Working Paper

Witek ten Hove[1,2]☯, Ger Koole[2]☯, Joost Berkhout[2]☯

1 Lectoraat Logistiek en Allianties, HAN University of Applied Sciences, Arnhem, The Netherlands,
2 Department of Mathematics, VU, Amsterdam, The Netherlands,

☯These authors contributed equally to this work.

¤Current Address: Dept/Program/Center, Institution Name, City, State, Country

†Deceased

¶Membership list can be found in the Acknowledgments sections

Abstract

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Curabitur eget porta erat. Morbi consectetur est vel gravida pretium. Suspendisse ut dui eu ante cursus gravida non sed sem. Nullam sapien tellus, commodo id velit id, eleifend volutpat quam. Phasellus mauris velit, dapibus finibus elementum vel, pulvinar non tellus. Nunc pellentesque pretium diam, quis maximus dolor faucibus id. Nunc convallis sodales ante, ut ullamcorper est egestas vitae. Nam sit amet enim ultrices, ultrices elit pulvinar, volutpat risus.

Author summary

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Curabitur eget porta erat. Morbi consectetur est vel gravida pretium. Suspendisse ut dui eu ante cursus gravida non sed sem. Nullam sapien tellus, commodo id velit id, eleifend volutpat quam. Phasellus mauris velit, dapibus finibus elementum vel, pulvinar non tellus. Nunc pellentesque pretium diam, quis maximus dolor faucibus id. Nunc convallis sodales ante, ut ullamcorper est egestas vitae. Nam sit amet enim ultrices, ultrices elit pulvinar, volutpat risus.

Introduction

(Verbatim Ger Koole) Many decision problems have a dynamic nature, the consequences of our decisions become available step by step over time, and can only be simulated or calculated as a Markov chain. Decisions have long-term consequences, and these consequences are also often of a stochastic nature. To “remember” these consequences the “state” of the system plays a crucial role. Decision problems can roughly be divided in two types of problems: those where the decisions are taken on the fly and are accounted for through a state change, and those where decisions are taken upfront. The first category falls into the framework of stochastic dynamic programming and is currently immensely popular in AI under the name reinforcement learning. The second is equally important but receives much less attention. Examples are the scheduling of people in service centers such as health clinics and call centers. Employees have to be scheduled well in advance, but the consequences in terms of for example waiting times can only be modeled through a stochastic process, for which simulation and Markov chain analysis are the two prime solution methods.

Other examples are the design of energy systems and appointment scheduling, but the list of possible applications is endless. Note that many service systems have both types of decision problems: for example long-term capacity and employee scheduling problems, and short-term task scheduling and re-adjustments to the schedule.

The focus of the project is on the second type of problem. Simulation, and to a lesser extend Markov chain analysis, are computationally costly solution methods, and they have to be executed for multiple decisions. Because the decision space is often multi-dimensional enumeration is not possible. Local search can only find local optima and that is for a fixed computational budget not even guaranteed.

Smarter methods are needed, a very interesting candidate is fitting a machine learning model to a limited set of solutions and then try to find the a (local) optimum. This has the advantage that, once trained, it is much faster to use a ML model than simulation or Markov chain analysis. This is known in the literature as surrogate models and response surface methodology (to be checked), but the current developments in machine learning open possibilities for new versions of algorithms and new applications. A couple of things to look into:

- applications into for example appointment scheduling and shift scheduling

- does an iterative approach help, where the test set consists of points close to the optimum of the previous iteration? perhaps in combination with linear regression with squares and interactions which gives a global optimum?

- can knowledge about the problem (such as monotonicity in a parameter) be included in a smart way in the prediction model?

Applications for and purpose of outbound appointment systems

[1] distinguish between three types of decisions for designing Outpatient Appointment Systems (OASs):

- Strategic: long-term decisions that determine the determine the main structure of AOS.
- Tactical: medium-term decisions on how patients groups or subgroups are processed.
- Operational: short-term decisions related to efficient scheduling individual patients

	Decision level	Code	Name (in alphabetical order)
Design decisions	Strategic	S1	Access policy
		S2	Number of servers/resources

	Decision level	Code	Name (in alphabetical order)
		S3	Policy on acceptance of walk-ins
		S4	Type of scheduling
Planning decisions	Tactical	T1	Allocation of capacity to patient groups
		T2	Appointment interval (slot)
		T3	Appointment scheduling window
		T4	Block size
		T5	Number of appointments in consultation session
		T6	Panel size
		T7	Priority of patient groups
	Operational	O1	Allocation of patients to servers/resources
		O2	Appointment day
		O3	Appointment time
		O4	Patient acceptance/rejection
		O5	Patient selection from waiting list
		O6	Patient sequence

They labeled articles on OASs by decision type, solution method and modeling approach.

Survey dimension	Code				
Solution method	AM	Analytical method			
	SL	Numerical method	Accurate method (Exact or bounded-error method)	Solving by a general-purpose optimization software package	LINGO
	SC				CPLEX
	SG				GAMS
	SO				Other items
	BB			Branch and bound	
	BC			Branch and cut	
	CG			Column generation	
	BD			Benders decomposition	
	LD			L-shaped decomposition	
	ND			Nested decomposition	
	O			Other items	
	H		Inaccurate method	Heuristic	
	MH-TS			Metaheuristic	Tabu search
	MH-GA				Genetic algorithm
	MH-SA				Simulated annealing
	MH-O				Other items
	S-SBO			Approximate stochastic optimization	Simulation-based optimization
	S-SAA				Sample average approximation
	S-O				Other items
Modeling approach	LP	Linear programming			
	ILP	Integer linear programming			
	INLP	Integer nonlinear programming			
	MILP	Mixed-integer linear programming			
	SOCP	Convex conic programming	Second order cone programming		
	C-SDP		Semi-definite programming		
	PSP	Stochastic programming	Probabilistic (or chance-constraint) programming		
	1-SSP		Single-stage stochastic programming		
	2-SSP		Two-stage stochastic programming		
	M-SSP		Multi-stage stochastic programming		
	MDP		Stochastic dynamic programming	Markov decision process	
	SDP-O			Other items	
	SP-O		Other items, such as distributionally robust optimization (DRO)		
	DP	Dynamic programming			
	CP	Constraint programming			
	MCDM	Multi-criteria decision making			
	MPDM	Multi-person decision making (game theory)			
	O	Other items, such as queuing theory (QT), graph theory (GT), and network theory (NT)			

In this article we will be using a single-stage stochastic programming (1-SSP) modeling approach.

Table 2. Selection of articles taken from [1] with single-stage stochastic programming (1-SSP) as the modeling approach.

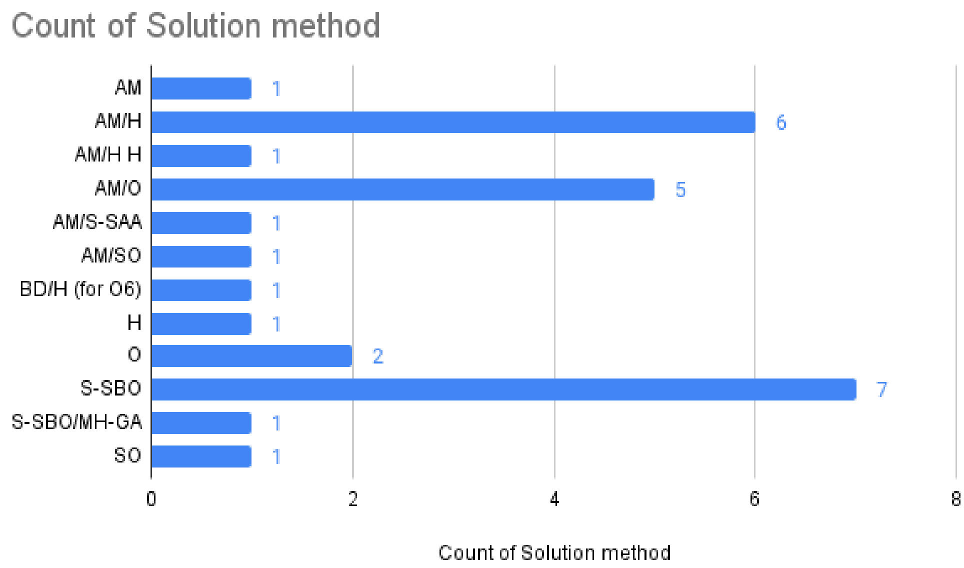
Reference	Tactical and operational decisions	Access policy	Type of scheduling: online (On), offline (Off)	Number of servers/resources: single (S), multiple (M)	Policy on acceptance of walk-ins: allowed (Yes), not allowed (No)	Objective: minimize (Min.), maximize (Max.)	Modeling approach	Solution method
(Begen & Queyranne, 2011)	O3 (OBA)	Traditional	Off	S	Yes (urgent)	Min. costs of waiting time, idle time, and overtime	1-SSP	AM
(Chakraborty et al., 2010)	O3/O4 (OBA) (integrated)	Open-access	On	S	No	Max. profit (revenue of patients seen – costs of patients overflowing between each two successive slots)	1-SSP	AM/H
(Chakraborty et al., 2013)	O3/O4 (OBA) (integrated)	Hybrid	On	S	No	Max. profit (revenue of patients seen – costs of waiting time and overtime)	1-SSP	AM/H
(LaGanga & Lawrence, 2012)	T4/T5 (integrated)	Traditional	-	S	No	Max. profit (revenue of patients seen – costs of waiting time and overtime)	1-SSP	AM/H
(Muthuraman & Lawley, 2008)	O3 (OBA)	Open-access	On	S	No	Max. profit (revenue of patients seen – costs of waiting time and overtime)	1-SSP	AM/H

Reference	Tactical and operational decisions	Access policy	Type of scheduling: online (On), offline (Off)	Number of servers/resources: single (S), multiple (M)	Policy on acceptance of walk-ins: allowed (Yes), not allowed (No)	Objective: minimize (Min.), maximize (Max.)	Modeling approach	Solution method
(Samorani & Ganguly, 2016)	O6 (for unpunctual patient) (OBA)	-	-	S	No	Min. costs of waiting time and idle time	1-SSP	AM/H
(Zacharias & Pinedo, 2014)	T4/O3/O6 (heterogeneous patients) (RBA) (integrated) T4/O3 (homogeneous patients) (OBA) (integrated)	Traditional	Off On	S	No	Min. costs of waiting time, idle time, and overtime	1-SSP	AM/H
(Zeng et al., 2010)	T1/T4/O3 (OBA) (integrated) O3/O4 (OBA) (integrated)	Traditional Open-access	Off On	S	No	Max. profit (revenue of patients seen – costs of waiting time and overtime)	1-SSP -	AM/H H
(Kaandorp & Koole, 2007)	T4/O3 (OBA) (integrated)	Traditional	-	S	No	Min. costs of waiting time, idle time, and overtime	1-SSP	AM/O
(Koeleman & Koole, 2012)	T4/O3 (OBA) (integrated)	-	-	S	Yes (Urgent)	Min. costs of waiting time, idle time, and overtime	1-SSP	AM/O

Reference	Tactical and operational decisions	Access policy	Type of scheduling: online (On), offline (Off)	Number of servers/resources: single (S), multiple (M)	Policy on acceptance of walk-ins: allowed (Yes), not allowed (No)	Objective: minimize (Min.), maximize (Max.)	Modeling approach	Solution method
(Kong et al., 2016)	O6	-	Off	S	No	Min. costs of waiting time and overtime	1-SSP	AM/O
(Qu et al., 2007)	T1	Hybrid	-	S	No	Max. # of patients seen	1-SSP	AM/O
(Yan et al., 2015)	O3/O4 (OBA) (integrated)	Traditional	On	S	No	Max. profit (revenue of patients seen – costs of waiting time, overtime, and idle time)	1-SSP	AM/O
(Begen et al., 2012)	O3 (OBA)	Traditional	Off	S	No	Min. costs of waiting time, idle time, and overtime	1-SSP	AM/S-SAA
(Luo et al., 2012)	T2/T5/O3 (OBA) (integrated)	Traditional	-	S	Yes (urgent)	Max. profit (revenue of patients seen – costs of waiting time and overtime)	1-SSP	AM/SO
(Chen & Robinson, 2014)	T2/O3/O6 (T2/O3: OBA, O6: RBA) (T2/O3: integrated, (T2/O3)/O6: sequential)	Hybrid	On	S	No	Min. costs of waiting time, idle time, and overtime	1-SSP	BD/H (for O6)

Reference	Tactical and operational decisions	Access policy	Type of scheduling: online (On), offline (Off)	Number of servers/resources: single (S), multiple (M)	Policy on acceptance of walk-ins: allowed (Yes), not allowed (No)	Objective: minimize (Min.), maximize (Max.)	Modeling approach	Solution method
(Vink et al., 2015)	T2/O3 (OBA) (integrated)	Traditional	Off	S	No	Min. costs of waiting time, idle time, and overtime	1-SSP	H
(Hassin & Mendel, 2008)	T2/O3 (OBA) (integrated)	Traditional	Off	S	No	Min. costs of waiting time and server availability	1-SSP	O
(Kim & Giachetti, 2006)	T5	Traditional	-	S	Yes (regular)	Max. profit (revenue of patients seen – costs of overtime and patient rejection)	1-SSP	O
(Anderson, Zheng, Yoon, & Khasawneh, 2015)	T2	Traditional	-	S	No	Min. costs of waiting time, idle time, and overtime	1-SSP	S-SBO
(Cayirli & Gunes, 2013)	T1/T4 (integrated)	Hybrid	On	S	Yes (regular)	Min. costs of waiting time, idle time, and overtime	1-SSP	S-SBO
(Huang & Zuniga, 2012)	T4/O3 (based on no-show threshold for each slot) (OBA)(integrated)	-	On	S	No	Min. costs of waiting time, idle time, and overtime	1-SSP	S-SBO

Reference	Tactical and operational decisions	Access policy	Type of scheduling: online (On), offline (Off)	Number of servers/resources: single (S), multiple (M)	Policy on acceptance of walk-ins: allowed (Yes), not allowed (No)	Objective: minimize (Min.), maximize (Max.)	Modeling approach	Solution method
(Huang, Hancock, & Herrin, 2012)	T2	-	-	S	No	Min. waiting time and idle time	1-SSP	S-SBO
(Klassen & Yoogalingam, 2009)	T2/O3 (OBA) (integrated)	Traditional	-	S	No	Min. costs of waiting time, idle time, and overtime	1-SSP	S-SBO
(Klassen & Yoogalingam, 2013)	T2/O3 (OBA) (integrated)	Traditional	-	S	No	Min. costs of waiting time and idle time	1-SSP	S-SBO
(Klassen & Yoogalingam, 2014)	T2/O3 (OBA) (integrated)	Traditional	On	S	No	Min. costs of waiting time and idle time	1-SSP	S-SBO
(Peng, Qu, & Shi, 2014)	T1/T4/O3 (OBA) (integrated)	Hybrid	On	S	Yes (regular)	Min. costs of waiting time, idle time, and overtime	1-SSP	S-SBO/MH-GA



Most solution methods for 1-SSP modeling approaches are analytical or simulation based (15 and 8 articles, resp.)

[2] mention four categories of OASs purposes:

1. Reducing costs
2. Increasing patient satisfaction
3. Lowering waiting time
4. Improving fairness

This will be reflected in the cost function where we will be attempting to minimize patient waiting times and physician over-time. Fairness is a rather subjective matter. The cost function will contain weights that can be adjusted to fit particular cases and/or preferences.

Problem description and solution methods

A schedule is a vector x consisting of T elements in which each element represents the number of patients scheduled at the interval starting at t:

$$x = [x_0, x_1, ...x_{T-1}]$$

with

$$\sum_{t=0}^{T-1} x_t = N,\ x_t \in \mathbb{N}_0$$

Patients are scheduled at fixed intervals with length d.

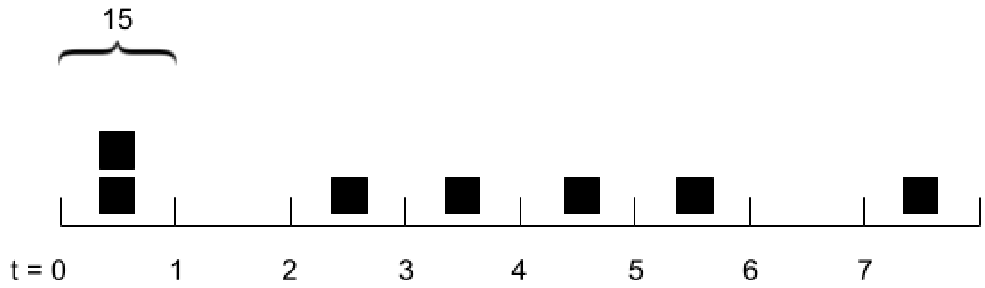


Figure 1. A schedule $x = [2, 0, 1, 1, 1, 1, 0, 1],\ T = 8,\ N = 7,\ d = 15$

Each patient has two endogenous features: type and service time, which are both independent and identically distributed variables. In our model we assume there are two patient types: standard and emergency. There is a probability q that a patient has an emergency. Service times have known distributions with mean β_s and β_e for standard and emergency patients respectively.

All standard patients are assumed to be punctual. The arrival rate of emergency patients has a Poisson distribution with rate λ per interval. Emergency patients get priority over standard patients that are waiting. If several emergency patients are waiting they are served in order of arrival.

The cost function consists of three elements:

1. The waiting time for patients: $W(x)$
2. The lateness or over-time of physicians: $L(x)$
3. The waiting or idle time for physicians: $I(x)$

and becomes:

$$C(x) = \alpha W(x) + \beta L(x) + \gamma I(x)$$

The weights $\alpha, \beta, \gamma \geq 0$ can be set to reflect the relative importance of each cost element.

The goal is to find a schedule x that minimizes the cost function $C(x)$:

$$min\{C(x) | \sum_{t=0}^{T-1} x_t = N, x_t \in \mathbb{N}_0\}$$

«Description solution method here»

Results

Discussion

Conclusion

Supporting information

Acknowledgments

References

1. Ahmadi-Javid A, Jalali Z, Klassen KJ. Outpatient appointment systems in healthcare: A review of optimization studies. European Journal of Operational Research. 2017;258(1):3–34. doi:10.1016/j.ejor.2016.06.064.

2. Ala A, Chen F. Appointment Scheduling Problem in Complexity Systems of the Healthcare Services: A Comprehensive Review. Journal of Healthcare Engineering. 2022;2022:1–16. doi:10.1155/2022/5819813.