UNIVERSITEIT GENT

FACULTY OF **ENGINEERING**

Optimization of Appointment Scheduling

Matthias Deceuninck Stijn De Vuyst Dieter Fiems

Industrial Systems Engineering and Product Design
Ghent University

Operations Research in Queuing Theory
7th November 2016

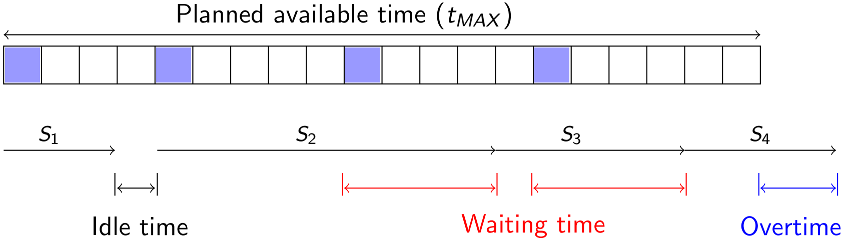
Introduction

- Interface between service providers and customers
- Arises in many service industries:
  - Healthcare: GP, dentist, outpatient clinics
  - Legal services
  - Accounting services
  - Loading/unloading ships

Static Appointment Scheduling Problem

1 Physician

K patients



Planned available time (t_{MAX})

S_1 S_2 S_3 S_4

Idle time

Waiting time

Overtime

Scheduling challenges

- Objective: minimize the weighted sum of
  - Patients' waiting times
  - Doctor's idle time
  - Doctor's overtime
- Complicating factors
  - uncertain service durations
  - unpunctual patients
  - uncertain patient demand
    - no-shows
    - emergencies
  - heterogeneous patients
  - doctor lateness / set-up activities

Contributions

Contributions

1. We propose a fast heuristic which can be applied under fairly general conditions.
2. We quantified the performance of our heuristic.
3. Patient heterogeneity affects optimal schedule and should be taken under consideration.

Solution approach

The optimization problem involves computational effort at three levels:

1. Evaluating the cost of a given schedule
2. Finding the optimal schedule for a fixed sequence of patients
3. Determining the optimal arrival sequence for customers having distinct attributes

Outline

1. Introduction

2. Research approach
   - Evaluation algorithm
   - Lower-level problem: optimize appointment times
   - Upper-level problem: sequencing

3. Numerical results
   - Homogeneous patients
   - Heterogeneous patients

4. Conclusions and Future Work

Outline

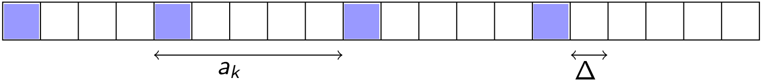
1. Introduction

2. Research approach
   - Evaluation algorithm
   - Lower-level problem: optimize appointment times
   - Upper-level problem: sequencing

3. Numerical results
   - Homogeneous patients
   - Heterogeneous patients

4. Conclusions and Future Work

Fast and versatile evaluation algorithm



Approach:

- modified Lindley's recursion
  - allows for explicit expressions for each patient

$$W_{k+1} = max(0, W_k + S_k - a_k)$$

- discrete time setting
  - discretised into fixed intervals of length Δ
  - trade-off accuracy and computational effort

Model description

Environmental factors

- Service time distribution $s_k(n)$
  - type of service, medical record, age, doctor
- No-show probability p_k
  - lead time, age, gender, type of service (5%-40%)
- Waiting cost c_{W_k}
  - urgency, opportunity cost
- Start time session $\theta(n)$
  - lateness doctor, set-up activities
- Idle time cost c_I / overtime cost c_0
  - availability service provider, equipment, assistants

Evaluation algorithm

Calculation of the performance measures (moments q up to any order):

$$\mathrm{E}\left[W_{k+1}^q\right] = -\ell_k^{(q)} + \sum_{r=0}^{q}\sum_{m=0}^{q-r}\binom{q}{r}\binom{q-r}{m}\mathrm{E}[S_k^m](-a_k)^{q-r-m}W_k^{(r)}\,,$$

$$\mathrm{E}[I_{k+1}^q] = (-1)^q\left(\mathrm{E}[(W_k - a_k + S_k)^q] - \mathrm{E}[W_{k+1}^q]\right) = (-1)^q\ell_k^{(q)}\,.$$

$$\mathrm{E}[O^q] = W_{K+1}^{(q)}\,.$$

with,

$$\ell_k^{(q)} = \sum_{r=0}^{a_k-1}\sum_{m=0}^{r} s_k(r-m)\,(r-a_k)^q\, w_k(m)\,.$$

$$w_{k+1}(n) = \sum_{m=0}^{n+a_k} s_k(n+a_k-m)\, w_k(m)\,,$$

Outline

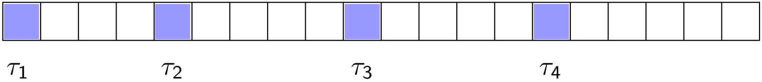
1. Introduction

2. Research approach
   - Evaluation algorithm
   - Lower-level problem: optimize appointment times
   - Upper-level problem: sequencing

3. Numerical results
   - Homogeneous patients
   - Heterogeneous patients

4. Conclusions and Future Work

Allocation of appointment times



GOAL: find $\boldsymbol{\tau} = \{\tau_1, \ldots, \tau_K\}$ which minimizes the objective function

Heuristic is needed: local search algorithm

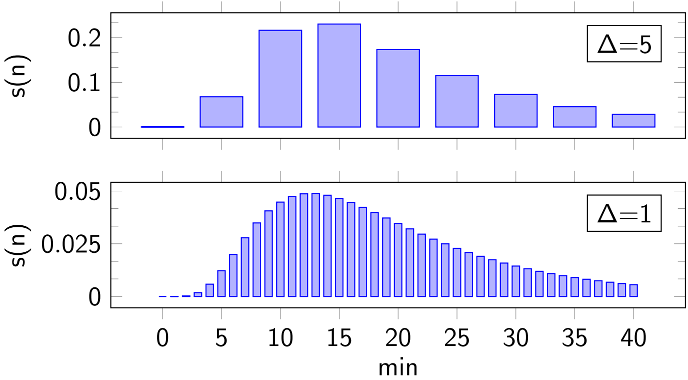
- Optimality can be guaranteed for homogeneous patients (Kaandorp and Koole, 2007; Koeleman and Koole, 2012)
- Unfortunately, not for heterogeneous patients (Tang et al., 2014)

Guidelines to reduce running times

1. Sensible choice of initial population
    - Some common heuristics in literature
        - 'Bailey's rule': schedule two patients at beginning
        - 'Fixed-interval': schedule patients with equal inter-arrival times
        - 'OFFSET': schedule first patients closer together

Guidelines to reduce running times

1. Sensible choice of initial population
2. Downscale the problem
    - Gradually decrease slot length Δ



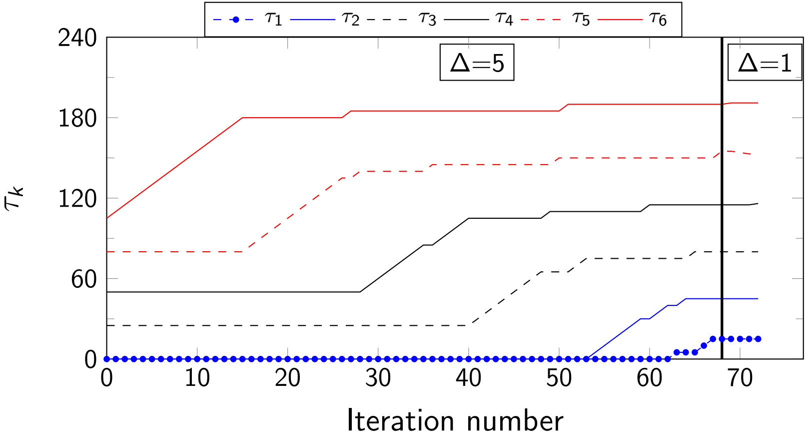
Guidelines to reduce running times

1. Sensible choice of initial population
2. Downscale the problem
3. Structural property:

Performance measures of kth patient only depend on τ_j for $j < k$

- Guidelines algorithm:
    - ⋆ Store all calculations in state variable **g**
    - ⋆ If τ_k changes, reuse information in **g** for $W_j, I_j, \ldots$ for $j < k$
    - ⋆ Start optimizing at the end of the schedule (reverse for-loop) (-30%)

Visualisation of search process



Outline

1. Introduction

2. Research approach
   - Evaluation algorithm
   - Lower-level problem: optimize appointment times
   - Upper-level problem: sequencing

3. Numerical results
   - Homogeneous patients
   - Heterogeneous patients

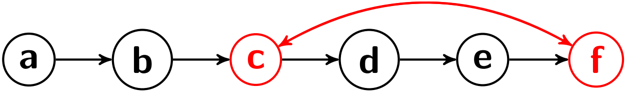
4. Conclusions and Future Work

Upper-level problem: sequencing

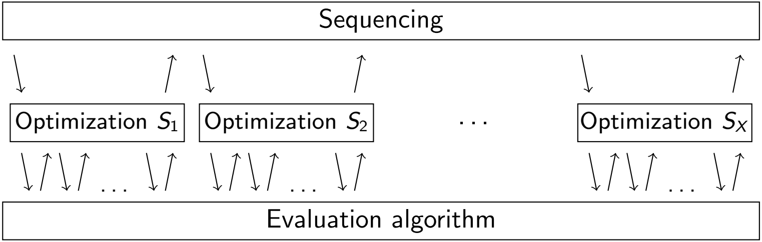
Patients may differ by:

- service time distribution
- no-show probability
- waiting cost

We perform a local search algorithm with pairwise interchange (swaps)



Global structure of optimization method



→ sequencing can be accelerated by multi-threading

Outline

1. Introduction

2. Research approach
   - Evaluation algorithm
   - Lower-level problem: optimize appointment times
   - Upper-level problem: sequencing

3. Numerical results
   - Homogeneous patients
   - Heterogeneous patients

4. Conclusions and Future Work

Outline

1. Introduction

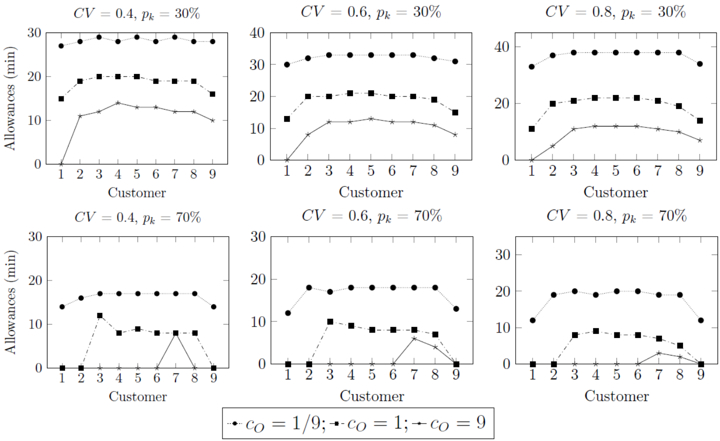
2. Research approach
   - Evaluation algorithm
   - Lower-level problem: optimize appointment times
   - Upper-level problem: sequencing

3. Numerical results
   - Homogeneous patients
   - Heterogeneous patients

4. Conclusions and Future Work

Structure of optimal solutions

In accordance with prior studies: solutions depict a dome-shaped pattern



Performance of heuristic for different neighbourhoods

K	Optimal Time (s)	$\Delta = 1$ Avg. gap (%)	Max. gap (%)	Time (s)	$\Delta = 5$ Avg. gap (%)	Max. gap (%)	Time (s)
6	6.9	0.001	0.079	0.9	0.123	3.229	0.12
8	34.4	0.003	0.240	1.6	0.212	4.140	0.15
10	168.9	0.006	0.346	2.6	0.442	23.775	0.19
12	818.0	0.009	0.408	4.3	0.506	6.804	0.24

Performance of heuristic for different neighbourhoods

K	Optimal Time (s)	$\Delta = 1$ Avg. gap (%)	Max. gap (%)	Time (s)	$\Delta = 5$ Avg. gap (%)	Max. gap (%)	Time (s)
6	6.9	0.001	0.079	0.9	0.123	3.229	0.12
8	34.4	0.003	0.240	1.6	0.212	4.140	0.15
10	168.9	0.006	0.346	2.6	0.442	23.775	0.19
12	818.0	0.009	0.408	4.3	0.506	6.804	0.24

→ We choose a smaller neighbourhood to reduce running times

Performance of heuristic for different neighbourhoods

		$\Delta = 1$			$\Delta = 5$		
K	Optimal Time (s)	Avg. gap (%)	Max. gap (%)	Time (s)	Avg. gap (%)	Max. gap (%)	Time (s)
6	6.9	0.001	0.079	0.9	0.123	3.229	0.12
8	34.4	0.003	0.240	1.6	0.212	4.140	0.15
10	168.9	0.006	0.346	2.6	0.442	23.775	0.19
12	818.0	0.009	0.408	4.3	0.506	6.804	0.24

Gaps depend on:

- overtime cost
- service time variability
- no-show probability

Outline

1. Introduction

2. Research approach
    - Evaluation algorithm
    - Lower-level problem: optimize appointment times
    - Upper-level problem: sequencing

3. Numerical results
    - Homogeneous patients
    - Heterogeneous patients

4. Conclusions and Future Work

Performance when patients are heterogeneous

Patient classification:

- 40% New patient: $S \sim \text{logN}(\mu = 20, \sigma = 16)$
- 60% Returning patient: $S \sim \text{logN}(10,4)$
- Average patient: $S \sim \text{logN}(14,8.4)$

No-show	Policy	Best-Case	Worst-Case	Δ Cost
0%	Heterogeneous	388.3	-	
	Homogeneous	549.0	594.5	41-53%
	Fixed-interval	484.3	673.5	25-73%
15%	Heterogeneous	358.0	-	
	Homogeneous	451.1	485.1	26-34%
	Fixed-interval	403.7	601.3	13-68%

Outline

1. Introduction

2. Research approach
   - Evaluation algorithm
   - Lower-level problem: optimize appointment times
   - Upper-level problem: sequencing

3. Numerical results
   - Homogeneous patients
   - Heterogeneous patients

4. Conclusions and Future Work

Conclusions and Future Work

- We introduced an efficient and effective heuristic for the static appointment scheduling problem
- The heuristic can be applied in a wide variety of environments
- Patient classification can result in significant cost reductions

- Future work
  - Data-mining techniques to effectively estimate service time distributions based on patient characteristics
  - Dynamic scheduling where not all patients are known in advance

Are there any questions?