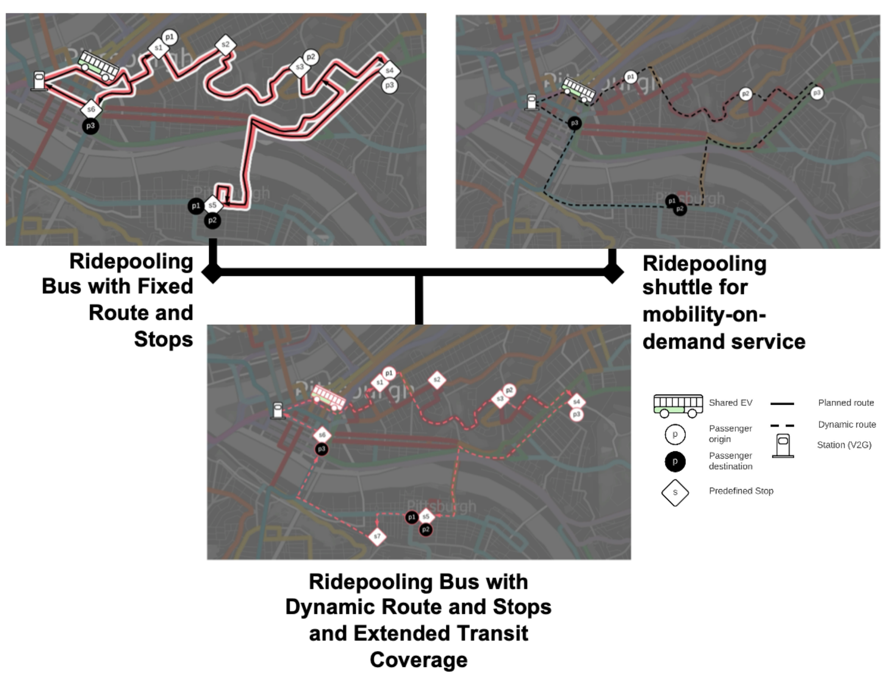
Equitable Ridepooling with Community-Driven Stops Using a Fleet of Autonomous Electric Vehicles

Rami Ariss, Ph.D. Civil & Environmental Engineering, Carnegie Mellon University, +1 (628) 224-9177, rariss@andrew.cmu.edu
Jeremy Michalek, Prof. Engineering and Public Policy, Carnegie Mellon University, +1 (412) 657-9879, jmichalek@cmu.edu
Matteo Pozzi, Prof. Civil & Environmental Engineering, Carnegie Mellon University, +1 (412) 992-8460, mpozzi@andrew.cmu.edu

Overview

The promising developments towards full driving automation will revolutionize transportation by presenting significant opportunities in self-driving vehicles, shared mobility, and electrification. However, it is critical that social responsibility be at the heart of this transition or else comfort and convenience for the privileged will come at the expense of marginalized groups and the environment. An equitable transportation revolution must ensure fleets of autonomous vehicles are electric and shared. Modernizing public transit for a shared autonomous future necessitates fleet management controls designed for ridepooling and handling the logistical complexities of electric vehicles. We formulate and address a multi-agent control problem where a fleet of multi-passenger autonomous electric vehicles (e.g. buses or shuttles) are coordinated by a central transit agency to best serve a community of riders who choose routes based on a predetermined set of stops. This novel autonomous ridepooling framework lays between existing public transit services with fixed station stops and routes and pure mobility-on-demand services (e.g. Uber, Lyft) with dynamic and adaptive stops and routes. The transit agency is expected to act in the best interest of society, managing the autonomous electric fleet to maximize social good by accounting for externalities in equitable and sustainable transportation. Rider data privacy and sociodemographic information collected and utilized by the transit agency will be nominal and socially responsible. We develop a Multi-Agent Reinforcement Learning (MARL) simulation environment and test various heuristic and RL trained policies that can control a fleet of autonomous electric vehicles for ridepooling with stochastic requests from a community. Using domains with different graph sizes (e.g. number of geographically distributed stop locations) and request distributions (i.e. for rider requests over each route), we observe that RL methods can be used to learn useful multi-agent policies that outperform benchmark heuristics, though performance is domain-dependent. Furthermore, trained autonomous RL agents can learn and follow policies that complement existing transit in our multi-agent framework, allowing them to extend the capacity and equity of legacy transportation systems.



Methods

We formulate a multi-agent ridepooling problem for a fleet of autonomous electric vehicles that adapt to community-driven requests to maximize social good. Vehicles serve as agents coordinated by a central agency that collects and distributes information on riders' origin-destination preferences over a set of potential stops and wait times. The formulation limits the amount of discriminatory information collected from passengers to uphold data privacy and social responsibility. Vehicle actions include driving to pick-up, pool, complete rides of multiple passengers, to stay and wait at a location, or to reach charging stations to recharge batteries. We model the system as a Markovian process whose evolution is dictated by the current state of the fleet and riders, autonomous electric vehicle and central agency decisions, and ride request and other exogenous information. The reward function incorporates ride profits, equity incentives, and time and distance-based operational and charging costs. The objective is to learn an optimal, dynamic multi-agent strategy that the central agency and individual vehicles can implement to maximize social good, including equity-based completed rides, wait times, and operational costs.

In order to model sequential decision-making under uncertainty, a Markov Decision Process (MDP) is used to frame stochastic control processes. Exact dynamic programming methods, such as value iteration and policy iteration, can be used to solve for an optimal policy that maximizes the discounted cumulative reward with guaranteed convergence based on known dynamics, known reward functions, and properties of the Bellman equations as a contraction operator. However, as the dynamics and reward functions are unknown in our ridepooling framework, these exact dynamic programming methods cannot be used. Furthermore, our state-action space is continuous and of high-dimensionality, so exact dynamic programming methods cannot be used and cannot scale. This necessitates our use of approximate dynamic programming methods, or reinforcement learning (RL) to learn policies over high-dimensional discrete and continuous state-action spaces in stochastic environments without prior knowledge of the dynamics or reward functions.

Within variations of our problem framework—incorporating passenger counts, passenger wait times, and trip or episodic time remaining—we explore different reward functions and demonstrate the ability for useful policies to be learned and outperform heuristics. As our ridepooling framework is novel, we propose reasonable heuristic policies that dictate a vehicles action for benchmarking. These heuristics include a random policy, fixed route policy (i.e. solving a eulerian circuit and cycling through stops), greedy policy (i.e. choosing the stop with the most outstanding requests), max state policy (i.e. choosing and completing a route with the most outstanding requests), and a dynamic programming lookahead policy (e.g. solving a three-step dynamic program lookahead for the current state and selecting a proxy optimal action using the learned value function).

Results

We utilize Double Deep Q-Networks (DDQN) as an RL algorithm to learn a useful neural network policy across various ridepooling domains. These domains vary by grid sizes (3-10 locations), geographic distribution of stops, request distributions (e.g. Skellam distribution for net requests; pulsed and time-cyclic requests), and reward functions (e.g. including distance-based costs; equity-scaled requests). We observe that RL outperforms most of our proposed heuristics over these domains with respect to mean episodic drop-offs to ± 2 statistically significant sample standard deviations. In terms of scalability, DDQN can learn a useful policy within reasonable computation time over these domains, while inference using a learned neural network policy is comparatively exceptionally quick and can be used for real-time execution. Multi-agency is also demonstrated where a single autonomous vehicle learns and adapts to an existing fleet of fixed route vehicles to collectively maximize the cumulative discounted reward.

Conclusions

To empower an equitable and sustainable transportation future, novel fleet management controls are needed in transportation systems to leverage autonomous, electric vehicles. We demonstrate a novel ridepooling framework as a Markov Decision Process and develop simulation environments that utilize an efficient set of state and action spaces that capture equity-based ridepooling with multiple agents to train RL-based policies on. In this environment, domain-specific performance varies by the number of locations, geographic distribution of stops, the set of route request distributions, and structure of the reward function. Over a variety of domains, we demonstrate the ability to learn useful policies under uncertainty using DDQN that can outperform reasonable heuristic policies (i.e. random, fixed route, greedy, max state, dynamic programming lookahead). We find that over the explored domains, the larger the grid size and the more heterogeneous the route request distributions are, the more likely that RL-based methods will outperform heuristic methods with statistical significance.