PRESENTATION at INFORMS ANNUAL MEETING 2021

informs ANNUAL MEETING

2021 ANAHEIM, CALIFORNIA

GAQ-EBkSP: A DRL-based Urban Traffic Dynamic Rerouting Framework using Fog-Cloud Architecture

Presenter: Runjia (Rayne) Du

Authors: Runjia Du, Sikai Chen*, Jiqian Dong, Pual (Youngjoun) Ha, Samuel Labi

Urban traffic congestion is a persistent problem

Increasing vehicle ownership

Total number of registered vehicles increase from 250 million ~ 273 million in 8 years (2010~2018)

Route-guidance methods in Intelligent Transport Systems (ITS)

Car-navigation systems: GoogleMap TomTom

Variable Message Signs (VMS): broadcast real-time traffic flow information

Problems with Car-navigation systems and VMS

identical guidance may provide for all vehicles with similar destinations

New occurrence of traffic congestion at other locations

Multi-route planning algorithm

route popularity

vehicle priority

→ **EBkSP**

Urban traffic network is extremely complex

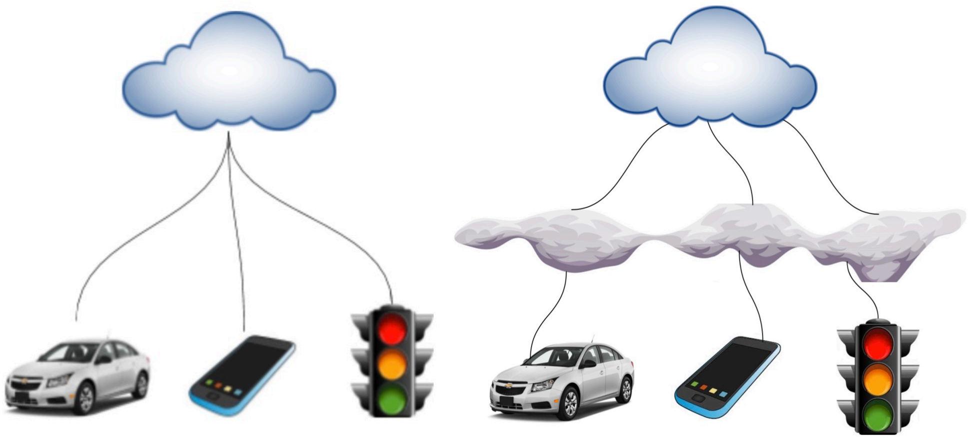
Information exchange needs to be highly efficient

Cloud computing environment

Information resources locate in the core of the network, the information exchange (communication) can be time-consuming (with high latency)

Fog computing environment

Information resources locate on the edge of the network, thus decreasing the distance from the core to the users and enhance the communication efficiency

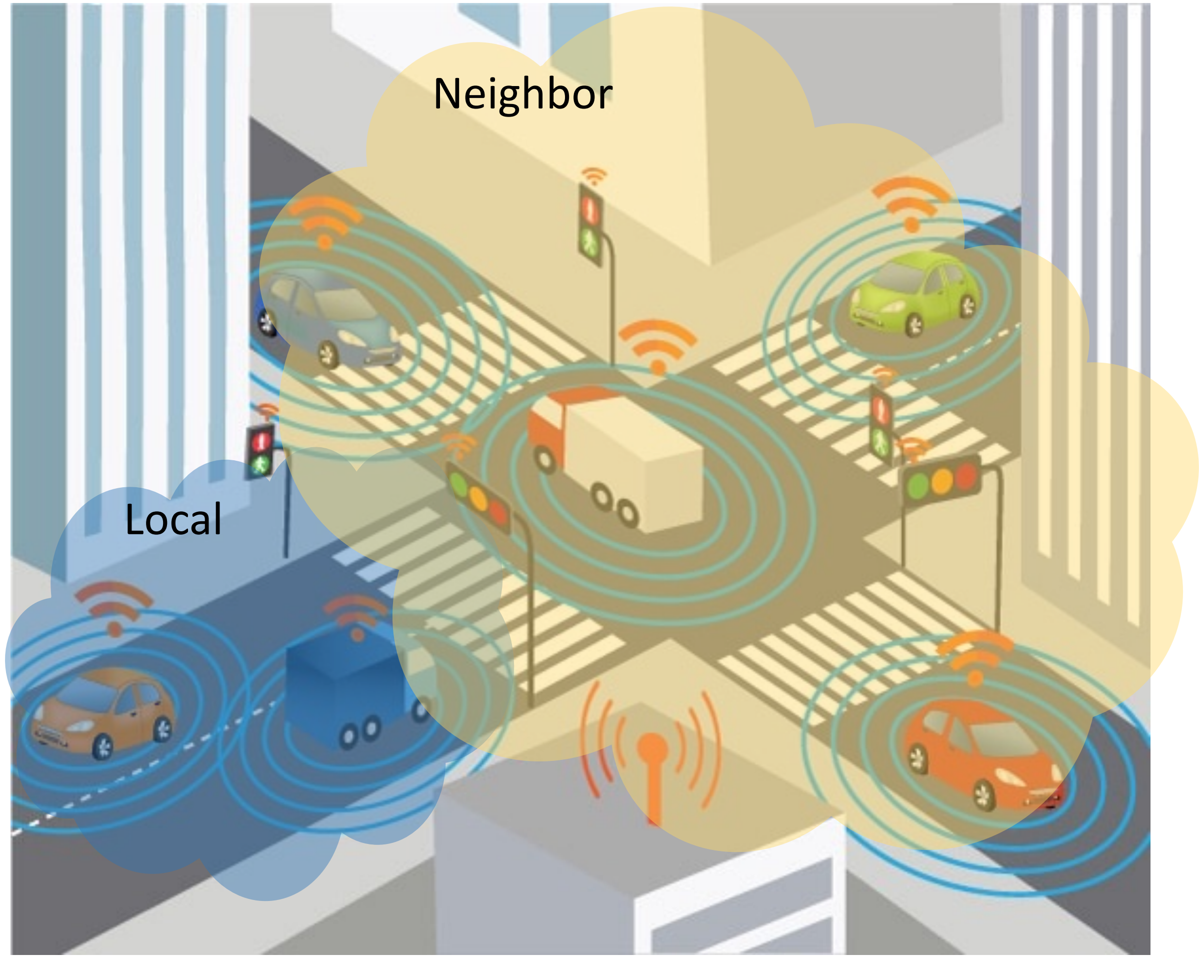


Urban traffic systems: Highly dynamic

- Deep Reinforcement Learning (DRL) is ideal for solving the dynamic rerouting problem

Urban traffic systems: large volume of information

- Both local and neighbor information are crucial for understanding the overall driving environment
- Differentiate the relative importance of input information based on the final decision can enhance the learning efficiency

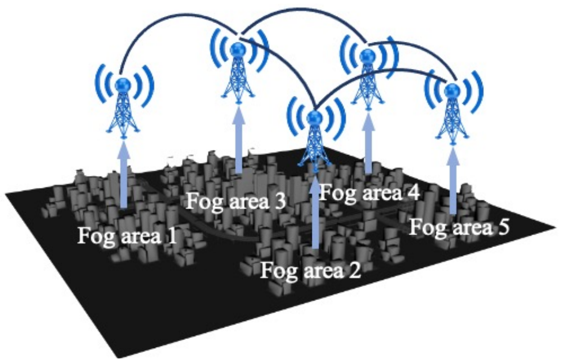


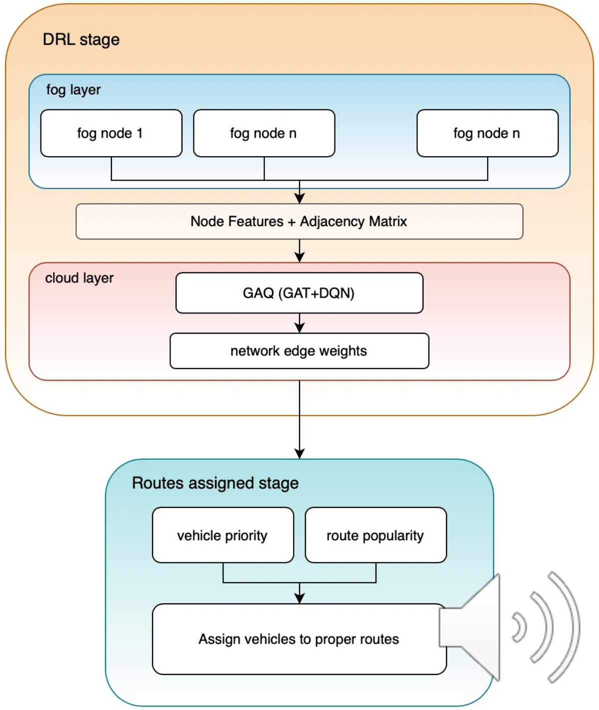
Design and test approaches for dynamically reroute the vehicles through the combination of DRL + attention mechanism and multi-route planning algorithm under the fog-cloud architecture

CENTER FOR CONNECTED AND AUTOMATED TRANSPORTATION

Two main stages

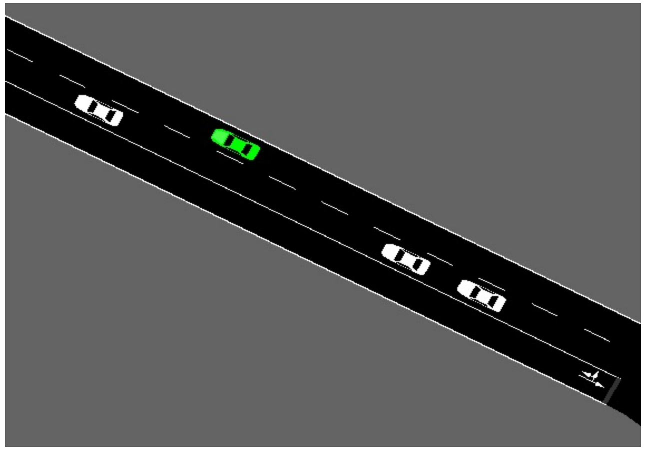
- DRL stage
  - Road network with fog paradigm is modeled as a graph (nodes are the fog nodes)
  - Road indexes for different fog regions are generated through Graph Attention Q-learning(GAQ)
- Routes assigned stage
  - Road weights are calculated by road indexes and road density
  - K alternative shortest routes are calculated through the road weights
  - Entropy balanced method is applied to assign the appropriate route to each vehicle considering the vehicles' priority and routes' popularity





Two types of vehicle

- Rerouting Vehicle (RV)
  - Rerouted by implementing the V2I technology
  - Colored green
- Background Vehicle (BV)
  - Not rerouted but incorporated to add randomness and dynamics to the network
  - Colored white



Agent

- Fog nodes are considered as agents
- At each time step t, agent i choose action a_t^i

State space

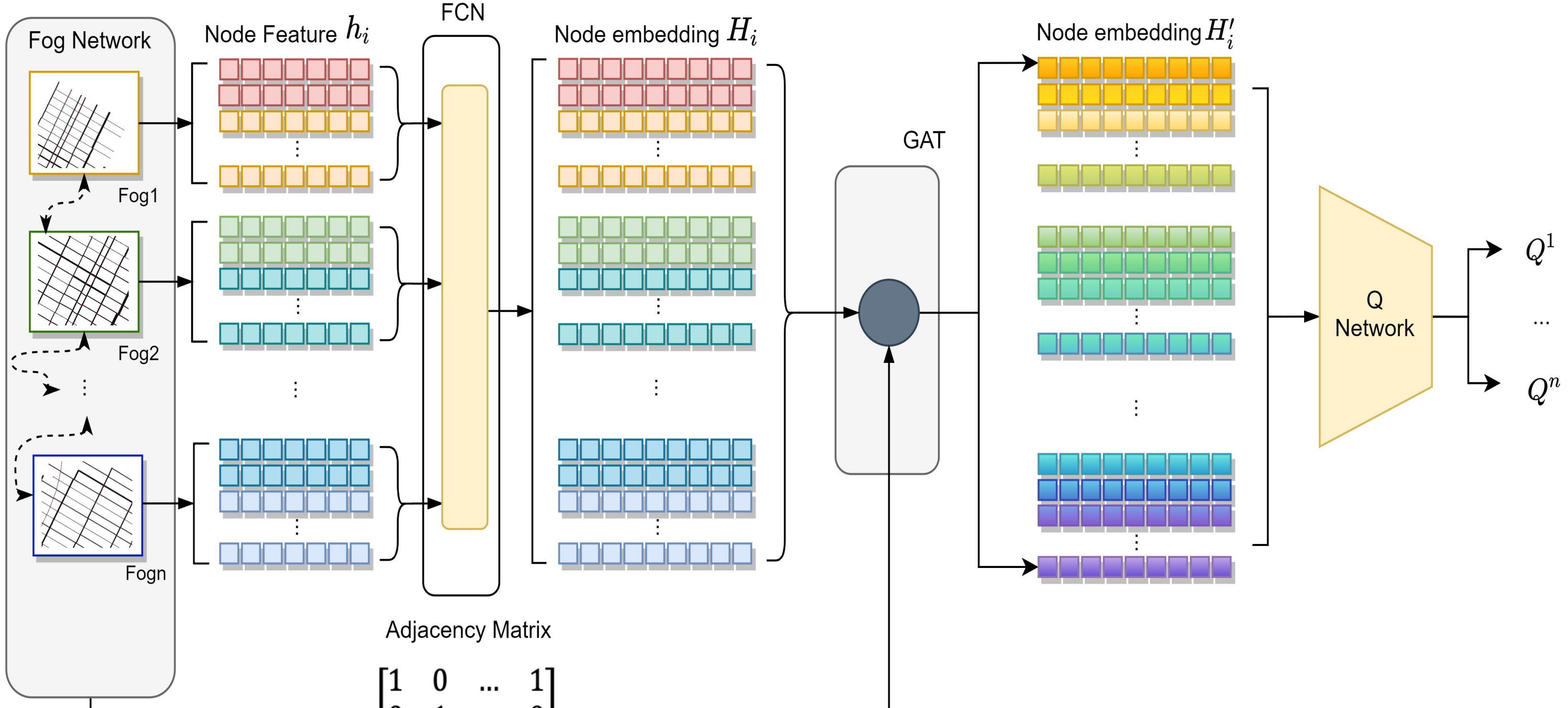
- Node feature at time step t: X_t
  - Average speed $\overline{v}_i$
  - Congestion condition of fog node area i c_i
- Adjacency matrix at time step t: A_t
  - Information topology and dependency of the fog nodes/fog agents

Action space

- Potential road index for fog nodes area
  - Discrete action space $a_i = \{0, 1, 2, 3, 4\}$

Reward function

- Reward
  - Speed increase compares to 11mph
- Penalty
  - Speed decrease compares to 11mph



$$\begin{bmatrix} 1 & 0 & \dots & 1 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & \cdots & 1 \end{bmatrix}$$

FCN encoder φ

Dense(32) + Dense(32)

GAT layer

GATConv(32)

Q network ρ

Dense(32) + Dense(32) + Dense(32) + Dense(64)

Output layer

Dense(5)

Two important factors

Vehicle priority

RVs' priority set $\mathcal{P}$ is obtained by the distance between their current location and the destination (D_{RV_i}) : $\mathcal{P} = (D_{RV_i}, RV_i)$

Vehicle that's nearer to the destination gets higher priority

Routes popularity

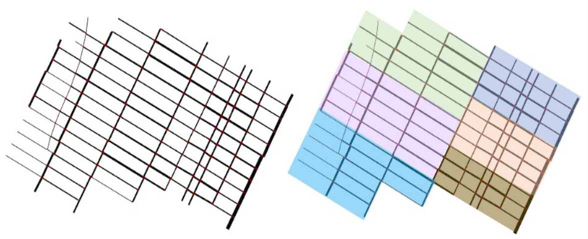
$$Pop(r_j) = e^{E(r_j)} \text{ with } E(r_j) = -\sum_{i=1}^{n} \left(\frac{fc_j^i}{N_{r_j}}\right) \ln\left(\frac{fc_j^i}{N_{r_j}}\right)$$

The final assigned route is chosen from the K shortest alternative routes, which prevents the vehicle with the final assigned route (the least popular route) from an excessively lengthy detour.

Simulator parameter

Network features:

- $5.926km^2$ area is extracted from the Manhattan area is used in this research, the network include 287 edges (roads) and 120 nodes (junctions).
- multiple road types in the network: 2-lane roads, 3-lane roads, 6-lane roads and 7-lane roads
- Each fog node covers about 50 edges (roads)



Scenario parameters:

- BVs (colored white) enter to the map from 8 roads
- RVs (colored green) enter to the map from the 3 roads
- BVs and RVs share a same destination and are randomly generated from the left side of the network

Vehicle control parameters

- BVs and RVs use SUMO’s built-in car-following and lane-changing controllers
- routing controller for RVs is based on the proposed GAQ-EBkSP model, while the routing controller for BVs is simply EBkSP.

Training parameter

Training steps

Total epochs trained: 1400

First 500 epochs as warm-up stage

Transition batches

transition batches of size 32 are sampled and put into the model

Optimization parameters

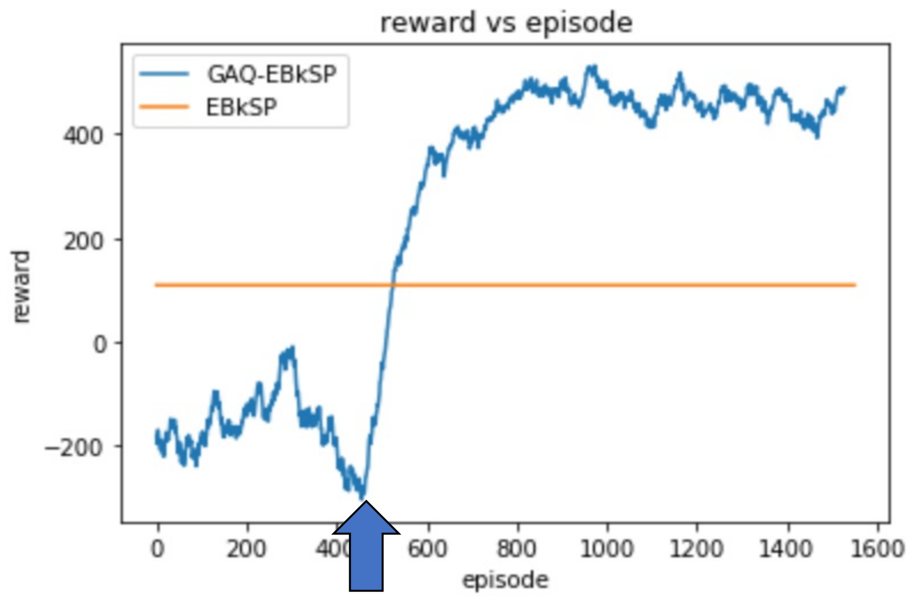
The optimization parameters used in this research is Adam, which has initial learning rate $\gamma = 10^{-4}$

Baseline model

In order to compare the performance with/without the RL training, the route assignment is also done using EBkSP

Training curve

First 500 episodes are random fully exploration



Training start

Two performance metrics

RVs' Average speed

RVs' Average travel time

Different scenarios

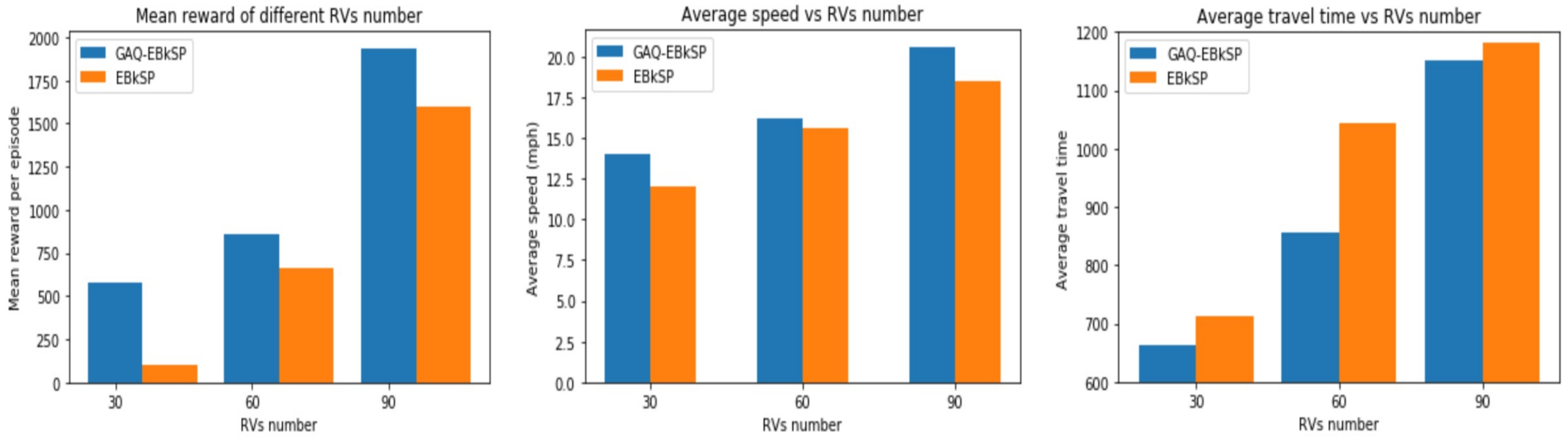
Trained in 30 RV scenario

Test from 30 RV to 90RV

RV number increase → higher average speed/ higher reward

RV number increase → longer average travel time

GAQ-EBkSP > EBkSP



Conclusion

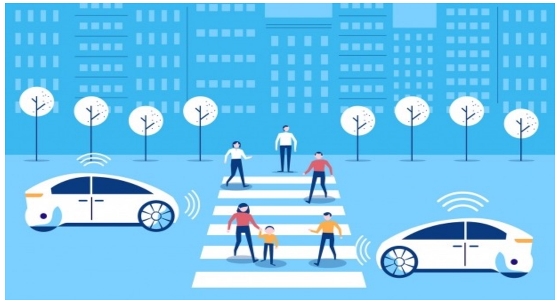
- DRL(GAQ)-EBkSP model based-on fog-cloud architecture is proposed to dynamically reroute the vehicles in large transportation networks
  - graph attention mechanism to fuse information and extract relevant information to enlarge the learning efficiency
- Proposed model outperforms baseline model (without learning) under different scenarios with different RV numbers
- High efficiency of the RVs in the network is achieved by the trained model

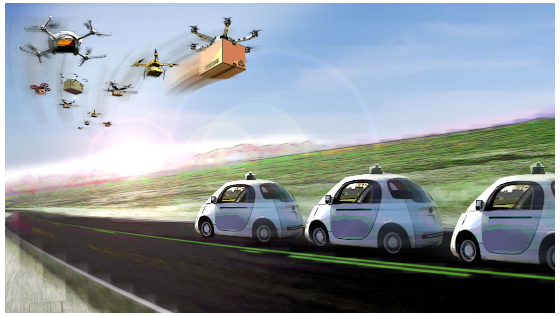
Future work

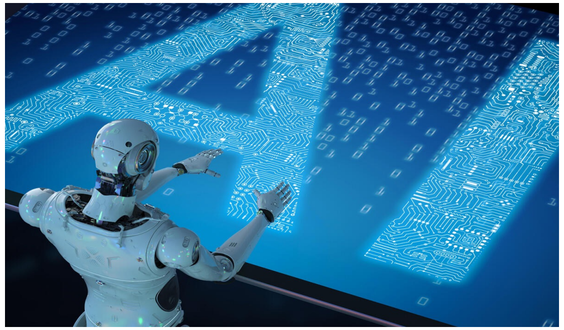
- Apply different priority standard to the framework
- Include different rerouting ratio (RV/BV) in the testing stage
- Use other similar RL-based models (LSTM-Q, GCQ) as the baseline models

PURDUE UNIVERSITY

CCAT
CENTER FOR CONNECTED AND AUTOMATED TRANSPORTATION







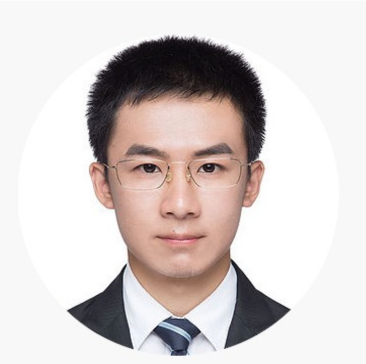
Purdue CCAT Director

Dr. Samuel Labi

Sky-Lab: AI & CAV Enabling Technologies

Dr. Sikai (Sky) Chen

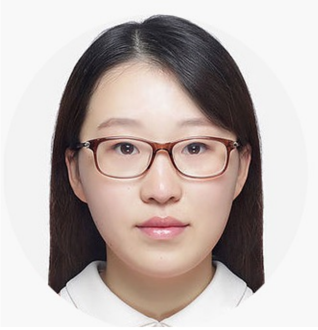
Sky
AI & CAV
Enabling Technologies



Jiqian Dong

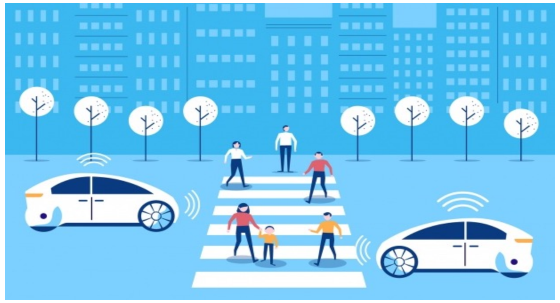
Rayne Du

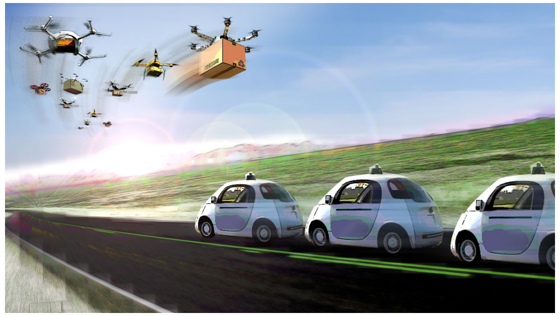
Paul (Young Joun) Ha

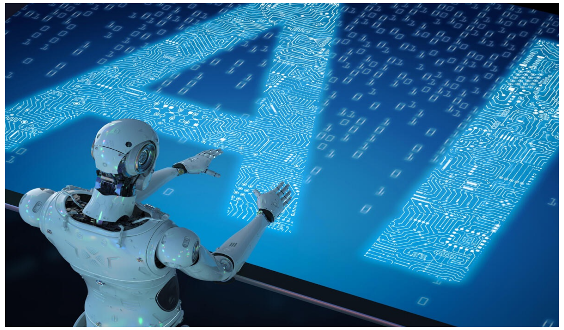


Yujie Li

Shuya Zong







QUESTIONES

COMMENTS

Sikai (Sky) Chen

Email: chen1670@purdue.edu

sikaichen@cmu.edu

Runjia Du

Email: du187@purdue.edu

SKY-Lab: https://www.sky-ai-cav.com/