Decentralized Online Scheduling of Combination-Appointments in Hospitals

Ivan Vermeulen[†], Sander Bohte[†], Sylvia Elkhuizen[‡], Piet Bakker[‡], Han La Poutré[†,§]
[†]Centrum Wiskunde & Informatica (CWI), Amsterdam, The Netherlands
[‡]Academic Medical Center (AMC), University of Amsterdam, The Netherlands
[§]Faculty of Technology Management, Eindhoven University of Technology, The Netherlands
Email: I.B.Vermeulen@cwi.nl

Abstract

We consider the online problem of scheduling combination appointments for outpatients. Scheduling multiple appointments on a single day is high on the list of outpatient preferences. It is hard to achieve for two reasons: first, due to the typical distributed authority in hospitals, scheduling combination appointments requires coordination between departments. Second, there is a trade-off between local scheduling efficiency and the fulfillment of patient scheduling preferences. We present a multi-agent approach, where patient agents coordinate with department agents. For individual departments, we design an efficient yet flexible local scheduling method with dynamic usage of capacity. Department agents use this method and use its flexibility to trade-off local efficiency against making single day combination appointments. We show in a stylized model of a real hospital setting that this multi-agent scheduling approach is highly effective and allows a hospital to set a desired level of efficiency versus fulfilled patient preferences.

Introduction

Efficiency is increasingly becoming a crucial aspect in hospitals. The demand for care is increasing, and rising costs will not fit within limited budgets. To cope with this, hospitals will have to improve on a strategic level as well as on an operational level (Vissers and Beech 2005). Furthermore, at the same time patients are asking, and hospitals are offering, more patient involvement and personalized care, medically and also logistically (Elkhuizen 2007). Patient service and resource efficiency will have to increase simultaneously.

Based on a practical case study, we research an online scheduling problem, where per patient one or more activities have to be scheduled at auxiliary departments. Departments have local scheduling authority, each department is primarily concerned with its own performance (Decker and Li 1998). This makes fulfilling patient preferences involving multiple departments difficult. We present an approach that allows fulfilling patient preferences while maintaining the desired performance levels at the local departments involved.

Optimizing local performance is often complex. A department receives appointment-requests from multiple in- and outpatient departments with varying medical properties and urgencies. With a fixed resource capacity, appointments must be scheduled such that for all urgency levels a satisfactory fraction of patients is scheduled on time. The typical approach to this problem is to classify patients into patient groups based on their urgency and medical properties, and allocate parts of the resource capacity to each patient group. Allocating capacity specifically can indeed improve performance depending on problem properties (van Dijk 2002). However, due to fluctuations in patient arrivals, initial capacity allocation can regularly mismatch current demands. Performance of scheduling approaches with a static allocation can therefore be significantly improved by using a dynamic optimization of allocation (Vermeulen et al. 2007a).

We consider multiple departments in the problem of scheduling combination appointments for outpatients. The service of scheduling a patient's appointments to a single day, a successful combination appointment, is high on the list of outpatient preferences and therefore of great importance to the hospital (Elkhuizen 2007). Due to distributed scheduling authority, making combination appointments requires coordination between departments each solving a complex local scheduling problem.

In current practice, coordination is very hard, because typically the only way for a department to maintain local schedule efficiency is to close off access to the resource calendar for external parties. Appointment-requests arrive on paper or by phone, and the actual scheduling is done by local schedulers that oversee the resource calendar. This makes coordination between departments impossible or requiring many phone calls. The incentive to allow external access to the resource calendar is limited; experience has shown this can greatly reduce local performance, due to the lack of overview of external schedulers on local efficiency.

In this paper, we present a multi-agent approach that respects the distributed nature of the hospital. Department agents represent local department scheduling objectives, and patient agents coordinate scheduling patient appointments by interacting with department agents. Department agents use a dynamic local schedule method, and can efficiently trade off local performance against the opportunity for the patient agent to make combination appointments. By setting the rules for their own local agent and for its interactions with other agents, departments can trust that their objectives are well maintained, while allowing the coordination of appointments.

We introduce a 'schedule-cost' function for the local problem of scheduling patients from multiple groups with a dynamic usage of capacity (which generalizes the ideas of dynamic schedule adjustments (Vermeulen et al. 2007a)). Department agents use this function for patient scheduling and evaluating trade-offs against performance. The schedule-cost function assigns a cost value to each timeslot given the current patient to be scheduled and the current state of the resource calendar, which allows coordination for scheduling combination appointments. Each department agent will make a set of timeslots up to a certain cost available to the patient agent, and the patient agent will select and combine timeslots to make combination-appointments. Theoretically, scheduling performance is maximized with an optimal cost function. Here, we design an approximately correct function by identifying the main properties of such a function.

The trade-off between local performance and successful combination appointment scheduling, is realized as a function of the maximal cost of timeslots offered by the department agents. Each department agent offers a set of appointment timeslots that are of acceptable cost for the efficiency of the department. Patient agents then select from these sets of timeslots such that their preferences are met. The higher the maximal cost of timeslots offered, the larger the change that the patient agent can make a successful combination appointment.

We demonstrate the success of the cost-based coordination approach in a wide range of experiments with a stylized patient scheduling model. The model is derived from a real-world case study at the Academic Medical Center (AMC) of the University of Amsterdam, and our approach is designed as an applicable method in practice. We show that cost-based coordination achieves high percentages of successful combination-appointments with a very small decrease in local scheduling efficiency. Furthermore, the cost-based local scheduling approach is shown to dominate approaches with static capacity by a large difference. Even in busy scenarios, cost-based coordinated scheduling achieves high service levels for all patient groups.

Related Work

Optimizing hospital logistics is a complex problem; logistic improvements are not easily applied. Traditional OR solutions in the hospital are often of limited help, as they are typically applied to static or centralized planning and scheduling problems (for an overview, see (Spyropoulos 2000)). Scheduling literature, specifically online scheduling (Pruhs, Torng, and Sgall 2004) and (open) shop scheduling problems (Brucker 2001), discuss properties related to our presented problem. However, most of this work does not consider local information, or online multi-objective scheduling. Complex multi-objective optimization are more commonly addressed with computational intelligence techniques such as genetic algorithms (Goldberg 1989).

Scheduling combination appointments involves complex local scheduling. Work such as (Patrick and Puterman 2007; Vermeulen et al. 2007a) discuss such problems. Also, in (Green, Savin, and Wang 2006) the authors discuss a local profit maximization problem of a MRI scheduling problem for three classes of patients. Their more abstract model requires setting specific revenue and penalty functions, for which the authors identify properties of an optimal solutions. The authors focus on local performance and do not consider a trade-off against patient preferences.

The distributed organization of typical hospitals suggests that a distributed scheduling system is most suitable to solve this problem. Multi-agent systems have been effectively used for coordination between autonomous parties (Weiss 1999; Nealon and Moreno 2003), including the patient scheduling problem (Paulussen et al. 2003; Decker and Li 1998; Vermeulen et al. 2007b). However, these works focus on either rescheduling or conflict resolution after an initial scheduling. Here we focus on achieving high performances of online initial scheduling, where patients are scheduled in turn without having to reschedule. In (Wellman et al. 2001) a first step is taken toward distributed online scheduling. The authors consider a virtual-market approach, and discuss different types of auction mechanisms. Their results show that it is hard to find a general solution.

Patient Scheduling Model

Our patient scheduling model is based on a case study in the AMC hospital, which treats more than 350,000 outpatients annually. Based on many site-visits and historical data, as well as discussions with medical experts we create a stylized model that captures the complexity of scheduling combination appointments at the AMC.

Patients

Between a patient's consults at an outpatient clinic, one or more tests or treatments are performed by auxiliary departments, see Figure 1. Especially in the diagnostic phase, multiple tests are often ordered to be done before the next consult. We focus on the problem of scheduling these tests at the different departments.

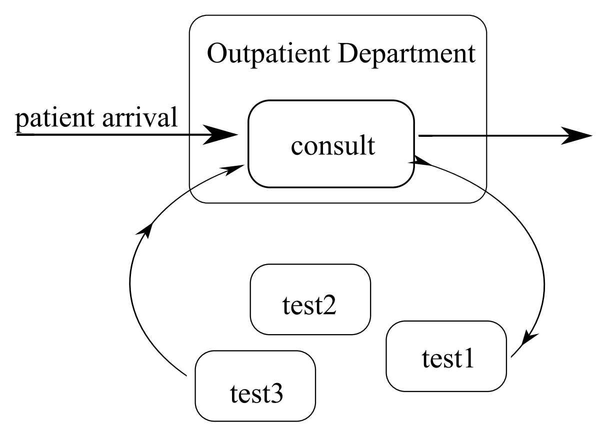

Figure 1: **One or more tests between consults, to be performed at different departments.**

Patients arrive over time, and have to be scheduled one by one. Each patient has a set s_i of one or more activities $\{a_{i,0}, ..., a_{i,n}\}$. Each activity $a_{i,j}$ has to be scheduled at the associated department $\text{DEP}(a_{i,j})$. Activities are taken to not have precedence constraints (which is usually the case for diagnostic tests). A patient's schedule window SCHEDWIN_i is the time-frame in which all his activities must be scheduled. The size of the schedule window expresses the urgency of a patient. Patients are divided into patient groups such that patients with the same urgency are in the same patient group $G(p_i)$.

To successfully schedule a combination appointment, for a patient with multiple activities, all activities must be scheduled on a single day, with a minimum transfer time (MINBTWN) and maximum waiting time (MAXBTWN) between following appointments. We additionally include a patient preferences model that determines which timeslot or combination appointment is preferred by a patient. Table 1 gives an overview of patient variables.

Table 1: Patient Variables

variable	**description**
p_i	patient i
s_i	set of activities of patient p_i
$a_{i,j}$	activity j of patient p_i
SCHEDWIN_i	schedule window of patient p_i
$G(p_i)$	patient group p_i belongs to
MINBTWN	minimum time between two appointments on a single day
MAXBTWN	maximal time between two appointments on a single day
P	patient preferences model

Departments

In our model we define a set of departments D. All activities $a_{i,j}$ with $\text{DEP}(a_{i,j}) = d \in D$ must be scheduled to exactly one timeslot on respective department d's resource calendar RC_d. Each day on the RC_d is partitioned into a number of non-overlapping timeslots ts, each defined by a start time, end time, day and department $[t_s, t_e, \text{day}, d]$. For RC_d, the total number of timeslots per day is TOTAL_d and the duration of each timeslot DUR_d. We abstract from general resource calendars by assuming all appointments at a department have the same duration.

Scheduling methods use a capacity allocation: on each recource calendar RC_d, a set of timeslots, with size $\text{SIZE}_{d,g}$, is allocated to patient group g on each day, with $\sum_g \text{SIZE}_{d,g} = \text{TOTAL}_d$. Table 2 gives an overview of the department variables.

Objective

The local scheduling objective in each department is to have an acceptable service level for each group, the service level $SL_{d,g}$ is defined as the fraction of patients from group g in department d scheduled within their schedule window. This is a typical performance indicator in the hospital. Hospital

Table 2: Department Variables

variable	**description**
D	set of departments
RC_d	resource calendar of department d
TOTAL_d	total number of timeslots
DUR_d	duration of a timeslot
$G(ts)$	patient group timeslot ts is allocated to
$\text{SIZE}_{d,g}$	number of timeslots allocated to group g

policy determines the importance of each group in the overall scheduling objective. Here a department performance is defined by the minimum service level of all its groups: $MSL_d = \min_g(SL_{d,g})$.

To evaluate local performance in scenarios with multiple departments, we take the average (**aMSL**) of the departments' MSL_d

$$aMSL = \frac{\sum_{d \in D} MSL_d}{|D|}$$

which measures the combined effect of a scheduling approach on local performance.

To measure successful combination appointment scheduling, we take the fraction (**CA**) of all partial plans with two or more activities all scheduled within their schedule window, that are successfully scheduled as a combination appointment.

To trade-off successful combination appointment scheduling against local performance, we define our overall multi-objective O as

$$O = \max \begin{bmatrix} \text{CA} \\ aMSL \end{bmatrix},$$

for which we will present a Pareto front of $(\text{CA}, aMSL)$ solutions (the trade-off between our two objectives).

Cost-Based Coordination

In our multi-agent system, department agents control local schedule efficiency, and patient agents coordinate combination appointments. The departments set rules for their department agents, and the hospital sets rules for patient agents, patients can provide preferences to their patient agent.

We consider the trade-off between local performance (aMSL) and making combination appointments (CA). Department agents rank all available timeslots with respect to schedule efficiency, and each department agent offers a set of the most efficient timeslots to the patient agent. The patient agent then selects and combines timeslots to make combination appointments. In our approach, timeslots are ranked based on a schedule-cost function.

Schedule-Cost Function

We define a schedule-cost function for ranking timeslots. It assigns a cost value to each timeslot given the current patient to be scheduled and the current state of the calendar. In theory, if a cost function is optimally defined, selecting the timeslot with the lowest cost for each patient maximizes local performance (MSL). We identify some of the main properties of such a function, which generalizes and extends the ideas from (Vermeulen et al. 2007a), and design a parameterized function with those properties.

Let a scheduling-cost function $SC_d(p, ts)$ give the cost of scheduling patient p to timeslot ts in department d. The cost function $SC_d(p, ts)$ includes the following considerations which we explain in more detail below:

1. timeslots on days at the beginning of the schedule window have reduced cost to avoid wasting capacity,
2. to balance the usage of timeslots over days, timeslots on days with many free timeslots have reduce cost,
3. the cost of a timeslot is dependent on the patient group the timeslot is allocated to.

We discuss the implementation of these properties next, and define how they are balanced in the overall schedule cost function. As an initial abstraction, all timeslots on the same day allocated to the same group have the same cost.

Earlier timeslots can be used by less future patients than later timeslots, and not using them can result in wasted capacity. We therefore make earlier timeslots cheaper, all else equal. We calculate this "lateness" of a day on the resource calendar as:

$$\text{LATE}_d(ts) = \min\left(1, \max\left(0, \frac{\text{DAY}(ts) - \text{FROM}_{\text{G}(ts)}}{\text{TILL}_{\text{G}(ts)} - \text{FROM}_{\text{G}(ts)}}\right)\right),$$

with timeslot ts on day $\text{DAY}(ts)$, and $\text{G}(ts)$ the group timeslot ts is allocated to, $\text{FROM}_{\text{G}(ts)}$ and $\text{TILL}_{\text{G}(ts)}$ are the beginning and end of the schedule window associated with group $\text{G}(ts)$.

To cope with unexpected peaks in demand, available capacity is best spread out evenly over days. To achieve this, we let timeslots on days with more free timeslots be cheaper, all else equal. We calculate this "fullness" of a day on the resource calendar as:

$$\text{FULL}_d(ts) = \max\left(0, 1 - \frac{\text{FREE}_{d,\text{G}(ts)}(\text{DAY}(ts))}{\text{SIZE}_{d,\text{G}(ts)}}\right)$$

with $\text{FREE}_{d,\text{G}(ts)}(\text{DAY}(ts))$ the number of free timeslots allocated to group $\text{G}(ts)$ on the day of ts.

Working from the findings of (Vermeulen et al. 2007a), we relax the definition of $\text{FULL}_d(ts)$. If there are any timeslots with cost 0 on $\text{DAY}(ts)$ allocated to any group, these timeslots can always be used and are therefor included in $\text{FREE}_{d,\text{G}(ts)}$. Furthermore, timeslots with cost 0, on the day before $\text{DAY}(ts)$ (if still in the schedule window) are also included in $\text{FREE}_{d,\text{G}(ts)}$ with a maximum of $\frac{\text{SIZE}_{d,\text{G}(ts)}}{2}$.

Our schedule cost function uses the allocation of timeslots to patient groups for maximizing performance (MSL_d). The cost of a timeslot is dependent on the group it is allocated to ($\text{G}(ts)$) and on the group of the current patient to be scheduled ($\text{G}(p)$). Timeslots before the beginning of the patient's schedule window must not be used, and therefore have cost sc_{max} (defined as a maximum cost value, at least twice as high as any other timeslot cost). Timeslots on days before the beginning of the schedule window of the patient group they are allocated to, have cost 0 to not waste capacity. We define overall cost $SC(p, ts)$ as

$$SC(p, ts) =$$

$$\begin{cases} sc_{\text{max}} & \text{if} \quad \text{DAY}(ts) < \text{FROM}_{\text{G}(p)} \\ 0 & \text{else if} \quad \text{DAY}(ts) < \text{FROM}_{\text{G}(ts)} \\ \beta_{d,\text{G}(ts)} * \left(\text{LATE}_d(ts)^{p1} + \text{FULL}_d(ts)\right)^{p2} & \text{else} \end{cases}$$

with parameter $p1$ scaling the trade-off between lateness and fullness of the day of timeslot ts, with typically $p1 \leq 1$. Parameter $p2$, with typically $p2 > 1$, sets the increase of marginal cost between timeslots ordered on increasing cost (if less timeslots are available cost increases more steeply). The factor $\beta_{d,\text{G}(ts)}$ scales the cost of a timeslots function per patient group. The value of $\beta_{\text{G}(ts)}$ is dependent on the size of the patient group ($\text{SIZE}_{d,\text{G}(ts)}$) (in MSL_d all groups are evaluated with equal weight, although they are of different size).

Based on our cost function, we define two scheduling methods: First Come Least Cost (**FCLC**) and First Come Maximal Relative Cost (**FCMRC(**mrc**)**). When scheduling patients with **FCLC**, the cheapest available timeslot is selected. If timeslots have equal cost, the earliest timeslot is selected.

When scheduling a patient with **FCMRC(**mrc**)** the patient agent can select the preferred timeslot from all timeslots with a cost maximal the cost of the cheapest timeslots plus the parameter mrc (maximal relative cost). Using FCMRC(mrc), the trade-off between scheduling most efficiently and freedom in selecting a timeslot is set by parameter mrc, with $mrc \leq \frac{sc_{\text{max}}}{2}$. If mrc is 0 only the most efficient timeslot(s) (with the lowest cost value) are available for scheduling. If mrc is $\frac{sc_{\text{max}}}{2}$, all timeslots allocated to the patient's group in or after the patient's schedule window (and cheaper timeslots allocated to other groups) are available for scheduling. Note that FCMRC($mrc = 0$) is almost equal to FCLC, the only difference is that if there are multiple timeslots with the lowest cost, in FCLC the earliest timeslot is selected, and in FCMRC($mrc = 0$) the patient agent can select a preferred timeslot from those with the lowest cost.

We will separately present results of our cost-based scheduling approach for a single department, including the trade-off between local efficiency and fulfilling patient preferences.

We use two benchmark approaches with a static allocation of capacity to patient groups where patients are not allowed to make use of capacity not allocated to their group: First Come First Serve (**FCFS**), and First Come Randomly Served (**FCRS**). When scheduling patients with FCFS, the earliest available timeslot allocated to group $\text{G}(p)$ is selected. Given optimal allocation of capacity to patient groups, FCFS is the most efficient static scheduling approach (Hopp and Spearman 2000). When scheduling a patient with **FCRS**, the patient agent can select a preferred timeslot from all available timeslots within SCHEDWIN_p allocated to group $\text{G}(p)$. If there are no such timeslots available, the activity is scheduled to the earliest timeslot allocated to $\text{G}(p)$ after the SCHEDWIN_p. Such static settings are typical in many practices and approaches (Vissers and Beech 2005).

Coordination

Patient agents interact with department agents to coordinate scheduling combination appointments. A patient agent initiates scheduling by requesting a set of timeslots from each department agent corresponding to an activity of the current patient. Given the cost based local scheduling method above, we present a coordination method that allows making combination-appointments at high local efficiency, and where additionally the trade-off between local efficiency and successful combination appointments can be set.

Department agents use FCMRC(mrc). For the current patient to be scheduled, each department makes a set of timeslots (with at most the cost of the cheapest timeslot plus mrc) available to the patient agent. The patient agent selects timeslots to make combination-appointments. Our experimental results will show that higher values of mrc will increase opportunities for making combination appointments, but also reduces local efficiency.

If no combination appointment can be made, the activities from the patient's partial plan will be scheduled independently. We consider two methods for selecting a combination-appointment if there are multiple possibilities available.

COOR-LC The patient agent selects the combination-appointment with the lowest summed cost of individual timeslots, ensuring that the most efficient combination-appointment is selected.

COOR-RC The patient agent selects the preferred combination from all combination-appointments. This approach takes patient preferences into account but is less efficient than COOR-LC.

As a benchmark approach for coordination we consider COOR-RC with FCRS: each department agent makes all available timeslots within the patient's SCHEDWIN_p allocated to $\text{G}(p)$ available to the patient agent. The patient agent combines timeslots into combination appointments, and the selects the preferred combination.

Experiments

Experimental Setup

We present results for a wide range of settings of a typical case in hospital practice. To capture the complexity of combination appointment scheduling, we define a scenario with four departments, differing in number of timeslots per day. We present results averaged over 50 runs per setting. Each run simulates 40 weeks of patient scheduling, of which we measure over the last 16 weeks. Results of conducted experiments with other settings are comparable to those presented here.

We define a Poisson arrival process with rate λ for patient arrival. We compare results between different patient arrival rates in the experiments discussed below, we report the utilization level which is a direct result of the value of λ. Each newly arrived patient p_i is 'diagnosed' with a set of activities s_i and urgency SCHEDWIN_i, given a distribution over all combinations of allowed activity sets and urgencies. Each department has to schedule patients with three different schedule windows: non-urgent (schedule between 2-14 days after the current day), semi-urgent (2-7), urgent (0-2). At all departments, the non-urgent group is the largest group (46% of all patients), and the urgent group the smallest (24%).

Non-urgent and semi-urgent patients can have one, two, or three activities to be scheduled at different departments. Urgent patients only have single activities. As typical practical values, 80% of all patients have a single activity, 14% have two activities (2-combi), and 6% have three activities (3-combi). Note that this corresponds with the percentage of appointments part of a combination varying between 60% and 46% for non-urgent groups, and between 35% and 30% for semi-urgent groups. The minimum transfer time between appointments scheduled on a single day is 45 minutes, and the maximum waiting time is 135 minutes.

The patient preference model P is taken to be the following: for non-urgent patients, we assume the aggregated patient preferences over timeslots is a uniform random distribution. A patient will select a random timeslot from the set of timeslots offered. The size of the set of offered timeslots is dependent on the value of mrc in our FCMRC(mrc) approach. We can measure how well patient preferences, apart from combination appointments, can be fulfilled by the number of different timeslots offered to a patient. For urgent patients the most efficient timeslot is selected (experiments where urgent patients could also select a preferred timeslot, had a slightly worse aMSL value, at most 0.05 difference). Similarly, in COOR-RC we use P (uniform random distribution) to select a preferred combination appointment.

The four departments in our scenario respectively, have 40, 30, 30, and 12 timeslots per day. Less timeslots means higher variability in demand, which reduces efficiency. For each department, we experimentally determined the static capacity allocation to patient groups ($\text{SIZE}_{d,g}$ values), used by approach FCFS and FCRS, and as initial allocation in FCMRC(mrc). Due to the discreteness and limited number of timeslots this search space is not large. Resource utilization in our simulations is higher than in hospital practice, which has additional stochastic elements that are of only limited direct influence on the scheduling complexity of combination appointments.

All department agents use FCMRC(mrc) for local scheduling, and interact with patient agents that use COOR-LC or COOR-RC. We have manually determined the set of cost-function parameters that worked well for the whole range of settings of our presented experiments. We vary the parameter mrc between $[0, \frac{sc_{\max}}{2}]$ to trade-off the opportunity for making combination appointment against local efficiency. We show the resulting trade-off by presenting Pareto fronts of $(CA, aMSL)$ solutions. We benchmark against FCFS (without coordination) and FCRS (with coordination).

Local Scheduling

We first evaluate our cost-based scheduling approach FCMRC(mrc) for local scheduling at a single department. By varying the mrc value the number of timeslots offered to a patient (x-axis) increases, but the department performance

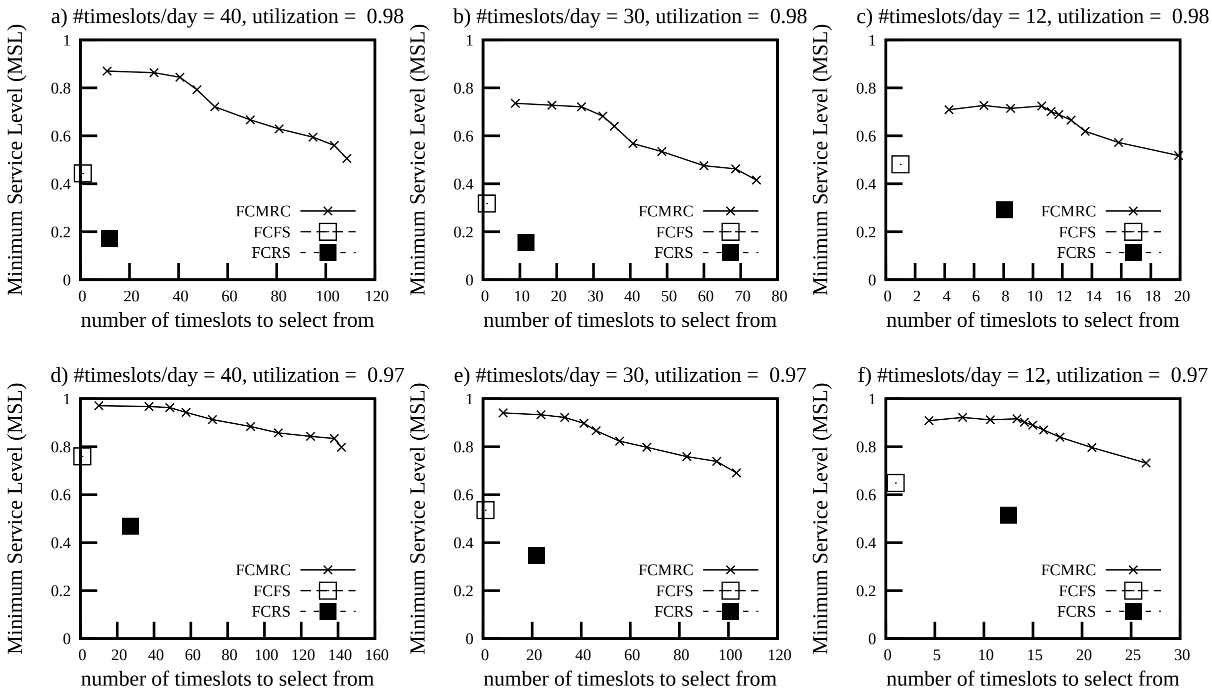

Figure 2: **Performances for six cases of single department scheduling. Departments have either 40 [a,d], 30 [b,e], of 12 [c,f] timeslots per day, and utilizations is either** 0.98 **[a-c] or** 0.97 **[d-f]**

(MSL, y-axis) decreases. In Figure 2a-f, we show performance results for 6 different cases: three departments with a different number of timeslots per day, and two utilization levels.

In all experiments FCMRC(mrc) dominates static benchmark approaches FCFS and FCRS with a large difference. It solves the local dynamic scheduling problem very efficiently. FCMRC(mrc) achieves the highest MSL_d levels, a high service level for all groups, and has an efficient trade-off between MSL_d and timeslot selection freedom. In all cases FCMRC($\frac{sc_{\max}}{2}$) with the highest timeslots selection freedom (extreme right data point of FCMRC) still has a better MSL_d performance than FCFS.

The performance of FCLC is not presented separately, it is very close to FCMRC(0) (extreme left data point of FCMRC). We have also implemented the approach of (Vermeulen et al. 2007a) in our simulations, which can not make a trade-off between objectives. Its performance is similar to FCMRC($\frac{sc_{\max}}{2}$).

Combination-Appointments

We schedule combination appointments with COOR-LC and COOR-RC, given departments using FCMRC(mrc). In Figures 3a-f, we show our approach achieves a high rate of successful combination appointments, with limited decrease in local performance. In the six figures of 3, we present results for scheduling 2-combi [a] and 3-combi [b] appointments, and specifically non-urgent 2-combi [c] and semi-urgent 2-combi [d]. Figures 3e,f show the results of scheduling 2-combi [e] and 3-combi [f] appointments at a lower utilization rate.

FCMRC(mrc) with varying mrc values, successfully finds efficient schedules, and allows a high performance trade-off between objectives. Our approach allows coordination of appointments while maintaining desired performance levels at the local departments involved. With increasing mrc values, patients have more selection freedom, also in selecting a preferred combination appointment with COOR-RC. The difference in performance between COOR-LC and COOR-RC is only significant when high levels of CA are desired.

Due to efficient trade-off, there is only a small difference in results of scheduling 2-combi's and scheduling 3-combi's (Figures 3a and 3b). Furthermore, both non-urgent and semi-urgent combination appointments (Figures 3c and 3d) are efficiently scheduled. Scheduling semi-urgent combinations is more difficult if non-urgent appointments can use of semi-urgent timeslots due to high mrc values,

Conclusions

For single day combination appointments, the online patient scheduling problem is a dynamic and distributed multi-objective optimization problem for which, to the best of our knowledge, there is no general solution.

We present a distributed approach where parties each have a local scheduling method that computes an efficiency cost for scheduling a certain patient to a certain timeslot. The cost function we design generalizes the dynamic scheduling method of (Vermeulen et al. 2007a). By allowing pa-

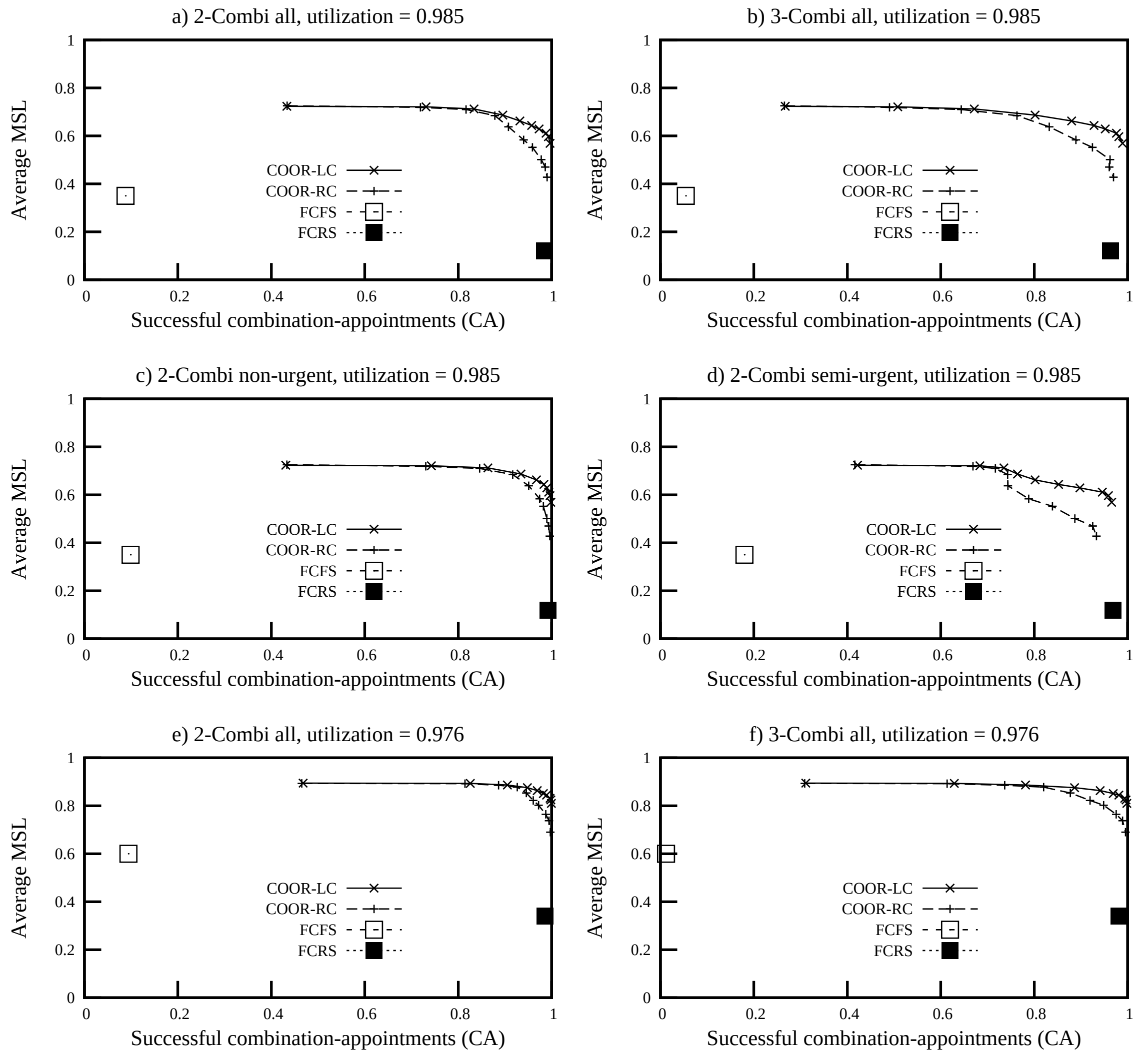

Figure 3: **Results for two utilization levels:** 0.985 **[a-d] and** 0.976 **[e,f], 2-combi's [a,e] and 3-combi's [b,f], and 2 combi's non-urgent [c] and semi-urgent [d].**

tient agents to choose timeslots over a range of costs, the patient agent can schedule single day combination appointments. Our results show that this method works efficiently, and can achieve a Pareto front of solutions trading-off timeslot selection freedom and schedule efficiency, as a function of the cost-range of the timeslots offered to patient agents.

Our approach is designed with its practical application in mind. Our problem model is stylized based on many site-visits and discussions with hospitals experts, and captures the complexity of scheduling combination appointments. Our assumptions and case settings are based on practical values, and practical validation is partly based on the expertise of the consulted medical experts and proposed users of our system. Furthermore, the experimental results are robust for a wide range of scenarios, and clarify the potential effectiveness within a real scenario.

The scheduling problem we consider is inherently distributed; there is always a trade-off between objectives of different departments and between departments and patient preferences. A centralized approach, although not applicable in practice, will have to make the same trade-off choices.

In future work, we would like to refine the trade-offs between scheduling objectives even further. This will include a more detailed patient preference model, such that individual preferences can be considered and valuated, and the cost-range of scheduling will be dependent on individual patient attributes. We will also continue work on generalizing our cost-based approach for complex local scheduling, including online optimization of the cost-function parameters.

Our approach allows a hospital or department to set a desired level of efficiency versus fulfilling patient preferences. In future work we additionally consider online mechanisms for dynamically setting this trade-off at each department.

References

Brucker, P. 2001. *Scheduling Algorithms*. Springer-Verlag.

Decker, K., and Li, J. 1998. Coordinated hospital patient scheduling. In *In Proceedings of the Third International Conference on Multi-Agent Systems (ICMAS98)*, 104–111.

Elkhuizen, S. 2007. *Patient Oriented Logistics*. Academic Medical Center, University of Amsterdam.

Goldberg, D. E. 1989. *Genetic Algorithms in Search, Optimization and Machine Learning*. Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc.

Green, L.; Savin, S.; and Wang, B. 2006. Managing patient service in a diagnostic medical facility. *Operations Research* 54:11–25.

Hopp, W. J., and Spearman, M. 2000. *Factory Physics: The Foundations of Manufacturing Management*. McGraw-Hill.

Nealon, J., and Moreno, A. 2003. Agent-based applications in health care. In Nealon, J., and Moreno, A., eds., *Applications of Software Agent Technology in the Health Care Domain*. Birkhueser Verlag. 3–18.

Patrick, J., and Puterman, M. 2007. Improving resource utilization for diagnostic services through flexible inpatient scheduling: A method for improving resource utilization. *Journal of the Operational Research Society* 58(Feb):235–245.

Paulussen, T. O.; Jennings, N. R.; Decker, K.; and Heinzl, A. 2003. Distributed patient scheduling in hospitals. In Gottlob, G., and Walsh, T., eds., *Proceedings of the Eighteenth International Joint Conference on Artificial Intelligence*. Morgan Kaufmann. 1224–1232.

Pruhs, K.; Torng, E.; and Sgall, J. 2004. Online scheduling. In Leung, J. Y.-T., ed., *Handbook of Scheduling: Algorithms, Models, and Performance Analysis*. CRC Press. chapter 15, 15.1–15.41.

Spyropoulos, C. D. 2000. Ai planning and scheduling in the medical hospital environment. *Artificial Intelligence in Medicine* 20(2):101–111.

van Dijk, N. M. 2002. Making simulation relevant in business: to pool or not to pool? ”the benefits of combining queuing and simulation”. In *WSC ’02: Proceedings of the 34th conference on Winter simulation*, 1469–1472. Winter Simulation Conference.

Vermeulen, I.; Bohte, S.; Elkhuizen, S.; Lameris, J.; Bakker, P.; and La Poutré, J. 2007a. Adaptive optimization of hospital resource calendars. In Bellazzi, R.; Abu-Hanna, A.; and Hunter, J., eds., *11th Conference on Artificial Intelligence in Medicine*, volume 4594 of *Lecture Notes in Computer Science*. Springer. 305–315.

Vermeulen, I.; Bohte, S.; Somefun, D.; and La Poutré, J. 2007b. Multi-agent pareto appointment exchanging in hospital patient scheduling. *Service Oriented Computing and Applications* 1(3):185–196.

Vissers, J., and Beech, R. 2005. *Health Operations Management*. Routledge.

Weiss, G., ed. 1999. *Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence*. Cambridge, MA: The MIT Press.

Wellman, M.; Walsh, W.; Wurman, P.; and MacKie-Mason, J. 2001. Auction protocols for decentralized scheduling. *Games and Economic Behaviour* 35:271–303.