Sparse Flexible Design for Radiation Therapy: A Machine Learning Approach

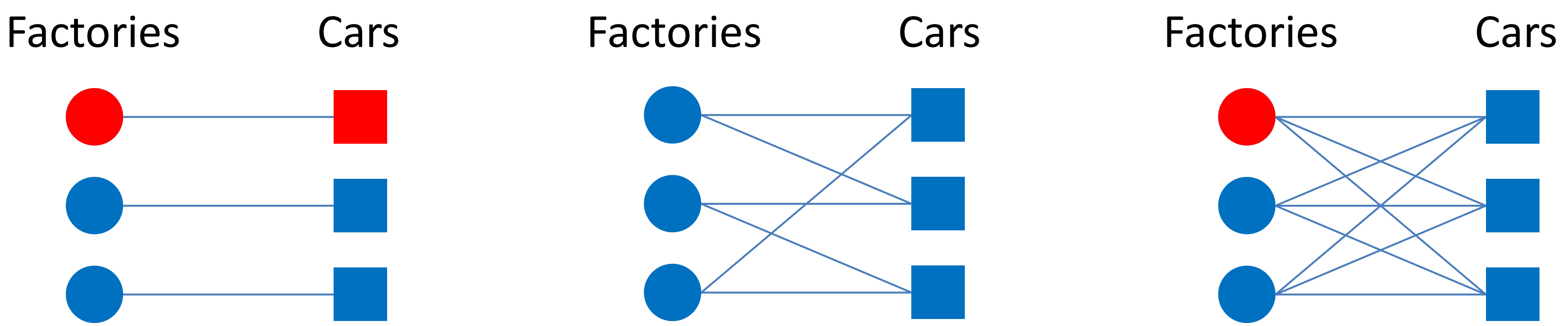
Timothy Chan
University of Toronto

Joint with
Daniel Letourneau, Benjamin Potter

Rotman Healthcare Analytics Roundtable 2020
March 9, 2020

Process flexibility

- Well-known and well-studied, especially within the manufacturing context



No flexibility   Some flexibility   Full flexibility

- Broadly applicable concept
  – Radiation therapy treatment network design

Radiation therapy

Treatment flexibility problem

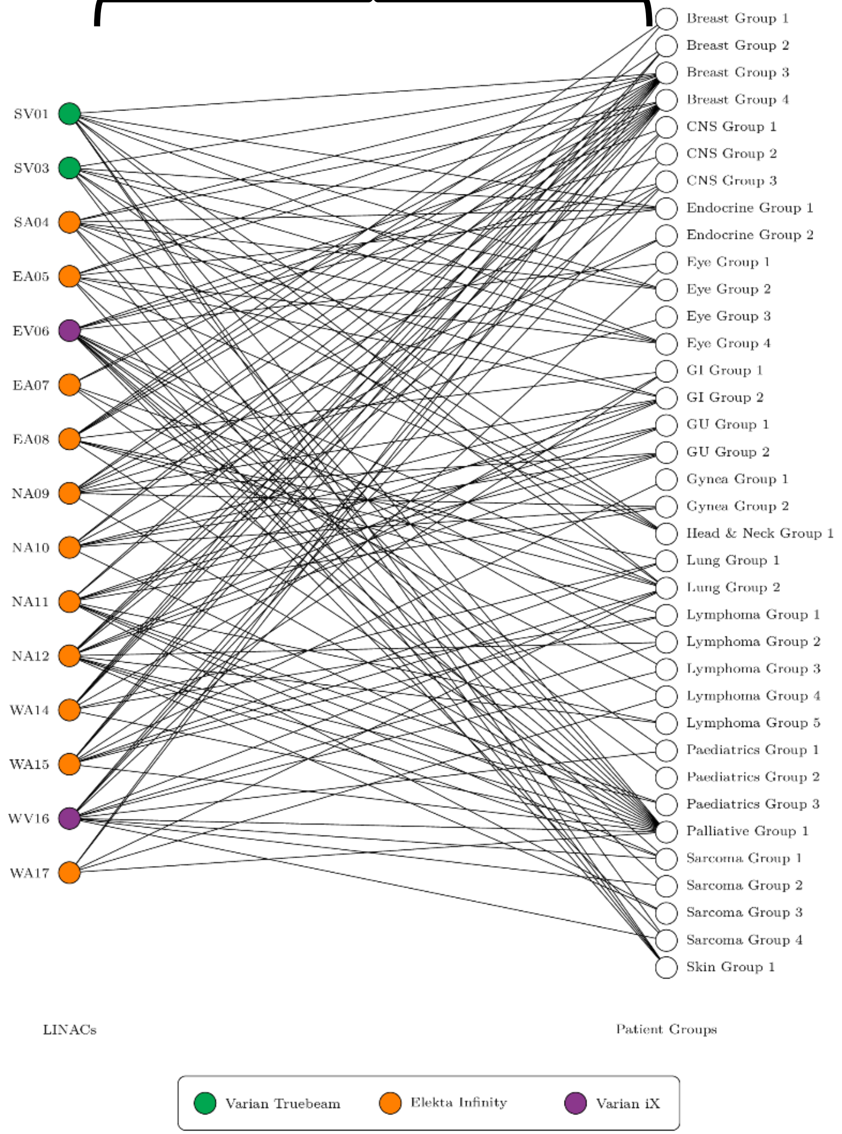
- Linacs are flexible machines and large cancer centers have many of them
  - The good: Ability to deal with supply/demand uncertainty
  - The bad: Device and training costs to being overly flexible
- Operational goal: Minimize overtime while satisfying daily demand
- Complicating factor: machine downtime
- Contribution: a machine learning-based method to design sparse (treatment) networks

The linac network at Princess Margaret

136 arcs

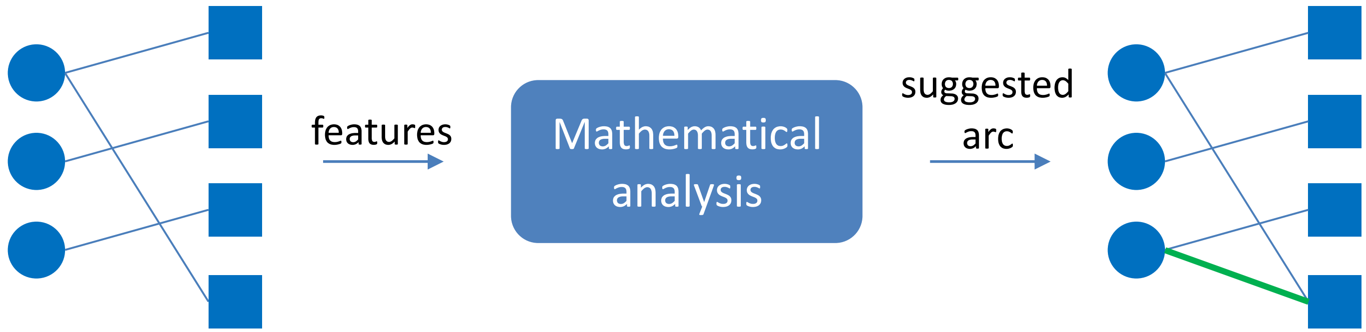
15 linacs

36 patient groups



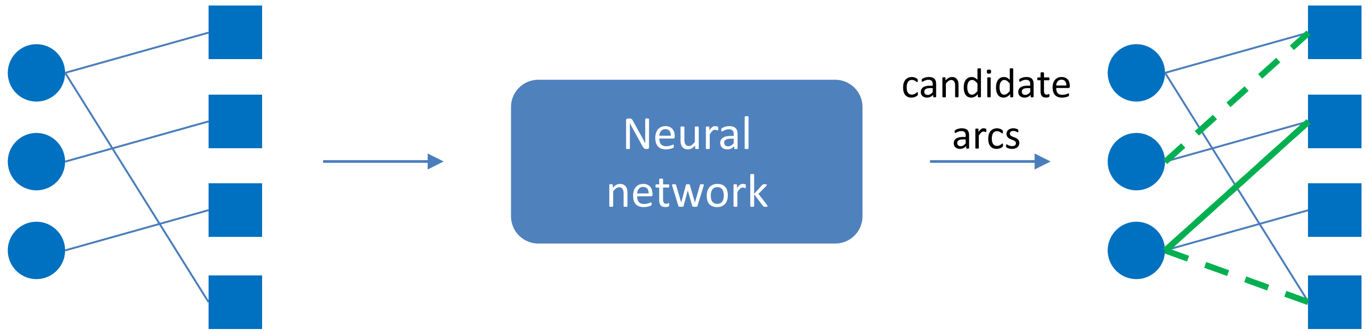
Existing research in process flexibility

- Symmetric/balanced: focus is typically on theory
  - E.g., optimality of the long chain, Simchi-Levi and Wei 2012
- General: focus on heuristics to design sparse networks, guided by deep theoretical insights
  - Chou et al. 2011, Simchi-Levi and Wei 2015, Feng et al. 2017, Yan et al. 2017



Our idea: Part 1

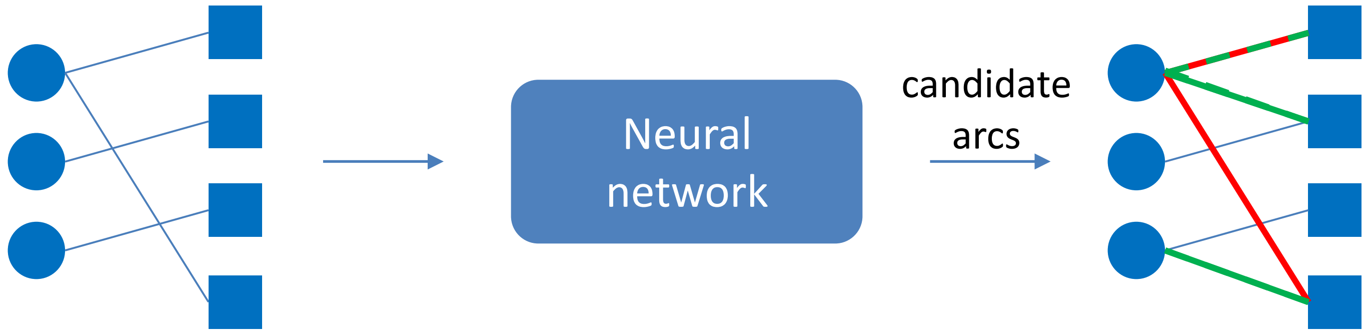
- Replace the deep mathematical analysis with simple machine learning idea
  - Bello et al. 2016, Khalil et al. 2017, Larsen et al. 2018, etc.



- "Predict and search" algorithm (PS)
  - Search size: Full optimization for top candidates
  - Batch size: Number of arcs to add

Our idea: Part 2

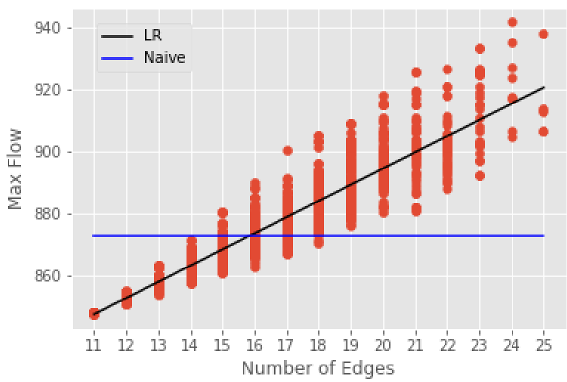
- Prediction enables quick search over huge neighborhood, including backtracking

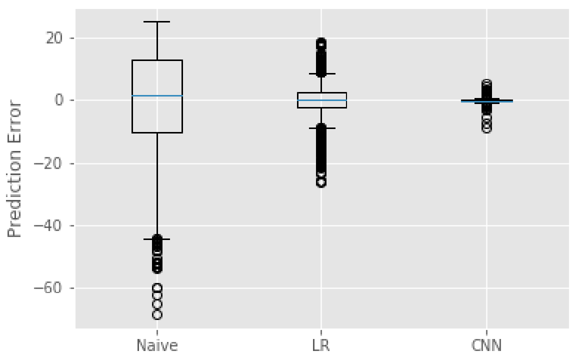


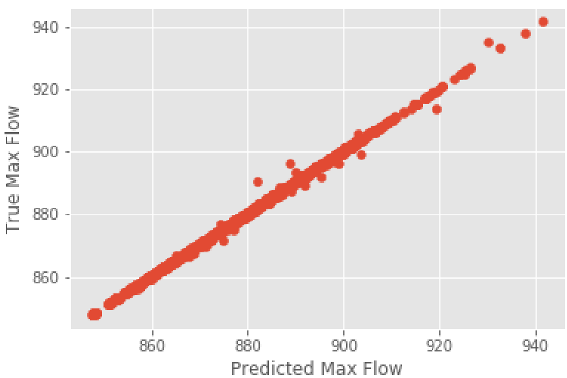
- “Predict and search with revisionist history” (PSRH)
- This is particularly useful if an intermediate arc makes a small, closed chain

Neural network model

- Train using adjacency matrix and max flow value
- Example performance on 10x10 network:
  - 5000 demand realizations from truncated N(100,40)
  - 20,000 flexibility designs with between 11-25 arcs
  - 90/10 training/testing





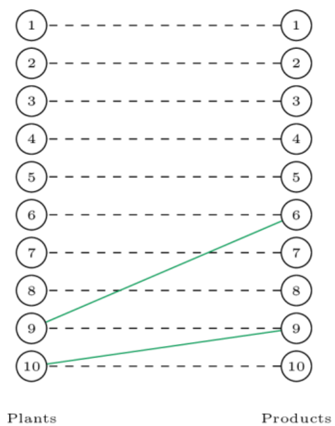


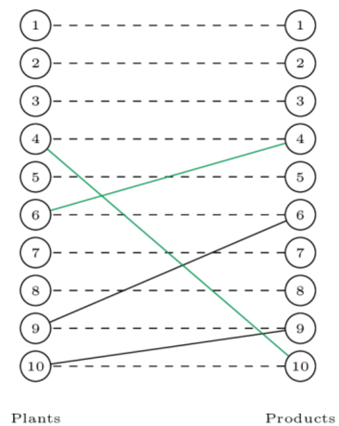
Visualizing algorithm progress

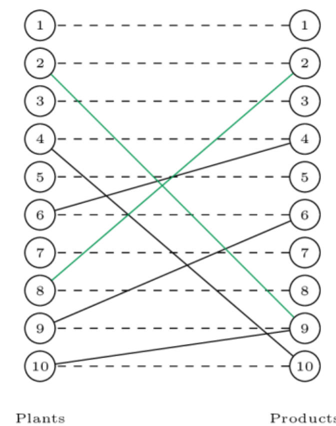
Closed chain

PS

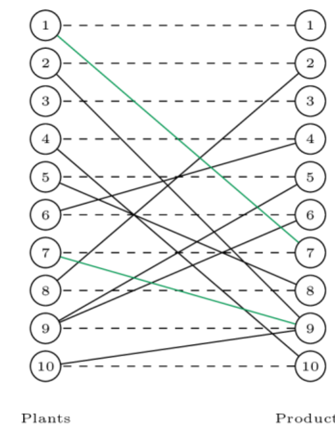
(+2)







…



Iteration 5

Iteration 1

Iteration 2

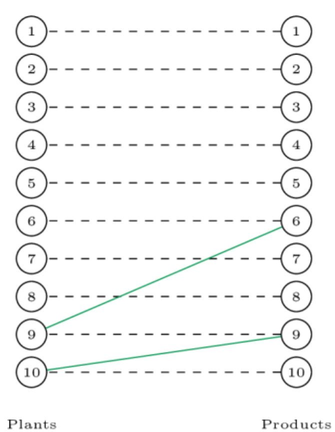
Iteration 3

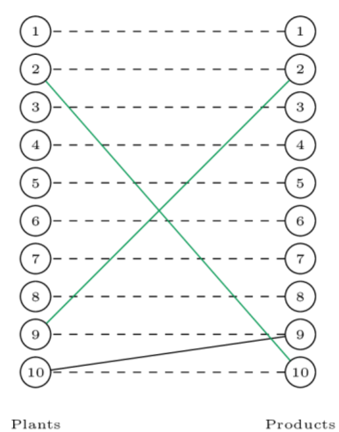
…

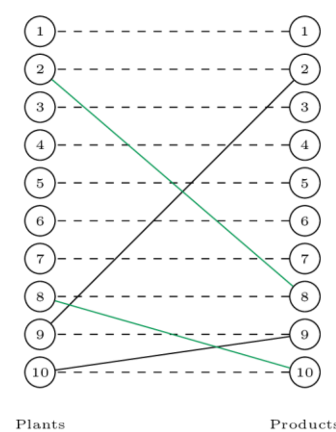
Iteration 10

PSRH

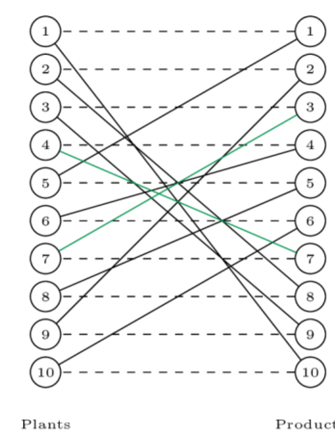
(-1/+2)







…



Closed chain

Breaks/forms new chain!

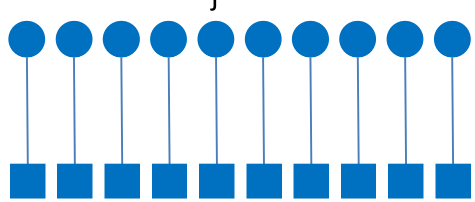
Long chain!

Comparison against existing approaches

- Existing approaches
  - Expander (Chou et al., 2011)
  - UW-PCI/W-PCI (Simchi-Levi and Wei, 2015)
  - MDEP (Feng et al., 2017)
  - DVBH (Yan et al., 2017)
- Three test settings (Simchi-Levi and Wei, 2015)

<u>Test setting 1</u>

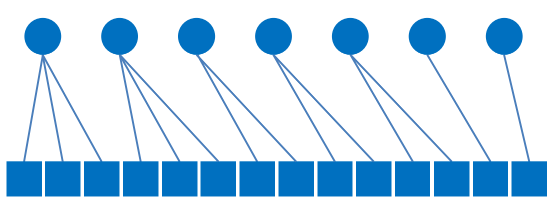
$c_j = 1$



$d_i \sim N(1,0.5)$

<u>Test setting 2 & 3</u>

3 3 2 2 2 1 1



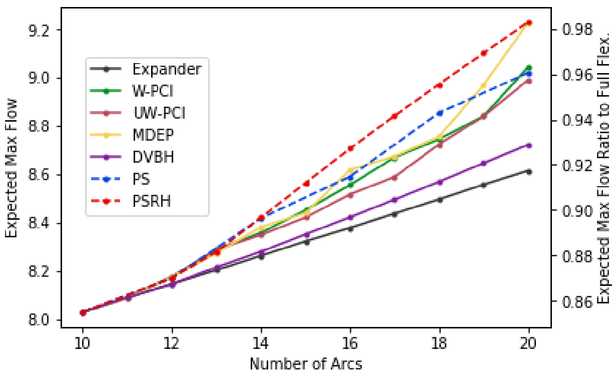
$d_i \sim N(\mu_i,0.5)$

$\mu_i = 1$ or $\sim U(0.5,1.5)$

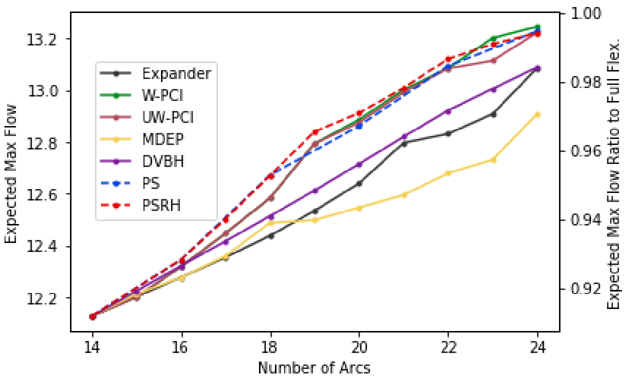
Comparison against existing approaches

- PS and PSRH are consistently competitive with best approaches
- Some existing approaches perform well in certain test settings but poorly in others
- Similar results for worst case, 10th percentile, etc.

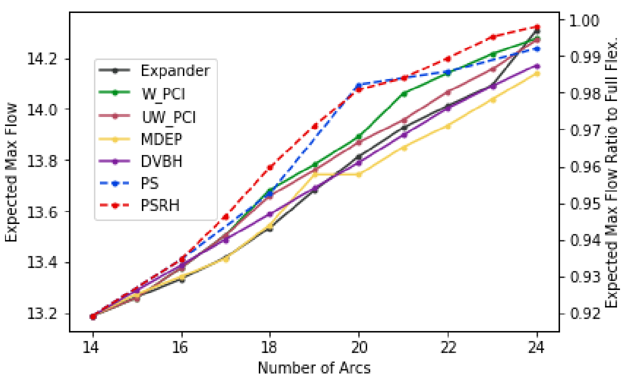
Test setting #1



Test setting #2



Test setting #3



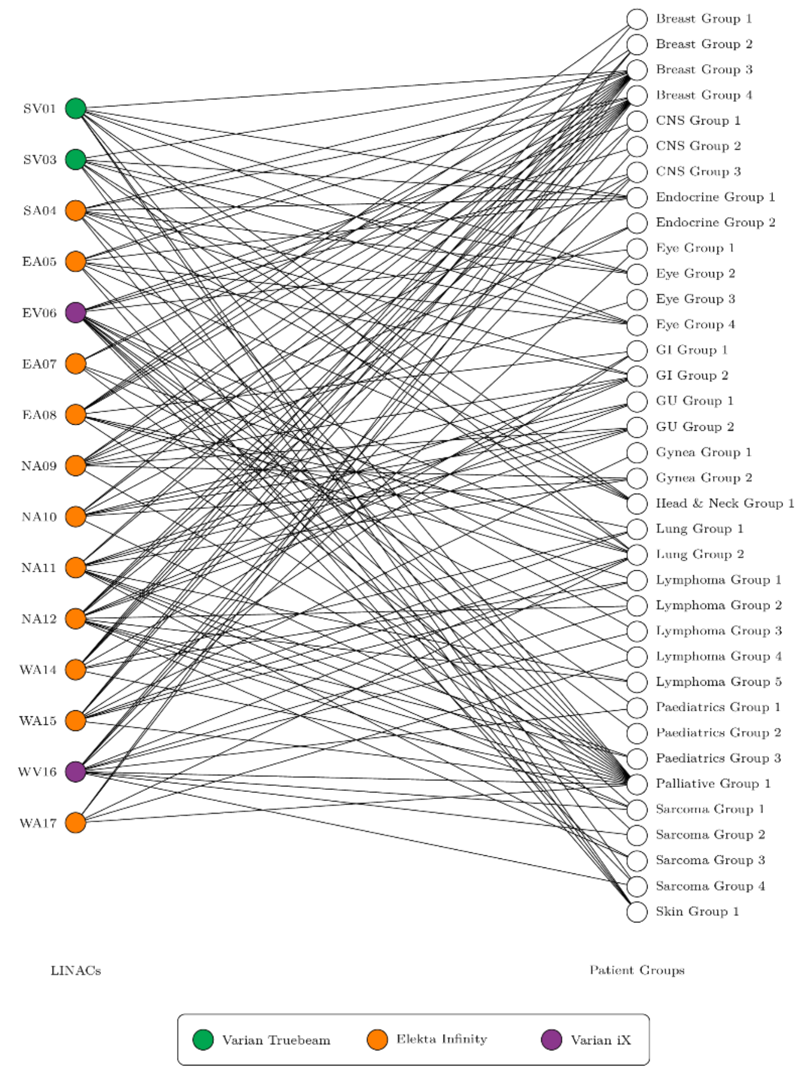
Comparison against existing approaches

Table 3 Flexibility Design Heuristics Performance Comparison

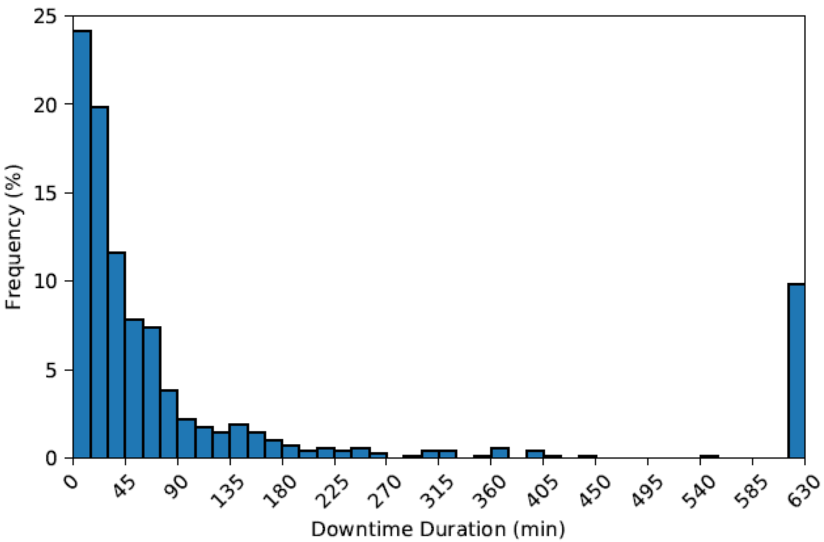
	Network	Test Setting 1 Avg.	Test Setting 1 10^{th} pct.	Test Setting 1 Worst ratio	Test Setting 2 Avg.	Test Setting 2 10^{th} pct.	Test Setting 2 Worst ratio	Test Setting 3 Avg.	Test Setting 3 10^{th} pct.	Test Setting 3 Worst ratio
Baselines	Initial	8.03	6.87	0.66	12.13	10.68	0.71	13.19	11.83	0.75
	Avg Training	8.67	7.44	0.69	12.84	11.31	0.78	13.90	12.38	0.79
	Best Training	9.07	7.87	0.80	13.18	11.65	0.87	14.27	12.74	0.88
	Full Flexibility	9.39	8.06	—	13.30	11.72	—	14.35	12.83	—
Existing Heuristics	Expander	8.61	7.40	0.73	13.08	11.48	0.82	14.31	12.82	0.86
	W-PCI	9.04	7.82	0.74	13.24	11.70	0.83	14.28	12.82	0.85
	UW-PCI	8.99	7.68	0.73	13.22	11.67	0.83	14.27	12.77	0.88
	MDEP	9.23	7.96	0.83	12.91	11.35	0.77	14.14	12.55	0.82
	DVBH	8.72	7.53	0.73	13.09	11.56	0.88	14.17	12.67	0.90
ML-based Heuristics	PS	9.02	7.72	0.77	13.23	11.69	0.82	14.24	12.82	0.87
	PSRH	9.23	7.95	0.79	13.22	11.68	0.79	14.32	12.82	0.89

Back to the linac flexibility problem

- Can we design a sparse network with comparable performance to the existing network?
- What is the value of homogeneous linacs?



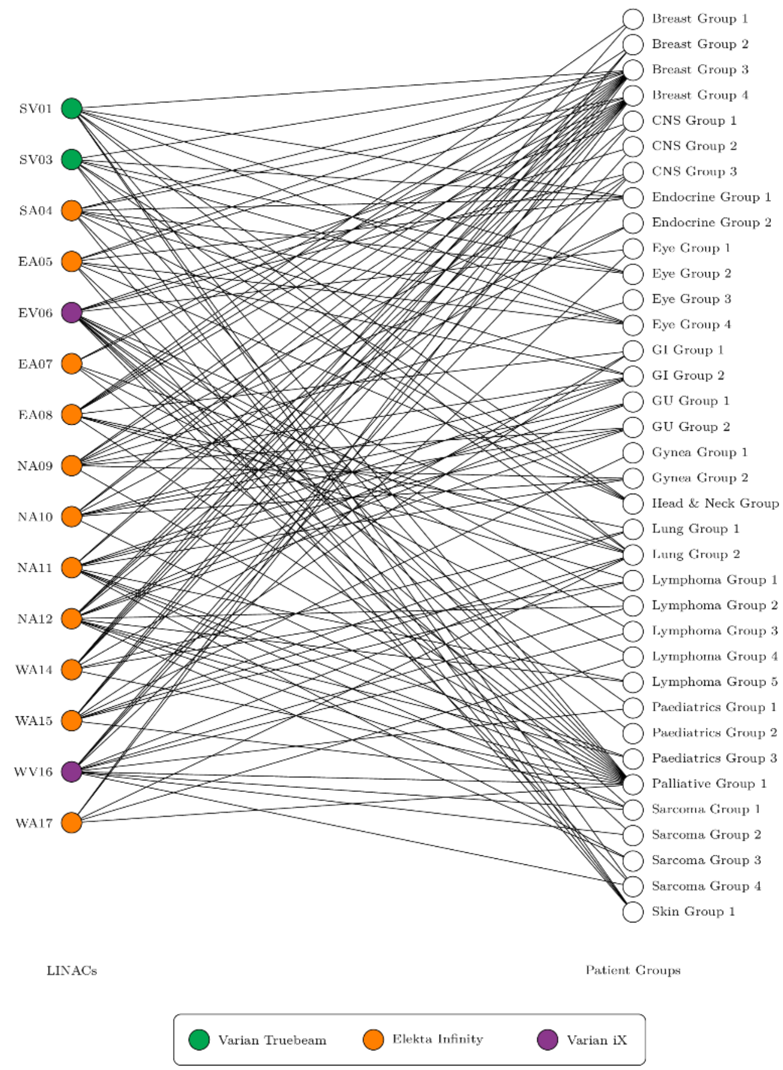
Network parameters



- 15 linacs, 3 models
- 36 patient groups
- 136 arcs
- Capacity: 8AM – 630PM
  - Downtime data 2015-2017
  - 70% of days at least one machine experienced downtime
    - MLC malfunction, software frozen, etc.
  - Each linac has ~9% chance of downtime on a given day
- Demand:
  - Curative treatment data from 2015-2016
  - Palliative demand assumed to be ~31% of total tx time

Specializing the method to radiation therapy

- Two constraints added to standard network flow model
  - Reshuffling: upper bound on total demand that can be moved to another linac following a breakdown
  - Linac heterogeneity: patients need to be treated on same type of linac if moved
- Same neural network-based heuristics
  - 95% of NN prediction errors within 5%



Designing a sparse treatment network

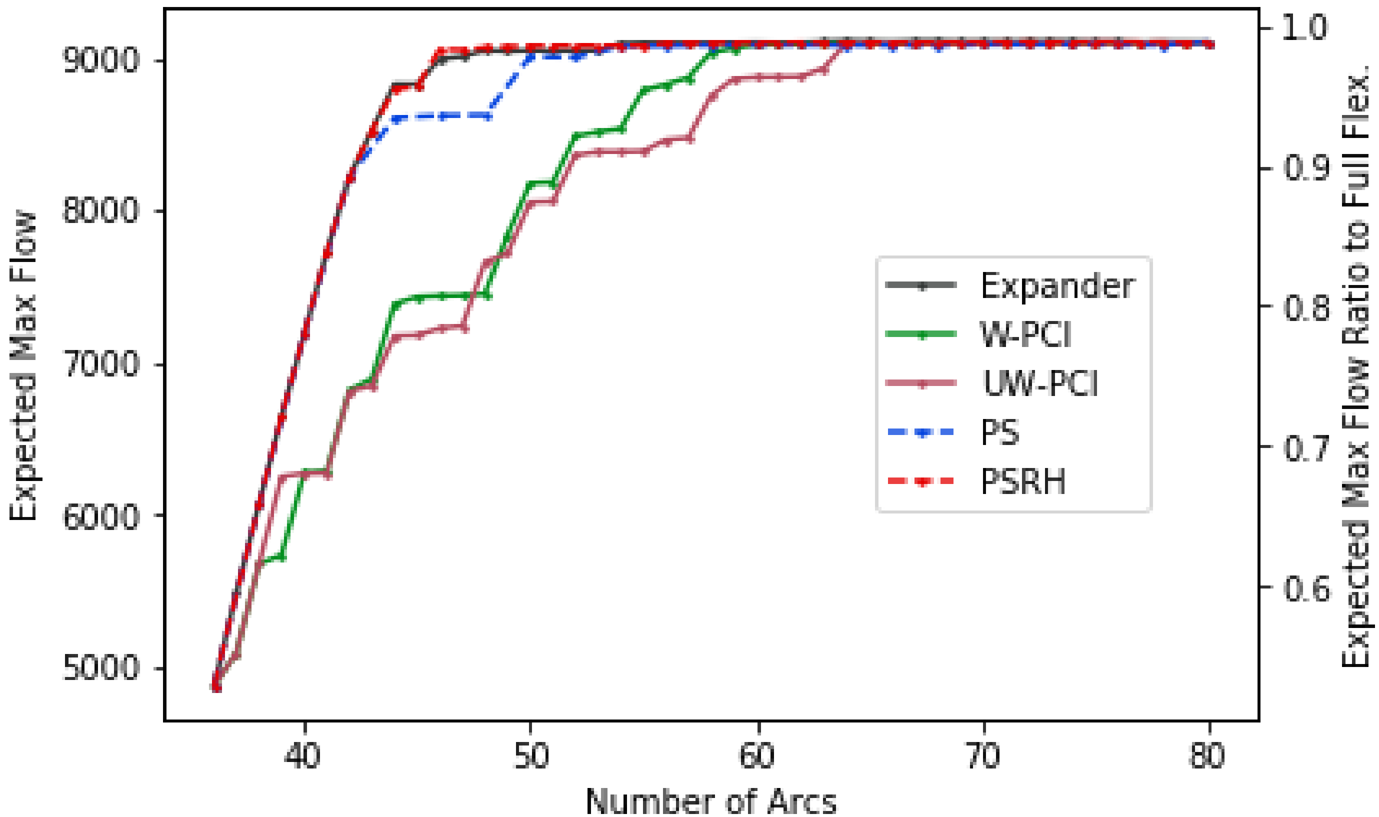


Table 5 **Comparison of Network Designs with Heterogenous LINACs.**

Network	Arcs	Expected Max Flow	10^{th} Percentile	Worst Ratio (%)
PSRH	46	9058.36 (9.68)	8662.52 (45.32)	96.66 (0.07)
PSRH	56	9100.22 (17.85)	8706.25 (73.63)	98.17 (0.81)
PSRH	80	9103.69 (16.51)	8734.09 (81.14)	98.08 (0.98)
Existing PM	136	9105.96 (17.64)	8748.79 (84.53)	98.91 (1.36)
Fully Flexible	540	9110.47 (17.74)	8755.09 (82.69)	—

Homogeneous linacs

- Homogeneous linacs can reduce average overtime by 27% and variability in overtime by order of magnitude

Table 5 **Comparison of Network Designs with Heterogenous LINACs.**

Network	Arcs	Expected Max Flow	10^{th} Percentile	Worst Ratio (%)
PSRH	46	9058.36 (9.68)	8662.52 (45.32)	96.66 (0.07)
PSRH	56	9100.22 (17.85)	8706.25 (73.63)	98.17 (0.81)
PSRH	80	9103.69 (16.51)	8734.09 (81.14)	98.08 (0.98)
Existing PM	136	9105.96 (17.64)	8748.79 (84.53)	98.91 (1.36)
Fully Flexible	540	9110.47 (17.74)	8755.09 (82.69)	—

Table 7 **Comparison of Network Designs with Homogenous LINACs**

Network	Arcs	Expected Max Flow	10^{th} Percentile	Worst Ratio (%)
PSRH	46	9055.28 (0.06)	8593.12 (0.00)	96.65 (0.09)
PSRH	56	9120.72 (4.25)	8711.51 (36.92)	97.02 (0.12)
PSRH	80	9131.37 (1.76)	8790.01 (8.92)	96.94 (0.15)
Existing PM	136	9130.61 (2.10)	8786.54 (10.32)	96.94 (0.10)
Fully Flexible	540	9139.09 (0.31)	8820.00 (0.00)	—

Other insights

- Initial schedule matters a lot, due to presence of reshuffling constraint
  - Can get up to 40% of value of homogeneous linacs just by adjusting initial schedule
  - Suggests that two-stage approach is important

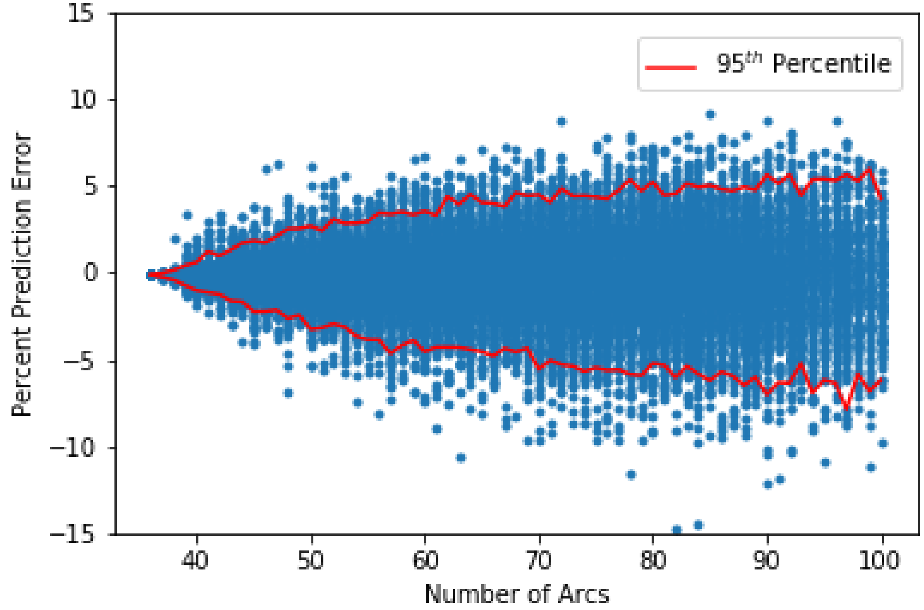
Summary

- Process flexibility is a useful lens through which to view many different problems
- Developed novel machine learning-based heuristic to design sparse treatment network
  - Likely to be most useful for large problems where decisions need to be made often, perhaps in real time

Thanks for listening!

Questions?

Timothy Chan
University of Toronto
tcychan@mie.utoronto.ca


95th Percentile
Percent Prediction Error
Number of Arcs
15
10
5
0
−5
−10
−15
40
50
60
70
80
90
100