Lecture Notes in Management Science (2016) Vol. 8, 72–77 ISSN 2008-0050 (Print), ISSN 1927-0097 (Online)

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

Proc. ICAOR 2016
Rotterdam, The Netherlands

Keywords:
Advance admission scheduling
Appointment scheduling
Mixed integer linear programming
Patient preferences
Simulation

Abstract

Advance admission scheduling in the field of health care is an important and complex problem. Often, exact models of realistic size cannot be solved due to the curse of dimensionality and heuristics have to be used. 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 perform a simulation in order to test the mixed integer linear program and to compare it to simpler online heuristics. We develop different scenarios and show that using the mixed integer linear program to schedule patients is beneficial.

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 (European Commission, 2015). One of the important parameters to increase the efficiency of clinical processes in outpatient clinics is the design of their appointment scheduling system. Apart from a few exceptions, the clear majority of publications on appointment scheduling models does not simultaneously take patient time preferences and different service time lengths into consideration.

The studies of (Gerard *et al*., 2008) and (Cheraghi-Sohi *et al*., 2008) underline the relevance of individual patient time preferences when arranging appointments. According to surveys of Klassen and Rohleder a quarter of the patients has concrete time preferences (Klassen & Rohleder, 1996). Considering patient time preferences when arranging appointments even has positive effects for the outpatient clinics: Patients are highly satisfied, so that both the number of no-shows and the number of patients migrating to other clinics decrease (Feldman *et al*., 2014). Then again, the resulting appointment flexibility for patients can lead to a high variance of 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).

Service durations in outpatient clinics depend on different aspects such as patient specific characteristics (e.g. age, degree of disease progression, cultural background) (Gupta & Denton, 2008) and service specific characteristics (e.g. medical problem, new vs. return patients) (Cayirli *et al*., 2003).

Literature review

Advance admission scheduling, i.e., the problem of assigning appointments for future days dynamically, is difficult. Considering patients' choice makes the problem even more complex. Often, exact models of realistic size cannot be solved due to the curse of dimensionality (Liu & Ryzin, 2008). Hence, heuristics and simulation are used to tackle the problem.

Our procedure of assigning appointment requests to free time slots is based on the dynamic optimization model of (Hahn-Goldberg, 2014). Firstly, a proactive template is generated based on the expected appointment requests for the following day. Then, 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 current request the template is updated with consideration of the already scheduled appointments and the current request.

In (Rohleder & Klassen, 2000) the authors integrate patient time preferences into their simulations. They compare different appointment scheduling rules with regard to 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. Two performance criteria are looked at: 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 (Wang & 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 includes 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.

In (Feldman *et al*., 2014) a set of appointment days is offered to the patient who then 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. They maximize the number of patients showing up for their appointments with regard to the clinic capacity where the show-up probability depends on the patient's time preferences. Feldman *et al*. solve their dynamic model by means of a heuristic. In (Wang & Fung, 2015) the appointment length is the same for all patients. The clinic offers a set of time-physician-combinations to each patient with the aim of maximizing the expected return of the clinic. The dynamic program is solved via an approximate dynamic programming approach using an LP-formulation with an affine approximation of the value function.

In contrast to the existing appointment scheduling models we consider both patient time preferences and different service time lengths. Additionally, we determine appointment times not only appointment days. The main difference to the scheduling procedure of Hahn-Goldberg is that the presented model generates a new template schedule for every request.

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 an (inhomogeneous) Poisson process. We assume 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 future demand into account. 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. 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 requests by reserving appointments for every patient type. The objective function maximizes the expected utilization of the schedule at the end of the boking horizon assuming that every appointment is offered to one patient of the corresponding type. In the objective function the probability P_{kt} of accepting an offered appointment is set to 1 for already assigned appointments. The resulting reserved appointments for the requesting patient's type not yet assigned are potential appointments to offer to him or her. In Table 1 the sets, parameters and variables are defined. Our model (MILP) then results in:

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

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

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

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

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

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

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

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

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

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 ensure that no appointment can last longer than the whole 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.

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

Sets	
$\boldsymbol{K}$	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
$\boldsymbol{T_{kt}}$	$= \{\max(0, (t - D_k + 1)), \ldots, t\}$
$\boldsymbol{a}$	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 ($d_k \leq 0$)

Numerical experiments

In order to test our model we performed a simulation. Patient requests are generated according to a Poisson process (this could easily be extended to inhomogeneous Poisson processes). 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 patient's 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 $P_{kj} = \frac{\exp(V_{kj})}{\sum_{i \in J} \exp(V_{ki})}$, 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 unfairness. For every patient type k we define unfairness 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 unfairness is the sum of the absolute values of those deviations. To validate our model, we compare the simulation results of our model to the results of two online scheduling heuristics we developed and to the offline MILP. 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). We believe online heuristic 2 is similar to the appointment assignment procedure in many practices. Before the beginning of the booking horizon, the offline model knows how many patients of each patient type will show up. In this case the MILP only needs to be solved once in the beginning and then the appointments are assigned according to the known procedure. Interviews of Klassen and Rohleder with receptionists of outpatient clinics showed that clinics scheduled a three-hour morning session and a four-hour afternoon session and that the appointments were scheduled in ten-minutes-intervals (Klassen & Rohleder, 1996). Therefore we use 42 time intervals per day in our numerical experiments. In (Cayirli *et al*., 2003) new and return patients are distinguished and they detect that the service time of new patients is twice as long as the service time of return patients. We experimented with different time preferences and service lengths. Here, we present scenarios that especially show the benefit of our model. The first scenarios are characterized by service lengths that are not multiples of each other. We combine two different service lengths - two and three time intervals - with three different time preferences: morning, afternoon and all day and obtain 6 patient types. 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 2 whereas patient types 4, 5 and 6 have a service time length of 3. We consider 4 constructed scenarios with different overall demand levels (defining the Poisson process parameter) and different demand proportions as can be seen in Table 2. We consider two more scenarios which are characterized by a patient group that accepts only morning appointments and a flexible patient group which prefers the morning appointments to the afternoon appointment. In contrast to the first 4 scenarios, the flexible patient group with morning preferences is characterized by $V_{kt} = 4.1$ for morning time intervals and $V_{kt} = 4$ for afternoon intervals.

Table 2. Scenarios

Scenario	*Service lengths*	*Time preferences*	*Expected number of patients over the booking horizon per type*	*Expected overall demand in time slots*
1	2, 3	morning, afternoon, all day	[3, 3, 3, 2, 2, 2]	36
2	2, 3	morning, afternoon, all day	[6, 6, 3, 4, 4, 2]	60
3	2, 3	morning, afternoon, all day	[6, 6, 6, 4, 4, 4]	72
4	2, 3	morning, afternoon, all day	[9, 9, 9, 6, 6, 6]	108
5	1, 2	morning, all day with morning preference	[9,9,6,6]	42
6	1, 2	morning, all day with morning preference	[18,18,12,12]	84

The first scenario corresponds to an expected under-utilization, the fifth scenario corresponds to an expected full-utilization whereas the other scenarios correspond to an expected exceed of the daily time capacity. As a solver we use IBM ILOG CPLEX 12.6.2. As a programming environment for the simulation we use the IBM ILOG CPLEX Optimization Studio. Every considered scenario was simulated several times until the 95% confidence interval of the number of unused time intervals and unfairness reached a predefined length (e.g. 0.05 for unfairness) which was defined as small as possible considering an upper bound of 40 to 60 minutes of simulation time. Further, it takes around 3 seconds to solve the mixed integer linear program once.

In the following, we compare the simulation results of our model with the simulation results of the online heuristics and the offline model. Here, we set the fairness parameter to $\alpha = 2$ (such that the unfairness values of the mixed integer linear program (MILP) are of a similar size as the unfairness values of online heuristic 1). In Table 3 you can see the average values for the number of unused time intervals (Un. time inter.) and unfairness for the 6 scenarios.

Table 3. Results of the numerical experiments for unused time intervals and unfairness

Scenarios	*MILP*		*Online 1*		*Online 2*		*Offline*	
	Un. time inter.	*Unfairn.*	*Un. time inter.*	*Unfairn.*	*Un. time inter.*	*Unfairn.*	*Un. time inter.*	*Unfairn.*
1	9.24±0.77	0.07±0.01	11.38±0.51	0.11±0.01	18.58±1.15	0.62±0.05	9.07±0.74	0.15±0.01
2	1.77±0.27	0.22±0.01	5.89±0.28	0.27±0.01	10.20±0.75	0.62±0.04	2.81±0.32	0.24±0.01
3	0.78±0.08	0.24±0.01	4.51±0.13	0.31±0.01	2.00±0.50	0.40±0.02	1.24±0.25	0.28±0.01
4	0.77±0.72	0.29±0.01	4.94±0.16	0.31±0.01	0.40±0.07	0.41±0.02	0±0.00	0.27±0.01
5	4.84±0.63	0.08±0.01	6.91±0.58	0.15±0.01	11.58±0.56	0.28±0.01	4.69±0.69	0.10±0.01
6	0.02±0.02	0.25±0.01	0.09±0.06	0.29±0.01	0.8±0.28	0.54±0.02	0.36±0.09	0.19±0.01

First of all, we can see in Table 3 that the MILP and the offline MILP yield similar results. Therefore, the exact demand knowledge does not lead to significantly better results compared to only having expected demand values when using this MILP. Further, we see that online heuristic 2 in general yields significantly worse results than the MILP. It tries to avoid gaps in the schedule through offering only the first fitting appointment. But as it does not consider preferences, some patients reject the offer and in the end more time intervals are left unused. The MILP also yields significantly better results than online heuristic 1. In our opinion this is due to the fact that the MILP considers the expected future demand. In particular, for the scenarios 1 to 4 the MILP considers the service time lengths of the future demand whereas online heuristic 1 leaves gaps of one time interval in the schedule which cannot be assigned to any patient type. For the scenarios 5 and 6, the MILP especially considers the time preferences of the future demand while online heuristic 1 schedules a flexible patient with preferences for the morning to a morning time slot. In this way possible appointments for the patient group with only morning preferences are blocked.

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 sensitivity analyses considering the model parameters. In addition, the presented model can be extended in several ways. The assumption that the type of a patient is known can 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

Cayirli, T., Veral, E., & Rosen, H. (2006). Designing appointment scheduling systems for ambulatory care services. Health Care Management Science, 9(1), 47–58.

Cheraghi-Sohi, S., Hole, A. R., Mead, N., McDonald, R., Whalley, D., & Bower, P. (2008). What Patients Want From Primary Care Consultations: A Discrete Choice Experiment to Identify Patients' Priorities. Annals Of Family Medicine, 6(2), 107–115.

European Commission. (2015). The European Union explained: Public health.

Feldman, J., Liu, N., Topaloglu, H., & Ziya, S. (2014). Appointment Scheduling Under Patient Preference and No-Show Behavior. Operations Research, 62(4), 794–811.

Gerard, K., Salisbury, C., Street, D., Pope, C., & Baxter, H. (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.

Gupta, D., & Denton, B. (2008). Appointment scheduling in health care: Challenges and opportunities. IIE Transactions, 40(9), 800–819.

Hahn-Goldberg, S. (2014). Dynamic Optimization Addressing Chemotherapy Outpatient Scheduling. University of Toronto.

Hahn-Goldberg, S., Carter, M. W., Beck, J. C., Trudeau, M., Sousa, P., & Beattie, K. (2014). Dynamic optimization of chemotherapy outpatient scheduling with uncertainty. Health Care Management Science, 17(4), 379–392.

Klassen, K. J., & Rohleder, T. R. (1996). Scheduling outpatient appointments in a dynamic environment. Journal of Operations Management, 14(2), 83–101.

Liu, Q., & Ryzin, G. Van. (2008). On the Choice-Based Linear Programming Model for Network Revenue Management. Manufacturing & Service Operations Managementervice, 10(2), 288–310.

McFadden, D. (1973). Conditional logit analysis of qualitative choice behavior. Frontiers in Econometrics.

Rohleder, T. R., & Klassen, K. J. (2000). Using client-variance information to improve dynamic appointment scheduling performance. Omega, 28(3), 293–302.

Wang, J., & Fung, R. (2015). Dynamic appointment scheduling with patient preferences and choices. Industrial Management & Data Systems, 115(4), 700–717.

Wang, W.-Y., & Gupta, D. (2011). Adaptive Appointment Systems with Patient Preferences. Manufacturing & Service Operations Management, 13(3), 373–389.