Stochastic Resource Optimization over Heterogeneous Graph Neural Networks for Failure-Predictive Maintenance Scheduling

Zheyuan Wang*, Matthew Gombolay*
*CORE Robotics Lab, School of Interactive Computing, Georgia Institute of Technology, Atlanta, GA, USA
https://core-robotics.gatech.edu/ E-mail: {pjohnwang,mgombolay3}@gatech.edu

Georgia Institute of Technology

Cognitive Optimization & Relational Robotics Lab

icaps

Introduction

- ❑ Resource optimization for predictive maintenance requires inferring and reasoning over stochastic failure models and dynamically allocating repair resources.
- ❑ Predictive maintenance scheduling is typically performed with ad hoc, hand-crafted heuristics and manual scheduling by human experts, which is time-consuming, laborious and hard to scale.
- ❑ Recent advances in AI have leveraged deep neural networks to solve operations research problems, but only touched static, deterministic setting.
- ❑ We build the scheduling policy network directly on a heterogeneous graph representation of scheduling problems and develop an RL-based policy optimization procedure to enable robust learning in highly stochastic environments.

Aircraft Maintenance Environment

- ❑ We develop a virtual predictive-maintenance scheduling environment, called AirME.
  - Heterogeneous aircraft, $\{p_i\}$
  - Homogeneous maintenance crews, $\{c_j\}$
  - A maintenance decision, $d = < p_i, c_j >$
  - An hour-based time system with discrete time steps
- ❑ Hybrid probabilistic failure model
  - Model component/part failure using Weibull distributions based on aircraft usage

$$p(x;\lambda,k) = \frac{k}{\lambda}\left(\frac{x}{\lambda}\right)^{k-1} e^{-(x/\lambda)^k}, x \geq 0$$

  - Grounded with "broken" status when failure happens for at least one of its components
  - Model parameters hidden from scheduling policies
- ❑ Stochastic aircraft maintenance task
  - Duration and cost generated on-the-fly
- ❑ Sample flying operations based on usage rate
  - Different planes earn different hourly income
- ❑ Scheduling objectives
  - O1: overall profit; O2: revenue only; O3: fleet availability
- ❑ POMDP formulation

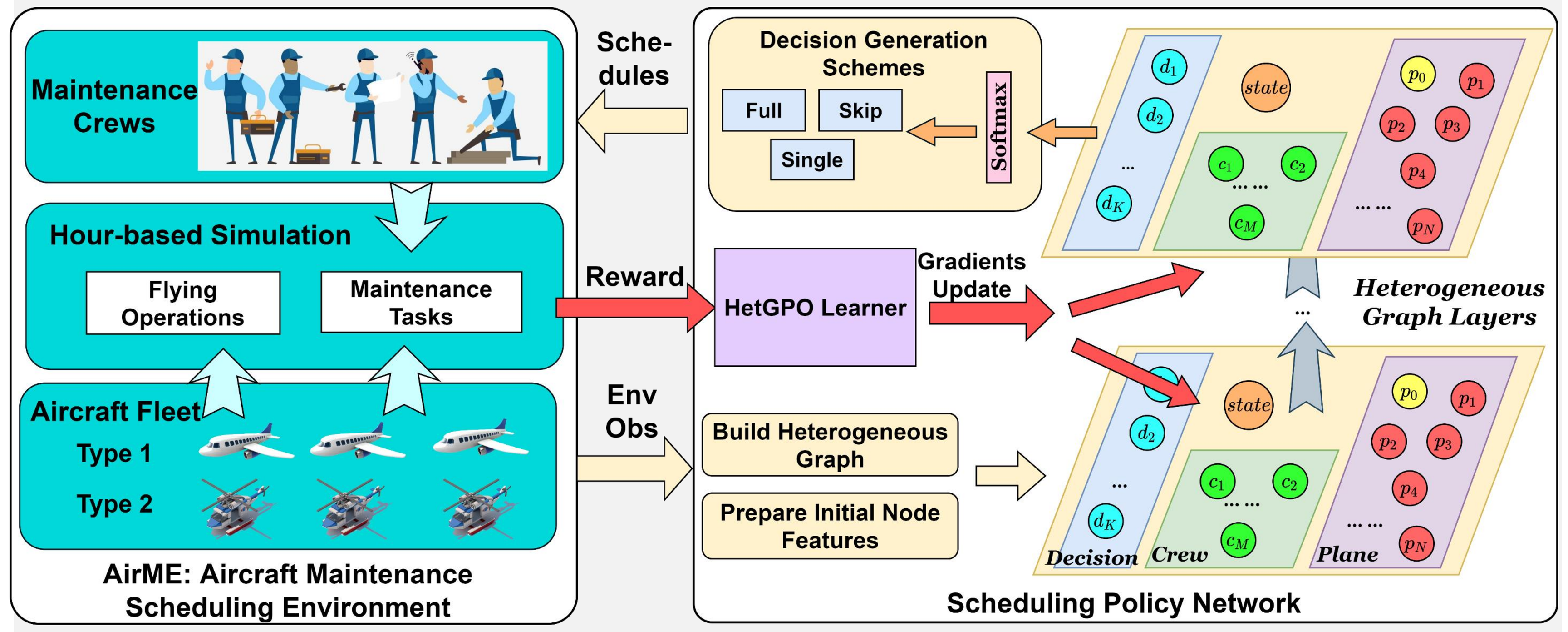

Fig. 1. Overview of AirME and Scheduling Policy Network

Stochastic Scheduling with Graphs

- ❑ Scheduling Policy Network, $\pi_\theta(u|o)$
  - Fully graph convolutional structure
  - Action $u_t = \{d_1, d_2, \ldots, d_n\}$
  - Conditional policy

$$p_\theta(u_t|o_t) = \prod_{i=1}^{n} p_\theta(d_i|o_t, d_{1:i-1})$$

  - Decision generation schemes: Full, Skip, Single
- ❑ Heterogeneous Graph Representation
  - Joint learning the problem representation and the policy
  - Model entities in the environment
    - Planes, crews
  - Model RL components
    - State, decisions
- ❑ A novel heterogeneous graph layer
  - Building block of our scheduling policy network
  - Computation Step
    - 1) per-edge-type message passing
    - 2) per-node-type feature reduction
  - Softmax-based attention for state summary node
    - Multi-head adaptation

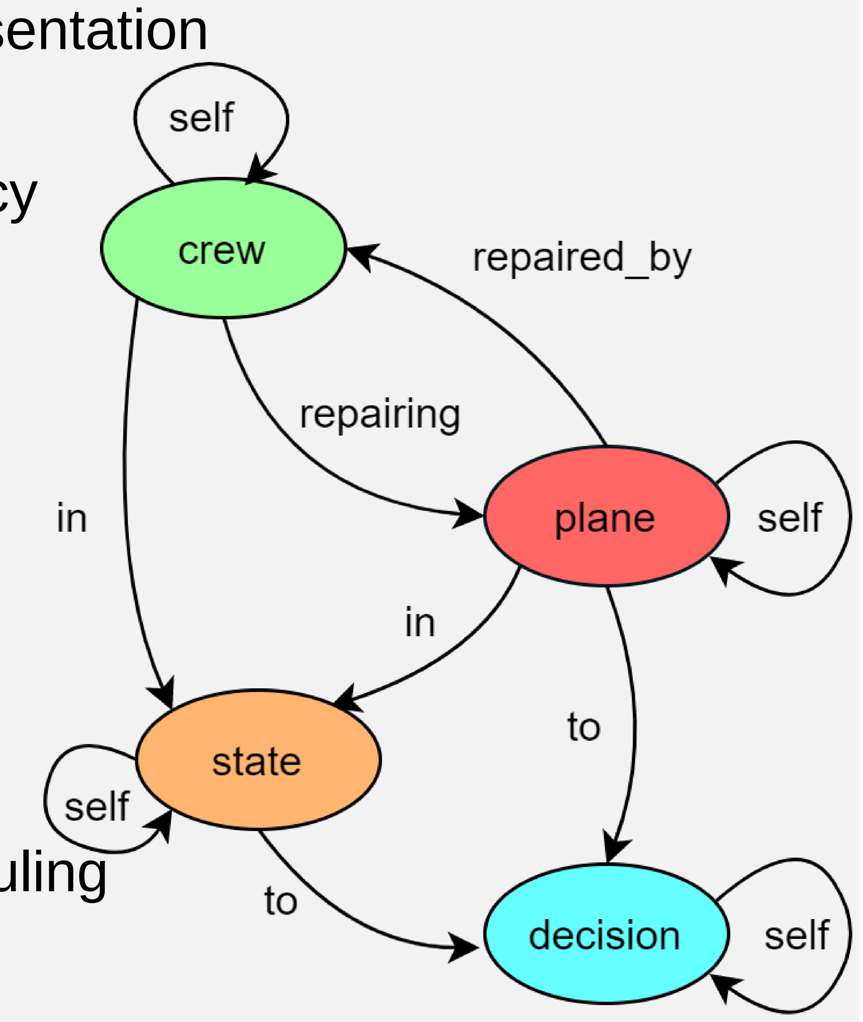

Fig. 2. Metagraph

Stochastic Policy Learning Methods

- ❑ HetGPO: heterogeneous graph-based policy learning framework
  - Developed from Proximal Policy Optimization (PPO) with several adaptations for stochastic scheduling
- ❑ Clipped surrogate objective

$$L(\theta) = \mathrm{E}_t[\min(r_t(\theta)A_t, clip(r_t(\theta), 1-\epsilon, 1+\epsilon)A_t)]$$

$$r_t(\theta) = \frac{\pi_\theta(u|o)}{\pi_{old}(u|o)} \qquad A_t\text{: Advantage term}$$

- ❑ Step-based baseline for variance reduction
  - More accessible and efficient than state-based baselines
- ❑ Use idea of curriculum learning

Experimental Results

- ❑ We evaluate the utility of HetGPO against the following algorithms
  - Hand-crafted heuristics, Model-based planning
  - Machine learning-based methods
- ❑ Dataset: small, medium and large AirME instances
- ❑ Evaluation on O1, O2, O3 using two metrics
  - M1: normalized objective value
  - M2: % improvement over the Corrective Scheduler
- ❑ Ablation studies

Table I. Evaluation results on O1: profit

Methods	Small		Medium		Large	
	M1	M2 (%)	M1	M2 (%)	M1	M2 (%)
Random	0.522 ± 0.025	-2.87 ± 5.65	0.532 ± 0.021	-2.65 ± 4.22	0.533 ± 0.016	-2.23 ± 3.51
Corrective	0.539 ± 0.023	0.0 ± 0.0	0.547 ± 0.016	0.0 ± 0.0	0.546 ± 0.016	0.0 ± 0.0
Condition-based	0.656 ± 0.050	21.7 ± 6.41	0.661 ± 0.041	20.8 ± 5.87	0.648 ± 0.051	18.6 ± 7.03
Periodic	0.599 ± 0.051	11.1 ± 6.96	0.598 ± 0.047	9.38 ± 7.00	0.587 ± 0.048	7.52 ± 6.83
Model-based	0.669 ± 0.052	24.0 ± 6.99	0.671 ± 0.044	22.8 ± 6.45	0.658 ± 0.054	20.5 ± 7.68
DeepRM	0.533 ± 0.015	-0.88 ± 2.57	0.538 ± 0.011	-1.47 ± 1.77	0.539 ± 0.013	-1.11 ± 0.97
Decima	0.651 ± 0.021	21.1 ± 6.42	0.660 ± 0.017	20.9 ± 4.49	0.663 ± 0.014	21.6 ± 4.51
HetGPO-Single	0.680 ± 0.012	26.4 ± 4.32	0.676 ± 0.011	23.7 ± 3.17	0.666 ± 0.011	22.3 ± 3.77
HetGPO-Skip	**0.695 ± 0.010**	**29.1 ± 4.09**	**0.697 ± 0.009**	**27.5 ± 2.70**	**0.695 ± 0.008**	**27.5 ± 2.72**
HetGPO-Full	0.693 ± 0.011	28.8 ± 4.01	0.694 ± 0.009	27.1 ± 2.62	0.693 ± 0.008	27.1 ± 2.68

See paper for full results on O2, O3

Conclusion

- ❑ HetGPO outperforms both heuristics and learning-based counterparts across various objectives and problem scales.
- ❑ Our future work involves applying HetGPO to a broader class of stochastic resource optimization problems, such as Patient Admission Scheduling (PAS) problems.

June 2022