Social Behaviour Modeling and Optimization through Big Data and Reinforcement Learning

Alina Vereshchaka and Wen Dong

University at Buffalo
Department of Computer Science and Engineering

University at Buffalo
The State University of New York

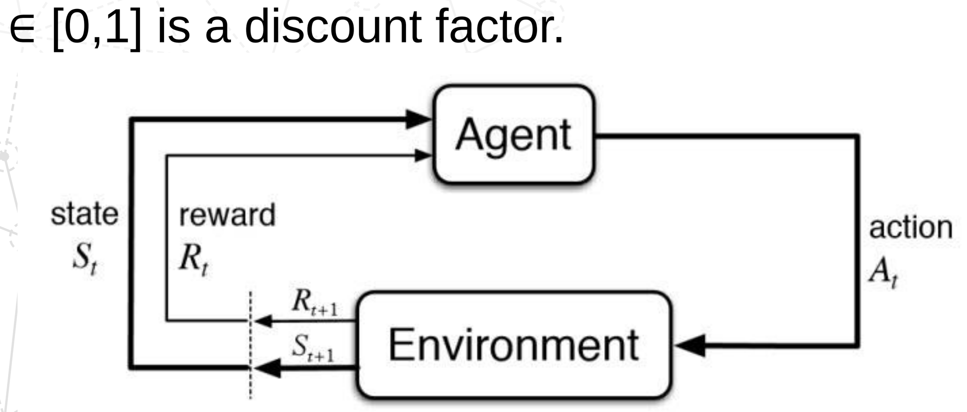
Overview

Solving complex real-world modeling and optimization problems presents several key challenges, including balancing the huge number of agents and system states, considering the complicated interactions between agents, and accounting for constant change in both the environment and the specification of optimality.
We propose two approaches to this problem that utilize big data and the reinforcement learning framework. These approaches can be applied to complex social behaviour problems, such as allocating limited resources during natural disaster events and optimizing dynamics in critical infrastructure.

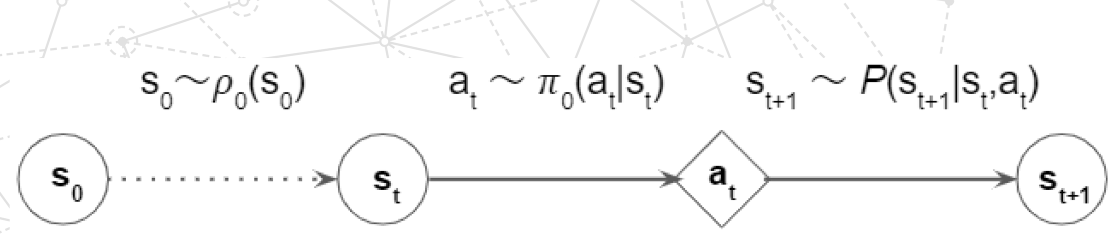
Methodology

Markov decision process (MDP), defined by the tuple ⟨*s, a, o,* P, *r*, ρ_0, γ⟩, where

- *s* ∈ S denotes states, describing the possible configurations of all agents;
- *a* ∈ A denotes actions, which can be discrete or continuous;
- P: S x A x S → R is the states transition probability distribution, where states evolve according to the stochastic dynamics $p(s_{t+1}|s_t,a_t)$;
- O is a set of observations for each agents;
- *r*: S → R is the reward function;
- ρ_0 : S → [0,1] is the distribution of the initial state;
- γ ∈ [0,1] is a discount factor.



Initial state s_0 is sampled from prior state visitation distribution $\rho_0(s_0)$; action a_t is sampled from the policy π_0 and the next state $s_{t+1}$is chosen based on the transition probability distribution $P(s_{t+1}|s_t, a_t)$.



Dynamic Resources Allocation During Natural Disasters

Natural disasters are important and devastating events for a country. According to the United Nations Office for Disaster Risk Reduction natural disasters are happening more frequently, which has caused a rise of 151% over the last twenty years in economic losses from climate-related disasters.

Our goal

Efficiently allocate resources during the natural disaster in real-time for mitigation efforts.

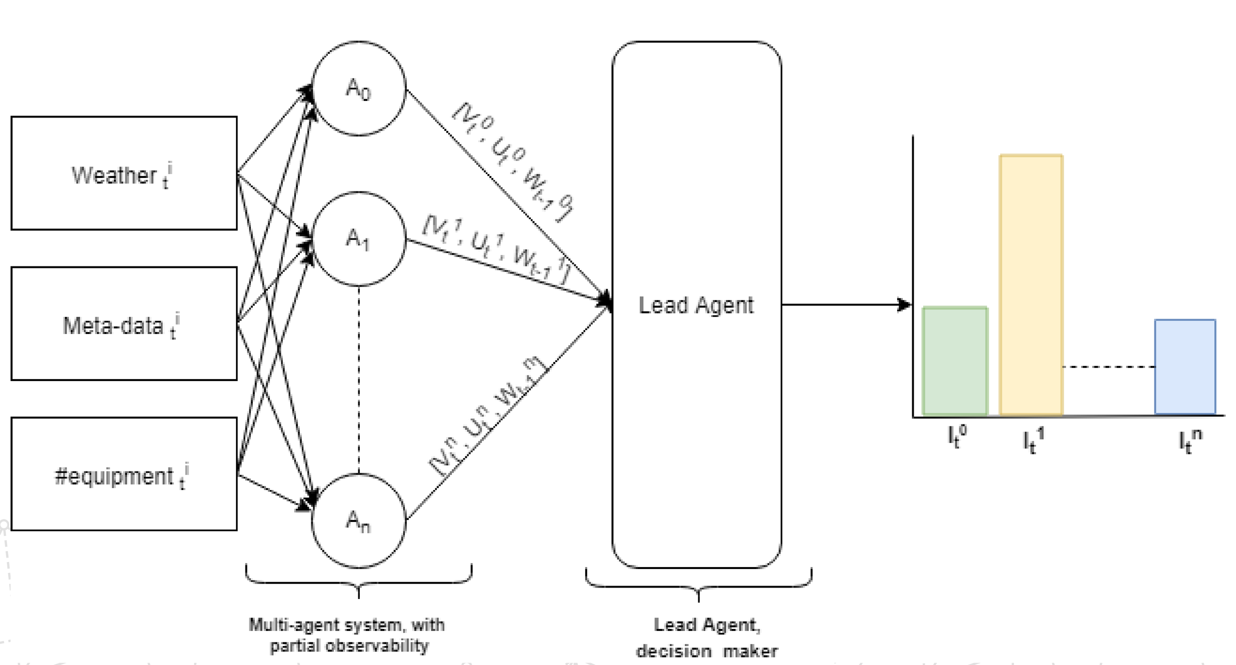
Our approach

We propose a hierarchal multi-agent real-time resource allocation framework to respond to the disaster scenarios, based on the reinforcement learning.

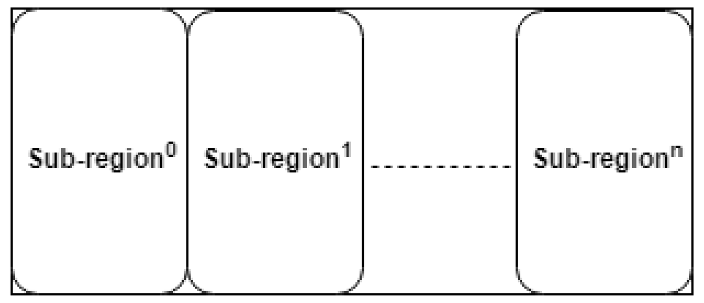
Our Solution

There are two levels in our framework.

- On the lower level (left side) there is a set of multi-agents (A_0,A_1,...,A_n). Each agent predicts the volume of hazards and the importance, taking as input an observation from the time-series environment.
- On the higher level (right side) there is a lead agent, which is trained to make allocation decisions for assigning volume distribution of resources.



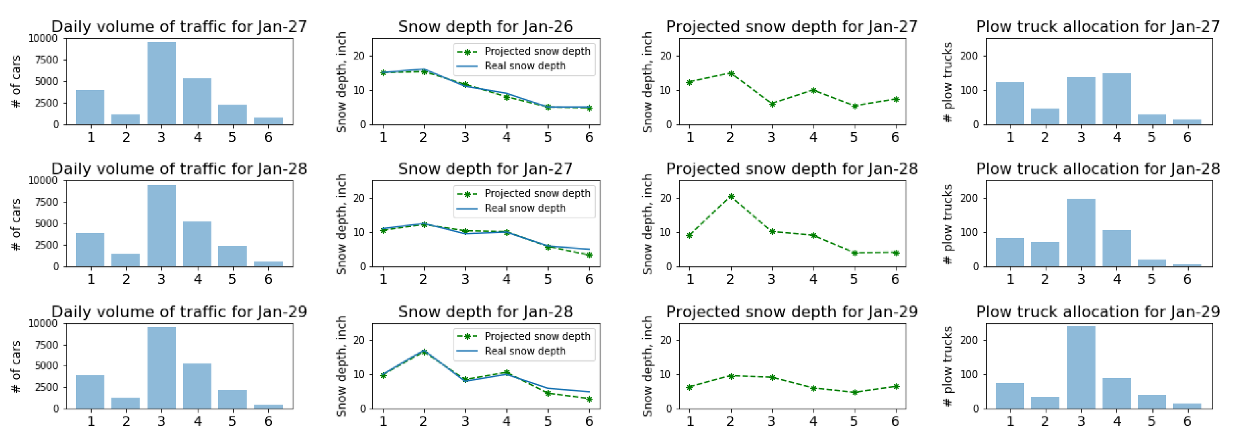
We define the region that is affected by the disaster where the limited resources have to be allocated. Further, we divide the affected region into sub-regions to narrow down the areas.



Results

We applied our framework to the snowstorm disaster in the Western New York area for the period of January - February 2019.

Plow truck allocation among six regions in NY, USA: Buffalo-Niagara, East Aurora, Hamburg, Kenmore, Tonawanda, Lockport



Social Behaviour Modeling in Critical Infrastructure

Emerging data that track the dynamics of large populations bring new potential for understanding human decision-making in a complex world and supporting better decision-making through the integration of continued partial observations about dynamics.

Our goal

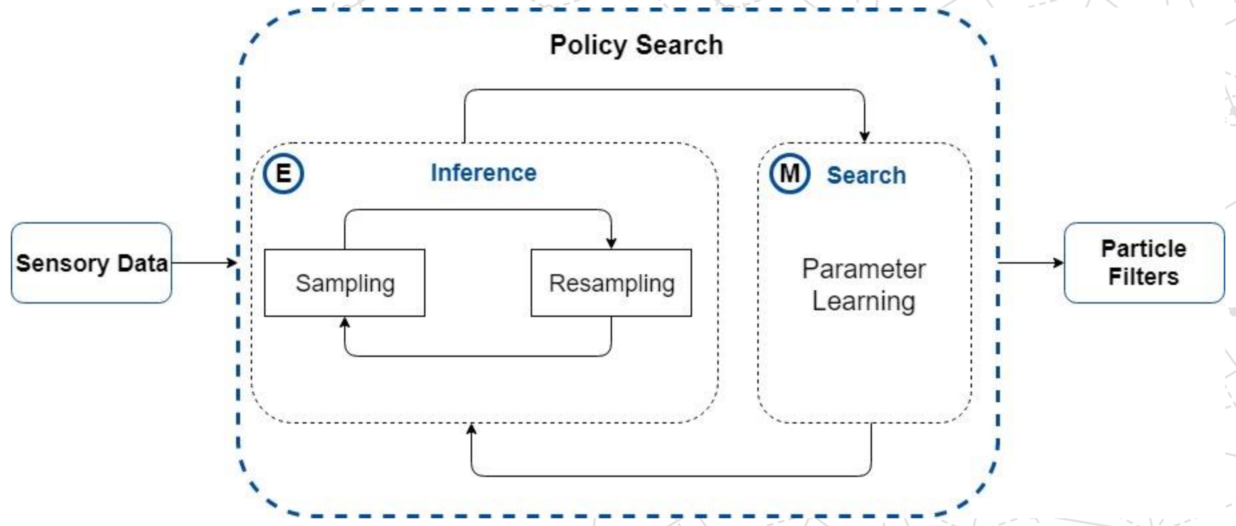
Model the diverse dynamics of agent interactions and decision-making in a complex network.

Our approach

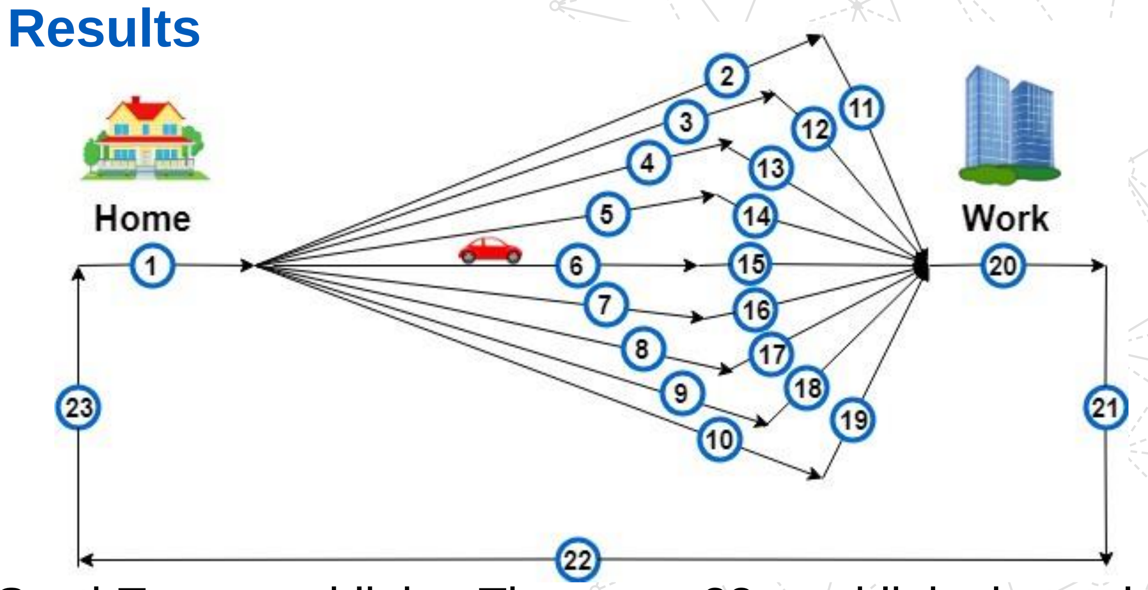
- We developed a particle filter algorithm to model how the networked system continually tracks the current state of itself and the environment using noisy observation streams, and to learn how to make near-optimal plans by adjusting the rates at which interaction events happen.
- Our learning algorithm works through policy search that maximizes the expected log-likelihood over agent policies in a mixture of dynamic Bayesian networks.

Overview of our policy search algorithm

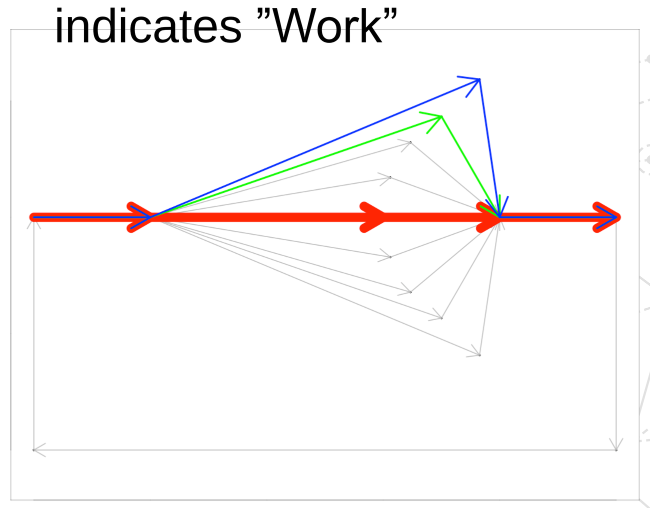
The algorithm is done in two parts. During the Inference we estimate the state density and reward associated with each particle. During the Search step we update policy parameters according to the Inference results.



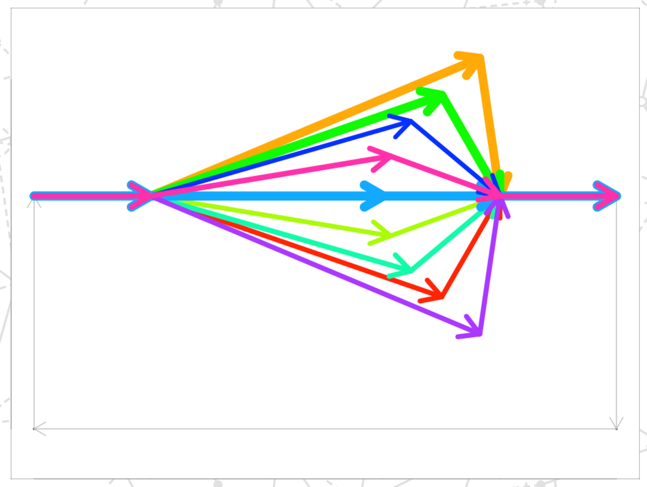
Results



SynthTown road links. There are 23 road links in total, where segment 1 indicates "Home" and segment 20 indicates "Work"



Baseline



Our algorithm

References

1. Alina Vereshchaka and Wen Dong. Dynamic Resources Allocation During Natural Disasters Using Multi-agent Environment. Proceedings of the International Conference on Social, Cultural, and Behavioral Modeling & Prediction and Behavior Representation in Modeling and Simulation, 2019

2. *To appear:* Fan Yang, Alina Vereshchaka and Wen Dong. Optimizing Complex Interaction Dynamics in Critical Infrastructure with a Stochastic Kinetic Model. In Proceedings of the 2019 Winter Simulation Conference.

Contact Info: avereshc@buffalo.edu