

## A TWO-STAGE SOLUTION APPROACH FOR THE LARGE-SCALE HOME HEALTHCARE ROUTEING AND SCHEDULING PROBLEM

M. Erdem<sup>1,2</sup> & S. Bulkan<sup>2</sup>

### ARTICLE INFO

#### Article details

Submitted by authors 31 Mar 2017  
Accepted for publication 25 Oct 2017  
Available online 13 Dec 2017

#### Contact details

\* Corresponding author  
mehmet.erdem@atilim.edu.tr

#### Author affiliations

- 1 Department of Industrial Engineering, Faculty of Engineering, Atılım University, Turkey
- 2 Department of Industrial Engineering, Faculty of Engineering, Marmara University, Turkey

#### DOI

<http://dx.doi.org/10.7166/28-4-1754>

### ABSTRACT

The purpose of this study is to introduce a two-stage solution approach for a large-scale home healthcare routeing and scheduling problem (HHCRSP). In the first part of the two-stage solution approach, a cluster-assign algorithm is employed, based on the home location and the time to obtain feasible clusters. In the second stage, using these clusters, route construction heuristics start to create schedules and routes, taking the side constraints of the model into consideration. Using the novelty of this two-stage solution approach, higher diversification is achieved with a series of newly-developed cross movement strategies. The computational results show that our solution approach offers certain advantages, such as an increase in the efficient use of human resources, and a decrease in the working time of nurses.

### OPSOMMING

Die doel van hierdie studie is om 'n tweeledige oplossingsbenadering vir grootskaalse tuissorg roetebeplanning en skedulering voor te stel vir gesondheidsorg. In die eerste deel van die benadering word 'n groepering-algoritme, gebaseer op die ligging van die tuiste en die tyd om vatbare groepe te vorm, toegepas. Tydens die tweede stap van die benadering word die groepe gebruik saam met roetebeplanning-heuristiek om skedules en roetes te skep terwyl die randvoorwaardes van die model in ag geneem word. Die tweeledige benadering bewerkstellig hoër diversifikasie met 'n reeks nuut ontwikkelde kruisbeweging strategieë. Die gesimuleerde resultate toon dat die voorgestelde oplossingsbenadering sekere voordele inhou, onder andere die 'n verhoging in die effektiewe aanwending van menslike hulpbronne en 'n afname in die werktyd van verpleegsters.

## 1 INTRODUCTION

Health can be briefly defined as the general physical, mental, and social conditions of being well and free from disease [1]. Although there are different and contradictory definitions of health in the literature [2, 3], the general importance of health cannot be ignored. Today, as the population gets older, there is an even greater demand for public healthcare. The World Health Organization (WHO) predicts that the allocation of human resources will be a crucial challenge for the healthcare industry in the next ten years as the population ages. Moreover, the growth rate in healthcare expenses exceeds that of the gross domestic product (GDP) in many countries; thus this issue is crucial for the financial sustainability of healthcare systems. In the US, healthcare spending was 17.5 per cent of GDP (\$3 trillion) [4] in 2014, and is predicted to reach 19.3 per cent of GDP (about \$4.5 trillion) in 2019. In France, healthcare expenditure is expected to reach €70 billion in 2020 [5]. Many countries are thus looking for ways to improve their fiscal support of healthcare. Furthermore, with labour costs representing nearly half of a hospital's operating costs, executives struggle to manage their healthcare workforce efficiently [6].

The home healthcare routing and scheduling problem (HHCSP) comprises the aspects of two well-known problems: routing and rostering. The HHCSP gives patients the opportunity to be nursed at home, which is preferable to visiting clinics or hospitals. Furthermore, home care increases the level of service quality and becomes more cost-efficient. On the other hand, designing schedules is difficult due to factors such as patients' or nurses' preferences, policies, time intervals, and travel time according to transportation modes [7].

In this paper, we attempt to address and solve this large-scale HHCSP. The main motivation of this work is to develop a solution approach to be used to create appropriate solutions while taking all restrictions, preferences, regulations, etc. into consideration. Our two-stage solution approach is based on the 'cluster first and schedule route second' procedure. In the first stage, a cluster-assign procedure is developed, based on the two-step cluster (TSC) algorithm; and in the second stage, the variable neighbourhood descend (VND) algorithm with a series of newly developed strategies is implemented for systematic exploration. Moreover, it is free from a parameter selection/tuning phase; the proposed schedule-route construction heuristics do not necessarily cover all the jobs at the beginning of the second stage; and a series of newly developed cross movement neighbourhood structures provide for higher diversification for the large-scale HHCSP. Synchronisation is also considered within the proposed model. A two-stage solution approach considers not only the preferences of the nurses and clients, but also those of the healthcare institutions.

This paper is organised as follows. In Section 2, a literature review on the HHCSP and, in Section 3, the definition of the problem and the model are presented. The proposed solution approach appears in Section 4. The results of the large-scale HHCSP and comparisons are reported in Section 5. The paper is concluded and future research is discussed in the last section.

## 2 LITERATURE REVIEW

In this section we review some of the works that tackle the short-term HHCSP. Cheng and Rich [8] present a mixed integer programming (MIP) model and heuristic approach to deal with the HHCSP as a multiple vehicle routing problem with time windows (VRPTW). Full-time and part-time home carers are distinguished, and the lunch break is also considered within the time windows. The authors consider minimising the cost of home carers as an objective. A two-step heuristic approach is proposed. In the first step, a randomised greedy heuristic finds an initial solution. During the next step, the solution is improved using a local search algorithm. The approach is tested using randomly-generated small test examples (four home carers and 10 patients) and larger (up to 300 home carers and 900 patients).

Eveborn, Flisberg and Rönnqvist [9] propose a decision support system (DSS) to solve daily schedules for an organisation in Sweden. They use a set-partitioning problem (SPP) model and improve it by using a repeated matching heuristic approach. The objective function consists of the cost related to the travelling time, and the constraints are time windows for visits or breaks for meals, and the competence level of the staff. Their model is evaluated with small test examples (21 home carers and 123 tasks).

Bertels and Fahle [10] develop a model to minimise transportation costs while maximising the satisfaction level of home carers and patients. The patients' preferences for certain visits, the working time limitations for home carers, and their competence level are considered as constraints. In order to solve single-day problem instances (20-50 nurses and 111-326 tasks), the authors apply a combination of exact approaches (linear programming (LP) and constraint programming (CP)) and metaheuristic approaches (simulated annealing (SA) and tabu search (TS)). The initial solutions are constructed by combining the exact methods, and then the metaheuristics start to improve the solutions. Akjiratikarl, Yenradee and Drake [11] present particle swarm optimisation (PSO) and local improvement procedures to solve five instances (12 nurses, 50 patients, and more than 100 visits). Bredström and Rönnqvist [12] introduce a mathematical programming model that employs additional synchronisation and precedence constraints. They apply a local branching heuristic approach to solve small to medium size test instances.

Rasmussen, Justesen, Dohn and Larsen [13] model the HHCSP as a SPP with additional constraints, and the problem is generalised as a VRPTW. The branch-and-price (BP) algorithm is proposed as an exact solution approach. The objective function consists of the total travelling cost, uncovered visits, and the preferences of the home carers. The constraints include time windows and a

generalised precedence employed to model the different types of temporal dependencies. In order to decrease the dimension of the problem and reduce the run time, a visit-based clustering arrangement is devised, depending on the preferences. However, the use of cluster analysis based on preferences is less effective than the geographical location-based clustering.

Castillo-Salazar, Landa-Silva and Qu [14] introduce a literature survey of workforce rostering and routeing problems. They investigate the characteristics of problem-solution approaches. The authors propose a model that is based on the same model as in the work done by Bredström and Rönnqvist [12].

Trautsamwieser and Hirsch [15] propose a variable neighbourhood search (VNS) approach to optimising the daily scheduling of Austrian Red Cross (ARC) home carers. Their real-world examples cover urban areas (13 nurses, 140 clients, and 140 tasks) and suburban (75 nurses, 420 clients, and up to 512 tasks) areas. The objective is to minimise the travelling and waiting time, and to cover the preferences of clients and nurses. The constraints are working time and breaks. Their proposed method is able to solve small generated examples (four nurses and 20 tasks) optimally. Hiermann, Prandtstetter, Rendl, Puchinger and Raidl [7] present a multi-modal HHCRSP model (driving a car, using public transport) to allocate nurses to the clients of an Austrian HHC provider. Moreover, while Trautsamwieser and Hirsch [15] use time-dependent travelling time, Hiermann *et al.* [7] use stochastic travelling time estimations. In both of the mathematical models stated above, no synchronisation constraints exist. The objective of Hiermann *et al.* [7] is to minimise the deviation of different constraints; however, computational comparison with this work is a challenging issue [16]. As in the Trautsamwieser and Hirsch [15] work, Hiermann *et al.* [7] propose a two-step heuristic solution. Firstly, the initial solutions are generated randomly or via CP. Secondly, these generated solutions are improved by employing different heuristics such as VNS, memetic algorithm (MA), scatter search (SS), and simulated annealing hyper-heuristic (SAHH). In local search procedures, a VND is embedded in VNS, and the neighbourhood structure comprises three movements: swapping nurses, swapping jobs, and repositioning jobs. Moreover, the authors employ a dynamic neighbourhood change/reordering (in VND) based on the improvement of the objective function value.

In a more recent study, Fikar and Hirsch [17] developed a solution procedure for daily schedules by using a dial-a-ride-problem (DARP). In this work, more flexibility is offered by walking to the clients' homes. The objective is to minimise driving-working time; and the constraints are time windows, qualification level, interdependencies, and the capacity of vehicles. They offer a two-stage solution approach that is initialised by building a walking route using the SPP, and then optimising a generated solution via a TS heuristic. The proposed solution procedure is evaluated with a real-life set of data provided by the ARC.

Only small instances (four nurses and 20 clients [15], nine nurses and 45 clients [18]) can be solved optimally; hence, to deal with larger instances, a heuristic approach is required in this framework.

Our solution approach starts with the cluster-assign procedure based on time and location; this leads to improvements in the objectives, and does not sacrifice optimality – as happens with Rasmussen *et al.* [13]. To our knowledge, this procedure is the first in the field of HHC to consider the location of clients and nurses; more than ten different working time intervals; the preferred starting time of jobs; and the distance between a pair of clusters. Furthermore, within this procedure, the TSC algorithm can yield the number of clusters automatically. This makes our cluster-assign procedure free from parameter-tuning processes. Rasmussen *et al.* [13] developed a model in which each client demands one job. On the other hand, Hiermann *et al.* [7] define only the formulation of the problem. In our proposed multi-modal mathematical model, clients can demand more than one job with different competence levels, different durations, and different time intervals. Although the developed model does not involve the term 'the number of nurses', as mentioned before, the proposed approach covers both construction and improvement strategies to reduce the number of nurses. Instead of directly defining a term in the mathematical model, we tackle this issue by using a series of newly-developed neighbourhood strategies.

The proposed solution approach has the following differences from, or improvements on, the work of Hiermann *et al.* [7]:

- The first stage of our proposed solution approach starts with a clustering algorithm, and then the assignment is carried out by considering the location, time intervals, and travelling distance.
- At the beginning of the second stage, we generate the initial solutions by using three different construction strategies: deterministic, random, and semi-random. Construction strategies do not necessarily cover all the jobs.
- A series of newly-developed neighbourhood strategies proposes a higher diversification ability, while the work of Hiermann *et al.* [7] only employs three neighbourhood strategies.
- We employ the VND algorithm, which has two parameters: the order of neighbourhood structure, and the maximum number of iterations in the second stage. However, Hiermann *et al.* [7] used different heuristics that have more than two parameters, making it necessary to apply parameter tuning processes that result in increasing the run time.
- Synchronisation is also considered in our proposed model.
- We focus on large-scale problems in the field of HHC.
- Although Hiermann *et al.* [7] defined the time windows as soft constraints, in our case a nurse is not allowed to start a job before the time window, and must start a job within the time window.

### 3 PROBLEM FORMULATION

In this section, a mathematical model and its components representing the HHCRSP will be formulated. The notations for the mathematical model are summarised in Table 1.

Table 1: Notation of the model

Notation	Definition
<b>Sets</b>	
$J=\{1,\dots,J\}$	Set of jobs
$N=\{1,\dots,N\}$	Set of nurses
$JA_n=J \cup \{sl_n,el_n\}$	Set of all possible jobs for the nurse $n$
$P$	Set of synchronised jobs
<b>Parameters</b>	
$sl_n,el_n$	Artificial jobs represent starting and ending locations of nurse $n$
$na_n, nb_n$	Starting and ending working time of nurse $n$
$ja_i, jb_i$	Starting and ending time of job $i$
$pt_i$	Preferred starting time of the job $i$
$s_{ijn}$	Nurse $n$ 's travelling time from job $i$ to job $j$ by using her/his car
$s'_{ijn}$	Nurse $n$ 's travelling time from job $i$ to job $j$ by using public transportation
$d_i$	Duration of job $i$
$jq_i, nq_n$	Competence level of job $i$ and nurse $n$ , respectively
$nr_n$	Characteristics of nurse $n$
$jr_i$	Characteristics of job $i$ (for each client)
$TR_n$	The mode of transportation of nurse $n$
$M$	Big number
<b>Variables</b>	
$c_{ijn}$	Nurse $n$ 's time costs travelling from job $i$ to job $j$ including the duration of job $i$ and idle time
$x_{ijn}$	Binary decision variable, 1 if nurse $n$ moves to job $j$ after carrying out job $i$ ; 0 otherwise
$ux_i$	Binary decision variable, 1 if a job is uncovered; 0 otherwise
$t_{in}$	Nurse $n$ 's starting time of the job $i$
$o_n$	Overtime work of a nurse $n$
$t_n$	Nurse $n$ 's completion time of the last assigned job
$z_i$	The deviation time from the preferred starting time of the job $i$
$v_{in}$	Nurse $n$ 's arriving time to the job $i$
$e_{in}$	Nurse $n$ 's idle time of the job $i$
$y_{ij}$	Binary decision variable, 1 if synchronised jobs $(i,j)$ are uncovered; 0 otherwise

The set of nurses and jobs are denoted by  $N = \{1, \dots, N\}$  and  $J = \{1, \dots, J\}$  respectively.  $JA_n$  is the set of all possible jobs for the nurse  $n$ . In order to perform some of the jobs, proper coordination may be required. Hence,  $P$  is defined as the set of synchronised jobs. The home location of nurses is the starting and ending point of their routes. Thus we introduce two artificial jobs for each nurse. These artificial jobs are  $\{sl_n, el_n\}$ , which represent the starting and ending home locations of nurse  $n$ . The working time interval of each nurse is defined by the working time window  $[na_n, nb_n]$ , where nurses cannot begin jobs before  $na_i$  and should finish the assigned job(s) before  $nb_i$ . If any nurse starts the service before the end of working hours, s/he can continue to work. The clients may demand more than one job represented in the set. For each job  $i \in J$ , a time window is defined by  $[ja_i, jb_i]$ , where  $ja_i$  and  $jb_i$  correspond to the starting and ending time of the job  $i$  respectively. In addition, for each job  $i$ , the preferred starting time  $pt_i \in [ja_i, jb_i]$  and the duration (service time) of the job  $d_i$  are determined. The time window of the artificial jobs equals the working time windows  $[na_{sl_n}, nb_{sl_n}] = [na_{el_n}, nb_{el_n}] = [na_n, nb_n]$ , and the duration and preferred starting time of the artificial jobs are set at zero.

The travelling time between two jobs is based on the parameters of either  $s_{ijn}$  (car) or  $s'_{ijn}$  (public transport). Moreover, the travelling time between job  $sl_n$  and job  $j$  and the travelling time between job  $j$  and job  $el_n$  are defined as zero for the nurse set. The mode of transportation (car or public transport) of each nurse  $n$  is defined by  $TR_n$ . If any nurse prefers to use public transport, then  $TR_n$  is one; otherwise, it becomes zero. Each nurse must use her/his own modality, and there is no switch while travelling between jobs. There are five different competence levels, defined from one to five, representing the qualifications of the nurses ( $nq_n$ ). In order to perform a job, it is necessary to satisfy the required level of competence. Nurses can only carry out a job as long as they have a higher or equal level. Furthermore, nurses and clients have some features such as gender, smoking habits, and pet ownership. These features must be considered when creating schedules.

The time cost ( $c_{ijn}$ ) consists of travelling time ( $s_{ijn}$  or  $s'_{ijn}$ ) based on the transportation mode; the duration of job  $i$  ( $d_i$ ); and idle time ( $e_{in}$ ) of nurse  $n$  if s/he arrives before  $ja_i$ . For the mathematical model, the following decision variables are defined: the binary routeing variables  $x_{ijn} \in \{0, 1\}$ ; the binary coverage variables  $ux_i \in \{0, 1\}$ ; and the scheduling variables  $t_{in}$ ,  $t_n$ , and  $o_n$ . If nurse  $n$  travels to job  $j$  after handling job  $i$ , a binary variable  $x_{ijn}$  is one; otherwise it is zero. The binary coverage variable is one if job  $i$  is uncovered within the solution; otherwise it is zero. The scheduling variable  $t_{in}$  is nurse  $n$ 's starting time for job  $i$ .  $v_{in}$  is defined as nurse  $n$ 's arriving time for job  $i$ . Nurse  $n$ 's completion time of the last assigned job and her/his overtime work are denoted as  $t_n$ , and  $o_n$  respectively.

The general formulation of the HHCSRPTW is equal to the uncapacitated multi-depot VRPTW [12]. Some of the modifications that cover the topics, such as synchronisation, uncovered visits [13], multimodality, and preferences [7], are added to our proposed model.

The objective function is to minimise the total time costs, uncovered visits, overtime work, and deviation from the preferred time of the job (constraint (1)). The objective function terms are prioritised by assigning time weights ( $w_1$  and  $w_2$ ), which are  $J$  (the number of jobs) and  $J/2$  respectively. Constraint (2) means that visits are either covered or left unassigned; and similarly constraints (3) and (4) are defined to calculate uncovered synchronised visit constraints ( $M$  is considered a big number). Nurses can only perform the job if the competence level is satisfied by constraint (5). Constraints (6) and (7) ensure that the nurses start and finish their jobs from their own home. Constraint (8) is introduced to maintain the flow. The time windows must be respected (constraints (9) and (10)), and travelling times are considered via constraint (11). The synchronisation is introduced by constraint (12). Constraint (13) guarantees that a synchronised job can be performed by different nurses. The overtime work of the nurses is calculated via constraint (14). Constraint (15) ensures the characteristics of nurse-client match. Constraints (16) and (17) express a piecewise linear function of the deviation from the starting time of the activity. Constraint (18) determines the arriving times of nurses according to the starting time of the previous jobs, the travelling times based on the transportation mode, and the duration of the jobs. The idle times of

nurses are computed via constraints (19) and (20). Similarly, time cost (constraint (21)) is defined as the summation of durations, the travelling times based on transportation mode, and idle times. Constraint (22) is a non-negativity constraint. Constraints (23) to (27) set the domains of the decision variables.

$$\text{Min. } \sum_n^N \sum_i^{JA} \sum_j^{JA} c_{ijn} x_{ijn} + w_1 \sum_i^J ux_i + w_2 \sum_n^N o_n + \sum_i^J z_i \quad (1)$$

$$\text{st. } \sum_n^N \sum_j^{JA} x_{ijn} + ux_i = 1 \quad \forall i \in J \quad (2)$$

$$ux_i + ux_j \leq My_{ij} \quad \forall (i, j) \in P \quad (3)$$

$$-ux_i - ux_j \leq M(1 - y_{ij}) \quad \forall (i, j) \in P \quad (4)$$

$$\sum_j^{JA} x_{ijn} j q_i \leq n q_n \quad \forall n \in N, \forall i \in J \quad (5)$$

$$\sum_j^{JA} x_{sl_n j n} = 1 \quad \forall n \in N \quad (6)$$

$$\sum_i^{JA} x_{iel_n n} = 1 \quad \forall n \in N \quad (7)$$

$$\sum_i^{JA} x_{ign} = \sum_j^{JA} x_{gjn} \quad \forall n \in N, \forall g \in J \quad (8)$$

$$j a_i \sum_j^{JA} x_{ijn} \leq t_{in} \leq j b_i \sum_j^{JA} x_{ijn} \quad \forall n \in N, \forall i \in J \cup \{sl_n\} \quad (9)$$

$$n a_n \leq t_{in} \leq n b_n \quad \forall n \in N \quad (10)$$

$$t_{in} + s_{ijn} TR_n x_{ijn} + s'_{ijn} (1 - TR_n) x_{ijn} + d_i x_{ijn} \leq t_{jn} + j b_i (1 - x_{ijn}) \quad \forall n \in N, \forall i, j \in JA \quad (11)$$

$$\sum_n^N t_{in} = \sum_n^N t_{jn} \quad \forall (i, j) \in P \quad (12)$$

$$\sum_l^{JA} x_{lin} + \sum_l^{JA} x_{ljn} \leq 1 \quad \forall (i, j) \in P, \forall n \in N \quad (13)$$

$$o_n \geq (t_n - n b_n) \quad \forall n \in N \quad (14)$$

$$\sum_j^{JA} x_{ijn} (n r_n + j r_j) \leq 1 \quad \forall n \in N, \forall i \in J \quad (15)$$

$$\left( \sum_n^N t_{in} - p t_i \right) \leq z_i \quad \forall i \in JA \quad (16)$$

$$-\left( \sum_n^N t_{in} - p t_i \right) \leq z_i \quad \forall i \in JA \quad (17)$$

$$v_{in} = t_{kn} + d_k + s_{kin} TR_n + s'_{kin} (1 - TR_n) \quad \forall n \in N, \forall i \in J, \forall k \in J \cup \{sl_n\} \quad (18)$$

$$e_{in} \geq t_{in} - v_{in} \quad \forall n \in N, \forall i \in J \quad (19)$$

$$e_{in} \geq 0 \quad \forall n \in N, \forall i \in J \quad (20)$$

$$c_{ijn} = d_i + s_{ijn}TR_n + s'_{ijn}(1 - TR_n) + e_{in} \quad \forall n \in N, \forall i, j \in J \quad (21)$$

$$o_n \geq 0 \quad \forall n \in N \quad (22)$$

$$x_{ijn} \in \{0, 1\} \quad \forall n \in N, \forall i, j \in JA \quad (23)$$

$$u_{xi} \in \{0, 1\} \quad \forall i \in JA \quad (24)$$

$$y_{ij} \in \{0, 1\} \quad \forall (i, j) \in P \quad (25)$$

$$t_n, t_{in} \in N \quad \forall i \in JA, \forall n \in N \quad (26)$$

$$z_i \in N \quad \forall i \in JA \quad (27)$$

## 4 SOLUTION APPROACH

In order to deal with the large-scale HHCRSP, a two-stage solution approach is developed, based on the ‘cluster-first route-second’ approach (Figure 1). Each component of the framework will be presented in this section in turn.

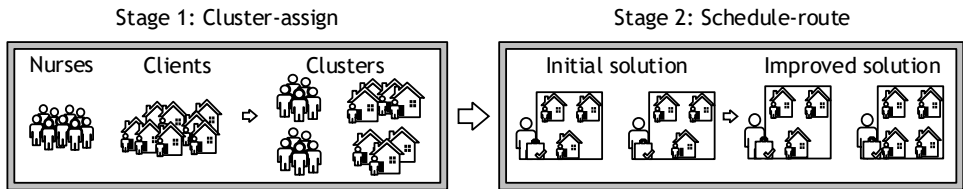


Figure 1: A two-stage solution approach

### 4.1 The cluster-assign procedure (Stage 1)

‘Cluster analysis’ or ‘clustering’ is the process of grouping a given set of data into classes. The aim of this process is to consider similar items within the same group and those that are dissimilar to the ones in the other groups [19]. It can also be used to decrease the size of the problem. In our case, the cluster-assign procedure is developed for this perspective.

- Apply the two-step cluster (TSC) algorithm depending on the geospatial data, and construct the clusters for the nurse ( $c_{nurse}$ ) and job ( $c_{job}$ ) sets automatically.
- Calculate the minimum number of cluster sets ( $\min \{c_{nurse}, c_{job}\}$ ) and re-construct the maximum number of cluster sets for the equalisation of each cluster ( $c_{nurse}=c_{job}$ ) via TSC.
- Compute the appropriate time interval between the clusters of jobs and the clusters of nurses, considering the average working time interval (for the clusters of nurses) and the average preferred time (for the clusters of jobs).
- Compute the travelling time between each cluster.
- Assign the clusters of jobs to the clusters of nurses for an appropriate time interval and travelling time.

#### 4.1.1 Two-step cluster analysis

In the proposed solution approach, TSC analysis is chosen to identify groups of homes in a city according to location. TSC was proposed by Chiu, Fang, Chen, Wang and Jeris [20] as a method that can be handled with large-scale data, and without considering categorical and continuous variables. Furthermore, one of the benefits of the TSC is to assign the number of clusters automatically. The TSC algorithm consists of two steps: pre-clustering and clustering. In the first step, each case of the data set is scanned, and then a decision is made about whether the case is to merge with the previously created clusters, or whether to construct a new cluster based on the distance criterion. In this step, the cluster feature (CF) tree is constructed to hold the summary statistics of the pre-clusters. This also helps to identify the pre-clusters. One of the advantages of the first step is to reduce the dimension of the initial data by constructing a new data matrix that has fewer cases [20, 21]. In the second step of the algorithm, each pre-cluster is considered as an entity, and an agglomerative hierarchical clustering method is applied to the pre-clusters. This method uses a bottom-up strategy, and begins with each of the sub-clusters obtained in the first step. The

algorithm then merges the two closest clusters into a single cluster until all clusters are members of the same cluster [22]. The number of clusters is automatically determined in the second step by considering the lowest Bayesian information criterion (BIC) and the highest ratio of distance measure [20, 21]. After creating the cluster, the next step is to generate the schedules via a metaheuristic approach.

## 4.2 The schedule-route procedure (Stage 2)

The next stage is the generation of the initial solution. Improving the obtained solution will be discussed later (Figure 1).

### 4.2.1 Variable neighbourhood search

A VNS is a metaheuristic approach developed by Mladenovic and Hansen [23]. VNS is based on the concept of systematic neighbourhood change that is proposed not only to find a local minimum, but also to avoid being stuck in the area of local minima. VNS advances by using a descent method to reach a local minimum. VND [24] is a version of the VNS method that is employed for improvement. The difference between VNS and VND is the deterministic neighbourhood change: VND does not cover the shaking step, whereas the basic VNS method does. The general scheme of VNS offers simplicity, robustness, and user-friendliness [24]. VND has two parts: the generation of the initial schedule, and the neighbourhood structure, as explained below.

#### i. Generation of initial schedule

Our solution approach starts with an heuristic that selects jobs in some of the predetermined orders until a solution has been obtained. Jobs are assigned to nurses (routes) based on some criteria by satisfying a set of constraints such as time windows, competence level, etc. [25]. The proposed construction heuristics do not necessarily cover all the jobs at the beginning.

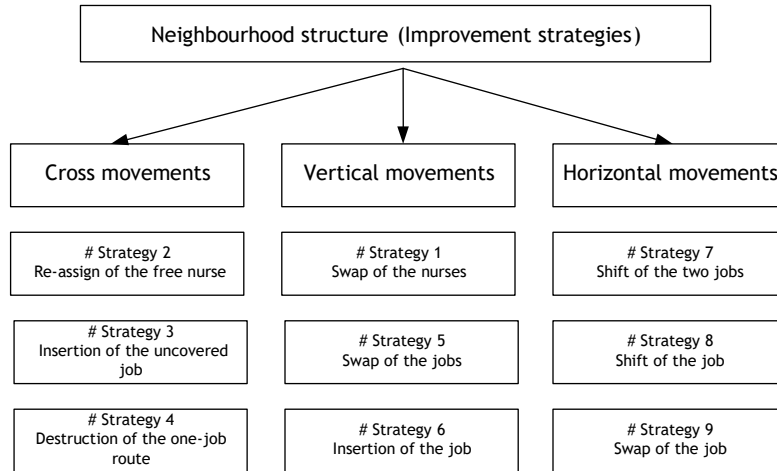
In order to generate an initial schedule, three different approaches are developed in the proposed method. Construction strategy (CS) 1 starts with the smallest index-employee and tries to assign the smallest index-job to this employee if the set of constraints is satisfied. It can be said that CS 1 generates solutions deterministically. CS 2 is the second strategy, and randomly assigns selected jobs to randomly selected nurses. In this way, it can be viewed as a random-solution generation heuristic. The last strategy is CS 3, which can be considered a hybrid of the first two strategies. It starts with the smallest index-nurse and assigns randomly selected jobs to it.

#### ii. Neighbourhood structures

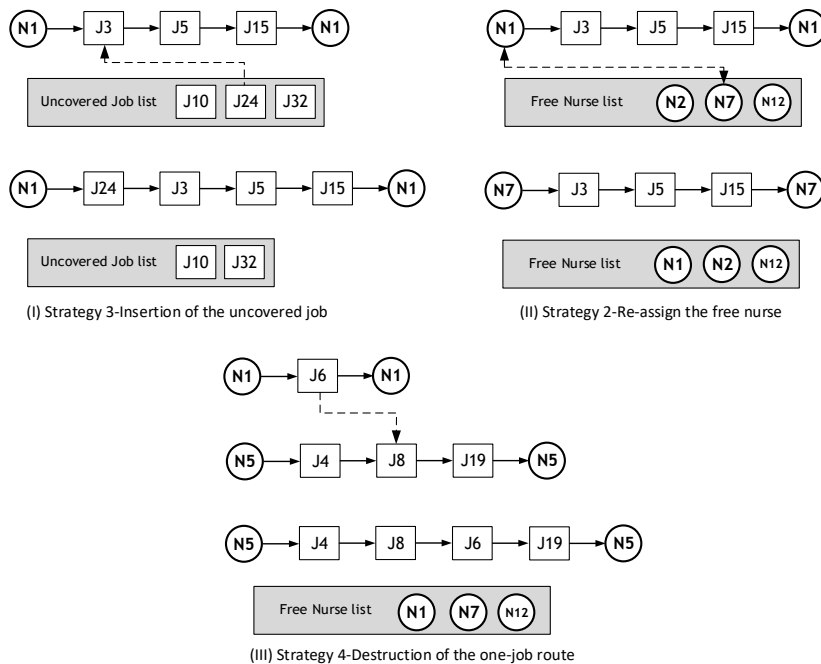
Different neighbourhood structures are developed to improve the solution constructed in the initialisation phase. These can be classified into three groups: cross, vertical, and horizontal, according to the structure of the HHCRSP (Figure 2). The neighbourhoods are ordered from the smallest to the largest, based on the number of the neighbourhoods [26]. The order of this structure numbers each strategy numerically from 1 to 9. The best solutions are generated by this order of neighbourhoods in the initial runs.

*Cross movements:* This group covers a series of newly developed neighbourhood structures. The aim of the newly developed cross movement strategies is to achieve an improved ability to diversify. Once the initial schedule is generated, some of the jobs might be left unassigned, and similarly some of the nurses might not be a part of the schedule. Thus, in order to cover all the jobs and to use the human resources efficiently, three movements or neighbourhoods are developed (Strategies 3, 2, and 4). These are illustrated in Figure 3, where circular and square nodes correspond to the nurses and the jobs respectively. Strategy 3 chooses one of the jobs randomly from the uncovered job list, and assigns this job to a randomly selected route/nurse. The aim of Strategy 2 is to benefit more efficiently from the human resources. When the free nurse list has an element, Strategy 2 starts to work. A randomly selected nurse from the free nurse list is exchanged with a randomly selected working nurse if there is a reduction in the objective function value. Strategy 4 is employed to remove one-job routes. This strategy looks for one-job routes and checks whether or not the element (job) of the route can be inserted into randomly selected multi-job routes. If this movement results in a reduction in the objective function value, the removal of the one-job route takes place during this step. Strategy 4 is developed to cover all of the jobs with the minimum number of nurses.





**Figure 2: Neighbourhood structure**



**Figure 3: Cross movements: I-Strategy 3, II-Strategy 2, and III-Strategy 4**

*Vertical movements:* This group consists of three strategies: Strategy 1, Strategy 5, and Strategy 6, as illustrated in Figure 4. Strategy 1 is proposed to swap two randomly selected employees if there is an improvement in the objective. Similar to Strategy 1, Strategy 5 is developed to swap these two randomly selected jobs. In this vertical movement, the locations of the two are also chosen by chance. The last movement in this group is Strategy 6, which employs insertion.

*Horizontal movements:* This group covers the basic neighbourhood structures: swap and shift. They are developed to find the best sequence/location of jobs. These locations are chosen by chance, and if there is an improvement, the swap occurs in the schedule during the Strategy 9 process. The difference between Strategy 9 and Strategy 8 is the shift mechanism; the working principle is the same for this improvement perspective. The last member of the horizontal movements is Strategy 7; this is the two-job shift version of Strategy 8. These strategies are illustrated in Figure 5.

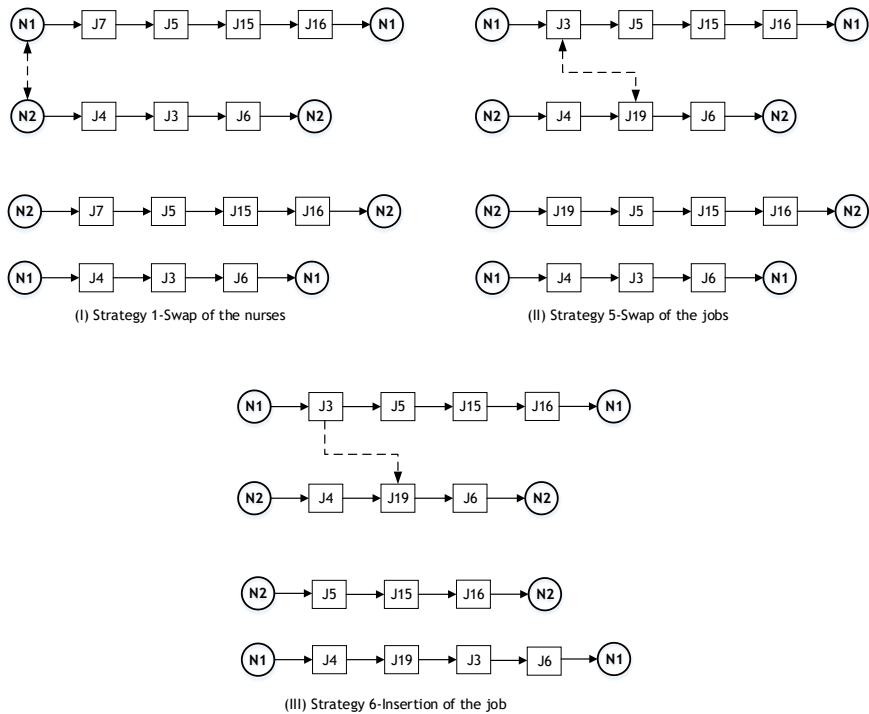


Figure 4: Vertical movements: I-Strategy 1, II-Strategy 5, and III-Strategy 6

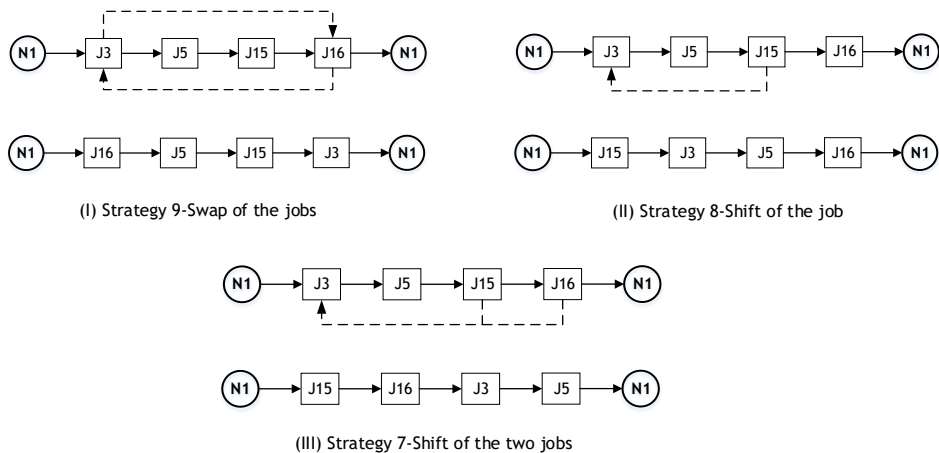


Figure 5: Horizontal movements: I-Strategy 9, II-Strategy 8, and III-Strategy 7

### 4.3 Test instances

Initially, the proposed solution approach is tested on the four instances (Days 1, 2, 3, and 4) provided by Hiermann *et al.* [7]. (The test instances and their detailed explanation can be found at [https://www.ads.tuwien.ac.at/w/Research/Problem\\_Instances](https://www.ads.tuwien.ac.at/w/Research/Problem_Instances).) No benchmark data is available, so we have generated a series of large-scale instances in accordance with these publicly available instances. Moreover, because of the lack of synchronisable jobs, five per cent of the jobs are added for each instance. The data instances involve more than 10 different working time windows and varied specified time windows for the jobs. The time in these instances is represented by a discrete time unit, which is set at 5. Assume that a nurse travels from one customer to another; this then takes six time units, which is a total of 30 minutes. The time windows are also computed in a similar manner.

## 5 COMPUTATIONAL RESULTS

In this section, the proposed solution approach is run on the four publicly available instances, and the results are compared with the solution method used by Hiermann *et al.* [7]. The proposed algorithm is then run on the series of generated large-scale instances. As stated in Cissé, Yalçındağ, Kergosien, Şahin, Lenté and Matta [27], there is no benchmark instance in the literature: all the published papers have used different models and different instances. In this study, we modified the VNS solution approach of Hiermann *et al.* [7] to solve our instances and compare them with the proposed solution approach.

The algorithm is run 10 times for all test instances on an Intel i7-5500U CPU 2.4GHz PC with 8GB RAM. The maximum number of iterations is set to 30,000 for each run. In the first stage of the proposed method, the cluster-assign procedure runs and forms the clusters. The second stage of the proposed method is to generate schedules and routes for each cluster with a VND heuristic, in which nine different neighbourhood structures are employed for a higher diversification ability.

The computational comparisons with the work of Hiermann *et al.* [7] are a challenge [16], since the comparisons have been carried out based on our model and considerations for the HHCRSP. In order to investigate the efficiency of the proposed VND approach, the computational results will be compared with the VNS-VND embedded in the three dynamic neighbourhood strategies proposed by Hiermann *et al.* [7]. Although they generated initial solutions via CP and a random construction heuristic, here the VNS-VND algorithm starts with the three proposed construction strategies. The performance measures by which the solutions are evaluated are the number of uncovered jobs and assigned (or working) nurses, the total time costs, the total overtime work, and the deviation from the preferred time. The average computational results and the details of these results are presented in Table 2 and Table 3 respectively. The objective function and its components, the number of nurses and jobs, the number of working and free nurses, and the number of clusters for each instance are represented in rows. The best solution for each instance is shown in bold.

Although using different neighbourhood structures increases the run time, the proposed approach provides better results than the VNS-VND method in terms of objective function value for all instances. When the initial solution strategies are considered, CS 3 uses a smaller healthcare workforce for each instance. On the other hand, CS 1 and 2 schedule more nurses than CS 3, depending on the deterministic and random insertion mechanisms.

When Table 2 is investigated further, it can be said that all the jobs are a part of the schedule. Thus the total duration time is fixed, and there is no room for the minimisation of this term. In order to minimise the total time costs, we should deal with the travelling time components. The proposed solution approach yields reduced total time costs, which means that the assigned nurses take less time. This improvement is accomplished because of the reduced total travelling time and idle (waiting) time. The VNS-VND approach of Hiermann *et al.* [7] generates schedules for the HHCRSP with more working time than the results presented in this paper. Similarly, the solutions of the VNS-VND approach propose more overtime work than the proposed method.

The last measure of quality of the results is the preferred time deviation. It can easily be seen that our solution approach causes higher time deviations for a number of reasons. First, the initial aim is to complete all of the sets of jobs with the minimum number of nurses, causing an increase in the deviation of time. That is, the generated schedules assign more jobs to fewer nurses. At first glance, this solution approach seems to increase the nurse-to-patient ratio; however, it considers the safety issues for the HHC system. This ratio should not exceed 1:8 [28, 29] for hospitals. Moreover, if the jobs are regarded as customers, or if we convert the nurse-to-patient ratio to a nurse-to-job ratio, this ratio does not exceed 1:8 for the HHC system. Secondly, time windows for the jobs are not considered as soft constraints. In other words, the nurses must wait before the opening time window of a job, and then continue to work if they start to complete the job before the closing time window.

The results of the proposed approach employ the minimum number of nurses; in fact, it also does not send overqualified nurses to the clients. Indeed, the proposed approach – including especially CS 3 – is the leading method for assigning the minimum number of nurses. The distribution of working nurses for four days according to their competence is illustrated in Table 4. The results mean that our approach provides more nurse satisfaction than the VNS-VND approach.

**Table 2: Comparison results**

	VNS-VND [7] (3 dynamic strategies)			The proposed method (VND-9 strategies)		
	Construction strategy					
	CS 1	CS 2	CS 3	CS 1	CS 2	CS 3
<b>Day 1</b>						
Total time costs	9,546.5	9,632.7	9,877.8	8,933.7	8,963	8,977
Total deviation from preferred time	16,702.2	17,279.5	15,893.7	20,409.6	20,882.7	21,882.1
Total overtime	73.8	52.6	252.5	28.4	11.2	18.2
No. of nurses	499	499	499	499	499	499
No. of free nurses	222.2	219.1	328.2	339.8	343.4	372.4
No. of working nurses	276.8	279.9	170.8	159.2	155.6	126.6
No. of jobs	781	781	781	781	781	781
No. of clusters	4	4	4	4	4	4
Run time (s)	129.31	136.25	132.54	168.21	173.74	174.87
Objective function	44,661.8	40,035.9	88,770.3	36,429.1	<b>32,640.1</b>	35,400
<b>Day 2</b>						
Total time costs	8,381.9	8,442.5	8,804.1	7,851.6	7,848	7,829.2
Total deviation from preferred time	13,773.3	15,104.4	15,628.9	17,998.6	18,217.3	19,837.2
Total overtime	62.2	83.6	180.9	4.4	18.7	12.8
No. of nurses	518	518	518	518	518	518
No. of free nurses	238.1	249.2	363.8	366.4	365.4	402
No. of working nurses	279.9	268.8	154.2	151.6	152.6	116
No. of jobs	703	703	703	703	703	703
No. of clusters	3	3	3	3	3	3
Run time (s)	129.98	136.50	135.12	173.17	176.31	178.49
Objective function	38,265.0	45,199.3	71,286.1	<b>26,989.8</b>	30,908.6	30,981.6
<b>Day 3</b>						
Total time costs	9,560	9,610.8	9,847	8,950.9	8,991.4	8,967.9
Total deviation from preferred time	16,991.5	17,430.3	18,003.3	20,983.5	20,341.2	22,516.1
Total overtime	93.8	73.6	250.8	51.5	49	55.4
No. of nurses	493	493	518	493	493	493
No. of free nurses	222.1	221.5	349.5	341.4	328.2	371.4
No. of working nurses	270.9	271.5	168.5	151.6	164.8	121.6
No. of jobs	779	779	779	779	779	779
No. of clusters	4	4	4	4	4	4
Run time (s)	126.95	133.00	131.56	171.09	175.37	174.93
Objective function	49,673.2	45,183.5	89,672.5	42,629.2	<b>41,411.1</b>	45,140.1
<b>Day 4</b>						
Total time costs	9,190.8	9,163.1	9,136.2	8,476	8,518.7	8,528.6
Total deviation from preferred time	17,277.4	17,355.8	17,378.2	20,505	20,551.3	23,630.7
Total overtime	86.9	75.5	93.1	36.7	24.7	83.3
No. of nurses	509	509	509	509	509	509
No. of free nurses	246.8	238.8	236.5	345.7	351.2	392.1
No. of working nurses	262.2	270.2	272.5	163.3	157.8	116.9
No. of jobs	781	781	781	781	781	781
No. of clusters	4	4	4	4	4	4
Run time (s)	128.34	145.75	138.27	177.05	180.99	203.55
Objective function	48,584.3	45,733.7	50,208.4	38,321.2	<b>35,356.2</b>	53,359.2

**Table 3: Details of computational results**

	VNS-VND [7] (3 dynamic strategies)			The proposed method (VND-9 strategies)		
	CS1	CS2	CS3	CS1	CS2	CS3
<b>Day 1</b>						
Objective (Average of ten runs)	44,661.8	40,035.9	88,770.25	36,429.1	<b>32,640.1</b>	35,400
Objective (Best of ten runs)	32,256	33,121.5	63,317.5	29,940	27,409	31,193.5
Std deviation	9,283.69	5,145.38	24,201.75	45,58.04	3,280.04	3,306.57
<b>Day 2</b>						
Objective (Average of ten runs)	38,265	45,199.30	71,286.1	<b>26,989.8</b>	30,908.6	30,981.6
Objective (Best of ten runs)	30,093	28,358	43,127	24,362	25,698	26,819
Std deviation	7,125.09	12,851.97	19,238.1	20,36.93	7,474.77	5,923.48
<b>Day 3</b>						
Objective (Average of ten runs)	49,673.2	45,183.5	89,672.5	42,629.15	<b>41,411.1</b>	45,140.1
Objective (Best of ten runs)	41,869.5	38,472.5	76,359.5	33,492	32,957	36,553
Std deviation	5,355.02	54,48.62	8,997.79	81,19.76	9,151.69	6,546.44
<b>Day 4</b>						
Objective (Average of ten runs)	48,584.25	45,733.65	50,208.35	38,321.15	<b>35,356.15</b>	53,359.15
Objective (Best of ten runs)	41,732	35,956	34,162	27,106	28,353.5	36,688
Std deviation	6,875	7,095.73	9,627.7	7,297.73	4,661.27	14,492.11

**Table 4: Competence level of working nurses**

	Day 1					Day 2					
	All nurses					All nurses					
Competence	1	2	3	4	5	1	2	3	4	5	
No. of nurses	4	23	384	37	51	2	27	399	41	49	
	Working nurses					Working nurses					
	CS 1	0	2.3	221.7	21.8	31	0	4.1	216.6	26.2	33
VNS-VND [7]	CS 2	0.2	2.6	221.3	22	33.8	0	3.9	205.3	25.4	34.2
	CS 3	0	1.5	135	13.9	20.4	0	3.3	119.2	13.4	18.3
	Working nurses					Working nurses					
	CS 1	0	1	124	12.3	21.9	<b>0</b>	<b>0.9</b>	<b>110.9</b>	<b>17.1</b>	<b>22.7</b>
The proposed method	CS 2	<b>0.1</b>	<b>0.6</b>	<b>119.1</b>	<b>13.4</b>	<b>22.4</b>	0	0.9	111.5	18.1	22.1
	CS 3	0	0.8	93.3	12.3	20.2	0	0.7	83.5	15.4	16.4
	Day 3					Day 4					
	All nurses					All nurses					
Competence	1	2	3	4	5	1	2	3	4	5	
No. of nurses	5	24	382	39	43	5	30	375	43	56	
	Working nurses					Working nurses					
	CS 1	0	3.9	214	25.6	27.4	0	5.8	199.3	25.3	31.8
VNS-VND [7]	CS 2	0	3.8	214.6	25.9	27.2	0	4.8	207.2	24.5	33.7
	CS 3	0	4.1	127.7	18.3	18.4	0	4.9	207.1	26.4	34.1
	Working nurses					Working nurses					
	CS 1	0	0.6	119.8	15	16.2	0	3.4	123.4	15.7	20.8
The proposed method	CS 2	<b>0</b>	<b>0.9</b>	<b>128.5</b>	<b>17</b>	<b>18.4</b>	<b>0</b>	<b>1.8</b>	<b>118</b>	<b>14.9</b>	<b>23.1</b>
	CS 3	0	0.4	92.7	12.3	16.2	0	1.9	85	13.7	16.3

In order to investigate the efficiency of the newly developed cross movement structures, more computations have been carried out. The algorithm was tested on the Day 1 instance; the results are compared in Figure 6. Here, in case the initial construction strategies cannot cover all of the jobs, only Strategy 6 will run. It is obvious that the newly developed cross movement strategies offer higher diversification and improvement without considering construction heuristics. In a similar way, the efficiency of the clustering approach is examined; the results are illustrated in Figure 7. The number of different clusters is shown on the horizontal axis, with the lines corresponding to the objective function, and the columns represent the run time. Although clustering increases the run

time, the objective function begins to improve by applying clustering. The important point here is that, to determine the appropriate number of clusters, we assign the number of clusters according to the  $\min\{c_{nurse}, c_{job}\}$  in the cluster-assign procedure; for the Day 1 instance, the number of clusters is set at four, and this four-cluster formation proposes the minimum objective function values for all the construction strategies. The same pattern is followed for the Day 2 instance. As a result, it is reasonable to apply clustering in the first stage, and the solution quality is not deteriorated by the use of clustering.

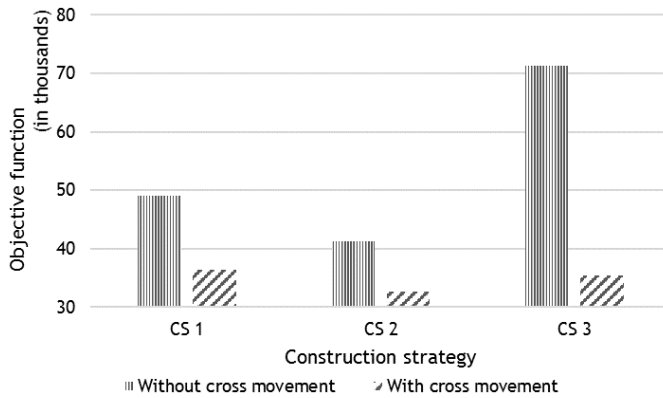


Figure 6: Efficiency of the newly developed cross movement strategies on Day 1 instance

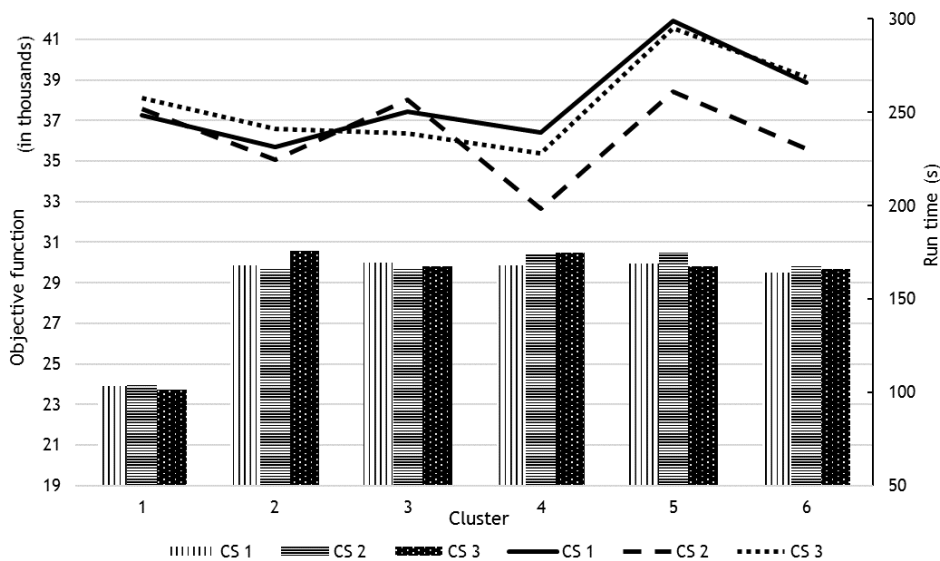


Figure 7: Efficiency of clustering on Day 1 instance

Table 5 provides an overview of the generated large-scale instances consisting of different numbers of nurses and jobs, and summarises the computational results. Synchronisation is also considered for the generated instances. Here, the algorithm covers all the jobs, and the results based on three different construction strategies represent the average value of 10 runs. For Instance 1, with the initial solution CS 1, the objective function value is calculated to be 16,653.30, the run time is 132.99 s, and the number of working nurses is 88.8. According to the objective function value, CS 2 initialisation generally yields better solutions. The algorithm initialised with CS 1 generally runs faster than the others. CS 3 constructs solutions with the minimum number of nurses.

**Table 5: Computational results**

I	No. of		Objective function			Time (S)			No. of working nurses		
	N	J	CS 1	CS 2	CS 3	CS 1	CS 2	CS 3	CS 1	C S2	CS 3
1	250	393	16,653.30	<b>16,590.30</b>	17,750.00	132.99	148.96	134.49	88.80	86.90	70.20
2	245	388	17,888.65	<b>16,664.90</b>	17,045.75	133.54	148.26	131.41	85.30	81.90	62.70
3	260	404	17,954.50	<b>17,311.10</b>	17,794.30	134.35	144.70	131.51	86.40	84.10	65.90
4	260	352	15,339.70	<b>14,481.80</b>	17,332.40	144.40	150.04	146.27	73.50	70.50	52.10
5	270	357	<b>17,747.30</b>	19,475.80	18,259.70	138.98	143.36	144.00	79.10	77.80	55.10
6	254	338	15,920.30	<b>14,688.70</b>	15,248.00	140.68	148.86	139.26	70.60	70.70	49.80
7	247	388	19,912.55	<b>15,563.70</b>	19,342.30	126.42	132.46	130.57	84.60	77.70	60.60
8	240	383	<b>16,186.10</b>	17,146.70	20,521.00	133.46	139.46	131.13	87.50	80.90	61.80
9	255	393	24,250.00	22,189.00	<b>21,123.50</b>	154.32	150.39	143.93	77.80	82.20	66.10
10	255	388	<b>18,544.15</b>	19,485.35	20,287.75	135.07	135.61	137.18	78.70	77.60	56.30
11	267	400	20,811.45	<b>17,652.70</b>	17,990.95	140.02	141.18	139.80	80.00	82.10	60.10
12	250	379	<b>16,074.00</b>	20,500.40	19,076.90	130.60	131.93	133.50	76.50	72.50	57.70
13	549	858	<b>40,060.65</b>	63,176.35	44,336.05	177.78	179.35	183.45	163.10	161.50	137.20
14	538	855	40,887.60	<b>38,830.50</b>	44,616.20	179.55	179.19	181.51	170.10	170.30	138.80
15	520	815	39,046.80	<b>35,933.90</b>	36,978.30	175.96	178.23	178.49	166.00	163.90	133.90
16	500	678	<b>28,787.00</b>	32,563.50	31,365.30	173.51	178.15	177.20	151.30	143.30	102.20
17	520	707	<b>29,885.20</b>	32,380.10	31,488.40	174.99	195.14	179.51	152.50	153.10	110.50
18	540	733	31,302.00	30,882.90	<b>30,593.90</b>	187.65	184.38	184.08	151.50	157.30	112.40
19	500	790	41,001.30	<b>35,374.50</b>	50,101.10	170.81	175.50	178.94	160.10	152.80	122.90
20	520	823	<b>35,470.30</b>	36,133.40	40,928.00	178.98	190.88	223.90	167.90	159.30	127.20
21	480	759	<b>38,479.70</b>	40,175.20	41,395.30	184.60	187.01	189.14	151.10	150.30	116.60
22	517	795	<b>37,040.00</b>	37,053.00	45,109.25	176.15	177.75	181.53	155.40	159.80	118.80
23	530	814	43,926.10	<b>38,488.40</b>	50,559.20	179.07	179.22	193.64	164.80	159.30	120.80
24	495	757	<b>35,769.90</b>	39,423.35	52,695.40	173.92	178.36	181.18	144.40	147.70	108.80
25	1000	1564	<b>72,714.00</b>	74,253.70	79,325.80	226.47	248.30	250.00	311.90	312.50	258.10
26	1020	1575	82,761.50	<b>82,647.40</b>	85,421.00	237.60	241.91	246.41	316.50	309.10	252.10
27	1017	1573	<b>96,690.20</b>	102,199.45	115,543.65	223.30	238.92	243.81	336.60	318.60	285.90
28	1030	1396	<b>63,569.40</b>	67,259.40	67,392.80	231.33	245.54	247.91	285.70	284.10	226.60
29	1025	1407	76,401.95	<b>69,503.30</b>	79,389.10	235.94	253.68	253.46	269.40	275.50	207.20
30	1030	1401	72,303.40	<b>62,447.40</b>	64,311.00	239.44	258.49	260.90	285.00	275.50	224.00
31	985	1554	84,906.25	<b>78,445.15</b>	101,058.65	227.88	242.96	247.15	269.60	260.20	231.40
32	991	1546	86,967.30	<b>83,657.45</b>	124,337.50	228.79	231.82	242.08	291.80	292.90	243.80
33	990	1549	87,135.00	<b>81,696.40</b>	97,663.50	227.97	244.19	248.90	283.40	269.40	239.30
34	1020	1564	91,432.20	89,573.80	146,297.10	229.54	248.37	249.09	274.60	263.90	230.90
35	1015	1552	74,930.65	<b>72,740.70</b>	86,071.75	244.03	265.73	279.91	294.80	285.90	232.10
36	1020	1548	<b>75,608.60</b>	82,716.60	109,084.90	231.13	245.79	252.49	297.50	291.90	227.10
37	1490	2341	129,761.40	<b>122,098.00</b>	144,863.50	560.90	558.80	563.30	444.40	460.00	383.50
38	1471	2316	<b>110,788.65</b>	117,134.35	132,666.45	572.54	565.06	560.83	455.30	454.70	373.70
39	1500	2362	<b>128,192.60</b>	133,031.20	155,599.80	546.66	575.21	563.87	422.70	473.50	375.00
40	1510	2447	126,657.20	<b>124,806.10</b>	161,009.40	526.66	566.44	568.32	467.10	479.70	402.80
41	1580	2709	172,721.30	<b>171,565.50</b>	190,027.80	591.20	637.34	641.81	468.70	468.30	435.30

Note: I, N, and J indicate instance, nurse, and job respectively.

## 6 CONCLUSION

The work introduces a two-stage solution approach for a large-scale HHCRSP. The approach integrates cluster analysis and heuristics into an optimisation framework. Some of the advantages are that it is free from optimal parameter selection in each stage, it creates feasible clusters, and has a higher diversification ability covering different strategies.

Our model considers nurses, clients, and healthcare institutions; thus there is no violation of the preferences of the clients and nurses. Only the preferred starting time of the jobs deviates. The numerical results indicate profitable outcomes. The decrease in the total time helps nurses to arrive at the clients' homes on time because of the reduced travelling time. Minimising the length of the schedules also increases the nurses' satisfaction. Furthermore, choosing an appropriate mode of transport and minimising the total travelling time leads to a lower fuel and energy consumption by vehicles and nurses. From the perspective of the healthcare institutions, they have the opportunity to create schedules with a minimum number of healthcare workers and to decrease personnel costs without being understaffed.

For future research, we intend to apply the route-first cluster-second method (based on the giant tour), which is the opposite of the approach proposed here. The extension of the daily planning horizon to the medium term is another subject for the HHCRSP. Because of the modularity of this two-stage solution approach, it can employ different clustering algorithms in the first stage, and different metaheuristics can be used for the improvement stage without ignoring the newly developed neighbourhood structures.

## REFERENCES

- [1] WHO. 1948. *WHO definition of health*. <http://www.who.int/about/definition/en/print.html> [accessed 01/06/16].
- [2] Huber, M. 2015. Health: Definitions A2, in James D. Wright (ed.), *International Encyclopedia of the Social & Behavioral Sciences (second edition)*. Oxford: Elsevier.
- [3] Song, M. & Kong, E.-H. 2015. Older adults' definitions of health: A metasynthesis. *International Journal of Nursing Studies*, 52(6): pp. 1097-1106.
- [4] Leonard, K. 2015. *Obamacare, drug prices behind health spending growth*. <http://www.usnews.com/news/articles/2015/12/02/rate-of-spending-on-health-care-is-on-its-way-up-again> [accessed 09/06/16].
- [5] Mattke, S., Klautzer, L., Mengistu, T., Garnett, J., Hu, J. & Wu, H. 2010. *Health and well-being in the home: A global analysis of needs, expectations, and priorities for home health care technology*. Santa Monica: RAND Corporation.
- [6] Maenhout, B. & Vanhoucke, M. 2013. An integrated nurse staffing and scheduling analysis for longer-term nursing staff allocation problems. *Omega*, 41(2): pp. 485-499.
- [7] Hiermann, G., Prandtstetter, M., Rendl, A., Puchinger, J. & Raidl, G. 2015. Metaheuristics for solving a multimodal home-healthcare scheduling problem. *Central European Journal of Operations Research*, 23(1): pp. 89-113.
- [8] Cheng, E. & Rich, J.L. 1998. *A home health care routing and scheduling problem*.
- [9] Eveborn, P., Flisberg, P. & Rönnqvist, M. 2006. Laps care: An operational system for staff planning of home care. *European Journal of Operational Research*, 171(3): pp. 962-976.
- [10] Bertels, S. & Fahle, T. 2006. A hybrid setup for a hybrid scenario: Combining heuristics for the home health care problem. *Computers & Operations Research*, 33(10): pp. 2866-2890.
- [11] Akjiratikarl, C., Yenradee, P. & Drake, P.R. 2007. PSO-based algorithm for home care worker scheduling in the UK. *Computers & Industrial Engineering*, 53(4): pp. 559-583.
- [12] Bredström, D. & Rönnqvist, M. 2008. Combined vehicle routing and scheduling with temporal precedence and synchronization constraints. *European Journal of Operational Research*, 191(1): pp. 19-31.
- [13] Rasmussen, M.S., Justesen, T., Dohn, A. & Larsen, J. 2012. The home care crew scheduling problem: Preference-based visit clustering and temporal dependencies. *European Journal of Operational Research*, 219(3): pp. 598-610.
- [14] Castillo-Salazar, J.A., Landa-Silva, D. & Qu, R. 2014. Workforce scheduling and routing problems: Literature survey and computational study. *Annals of Operations Research*, pp. 1-29.
- [15] Trautsamwieser, A. & Hirsch, P. 2011. Optimization of daily scheduling for home health care services. *Journal of Applied Operational Research*, 3: pp. 124-136.
- [16] Fikar, C. & Hirsch, P. 2017. Home health care routing and scheduling: A review. *Computers & Operations Research*, 77: pp. 86-95.
- [17] Fikar, C. & Hirsch, P. 2015. A matheuristic for routing real-world home service transport systems facilitating walking. *Journal of Cleaner Production*, 105: pp. 300-310.
- [18] Trautsamwieser, A. & Hirsch, P. 2014. A branch-price-and-cut approach for solving the medium-term home health care planning problem. *Networks*, 64(3): pp. 143-159.



- [19] Han, J. & Kamber, M. 2006. *Data mining: Concepts and techniques*. The Morgan Kaufmann Series in Data Management Systems, 2nd edition. Morgan Kaufmann.
- [20] Chiu, T., Fang, D., Chen, J., Wang, Y. & Jeris, C. 2001. A robust and scalable clustering algorithm for mixed type attributes in large database environment. In *Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining*: pp. 263-268. San Francisco: ACM.
- [21] Michailidou, C., Maheras, P., Arseni-Papadimitriou, A., Kolyva-Machera, F. & Anagnostopoulou, C. 2009. A study of weather types at Athens and Thessaloniki and their relationship to circulation types for the cold-wet period, part I: Two-step cluster analysis. *Theoretical and Applied Climatology*, 97(1-2): pp. 163-177.
- [22] Zaki, M.J. & Meira Jr., W. 2014. *Data mining and analysis: Fundamental concepts and algorithms*. 1st edition. USA: Cambridge University Press.
- [23] Mladenovic, N. & Hansen, P. 1997. Variable neighborhood search. *Comput. Oper. Res.*, 24(11): pp. 1097-1100.
- [24] Hansen, P., Mladenović, N., Brimberg, J. & Pérez, J.A.M. 2010. Variable neighborhood search, in M. Gendreau and J.-Y. Potvin (eds), *Handbook of metaheuristics*. Boston: Springer US.
- [25] Bräysy, O. & Gendreau, M. 2005. Vehicle routing problem with time windows, Part I: Route construction and local search algorithms. *Transportation Science*, 39(1): pp. 104-118.
- [26] Hansen, P. & Mladenovic, N. 2002. Developments of variable neighborhood search, in C.C. Ribeiro and P. Hansen (eds), *Essays and surveys in metaheuristics*. Boston: Springer US.
- [27] Cissé, M., Yalçındağ, S., Kergosien, Y., Şahin, E., Lenté, C. & Matta, A. 2017. OR problems related to home health care: A review of relevant routing and scheduling problems. *Operations Research for Health Care*,
- [28] Aiken, L., Clarke, S., Sloane, D., Sochalski, J. & Silber, J. 2002. Hospital nurse staffing and patient mortality, nurse burnout, and job dissatisfaction. *JAMA*, 288(16): pp. 1987-1993.
- [29] Paspap. 2016. *Studies on nurse-to-patient ratios*. <http://www.pennanurses.org/pac/ratio-research/> [accessed 21/06/16].