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

\[
x^* = \min_{x} \mathbf{c}^\top x
\]

s.t. Discrete constraint on nodes or edges

Constraints defined on Graph, Hypergraph, or Simplicial Complex

Many practical COPs are repetitive!

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

- 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

E.g., COP instances coming at data or video frame rates in wireless link scheduling or computer vision.
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
We propose a hybrid pipeline: GDPG-Twin

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**
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.

Approximation ratio

Execution local complexity

Training complexity

Benchmark: ZOO (zeroth-order optimization)

May 1-5, 2023
Paper ID: 4014, ICLR 2023
Generalize to other independent R-COPs

**Minimum Weighted Dominating Set**

Source: J. Abernethy, CS 3510

**Node Weighted Steiner Tree**

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

**Minimum Weighted Connected Dominating Set**

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

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.

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.

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.

Figures and tables show performance comparisons across different algorithms and network sizes, highlighting the effectiveness of GCNN-based heuristics in solving various combinatorial optimization problems.
R-COP in graph-based MDP: Delay-oriented link scheduling

The ML pipeline is supposed to improve delay on centralized graphs

\[
q(t+1) = q(t) + a(t) - x(t) \odot \min(l(t), q(t))
\]

Minimize

\[
\mathbb{E}_{i \in V, t \leq T} [q_i(t)]
\]

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


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.
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

Z. Zhao, A. Swami, S. Segarra, “Graph-based deterministic policy gradient for repetitive combinatorial optimization problems,” ICLR 2023

Paper URL

https://openreview.net/forum?id=yHIIM9BgOo