JOHNS HOPKINS
INSTITUTE for ASSURED AUTONOMY

Assuring City Scale Infrastructure Systems

PI: Yair Amir, WSE

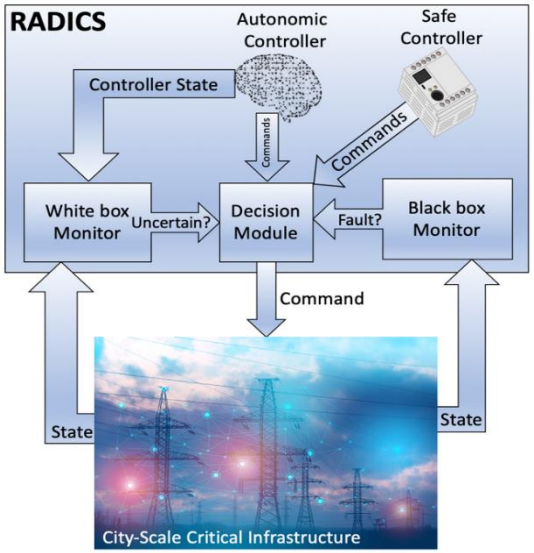
PI: Tamim Sookoor, APL

Total Budget: $750K (2 years)

Date: 06/25/2020

Project Summary

- AI systems are optimized for the average case. They cannot be used in critical systems that need to guarantee safety in worst case scenarios. The problem with AI systems is the long tail of edge cases that lead to failure situations. We want to gain the benefits of AI on the average case without incurring failures due to the long tail edge cases.
- Reinforcement learning (RL) algorithms are difficult to reason about and have non-intuitive behavior. In addition to challenges associated with deep learning, RL algorithms break in nonintuitive ways due to phenomena such as reward hacking and specification gaming.
- Runtime Assurance of Distributed Intelligent Control Systems (RADICS) combines an invariant-based Black-Box Monitor with a White-Box Monitor that evaluates the confidence of the machine learning algorithm. The autonomous system is assured through these two monitors.



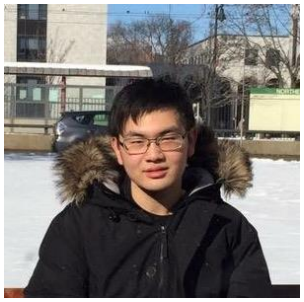
Project Team

Tamim Sookoor (PI)

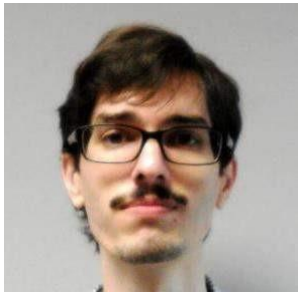
Yair Amir (PI)

Brian Wheatman
Traffic Control Testbed (L)
Blackbox Monitor (L)

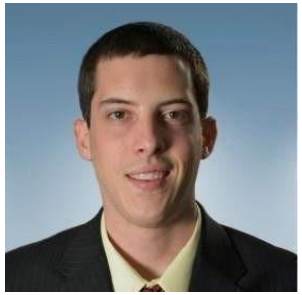
Sahiti Bommareddy
Smart Grid Testbed (L)



Jerry Chen
Traffic Control Testbed



Sebastian Zanlongo
Traffic Control Testbed (L)



Brad Potteiger
Traffic Control Testbed

Tim Krentz
Traffic Control Testbed
Smart Grid Testbed

Christina Selby
Whitebox Monitor (L)

Paul Wood
Whitebox Monitor

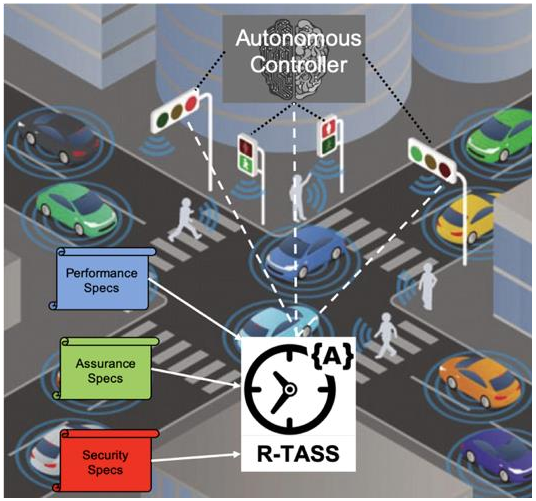
Nick Sarfaraz
Whitebox Monitor

Legend
WSE
APL
(L)ead

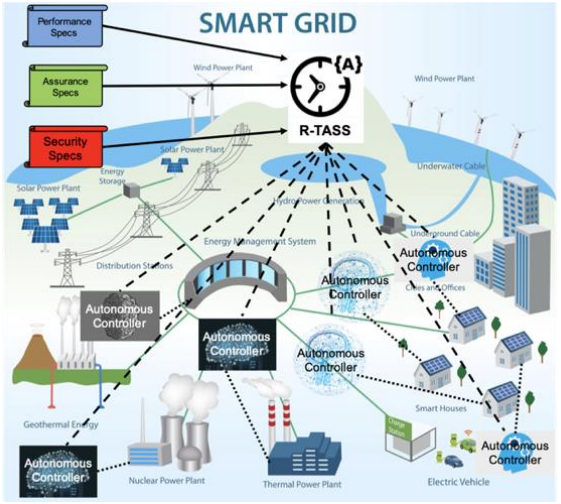
JOHNS HOPKINS
INSTITUTE for
ASSURED AUTONOMY

Flagship Use-case Projects

- Two ecosystem testbeds targeted at transportation and public safety domains



Transportation – Intelligent Traffic Control



Public Safety – Smart Power Grid

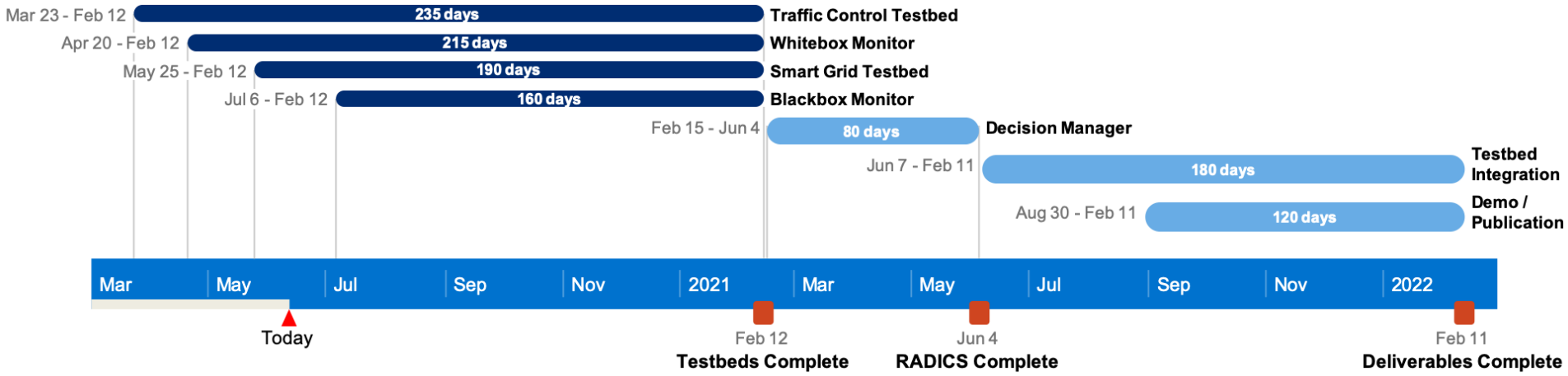
Testbeds could be used for transportation and public safety domains

Major Project Tasks

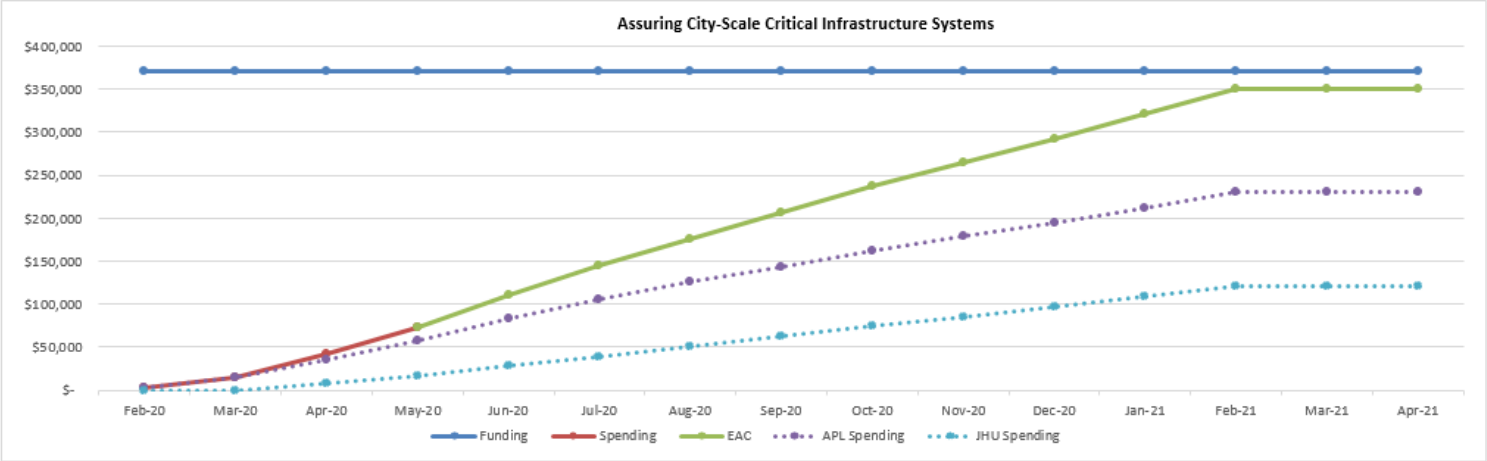
- Traffic Control Testbed (03/2020 – 02/2021)
  - Use the Simulation of Urban Mobility (SUMO) to model and simulate a city-scale highway network and investigate an autonomic controller (AC) and safe controller (SC)
  - Accomplishments:
    - Implemented testbed using SUMO and the FLOW deep reinforcement learning framework
    - Implemented a timer-based controller to act as a proxy for the safe controller
    - Trained a preliminary reinforcement learning (RL) based controller that outperforms the safe controller
    - Currently only a limited setup in terms of generality and realism
- Whitebox Monitor (04/2020 – 02/2021)
  - The whitebox monitor evaluates the confidence of the RL algorithms making control decisions
  - Accomplishments:
    - Investigating a number of potential approaches to assess RL confidence and estimate its competence:
      - Ensemble-based approach to identify uncertainty (made some progress already)
      - Predictor of the reward given the current state
      - Unscented transform-based approach to identify sensitivity to dataset shift
- Smartgrid Testbed (05/2020 – 02/2021)
  - Extend the Spire system, a SCADA system that actually operated within a power company grid (www.dsn.jhu.edu/spire), in order to model a city-scale energy utility
  - Currently investigating a number of economic dispatch simulators
- Blackbox Monitor (07/2020 – 02/2021)
  - Invariant-based monitor that detects faults at the system level by building on the vast knowledge of how traditional non-autonomous systems have been assured over the last five decades
  - Preliminary proof of concept demonstration in the traffic control testbed

Task Progress

Status as of 06/25/2020



Budgets: Planned vs. Actual



		Feb-20	Mar-20	Apr-20	May-20	Jun-20	Jul-20	Aug-20	Sep-20	Oct-20	Nov-20	Dec-20	Jan-21	Feb-21	Mar-21	Apr-21
	Funding	$ 370,114	$ 370,114	$ 370,114	$ 370,114	$ 370,114	$ 370,114	$ 370,114	$ 370,114	$ 370,114	$ 370,114	$ 370,114	$ 370,114	$ 370,114	$ 370,114	$ 370,114
	Spending	$ 3,037	$ 15,549	$ 43,120	$ 73,612											
Cumulative	EAC				$ 73,612	$ 110,650	$ 144,667	$ 176,543	$ 205,770	$ 236,627	$ 264,461	$ 291,932	$ 320,854	$ 350,381	$ 350,381	$ 350,381

		Feb-20	Mar-20	Apr-20	May-20	Jun-20	Jul-20	Aug-20	Sep-20	Oct-20	Nov-20	Dec-20	Jan-21	Feb-21	Mar-21	Apr-21
	Funding	$ 249,880	$ 249,880	$ 249,880	$ 249,880	$ 249,880	$ 249,880	$ 249,880	$ 249,880	$ 249,880	$ 249,880	$ 249,880	$ 249,880	$ 249,880	$ 249,880	$ 249,880
	APL Spending	$ 3,037	$ 15,549	$ 34,810	$ 56,992	$ 82,518	$ 105,021	$ 125,385	$ 143,100	$ 162,444	$ 178,765	$ 194,724	$ 212,133	$ 230,147	$ 230,147	$ 230,147
	Actuals	$ 3,037	$ 12,512	$ 19,261	$ 22,182	$ -	$ -	$ -	$ -	$ -	$ -	$ -	$ -	$ -	$ -	$ -
APL FFPB0003	Plan	$ -	$ 22,331	$ 22,331	$ 23,587	$ 25,526	$ 22,503	$ 20,364	$ 17,714	$ 19,344	$ 16,321	$ 15,959	$ 17,409	$ 18,014	$ -	$ -

		Feb-20	Mar-20	Apr-20	May-20	Jun-20	Jul-20	Aug-20	Sep-20	Oct-20	Nov-20	Dec-20	Jan-21	Feb-21	Mar-21	Apr-21
	Funding	$ 120,234	$ 120,234	$ 120,234	$ 120,234	$ 120,234	$ 120,234	$ 120,234	$ 120,234	$ 120,234	$ 120,234	$ 120,234	$ 120,234	$ 120,234	$ 120,234	$ 120,234
	JHU Spending	$ -	$ -	$ 8,310	$ 16,620	$ 28,133	$ 39,645	$ 51,158	$ 62,671	$ 74,183	$ 85,696	$ 97,209	$ 108,721	$ 120,234	$ 120,234	$ 120,234
JHU Project ID	Actuals	$ -	$ -	$ 8,310	$ 8,310	$ -	$ -	$ -	$ -	$ -	$ -	$ -	$ -	$ -	$ -	$ -
80052269	Plan					$ 11,513	$ 11,513	$ 11,513	$ 11,513	$ 11,513	$ 11,513	$ 11,513	$ 11,513	$ 11,513		

JOHNS HOPKINS
INSTITUTE for
ASSURED AUTONOMY

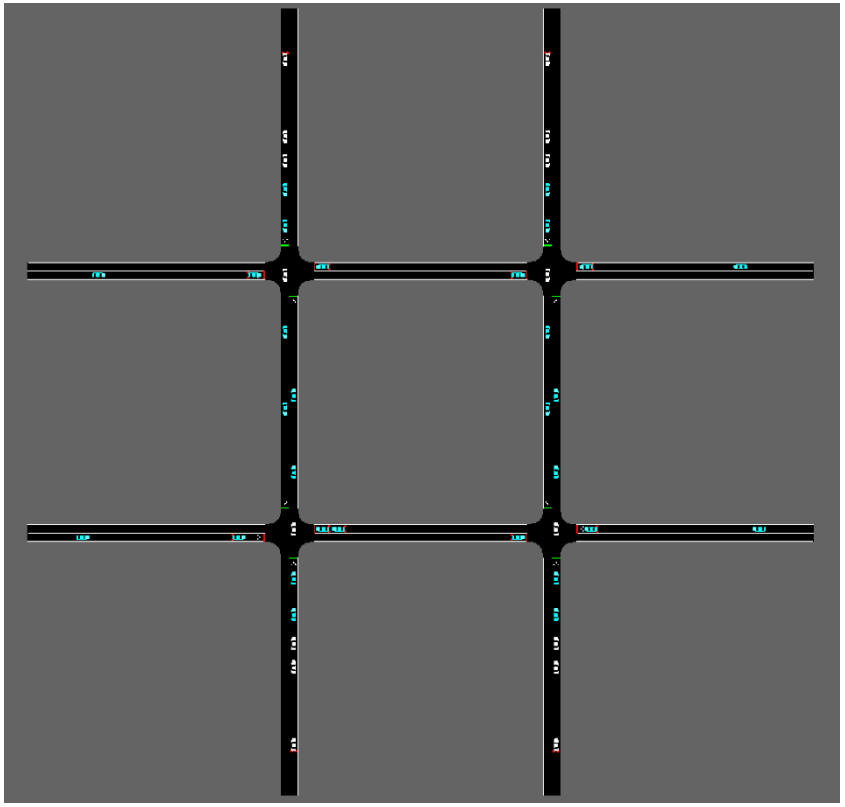
Traffic Control Testbed

- Definitions
  - Simulation Step: simulated 0.1s
  - Iteration: 3000 simulation steps, which is 300s.
  - Traffic Inflow Rate: 500 vehicles per hour on every outside edge. We have 8 outside edges in the traffic grid. Therefore, we have a total of 4000 vehicles per hour in the whole system.
- Benchmark
  - Safe Controller: Average waiting time is 9.97 seconds.
  - RL Model trained with 500 vehicles per hour on each outside edge:

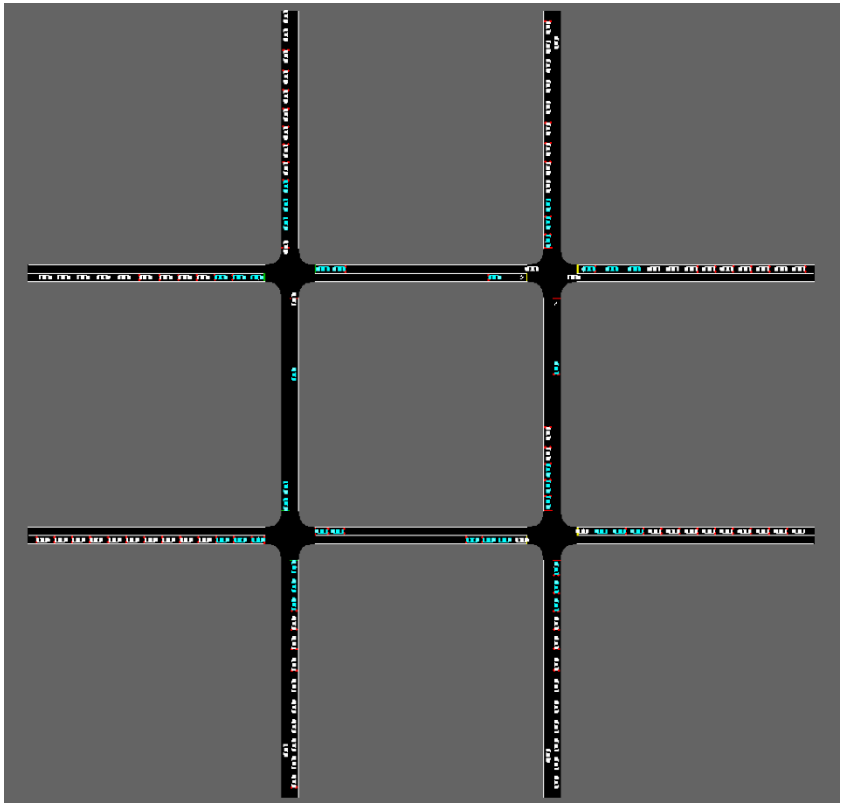
Number of Steps / Iterations	Average Waiting Time (s)
50K / 16	37.88
1M / 333	15.55
3M / 1,000	5.17
10M / 3,333	3.63
30M / 10,000	0.29

Traffic Control Testbed – Demo

Safe Controller

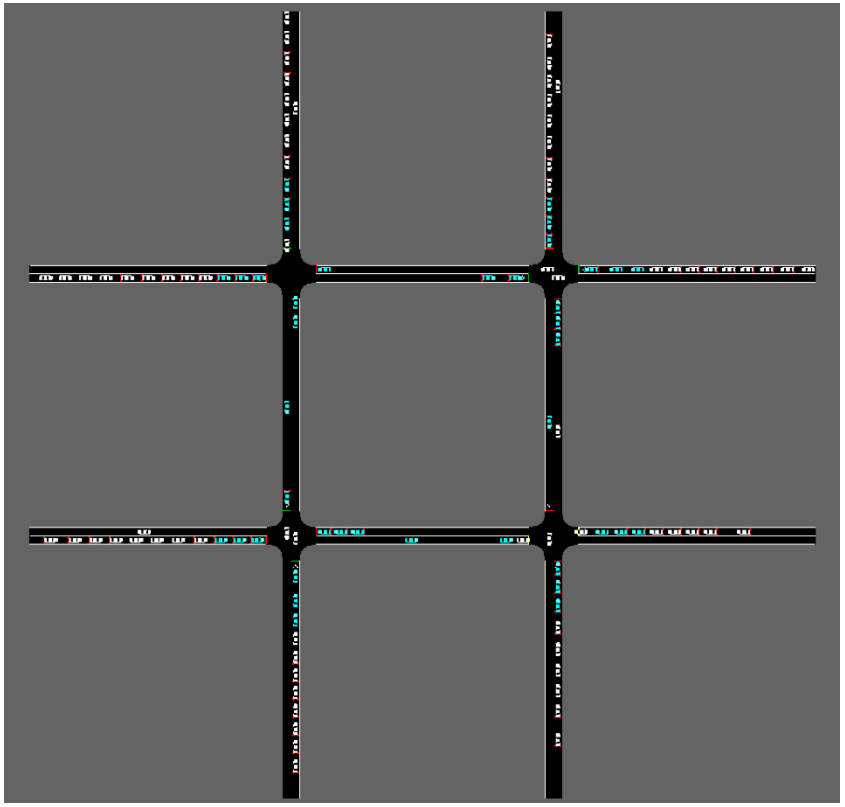


RL Model after 50K steps / 16 iterations

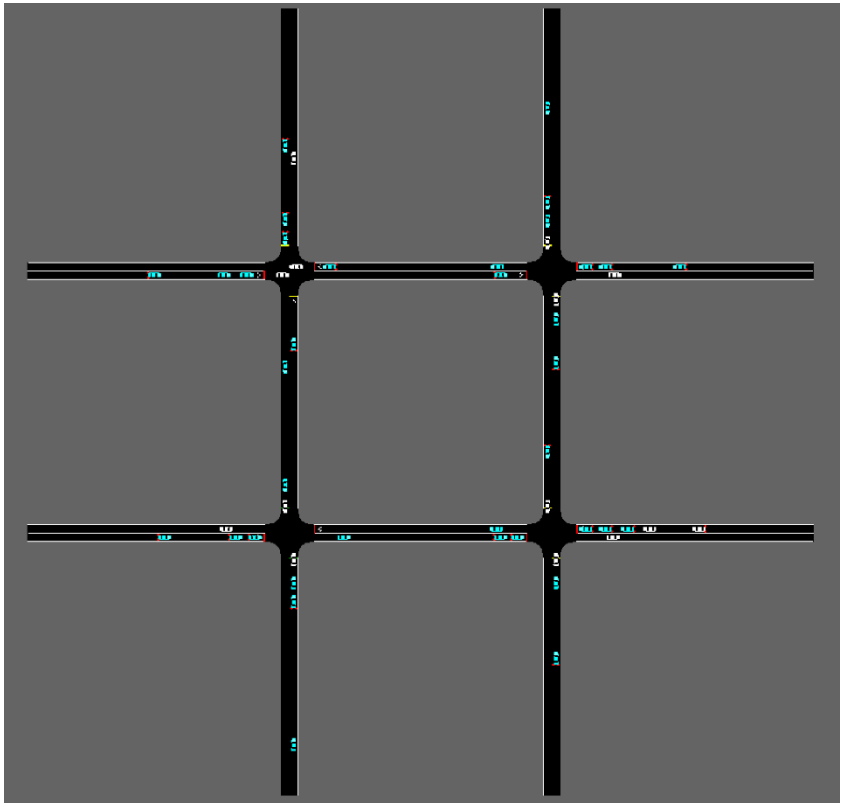


Traffic Control Testbed – Demo

RL Model after 1M steps / 333 iterations

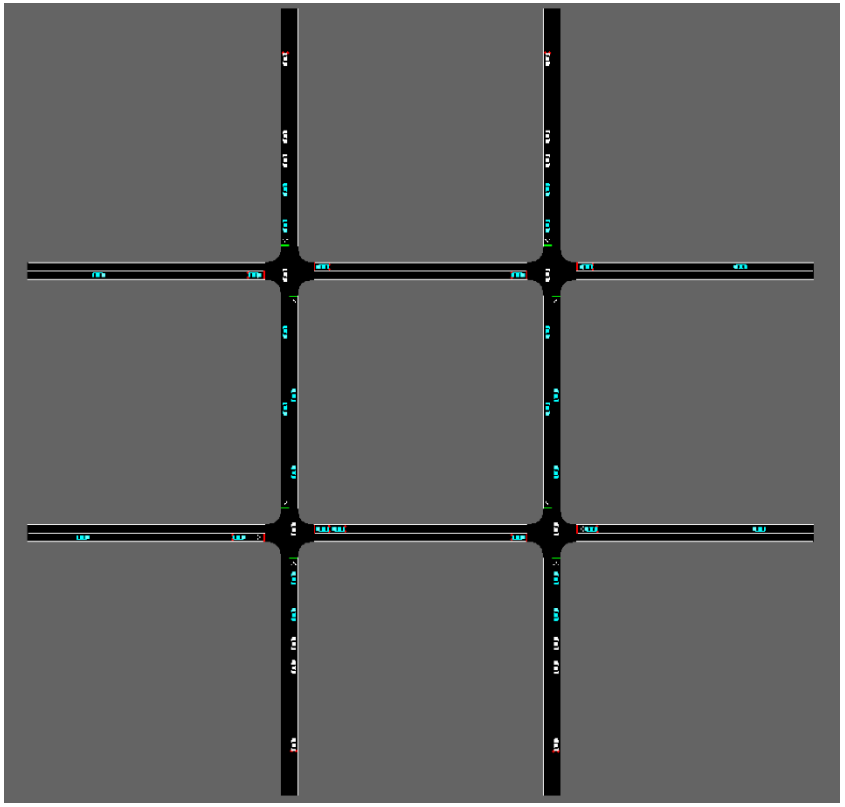


RL Model after 3M steps / 1000 iterations

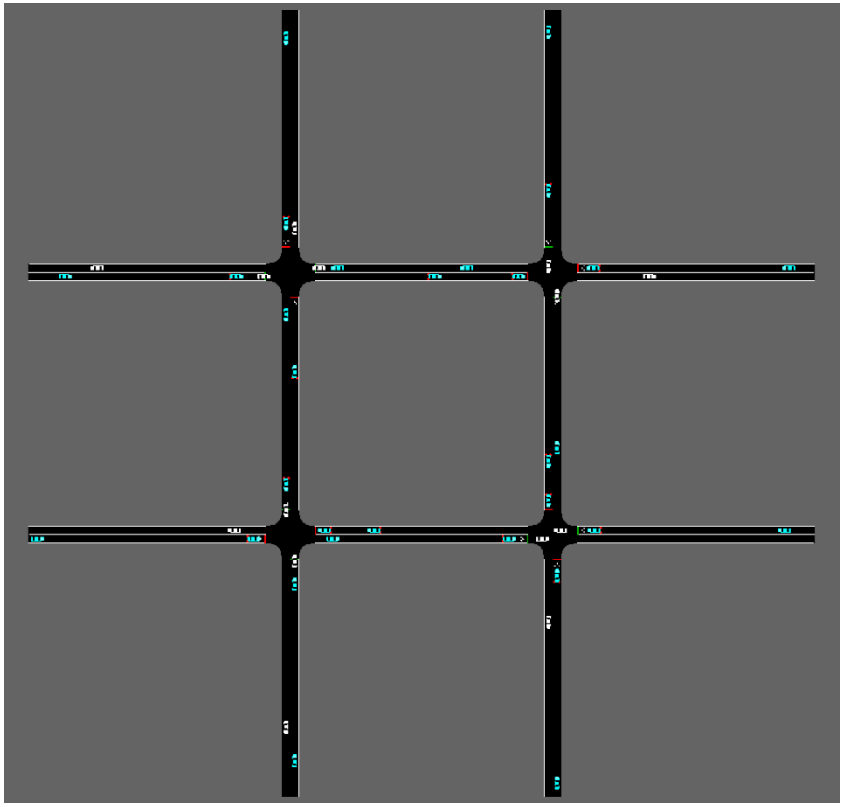


Traffic Control Testbed – Demo

Safe Controller

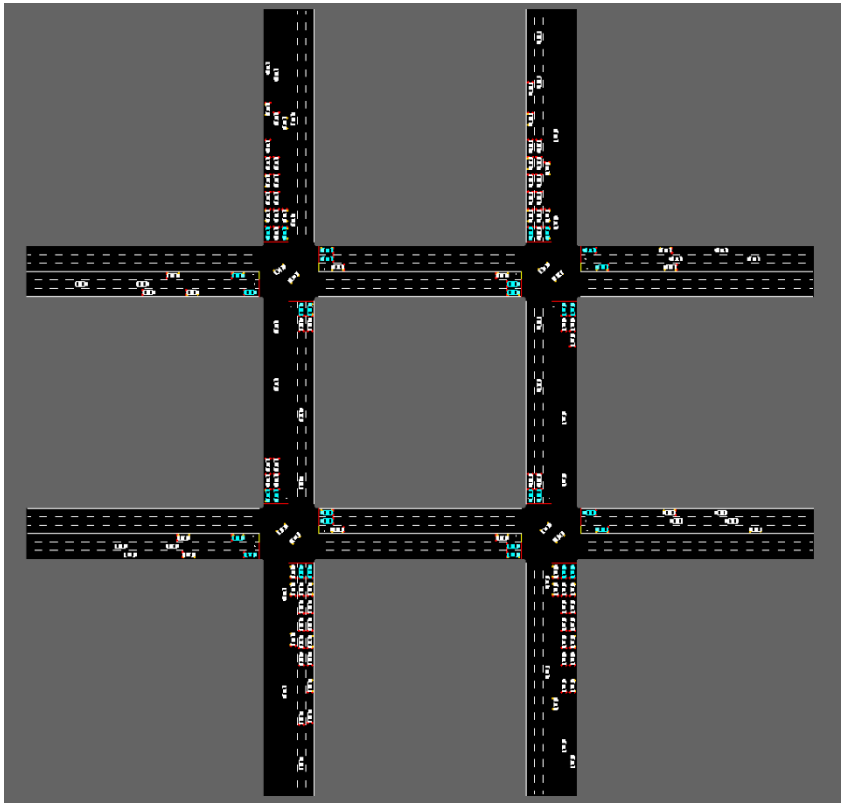


RL Model after 30M steps / 10000 iterations



Traffic Control Testbed – Demo

A More Realistic Traffic Grid



Smart Grid Testbed – Demo

Spire

- Spire is an open-source intrusion-tolerant SCADA system for the power grid. Spire includes a SCADA Master and a PLC proxy designed from scratch to support intrusion tolerance, as well as a Human Machine Interaction (MHI) based on pvbrowser. Spire emulates power grid management under a number of distribution and generation scenarios
- We plan to extend Spire to include an Economic Dispatch module that predicts the demand for power and optimizes the cost of power generation to meet the predicted demand. We will design a Simplex architecture with the corresponding autonomous controller, black box and white box monitors, and decision module

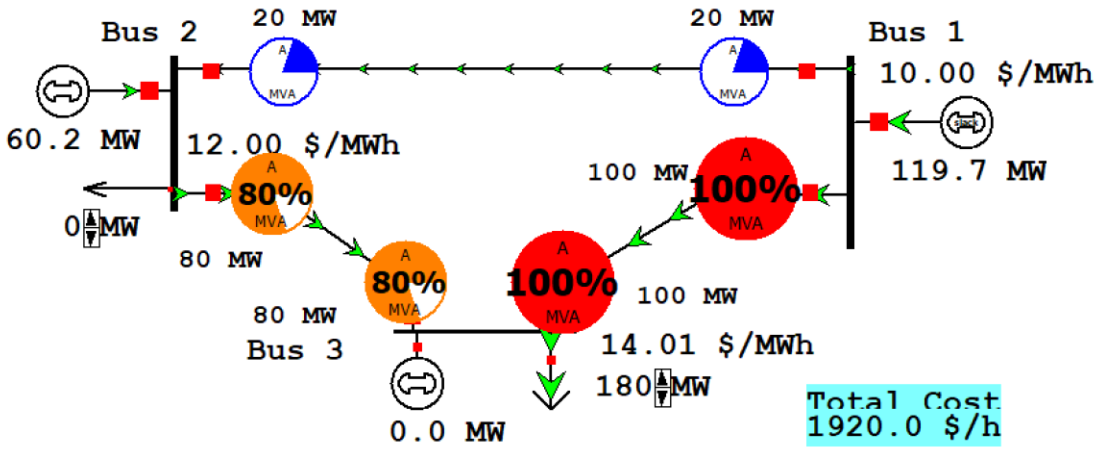
Smart Grid Testbed – Demo

Economic Dispatch Simulators

Currently exploring economic dispatch simulators to realistically model power generation systems

Simulators under consideration:

- Powerworld - https://www.powerworld.com/
- GridLAB-D - https://www.gridlabd.org/
- Etap - https://etap.com/

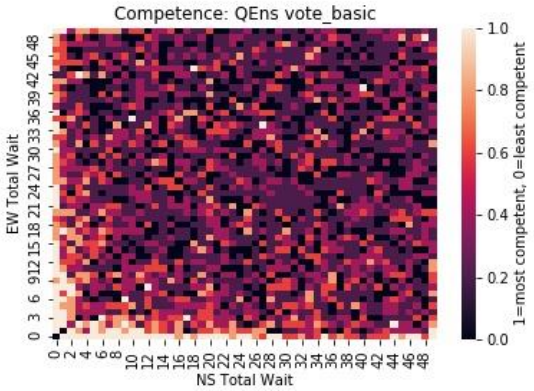


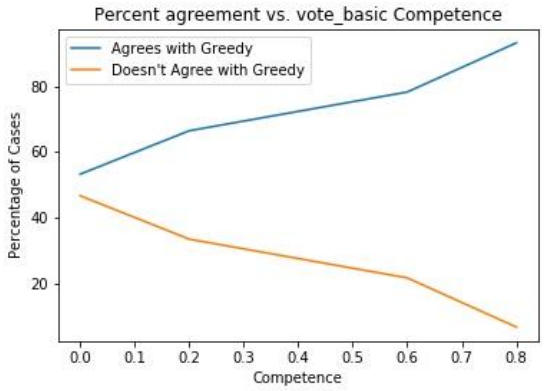
White Box Monitor Progress

- A two-queue simulation with known optimal solution for minimizing total wait time has been used to develop and validate the use of ensembling approaches for Q-table policies.
- This methodology is being extended to scenarios for which an optimal solution is not known by analysis of correlations between proposed competence metrics and performance-based metrics.
- The methodology is also being extended to reinforcement learning policies that utilize deep neural networks.

Optimal (Greedy) Policy

Example Q-Ensemble Policy





As desired, agreement with optimal policy increases with competence estimate.

Insights and Challenges

- Insights
  - RL algorithms that work well on a few intersections may not scale well to many intersections
    - The state space for traffic is actually very large (#lights, #light_states, #lanes, #sensor_types, vehicle_information, traffic_rules: turn or not, type of turn, etc.)
    - Long training times necessary to properly train model
    - Exploring redefining state space to allow transfer learning between intersections
  - Challenges and opportunities with centralized vs decentralized control of traffic light network
    - Single vs multi-agent RL as possible solutions
  - There could be situations from which recovery would be difficult if model has not been sufficiently trained
  - Reverse engineering Flow
    - Flow does not do what we thought it did: interface to openAI Gym and not specialized SUMO models

- Challenges
  - Training a generalizable RL traffic controller – Trade-off between generalizability and end-to-end-solution
  - Making the model more realistic (adding turns, turn lanes, etc.)
  - Definition of RL competence/safety – Investigating multiple safety metrics to identify a usable subset
  - Building scenarios and models simple enough to experiment on, but now too far removed from reality

Publication Status

- Reinforcement Learning Testbed for Prototyping and Evaluating Intelligent Traffic Control
  - With the increase in embedded computing power and integrations with cloud resources, it is now easier than ever to perform sophisticated computations at the edge. As such, traditional Cyber-Physical Systems such as traffic controllers are moving away from fixed time algorithms to innovative AI and Machine Learning based approaches. By leveraging techniques such as Reinforcement Learning within next-generation traffic controllers, city managers can optimize traffic flows continuously through online and offline learning, resulting in a decrease in backups, increased resilience to cyber-attacks and crashes, and minimization of emergency vehicle response times. However, as currently illustrated with the rise of autonomous vehicles, it is imperative to conduct an extensive evaluation to guarantee correct and safe behavior when introducing novel technology into safety-critical infrastructure. This talk focuses on the development of our reinforcement learning testbed for the rapid prototyping and evaluation of intelligent traffic control algorithms within smart cities. Further, we discuss how metrics collected through simulation experiments can be leveraged within a broader context to develop a novel Simplex based architecture for assuring the safety, security, and resiliency of urban traffic systems.
  - APL AI Workshop

- Runtime Assurance of Reinforcement Learning Algorithms
  - The integration of autonomic controllers into safety-critical cyber-physical systems requires the ability to estimate the competence of a given reinforcement learning algorithm. A general method for competence estimation of reinforcement learning policies does not exist. This talk will discuss progress made on this problem as part of the Institute for Advanced Autonomy's "Assuring City Scale Infrastructure Systems" IRAD. A critical task of this project is the development and use of a "White Box Monitor" in a reinforcement learning testbed for prototyping and evaluating intelligent traffic control.
  - APL AI Workshop

External Collaboration Status

- Partners for Automated Vehicle Education (PAVE)
  - Brad Potteiger is a member of the Academic Advisory Council
  - Details: TBD
  - Potential funding if collaborator is a potential funding source: N/A
- National Institute for Standards and Technology (NIST)
  - Preliminary discussions with Keith Stouffer and Chee Tang of the Railway Infrastructure lab on opportunities for collaboration
  - Details: TBD
  - Potential funding if collaborator is a potential funding source: N/A
- DOE Grid Modernization Lab Consortium (GMLC) Byzantine Security
  - New project started at the DSN lab led by Pacific Northwest National Lab (PNNL)
  - Key participants include LBL and SNL national labs as well as industry manufacturers (ABB, GE, Siemens) and power companies (HECO, PNM, WAPA)
  - Potential source of information about power grids, demand management and economic dispatch
  - Perhaps potential source of future funding: TBD
- DARPA Assured Autonomy Program
  - Preliminary discussions with Dr. Sandeep Neema (PM) on opportunities for collaboration
  - Details: TBD
  - Potential funding if collaborator is a potential funding source: TBD

Broader Impacts

- Applications
  - Unmanned Aerial Systems
  - Power Grid
  - Industrial Control Systems
  - Railway
- IAA Collaboration
  - Physical Adversarial Machine Learning – HIL Integration

JOHNS HOPKINS

INSTITUTE *for*
ASSURED AUTONOMY