Chair of Communication Networks
Department of Electrical and Computer Engineering
Technical University of Munich

o‘zapft is: Tap Your Networking Algorithm‘s Big Data!

Andreas Blenk

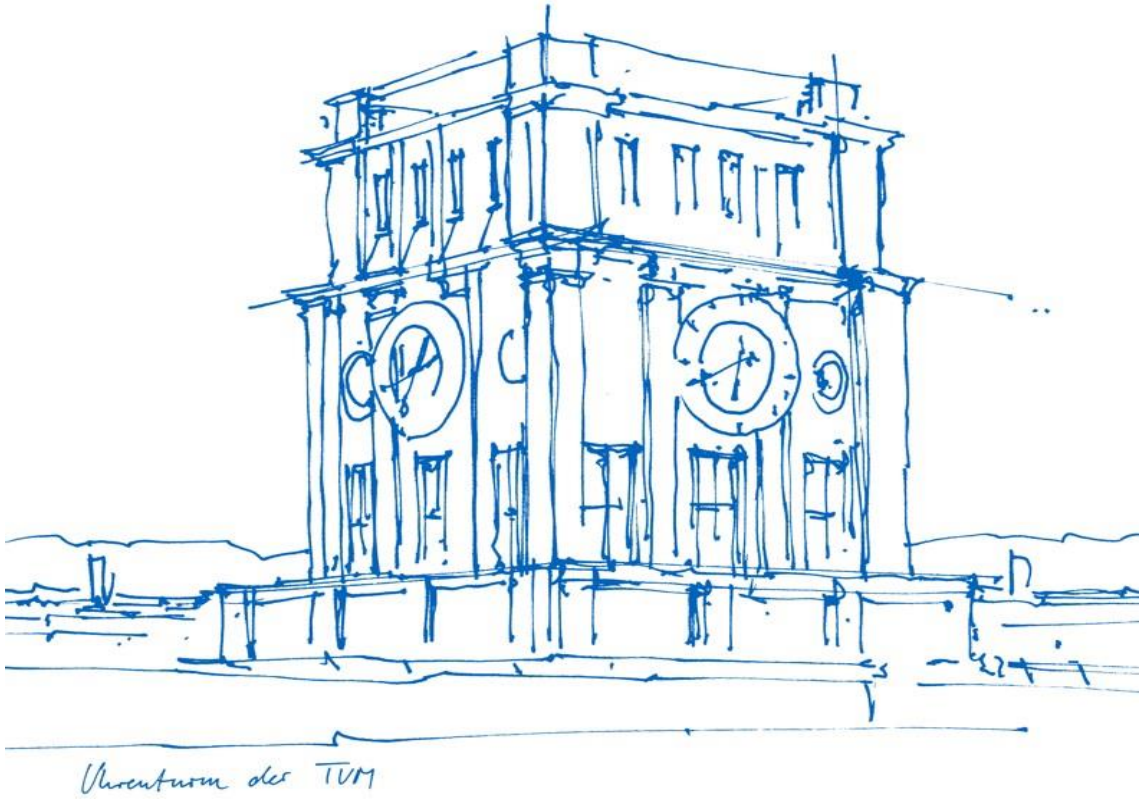
Patrick Kalmbach

Patrick.Kalmbach@tum.de

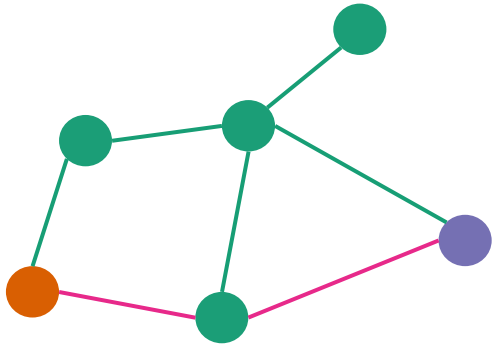
Wolfgang Kellerer

Stefan Schmid

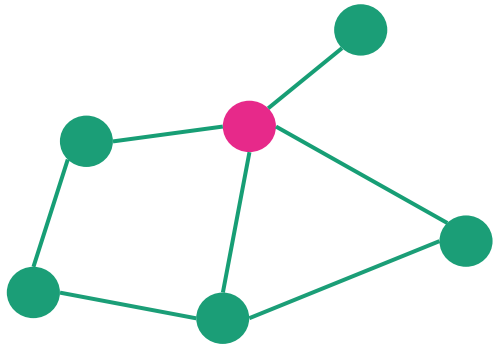
schmiste@cs.aau.dk



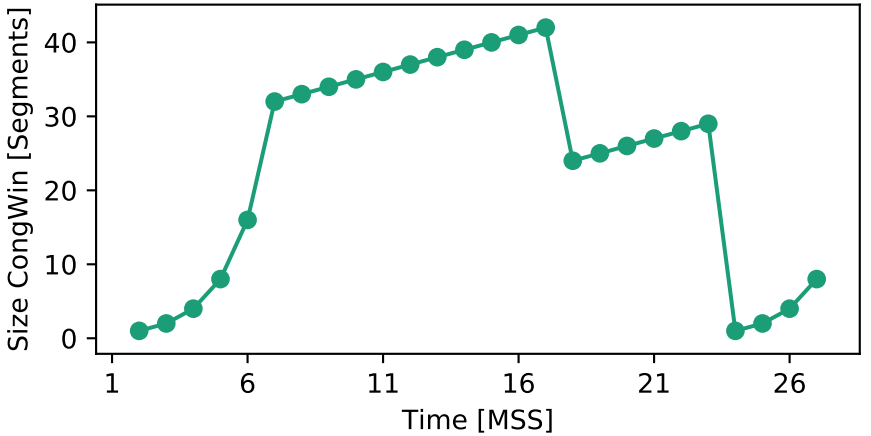
Network Algorithms are Ubiquitious



Routing



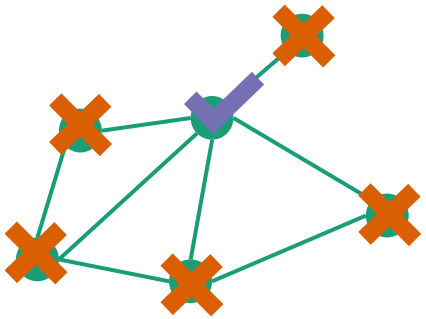
Cache/Content
Placement



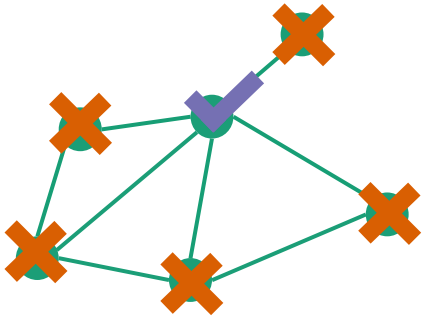
Congestion Control

Two classes of optimization problems: **Packing** and **Covering** problems

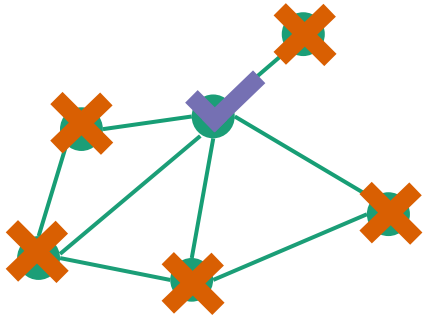
The Limitation – Fire and Forget



Place Cache



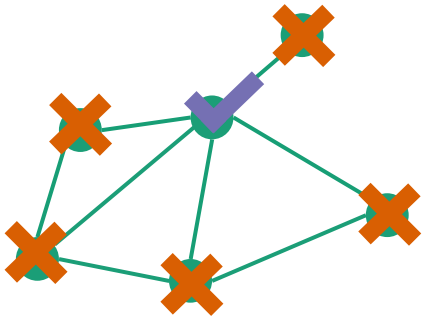
Place Cache



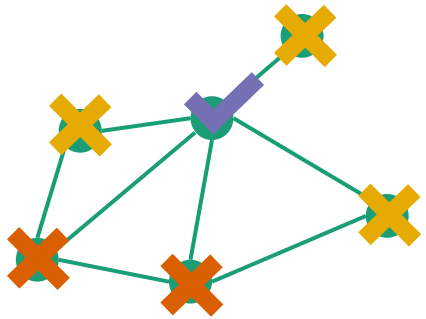
Place Cache

Algorithms repeatedly solve similar problems **from scratch.** This is not only boring for the algorithm but also a waste of information and resources

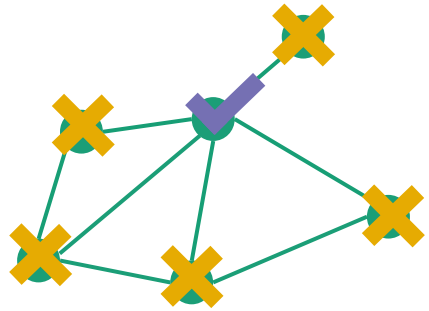
The Opportunity – Tap into your Algorithm's Big Data



Place Cache

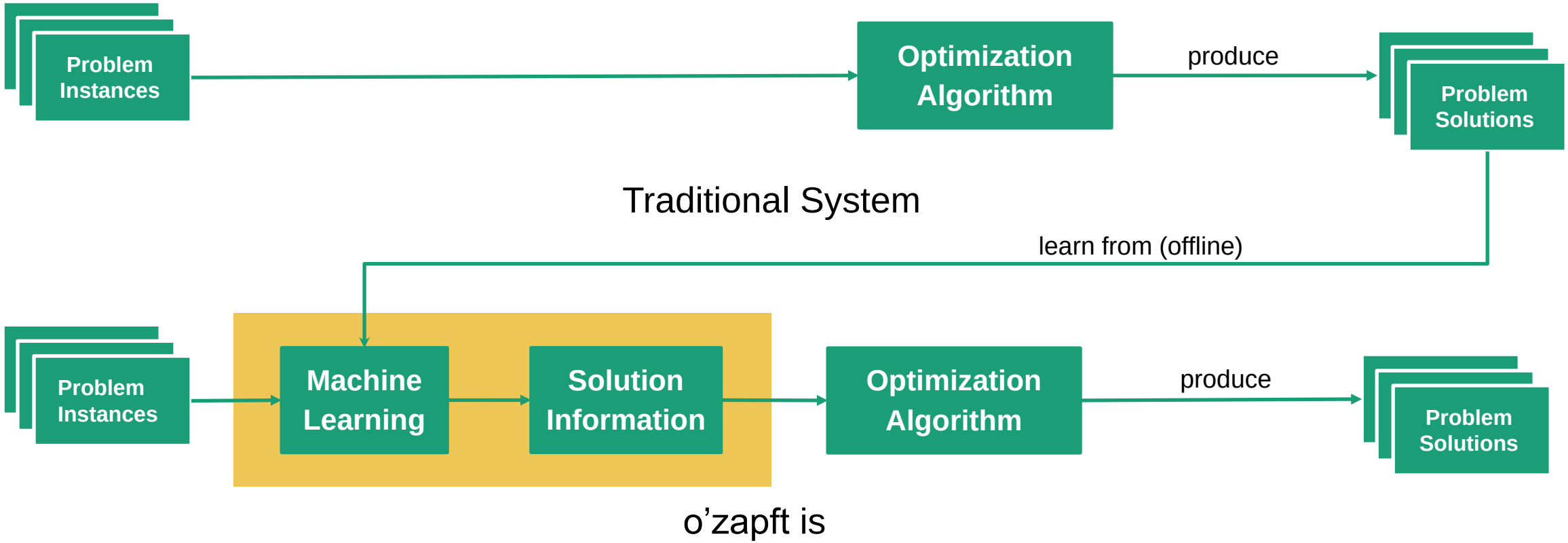


Place Cache



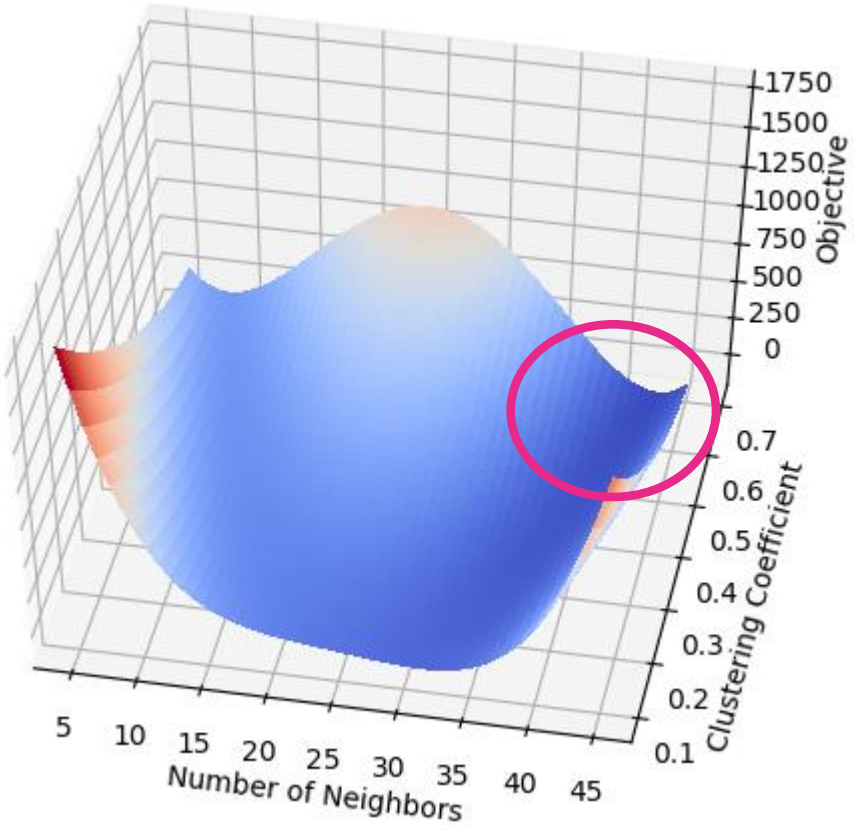
Place Cache

Traditional vs. Proposed System

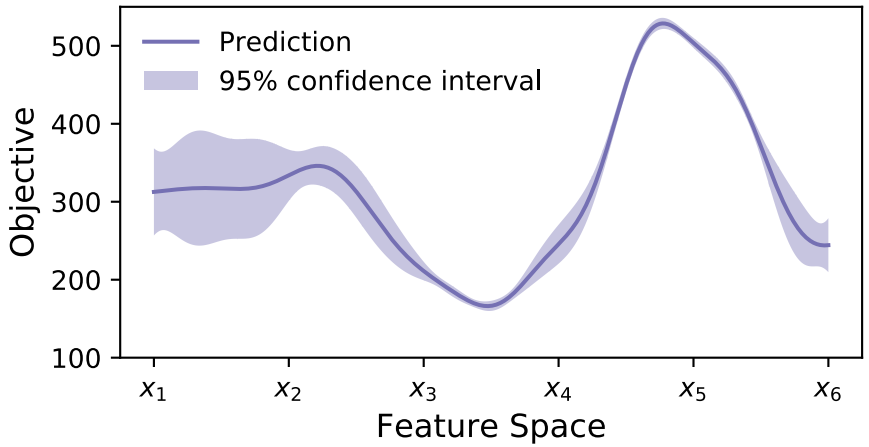


Data Available at: Patrick Kalmbach, Johannes Zerwas, Michael Manhart, Andreas Blenk, Stefan Schmid, and Wolfgang Kellerer. 2017. Data on "o'zapft is: Tap Your Network Algorithm's Big Data!". (2017). https://doi.org/10.14459/2017md1361589

Potentials



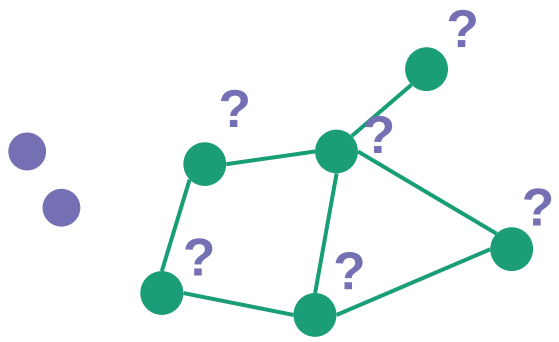
Search Space Reduction reduction/Initial Solutions



Predict Value of Objective Function

Facility Location Use-Case

Facility Location (Controller Placement) – Guess Initial solutions

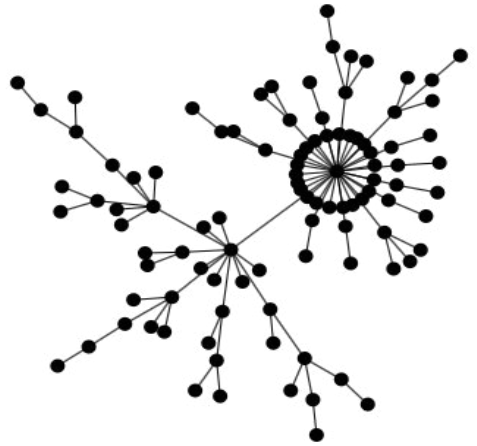


Problem: Given a network and a number of controllers, where to place the controllers?

Data Generation – Facility Location

Optimization Algorithms:

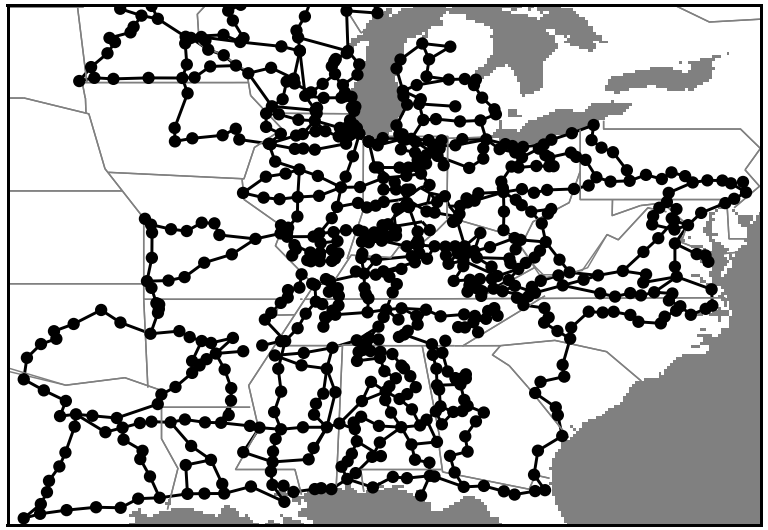
- Random Placement (Rnd)
- Mixed Integer Program (MIP)
- Greedy (Gdy)



Barabasi-Albert Graph [3]

Substrates

- Barabasi-Albert (BA) [2] 40 nodes
- Whole Topology Zoo (TZ) [1] 20-800 nodes



Kentucky Data Link

Objective:
Minimize maximum latency between node and controller

[1] Knight et al., The Internet Topology Zoo. *IEEE J. on Sel. Areas in Communica-tions 29*, 9 (2011).
[2] Saino et al., A Toolchain for Simplifying Network Simulation Setup, in *Procs. SIMUTOOLS '13*, Cannes, France, March 2013
[3] Picture taken from http://graphstream-project.org/media/img/generator_overview_barabasi_albert.png

Learning to Place

Library:

- Sci-Kit Learn [1]

Features:

- Node degree
- Closeness
- Betweeness
- Spectral Features

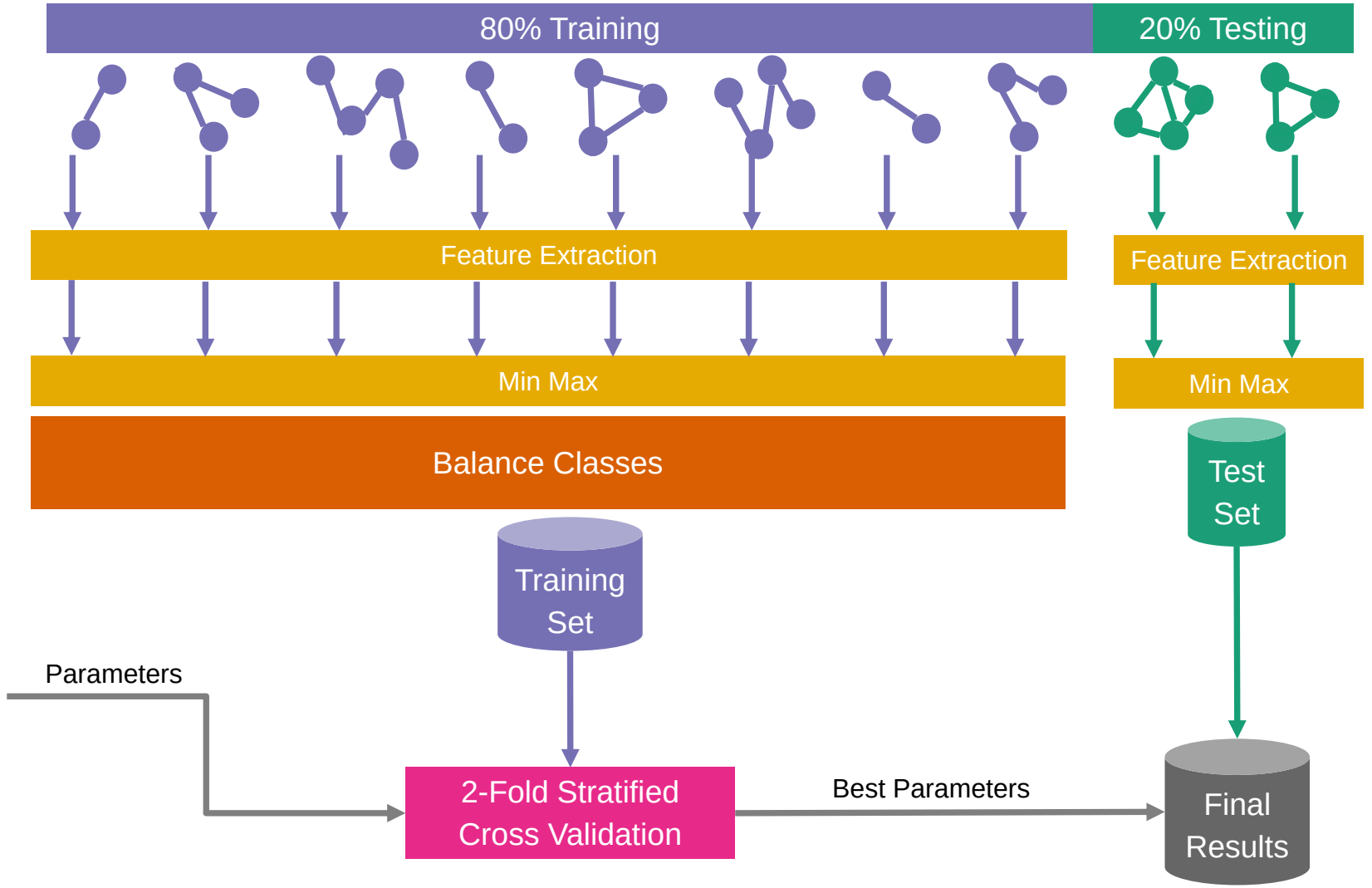
Measures:

- F1 Score

Classifier:

- Logistic Regression
- Support Vector Machine
- Random Forest
- **Extra Tree**
- AdaBoost

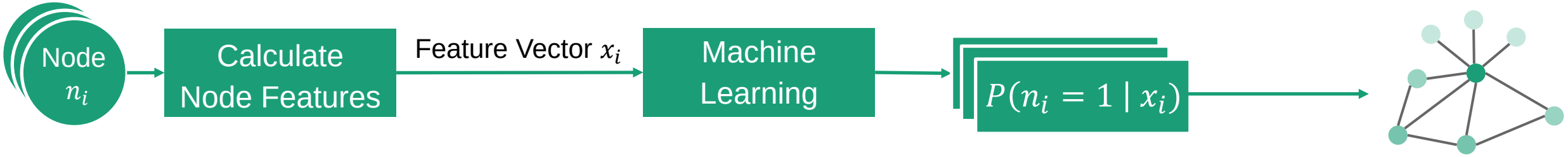
Model Training and Selection:



[1] Scikit-learn: Machine Learning in Python, Pedregosa et al., JMLR 12, pp. 2825-2830, 2011.

Use Case: Facility Location

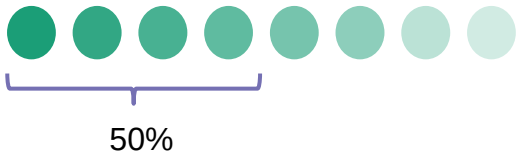
Problem: Place Facility on network node such that maximum latency is minimized

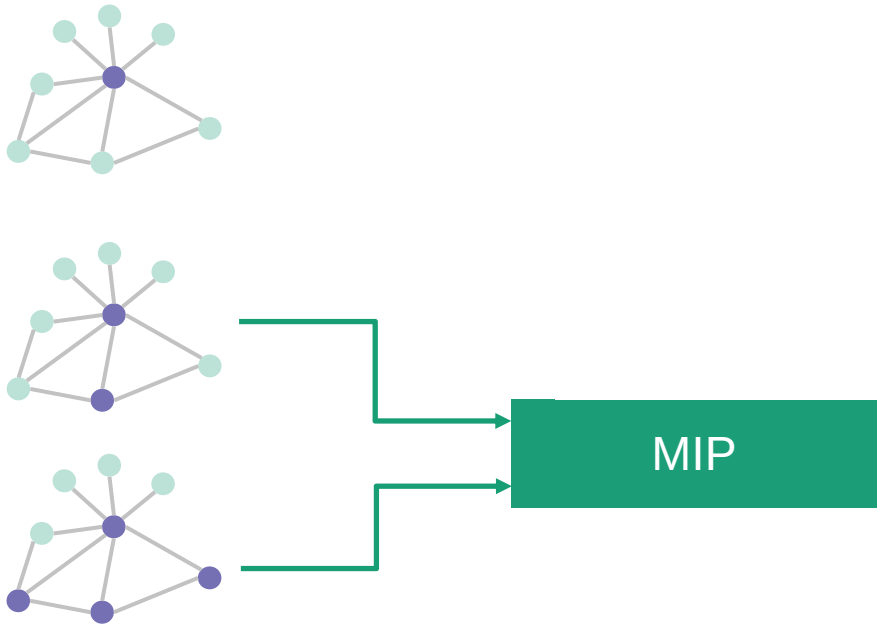


o'zapft is-Exact $argmax_{n_i}\{P(n_i = 1 \mid x_i)\}$

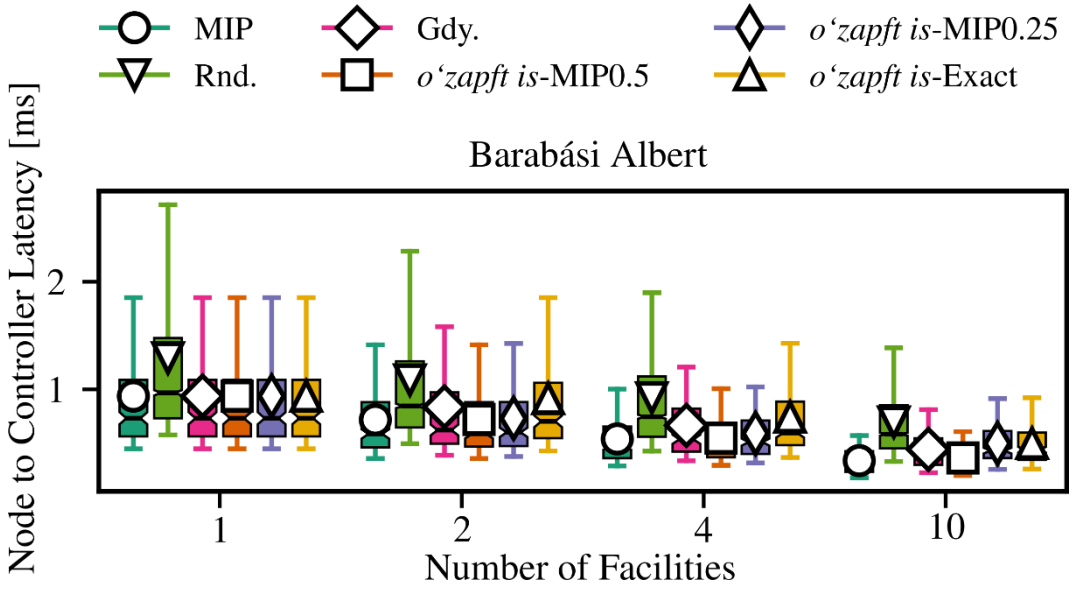
*o'zapft is-MIP*0.25

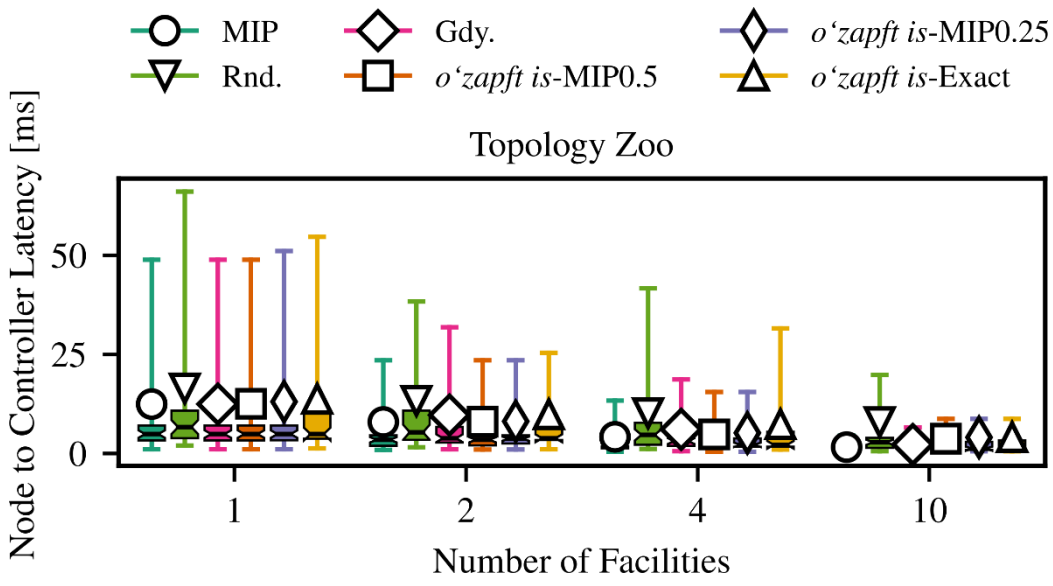
*o'zapft is-MIP*0.5





Use Case: Facility Location

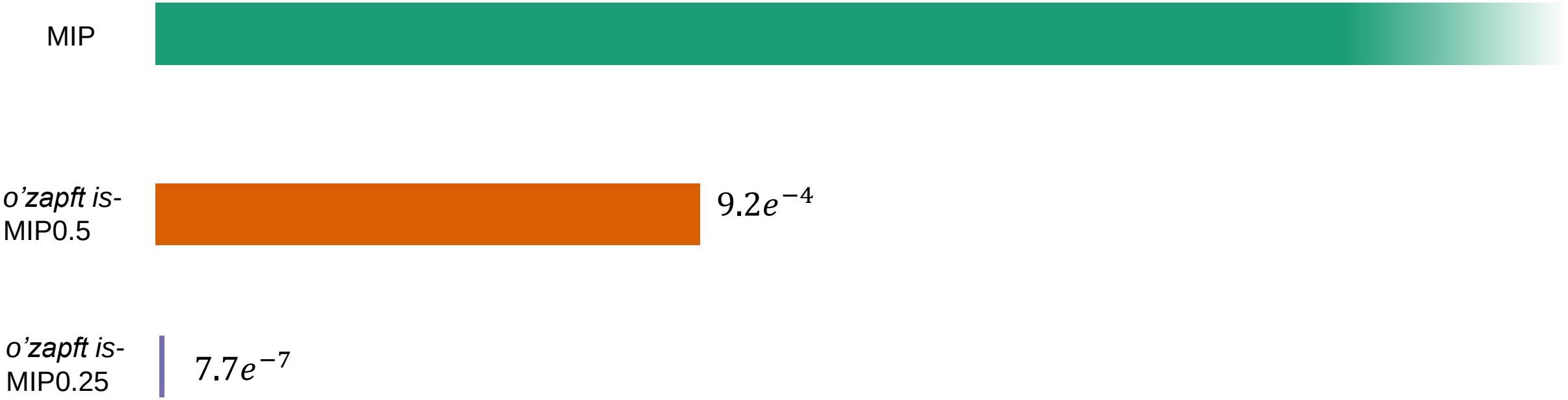




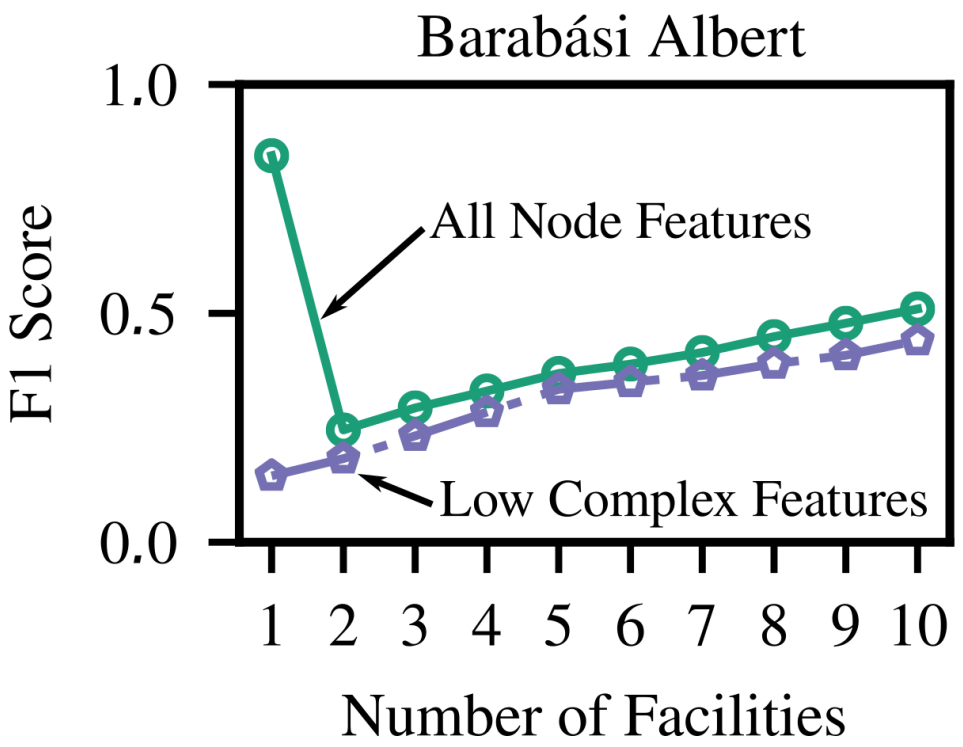
Large reduction of solution space with only small performance degradation

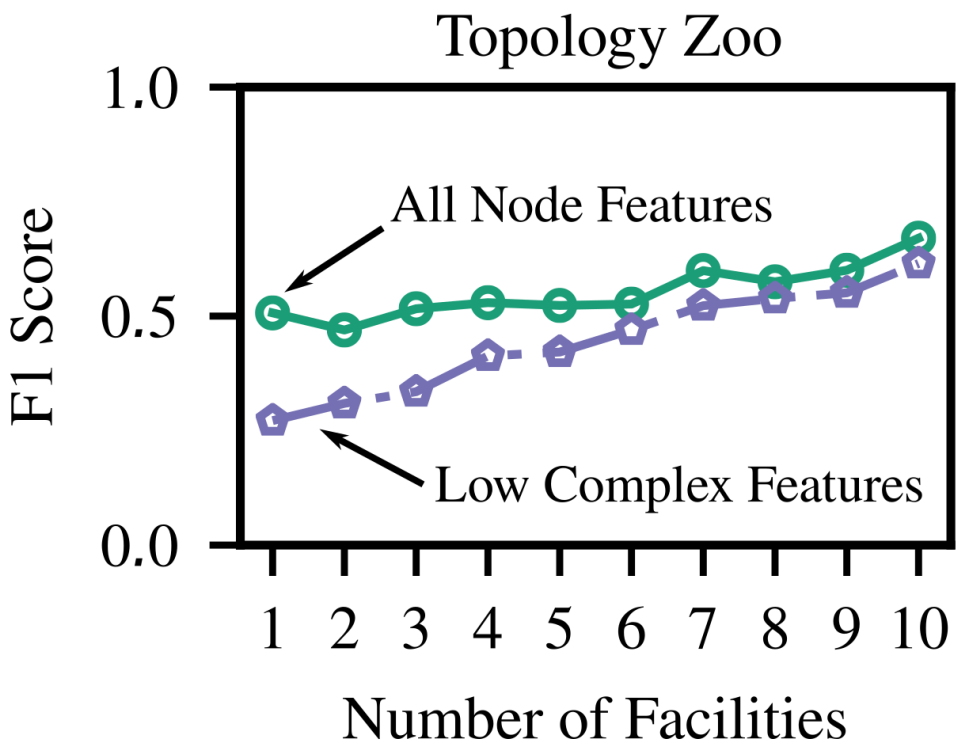
Search Space Reduction

Substrate	Kentucky Data Link
Number of Nodes	734
Number Facilities	10



Impact of Complex Features

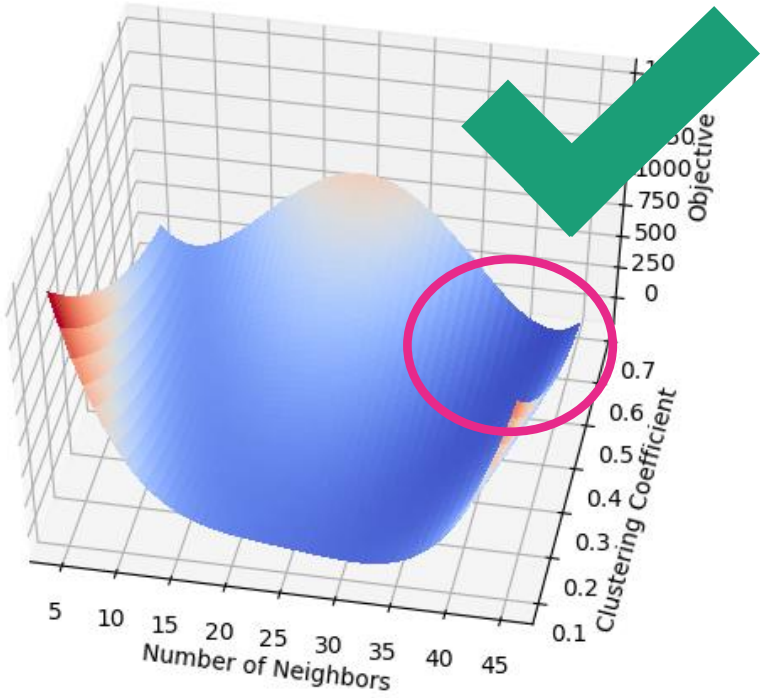




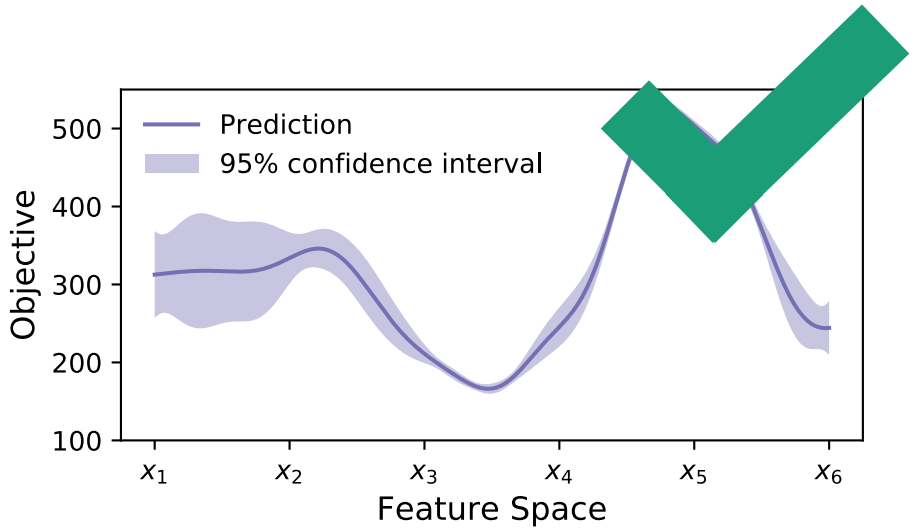
Good classification performance for features with low complexity

Conclusion

It is possible to predict the behavior of networking algorithms from past problem-solution-pairs



Search Space Reduction reduction/Initial Solutions



Predict Value of Objective Function

Future Work

- Transfer Learning
- Investigate the size of the minimal search space
- Investigate whether heuristics improve on reduced search space
- Investigating the applicability of Deep Learning