CAREER: Learning for Generalization in Large-Scale Cyber-Physical Systems

PI: Cathy Wu (MIT)

Take a picture to **download** the **full papers**

Idea: **Hybridize** methods to address **heterogeneity** in transportation CPS

BROADER IMPACT

- **Heterogeneity** is fundamental to the design & analysis of **future transportation systems.**

Source of heterogeneity	Examples
Diverse stakeholders	Pedestrian, car, truck; local, state, federal
Geographical contexts	Rural, urban; OECD, non-OECD
Rich objectives	Climate, equity, safety, congestion
Emerging technology	Connectivity, automation, electrification

BACKGROUND

- Under heterogeneity, **existing methods often fail**, either finding good solutions slowly (**inefficient**) or not finding them at all (**non-robust**) [1, 2, 7].
- **Hybridize** to achieve the best of both worlds?

Method class	Robustness	Efficiency
Model-based	✓	X
Model-free	X	✓
Hybrid	✓?	✓?

TECHNICAL APPROACH

- Use a **Contextual Markov Decision Process** (cMDP) to capture **heterogeneity** within a problem class
- cMDP: $\mathcal{M} := \{M_c\}_{c\sim p_C} = \{(\mathcal{S}_c, \mathcal{A}_c, \mathcal{T}_c, r_c, \rho_c)\}_{c\sim p_C}$
- **Aim: Design efficient & robust hybridized methods.**

SCIENTIFIC IMPACT

- Address **increasing system complexity** in CPS.
- Bridge gap between model-based & model-free methods.

CHALLENGES

- Challenges stem from numerous parts of a generic solution framework:

$$\pi_c(s) = \arg\max_{a\in A_c} Q_c(s,a), \quad c \in \mathcal{C}$$

Training policies π_c often fragile due to non-robustness of deep reinforcement learning

Existing solvers emit value function $Q_c(s,a)$ but are often slow under new contexts

State s often large due to dense spatio-temporal constraints

Context space $\mathcal{C}$ often large due to heterogeneity in transportation

Action space A_c often large due to the scale of transportation

RESULTS: Simple yet effective hybridization strategies, validated on challenging heterogeneous transportation CPS.

Strategy 1: Accelerate model-based solvers [3-5]

Slow model-based solver

$$\boldsymbol{Q_\theta}(s,a) \approx Q(s,a)$$

Fast inference

Strategy 2: Robustify model-free learning [6-7]

Non-robust policy learning $\boldsymbol{\pi_\theta}(s,a|c)$

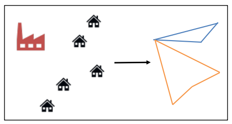
Leverage solutions to related problems

$\pi^*(s,a|c')$ for $c' \approx c$

VALIDATION ILLUSTRATIONS (APPLICATIONS)

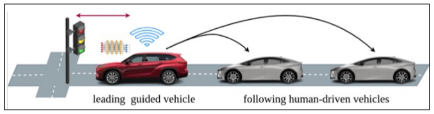
Large VRPs (1000+ customers) [3]

A: Neighborhoods, $|A| \approx 200$
$\mathcal{C}$: Location dist.; problem size; VRP constraints (e.g., time windows)
Q_c: LKH-3

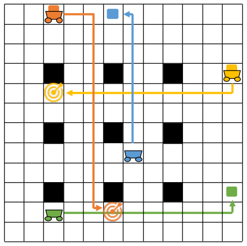


Cooperative eco-driving [6]

A: Longitudinal acceleration
$\mathcal{C}$: Inflow rate; lane length & count; speed limit; phase lengths & count

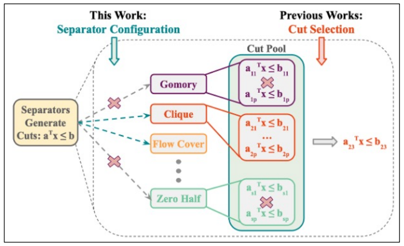


Warehouse MAPF [4]



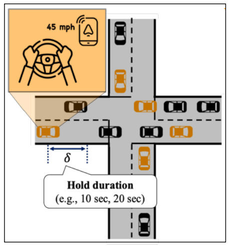
A: Neighborhoods $|A| \approx 100$
$\mathcal{C}$: Floor map (incl. size, # of agents)
Q_c: Priority-based search (PBS)

MILP [5]



A: Separators $|A| \approx 2^{17}$, $|A| \approx 20$;
$\mathcal{C}$: MILP classes (Ex. bin packing, facility location, MIPLIB, load balancing); Q: Branch-and-cut

Advisory autonomy [7]



A: Longitudinal acceleration
$\mathcal{C}$: Advisory duration (0.1-40s);
Traffic network

REFERENCES
[1] F. Lamnabhi-Lagarrigue, et al. “Systems & control for the future of humanity, research agenda: Current and future roles, impact and grand challenges,” Annual Reviews in Control, 2017.
[2] V. Jayawardana, et al., and **C. Wu**, “The impact of task underspecification in evaluating deep reinforcement learning,” in Advances in neural information processing systems (NeurIPS), 2022.
[3] S. Li*, Z. Yan*, and **C. Wu**, “Learning to delegate for large-scale vehicle routing,” Advances in Neural Information Processing Systems (NeurIPS), 2021. **Spotlight (<3%)**.
[4] Z. Yan and **C. Wu**, “Neural Neighborhood Search for Multi-agent Path Finding,” International Conference on Learning Representations (ICLR), 2024.
[5] S. Li*, W. Ouyang*, M. Paulus, and **C. Wu**, “Learning to Configure Separators in Branch-and-Cut,” in Advances in Neural Information Processing Systems (NeurIPS), 2023.
[6] V. Jayawardana, S. Li, **C. Wu**, Y. Farid, and K. Oguchi, “Generalizing Eco-Lagrangian Control via Multi-residual Task Learning,” IEEE International Conference on Robotics and Automation (ICRA), 2024.
[7] J.-H. Cho, S. Li, J. Kim, and **C. Wu**, “Temporal Transfer Learning for Traffic Optimization with Coarse-grained Advisory Autonomy,” In review.

Award ID#: CNS - 2239566