Data-Driven Models for Efficient Cancer Treatment Delivery

LOUIS-MARTIN ROUSSEAU

CANADA RESEARCH CHAIR IN
HEATLHCARE ANALYTICS AND LOGISTICS
HANALOG

POLYTECHNIQUE
MONTRÉAL

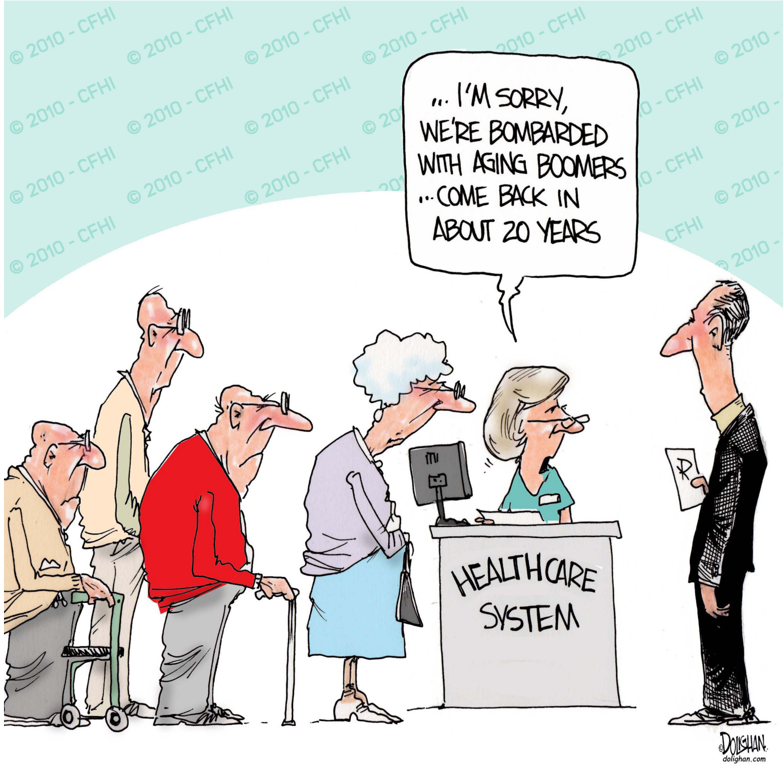
The impact of cancer

One person is diagnosed with cancer every
3 minutes in Canada, 20 seconds in USA.

One person dies from cancer every
7 minutes in Canada, 1 minute in USA.

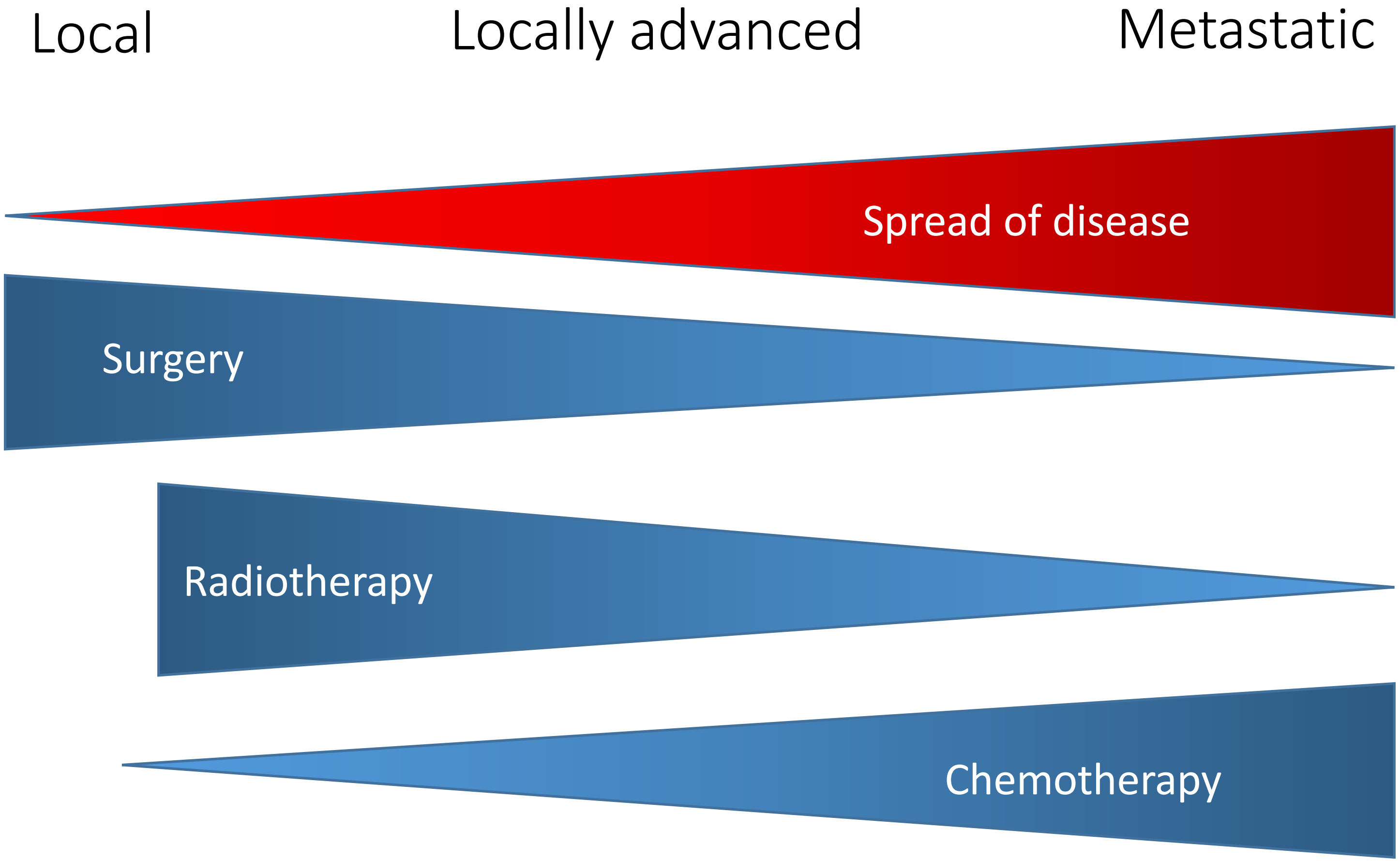
First cause of mortality in Canada (30%):
  45% of Canadian will develop cancer
  5 year survivability 66%

Ever increasing of new cancer cases:
  12% within 4 years
  Aging of population;
  Demographic growth.



How to treat all these patients while keeping excellent care ?

What are your treatment options ?



Local

Locally advanced

Metastatic

Spread of disease

Surgery

Radiotherapy

Chemotherapy

About 50% of cancer patients will receive radiotherapy

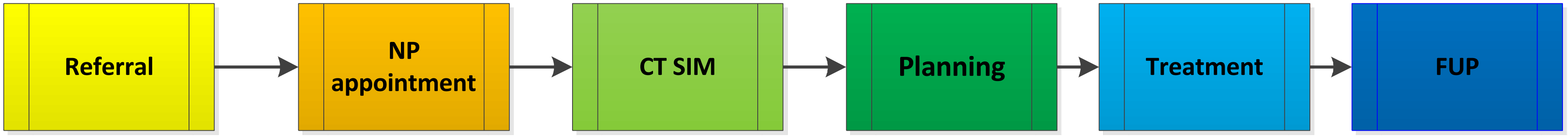
Radiation Therapy

Vickers 6 Prototype Newcastle-on-Tyne 1960

Teams

	Chemotherapy	Radiotherapy
Prescribes	Oncologist	Radiation Oncologist
Prepares	Pharmacist	Physicist
Delivers	Nurse	Therapist

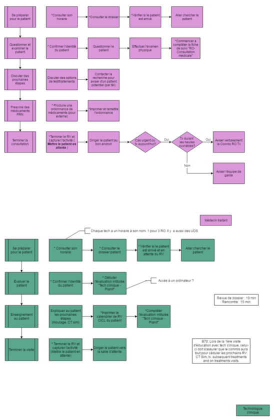
Care Trajectory



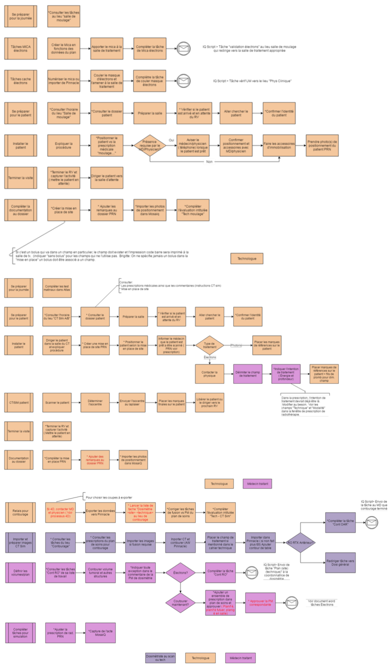
Referral → NP appointment → CT SIM → Planning → Treatment → FUP

Care Trajectory in details

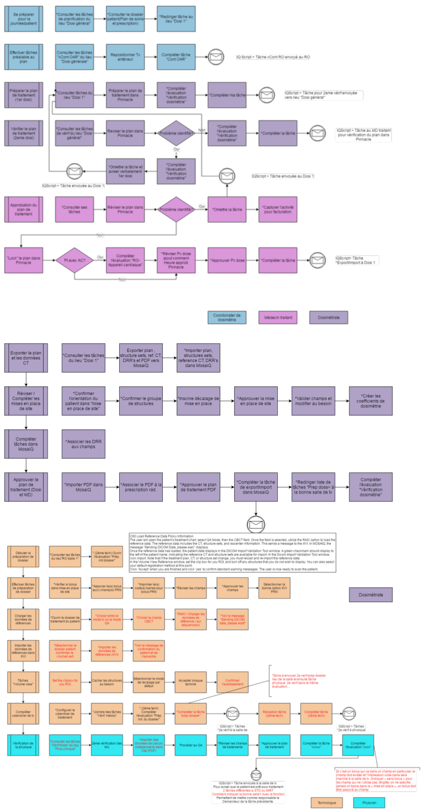
NP appointment



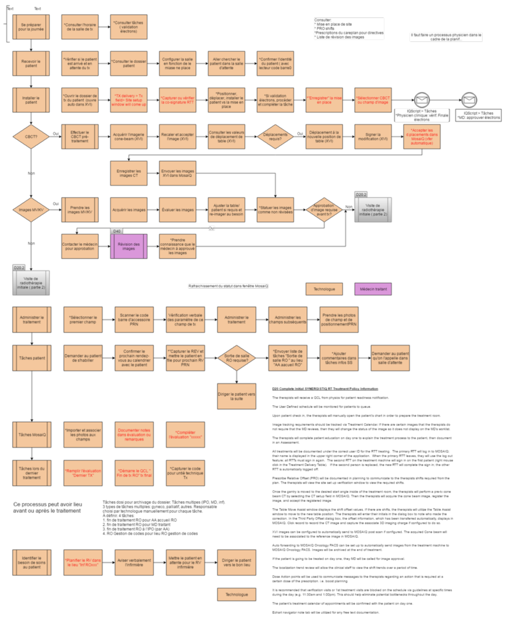
CT SIM



Planning



Treatment



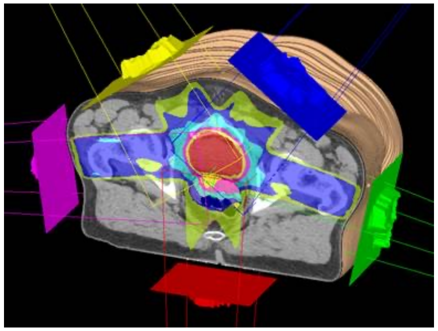
Important steps

Simulation:

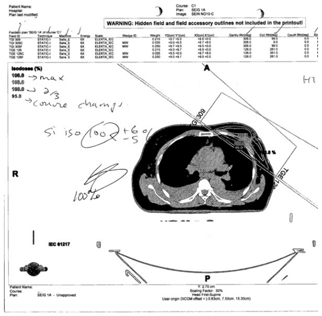
- Uses: CT, MRI, PET-CT
- Used for treatment planning purposes
- 3D model of the human body

Treatment Planning

- Calculates radiation deposition in the human body
- Multi-criteria optimization solver
- Server farm, GPU calculations, etc.



Plan approval



Patient Name:
Hospital:
Plan last modified:

Course: C1
Plan: SEIG 1A
Image: SEIN NO10-C

WARNING: Hidden field and field accessory outlines not included in the printout!

Fields in plan 'SEIG 1A' of course 'C1'

Field ID	Technique	Machine	Energy	Scale	Wedge ID	Weight	Y2[cm] Y1[cm]	X2[cm] X1[cm]	Gantry Rtn[deg]	Coll Rtn[deg]	Couch Rtn[deg]
TGI 309	STATIC-I	Salle_E	6X	ELEKTA_IEC		0.215	+9.7 +9.0	+9.5 +0.0	306.0	99.0	0.0
TGI 309C	STATIC-I	Salle_E	6X	ELEKTA_IEC	MW	0.035	+0.0 +9.5	+9.7 +9.0	306.0	9.0	0.0
TGI 309F	STATIC-I	Salle_E	6X	ELEKTA_IEC	MW	0.250	+9.7 +9.0	+9.5 +0.0	306.0	99.0	0.0
TGE 126	STATIC-I	Salle_E	6X	ELEKTA_IEC		0.215	+9.0 +9.7	+9.5 +0.0	126.0	261.0	0.0
TGE 126C	STATIC-I	Salle_E	6X	ELEKTA_IEC	MW	0.035	+9.5 +0.0	+9.7 +9.0	126.0	351.0	0.0
TGE 126F	STATIC-I	Salle_E	6X	ELEKTA_IEC	MW	0.250	+9.0 +9.7	+9.5 +0.0	126.0	261.0	0.0

Isodoses (%)
106.0 → max
103.0
100.0 → 2/3
95.0
→ courbe changé

si iso 100% +6 -5

100%

Patient Name:
Course:
Plan: SEIG 1A - Unapproved

Y: 2.70 cm
Scaling Factor: 30%
Head First-Supine
User origin DICOM offset = (-0.63cm, 7.53cm, 15.30cm)

Linear accelerator

- mm accuracy
- 100x more powerful than a radiology X-ray

Outline

- Dynamic Radiation Therapy Patient Booking
    - Online stochastic combinatorial optimization
    - Prediction-based scheduling

- Radiation Therapy Treatment Planning
    - Unsupervised learning to reduce problem size
    - Trajectory optimization for Cyberknife

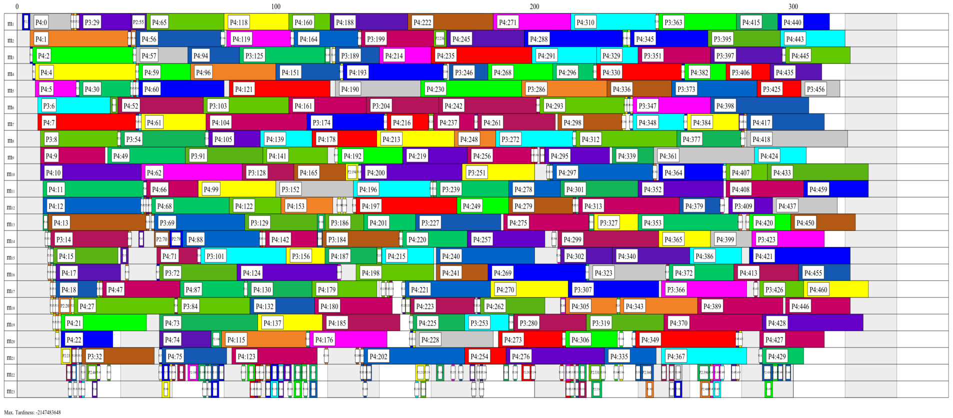
When to book a patient ?

Patient diagnosed

how much time ?

Patient is treated

Considering existing calendar…



… and patient priorities

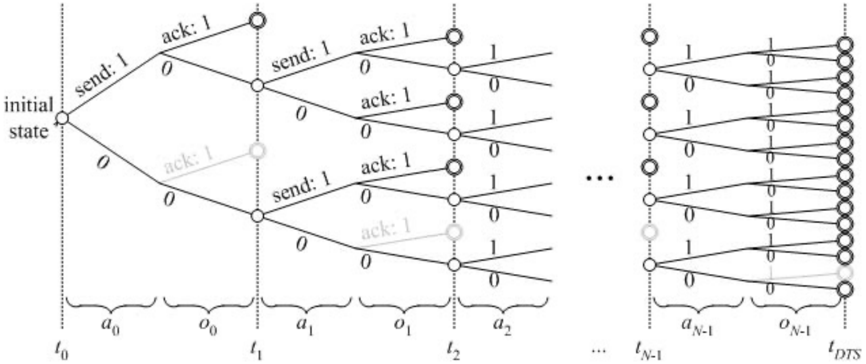
Palliative (P2)	Curative 1 (P3)	Curative 2 (P4)
< 3 days	< 14 days	< 28 days

Different possible approaches

Stochastic Optimization

$$\begin{array}{ll} \min\limits_{x\in\mathbb{R}^n} & g(x) = c^T x + E[Q(x,\xi)] \\ \text{subject to} & Ax = b \\ & x \geq 0 \end{array} \quad \longleftrightarrow \quad \begin{array}{ll} \min\limits_{y\in\mathbb{R}^m} & q(\xi)^T y \\ \text{subject to} & T(\xi)x + W(\xi)y = h(\xi) \\ & y \geq 0 \end{array}$$

Markov Decision Process



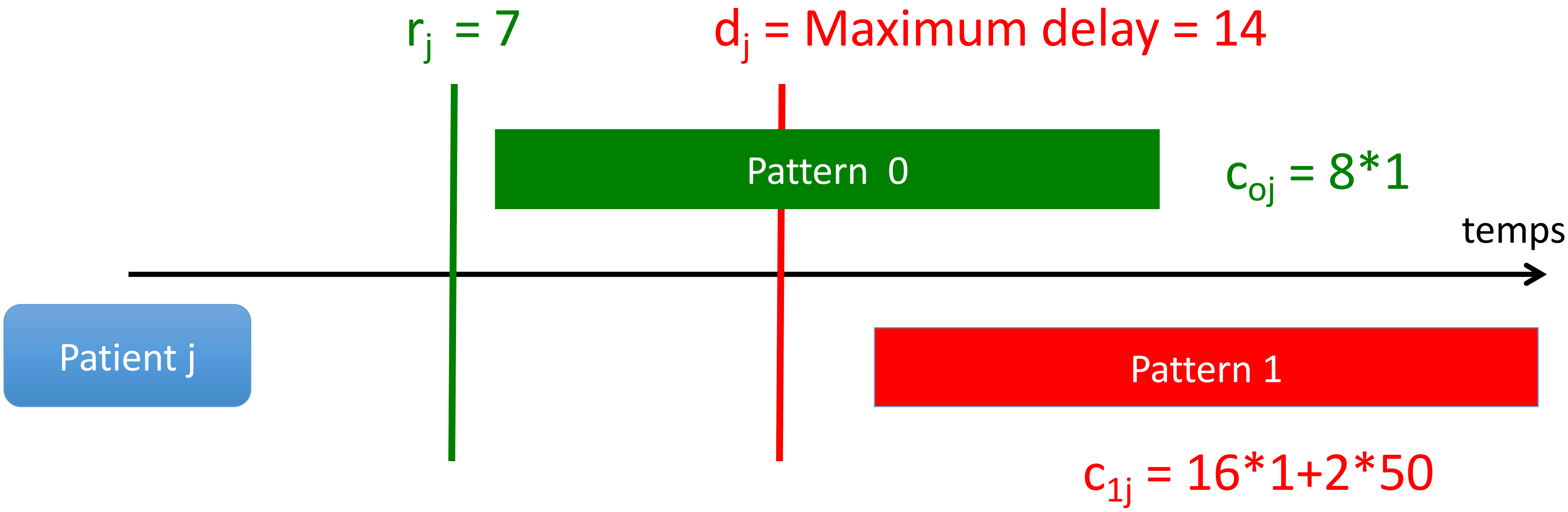
Online Optimization

RT cancer patient booking

- Online stochastic combinatorial optimization:
  1. For each solution, we compute :
     1. A utilization cost (by day and by linac) for a time slot;
  2. We choose the appointment of minimum cost:
     1. Waiting time cost (depending of the priority) ;
     2. Expected utilization cost.

- Booking model -> Dantzig-Wolfe decomposition;
- Uncertainties -> Benders decomposition.

Structure of the model



$r_j = 7$

d_j = Maximum delay = 14

Pattern 0

$c_{oj} = 8*1$

temps

Patient j

Pattern 1

$c_{1j} = 16*1+2*50$

Treatment planning fix to 7 days

Stochastic Programming Model

$$\min \sum_{i\in S_j} c_{ij}x_{ij} + \mathbb{E}_{\omega\in\Omega_j}[\sum_{l\in\mathcal{P}^\omega}\sum_{i\in S_l} c_{il}y_{il}^\omega + \sum_{k\in H}\sum_{m\in M} c^o z_{mk}^\omega]$$

subject to:

$$\sum_{i\in S_j} x_{ij} = 1$$

$$\sum_{i\in S_l} y_{il}^\omega = 1, \qquad \forall \omega\in\Omega_j, \forall l\in\mathcal{P}^\omega$$

Approximated utilization cost of a given initial treatment time slot

$$\boxed{\sum_{i\in S_j} a_{ijk}^m x_{ij} + \sum_{l\in\mathcal{P}^\omega}\sum_{i\in S_l} a_{ilk}^m y_{il}^\omega \le F_k^m + z_{mk}^\omega, \qquad \forall m\in M, \forall k\in H, \forall\omega\in\Omega_j}$$

= dual variable of this constraint

Choose greedily the pattern with the best reduced cost

$$\mathbb{1}_{\mathcal{P}_p}(j)\sum_{i\in S_j} a_{ijk}^m x_{ij} + \sum_{l\in\mathcal{P}_p^\omega}\sum_{i\in S_l} a_{ilk}^m y_{il}^\omega \ge z_{mk}^\omega, \qquad \forall m\in M, \forall k\in H, \forall\omega\in\Omega_j$$

$$\sum_{k=b}^{b+4} z_{mk}^\omega \le O_{week}, \qquad \forall m\in M, \forall b\in\mathcal{B}, \forall\omega\in\Omega_j$$

$$z_{mk}^\omega \in [0, O_{day}], \qquad \forall m\in M, \forall k\in H, \forall\omega\in\Omega_j$$

$$x_{ij}\in\{0,1\}, \qquad \forall i\in S_j$$

$$y_{il}^\omega\in\{0,1\}, \qquad \forall l\in\mathcal{P}^\omega, \forall i\in S_l, \forall\omega\in\Omega_j$$

Initial Results – Regional Hospital

	Due date violations			Average waiting time			Utilization	Overtime
	>3	>14	>28	Palliative	Curative 1	Curative 2		
CICL	14	16	0	2,07	14,38	12,98	88,3%	44
OSCO - 1	9	6	0	1,05	10,57	15,98	88,0%	6

CICL real data:
- 170 patients ;
- 120 days;
- 2 linacs with 23 slots.

Very small and simple with homogenous appointement times

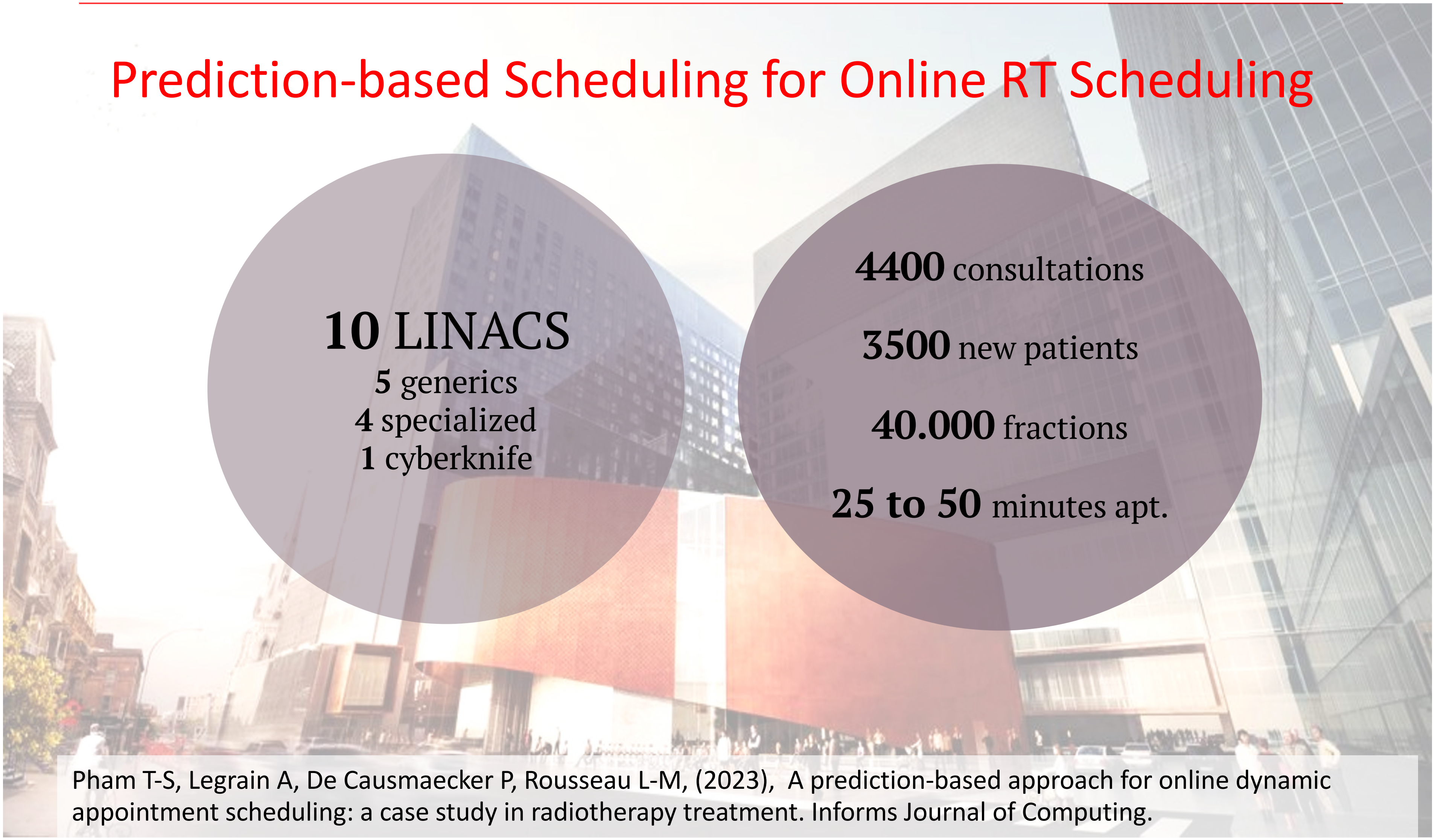
Legrain A, Fortin MA, Lahrichi N, Rousseau L-M (2015) “Online Stochastic Optimization of Radiotherapy Patient Scheduling”, *Healthcare Management Science*, 18, 110-123.

Outline

- Dynamic Radiation Therapy Patient Booking
    - Online stochastic combinatorial optimization
    - Prediction-based scheduling

- Radiation Therapy Treatment Planning
    - Unsupervised learning to reduce problem size
    - Trajectory optimization for Cyberknife

UNIVERSITY OF MONTRÉAL HOSPTITAL CENTER (CHUM)



Prediction-based Scheduling for Online RT Scheduling

10 LINACS
5 generics
4 specialized
1 cyberknife

4400 consultations

3500 new patients

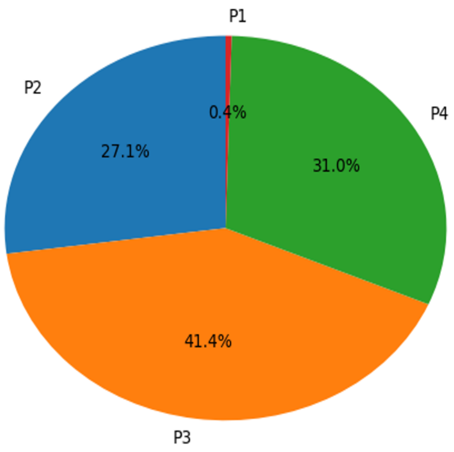
40.000 fractions

25 to 50 minutes apt.

Pham T-S, Legrain A, De Causmaecker P, Rousseau L-M, (2023), A prediction-based approach for online dynamic appointment scheduling: a case study in radiotherapy treatment. Informs Journal of Computing.

CHUM - 2019

	Category	Treatment deadline (days)	Percentage of overdue treatment (%)	Average waiting time (days)
Palliative	**P1**	1	14.29	1.09
	P2	3	79.89	6.91
Curative	**P3**	14	74.55	18.11
	P4	28	29.89	22.59

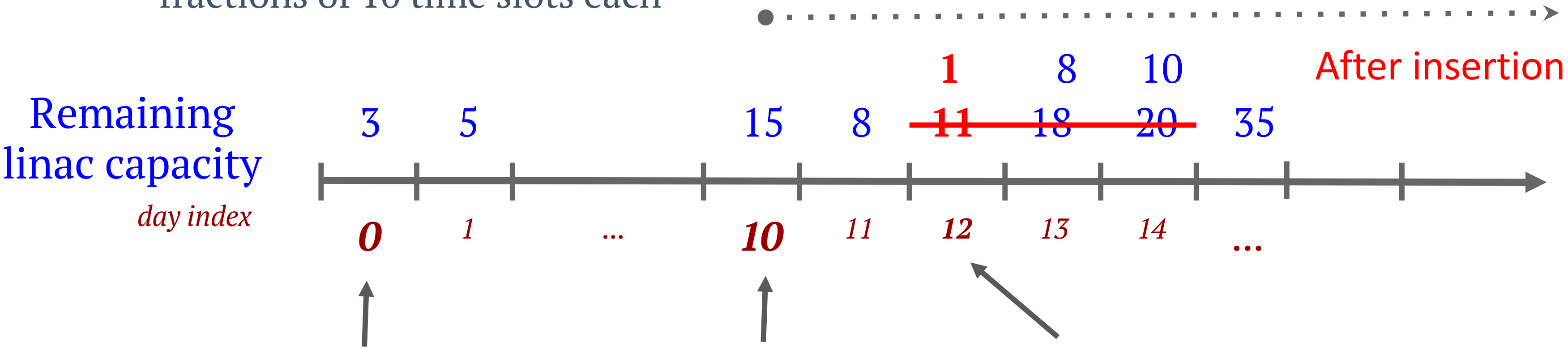


Objective: minimizing overdue treatment and waiting time

Online Scheduling with a Greedy Heuristic

- 1 linac, capacity 120 time slots
- a curative patient (P4) with 3 fractions of 10 time slots each

Looking for the first eligible date that can accommodate the whole treatment

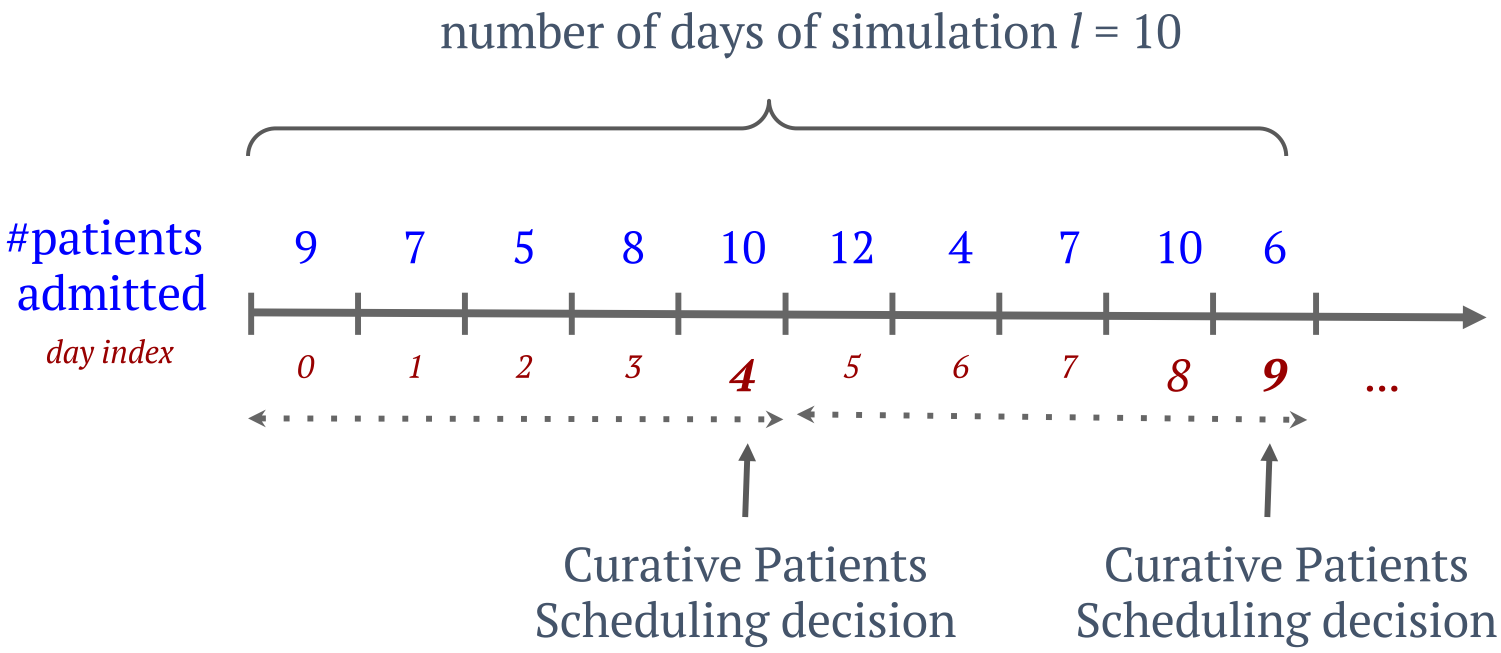


Patient admitted

Start looking at one/two weeks after admission

The first eligible date

Batch Scheduling



number of days of simulation $l = 10$

#patients admitted

9 7 5 8 10 12 4 7 10 6

day index

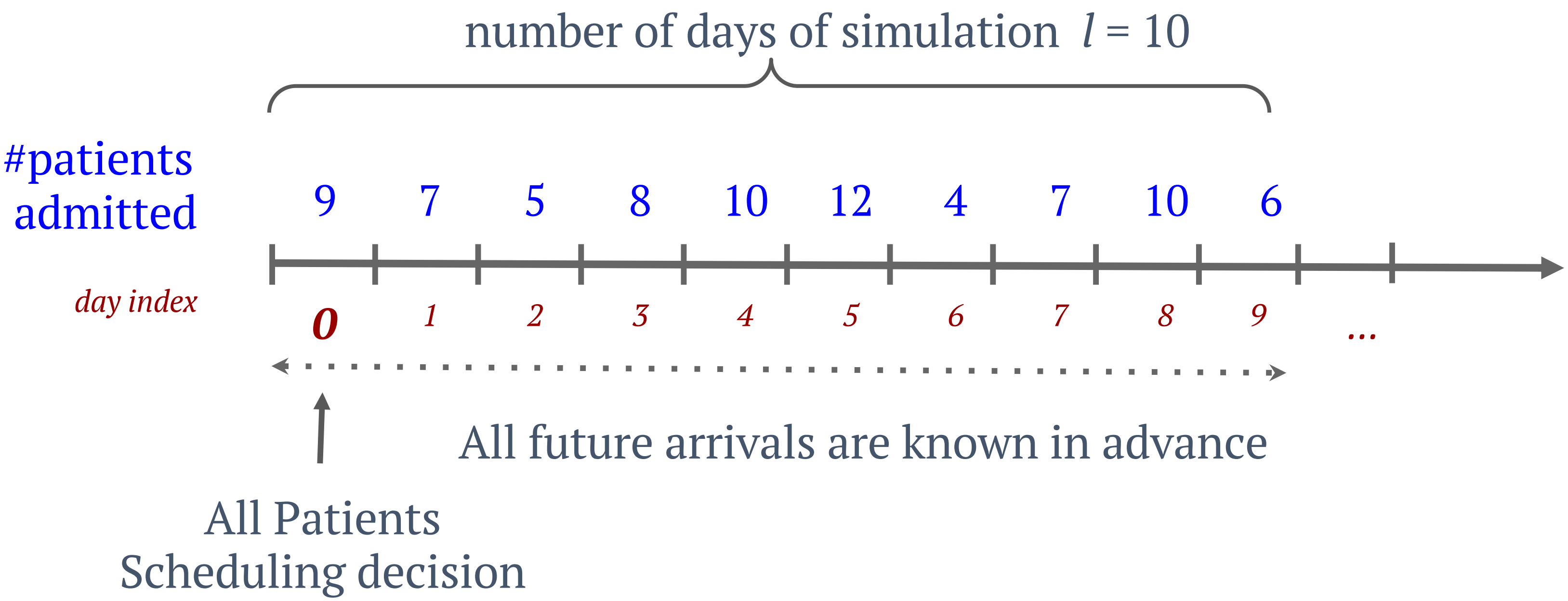
0 1 2 3 ***4*** *5 6 7* ***8*** ***9*** …

Curative Patients Scheduling decision

Curative Patients Scheduling decision

Palliative patients: schedule at arrival

Offline Scheduling – with Perfect Information



number of days of simulation $l = 10$

#patients admitted	9	7	5	8	10	12	4	7	10	6	
day index	***0***	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	…

All future arrivals are known in advance

All Patients
Scheduling decision

A MIP Model for Batch\Offline Scheduling

$$x_{tl}^{i} = \begin{cases} 1 & \text{if patient } i \text{ receives their treatment on day } t\text{, linac } l \\ 0 & \text{otherwise} \end{cases}$$

$$\begin{aligned} \text{minimize} & \sum_{i \in \hat{\mathcal{P}}} \sum_{t \in \mathcal{T}, t > a_i} \sum_{l \in \mathcal{L}} \omega_1 \boxed{(t - a_i) log(t - a_i + 1) x_{tl}^{i}} \\ & + \sum_{i \in \hat{\mathcal{P}}} \sum_{t \in \mathcal{T}, t > d_i} \sum_{l \in \mathcal{L}} \omega_2 \boxed{(t - d_i) log(t - d_i + 1) x_{tl}^{i}} \end{aligned}$$

waiting time

overdue time

A MIP Model for Batch\Offline Scheduling

$$\sum_{t\in\mathcal{T}}\sum_{l\in\mathcal{L}} x^i_{tl} = 1 \qquad \forall i \in \hat{\mathcal{P}}$$

assignment constraint

$$x^i_{tl} = 0 \qquad \forall i \in \hat{\mathcal{P}}, l \in \mathcal{L}, t \in \{0, \dots, r_i - 1\}$$

ready date

$$\sum_{i\in\hat{\mathcal{P}}}\sum_{t'=max\{0,t-I_i+1\}}^{t} p_i x^i_{t'l} \leq \hat{C}^t_l \qquad \forall t \in \mathcal{T}, l \in \mathcal{L}$$

capacity constraints

$$\sum_{i\in\mathcal{P}^C}\sum_{t'\in\{t-I_i+1,\dots,t\}} p_i x^i_{t'l} \leq max\{0, \hat{C}^t_l - \gamma C^t_l\} \qquad \forall t \in \mathcal{T}, l \in \mathcal{L}$$

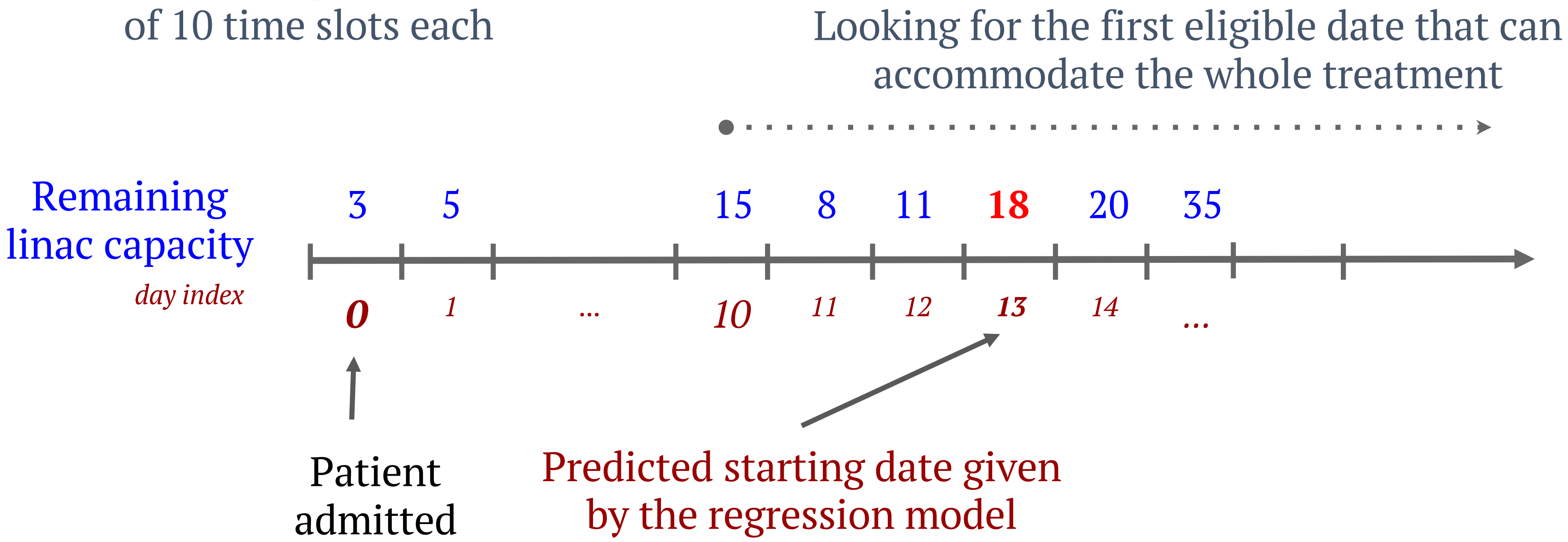
reserved capacity

$$x^i_{tl} \in \{0, 1\} \qquad \forall i \in \hat{\mathcal{P}}, t \in \mathcal{T}, l \in \mathcal{L}$$

WARNING: CURRENTLY IGNORES INTRA-DAY SCHEDULING

Prediction-based Scheduling

- 1 linac, capacity 120 time slots
- a curative patient with 3 fractions of 10 time slots each



Looking for the first eligible date that can accommodate the whole treatment

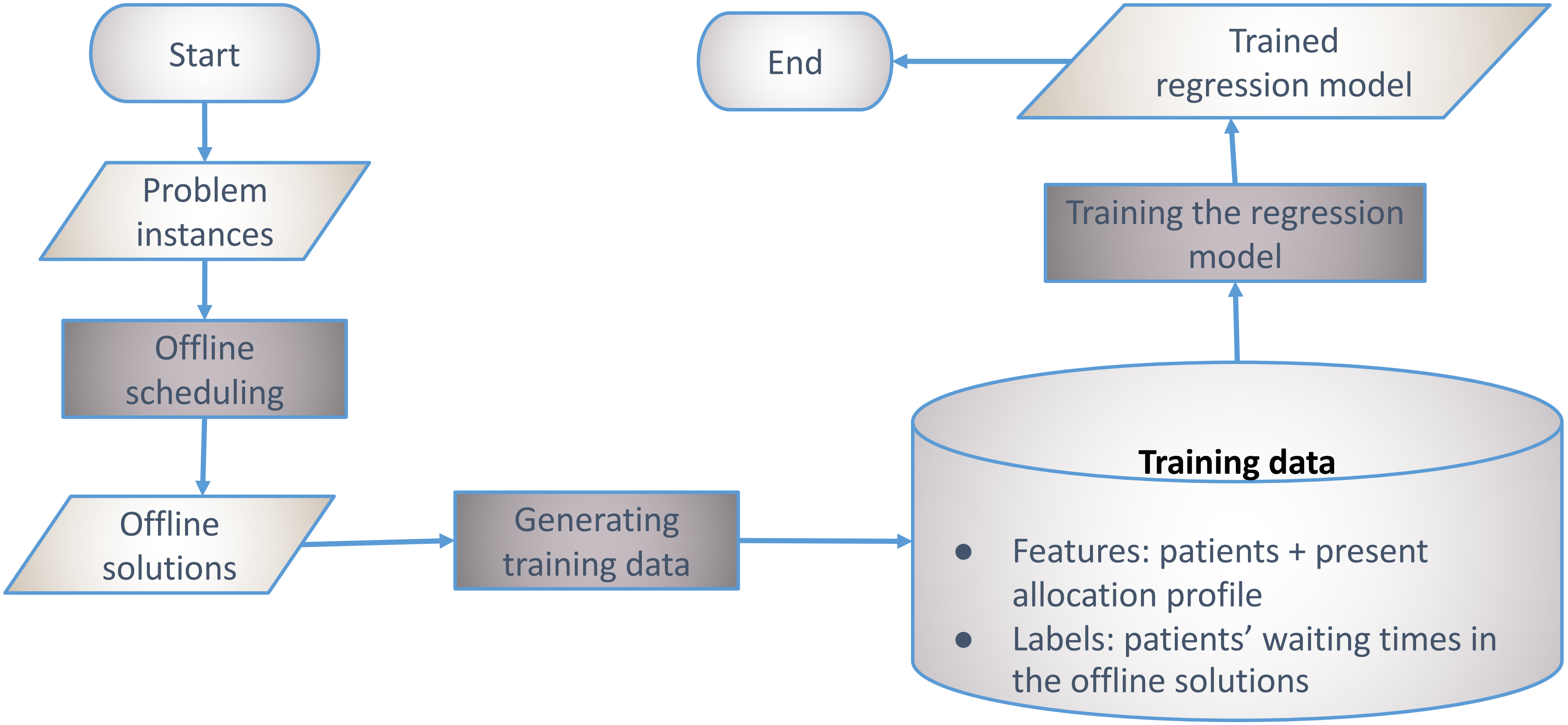
Remaining linac capacity	3	5		15	8	11	**18**	20	35
day index	***0***	*1*	…	*10*	*11*	*12*	***13***	*14*	…

Patient admitted

Predicted starting date given by the regression model

How do we predict a "good" starting date for a patient?

Training the Regression Model for Scheduling



Training data

- Features: patients + present allocation profile
- Labels: patients' waiting times in the offline solutions

Constructing Training data from an Offline Solution

Notation	Explanation
$\mathcal{P}$	set of patients
C	total linac capacity
α_i	admission time point of patient i
d_i	due date of patient i ($d_i \leq r_i$)
I_i	number of fractions of patient i
p_i	fraction's length of patient i
$\hat{c}_d^\phi$	available capacity of all linac on day d, measured at time point ϕ
$\hat{C}^\phi$	set of *present capacity* of all days in the *sample horizon* $\mathcal{D}^\phi$
w_i^*	waiting time of patient i in the offline solution s_i^*

- Datapoints: $X_i = \{r_i, I_i, d_i, p_i, \hat{C}^{\phi,\phi=\alpha_i}\} \Rightarrow \mathcal{X}$
- Labels: $y_i = w_i^* \Rightarrow \mathcal{Y}$
- Estimate: $\xi : \mathcal{X} \rightarrow \mathcal{Y}$

MIP as a
? Labelling
Machine

- Objective: $$\mathcal{L}(\phi) = \underbrace{\sum_i l(\hat{y}_i, y_i)}_{\text{Loss function}} + \underbrace{\sum_k \Omega(f_k)}_{\text{Regularization}}$$

Constructing Training data from an Offline Solution

Notation	Explanation
$\mathcal{P}$	set of patients
C	total linac capacity
α_i	admission time point of patient i
d_i	due date of patient i ($d_i \leq r_i$)
I_i	number of fractions of patient i
p_i	fraction's length of patient i
$\hat{c}_d^{\phi}$	available capacity of all linac on day d, measured at time point ϕ
$\hat{C}^{\phi}$	set of *present capacity* of all days in the *sample horizon* $\mathcal{D}^{\phi}$
w_i^*	waiting time of patient i in the offline solution s_i^*

MIP as a Episode Builder ?

STATE ?

- Datapoints: $X_i = \{r_i, I_i, d_i, p_i, \hat{C}^{\phi,\phi=\alpha_i}\} \Rightarrow \mathcal{X}$
- Labels: $y_i = w_i^* \Rightarrow \mathcal{Y}$ ACTION ?
- Estimate: $\xi : \mathcal{X} \rightarrow \mathcal{Y}$
- Objecti

$$\mathcal{L}(\phi) = \underbrace{\sum_i l(\hat{y}_i, y_i)}_{\text{Loss function}} + \underbrace{\sum_k \Omega(f_k)}_{\text{Regularization}}$$

Data Generation

Patient arrivals: Poisson distribution

Treatment plans: based on historical data

Instance setting

- Number of linacs
- Arrival rate (average daily number of patients)

For each instance setting: 500 instances

- 400 for training the regression model
- 100 for testing

PREDICTIVE MODELS

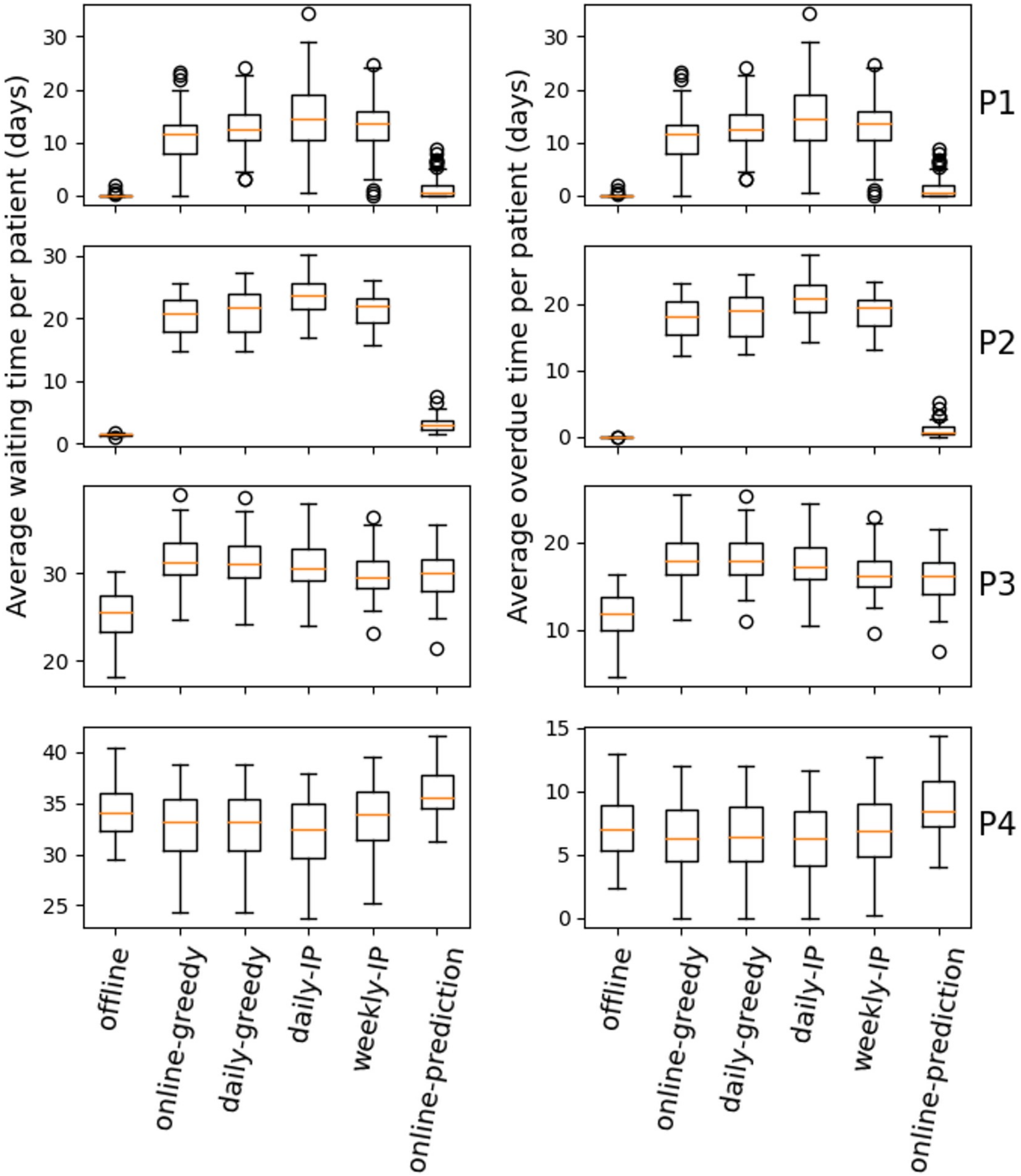
	Training time	Training		Testing	
		MSE	MAE	MSE	MAE
MLP	116.19	3.45	1.32	3.33	1.29
SGD	0.35	6.06	1.84	5.61	1.77
Lasso	0.44	5.97	1.81	5.52	1.74
ElasticNet	0.25	6.26	1.85	5.83	1.8
SVR	43.16	3.19	1.07	3.12	1.07
Decision Tree	0.84	2.41	0.48	6.59	1.4
Random forest	51	**0.38**	**0.39**	2.64	1.03
XGBoost	7.71	0.96	0.66	**2.44**	**0.97**

DYNAMIC SCHEDULING STRATEGIES

	Scheduling strategy	Scheduling palliative patients	Scheduling curative patients
Batch scheduling	**Offline**	Scheduling once with all future arrivals known in advance	
	Daily	Every day	Every day
	Weekly	Every day	Every Friday
	Daily greedy	Every day	Every day
Online scheduling	**Greedy**	At admission	At admission
	Prediction-based	At admission	At admission

8 LINACS

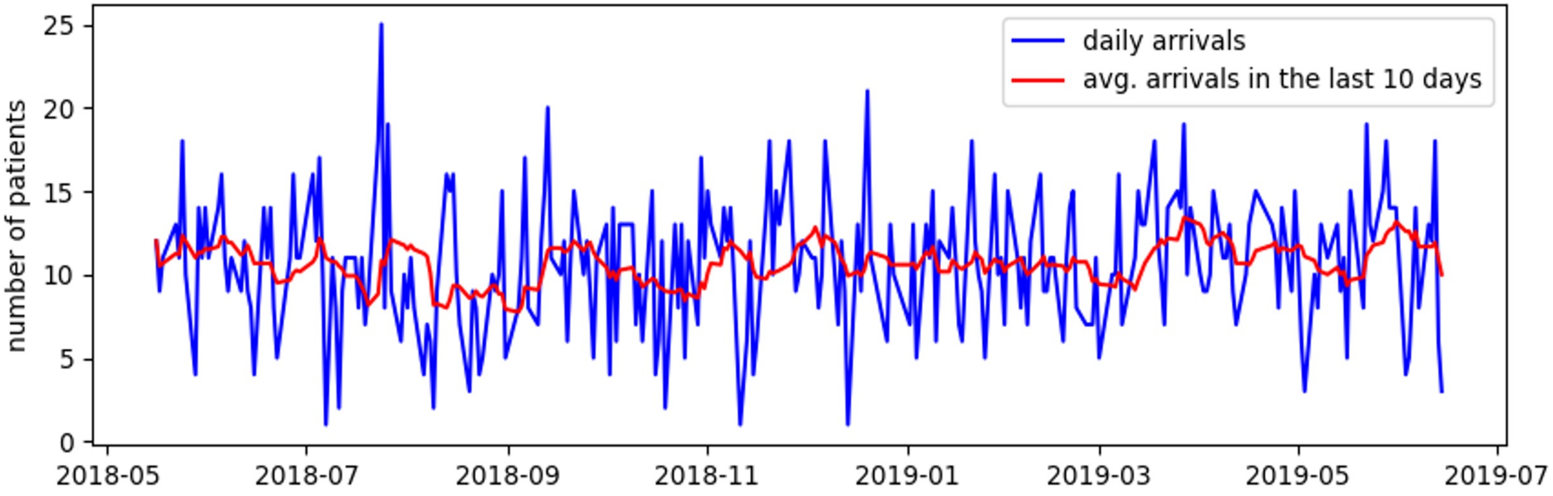
Arrival rate of 12.0



P4 patients are slightly delayed to create enough room for P1 & P2

Experiment on a real CHUM data

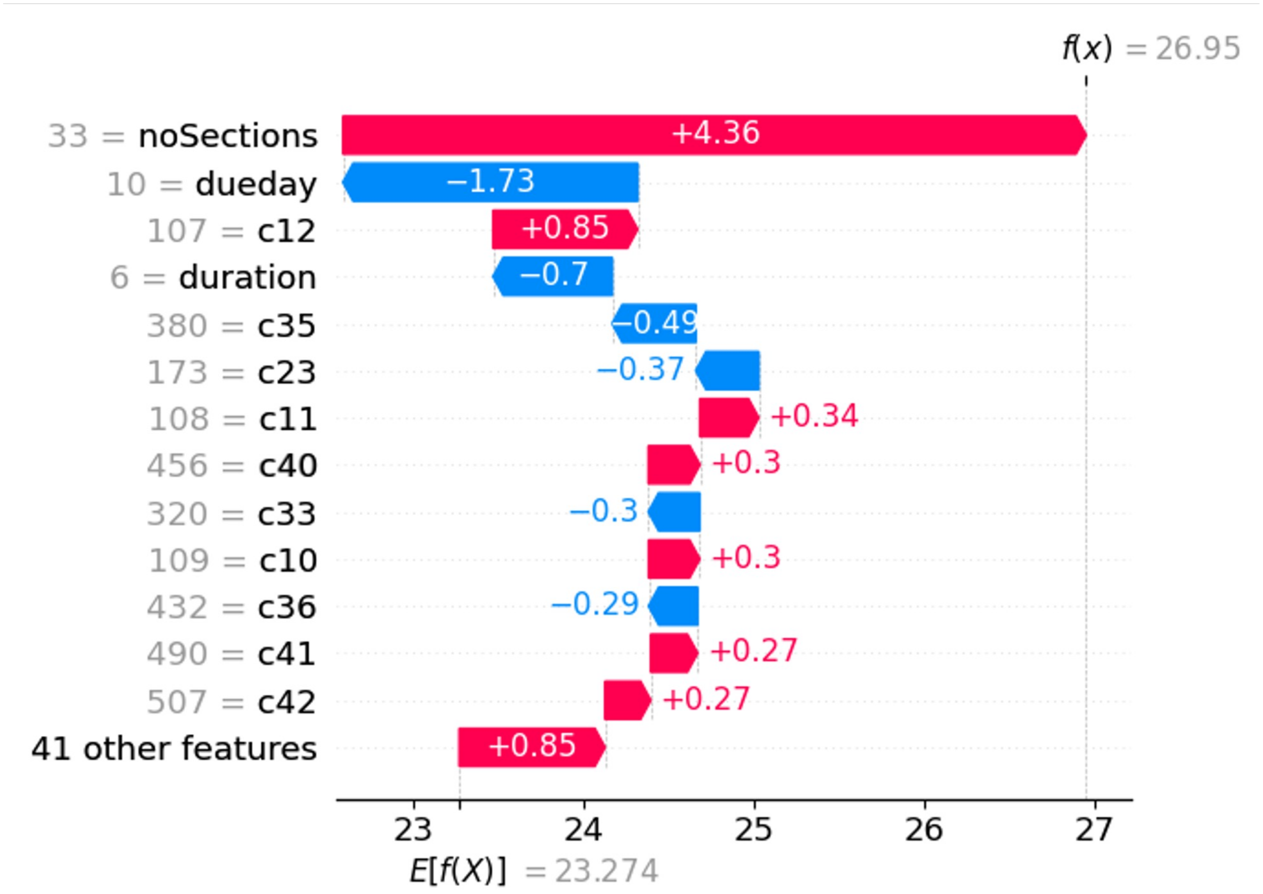
- 7 linacs operating 8 hours/day
- High fluctuation in arrival rate
    - Instance setting for training: arrival rate of 10.1 patients/day



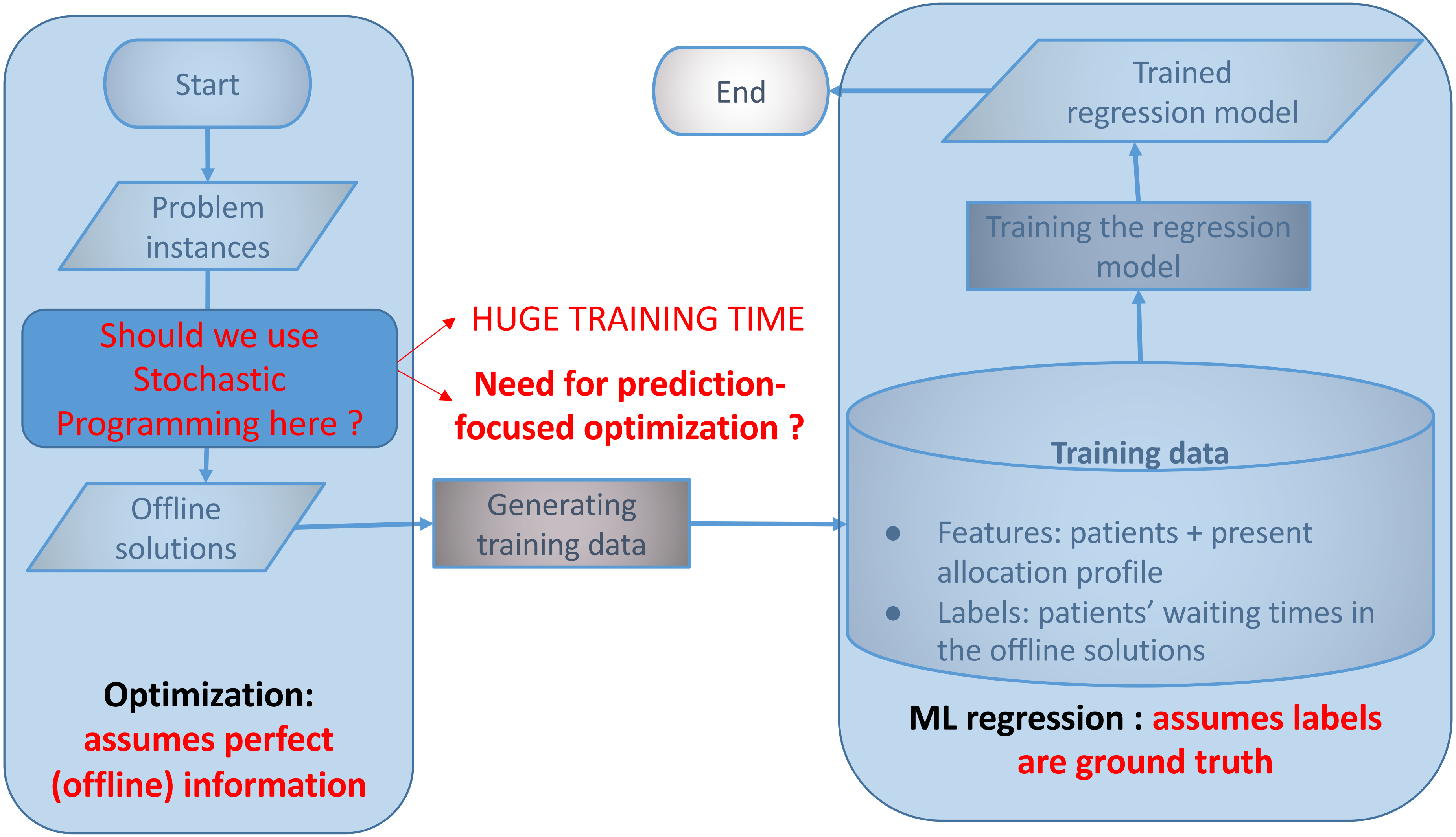
RESULTS ON THE REAL INSTANCE

Scheduling strategy	Avg. occupancy (%)	Waiting time (days)					Overdue time (days)				
		overall	P1	P2	P3	P4	overall	P1	P2	P3	P4
online-greedy	97.45	33.02	5.14	6.13	43.67	44.02	44.02	5.14	3.91	29.74	16.18
daily-greedy	97.51	32.91	6.00	6.23	43.48	43.80	17.71	6.00	3.99	29.58	16.00
daily	97.72	33.53	9.79	9.63	42.87	**43.44**	18.25	9.79	7.15	28.93	**15.65**
weekly	97.61	33.04	7.86	7.72	**42.42**	44.10	17.76	7.86	5.37	**28.51**	16.19
prediction-based	97.14	32.93	**3.29**	**4.05**	44.21	44.94	17.69	**3.29**	**1.99**	30.22	16.96

Explainability



Training the Regression Model for Scheduling



Start

Problem instances

Should we use Stochastic Programming here ?

Offline solutions

Optimization: assumes perfect (offline) information

HUGE TRAINING TIME

Need for prediction-focused optimization ?

Generating training data

End

Trained regression model

Training the regression model

Training data

- Features: patients + present allocation profile
- Labels: patients’ waiting times in the offline solutions

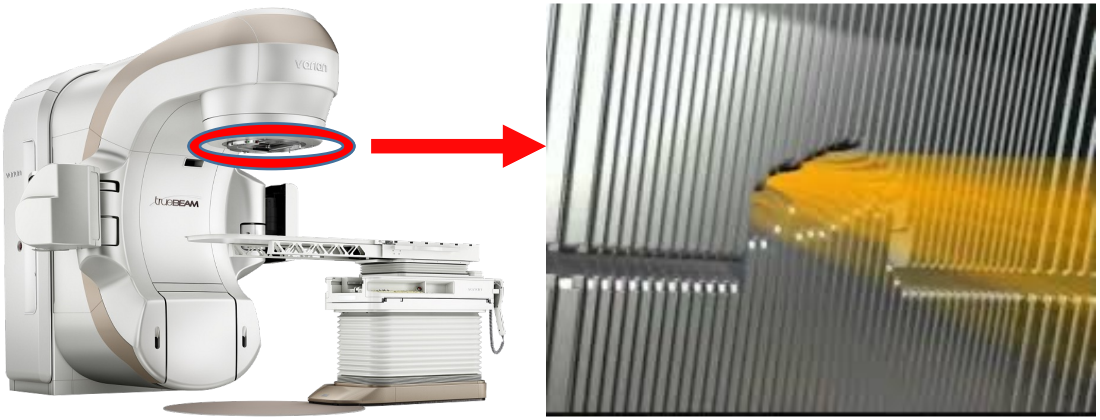
ML regression : assumes labels are ground truth

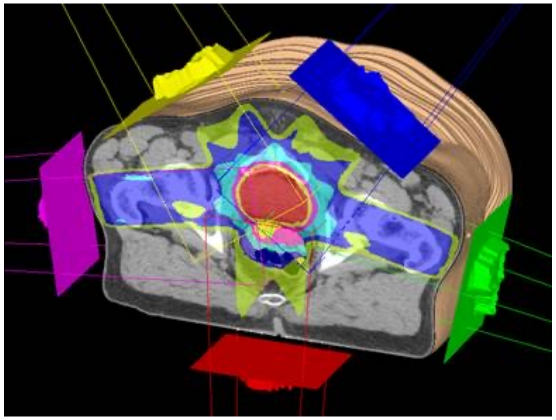
Outline

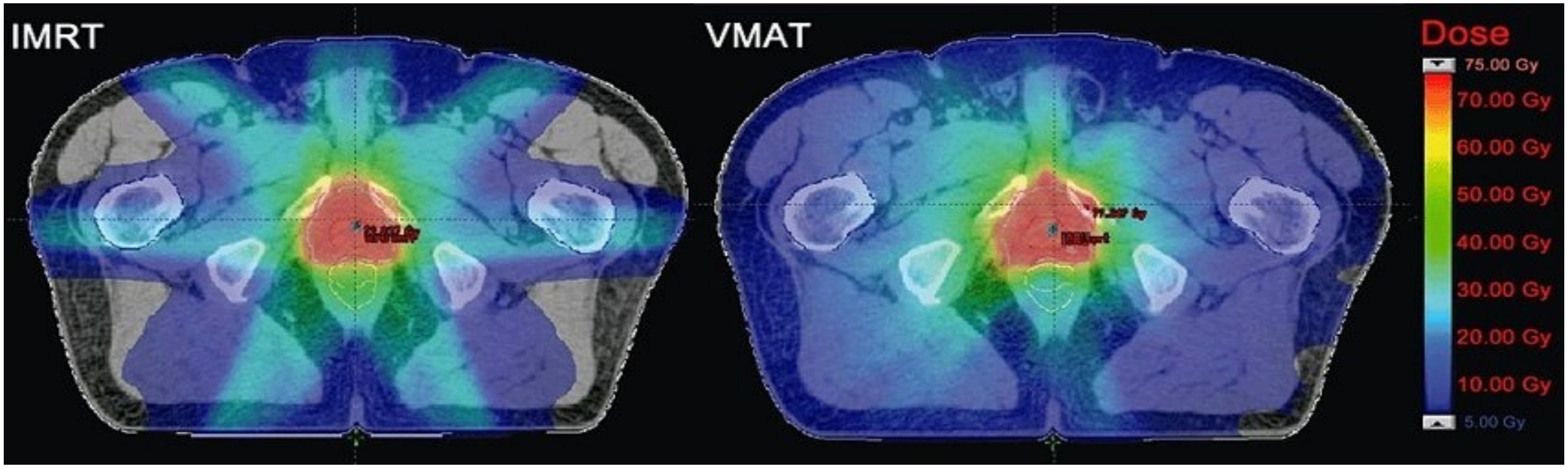
- Dynamic Radiation Therapy Patient Booking
    - Online stochastic combinatorial optimization
    - Prediction-based scheduling

- Radiation Therapy Treatment Planning
    - Unsupervised learning to reduce problem size
    - Trajectory optimization for Cyberknife

Planning Treatment for Radiation Therapy







Mahnam M, Gendreau M, Lahrichi N, Rousseau L-M, (2017), “Simultaneous delivery time and aperture shape optimization for the volumetric-modulated arc therapy (VMAT) treatment planning problem”, *Physics in Medicine and Biology*, 62,

VMAT: Delivery time & aperture shape optimization

- Vocabulary:
  - **Voxel**: a cube in the body (a 3D pixel)
  - **Sector**: is a position (angle) around the body.
  - **Aperture**: a configuration of the tungsten leafs
  - **Beamlet:** the smallest possible beam
  - **Dose**: the amount of energy deposited in a voxel (in Gray)

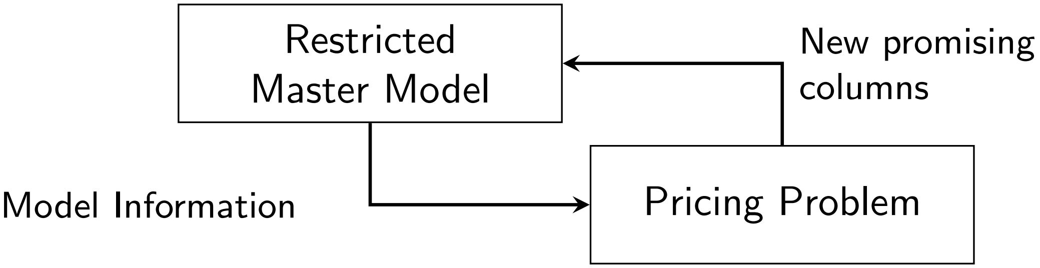
- **Decisions**:
  - Selecting a **sequence** (each 2°) of apertures.
  - Determining the beam energy & rotation speed.

- **Objectives:**
  - Maximize plan quality (deposited dose match prescribed dose)
  - Minimize treatment time

VMAT: Delivery time & aperture shape optimization

- Highly combinatorial problem:
    - In a small case with (5 × 10) beams and 100 sectors,
    - there are 7.1×10^{251} apertures shapes.
    - Real problem is (80x80) x 180 sectors

- Using Column Generation (CG): a Mathematical Optimization technique for solving large-scale problems
    - Exploits decomposable structures
    - Handles large number of variables

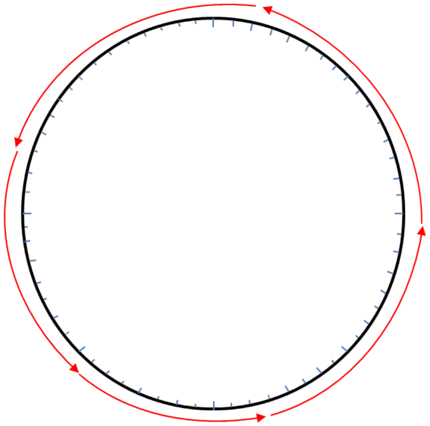


Restricted Master Model

New promising columns

Model Information

Pricing Problem



Master Model: arc and intensity selection

Objective function

- quadratic voxel-based penalty function + delivery time

Constraints

1. Calculating the dose deviation from prescribed thresholds
2. Each sector should be covered by at most one arc
3. Restricting the change of dose rate between adjacent sectors
4. Restricting the dose rate to the max R
5. The gantry speed at each sector should be enough for leaf motions of the assigned arc
6. Restricting the change of sector time between adjacent sectors
7. Restricting the sector time to lower and upper bounds
8. Restricting the maximum total treatment time.

Master Model: arc and intensity selection

$$\mathbf{GP}: \min \quad \boxed{\mathsf{F}(z) + w\,T_{max}}$$

Weighted quality and time objective

$$z_j = \sum_{k \in K} \sum_{h \in H_k} D_{jh}(A_h^k)\, y^k\, \rho_h\, t_h \qquad \forall j \in \mathcal{V}$$

$$\sum_{k \in K} a_h^k\, y^k \leq 1 \qquad \forall h \in H$$

$$\mid \rho_{h+1} - \rho_h \mid \leq \Delta_\rho \qquad \forall h = 1, 2, \cdots, |H| - 1$$

$$0 \leq \rho_h \leq R \qquad \forall h \in H$$

$$\sum_{k \in K} \tau_{h,h+1}^k\, y^k \leq t_h \qquad \forall h \in H$$

$$\mid t_{h+1} - t_h \mid \leq \Delta_t \qquad \forall h = 1, 2, \cdots, |H| - 1$$

$$\underline{T} \leq t_h \leq \overline{T} \qquad \forall h \in H$$

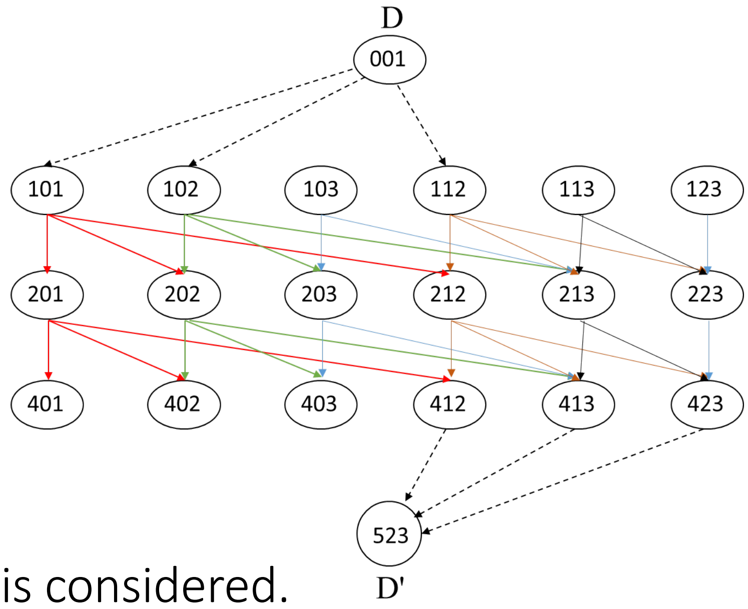
$$\sum_{h \in H} t_h \leq T_{max}$$

$$y^k \in \{0, 1\} \qquad \forall k \in K$$

Subproblem: building new arcs

The situation of each row in each sector is indicated as a node (h, l, r);

- e.g. node $(90, 0, 4)_5$ is the position of leaves of row 5 in sector 90:



Constraints include:

1. Maximum leaf motion constraint is considered.
2. Conflicting trailing and leading leaves are avoided, i.e. t + 1 ≤ r
3. Cost of nodes and arcs based on the Master Model (dual values)

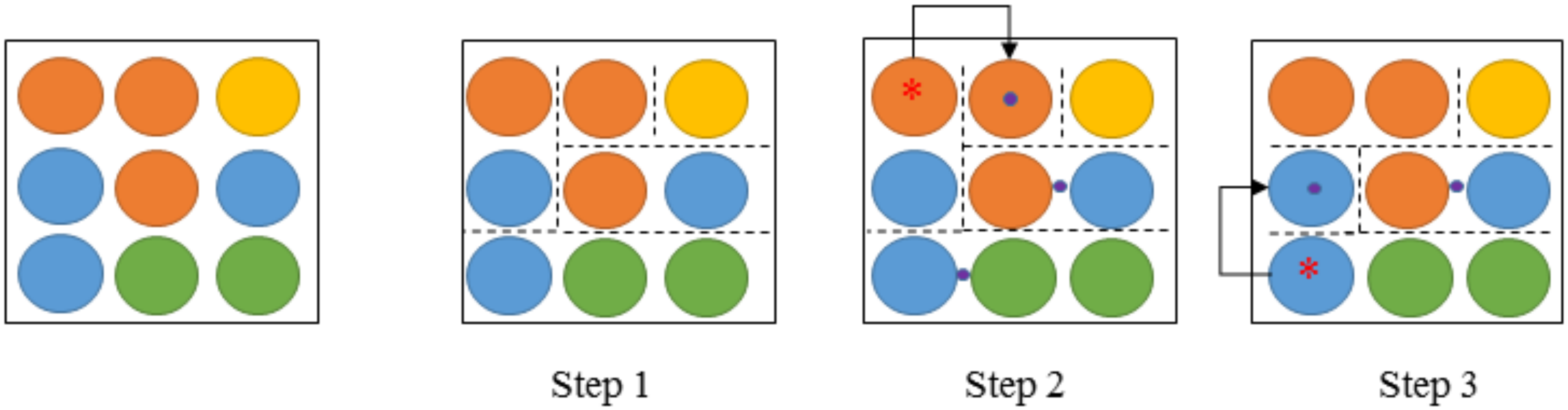
- Polynomial shortest path algorithm easily obtain the best solution.

Data-drive size reduction

Random down-sampling is a usual approach (Kufer et al. 2003)

We propose an unsupervised learning method:

- Observation : similar voxels would be considered in a cluster.
- Each voxel is associated with feature tensor based the dose received from each beamlet, assuming fully opened aperture in all sectors.
- We then apply a variant of the K-Means algorithm.



Reducing the problem size

Set # clusters to 5% of initial number of voxels for normal tissue voxels, 15% for tumor voxels

Voxel aggregation computational results.

Iter	# Transfer Iter	Time (Sec.)	Avg.Dist
0			54.27054
1	19265	0.952947	11.57258
2	2160	0.74742	10.51782
3	301	0.733188	10.44499
4	40	0.729702	10.44053
5	13	0.738105	10.43969

Experimental evaluation

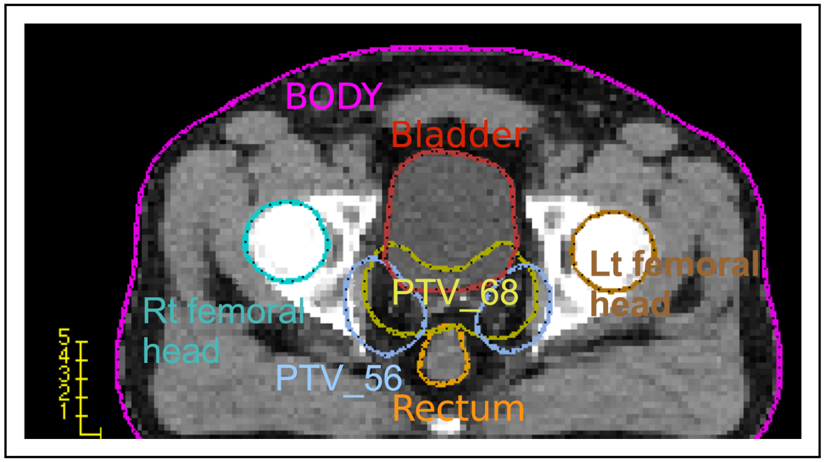
Case Characteristics

Total # beamlets	25,404
Beamlet size (mm)	1 × 1
Voxel resolution (mm)	3, 3, 3
# Target voxels	9491
# Body voxels	690,373

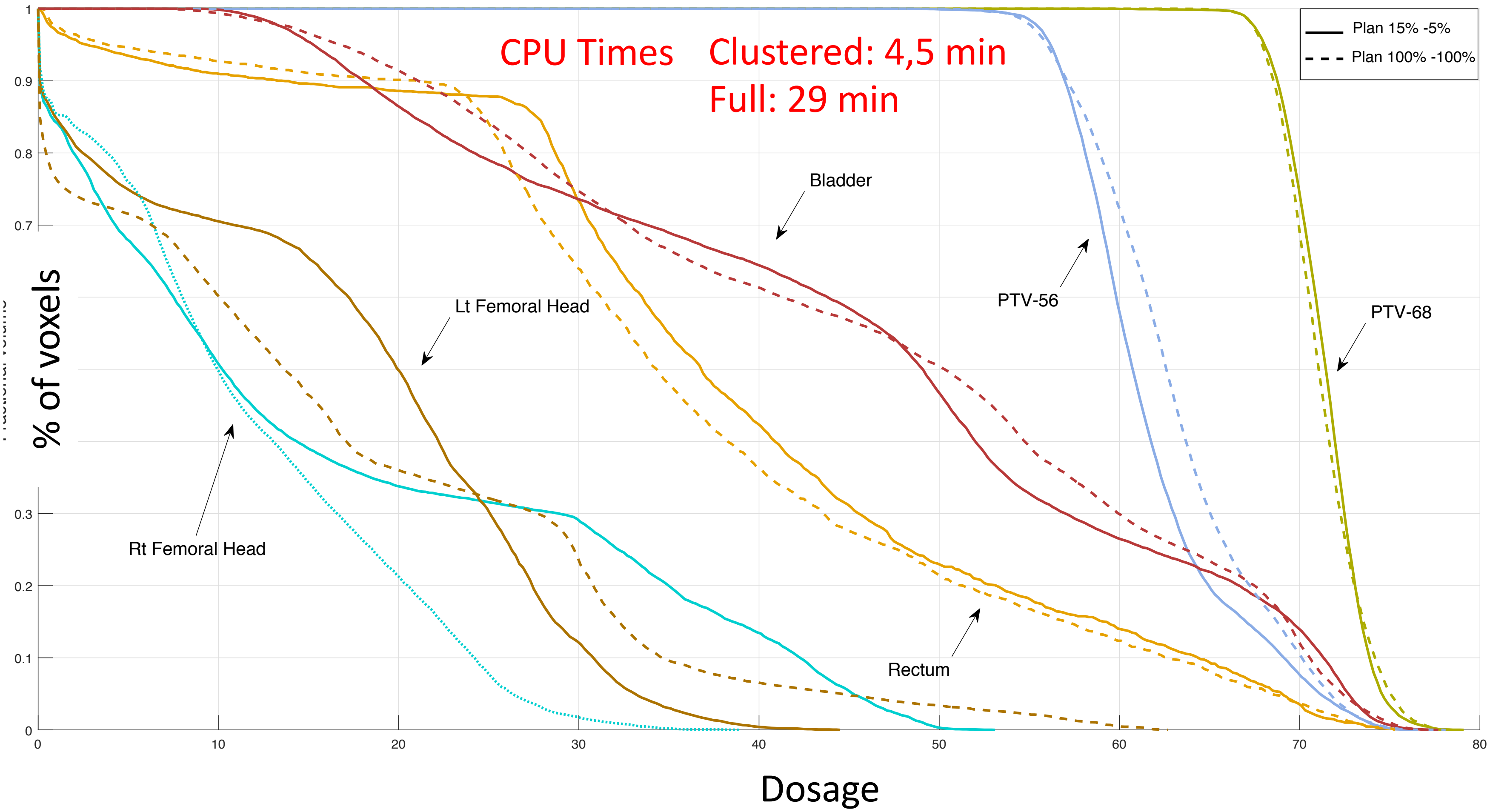
Algorithm Parameters

Max dose rate	600 MU/min
Max leaf speed	3 cm/sec
Max fluence change	2 MU/s
Max time change	2 s
Gantry speed	$[1 \quad 6]^\circ$/sec

- CORT dataset (Craft et al, 2014)
- 180 equispaced sectors
- Algorithm is implemented in C++/CPLEX

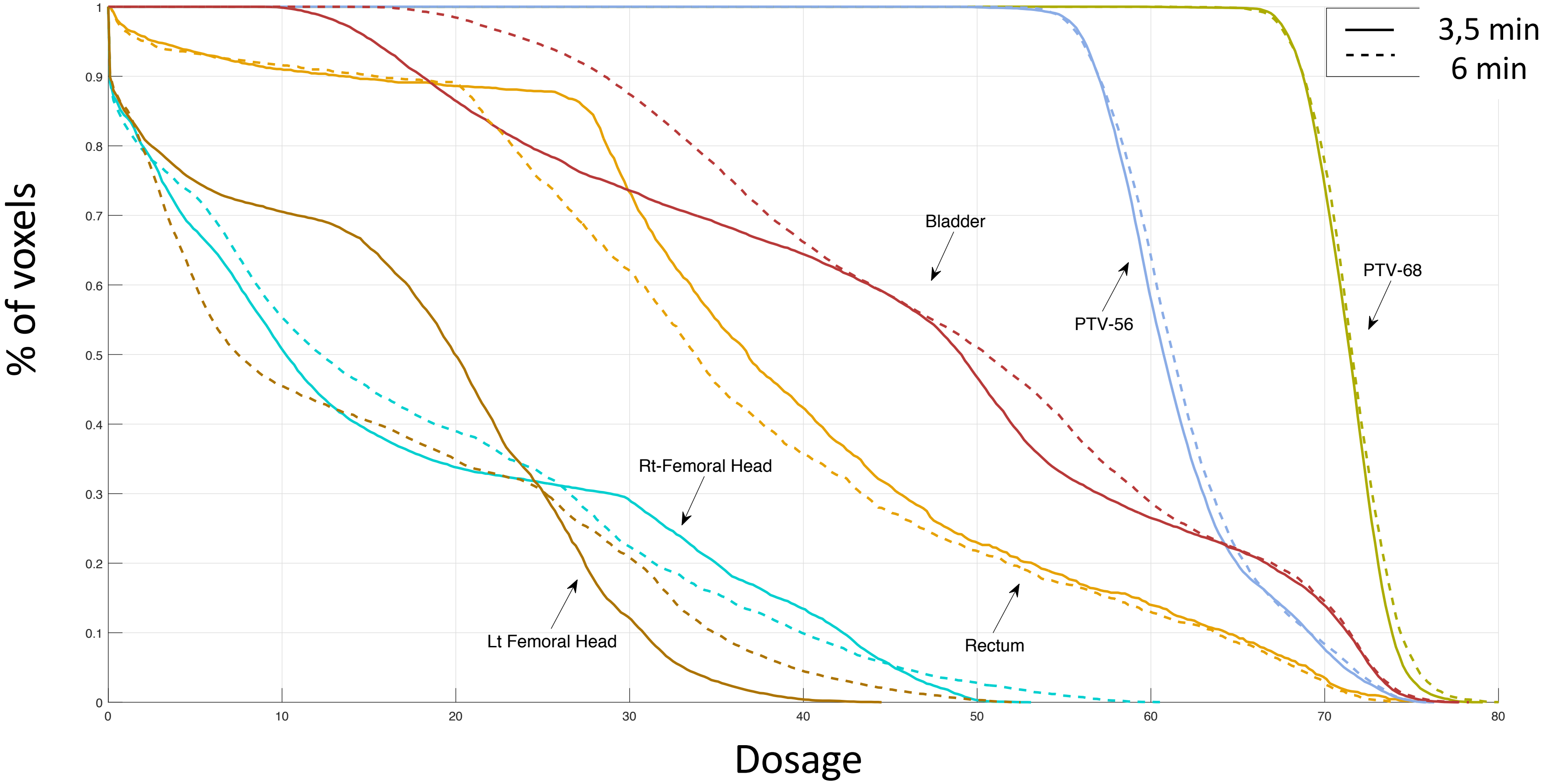


Effect of ML-based aggregation



CPU Times Clustered: 4,5 min
Full: 29 min

Effect of delivery time



CHAIRE DE RECHERCHE DU CANADA EN ANALYTIQUE ET LOGISTIQUE DE SOINS DE SANTÉ
HANALOG

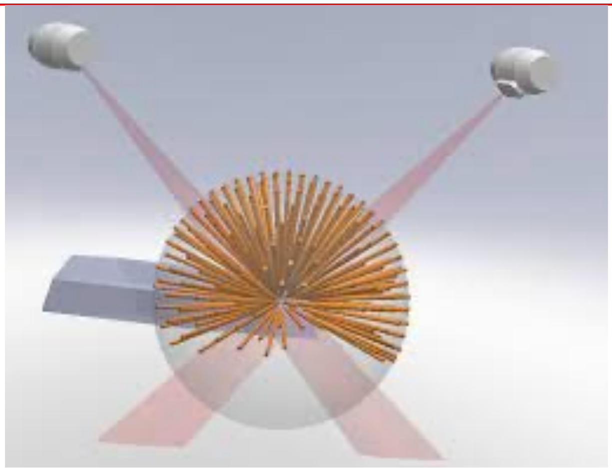
POLYTECHNIQUE MONTRÉAL

Outline

- Dynamic Radiation Therapy Patient Booking
    - Online stochastic combinatorial optimization
    - Prediction-based scheduling

- Radiation Therapy Treatment Planning
    - Unsupervised learning to reduce problem size
    - Trajectory optimization for Cyberknife

Cyberknife (SBRT)



- High quality beam in terms of dose conformity
- BUT Long treatment time (1 hour)
- Up to 70% of treatment time corresponds to the robotic arm movement between nodes

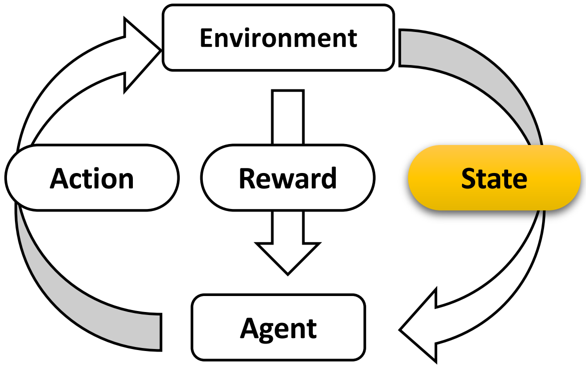
Objectives

1. Minimize the distance covered by the robotic arm
2. Maximize the conformity of delivered dose to the prescribed dose
3. Scatter the beams around the patient to avoid clusters

Kafaei, P., Cappart, Q., Renaud, M. A., Chapados, N., & Rousseau, L. M. (2021). Graph neural networks and deep reinforcement learning for simultaneous beam orientation and trajectory optimization of Cyberknife. Physics in Medicine & Biology, 66(21), 215002.

Reinforcement Learning Framework

A complete acyclic graph between each shooting position (nodes)

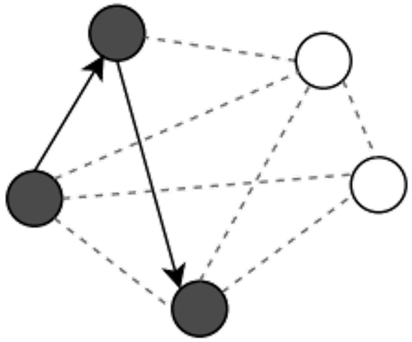


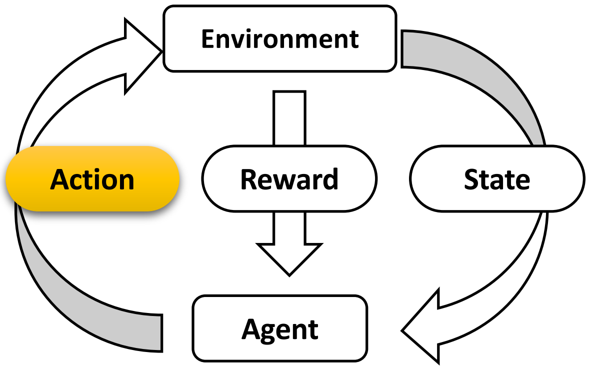
Features

- (x, y, z): coordinates of the nodes
- d_{tar}: dose deposited in the tumor at unit intensity
- d_{oar}: dose deposited in other tissues at unit intensity
- A set of the neighbors of each node

Reinforcement Learning Framework

Next node to be selected





Environment

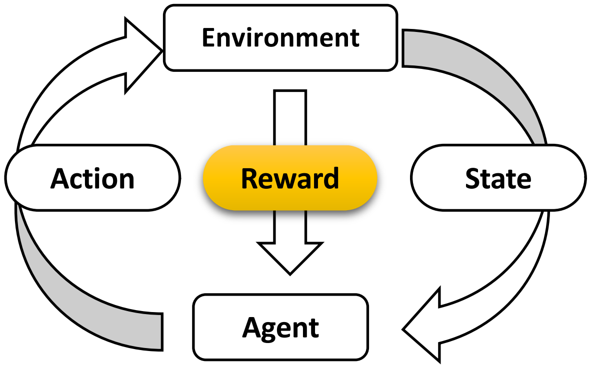
Action

Reward

State

Agent

Reinforcement Learning Framework



Non-terminal State

$$R(s_t, a_t) = -(r_1 + r_2),\ s_t \neq s_\emptyset$$

Terminal State

$$R(s_t, a_t) = -(r_1 + r_2 + r_3),\ \ s_t = s_\emptyset$$

$r_1 \rightarrow$ the Euclidean distance between m and n

$r_2 \rightarrow \frac{d_{oar}}{d_{tar}}$ for beam n (doses are pre-computed with a MC simulation engine)

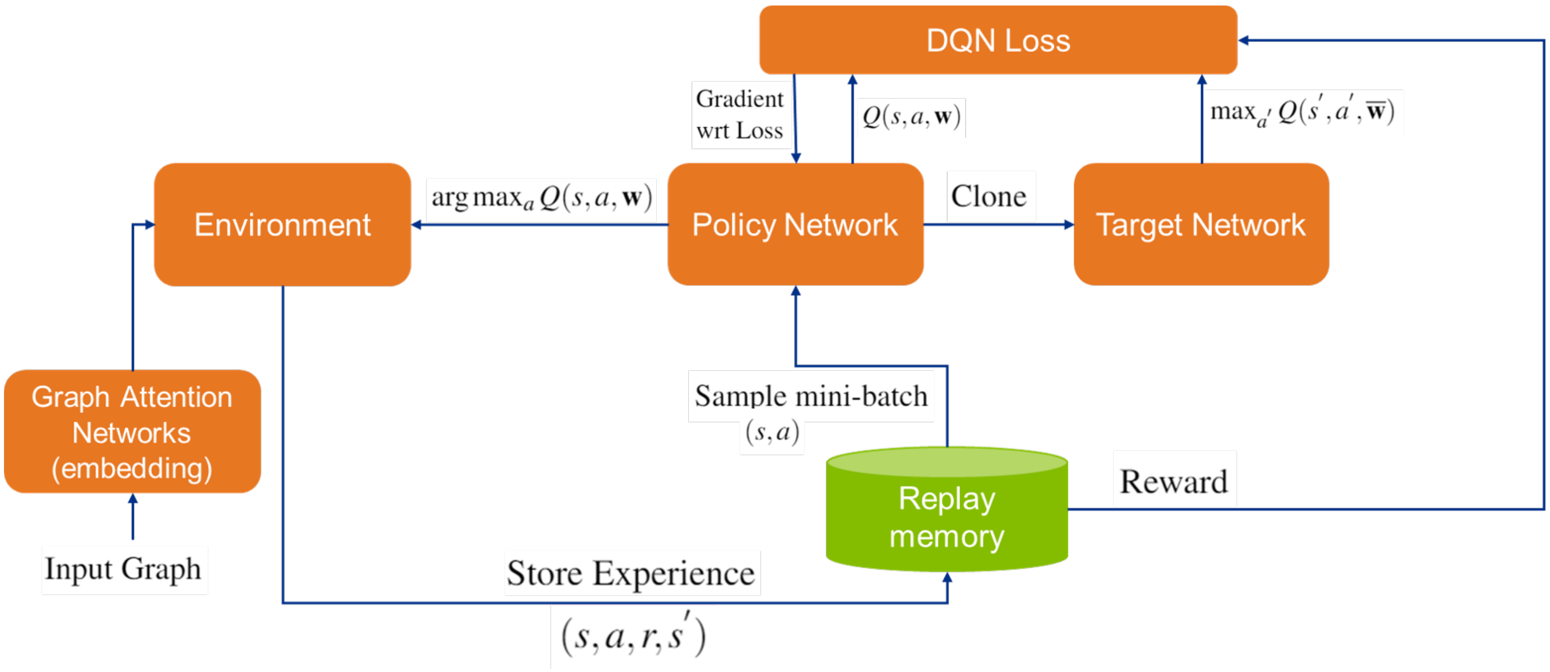
$r_3 \rightarrow$ maximum separation between selected nodes, defined as:

$$\sum_{i,j \in E_s} K\left(1 - \cos\alpha_{ij}\right)^{-1}$$

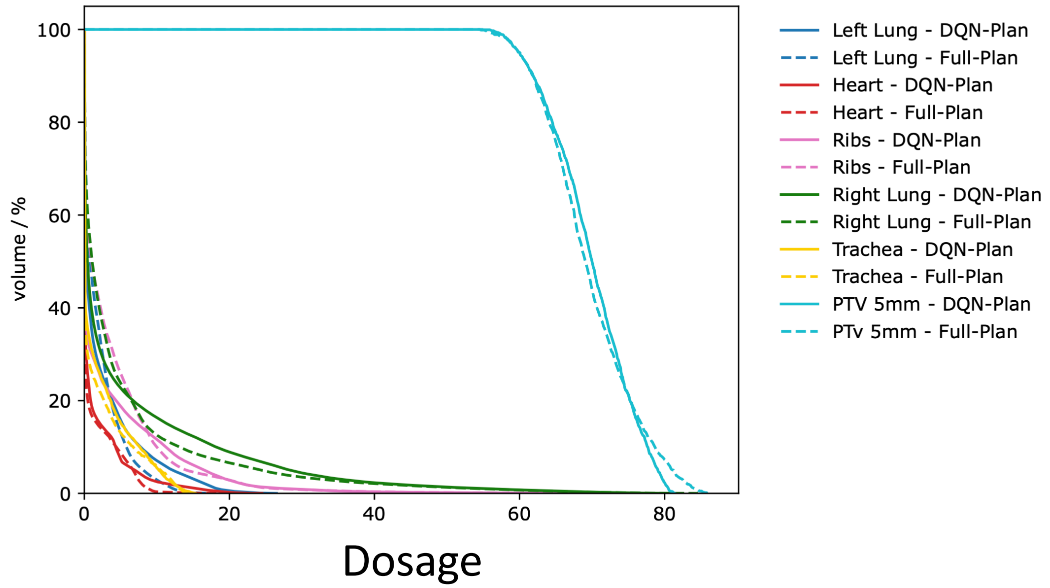
m: the last node added to the trajectory in state s_t

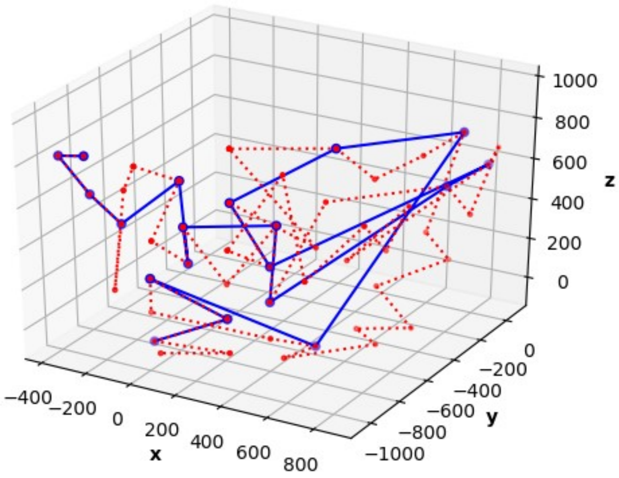
n: next node selected $(a_t = n)$

Reinforcement Learning Framework



Experimental Results (Patient #1)





Method	Obj.	Execution time (s)	Total arm Distance	Time (min)
DQN	3.53	1.09	5,529	35
Gurobi	3.27	3600.00	6,350	37
Heuristic	4.50	2.47	5,494	35
Random	4.65	0.15	2,696	31
Clinical	NA[a]	NA[a]	17,726	54

In conclusion

Pictures generated by DAL-E

- Long waiting times make patients (very) anxious
- Making the system more efficient is thus not only important from a cost perspective