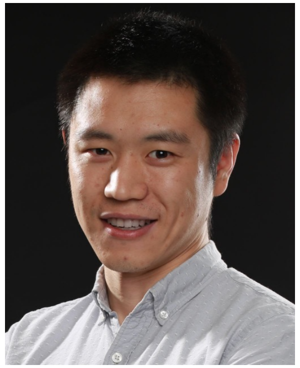
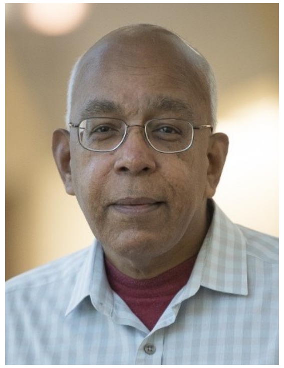
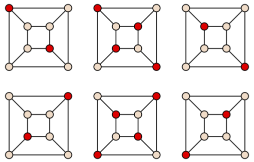
ICLR 2023

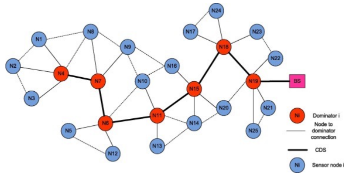
Paper ID: 4014

Graph-based Deterministic Policy Gradient for Repetitive Combinatorial Optimization Problems





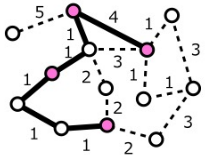




Zhongyuan Zhao*, *Ananthram Swami* †, *Santiago Segarra**

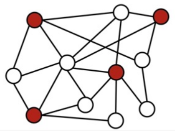
*Rice University, USA

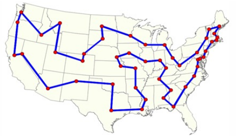
† US Army's DEVCOM Army Research Laboratory, USA



The 11th International Conference on Learning Representations (ICLR)

May 1-5, 2023





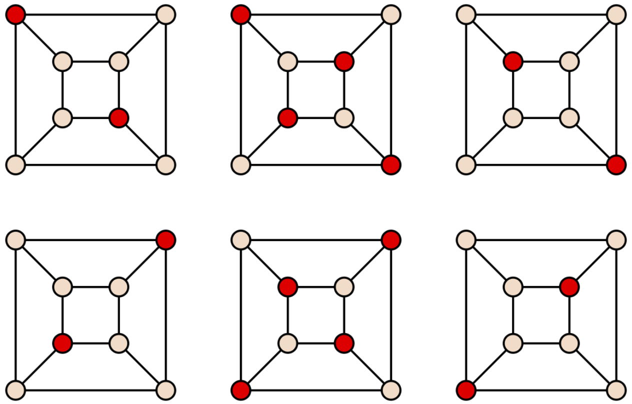
Combinatorial Optimization Problems (COP)

- Characters
  - Discrete (Integer) constraints
  - Relational constraints
  - Minimize total cost (maximum total utility)
  - Non-convex, often *NP-hard*!
- Example: Maximum Weighted Independent Set (MWIS)
  - Independent set: a set of disconnected vertices
  - Maximum total weight

$$\mathbf{x}^* = \min_{\mathbf{x}} \ \mathbf{c}^\top \mathbf{x}$$ COP

$s.t.$ *Discrete constraint on nodes or edges*

Constraints defined on Graph, Hypergraph, or Simplicial Complex



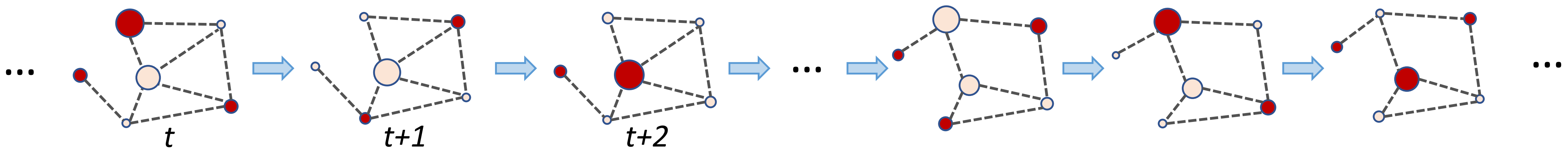
Source: Wikipedia – Maximal independent set

Many practical COPs are repetitive!

Nodes *Edges* *Cost vector*

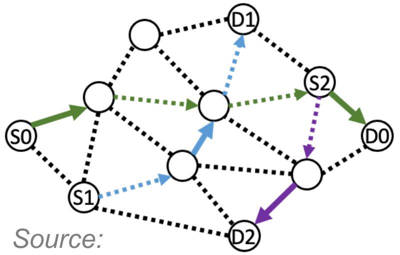
weighted graph $(\mathcal{V}(t), \mathcal{E}(t), \mathbf{c}(t))$

Graph-based
Markov decision process (MDP)



Applications

Routing & Scheduling in **communication networks**



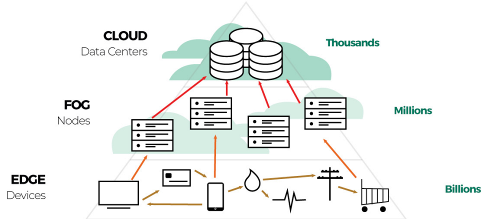
Source: zhongyuanzhao.com

Multi-object tracking in **computer vision**

Vehicle routing problems in **distribution networks**

Source: caliper.com

Resource allocation & job scheduling in **cloud, frog, edge computing**

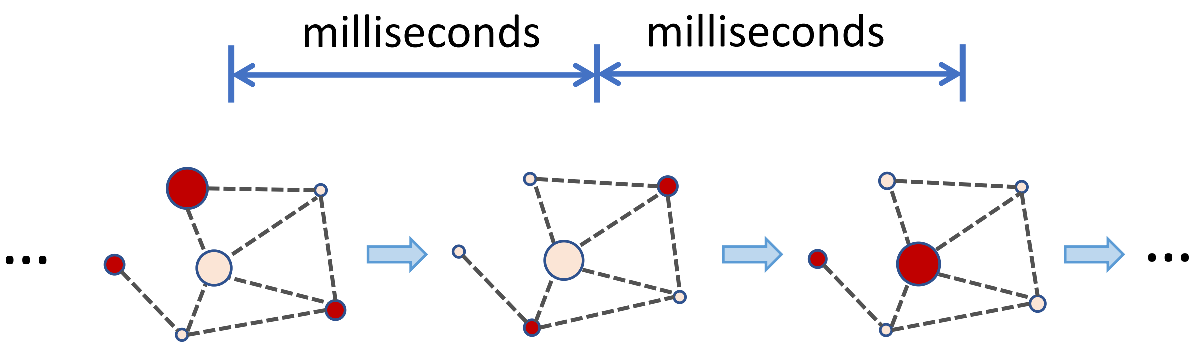


Source: e-ziggurat.com

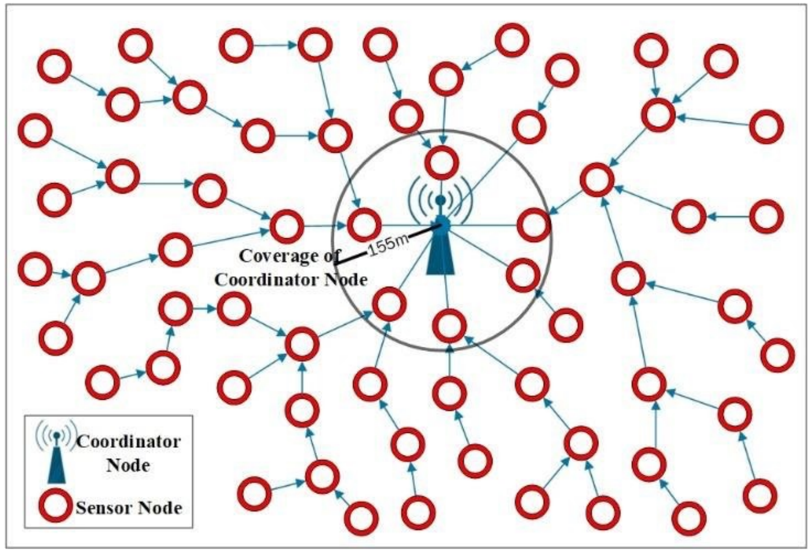
Characters & Challenges

1. Network state of *t+1* depends on the decisions at *t*
2. Cost vector ***c*** changes rapidly compared to network topology
3. Dynamic network topology
4. Practical restrictions: limited runtime and/or distributed execution

Practical restrictions: runtime & distributed execution



e.g., COP instances coming at data or video frame rates in wireless link scheduling or computer vision



Source: (D. Arı , M. Çıbuk and F. Ağgün , 2017)

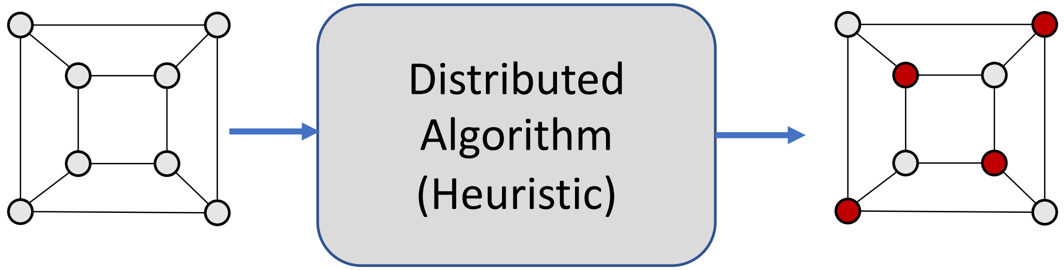
- Centralized COP solver
  - High communication overhead → Network state changes before being collected to a server
  - High computational complexity → Scales up quickly by network size
  - Single-point-of-failure
- Distributed COP solver → only needs neighborhood information, fast, robust

Why don't just use a GNN?

- Graph neural networks (GNNs)
  - Distributed execution, fast
  - Generalize to different topologies
  - bad at encoding relational constraints in COPs
- Worst-case example for MWIS
  - a regular graph
  - every node has the same weight

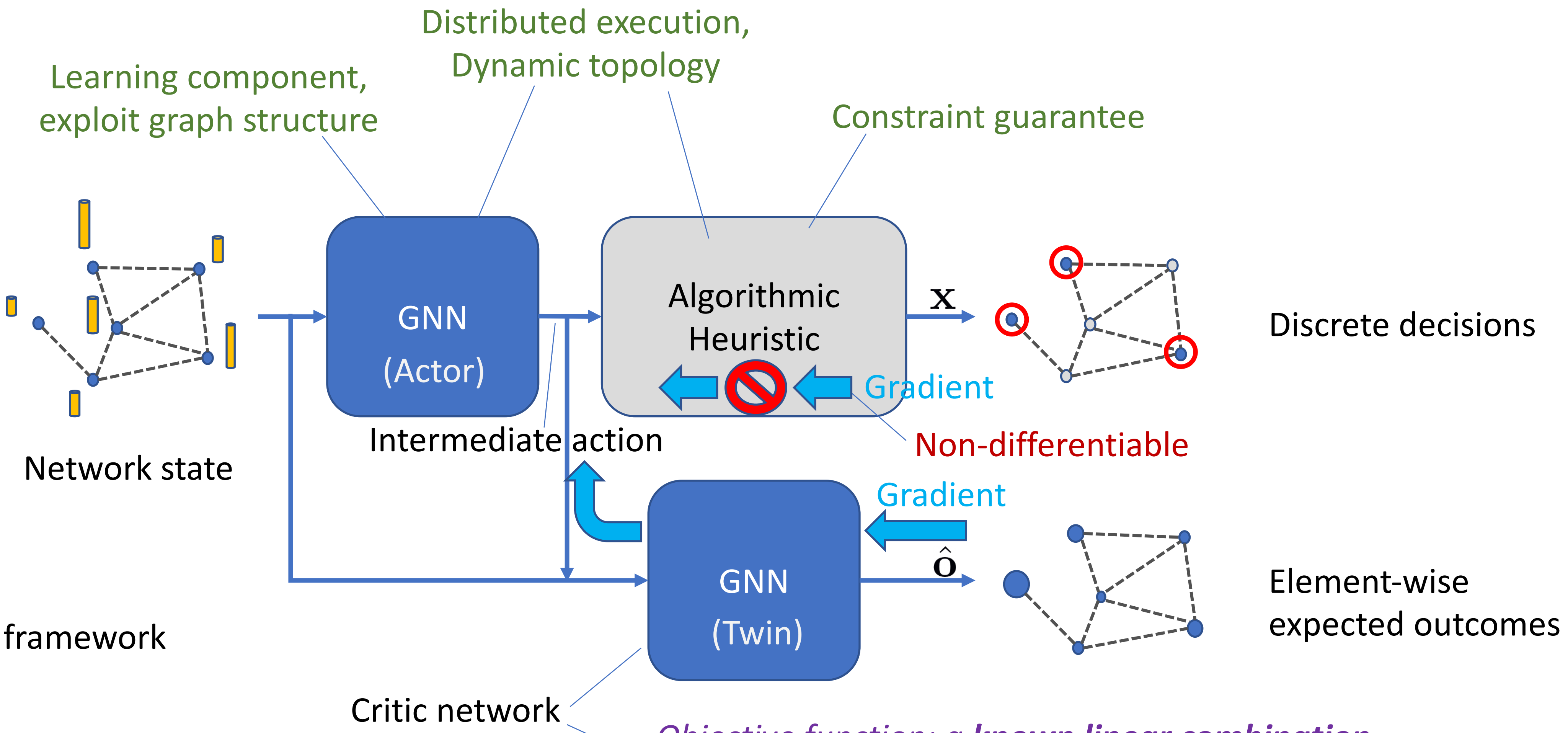
Identical input → identical prediction

Constraint violations



Tie-breaking rules

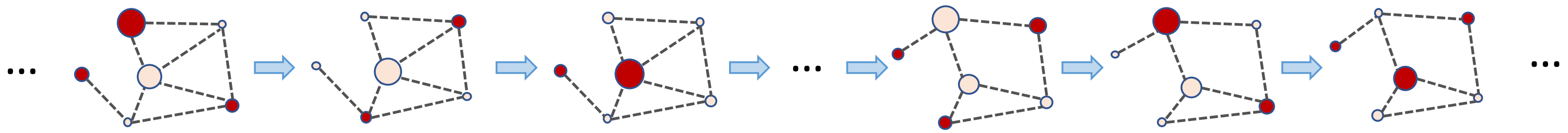
We propose a hybrid pipeline: GDPG-Twin



Actor-critic framework

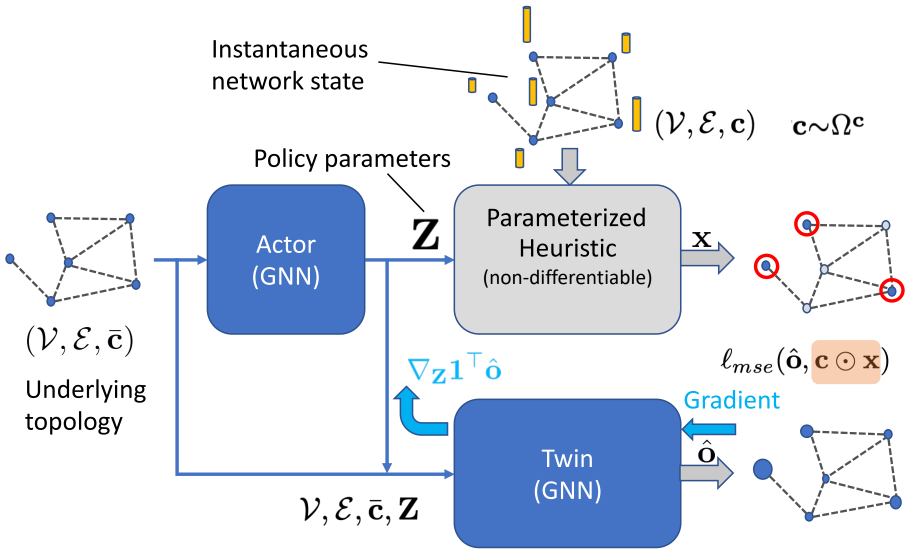
*Objective function: a **known linear combination** of element-wise expected outcomes!*

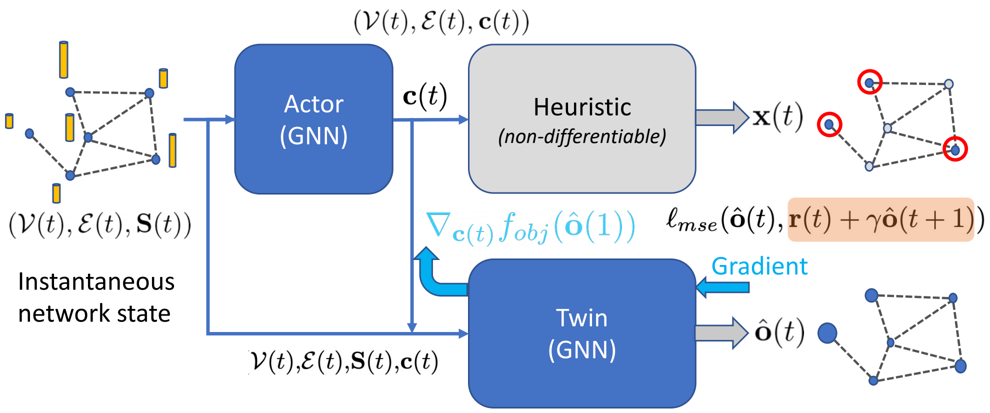
Solutions for two types of R-COPs



- Independent R-COP
  - Optimize each instance individually
- Goal: reduce optimality gap with minimal overhead

- R-COP in graph-based Markov decision process
  - Inter-state dependency MUST be considered
- Goal: achieve long-term system-level objective

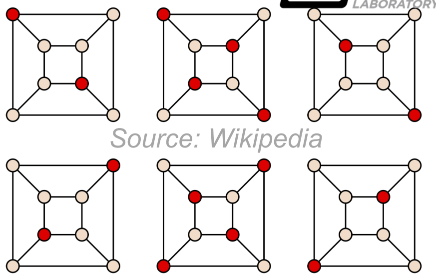




Scalar reward in standalone system → vector of element-wise rewards in network settings

Independent R-COP

Maximum Weighted Independent Set



On 500 random graphs from **Erdős–Rényi** model

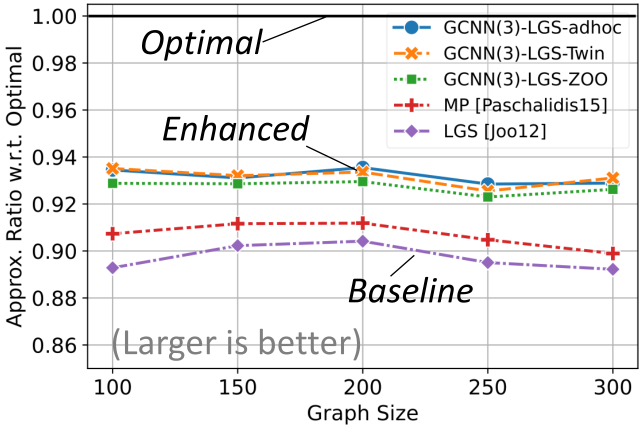


Figure 1: Approximation ratios (Larger is better) of the vanilla and GCNN-enhanced distributed heuristics for MWIS problem (max), w.r.t. the optimal solver.

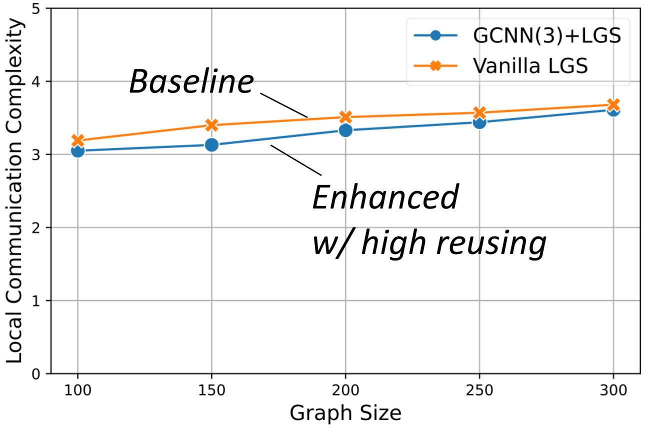


Figure 3: Average local communication complexity of GCNN-enhanced and vanilla LGS-MWIS solvers per instance, in rounds, excluding the GCNN ($N = \infty$).

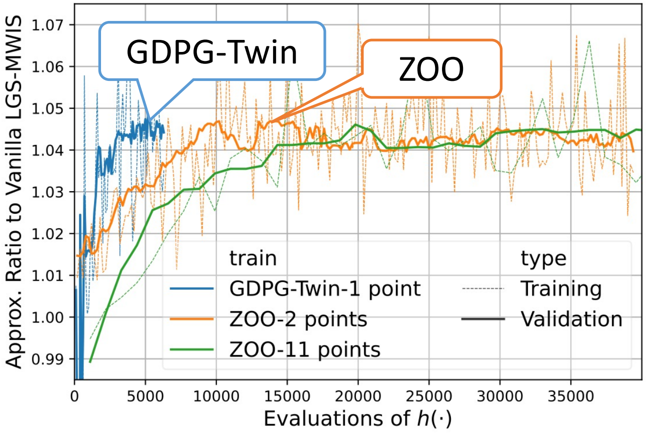
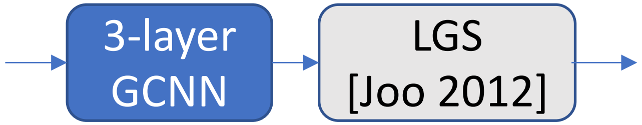


Figure 8: Performance trajectories of GCNN-enhanced LGS-MWIS trained by GDPG-Twin and ZOOs with 2-point and 11-point gradient estimations. Larger is better. GDPG-Twin needs fewer evaluations of $h(\cdot)$.

Approximation ratio

Execution local complexity



Baseline

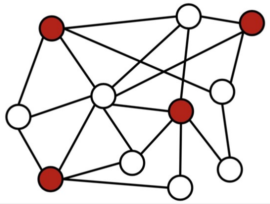
Training complexity

Benchmark: ZOO (zeroth-order optimization)

C. Joo and N. B. Shroff, "Local Greedy Approximation for Scheduling in Multihop Wireless Networks," in *IEEE Transactions on Mobile Computing*, vol. 11, no. 3, pp. 414-426, March 2012, doi: 10.1109/TMC.2011.33.

Generalize to other independent R-COPs

Minimum Weighted Dominating Set



Source:
J. Abernethy,
CS 3510

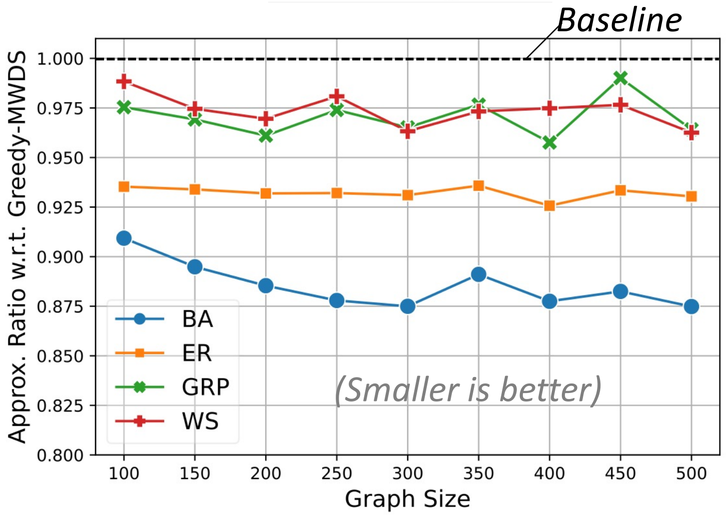
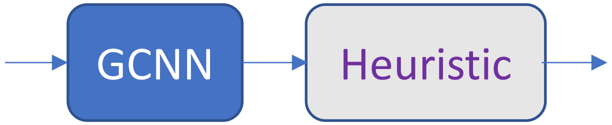
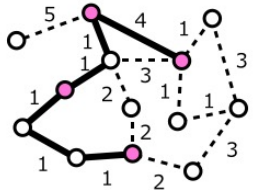


Figure 2: Approximation ratio (Smaller is better) of the GCNN-enhanced w.r.t. the vanilla Greedy-MWDS for MWDS problem (min) on 4 sets of random graphs.



Centralized Greedy-MWDS

Node Weighted Steiner Tree



Source:
(M. Fujita, T. Kimura, & K. Jin'no, 2016)

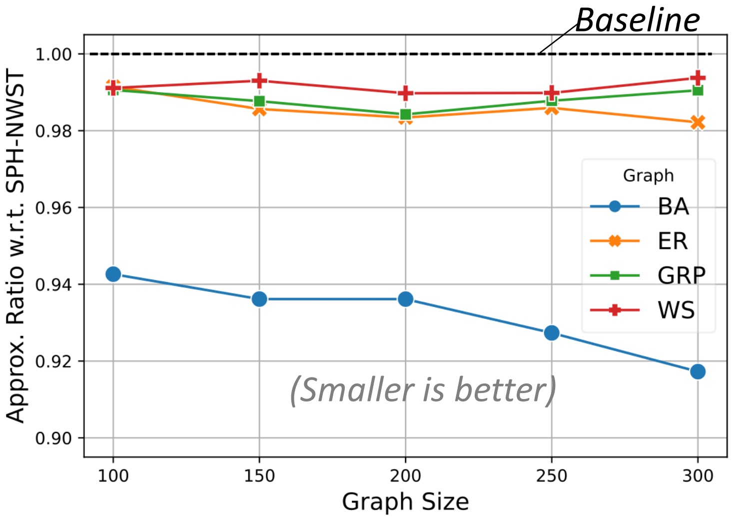
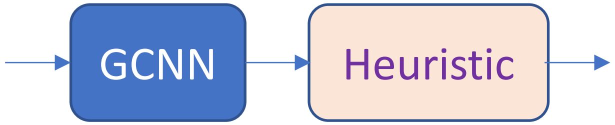
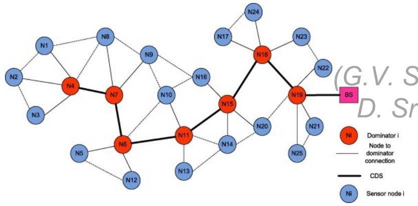


Figure 5: Approximation ratio (Smaller is better) of the GCNN-enhanced w.r.t. vanilla K-SPH-NWST for NWST problem on 4 sets of random graphs. NWST is a minimization (min) problem.



Distributed K-SPH-NWST

Minimum Weighted Connected Dominating Set



Source:
(G.V. Shaamili Varsa, D. Sridharan, 2019)

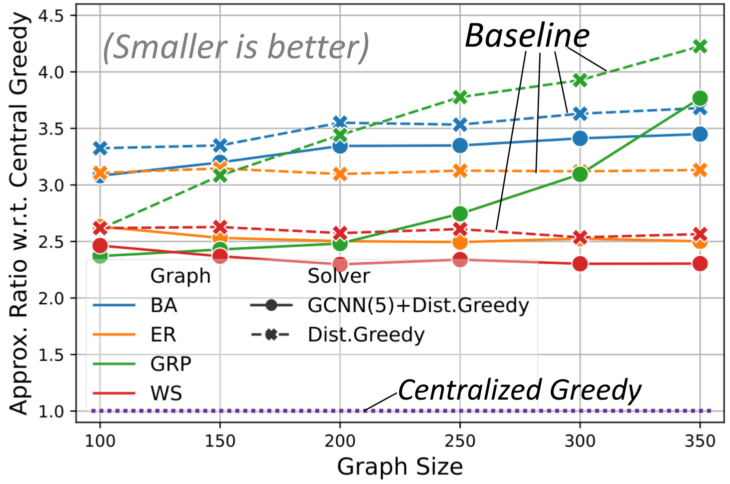
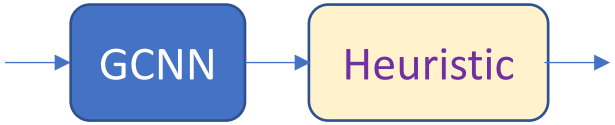


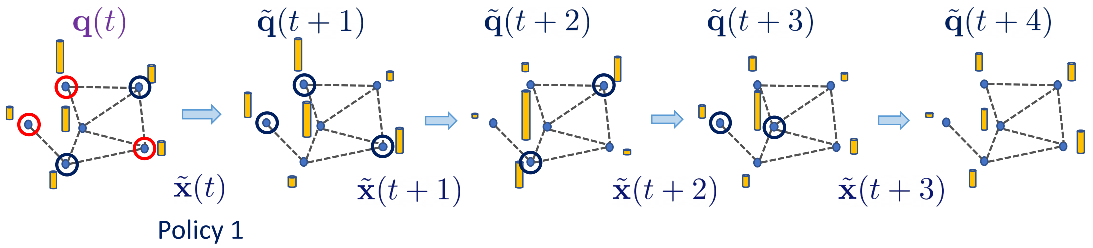
Figure 6: Approximation ratios (Smaller is better) of the vanilla and GCNN-enhanced distributed heuristics w.r.t. a centralized heuristic for MWCDS problem on 4 sets of random graphs. MWCDS is a min. problem.



Distributed Greedy-MWCDS

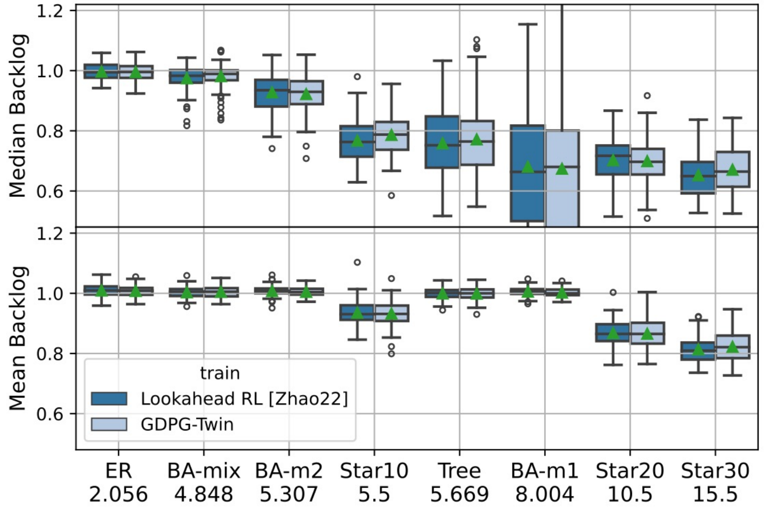
R-COP in graph-based MDP: Delay-oriented link scheduling

The ML pipeline is supposed to improve delay on centralized graphs



$$\mathbf{q}(t+1) = \mathbf{q}(t) + \mathbf{a}(t) - \mathbf{x}(t) \odot \min(\mathbf{l}(t), \mathbf{q}(t))$$

Minimize $\mathbb{E}_{i\in\mathcal{V}, t\leq T}\left[\mathbf{q}_i(t)\right]$



Measure of graph centralization

Figure 7: GDPG-Twin achieves similar network-wide mean and medium backlogs (smaller is better) of lookahead RL (Zhao et al., 2022b) in training a distributed link scheduler, using only 1/5 evaluations of $h(\cdot)$ of it.

GDPG-Twin can do the same job of an <u>ad-hoc RL scheme</u>* at 1/5 computational cost

* Z. Zhao, G. Verma, A. Swami and S. Segarra, "Delay-Oriented Distributed Scheduling Using Graph Neural Networks," *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Singapore, Singapore, 2022, pp. 8902-8906, doi: 10.1109/ICASSP43922.2022.9746926.

Conclusion

- Repetitive Combinatorics
  - Shared topology, different costs
  - Graph-based Markov decision process
  - Limited runtime, distributed execution

- A general actor-critic framework
  - Reduce optimality gap with min overhead
  - Enable long-term goal seeking
  - *Beyond COP → general network processes*

ICLR 2023, Paper ID: 4014

Zhongyuan Zhao, Ananthram Swami, Santiago Segarra,
Graph-based Deterministic Policy Gradient for Repetitive Combinatorial Optimization Problems

https://openreview.net/forum?id=yHIlM9BgOo

Paper URL