ASCE T&DI
Technical Committee on
Artificial Intelligence
Student Competition
November 18, 2021

Connected Automated Vehicle Cooperative Control with a Deep Reinforcement Learning Approach in a mixed traffic environment

Haotian Shi

WISCONSIN
UNIVERSITY OF WISCONSIN–MADISON

INTRODUCTION

Background:

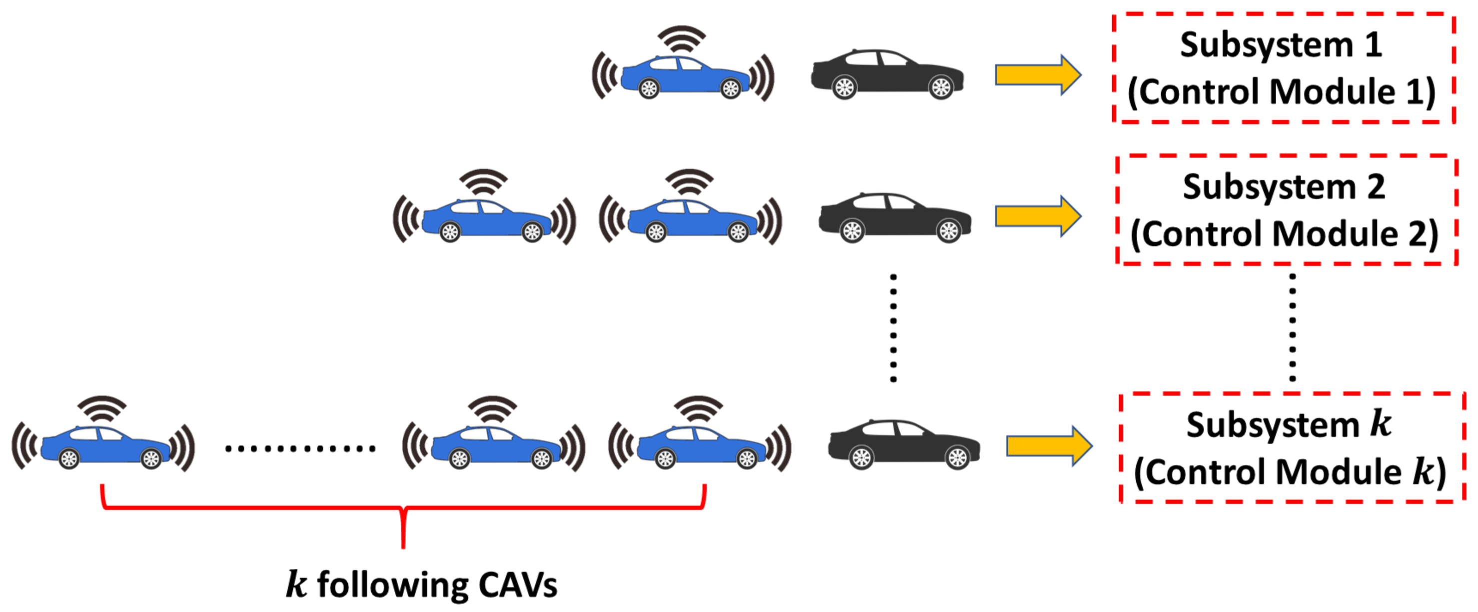
- **Scenario:** Connected automated vehicles longitudinal control in **mixed traffic** of CAVs and human-driven vehicles (HDVs).
- **Research gaps:**
  - A **fast-computing, generalizable, stable cooperative multi-objective** CAVs control algorithm with incorporating **HDV's car-following characteristics**.
  - Hard to optimize mixed traffic flow considering the **different CAV-HDV combinations**.
- **Deep reinforcement learning:**
  - Suitable for capturing complex and **stochastic** system characteristics.
  - Rapidly implemented in online process.
- **Equilibrium concept from control theory:**
  - Facilitates the stability analysis (i.e., **local stability and string stability**).
  - Gives DRL an exploration direction in the training process to **improve the convergence**.

Objectives:

- Integrate the merits of **DRL algorithm** (suitable for complex characteristics and rapid online implementation) and optimal control methods (MPC) (constrained optimization framework which enables **multiple objectives** and constraints based on the concept of **equilibrium point**).
- Achieve **cooperative** CAV longitudinal control with **incorporating characteristics of HDVs**.

ENVIRONMENT SETTING

- **Decompose** mixed traffic into **multiple subsystems** where each subsystem comprises HDV followed by cooperative CAVs.
- CAVs within each subsystem are **cooperatively controlled** by a corresponding DRL-based controller.



Model Development

Policy update algorithm: Distributed Proximal Policy Optimization (DPPO) algorithm

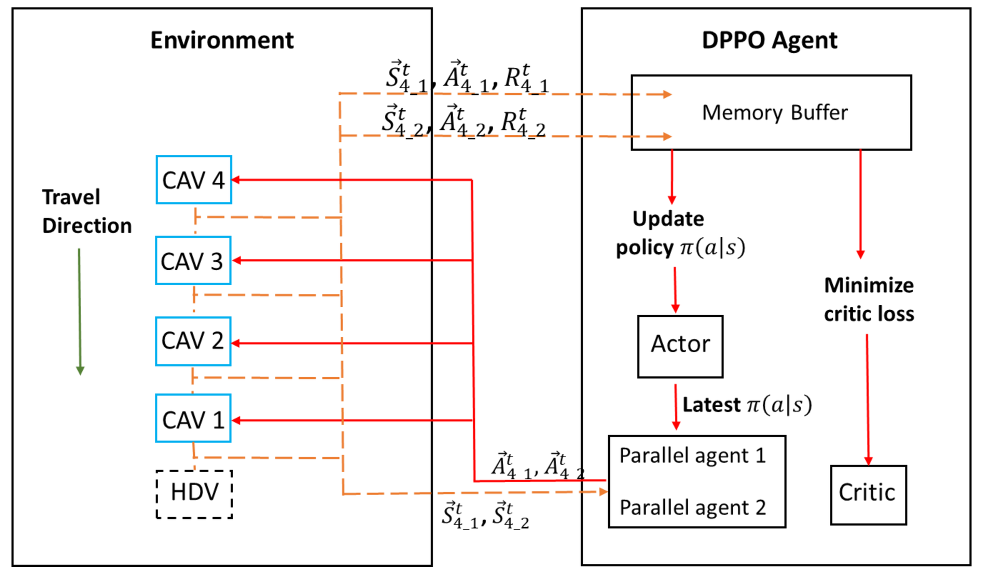
State $\vec{S}_k^t = \left[\vec{s}_{k,1}^t\ \vec{s}_{k,2}^t \ldots\ \vec{s}_{k,i}^t\ \ldots\ \vec{s}_{k,k}^t\right]^T, \vec{s}_{k,i}^t = \left[v_i^t\ \ g_i^t\ \ \Delta v_i^t\ \ \Delta d_i^t\right]^T$

Action: $\vec{A}_k^t = \left[a_{k,1}^t\ a_{k,2}^t, \ldots a_{k,i}^t \ldots a_{k,k}^t\right]^T$

Multiple Objectives:

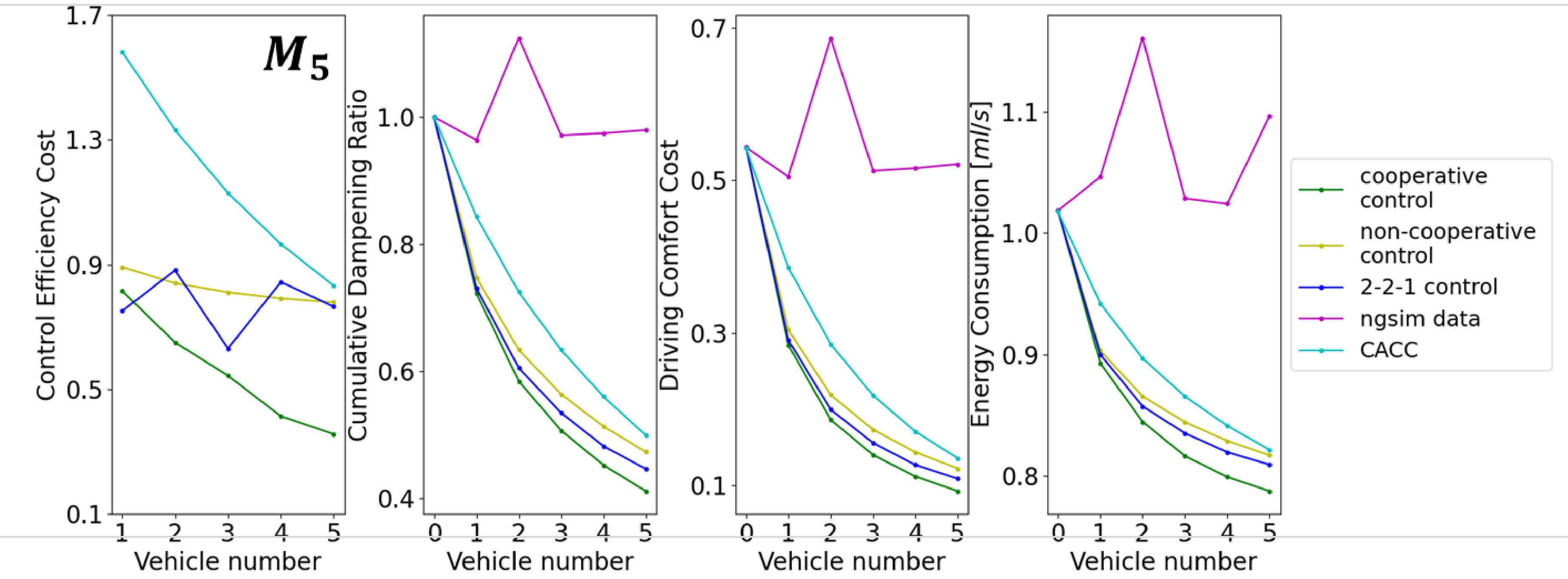
- Car following efficiency: $c_i^t\left(\vec{x}_i^t\right) = \left(\vec{x}_i^t\right)^T Q_i \vec{x}_i^t, \vec{x}_i^t = \left[\Delta d_i^t,\ \Delta v_i^t\right]$
- String stability: $d_{p,i} = \dfrac{\|a_i^t\|_2}{\|a_0^t\|_2} = \dfrac{\left(\sum_{t=0}^N |a_i^t|^2\right)^{\frac{1}{2}}}{\left(\sum_{t=0}^N |a_0^t|^2\right)^{\frac{1}{2}}}$
- Energy consumption: $e_i^t\left(v_i^t, a_i^t\right)$ based on VT-micro model.

Training:
Embed the **ground-truth NGSIM dataset** in the training process

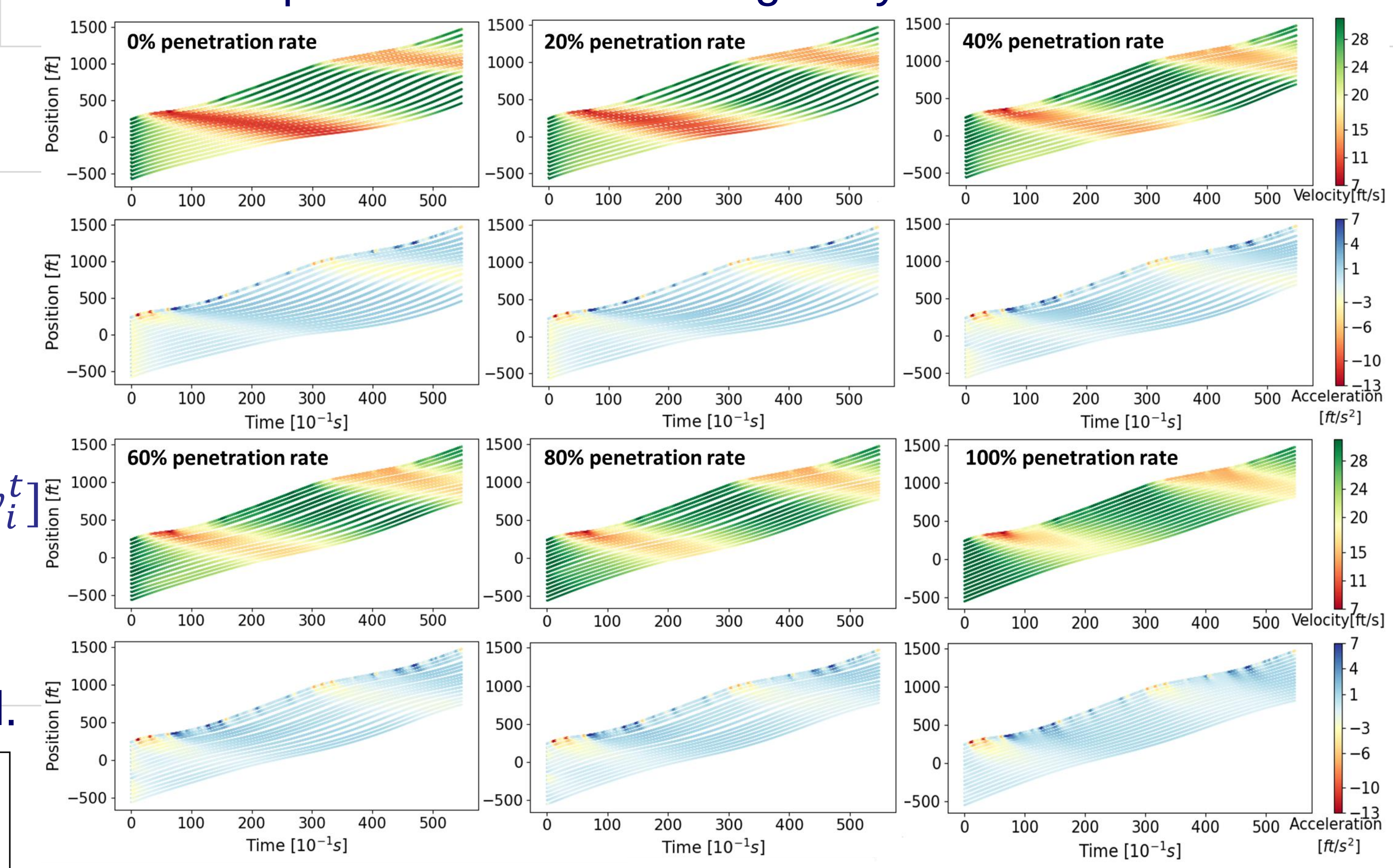


SIMULATION EXPERIEMENTS

Model performance validation: compared with NGSIM data, decentralized control, and linear-based CACC strategy.



Models applied in the mixed traffic with different penetration rates: dampen traffic oscillations greatly



Shi, H., Zhou, Y.*, Wu, K., Wang, X., Lin, Y., & Ran, B. (2021). Connected automated vehicle cooperative control with a deep reinforcement learning approach in a mixed traffic environment. *Transportation Research Part C: Emerging Technologies*, *133*, 103421.