

---

# Control of Autonomous Electric Fleets for Ridehail Systems

Nicholas Kullman<sup>\*1,2</sup>, Martin Cousineau<sup>3</sup>, Justin Goodson<sup>4</sup>, and Jorge Mendoza<sup>2,3</sup>

<sup>1</sup>Laboratoire d'Informatique de l'Université de Tours (LIFAT) – Université de Tours : LIFATEA 6300, CNRS, ROOT ERL CNRS 7002 – 64 Avenue Jean Portalis, 37200 Tours, France

<sup>2</sup>Centre Interuniversitaire de Recherche sur les Réseaux d'Éentreprise, la Logistique et le Transport (CIRRELT) – Pavillon André-Aisenstadt, bureau 3520 2920, Chemin de la Tour Montréal (Québec) H3T 1J4 CANADA, Canada

<sup>3</sup>HEC Montréal – 3000 Chemin de la Côte-Sainte-Catherine, Montréal, QC H3T 2A7, Canada

<sup>4</sup>Saint Louis University Richard A. Chaifetz School of Business (SLU) – 3674 Lindell Blvd, St. Louis, MO 63108, United States

## Abstract

Operators of ridehail platforms such as Lyft and Uber will likely be early-adopters of autonomous electric vehicles (AEVs), since AEVs promise to reduce costs, be safer, and more efficient. While studies on the operation of ridehail systems with AEVs exist, nearly all have ignored the need to recharge the vehicles during operation. We address this here in our work on the ridehail problem with AEVs (RP-AEV).

In the RP-AEV, a decision maker (DM) operates a fleet of AEVs that serve requests arising randomly throughout a region. The DM is responsible for assigning AEVs to requests, as well as repositioning and recharging AEVs in anticipation of future requests. We model the RP-AEV as a Markov decision process.

We compare classical approximate dynamic programming (ADP) solution methods with those of deep reinforcement learning (RL), which have garnered enthusiasm but achieved only limited success to date in operational problems. From ADP, we explore novel heuristic policies, both alone and combined with lookaheads. From RL, we build on the approach from Holler et al. (2018). We employ neural-networks (NNs) both to determine the state representation (with single-layer NNs) and to learn state-action value functions (with deep NNs) using Q-learning.

Additionally, we establish a dual bound to gauge the effectiveness of these approaches by calculating the expected value with perfect information. With perfect information, the RP-AEV may be decomposed so as to permit a solution via Benders decomposition, where the master problem assigns AEVs to requests, and the subproblem provides instructions for repositioning and recharging.

**Keywords:** autonomous vehicles, electric vehicles, markov decision process, deep learning, reinforcement learning, approximate dynamic programming

---

\*Speaker