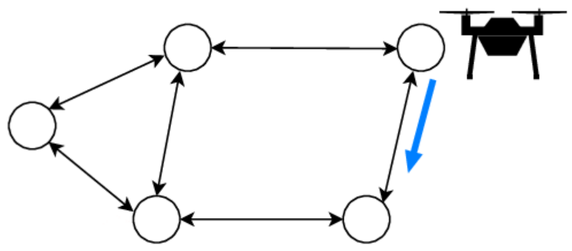
Reinforcement Learning based Persistent Surveillance by Robots

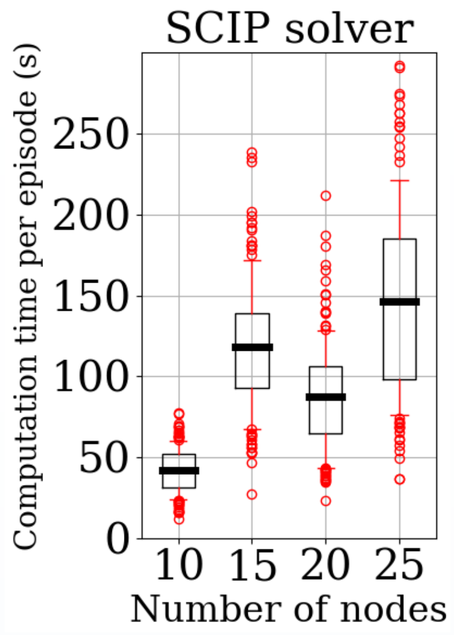
Aravind S

Advisor: Dr. Pavankumar Tallapragada

Robert Bosch Centre for Cyber-Physical Systems

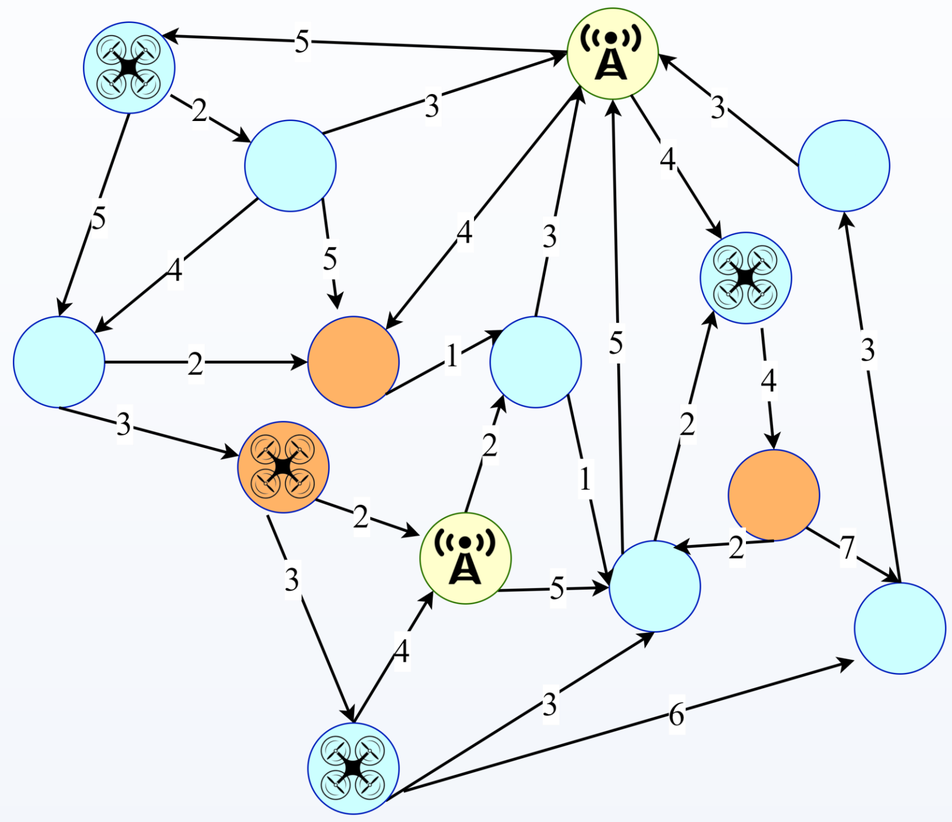
INTRODUCTION





- Surveillance problems posed as route optimization problems over graphs are computationally very expensive to solve.
- Computation time increases exponentially with horizon length and number of agents
- Using RL to solve these problems, can obtain **near-optimal** routing policies with **very little computation time**.

PROBLEM SETUP



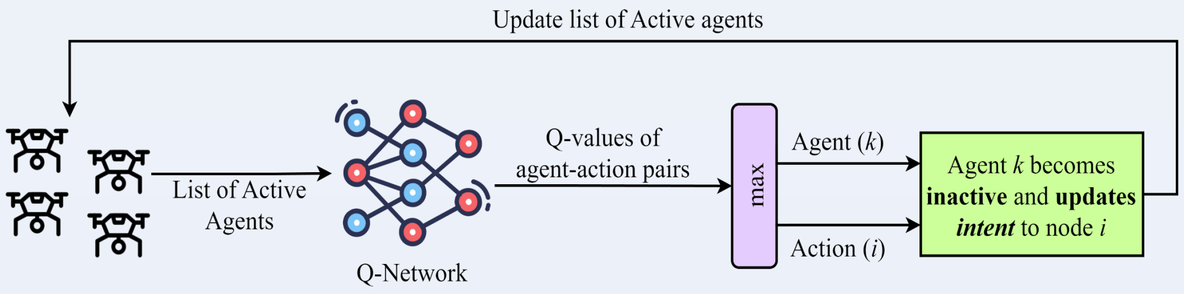
- Strongly connected graph with survey and base nodes
- Time varying priorities for each of the nodes
- The agents must visit the base nodes to upload collected surveillance data
- Each agent is energy constrained and has a finite battery
- The persistent surveillance problem is posed as a **rolling horizon optimal control problem**
- Minimize the weighted (by node priorities) graph idleness and time between visits to a base node.

REINFORCEMENT LEARNING FRAMEWORK

- Training is done using the **Double Dueling Q-Learning (D3QN)** algorithm
- The energy constraints are guaranteed to be satisfied by **projection of actions to a safe set** for each agent
- An action *j,* is chosen only if

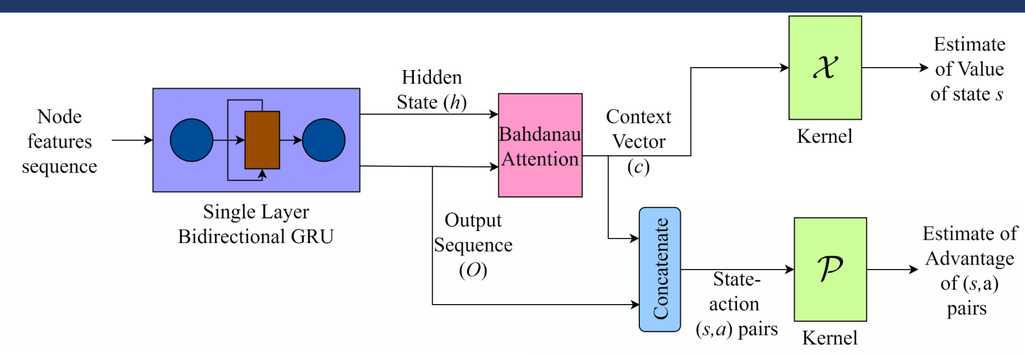
$$e_k(t) \geq w(l_k(t), j) + e_{\min}(j)$$

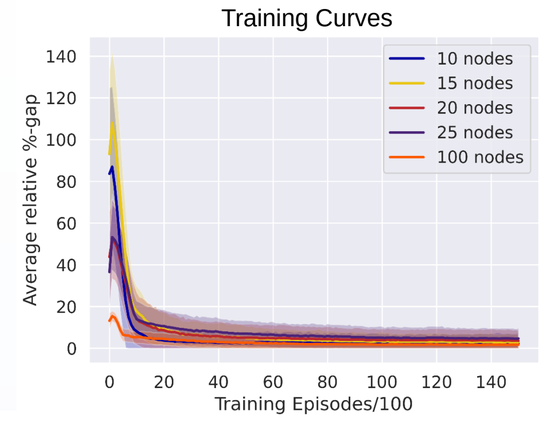
- To tackle to computation intractability of the large action space of the multi agent problem, we adopt a **Sequential Decision Making** approach

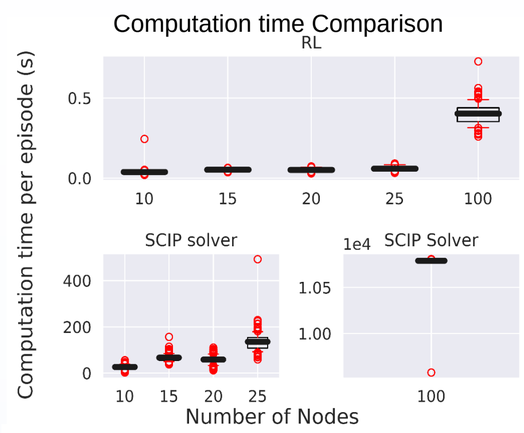


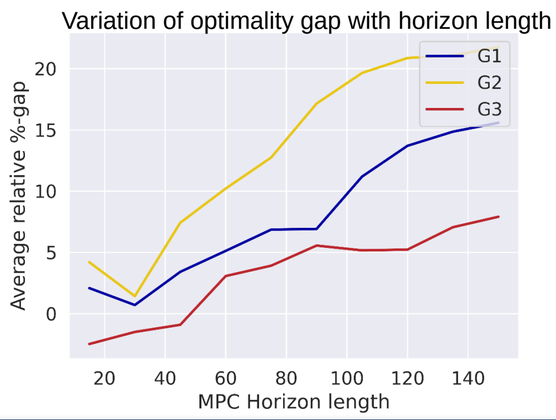
Mr. Nishchal Hoysal G, a PhD student in RBCCPS has contributed towards this project in the literature review, problem formulation and simulations without energy constraints. Mr. Lalit, an Intern at the CNS group has contributed to visualizing the motion of the agents on the graphs.

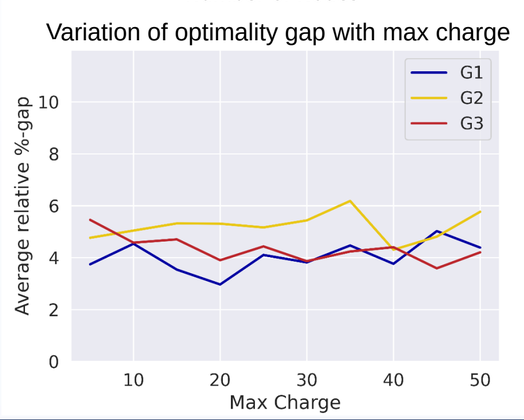
THE SINGLE AGENT PROBLEM



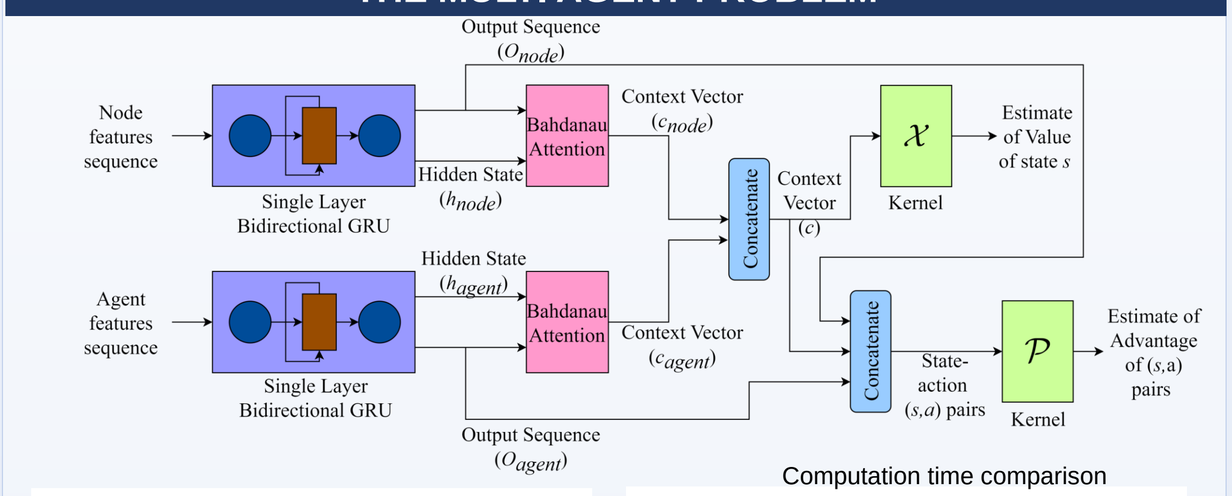


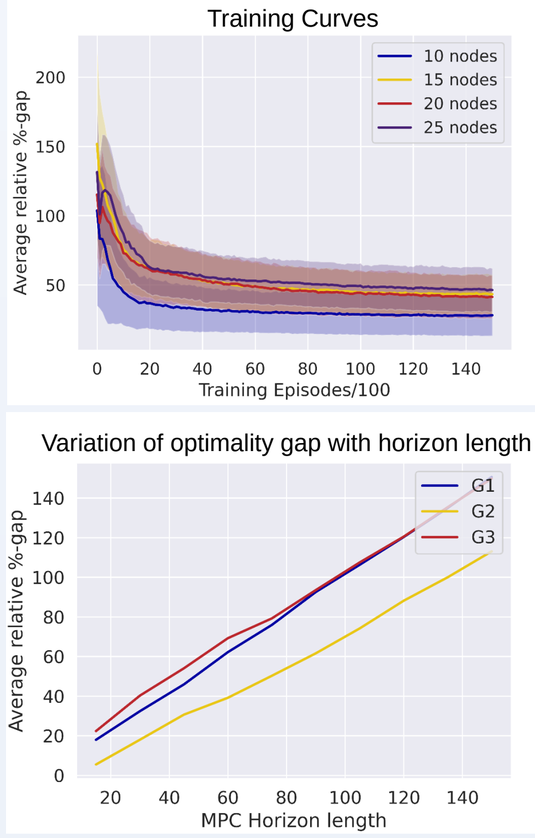


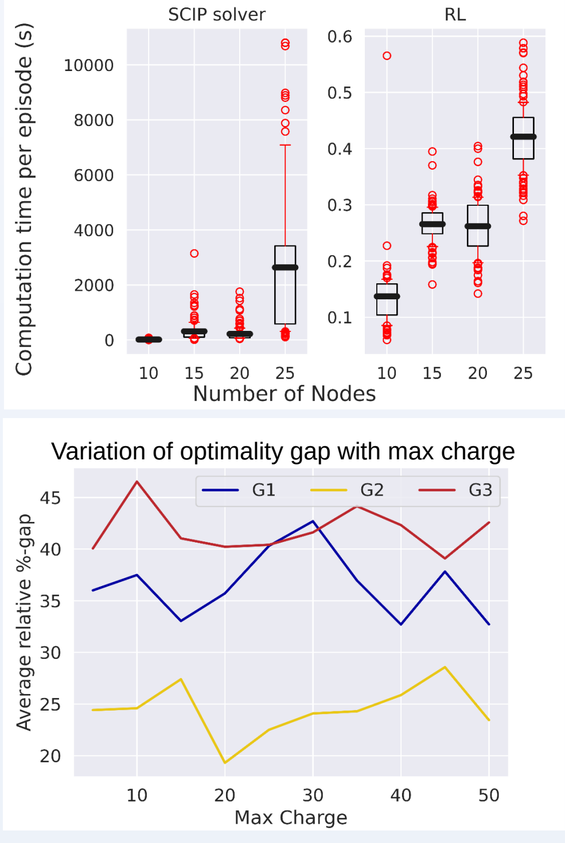




THE MULTI AGENT PROBLEM







CONCLUSION

- In this project, we proposed a rolling horizon based RL framework for persistent surveillance using energy constrained robots, where each robot has to not only perform surveillance but also report the collected data to base nodes.
- The RL policies produced solutions comparable to the SCIP solver for the single agent case but were highly suboptimal for the multi agent case.
- The method provides several orders of magnitude reduction in the computation time for both the single and multi agent problem.