Dynamic appointment scheduling with patient time preferences and different service time lengths

Anne Zander[1] and Uta Mohring[1]

[1]*Karlsruhe Institute of Technology, Kaiserstr.12, 76131 Karlsruhe, Germany*
anne.zander@*kit.edu; uta.mohring@student.kit.edu*

Abstract. In this paper we consider the appointment schedule of a physician's day. We assume patient types defined by different time preferences and service time lengths. Patient requests for the day are handled directly during a booking horizon. We present a mixed integer linear programming model to determine a set of appointments to offer a patient requesting an appointment. The objective is to schedule the requesting patient while also taking future demand into account. We want to maximize the overall utilization assuring a certain fairness level. We further implement a simulation in order to test the mixed integer linear program and to compare it to simpler online heuristics.

Keywords: Appointment Scheduling; Patient Preferences; Mixed Integer Linear Programming; Simulation

---

Introduction

To increase efficiency in outpatient clinics is one of the current issues in European health care systems. Due to the demographic development and the increase of chronic and psychological diseases outpatient clinics are facing an augmenting number of appointment requests that often goes beyond the clinics' capacities (Europäische Kommission 2014). One of the important adjusting screws to increase the efficiency of clinical processes in outpatient clinics is the design of their appointment scheduling system. A lot of publications develop appointment scheduling models for outpatient clinics. Apart from a few exceptions, the clear majority of these appointment scheduling models does not take patient time preferences into consideration.
The studies of (Gerard et al. 2008) and (Cheraghi-Sohi et al. 2008) surveying patients in British outpatient clinics underline the relevance of individual patient time preferences when arranging appointments. According to surveys of Rohleder and Klassen a quarter of the patients has concrete time preferences (Klassen and Rohleder 1996). Considering patient time preferences when arranging appointments has even positive effects for the outpatient clinics: Patients are highly satisfied, so that both the number of no-shows and the number of patient migrating to other clinics decrease (Feldman et al. 2014). The consideration of patient time preferences indicates a certain extent of appointment flexibility for the patients. This appointment flexibility can however result in a highly variable daily capacity utilization, so that outpatient clinics should carefully decide how much appointment flexibility they want to offer to their patients (Feldman et al. 2014).

Literature Review

The procedure of the presented appointment scheduling model for assigning appointment requests to free time slots is based on the dynamic optimization model of (Hahn-Goldberg 2014). Her main idea is to improve the appointment scheduling of a chemotherapy outpatient clinic by a combination of proactive and online optimization. Firstly, a proactive template is generated based on the expected appointment requests for the following day. Afterwards, this template is used to assign free appointments to the current appointment requests. In situations in which the template does not contain a suitable appointment for the actual request the template is updated with regard to the already scheduled appointments and the actual request. (Rohleder 2000) integrate patient time preferences into their simulations. They compare different appointment scheduling rules concerning their effects on patients' waiting times, physicians' idle times and the returns of the clinic. In order to consider patient time preferences, they distinguish between normal and special appointment requests. Special requests are characterized by a probability distribution describing their time preferences. Rohleder and Klassen consider two different probability distributions: a uniform distribution and a so called end of period distribution characterizing patients who prefer appointments at the end of the period. Two performance criterions are added to the simulations in order to evaluate the different scheduling rules concerning the consideration of patient time preferences: the proportion of special request patients receiving the specific appointment requested and the proportion of special request patients not receiving any appointment.

In the dynamic appointment scheduling model of (W.-Y. Wang and Gupta 2011) an appointment request is characterized by a set of preferred appointment times and a preferred physician. The clinic tries to assign an acceptable time-physician-combination to each appointment request while maximizing its returns. Since the clinic return function considers the probabilities of patients accepting their offered appointment, this model implicitly takes patient time preferences into consideration. Wang and Gupta solve this optimization problem with two different heuristics. (Feldman et al. 2014) start from the situation, where a certain set of appointment days is offered to each patient and then each patient chooses his preferred appointment day out of this set. The optimization model determines for each possible set of appointment days the optimal probability with which this set of days should be offered to the patients. In the present connection a probability distribution is called optimal if it maximizes the number of patients showing up for their appointments with regard to the clinic capacity. In this way patient time preferences are implicitly taken into consideration. Feldman et al. realize these ideas in a static model, which they solve exactly. In a further step the static model is extended to a dynamic one which additionally considers the current state of the scheduled appointments. This variation is solved by means of a heuristic. Feldman et al. only focus on the appointment days, they do not determine the concrete appointment times of the patients.

Similar to the model of Wang and Gupta, (J. Wang and Fung 2015) characterize a patient appointment request by an set of preferred appointment times and a preferred physician. Here the appointment length is the same for all patients. In contrast to Wang and Gupta the clinic does not offer a concrete time-physician-combination, but it offers some time-physician-combinations to each patient. The offered combinations maximize the expected return of the clinic with regard to the patient preferences and the current state of the scheduled appointments. Finally, the patient either chooses one of the offered combinations or withdraws his appointment request. This process is modeled as a dynamic program. The model is solved via an approximate dynamic programming approach using an LP-formulation with an affine approximation of the value function.

Model

In our model we consider the appointment schedule of a physician’s day. We assume that there exists a booking horizon during which patients are able to call or to go online to book an appointment for that day. We further assume that patients can be divided into different patient types. Here, a type is defined by a service time length (time of treatment that is needed for a patient of that type) and by time preferences with respect to the possible appointment slots. For every patient type we assume that the request arrival process for appointments during the booking horizon is Poisson. We presume that the patient type of every incoming request is known and that the request has to be handled right away. The considered day is divided into T equal time intervals. Every possible service time length is a multiple of the interval length. To handle a request, a set of appointments matching the patient’s service time length has to be offered (overlapping is not allowed). The patient then chooses one of these appointments or rejects and leaves. The challenge is to offer patients a set of fitting appointments such that the probability that this patient accepts one of them is high. At the same time we want to take into account future demand. The overall goal hereby is to maximize the utilization of the schedule or equivalently to minimize the unused time intervals at the end of the booking horizon. Here, we suppose that patients who accepted an appointment will show up. In order to determine a set of appointments to offer to an incoming request we solve a mixed integer linear program. This model considers already assigned appointments. Further, it tries to schedule the incoming request and the expected future request by reserving appointments for every patient type. The objective function maximizes the expected utilization of the schedule assuming that every reserved appointment is offered to one patient of the corresponding type. The resulting reserved appointments for the requesting patient's type are potential appointments to offer to him or her. In Table 1 the sets, parameters and variables are defined. Our model then results in:

$$\max \quad \sum_{k \in K} \sum_{t \in T} P_{kt} \cdot D_k \cdot x_{kt}$$

$$\text{s. t.} \quad \sum_{t \in T} x_{kt} + d_k = N_k + A_k \qquad \forall k \in K \qquad (1)$$

$$(t + D_k) \cdot x_{kt} \leq |T| \qquad \forall k \in K, t \in T \qquad (2)$$

$$\sum_{k \in K} \sum_{t' \in T_{kt}} x_{kt'} \leq 1 \qquad \forall t \in T \qquad (3)$$

$$d_k - \frac{N_k \cdot \left(\sum_{i \in K} D_i \cdot (A_i + N_i) - |T|\right)}{\sum_{i \in K} D_i \cdot N_i} \leq a \qquad \forall k \in K \qquad (4)$$

$$-a \leq d_k - \frac{N_k \cdot \left(\sum_{i \in K} D_i \cdot (A_i + N_i) - |T|\right)}{\sum_{i \in K} D_i \cdot N_i} \qquad \forall k \in K \qquad (5)$$

$$x_{kt} = 1 \qquad \forall (k,t) \in E \qquad (6)$$

$$x_{kt} \in \{0,1\} \qquad \forall k \in K, t \in T \qquad (7)$$

$$d_k \in R \qquad \forall k \in K \qquad (8)$$

Constraint (1) ensures that the number of appointments reserved for patient type k plus a deviation is given by the number of already assigned appointments plus the expected demand of that type. In constraint (2) we assert that no appointment can last longer than the end of the day. Constraint (3) makes overlapping of appointments impossible. It is assumed fair to schedule a number of appointments of type k proportional to the number of expected requests of type k. Constraints (4) and (5) ensure that there can only be a certain deviation from this proportion through setting parameter a. This can be seen through rearranging the assertion $N_k - d_k \approx \frac{N_k}{\sum_{i \in K} N_i} \cdot \sum_{i \in K}(N_i - d_i)$. Here, $\frac{N_k}{\sum_{i \in K} N_i}$ is the share of expected demand of type k and $(N_k - d_k)$ is the number of reserved appointments for patients of type k (not including already assigned appointments). Constraint (6) fixes the already assigned appointments. Constraints (7) and (8) are the domain constraints. In addition, to ensure the consideration of the current request of type k, we augment the expected demand from now until the end of the booking horizon of type k, which is $N_{k,}$ by 1.

Sets	Definition
$\boldsymbol{K}$	Set of types
$\boldsymbol{T}$	Set of time intervals of the day
$\boldsymbol{E}$	Set of tuples of already assigned appointments $(k,t), k \in K, t \in T$ where t is the first time interval of the appointment
Parameters	
$\boldsymbol{P_{kt}}$	Probability that a patient of type k accepts an appointment starting in time interval t if only this appointment is offered (Probability is set to one for already assigned appointments)
$\boldsymbol{D_k}$	Service length for a patient of type k
$\boldsymbol{N_k}$	Expected demand of type k from now until the end of the booking horizon
$\boldsymbol{A_k}$	Number of already assigned appointments of type k
$\boldsymbol{T_{kt}}$	$= \{\max\big(0, (t - D_k + 1)\big), \dots, t\}$
$\boldsymbol{a}$	Fairness parameter
Variables	
$\boldsymbol{x_{kt}}$	Binary variable that equals 1 if time interval t is the starting time interval of an appointment reserved for a patient of type k
$\boldsymbol{d_k}$	Demand of type k that is not considered ($d_k \geq 0$) or that is over considered ($d_k \leq 0$)

Table 1: Sets, parameters and decision variables of the model

Numerical Experiments

In order to test our model we implemented a simulation. Patient requests are generated according to a Poisson process. For every appointment request the mixed integer model is solved and the set of reserved appointments for the requesting patient's type is offered to the patient. We apply the logit decision model presented in (McFadden 1973) to model the patients choice. That means given a set J of decision alternatives, the probability of choosing alternative $j \in J$ for a patient of type k is given by $P_{kj} = \frac{e^{V_{kj}}}{\sum_{i \in J} e^{V_{ki}}}$, where V_{kj} is the expected benefit of a type k patient for choice j. In our case j either is given by an appointment (denoted by its starting time interval t) or by the choice not to accept any offered appointment. At the end of the booking horizon we count the number of unused time intervals. Besides, we measure fairness. For every patient type k we define fairness as the deviation of the proportion of assigned appointments for type k to the number of overall assigned appointments from the proportion of incoming requests of type k to the number of all incoming requests. The overall fairness is the sum of the absolute values of those deviations. In this sense a low fairness value is good. To validate our model, we compare the simulation results of our model to the results of two online scheduling heuristics. Online heuristic 1 offers all fitting appointments (with respect to service length). Online heuristic 2 only offers the earliest appointment that fits (with respect to service length). Online heuristic 2 is similar to the appointment assignment procedure in many practices. In our numerical experiments we use 42 time intervals per day and 6 different patient types which result from combining two service time lengths (one and two time intervals) and three time preferences (morning, afternoon and all day). To be more precise, we suppose that $V_{kt} = 4.1$ for time intervals t that correspond to the time preference of patients type k, otherwise we assume $V_{kt} = 0$. The benefit of rejecting any offered appointment is set to V_{kr} =0 for $\max_{t \in J} V_{kt} = 4.1$ and to V_{kr} =4.1 for $\max_{t \in J} V_{kt} = 0$. These settings result in very high accepting probabilities of 0.98 for fitting time slots. Patient types 1, 2 and 3 have a service time length of 1 whereas patient types 4, 5 and 6 have a service time length of 2. We consider 6 scenarios with different overall demand levels and different demand proportions as can be seen in Table 2.

Scenario	Expected demand per type
1	[3, 3, 3, 2, 2, 2]
2	[6, 6, 3, 4, 4, 2]
3	[6, 6, 6, 4, 4, 4]
4	[9, 9, 9, 6, 6, 6]
5	[12, 12, 12, 8, 8, 8]
6	[12, 12, 6, 8, 8, 4]

Table 2: Scenarios

The first two scenarios correspond to an expected under-utilization whereas the last three scenarios correspond to an expected exceed of the daily time capacity. As a solver we use IBM ILOG CPLEX. As a programming environment for the simulation we use the IBM ILOG CPLEX Optimization Studio. Every considered scenario was simulated several times until the accuracy was acceptable. No scenario simulation took longer than 40 min. Further, it takes around 10 seconds to solve the mixed integer linear program once.

First, we investigated the influence of the fairness parameter a. To this end only scenario 3 was considered for different values of a. The results can be seen in Table 3.

Parameter a	Number of unused time intervals	Fairness
1	5.06±1.1	0.06±0.02
2	4.23±0.75	0.07±0.01
3	3.5±0.86	0.08±0.02

Table 3: Variation of the fairness parameter a

It can be seen that the number of unused time intervals is lower for bigger a but at the same time also the fairness value is going up as expected. For very small a the integer linear program becomes infeasible.

In the following, we compare the simulation results of our model with the simulation results of the online heuristics. Here, we set the fairness parameter to $a = 2$ (such that the fairness values of the mixed integer linear program (MILP) are of a similar size as the fairness values of online heuristic 1). In Table 4 you can see the average values of the number of unused time intervals (Un. time inter.) and of the measure fairness for the 6 scenarios.

Scenarios	MILP		Online 1		Online 2	
	Un. time inter.	Fairness	Un. Time inter.	Fairness	Un. time inter.	Fairness
1	22.00 ±1.6	0.00	22.33±1.5	0.02±0.01	28.53±0.9	0.74±0.06
2	16.67±1.00	0.00	16.67±1.00	0.00	22.97±0.8	0.84±0.04
3	4.23±0.75	0.07±0.01	5.4±0.66	0.01±0.01	15.57±0.96	0.55±0.04
4	0.06±0.06	0.18±0.02	0.10±0.06	0.25±0.01	3.30±0.8	3.50±0.02
5	0.03±0.03	0.22±0.01	0.07±0.05	0.24±0.02	0.30±0.22	0.39±0.02
6	0.03±0.03	0.22±0.01	0.30±0.17	0.22±0.01	1.93±0.43	0.30±0.02

Table 4: Results of the numerical experiments

First of all, we can see in Table 4 that online heuristic 2 yields significantly worse results than the MILP and online heuristic 1. It tries to avoid gaps in the schedule through offering only the first fitting appointment. But as it doesn't consider preferences, some patients reject the offer and in the end more time intervals are left unused. The results for the MILP and online heuristic 1 don't differ significantly for scenarios 1 and 2 (expected under-utilization). The reason for this is probably that for expected under-utilization the consideration of upcoming demand is not as important because there are enough free time slots. Regarding scenario 3 the MILP has less unused time intervals and a little higher fairness value compared to online heuristic 1. In the cases of scenarios 4, 5 and 6 (exceeded daily capacity) the MILP yields slightly better results than online heuristic 1. In these cases, considering future demand seems to be beneficial.

Conclusion and Outlook

In this paper we presented a mixed integer linear programming model determining a set of appointments to offer to a patient with certain time preferences and a service length in order to schedule this patient while also taking future demand into account to maximize the overall utilization while assuring a certain fairness level.

Possible future work on the model includes more numerical experiments to determine scenarios where the MILP model yields an additional benefit compared to the online heuristics. Further, more sensitivity analyses considering the model parameters should be conducted.

In addition, the presented model can be extended in several ways. The assumption that the type of a patient is known could be relaxed. In addition, data from outpatient clinics about time preferences and service time lengths should be collected and clustered in order to find realistic patient types. Further, it could be beneficial to generate more than one schedule for every patient request in order to find even more appointments to offer. One could consider several days at the same time testing the limits of the mixed integer programming model. For large problems constraint programming could be applied as it has been done in (Hahn-Goldberg et al. 2014).

References

Cheraghi-Sohi, Sudeh, Arne Risa Hole, Nicola Mead, Ruth Mcdonald, Diane Whalley, and Peter Bower. 2008. "What Patients Want From Primary Care Consultations: A Discrete Choice Experi- Ment to Identify Patients' Priorities." *Annals Of Family Medicine* 6(2): 107–15.

Europäische Kommission. 2014. *Die Europäische Union Erklärt: Gesundheitswesen.*

Feldman, Jacob, Nan Liu, Huseyin Topaloglu, and Serhan Ziya. 2014. "Appointment Scheduling Under Patient Preference and No-Show Behavior." *Operations Research* 62(4): 794–811. http://pubsonline.informs.org/doi/abs/10.1287/opre.2014.1286.

Gerard, Karen, Chris Salisbury, Deborah Street, Catherine Pope, and Helen Baxter. 2008. "Is Fast Access to General Practice All That Should Matter? A Discrete Choice Experiment of Patients' Preferences." *Journal of health services research & policy* 13(2): 3–10.

Hahn-Goldberg, Shoshana. 2014. "Dynamic Optimization Addressing Chemotherapy Outpatient Scheduling." University of Toronto.

Hahn-Goldberg, Shoshana, Michael W Carter, J Christopher Beck, Maureen Trudeau, Philomena Sousa, and Kathy Beattie. 2014. "Dynamic Optimization of Chemotherapy Outpatient Scheduling with Uncertainty." *Health care management science* 17(4): 379–92. http://www.ncbi.nlm.nih.gov/pubmed/24477637 (February 4, 2014).

Klassen, Kenneth J., and Thomas R. Rohleder. 1996. "Scheduling Outpatient Appointments in a Dynamic Environment." *Journal of Operations Management* 14(2): 83–101.

McFadden, Daniel. 1973. “Conditional Logit Analysis of Qualitative Choice Behavior.” *Frontiers in Econometrics*: 105–42.

Rohleder, TR. 2000. “Using Client-Variance Information to Improve Dynamic Appointment Scheduling Performance.” *Omega* 28(3): 293–302. http://www.sciencedirect.com/science/article/pii/S0305048399000407 (January 12, 2015).

Wang, Jin, and Richard Fung. 2015. “Dynamic Appointment Scheduling with Patient Preferences and Choices.” *Industrial management & data systems* 115(4): 700–717.

Wang, Wen-Ya, and Diwakar Gupta. 2011. “Adaptive Appointment Systems with Patient Preferences.” *Manufacturing & Service Operations Management* 13(3): 373–89. http://pubsonline.informs.org/doi/abs/10.1287/msom.1110.0332 (November 1, 2013).