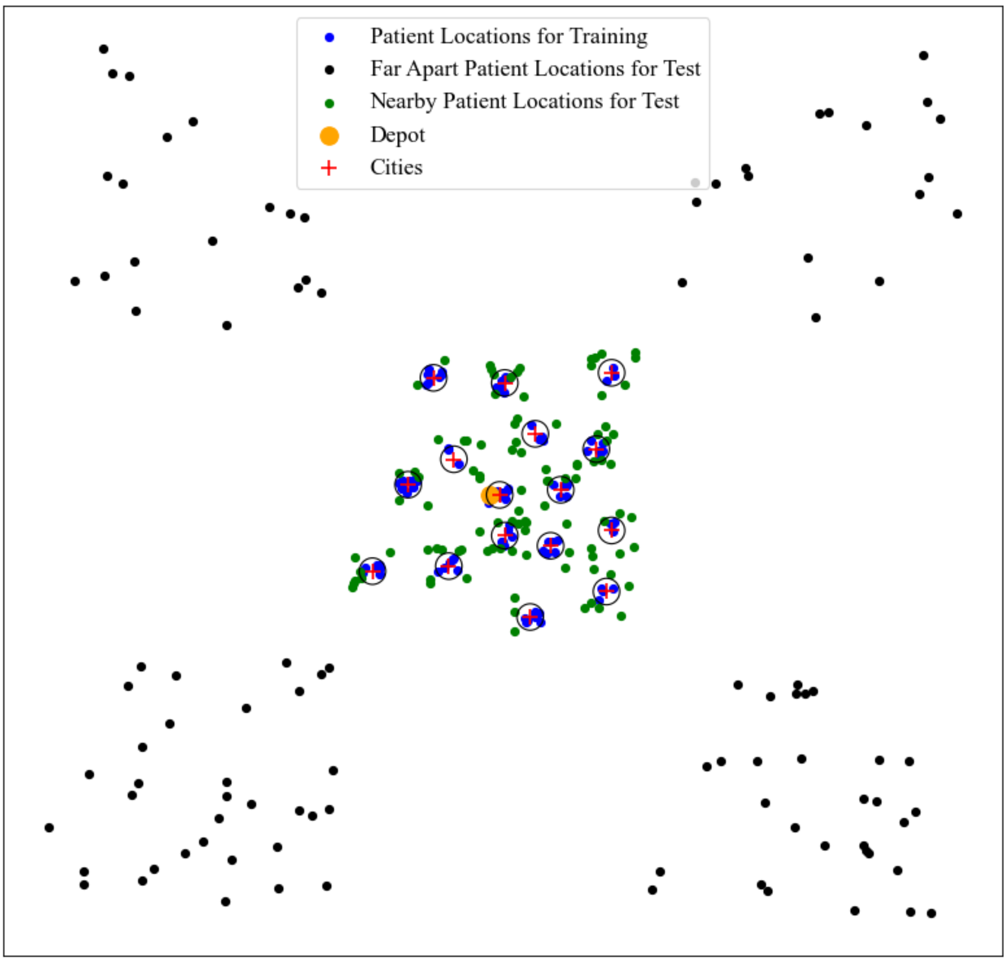

Figure 5.1. Illustration of a sample TSP instance generated using the described method

The terms "training data" and "historical patient locations" refer to the data used to build the NN model and the Kernel estimator in the KR technique. Therefore, these terms essentially refer to the same data and will be used interchangeably throughout the rest of the chapter. Furthermore, test data include locations of patients along a certain route for which the route duration must be estimated.

In order to generate the training data consisting of historical patient locations, first a number of cities are marked in a specific area. Then, the historical patient locations are generated such that their x and y coordinates are distributed around those of the predetermined city locations by a deviation sampled from a normal distribution with a mean of one and standard deviation of 0.5. In Figure 5.1, the cities are represented in red color, and the blue dots show the location of patients in the historical data (i.e., training data). Green dots denote the patient locations that are nearby the locations in the training data. Their coordinates are generated in a similar way to the training data itself, except this time the parameters for normal distribution are chosen as three and one for the mean and the standard deviation, respectively. Next, black dots show the patient locations further away from the historical data. A similar procedure is followed in the generation of those locations, and this time the corresponding parameters are set to 60 and 3.5 for the mean and the standard deviation. Lastly, the orange dot represents the depot. It is assumed that each route in the historical data starts and ends at the depot illustrated in the figure. The historical data used in training both estimation methods consist of the x and y coordinates of the patients visited on each route as input and the duration of the corresponding route as output. In the experiments, routes consisting of four and eight patients are considered to evaluate the performance of both methods on routes with different numbers of nodes. In addition to that, the effect of the size of the training data is also investigated. The problems can be summarized in Table 5.1 below.

Table 5.1. Summary of the problems used in the comparison of the estimation methods

Problem Instance	**Number of Patients on Each Route**	**Size of the Test Set**	**Size of the Training Set**
F-4	4	1000	250-500-1000
N-4	4	1000	250-500-1000
F-8	8	1000	250-500-1000
N-8	8	1000	250-500-1000

In the above table, the problem instances are named according to their distances from the locations in the training data and the number of patients they contain on each route. The instances with "F" in their names consist of locations further away from the historical data, whereas the letter "N" stands for instances consisting of locations nearby the historical data. The numbers denote the number of patients each instance has on each of its routes, and each problem is tested using a test set composed of 1000 rows (routes). Both estimation methods are trained with the same historical data of 250, 500, and 1000 rows and tested using the instances described above.

5.1.2. Test Results on the Comparison of Route Duration Estimation Methods

The rows in each problem instance are divided randomly into 10 groups, and 100 estimations are obtained for each group using both estimation methods. The estimation error is calculated using Mean Absolute Error (MAE) which is defined in Equation (5.1) where y_i represents the real route duration obtained by solving the TSP, x_i is the route duration estimate from the estimation method, and n is the size of the test data (100) for each group.

$$MAE = \frac{\sum_{i=1}^{n} |y_i - x_i|}{n} \tag{5.1}$$

MAE values of 10 groups of test data are obtained to compare the performances of the two methods. The comparisons are made using the average MAE values of 10 groups.

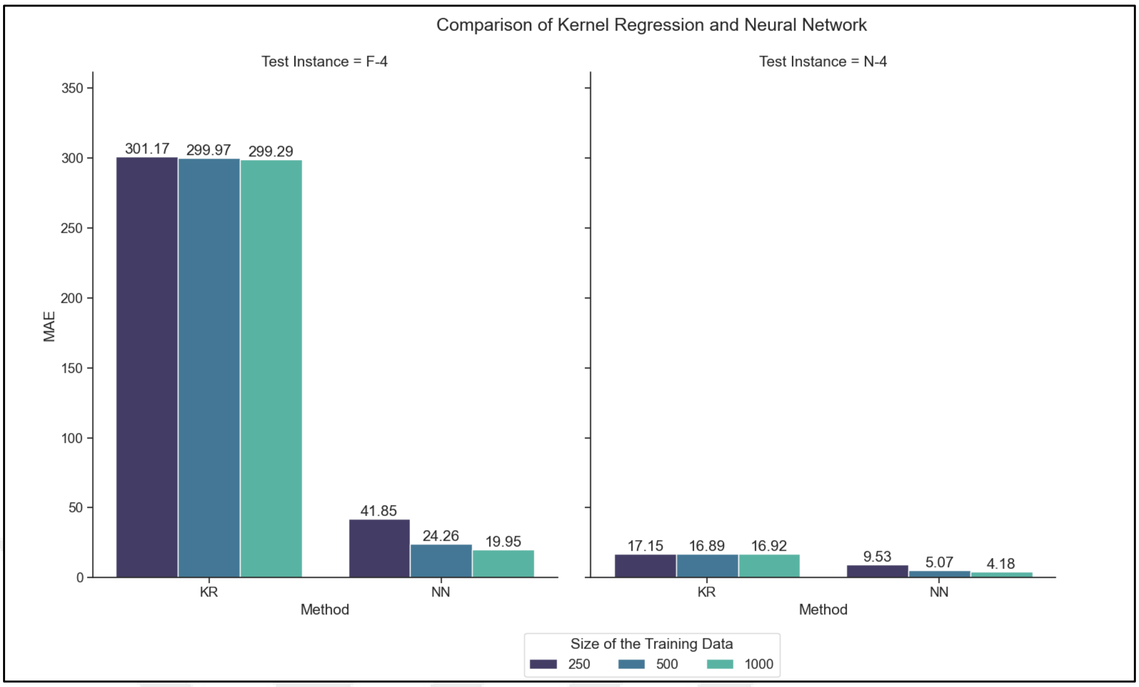

Figure 5.2. Performance comparison of route duration estimation methods on problems with four patients, F-4 and N-4

Above, Figure 5.2 reveals the significant gap in errors between KR and NN estimations for the problem instance F-4. As it was stated by Yalçındağ et al. [7], the KR technique cannot adapt to the difference between the locations well, therefore the error is much higher for instance F-4 than N-4. When the estimations are made for instances near the historical location data, the performances of both methods are similar. However, it is observed that NN estimations are better than KR regardless of the instance type. It also appears that the performance of NN is improved as the size of the training data increases for both problem instances whereas the improvement in the performance of the KR method with more training data is quite limited. This result suggests that the NN is more capable of taking advantage of additional data than KR.

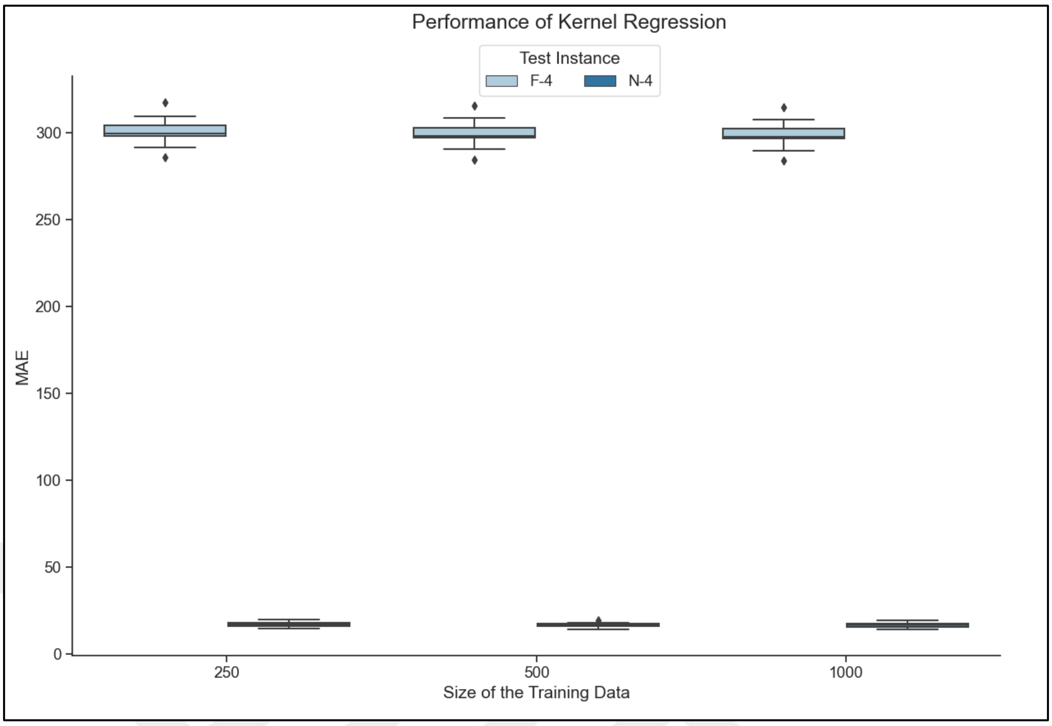

Figure 5.3. Boxplot of the MAE values of the KR method on problems F-4 and N-4

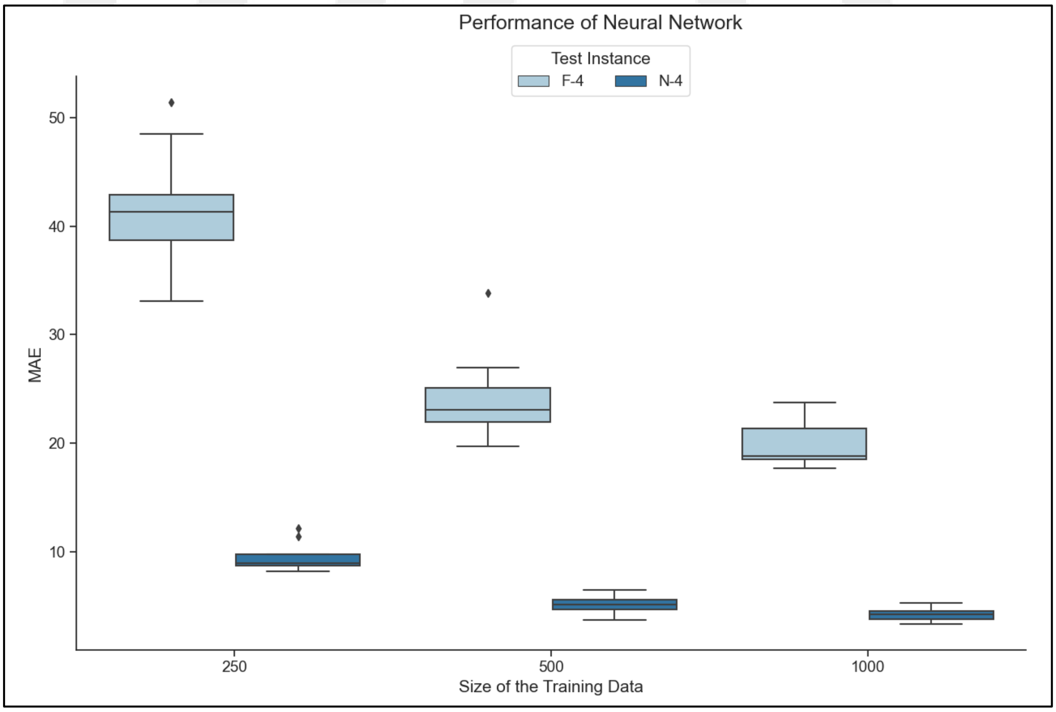

Figure 5.4. Boxplot of the MAE values of NN model on instances F-4 and N-4

As shown in Figure 5.3 and Figure 5.4, the KR approach has a lower range and interquartile range of error than the NN model overall. However, the range of errors in NN estimations decreases as the size of the training data is increased.

When the number of patients on each route is increased to eight, it is discovered that the error in the estimations of both methods also increases. In contrast to Figure 5.2, Figure 5.5 shows for the problem N-8, there is no significant difference between the methods with smaller training data. Although the performance of the NN seems to have worsened for these problem instances, when it is trained with a large enough data set, it surpasses the performance of the KR method.

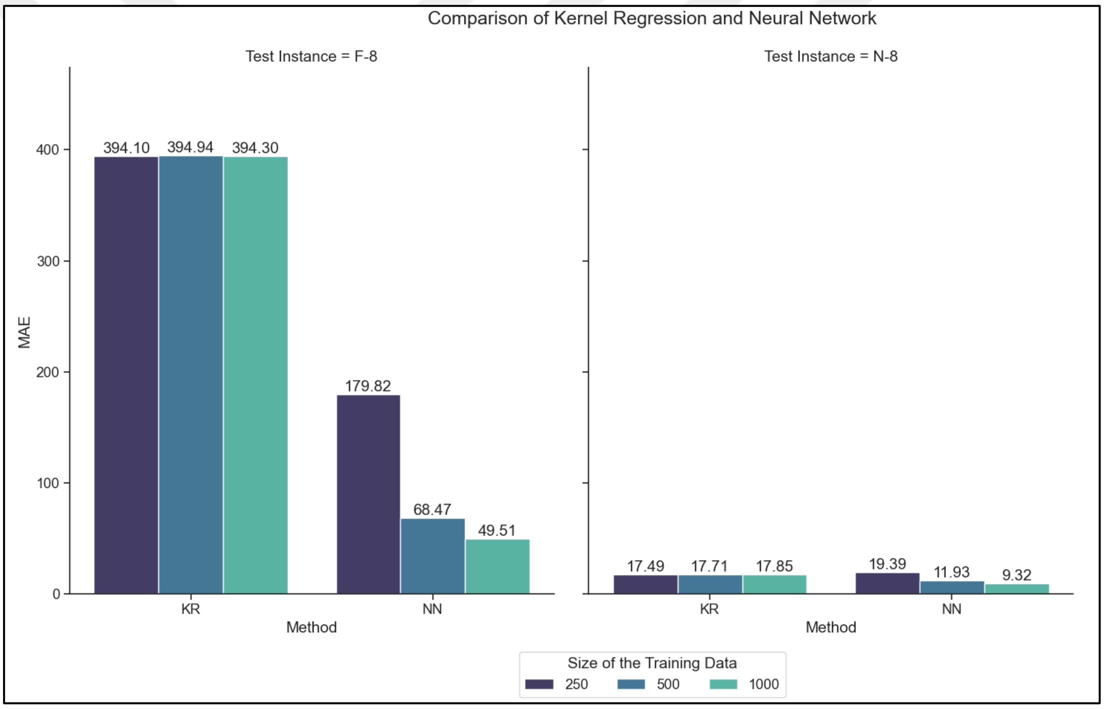

Figure 5.5. Performance comparison of route duration estimation methods on problems with eight patients, F-8 and N-8

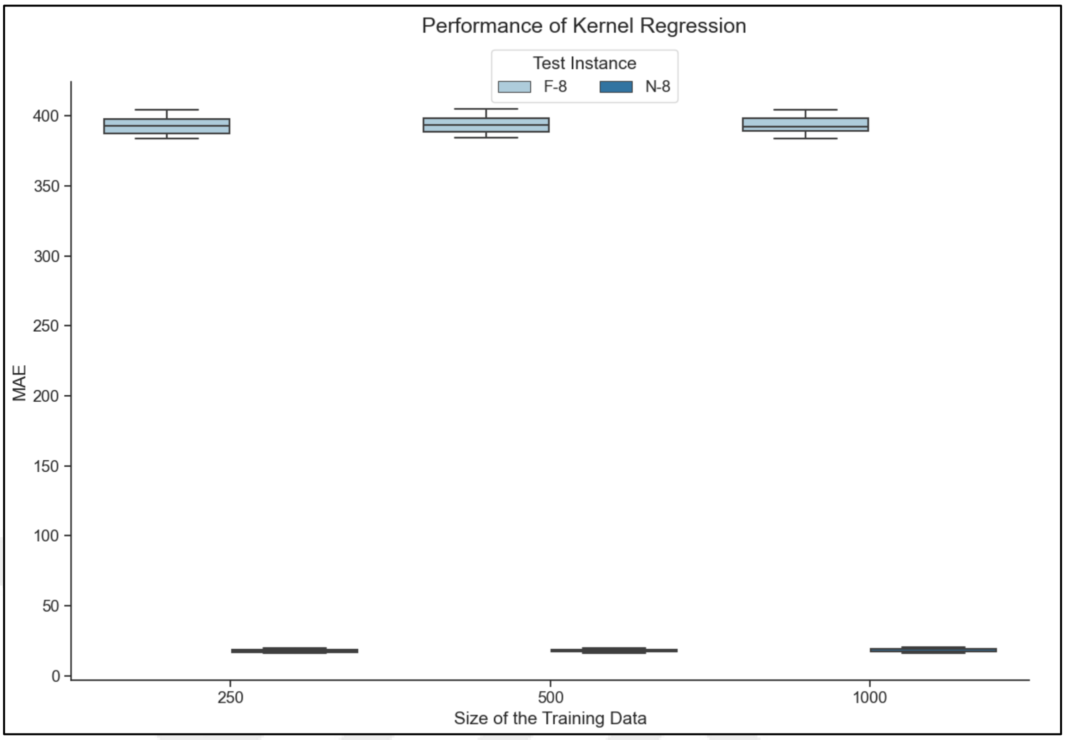

Figure 5.6. Boxplot of the MAE values of KR method on instances F-8 and N-8

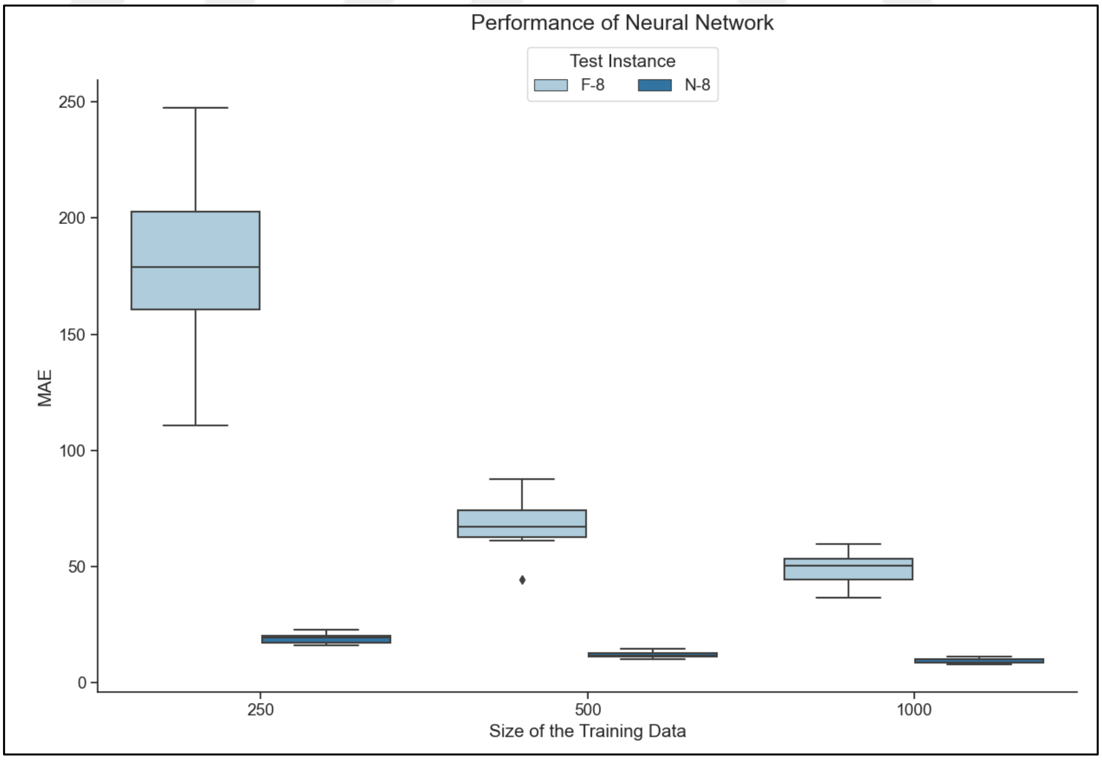

Figure 5.7. Boxplot of the MAE values of the KR method on instances F-8 and N-8

Figure 5.7 also illustrates that the NN model struggles in estimating route durations for with more stops on each route when patients are further away, and it is seen that the errors are not as stable compared to the performance presented in Figure 5.4. Nevertheless, the performance is increased if the NN model is provided with more training data.

The tests show that the NN model delivers better estimates of the route duration than KR whether the tested locations are far apart or near the historical locations. Moreover, KR technique cannot handle historical data that contain routes with different numbers of patients, therefore a separate estimation needs to be carried out for each set of historical data of different route size. However, the NN model can work with routes of different sizes. Hence, a huge amount of training data consisting of routes of varying lengths are used to train the NN model at the beginning of the matheuristic algorithm, and the estimations are obtained throughout the algorithm using the previously trained model.

5.2. PERFORMANCE OF THE MATHEURISTIC FRAMEWORK

In order to evaluate the performance of the proposed matheuristic framework, a total of six problem instances of different sizes are generated. This section is dedicated to the explanation of the problem instance generation procedure for the addressed problem (Section 5.2.1) and the results obtained regarding the performance of the presented matheuristic framework (Section 5.2.2).

5.2.1. Generation of Problem Instances

In order to generate problem instances of HHCP, the first step is to determine the patient locations. For random problem instances, the x and y coordinates of each patient are sampled from a uniform distribution with a lower bound of zero and an upper bound of 50. The location of the depot, or HHC center is selected as the center of the considered area (25, 25). Time windows of the patients are generated randomly according to a predefined ratio depicted below.

$$\frac{LFT_j - EST_j}{s_j} \tag{5.2}$$

For each patient, the ratio depicted in (5.2) is sampled from a normal distribution with a mean of three and a standard deviation of 0.2. Then, the values of LFT_j and EST_j are randomly generated according to the determined ratio. This is achieved by first randomly generating a ratio and an EST_j value, then the LFT_j is computed using (5.2). The service time s_j is assumed to be 45 minutes for all problem instances. In most of the real-life problems, the time windows of patients are not strict; meaning that the visits can usually occur within a large time interval during working hours. Therefore, the value range used for the ratio can be considered as a realistic demonstration of patients' time windows. The resulting time windows are visualized in Figure 5.8 below for one of the problem instances with 15 patients.

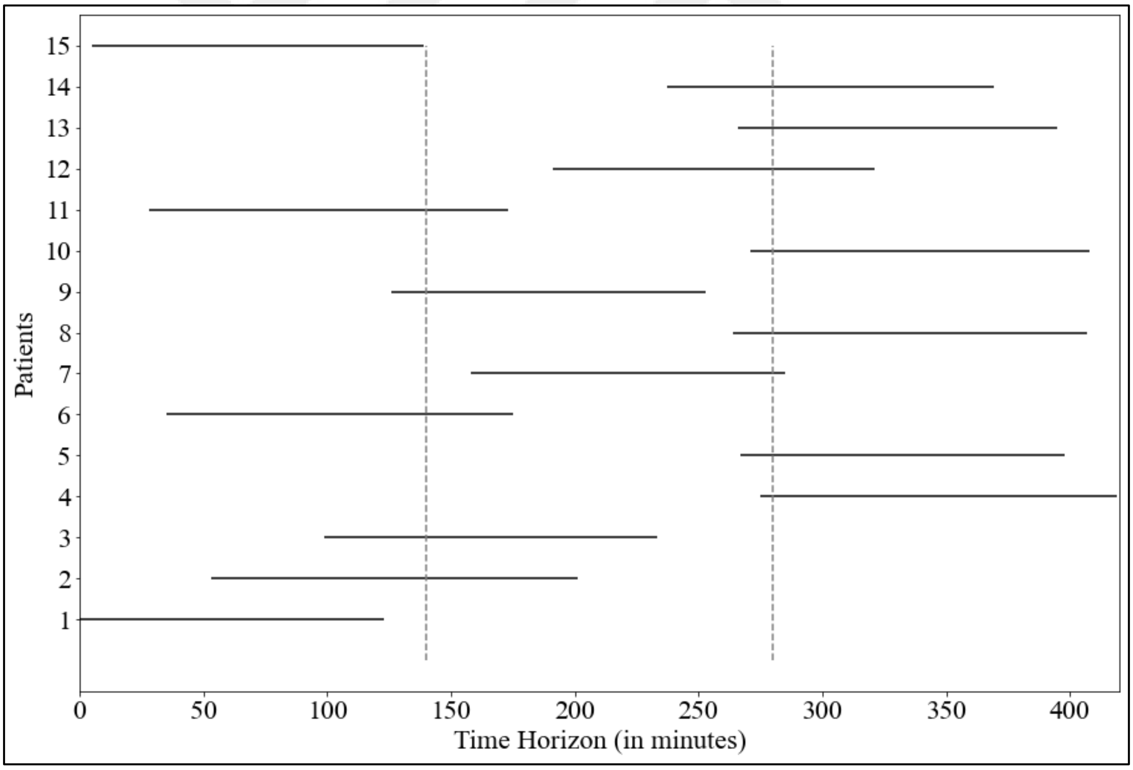

Figure 5.8. Visualization of time windows over daily working hours for problem instance PI-1

In the above figure, the time window of each patient is illustrated for one of the selected problem instances over the daily working hours, which is 420 minutes. The vertical dashed lines divide the 420-minute time horizon into three equal intervals, and each interval represents the morning, noon, and evening times of a 24-hour day period. Time windows are generated in a realistic way such that patients are usually available during one of the time intervals stated. The same distribution and parameters are used to generate time windows of the patients for all the problem instances used throughout this thesis.

The patterns for each problem instance are also generated in a random manner by considering the necessary patterns that would lead to a feasible solution. Since the patterns are responsible for the scheduling of the visits, they must be in line with the care plan of each patient. It is assumed in problem instances that a patient can require up to four days of visits from any operator. Therefore, three to four patterns are randomly generated for each possible care plan. As it was mentioned earlier, there exists a total list of patterns that can be utilized throughout the algorithm. Such restriction is needed to be able to compare the solution obtained from the matheuristic algorithm and the integrated model since the integrated model cannot deal with a large set of patterns as it increases the complexity of the problem. A subset of the entire pattern list is provided to the assignment-scheduling problem, so not all of the patterns are available at the beginning. It is up to the algorithm to find the most profitable patterns and incorporate them into the solution. Finally, the capacity of each operator is assumed to be equal, which indicates a daily shift of 420 minutes for all operators.

The size of the problem is described by the number of patients, operators, and patterns present in the problem. Two instances of each size are generated randomly using the method described above. Each problem instance is identified by a combination of the abbreviation "PI", and the instance number. The total of six problem instances are summarized in Table 5.2.

Table 5.2. Characteristics of the problem instances used in the evaluation

Problem Instance	Number of Patients	Number of Operators	Initial Number of Patterns	Total Number of Patterns
PI-1	15	2	5	11
PI-2	15	2	6	11
PI-3	20	2	7	13
PI-4	20	2	5	13
PI-5	25	3	6	15
PI-6	25	3	7	15

5.2.2. Performance of the Matheuristic Framework on Problem Instances

For each of the problem instances mentioned in the previous section, five independent replications are obtained. Each replication is obtained by running the algorithm on the instance for three hours. The algorithm is initialized by first solving the assignment-scheduling problem for a provided subset of patterns. Then, the routing problem is solved using the assignment and scheduling information obtained from the this solution. As it was mentioned earlier, the two-stage approach often yields an infeasible solution in the routing stage. This is the case for five out of six problem instances generated for the tests, as displayed in Table 5.3.

Table 5.3. The initial solutions obtained from the two-stage solution approach for each problem instance

Problem Instance	Number of Patients	Number of Operators	Initial Objective Function Value
PI-1	15	2	Infeasible
PI-2	15	2	Infeasible
PI-3	20	2	Infeasible
PI-4	20	2	0.9065
PI-5	25	3	Infeasible
PI-6	25	3	Infeasible

Except for the problem instance denoted with PI-4, the algorithm is initialized with an infeasible solution for all instances. As stated earlier, the algorithm can handle infeasible solutions within the VNS metaheuristic. The first feasible solution found by the algorithm may vary between replications due to the randomness within the algorithm.

In Table 5.4, the results obtained from the matheuristic framework are presented for each of the five replications of the six problem instances. In the cases for which the initial solution is infeasible, the improvement percentage is calculated by comparing the objective function values of the best-found solution and the first feasible solution obtained by the matheuristic algorithm. It can be interpreted from the table that even when the algorithm starts from different feasible solutions, it can converge to solutions with similar objective function values. Depending on the first feasible solution obtained within the algorithm, high improvement ratios up to 58% are observed.

Table 5.4. Performance of the matheuristic framework on each of the problem instances with five independent replications

Problem Instance	Replication	Objective Function Value of Initial Solution	Objective Function Value of the First Feasible Solution	Objective Function Value of Matheuristic Solution (in 3 hours)	Improvement (%)
PI-1	1	Infeasible	0.7482	0.6333	18.14
	2		0.7482	0.6676	12.07
	3		0.7482	0.6691	12.87
	4		0.7482	0.6786	10.26
	5		0.7482	0.6686	11.90
	AVERAGE	-	**0.7482**	**0.6634**	**13.05**
PI-2	1	Infeasible	0.8288	0.6726	23.23
	2		0.9707	0.6624	46.53
	3		0.9707	0.6605	46.96
	4		0.8288	0.6605	25.48
	5		0.9183	0.6605	39.03
	AVERAGE	-	**0.9035**	**0.6633**	**36.25**
PI-3	1	Infeasible	0.7944	0.7395	7.43
	2		0.7636	0.7135	7.02
	3		0.8255	0.7259	13.71
	4		0.8121	0.7259	11.87
	5		0.7623	0.7380	3.33
	AVERAGE	-	**0.7916**	**0.7286**	**8.67**
PI-4	1	0.9065	0.9065	0.8336	8.75
	2		0.9065	0.7889	14.90
	3		0.9065	0.8187	10.72
	4		0.9065	0.7922	14.44
	5		0.9065	0.7961	13.87
	AVERAGE	-	**0.9065**	**0.8059**	**12.54**
PI-5	1	Infeasible	0.8740	0.6342	37.82
	2		0.8740	0.6449	35.53
	3		0.8587	0.6291	36.50
	4		0.8740	0.6448	35.54
	5		0.8740	0.6376	37.07
	AVERAGE	-	**0.8710**	**0.6381**	**36.49**
PI-6	1	Infeasible	0.9481	0.6035	57.11
	2		0.9204	0.5914	55.63
	3		0.9481	0.5966	58.91
	4		0.9204	0.6031	52.60
	5		0.9204	0.6052	52.07
	AVERAGE	-	**0.9315**	**0.6000**	**55.26**

Table 5.5. Comparison of the performance of the matheuristic framework and the integrated model

Problem Instance	Replication	Objective Function Value of Matheuristic Solution (in 3 hours)	Time at Which the Best Solution is Obtained (min)	Objective Function Value of the Integrated Model	Integrated Model Solution Time (min)	Difference Between Matheuristic and the Integrated Model (%)
PI-1	1	0.6333	6.94	0.6333	4.23	0.00
	2	0.6676	51.81			5.42
	3	0.6691	14.16			5.66
	4	0.6786	1.06			7.15
	5	0.6686	7.75			5.58
PI-2	1	0.6726	8.09	0.6579	4.77	2.23
	2	0.6624	34.26			0.69
	3	0.6605	133.17			0.40
	4	0.6605	72.15			0.40
	5	0.6605	38.87			0.40
PI-3	1	0.7395	150.42	0.6969	141.75	6.11
	2	0.7135	114.66			2.37
	3	0.7259	15.48			4.16
	4	0.7259	14.87			4.16
	5	0.7380	60.41			5.89
PI-4	1	0.8336	6.66	0.7732	517.48	7.81
	2	0.7889	28.03			2.03
	3	0.8187	12.84			5.89
	4	0.7922	93.09			2.45
	5	0.7961	51.54			2.96
PI-5	1	0.6342	176.66	**0.6230***	540.00	1.80
	2	0.6449	89.67			3.53
	3	0.6291	95.08			0.99
	4	0.6448	105.07			3.51
	5	0.6376	51.33			2.36
PI-6	1	0.6035	39.42	**0.6008***	540.00	0.44
	2	0.5914	111.63			1.57
	3	0.5966	36.88			0.70
	4	0.6031	88.48			0.39
	5	0.6052	175.45			0.74

Table 5.5 depicts the comparison between the performances of the matheuristic framework and the integrated model. For the sake of a fair comparison, the integrated model is solved given the total list of allowed patterns. It should be noted that for problem instances PI-5 and PI-6, the objective function values of the integrated model solution marked with (*) have a

20% optimality gap since no optimal solution was obtained within nine hours, however, all other integrated model solutions are optimal with zero optimality gap. Overall, it is observed that the solution time of the integrated model grows longer as the problem size increases. For larger problem instances, e.g., problem instance PI-3 and beyond, the results indicate that solving the problem using the matheuristic framework leads to near-optimal solutions with at most a 7.81% difference from the integrated model solution in a shorter time (within 3 hours). On top of that, in the problem instance denoted with PI-6, two of the matheuristic solutions (replications two and three) yielded a better objective function value than that of the integrated model solution in less than two hours.

6. CONCLUSION

The primary contribution of this thesis is the introduction of a matheuristic framework for the HHCP. This problem assigns identically skilled operators to patients, schedules their visits, and establishes the routes in a time horizon of one week while taking the patients' time windows into consideration with the objective of minimizing the maximum operator workload. For large problem instances, the integrated model for the HHCP cannot be solved to optimality due to its NP-Hardness. In addition to that, the two-stage solution approach does not guarantee a feasible solution since the assignment and scheduling decisions are made without considering the routing information. Therefore, we propose a matheuristic framework that improves the solution created using the two-stage solution approach for HHCP by iteratively updating the solution through the VNS metaheuristic. Even if the initial solution found by the two-stage approach is infeasible, the framework can improve the existing solution and achieve near-optimal solutions.

Another important contribution is the development of an NN model that estimates the route durations for visiting a set of given locations. The estimations are used within the VNS algorithm during the evaluation stage of a particular move. The greatest advantage of using the matheuristic framework is to be able to reflect the past routing decisions made by the operators to the current solution. As an extension to the current NN model, more guided and comprehensive estimations can be obtained by providing information regarding the attributes of the patients or other distinctive information about the patient locations.

In order to assess the performance of the NN based estimation method and the matheuristic framework, a set of problem instances is generated. The developed NN model is compared to the KR estimation method developed in Yalçındağ et al. [7], and the results revealed that the NN model presented in this thesis provides better estimates. Furthermore, since the weights used in NN models are continuously updated, the learning process continues as new data are introduced to the model. Hence, in time, the model can make even better estimations regardless of the difference in geographical locations, or regions.

Moreover, the results obtained via the matheuristic framework are compared to the ones acquired by solving the integrated model, and it is observed that the framework provides optimal or near-optimal solutions in a reasonable time. Therefore, the framework can be used as a solution approach for the HHCP.

Another area of future research can focus on extending the neighborhood set within the VNS algorithm to include a neighborhood that considers the time windows of patients. This way, the algorithm may perform changes that are better guided towards creating feasible solutions.

REFERENCES

1. Palladino L. Public Investment in Home Healthcare in the United States During the COVID-19 Pandemic: A Win-Win Strategy. *Fem Econ* 2021; 27: 436–452.
2. On Birinci Kalkınma Planı 2019-2023; [cited 2023 12 June]. Available from: https://www.sbb.gov.tr/wp-content/uploads/2022/07/On_Birinci_Kalkinma_Plani-2019-2023.pdf.
3. Korte B, Vygen J. The Traveling Salesman Problem. 2012; 557–592.
4. Chaieb M, Jemai J, Mellouli K. A decomposition - construction approach for solving the home health care scheduling problem. *Health Care Manag Sci* 2020; 23: 264–286.
5. Gomes MI, Ramos TRP. Modelling and (re-)planning periodic home social care services with loyalty and non-loyalty features. *Eur J Oper Res* 2019; 277: 284–299.
6. Yalçindağ S, Cappanera P, Grazia Scutellà M, et al. Pattern-based decompositions for human resource planning in home health care services. *Comput Oper Res* 2016; 73: 12–26.
7. Yalçındağ S, Matta A, Şahin E, et al. The patient assignment problem in home health care: using a data-driven method to estimate the travel times of care givers. *Flex Serv Manuf J* 2016; 28: 304–335.
8. Cappanera P, Scutellà MG. Joint assignment, scheduling, and routing models to home care optimization: A pattern-based approach. *Transportation Science* 2015; 49: 830–852.
9. Fikar C, Hirsch P. Home health care routing and scheduling: A review. *Computers and Operations Research* 2017; 77: 86–95.
10. Cissé M, Yalçındağ S, Kergosien Y, et al. OR problems related to Home Health Care: A review of relevant routing and scheduling problems. *Operations Research for Health Care* 2017; 13–14: 1–22.

11. Euchi J, Masmoudi M, Siarry P. Home health care routing and scheduling problems: a literature review. *4OR* 2022; 20: 351–389.

12. Grieco L, Utley M, Crowe S. Operational research applied to decisions in home health care: A systematic literature review. *Journal of the Operational Research Society* 2021; 72: 1960–1991.

13. Liu R, Yuan B, Jiang Z. Mathematical model and exact algorithm for the home care worker scheduling and routing problem with lunch break requirements. *Int J Prod Res* 2017; 55: 558–575.

14. Frifita S, Masmoudi M. VNS methods for home care routing and scheduling problem with temporal dependencies, and multiple structures and specialties. *International Transactions in Operational Research* 2020; 27: 291–313.

15. Nasir JA, Kuo Y-H. A decision support framework for home health care transportation with simultaneous multi-vehicle routing and staff scheduling synchronization. Epub ahead of print 2020. DOI: 10.1016/j.dss.2020.113361.

16. Lin M, Ma L, Ying C. Matching daily home health-care demands with supply in service-sharing platforms. *Transportation Research Part E* 2021; 145: 102177.

17. Akbari V, Sadati İ, Salman FS, et al. Minimizing total weighted latency in home healthcare routing and scheduling with patient prioritization. *OR Spectrum*. Epub ahead of print 2023. DOI: 10.1007/S00291-023-00713-3.

18. Nickel S, Schröder M, Steeg J. Mid-term and short-term planning support for home health care services. *Eur J Oper Res* 2012; 219: 574–587.

19. Trautsamwieser A, Hirsch P. A branch-price-and-cut approach for solving the medium-term home health care planning problem. *Networks* 2014; 64: 143–159.

20. Grenouilleau F, Legrain A, Lahrichi N, et al. A set partitioning heuristic for the home health care routing and scheduling problem. *Eur J Oper Res* 2019; 275: 295–303.

21. Mosquera F, Smet P, Vanden Berghe G. Flexible home care scheduling R. *Omega (Westport)* 2019; 83: 80–95.

22. Grenouilleau F, Lahrichi N, Rousseau LM. New decomposition methods for home care scheduling with predefined visits. *Comput Oper Res*; 115. Epub ahead of print 1 March 2020. DOI: 10.1016/j.cor.2019.104855.

23. Gong X, Geng N, Zhu Y, et al. A Matheuristic Approach for the Home Care Scheduling Problem with Chargeable Overtime and Preference Matching. *IEEE Transactions on Automation Science and Engineering* 2021; 18: 282–298.

24. Cinar A, Salman FS, Bozkaya B. Prioritized single nurse routing and scheduling for home healthcare services. *Eur J Oper Res* 2021; 289: 867–878.

25. Bhattarai S, Correa-Martinez Y, Bedoya-Valencia L. A multi-objective home healthcare routing problem. *Int J Healthc Manag* 2022; 16: 311–325.

26. Yalçındağ S, Matta A. A decomposition approach for the home health care problem with time windows. In: *Springer Proceedings in Mathematics and Statistics*. Springer New York LLC, 2017, pp. 221–232.

27. Xie W, Liu T, Li X, et al. Robust homecare service capacity planning. *Comput Oper Res*; 154. Epub ahead of print 1 June 2023. DOI: 10.1016/j.cor.2023.106155.

28. Tanoumand N, Ünlüyurt T. An exact algorithm for the resource constrained home health care vehicle routing problem. *Ann Oper Res* 2021; 304: 397–425.

29. Gomes MI, Rodrigues T, Ramos P. Modelling and (re-)planning periodic home social care services with loyalty and non-loyalty features. *Eur J Oper Res* 2019; 277: 284–299.

30. Wand MP, Jones MC. Kernel Smoothing. *Kernel Smoothing*. Epub ahead of print 1 December 1994. DOI: 10.1201/B14876.

31. Müller Berndt, Reinhardt Joachim, Strickland MT. Neural Networks : an Introduction. 1995; 330.

32. Gao J, Sun L, Gen M. A hybrid genetic and variable neighborhood descent algorithm for flexible job shop scheduling problems. *Comput Oper Res* 2008; 35: 2892–2907.

33. Hemmelmayr VC, Doerner KF, Hartl RF. A variable neighborhood search heuristic for periodic routing problems. *Eur J Oper Res* 2009; 195: 791–802.

34. Fleszar K, Osman IH, Hindi KS. A variable neighbourhood search algorithm for the open vehicle routing problem. *Eur J Oper Res* 2009; 195: 803–809.

35. Mladenović N, Hansen P. Variable neighborhood search. *Comput Oper Res* 1997; 24: 1097–1100.