Traffic Control Systems

Using Reinforcement Learning

GROUP 12
Rohan Mistry
Harry Leung
Derek Xu

AGENDA

01

PROBLEM

...

02

EXISTING SOLUTIONS

...

03

APPROACH

...

04

CONCLUSION

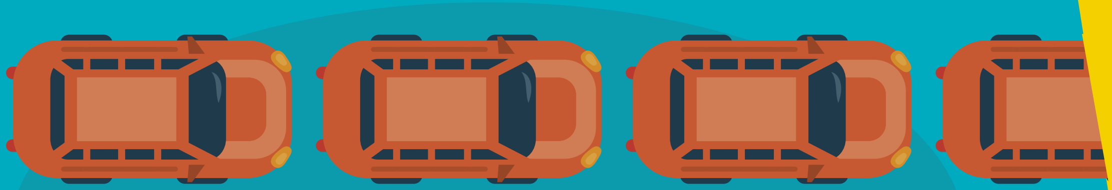
...

PROBLEM

- Traffic congestion is a major issue affecting urban areas worldwide, leading to increased travel times, air pollution, and commuter frustration
- Occurs often when road demand exceeds supply, particularly during peak hours

SIGNIFICANCE

- Far-reaching effects on economic efficiency, public health, and environmental sustainability
- Congested traffic systems contribute to wasted fuel and time, resulting in billions of dollars in economic losses



REAL-WORLD IMPACT

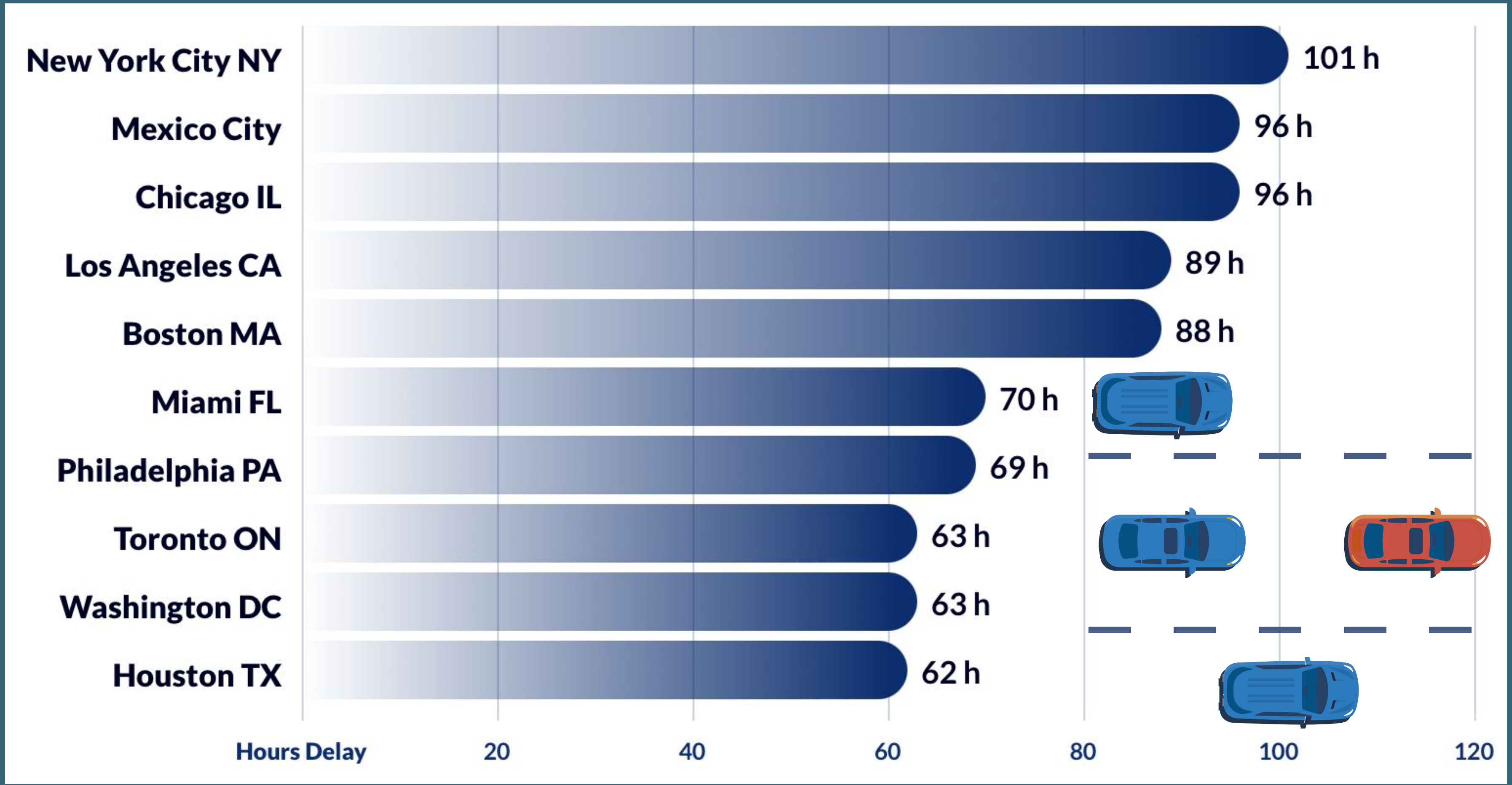
According to Petroski, traffic congestion and associated delays costs U.S. economy over $120 billion annually

U.S. traffic congestion resulted in average driver spending 51 hours annually in traffic, equating to nearly $1,000 in lost time and increased pollution

INCREASED VEHICLE EMISSIONS

AIR QUALITY DEGRADATION

CLIMATE CHANGE


New York City NY
101 h
Mexico City
96 h
Chicago IL
96 h
Los Angeles CA
89 h
Boston MA
88 h
Miami FL
70 h
Philadelphia PA
69 h
Toronto ON
63 h
Washington DC
63 h
Houston TX
62 h
Hours Delay
20
40
60
80
100
120

US Cities Congestion

View cities by:

- ○ 2023 Congestion Delay
- ◉ % Change from 2022
  - Increased Congestion %
  - Decreased Congestion %

EXISTING SOLUTIONS

FIXED TRAFFIC LIGHT TIMINGS

Pre-defined, static schedules based on historical data, timings set for different times of day and traffic patterns, no real-time adjustments

Simple to implement, cost-effective

Inflexible, unadaptable

ADAPTIVE SIGNAL CONTROL SYSTEMS

Dynamically adjusts timings based on traffic data collected from sensors

Adapts to traffic conditions

Does not learn in real time

VEHICLE – TO – INFRASTRUCTURE

Direct communication between vehicles and traffic infrastructure (vehicles share speed, location, destination, etc.)

Real-time adjustments based on vehicle data

Relies on vehicle compliance, not universal

HIGHWAY INFRASTRUCTURE INVESTMENTS

Investing in new roads and highways via government fundings

Expensive, inefficient, ineffective

PREVIOUS APPROACHES

DEEP Q-NETWORKS (DQN)

APPROACH (study by Van der Pol & Oliehoek, 2016):

- Applied by representing intersections as agents that learn optimal traffic signal policies
- Each intersection makes decisions based on current traffic conditions, aiming to reduce average waiting time for vehicles
- Promising results in reducing delays at intersections
- Allows system to approximate complex policies and learn directly from high-dimensional inputs, like real-time traffic flow data

LIMITATIONS:

- Struggles in highly dynamic environments due to issues with stability and convergence, especially if traffic patterns change frequently

GRAPH NEURAL NETWORKS (GNN)

APPROACH (study by Li et al., 2020):

- Models traffic network as graph where intersections are nodes and roads are edges
- By combining GNNs with RL, each node leverages graph structure to make better-informed decisions based on state of neighboring intersections
- Enables agents to generalize learned policies across different parts of traffic network
- Can learn spatial relationships within traffic grid

LIMITATIONS:

- Computationally intensive, especially in large urban networks
- Requires significant tuning to achieve stable training

PREVIOUS APPROACHES

ACTOR-CRITIC ALGORITHMS

APPROACH (study by Mohammad Aslani et al., 2017):

- Critic's feedback helps actor make better decisions over time, useful for traffic systems where optimal actions vary frequently
- By adjusting actor based on feedback from critic, algorithm can efficiently balance exploring new actions and exploiting known strategies to reduce traffic congestion
- Can adapt to changing traffic patterns over time

LIMITATIONS:

- Errors in value function estimation can cause instability which can lead to fluctuations in traffic control policies
- Convergence issues in non-stationary environments where conditions change often

PROXIMAL POLICY OPTIMIZATION (PPO)

APPROACH (study by Liang et al., 2019):

- Clipping mechanism useful for traffic environments prone to abrupt changes in traffic flow by restricting how much policy can change at each update
- Uses mini-batch training and reuses experience data across multiple updates
- Adjusts traffic light durations dynamically, handles fine-grained decisions such as how long to keep light green or red

LIMITATIONS:

- Requires large amounts of training data and significant computational resources for complex/large-scale traffic networks
- Not as flexible when using single agent approach

ENVIRONMENT

Can use either, both compatible with Gym for integrated environment

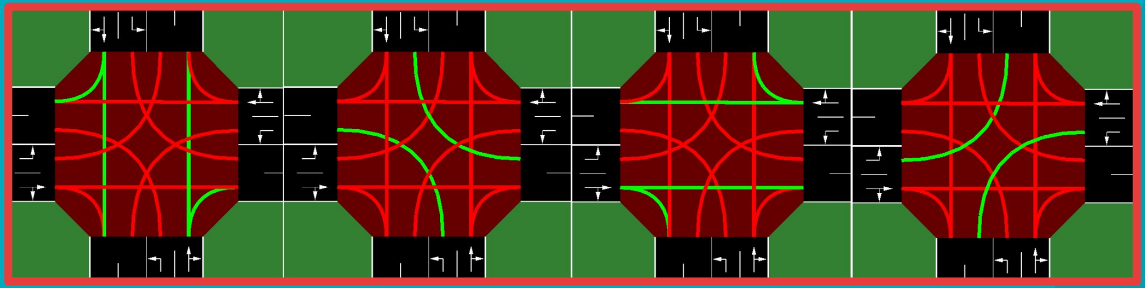
CITYFLOW	SUMO-RL WITH FLOW

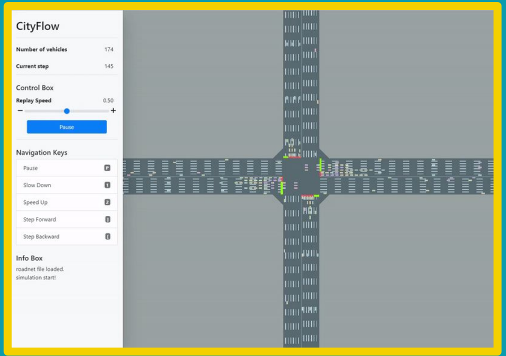
Two multi-agent RL environments for large-scale city traffic scenarios, with microscopic traffic simulator

Observation Space for each intersection: vehicle count, vehicle queue lengths, traffic light status/direction, waiting times, vehicle speed, arrival/departure rates

Action Space: switching traffic light status or direction, discrete actions

Reward Function Factors: queue length reduction, waiting time reduction, throughput maximization





APPROACH

Train algorithm to handle multiple independent agents representing interdependent intersections which coordinate with neighboring agents

ALGORITHM – MULTI AGENT PPO

Plan to implement Multi-Agent Proximal Policy Optimization (MAPPO) algorithm allowing for effective learning of complex policies in high-dimensional action spaces and learn optimal signal timings over time

Adapts PPO for multi-agent scenarios by treating each agent's policy separately while sharing experience or by using centralized training method

Can potentially use gradient-based updates and balancing exploration and exploitation to adjust signal timings dynamically and minimize vehicle delay/travel time, though this requires more computations and time

IMPLEMENTATION

1. Develop custom Gym environment simulating urban traffic
2. Train MAPPO agents using simulated traffic dataset which includes vehicle count, queue lengths, traffic light status, waiting time
3. Evaluate performance (waiting time reduction, throughput maximization) against existing traffic control methods

EXPECTATIONS

Anticipating improved throughput and reduced waiting time with coordination between intersections, will need several thousand training episodes for performance to stabilize

NOVELTY & CHALLENGES

NOVELTY

MARL allows for modeling each intersection/lane as individual agent which can coordinate/compete to optimize network-wide traffic flow

More flexible and higher potential for real-world applicability, built to handle scalability

Most implementations have focused on single agents

Field is still evolving with new multi-agent cooperative and competitive strategies, research into MARL for traffic applications still ongoing

Very well-suited for large-scale, multi-intersection scenarios compared to PPO, DQN, A3C, etc. which are typically implemented for single-agent settings

CHALLENGES

Complex setup because must design multiple agents that can either collaborate or compete

Requires additional mechanisms for inter-agent communication, cooperation protocols, and reward-sharing strategies

Training stability and computational cost are higher than simpler approaches because each agent is essentially running its own RL model that needs to interact with others

Reward tuning may be difficult

CONCLUSION

Proposed traffic control system uses multi-agent PPO for adaptive traffic management

Enables each intersection to act as coordinated agents and dynamically adjust signal timing based on real-time traffic flow

MAPPO algorithm: addresses limitations of static and single-agent systems; more scalable for complex urban networks

Expected benefits: reduced travel time and improved response to traffic variability

Potential to revolutionize traffic management in high population areas, enhancing overall economic efficiency and environmental sustainability