Models for Intra-Hospital Patient Routing

Belma Turan*, Verena Schmid* and Karl. F. Doerner**

* University of Vienna, Vienna, Austria
**Johannes Kepler University Linz, Linz, Austria

(belma.turan@univie.ac.at, verena.schmid@univie.ac.at, karl.doerner@jku.at)

***Abstract*—The aim of this work is to introduce a new model for intra-hospital routing of patients, considering both client- and management related issues. Patients in a hospital have a fixed appointment, such as x-rays or ultrasonic and due to medical reasons they cannot go on their own, so they will be escorted by porters. In our model logistical costs for the usage of porters and patient inconvenience will be minimized. We show that the different developed model variants are tractable for realistic problem instances of medium-sized hospitals.**

I. INTRODUCTION AND LITERATURE REVIEW

In the recent past transportation, scheduling and supply chain management oriented problems for health care related applications are gaining increased attention in the scientific community. Decision situations may arise on a strategic, tactical as well as an operational level. In the current paper we will concentrate on those occurring on the operational level. The underlying problems can formally be modeled in terms of combinatorial optimization problems coming from scheduling (for both personnel ([1],[2]), resources and rooms such as operating theatres ([2],[3]), transportation routing (of ambulances [4], nurses and doctors) and supply chain management (supply, delivery, reverse logistics of medical waste). In large hospitals, where different wards are typically spread across the site in so called pavilions, routing operations come at high costs. Those costs typically include pure routing (i.e. distance/travel time) related costs, but may also include additional costs coming from delays, missed appointments, inefficient usage of resources, etc. Hence finding good solutions to the underlying routing operations is highly essential. In the remainder of this paper we will focus on transportation routing problems arising in the field of patient routing within hospitals. Patients have a fixed appointment, such as x-rays, ultrasonic, blood testing or surgery. Due to medical reasons they cannot go on their own, but rather they will be escorted there (and back) by porters. Hence two transportation requests from the porters' point of view need to be scheduled, such that their routes are optimized.

To the best of our knowledge so far only classical patient routing problems (such as dial-a-ride problems) have been considered in the scientific community. See [4], [5], [6], [7], [8], [9] and [10] for additional details.

The problem belongs to the family of *pickup and delivery* problems. Ones of the first papers to address this issue have been developed by [5] and [7]. Their optimization was focused on minimizing the transportation costs. Client-centered aspects (such as waiting times observed by real persons) have been included by [11] and [10]. In their approaches maximum ride time restrictions have been included in the optimization. A formal description closely related to real-world patient transportation has been proposed in [12]. For a recent survey on different formulations see [6], [13] and [14]. A decision support system for real-world patient transportation has been described in [8], based on the problem and solution techniques proposed in [4]. Several variants of this problem have been studied recently (see [15], [16] and [17]). True Pareto optimization considering the two conflicting objectives (costs vs. user inconvenience) has been proposed in [9].

The problem under consideration is a special case of the classical *dial-a-ride* problem, arising within hospitals. The capacity of the vehicle (i.e. the porter) is set to one. Furthermore transportation requests are paired in a sense that every patient triggers two transportation requests and their inconvenience is supposed to be minimized. From the patients point of view it would be beneficial if the very same porter could escort them on both resulting transportation requests. Additional extensions under consideration are the porters themselves. Besides executing transportation requests they are also bound to fulfilling other tasks beyond the scope of this model. Hence their duty time should be compact in a sense, that they can also be deployed for those tasks for time slots of sufficient length.

Alternatively the problem at hand can also be considered a special type of the *stacker crane problem* (see [1], [18] and [19]), whereas the latter is an application of full-truck load movements. The stacker crane problem typically arises in port operations where containers need to be moved between different stacks, such that the costs associated with empty movements of the corresponding transporting device are minimized. In the context of our application containers correspond to patients, who need to be transported between known locations. Porters would correspond to the moving device (i.e. the crane) whose empty movements are supposed to be optimized.

The remainder of the paper is organized as follows. A detailed problem description as well as a mathematical formulation is given in Section II. We start with the formulation of a general model and extend it accordingly in order to capture the above-mentioned features from the porters' and patients' point of view. Section III gives a description of the data that are used to test the model, as well as detailed numerical results. We conclude this paper with a summary of the managerial implications and core findings of the investigated model in Section IV.

II. PROBLEM DESCRIPTION

In this section the patient routing and scheduling problem at a pavilion structured hospitals is going to be described. The problem focuses on the in-house transportation of patients, were patients have to be transported between different hospital units. Each patient has one medical examination or surgery scheduled at one hospital unit. From this follows, that we can distinguish between two different kinds of transportation requests for each patient. The first transportation request is an inbound transportation request, where the patient has to be brought from her hospital ward to the facility where the medical examination is scheduled. After the medical examination ends, the patient has to be picked up at this facility and be brought back to the hospital ward. This kind of transportation requests will be referred to as an outbound transportation request of the patient. The transportation requests of patients are done by the so called porters, who are non-medical staff members responsible for logistical operations within a hospital.

The model presented within this paper focuses on pavilion structured hospitals that are characterized by locally dispersed hospital units. These hospital units can be grouped into several hospital wards, where only patient beds are stationed, and several facilities where medical examinations (like blood tests, X-rays and other) or where surgeries take place. Furthermore there are designated buildings where porters are located and have to fulfill additional tasks whenever they are able to. These are called *home depots* of porters. The porters, besides their other duties, are assigned to escort patients on their in-house transportation. This duty of the porters is very important especially in case of elderly or disabled persons. The patients are transported either in stretchers or in wheelchairs, depending on their condition. Hence, the capacity of a porter is set to one, i.e. a porter can only take care of at most one patient at time.

The aim of this work is to introduce a new model for intra-hospital routing of patients, considering both client- and management related issues. The underlying objectives are typically conflicting by nature. A weighted sum approach is used in order to optimize the underlying tradeoff. Furthermore we show that the model at hand is tractable for realistic problem instances of medium-sized hospitals. We assume that this problem can be applied to many hospitals with pavilion structure and help to increase patient convenience, reduce costs and inefficient usage of resources.

A. *General problem*

Our research has started with the general problem. The general problem concentrates on finding the optimal assignment under consideration to minimize the travel time between different nodes at a pavilion structured hospital and to reduce patient inconvenience imposed by long waiting times. The total travel times at the hospital consist of the two types of travel times: the travel times when porters travel with the patient from the pickup to the delivery location, and when the porters travel without the patient (so called empty travel times). As the travel times between pickup and delivery location of a transportation request are fix, they are not considered in the optimization process. Therefore, the travel times that occur in the objective function are only empty travel times.

In the general problem, all medical assistants are in their home depots at the beginning of the work day. As the list with transportation requests that are scheduled for that day is already known in advance, the optimal assignment can be done ahead. The assistants leave the depot and go to complete their transportation requests in the order that was computed before. After they have completed their last transportation request in the tour, they go back to their home depots.

In this paper the problem of offline or static data, where the data are already known, is studied. All information about transportation requests is randomly generated. Pickup and delivery locations for each transportation request and the associated time point when pickup or delivery should take place are already given.

As already mentioned, transportation requests can be divided into two types, inbound and outbound transportation requests, with respect to whether the patient is being escorted to her medical examination, or being picked up afterwards. Division of the transportation requests into these two types was done because it is important to distinguish between two different time windows. In the first case, a patient cannot be brought to the delivery location after the given time point, i.e. the patient cannot come late to her medical examination. The patient can however be brought there earlier, but in this case waiting time is imposed, which is going to be penalized for each minute the patient has to wait. In the second case, by outbound transportation requests, an assistant should pick up the patient at the given time when the patient's medical examination ends. If an assistant comes late, and a patient has to wait for her, a penalty for patient's waiting is imposed. In this case, however, assistant is allowed to come earlier, but her waiting time is not penalized in the general model.

In our paper, all these facts are taken into consideration, and a model is going to be presented that relies on these limitations and objectives. Prior to presenting the mathematical model, the required notation is introduced. Different sets and numbers used are listed in Table I.

TABLE I: NOTATION FOR SETS USED

Abbreviation	Description
$\mathbf{R^I}$	set of inbound transportation requests
$\mathbf{R^O}$	set of outbound transportation requests
R	set of all transportation requests, $\mathbf{R} = \mathbf{R^I} \cup \mathbf{R^O}$
B	set of bed stations
O	set of medical examination rooms
D	set of depots
P	set of patients
S	set of porters
N	set of all nodes at the hospital, $\mathbf{N} = \mathbf{B} \cup \mathbf{O} \cup \mathbf{D}$
$\mathbf{R^*}$	set of all transportation requests and depots, $\mathbf{R^*} = \mathbf{R} \cup \mathbf{D}$

Furthermore, as shown in Table II, the following data associated with time and penalties are used in the model:

TABLE II: NOTATION ASSOCIATED WITH INPUT PARAMETERS

Abbreviation	Description
E_i	critical time point for transportation request *i*
T_{ij}	travel time between nodes *i* and *j* (in minutes)
D(s)	home depot of a porter *s*
A(i)	patient, to whom transportation request *i* refers to
S_i^T	service time for transportation request *i*
α	penalty for time when porter travels empty
β	penalty for waiting time (of patient)

With E_i the critical time point for transportation request *i* is represented. In the case of an inbound transportation request, E_i is the start time of the medical procedure, and in the case of an outbound transportation request, E_i is the time point when the medical procedure ends. Travel times between nodes at the hospital are represented with T_{ij}. The service time is represented with S_i^T. The service time includes the time that is necessary to prepare the patient for the transportation request and also the travel time between the pickup and the drop off location of her transportation request.

In our model, we have used several different decision variables. In the next section we will introduce the binary variables that stand for routing and scheduling decisions, assignment decisions and penalties.

The routing and scheduling decisions are represented by the binary flow variable x_{ijs}, which is equal to 1 if the porter *s* is assigned to transportation request *j* after she completes the transportation request *i*, and equal to 0 otherwise. The binary variable x_{ijs} is defined for $i, j \in \boldsymbol{R^*}$ and $s \in \boldsymbol{S}$.

The assignment binary variable y_{is} serves to represent the assignment of transportation requests to porters. The variable is defined for $i \in \boldsymbol{R^*}$ and $s \in \boldsymbol{S}$ and is equal to 1 if porter *s* is assigned to transportation request *i*, and is equal to 0 otherwise.

Besides the binary variables, we have also used variables that are mostly associated with time. With a_i the porter's arrival time to the pickup node of a transportation request *i* is represented. After the porter has arrived to the pickup location she picks up the patient, if the patient is ready (e.g. if the medical examination has already ended), and the transportation request can start. The time when transportation request *i* starts is represented with variable b_i. The actual travel time between pickup and delivery node of transportation request *i* is included in service time for each transportation request. The arrival time to the delivery location of transportation request is represented with c_i. After the porter has escorted the patient to her delivery location, she can leave the delivery node. The time when porter leaves the delivery node of transportation request *i* is represented with d_i.

All decision variables used are listed in Table III.

TABLE III: DECISION VARIABLES

Abbreviation	Description
a_i	arrival time to the pickup location of request *i*
b_i	start time of transport of transportation request *i*
c_i	arrival time to the delivery location of request *i*
d_i	time point when porter leaves the delivery location of transportation request *i*
x_{ijs}	binary flow variable
y_{is}	binary assignment variable

After all the data and indices used are introduced, and the decision variables explained, the mathematical model can be presented:

Minimize

$$\alpha \sum_{i \in \boldsymbol{R^*}} \sum_{j \in \boldsymbol{R^*}} \sum_{s \in \boldsymbol{S}} x_{ijs} T_{ij} + \beta \left(\sum_{i \in \boldsymbol{R^I}} (E_i - c_i) + \sum_{i \in \boldsymbol{R^O}} (b_i - E_i) \right) \quad (1)$$

subject to

$$\sum_{s \in \boldsymbol{S}} y_{is} = 1 \qquad \forall i \in \boldsymbol{R} \quad (2)$$

$$y_{\boldsymbol{D}(s)s} = 1 \qquad \forall s \in \boldsymbol{S} \quad (3)$$

$$y_{is} = \sum_{j \in \boldsymbol{R^*}} x_{ijs} \qquad \forall i \in \boldsymbol{R^*}, \forall s \in \boldsymbol{S} \quad (4)$$

$$y_{is} = \sum_{j \in \boldsymbol{R^*}} x_{jis} \qquad \forall i \in \boldsymbol{R^*}, \forall s \in \boldsymbol{S} \quad (5)$$

$$c_i = b_i + S_i^T \qquad \forall i \in \boldsymbol{R} \quad (6)$$

$$a_i \le b_i \qquad \forall i \in \boldsymbol{R} \quad (7)$$

$$c_i \le d_i \qquad \forall i \in \boldsymbol{R} \quad (8)$$

$$c_i \le E_i \qquad \forall i \in \boldsymbol{R^I} \quad (9)$$

$$E_i \le b_i \qquad \forall i \in \boldsymbol{R^O} \quad (10)$$

$$d_i + T_{ij} \le a_j + M(1 - x_{ijs}) \qquad \forall i,j \in \boldsymbol{R^*},\ s \in \boldsymbol{S} \quad (11)$$

$$d_i + T_{ij} \ge a_j - M(1 - x_{ijs}) \qquad \forall i,j \in \boldsymbol{R^*},\ s \in \boldsymbol{S} \quad (12)$$

$$x_{ijs} \in \{1,0\} \qquad \forall i,j \in \boldsymbol{R^*}, \forall s \in \boldsymbol{S} \quad (13)$$

$$y_{is} \in \{1,0\} \qquad \forall i \in \boldsymbol{R^*}, \forall s \in \boldsymbol{S} \quad (14)$$

$$a_i, b_i, c_i, d_i \ge 0 \qquad \forall i \in \boldsymbol{R^*} \quad (15)$$

To determine the objective function (1), we have used the weighted sum approach. Each of three different terms that are considered in the objective function has to be multiplied with a coefficient. However, these terms can be divided into two categories, so there are only two different coefficients. The coefficients were subjectively estimated by the authors.

The total travel time of porters is represented with the first term of the objective function. The optimal route should be created so that the travel times of porters are as short as possible. As already mentioned, the pickup and delivery node of one transportation request are paired, that means that the time travelled between a pickup location and its associated delivery location is fixed and cannot be changed or improved anymore. With the first term in the objective function, the total travel time of porters is measured, i.e. the travel time between the delivery node of transportation request i to the pickup node of transportation request j, if the porter s is assigned to transportation request j after transportation request i. As x_{ijs} is defined for set R^*, which includes depots besides requests, the first term then also contains travel times from the porter's home depot to her first transportation request (or rather to the pickup location of her first transportation request) and from the delivery location of her last transportation request to her home depot.

With the other group of terms it is made sure that the comfortableness of patients is also considered in the optimal plan. The second term is used for inbound requests. With this term the total waiting time for patients, from the time point when they arrive to facility where their medical procedure is scheduled until the time point when their medical procedure actually begins, is minimized. Similarly, the third term serves for the outbound requests, and ensures that patients have to wait as short as possible to be picked up after their medical procedure and brought back to their bed. Both terms are multiplied with the same penalty coefficient in the objective function, as it is equally important for us that patients don't have to wait long in both cases.

In order to obtain the optimal solutions, some restrictions have to be considered. Constraints (2) make sure that each transportation request is served once. With (3) it is ensured that each porter has to visit her home depot. Constraints (4) and (5) are in and out degree constraints.

Restrictions that make sure that every transportation request is completed punctually also need to be considered. With (6) the arrival time to the target point is calculated. Constraints (7) and (8) ensure that the start of the transportation request cannot begin if the porter has not arrived yet and that the porter cannot be available for the next transportation request before she has delivered the patient to the target location. A porter can be responsible for only one transportation request at a time. In case of an inbound request, with (9) is guaranteed that patient cannot come late for her medical procedure. On the other hand, in case of an outbound request, constraints (10) make sure that the beginning of the transport is after the end point of the surgery. If the porter comes earlier to pick up the patient after the medical procedure, the porter has to wait until the procedure is done, and the transport cannot start until then. Constraints (11) and (12) enable the connection between two consequent transportation requests, whereas (13) and (14) are binary constraints. Constraint (15) is a non-negativity constraint.

This mathematical model was implemented in XPRESS. In Section III we are going to introduce and explain the data used and the solutions we managed to obtain. There will be more words on how the different data were generated, which types of problems were considered and different solutions will be discussed.

B. *Patient-centered extension*

In this section the first extension of the model is going to be introduced. For this extension the general model was used and the necessary changes were applied.

The extension is based on the wish to provide the best service for patients. The quality of hospital service is measured through patients' satisfaction. Therefore their convenience and well-being should be of high priority for hospital management. In the general model the main goal was to reduce the patients' waiting time. Now the question imposes what else could be done in order to increase patients' convenience. It was already mentioned that every patient triggers two transportation requests, one from her hospital ward to the medical examination room and the other one from this examination room back to the hospital ward. As the medical examinations are stressful for the patients, and the necessary transportation further increases this stress, it would be reasonable to try to reduce this inconvenience by assigning the *same* porter to take care of the patient. If the same person picks up the patient and escorts her to her medical examination and afterwards picks her up and delivers back to bed, the patient could develop a feeling of trust toward this person. Knowing that there is one porter who is responsible for her could increase the comfort for the patient.

In order to obtain the resulting assignments and routes, the necessary changes mentioned above need to be implemented in the model. As we have based this extension on our general model, only the changes will be stated.

The data used and the decision variables remain the same as in the general model. There are also no changes in the objective function. However, we need additional constrains that will make sure that one porter is assigned to one patient. The following constraint

$$y_{is} = y_{js} \qquad \forall s \in \boldsymbol{S},\ i,j \in \boldsymbol{R} \cap \boldsymbol{A(i)} = \boldsymbol{A(j)} \quad (16)$$

makes sure that if the transportation request i is assigned to the same patient as the transportation request j, then porter s has to take care of both transportation requests. In this case, $A(i)$ stands for a patient to whom transportation request i refers to.

This model was also implemented in XPRESS and solved. At the end, all three models are compared. The analysis of the solution obtained is stated in Section III.

C. *Hospital-centered extension*

In this section we are going to introduce the second extension of the model. So far we assumed that medical assistants are assigned to complete transportation requests, they start their workday at their home depot, and then perform the transportation requests assigned. After they have completed all of their transportation requests, they go back to their home depots. This model however doesn't take into consideration the time that porters have to wait empty between two consecutive transportation requests. It is assumed that a porter completes one transportation request, picks up a patient and delivers her at the desired location and then goes to the pickup location of her next transportation request. If the porter arrives there earlier than planned, she has to wait. From the managements' point of view this however is suboptimal. Porters are also responsible for executing additional tasks beyond the transportation requests. So far (empty) waiting times occurred somewhere in the hospital compound and the management is not able to use their resources efficiently. Hence we now want to consider this issue explicitly. By sending porters temporarily back home to their home depots, porters could be assigned other tasks there. This may lead to a deterioration of the solution with respect to the distance travelled empty by porters as they may encounter a detour via their home depot. On the other hand this allows to efficiently use their resources for other tasks (e.g. collection of blood samples, delivery and supply of medical instruments, etc.) Porters however should only be sent back to their home depots if the resulting time spent there exceeds a certain minimum time span.

Taking all these information into consideration, we have changed and extended the model accordingly. As the general model was our basis, we are now going to introduce all the changes that were necessary to implement. One of the main changes was to introduce new decision variables that will force the model to send the porter back home if there is enough time and another variable to capture the actual time travelled empty.

The decision whether the porter goes back to her home depot between her two consequent transportation requests is represented by a binary variable w_i. This variable is defined for $\forall\, i \in \boldsymbol{R}$ and is equal to 1 if the porter temporarily goes back to her home depot after she has completed her transportation request i. There is a difference between w_i and $x_{iD(s)s}$, where the latter one refers to the last transportation request i of the porter s, after which completion the porter will finally go back to her home depot. In the first case, if w_i is equal to 1, the porter will only go to her home depot, spend some time there, but then leave the depot in order to complete some other transportation requests.

The actual travelling time spent empty between two transportation requests i and j by porter s is no longer constant. It will be represented by a variable t_{ijs}, capturing an eventual detour via her depot. The new data used are listed in Table IV.

TABLE IV:
DECISION VARIABLES AND PARAMETERS FOR 2nd EXTENSION

Abbreviation	Description
t_{ijs}	time travelled between locations i and j by porter s
w_i^H	time that porter spends in her home depot after transportation request i
w_i	binary variable equals 1 if porter is sent back home temporarily after transportation request i
T^W	minimal waiting time porter needs to spend at home depot
γ	penalty, if porter has to wait idle between transportation requests

As far as the mathematical model is concerned, we are now only going to state the changes of the general model that took places and list the additional constraints.

Minimize

$$\alpha \sum_{i\in \boldsymbol{R}^*} \sum_{j\in \boldsymbol{R}^*} \sum_{s\in \boldsymbol{S}} t_{ijs} + \beta \left(\sum_{i\in \boldsymbol{R}^I} (E_i - c_i) + \sum_{i\in \boldsymbol{R}^O} (b_i - E_i) \right) + \gamma \left(\sum_{i\in \boldsymbol{R}} (b_i - a_i) + \sum_{i\in \boldsymbol{R}} (d_i - c_i) \right) \quad (17)$$

subject to

$$w_i + x_{i\boldsymbol{D}(s)s} \le 1 \qquad \forall i \in \boldsymbol{R}, s \in \boldsymbol{S} \quad (18)$$

$$T_{i\boldsymbol{D}(s)} + T_{\boldsymbol{D}(s)j} + w_i^H \le a_j - d_i + M(2 - x_{ijs} - w_i) \qquad \forall i, j \in \boldsymbol{R}^*, s \in \boldsymbol{S} \quad (19)$$

$$T_{i\boldsymbol{D}(s)} + T_{\boldsymbol{D}(s)j} + w_i^H \ge a_j - d_i - M(2 - x_{ijs} - w_i) \qquad \forall i \in \boldsymbol{R}, j \in \boldsymbol{R}^*, s \in \boldsymbol{S} \quad (20)$$

$$w_i^H \ge T^W * w_i \qquad \forall i \in \boldsymbol{R} \quad (21)$$

$$T_{ij} + d_i \le a_j + M(1 - x_{ijs} + w_i) \qquad \forall i, j \in \boldsymbol{R}^*, s \in \boldsymbol{S} \quad (22)$$

$$T_{ij} + d_i \ge a_j - M(1 - x_{ijs} + w_i) \qquad \forall i, j \in \boldsymbol{R}^*, s \in \boldsymbol{S} \quad (23)$$

$$t_{ijs} \ge T_{i\boldsymbol{D}(s)} + T_{\boldsymbol{D}(s)j} - M(2 - x_{ijs} - w_i) \qquad \forall i \in \boldsymbol{R}, j \in \boldsymbol{R}^*, s \in \boldsymbol{S} \quad (24)$$

$$t_{ijs} \ge T_{ij}\,(x_{ijs} - w_i) \qquad \forall i, j \in \boldsymbol{R}^*, s \in \boldsymbol{S} \quad (25)$$

$$w_i \in \{1, 0\} \qquad \forall i \in \boldsymbol{R} \quad (26)$$

$$w_i^H \ge 0 \qquad \forall i \in \boldsymbol{R} \quad (27)$$

$$t_{ijs} \ge 0 \qquad \forall i, j \in \boldsymbol{R}^*, \forall s \in \boldsymbol{S} \quad (28)$$

There have been few changes in the objective function. Instead of the first sum in general model, where the total time travelled by the porter s was only depending on the time travelled between her two consequent transportation requests i and j, the new term (18) calculates the sum of time travelled between two transportation requests, and also adds up the travel time the porter needed to go to her home depot and back, if she goes there temporarily between these transportation requests. There was also one more group of goals added. The fourth and the fifth term in the objective function stand for the waiting time of the porter, either before or after having executed a transportation request. These waiting times are penalized with the same coefficient in the objective function, as these waiting times are equally important. However, it is important to say that these waiting times refer only to waiting times between two transportation requests in case when the porter doesn't go back to her home depot. The waiting time in the home depot is not penalized, as it is important that the porter spends as much time there as possible, so that she could be assigned to some other duties. The fourth term penalizes the waiting times for porters when they arrive to the pickup location and have to wait until the transportation request starts. Similarly, the fifth term penalizes the waiting time at the delivery location, when the porter has delivered a patient and waits there to be assigned to next transportation request.

There were also few changes in constraints. Instead of constraints (11) and (12) in the general model, new constraints (18) to (28) are added. With (18) it is made sure that porter can either temporarily go home between her two transportation requests, spend some time there and then leave the depot again to complete some other transportation requests, or the porter can finally go to her home depot at the end of her workday. Constraints (19) and (20) make sure that if a porter goes to her home depot temporarily between two transportation requests, there is enough time between two successive transportation requests. There should be enough time to go to her home depot after one transportation request, to spend some time there and then go to complete her next transportation request. Time between two transportation requests is measured as the difference between the time when porter is free from her first transportation request and the time when she should arrive to her next transportation request, if the porter is assigned to complete transportation request *j* after request *i*. (21) enables that, if the porter goes back to her home depot after transportation request *i*, she should spend at least some given time T^W there. Constraints (22) and (23) make sure that there is a connection between two transportation requests. With (24) and (25) the total travel time between two successive transportation requests is modeled. The travel time will include the travel time from the drop-off location of the porter's last transportation request to her home depot and from her home depot to the pick-up location of her next transportation request, in case that the porter goes back to her home depot after she has completed the transportation request. If the porter goes straight to her next transportation request without visiting the home depot, the travel time will be equal to the travel time from the drop-off location of her last transportation request to the pick-up location of her next transportation request. Constraints (26) are binary and constraints (27) and (28) are non-negativity constraints.

These changes are also implemented and the obtained solution as well as the comparison to the general model is discussed in the Section III.

III. NUMERICAL RESULTS

In this section, the data used in the model will be introduced and explained. We tested all instances with XPRESS until the optimal solution was found, or until the termination criterion was reached. For our purposes, the termination criterion was set to one hour of computation time. If the optimal solution is not found by then, the best solution found so far was used as an objective value and all analysis and comparisons is made using this value.

A. Instances

Four different classes of instances are generated, that differ in the number of transportation requests and the number of porters that are available. For the first three classes, ten different instances are created and tested. For the fourth class, the number of porters is fixed, and the number of transportation requests is being varied. The travel time matrix is drawn from the real-world data. The travel time matrix and the penalty coefficients remain the same for all instances used. The numbers of porters and transportation requests, and the penalty coefficients are set by the authors.

Instances differ from each other only in the critical time points, when patients' medical procedures are scheduled at and the time points when they finish, i.e. when the patient should be picked up and brought back to bed station. These critical time points are generated with random number generator. The idea was to generate different time points in the period from 8 am to 1 pm, for the first three classes, and the period from 8 am to 5 am for the fourth class. Time points were expressed in minutes. We also assumed that all of these medical procedures had a length of approximately 20 to 30 minutes. So adding these two randomly generated values up, we could determine the start and the end point of the medical procedure.

The pickup and delivery nodes are randomly chosen among nodes at the hospital that stand for bed stations and for the facilities where medical procedures are being done. Transportation requests are assigned to patients in ascending order.

For the first class, the number of porters was set to 2 and the number of patients to 3, respectively. Hence a total of 6 transportation requests have to be considered. Ten instances were created and tested for the general version of our model and for both extensions of the model. The solution for the general model, and also the comparison of the general model solution to the solutions of the two extensions is given in the next Section.

Similarly, for the second (third) class of instances the number of porters was set to 3 (4) and the number of patients was set to 5 (7). The solution for these two types is given in the next Section. For instances belonging to class II additional test runs are made to investigate how the solution changes if the minimal acquired time the porter has to spend in the home depot (T^W) varies. An overview on the size of problem instances under consideration can be found in Table V.

For the fourth class, the number of porters is fixed to 4 and the number of patients is varied from 8 to 20. The solution and the sensitivity analysis are given in the next Section.

The penalty coefficient for the time travelled α is set to be 1, penalty coefficient χ (waiting time of porters) is set to be 2, and the penalty coefficient for the waiting time of patients (β) is set to be 3. This can be interpreted as following: one minute that a patient (porter) has to wait is three (two) times more important than an additional minute that a porter has to travel. These coefficients can easily be adapted by the decision maker in order to reflect their true preferences.

TABLE V: INSTANCES CLASSES

Class	Number of porters	Number of patients
I	2	3
II	3	5
III	4	7
IV	4	8-20

B. General model

The obtained solution with XPRESS-Solver for the first class is given in the Table VI. The table contains all relevant data related to the patients' transportation request, e.g. average waiting time from both porters' and patients' point of view, total time travelled by porters and average time needed to compute the results. The average value of the objective value is 29.30, weighted over 10 instances. The total time travelled by porters is equal to the objective function value. As the value of the objective function is weighted sum of travel time and waiting times for patients, this leads to the conclusion that patients don't have to wait at all. On the other hand, porters do have to wait, either in case that they have to pick up the patient after the medical procedure, where the porter waits at the pickup node, or in case that they have just dropped off a patient, when they wait at the delivery node. The average number of patients that are assigned to one patient is 1.23, what can be interpreted that in approximately 2 out of 3 times patient will be assigned to only one porter. The model is solved to optimality for all ten instances. The model could be solved in only couple of centiseconds.

TABLE VI: GENERAL MODEL (FIRST CLASS)

N	f	tt	w (s)	w (p)	No (p)	GAP	time
1	28.00	28.00	17.67	0.00	1.33	0.00%	0.08
2	28.00	28.00	39.17	0.00	1.00	0.00%	0.06
3	21.00	21.00	112.67	0.00	1.33	0.00%	0.03
4	28.00	28.00	27.00	0.00	1.00	0.00%	0.05
5	47.00	47.00	11.00	0.00	1.33	0.00%	0.11
6	28.00	28.00	38.67	0.00	1.33	0.00%	0.03
7	36.00	36.00	53.67	0.00	1.33	0.00%	0.06
8	28.00	28.00	64.00	0.00	1.33	0.00%	0.03
9	28.00	28.00	26.67	0.00	1.00	0.00%	0.11
10	21.00	21.00	112.67	0.00	1.33	0.00%	0.03
avg	29.30	29.30	50.32	0.00	1.23	0.00%	0.06

The notation used in the table VI as well as the following tables is listed in the table VII.

TABLE VII: NOTATION

Abbreviation	**Description**
N	Instance number
C	Class of Instances
f	Value of objective function
tt	Total travel time (empty)
w(s)	Total porters' waiting time
w(p)	Total patients' waiting time
No(p)	Number of porters that are assigned to one patient
GAP	Gap between best solution found so far and the best bound in %
No (OS)	Number of optimal solutions found
time	Elapsed run time until termination

After the first class instances were tested, the instances for the class two and three were created and tested. A comparison is given in the table VIII. In the table VIII, *C* stands for the class number, and with *No(OS)* the number of optimal solutions found (summed over ten different instances for each class) is represented.

TABLE VIII: GENERAL MODEL – COMPARISON

C	f	tt	w (s)	w (p)	No (p)	GAP	No (OS)	time
1	29.30	29.30	50.32	0.00	1.23	0.00%	10	0.06
2	37.30	37.30	129.22	0.00	1.32	0.00%	10	0.30
3	49.40	49.40	111.94	0.00	1.24	0.00%	10	1.97
avg	38.67	38.67	97.16	0.00	1.27	0.00%	10	0.78

It can be concluded that the size of the instances tested influences the computation time. The time needed to find the optimal solution increases with the increase in the size of the instances. However, the optimal solution still could be found very fast.

C. *Patient centered extension*

After the general model was tested, the two extensions were also implemented in the XPRESS-Solver and solved. The first extension investigates the change in the objective function values in case when it is required that only one porter can be assigned to patient. In this case, the objective function value is slightly increased in comparison to the objective value of the general model. This can be interpreted as, that the costs to increase the patients' convenience and to make sure that they are taken care of by the same porter, are not much higher than in general case. The solutions for the first three classes are given in the Table IX.

The requirement of only using one porter per patient has a minor impact on the quality of the solution obtained, due to mediocre additional empty travel times by porters. On average the resulting empty travel times increase by 11.7% from 38.67 to 43.2 time units. From the patients point of view however the situation improves. Waiting times still do not occur. With this extension the average number porters in use per patient is forced to one (in contrast to 1.27 in the previous general case, where this feature has not been addressed explicitly). This additional constraint has only minor impact on the run times required. On average all instances can be solved to optimality within 0.8 seconds.

TABLE IX: PATIENT-CENTERED EXTENSION – COMPARISON

C	f	tt	w (s)	w (p)	No (p)	GAP	No (OS)	time
1	30.90	30.90	24.17	0.00	1.00	0.00%	10	0.03
2	42.80	42.80	26.61	0.00	1.00	0.00%	10	0.45
3	56.00	56.00	29.23	0.00	1.00	0.00%	10	2.05
avg	43.23	43.23	26.67	0.00	1.00	0.00%	10	0.84

D. *Hospital centered extension*

Up to now optimization from the porters' view was focused on their time travelled empty. Waiting times, i.e. time slots when they are off-duty, were not considered explicitly. This, however, seems to be a waste of resources, as porters have other tasks to fulfill. Waiting times occurred between serving consecutive transportation requests somewhere within the hospital complex. Within this extension we tried to include the minimization of waiting times (i.e. idle times) spent *somewhere* in the compound from the porters point of view, by sending them back to their home depots, where they are supposed to fulfill other tasks. Hence we now try to emphasize generating schedules for porters where idle times are connected, long enough and occur at the corresponding home depots.

The results for the hospital centered extension are given in the Table X. The minimal time that porters should spend at their home depot, in the case that she goes there between two transportation requests, was set to 15 minutes, for all classes tested. In this extension, where the porters were sent to their home depots in case that there was enough time, a significant increase in the objective function can be noticed. This increase however is due to the increase in the time travelled empty by porters. What can be noticed is that the porters' average waiting time is decreased significantly in comparison to the general model (7.16 vs. 97.16 time units). This is due to the fact that porters now can go temporarily back to their home depots between two consecutive transportation requests. The time they spend in the home depot is represented with w^H and the number of times they go to the home depot is represented with w_i. On average porters are send back to their home depots 4.70 times and spend 190.78 time units there.

Considering this feature comes at high costs. The solution quality deteriorates in a three-fold way. The average number of porters in use increases from 1.27 to 1.48. Patient inconvenience is further increased by additional waiting times that occur before or after the transportation requests. (These waiting times increase from 0.00 to 4.13 time units.) Furthermore empty travel times by porters increase from 38.67 to 87.80 time units. This effect is not surprising as sending porters to their home depots leads to increased travel times.

With this extension however the underlying combinatorial complexity increases dramatically. In contrast to the previous case this extension heavily influences the solvers capabilities of quickly finding good (optimal) solutions. Especially for the third class it can be noticed, that none of the instances tested could be solved to optimality in one hour of running time. The gap between the best solution found and the best bound after 3600 seconds is still at 95.90%.

TABLE X
HOSPITAL-CENTERED EXTENSION - COMPARISON

C	f	tt	w (s)	w (p)	w_i^H	w_i	No (p)	GAP	No (OS)	time
1	61.20	57.30	3.90	0.00	107.21	3.10	1.27	0.0%	10	1.03
2	98.00	89.30	8.20	0.50	202.41	4.70	1.52	0.0%	10	373.8
3	138.10	116.80	9.40	11.90	262.72	6.30	1.64	95.9%	0	3792.3
avg	99.10	87.80	7.16	4.13	190.78	4.70	1.48	32.0%	6.67	1389.1

For the second class, a sensitivity analysis is performed in order to depict how the required waiting time affects the value of the objective function.

The waiting time was varied between 5 minutes to 140 minutes. The sensitivity analysis is performed on one randomly chosen instance among those that were used for the second class. The values obtained are listed in the Table XI.

TABLE XI: SENSITIVITY ANALYSIS

T^W	f	w_i^H	w_i	w(s)	w(p)
5	110.0	193.0	7	0.0	0.0
10	110.0	193.0	7	0.0	0.0
15	110.0	193.0	7	0.0	0.0
20	132.0	185.0	7	0.0	7.0
25	149.0	193.0	5	17.0	10.0
30	169.0	218.0	5	12.0	18.0
35	177.0	225.0	4	3.0	24.0
40	177.0	225.0	4	3.0	24.0
45	177.0	225.0	4	3.0	24.0
50	179.0	223.0	3	10.0	24.0
55	179.0	223.0	3	10.0	24.0
60	179.0	223.0	3	10.0	24.0
80	193.0	168.0	2	77.0	11.0
100	201.0	120.0	1	101.0	25.0
120	201.0	120.0	1	101.0	25.0
140	208.0	0.0	0	122.0	9.0

The minimal time that is required for porter to spend at her home depot is denoted by T^W. The table shows the following things: with an increase in the minimum time that needs to be spent at the home depot if porters are sent back there temporarily between two successive transportation requests, the number of times porters finally go back home (w_i) decreases. Simultaneously the time spent at home (w_i^H) first increases and starts to decrease starting from $T^W = 50$, as the total time spent there is offset by the reduced number of visits at the home depot. Similarly the inconvenience of patients (measured in terms of their waiting times *w(p)*) first increases and starts to decrease again starting from $T^W = 80$, as it becomes less efficient to send porters back home.

E. Evaluation of the solution quality

To show how the value of the objective function and the best bound change over time, a random instance is chosen among those that were used for the third class of instances. The relation is shown in the Figure 1.

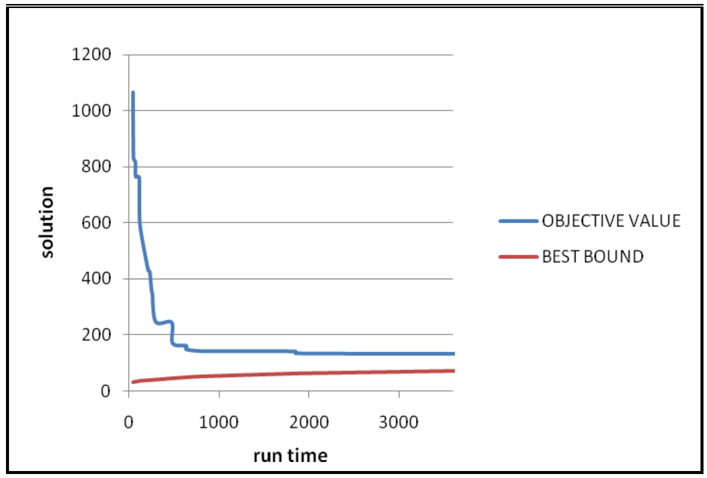

Figure 1. Objective value and best bound vs. time

Instances within the third class where among the largest instances under consideration, where the number of porters was set to four and 14 transportation requests had to be considered. For this particular instance the first feasible solution was found after 45.66 seconds, with an initial gap of 3363%. Within the first half of the total run time the solution quality could be improved by 86.6% to 142, the resulting gap decreased to 125%. Within the last 1800 seconds the solution (gap) can only be improved marginally. The problem under consideration is static and operational by nature. Hence the optimization could be executed over night in order to generate a solution for the next day. It could be observed however that solutions of reasonable quality already could be obtained within half an hour.

F. Larger Instances

In order to test the performance of the solver in use the following experiment is set up: the number of patients (and the resulting transportation requests) is gradually increased up to 20 (40). For the resulting fourth class of instances, the number of porters still is set to 4. The best solutions found within one hour of run time can be found in Table XII and XIII. The number of patients is represented with $|p|$. The XPRESS-Solver could solve the problem for maximally 40 transportation requests in one hour of running time. As the computing times were very large, only the general model and the first extension were tested for the fourth class. The solutions obtained for the general model are depicted in table XII.

TABLE XII: FOURTH CLASS – GENERAL MODEL

|p|	f	tt	w (s)	w (p)	No (p)	GAP	No (OS)	time
8	56.0	56.0	108.0	0.0	1.4	0.00%	1	2.0
9	40.0	40.0	138.9	0.0	1.3	0.00%	1	1.5
10	47.0	47.0	131.9	0.0	1.4	0.00%	1	3.6
11	56.0	56.0	53.7	0.0	1.2	0.00%	1	18.9
12	56.0	56.0	53.7	0.0	1.3	0.00%	1	18.0
13	48.0	48.0	82.6	0.0	1.3	0.00%	1	25.6
14	56.0	56.0	51.3	0.0	1.1	0.00%	1	27.4
15	56.0	56.0	37.0	0.0	1.1	0.00%	1	194.8
16	83.0	83.0	47.9	0.0	1.4	0.00%	1	228.7
17	83.0	80.0	34.5	3.0	1.4	8.68%	0	3645.3
18	64.0	64.0	76.1	0.0	1.2	0.00%	1	344.3
19	64.0	64.0	30.9	0.0	1.1	0.00%	1	2322.8
20	68.0	68.0	25.1	0.0	1.1	0.00%	1	1786.2

With the increase in the number of patients, the value of the objective function, as well as the computational time, tend to increase. The problems up to $|p| = 15$ (up to 30 transportation requests) could be solved within less than 30 seconds. Afterwards the computational time has increased rapidly, but the problem could still be solved to optimality for almost all instances tested (exception is a problem with 17 patients).

The fourth class was also tested with the extended model, namely with the patient centered extension. The obtained results are given in the Table XIII. Compared to the general model, this extension could be solved to optimality for all instances tested. The optimal solution could be obtained within less than 10 minutes. The values of the objective function (f) deteriorated up to 20% in comparison to the general model, which means that with slight increase in the travel time, the patient convenience (i.e. the same porter to take care of them) could be guaranteed. From the porters' point of view, the solution deteriorates slightly in the sense that the travel time, *tt*, increases, but the porters' waiting times even decrease for few instances tested.

TABLE XIII: FOURTH CLASS – PATIENT CENTERED EXTENSION

|p|	f	tt	w (s)	w (p)	No (p)	GAP	No (OS)	time
8	64.0	64.0	39.3	0.0	1.0	0.00%	1	6.0
9	48.0	48.0	39.3	0.0	1.0	0.00%	1	0.8
10	56.0	56.0	39.1	0.0	1.0	0.00%	1	3.8
11	56.0	56.0	49.7	0.0	1.0	0.00%	1	9.6
12	56.0	56.0	47.6	0.0	1.0	0.00%	1	8.4
13	56.0	56.0	24.8	0.0	1.0	0.00%	1	48.5
14	56.0	56.0	38.1	0.0	1.0	0.00%	1	19.3
15	56.0	56.0	47.8	0.0	1.0	0.00%	1	26.8
16	83.0	83.0	21.8	0.0	1.0	0.00%	1	107.3
17	83.0	80.0	37.6	3.0	1.0	0.00%	1	383.2
18	64.0	64.0	31.8	0.0	1.0	0.00%	1	167.1
19	64.0	64.0	44.2	0.0	1.0	0.00%	1	246.2
20	72.0	72.0	38.1	0.0	1.0	0.00%	1	598.2

Larger instances were also tested with XPRESS, where the computation time was increased to five hours. Already for the problem with 4 porters and 25 patients (i.e. 50 transportation requests) no feasible solution could be found within the given time.

IV. CONCLUSION AND OUTLOOK

The scope of this paper was to present a novel optimization model for considering intra-hospital routing, a very important aspect prevailing in health-care modeling. Besides considering classical objectives for vehicle routing (i.e. minimizing distance or travel time related costs) we are also focusing on the patients' perspective (i.e. the so called client-centered perspective, consisting of minimizing waiting times before and after their scheduled appointments, as well the number or porters assigned to patients).

Reducing patients' inconvenience by imposing a limit on the number of porters in use per patient can easily be implemented. The quality of the solution obtained deteriorates slightly due to additional travel times by porters (on average 11.7%). The effect on run times for obtaining the solution is neglectible. Hence the inconvenience of the patient can easily be reduced at low costs and should be considered whenever possible. The number of porters in use per patient can be reduced by 27%.

Considering concentrated waiting times of porters, such that they can easily be assigned to other tasks beyond the scope of this model is also possible. From the patients' point of view the solution deteriorates slightly. However one major drawback are the resulting run times. The complexity of the underlying model increases dramatically and the model becomes computationally intractable as the size of the problem instance increases. Within one hour of run time our large instances under consideration could not be solved to optimality. The resulting gap was still at 95.6%.

The proposed model can be solved within a reasonable amount of time for small problem instances. For larger instances including more than 40 transportation requests however the use and development of a (meta-)heuristic becomes unavoidable. We are planning to further extend this model in a two-fold way: First we would like to relax the assumption that every patient may only be escorted to/from one single appointment, by considering several appointments that need to be considered sequentially. Next we would like to combine the optimization of the underlying routing problem (given starting and ending times of appointments) with the scheduling problem at hand. By simultaneously optimizing both the underlying scheduling problem (for the individual appointments, by making sure that at most one appointment may be scheduled per room) and the resulting routing problem, we expect to achieve even better results.

REFERENCES

[1] M. Gronalt, R.F. Hartl, M. Reimann, "New savings based algorithms for time constrained pickup and delivery of full truckloads". *European Journal of Operational Research* (2003), vol. 151, no. 3, pp. 520-535.

[2] J. Beliën, E. Demeulemeester, and B. Cardoen. "A decision support system for cyclic master surgery scheduling with multiple objectives". *Journal of Scheduling* (2009), vol. 12, no. 2, pp. 147-161.

[3] B. Cardoen, E. Demeulemeester, and J. Beliën. "Operating room planning and scheduling: a literature review". *European Journal of Operational Research* (2010), vol. 201, no. 3, pp. 921-932.

[4] A. Beaudry, G. Laporte, T. Melo, S. Nickel, "Dynamic transportation of patients in hospitals". *OR Spectrum* (2010), vol. 32, no. 1, pp. 77-107.

[5] H. N. Psaraftis, "A dynamic programming solution to the single-vehicle, many-to-many immediate request dial-a-ride problem". *Transportation Science* (1980) vol. 14, no. 2, pp. 130–154.

[6] J. C. Cordeau, G. Laporte, "The Dial-a-Ride Problem: models and algorithms", *Annals of Operations Research* (2007) vol. 153, no.1, pp. 29-46.

[7] M.W.P. Savelsbergh, M. Sol, "The general pickup and delivery problem", *Transportation Science* (1998) vol. 29, pp. 17-29.

[8] T. Hanne, T. Melo, S. Nickel, “Bringing Robustness to Patient Flow Management through Optimized Patient Transports in Hospitals“, Interfaces (2009), vol. 39, no. 3, pp. 241-255.

[9] S. Parragh, K.F. Doerner, X. Gandibleux, R.F. Hartl, “A two-phase heuristic solution approach for the multiobjective dial-a-ride-problem”. *Networks* (2009), vol. 54, no. 4, pp. 227-242.

[10] P. Toth, D. Vigo, “Heuristic Algorithms for the handicapped persons transportation problem”. *Transportation Science* (1997), vol. 31, no. 1, pp. 60-71.

[11] O. B. G. Madsen, H. F. Ravn, J. M. Rygaard, “A heuristic algorithm for a dial-a-ride problem with time windows, multiple capacities, and multiple objectives”. *Annals of Operations Research* (1995) vol. 60, no. 1, pp. 193-208.

[12] J.F. Cordeau and G. Laporte. “A tabu search heuristic for the static multi-vehicle dial-a-ride problem”. *Transportation Research Part B: Methodological* (2003), vol. 37, no. 6, pp. 579-594.

[13] S. N. Parragh, K.F. Doerner and R.F. Hartl. “A survey on pickup and delivery problems Part II: Transportation between pickup and delivery locations”.*Journal für Betriebswirtschaft* (2008), vol. 58, no. 2, pp. 81-117.

[14] J.F. Cordeau and G. Laporte, “The Dial-a-Ride Problem (DARP): Variants, modeling issues and algorithms”. *4OR: A Quarterly Journal of Operations Research* (2003), vol. 1, no. 2, pp. 89-101.

[15] E. Melachrinoudis, A. B. Ilhan, H. Min,” A dial-a-ride problem for client transportation in a health-care organization”. *Computers & Operations Research* (2007) vol. 34, pp.742–759.

[16] C. Fiegl, C. Pontow, “Online scheduling of pick-up and delivery tasks in hospitals”. *Journal of Biomedical Informatics* (2009), vol. 42, pp. 624-632.

[17] R. W. Hall, “Patient Flow: Reducing Delay in Healthcare Delivery”. *International Series in Operations Research & Management Science* (2006),vol. 91, pp. 1-45.

[18] G. Righini, M. Trubian, “Data-dependent bounds for the General and the Asymmetric Stacker-Crane problems”. *Discrete Applied Mathematics* (1999), vol. 91, pp. 235-242.

[19] A. Coja-Oghlan, S.O. Krumke, T. Nierhoff, „A heuristic for the stacker crane problem on trees which is almost surely exact“. *Journal of Algorithms* (2006), vol. 61, no. 1, pp. 1-19.