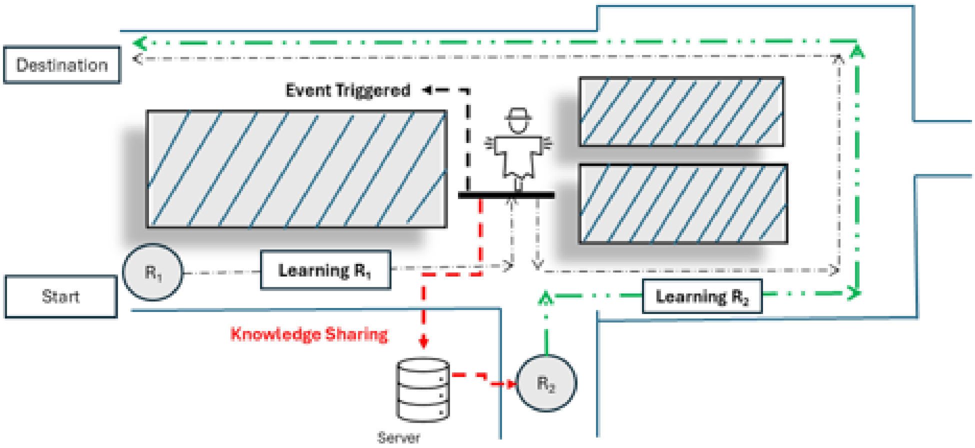
Dynamic Event-Triggered Learning for Cognitive Route Optimization in Time Federated Agile-Robotics

Muhammad Waqas Nawaz, Qammer H Abbasi, Muhammad Imran, Dr. Olaoluwa Popoola
University of Glasgow, UK

Objective:

We focus on distributed multi-autonomous system approach that enables real-time decision-making and knowledge sharing to minimize energy consumption and improve agility in transport operations, contributing to the aim of achieving net zero emissions in the transport sector. Temporal Event-Triggered Federated Learning threshold adds additional reliability to our communication strategy and contributes to our multi-robot navigation system's overall efficiency and accuracy. In this context, robots collaboratively acquire a shared collision avoidance model, utilizing local model updates communicated to a central server from the affected unit. Upon receiving updates, the central server aggregates information, adapting the global path planning model and broadcasting it back to all units. This ensures that all units benefit from collective knowledge and adaptations to navigate toward a destination while avoiding obstacles in object-restricted areas utilizing resources efficiently as shown below.



Our research aims to address the challenges of energy consumption and computational overhead in autonomous system by developing intelligent decision-making algorithms. Specifically, we seek to achieve the following objectives:

1. Develop a framework for knowledge aggregation among autonomous system to enable real-time information sharing.
2. Optimize energy usage and reduce computational overhead through intelligent decision-making algorithms.
3. Evaluate the performance and effectiveness of the proposed approach in simulation and real-world scenarios.

Methodology:

In our proposed approach using Proximal Policy Optimization (PPO) with temporal discounted rewards and communication-triggered Federated Learning for adaptive collision avoidance in a multi-robot system, the adjustment of the policy when the cumulative negative reward reaches its closest to obstacles with time step involves the learning process within the PPO algorithm.

Here's a high-level overview of how the cumulative negative reward can impact policy adjustment in the context of PPO:

$$S_{t+1} \sim P(\cdot | S_t, A_t)$$

$$\pi_\theta(a \mid s) = P(A_t = a \mid S_t = s, \theta)$$

$$G_t = \sum_{k=0}^{\infty} \gamma k R_{t+k+1} \times d_t$$

$$R_t = -\alpha_t \cdot d_t$$

where d_t is the proximity to obstacles, and α_t is a scaling factor at time t.

S_t: State at time step t.
A_t: Action taken at time step t.
$\pi_\theta(a|s)$: Policy function parameterized by θ, representing the probability of acting a given state s.
R_t: Immediate reward at time step t.
G_t: Cumulative discounted reward at time step t.
γ: Discount factor.
$L(\theta)$: PPO objective function.
δ_t: Surrogate advantage function

$$L(\theta) = \mathrm{E}_t\left[\min\left(\frac{\pi_\theta(A_t|S_t)}{\pi_{\theta\mathrm{old}}(A_t|S_t)} \cdot \delta_t, \mathrm{clip}\left(1-\epsilon, 1+\epsilon, \frac{\pi_\theta(A_t|S_t)}{\pi_{\theta\mathrm{old}}(A_t|S_t)} \cdot \delta_t\right)\right)\right] . \max\left(-1, -k. d_t\right)$$

$$\delta_t = R_t + \gamma \cdot V_\theta(S_{t+1}) - V_\theta(S_t)$$

$$V_\theta(S_t) \leftarrow \mathrm{E}_t[R_{t+1} + \gamma \cdot V_\theta(S_{t+1})]$$

$$\theta_i \leftarrow \mathrm{FederatedUpdate}(\theta_i, \{\theta_j\}_{j \neq i})$$

Where θ_i is the local policy parameters of agent i, and $\{\theta_j\}_{j \neq i}$ represents policies of other agents.

If $G_t \leq$Threshold, then trigger federated learning event. By incorporating the cumulative negative reward into the PPO learning process, the policy of the robot is adjusted to navigate away from obstacles, and communication events are triggered when the robot encounters situations that warrant updates to the shared knowledge base. This helps in achieving adaptive collision avoidance in the multi-robot system.

Training Process:

- Optimize θ using PPO for individual agents.
- Utilizing negative discount awards for learning obstacle avoidance.
- Trigger FL events based on cumulative discounted rewards.

The policy adjustment is typically performed through gradient descent, where the gradients of the policy with respect to the negative reward are used to update the policy parameters.

Results:

Evaluation results using DRL embedded PPO and DQN training strategies show success rates for path navigation and average completion times, with and without knowledge sharing, respectively, via event-triggered FL.

Examples	Test 1	Test 2	Test 3	Test 4	Test 5	Mean
DQN	85%	86%	92%	85%	89%	0.874%
Temporal-PPO	98%	95%	96.9%	98%	94%	0.965%

Figure 1: Success rate by comparing the accuracy.

Fig. 1 illustrates the evaluation results in the simulated environment shows the success rate for each model, the task was performed 10 times and the results average over 10 runs for 5 agents. The avg performance of models trained with Deep Q-network is the lowest. Fig. 2 shows the average completion time of each run with knowledge sharing strategy for intelligent navigation using Temporal-PPO.

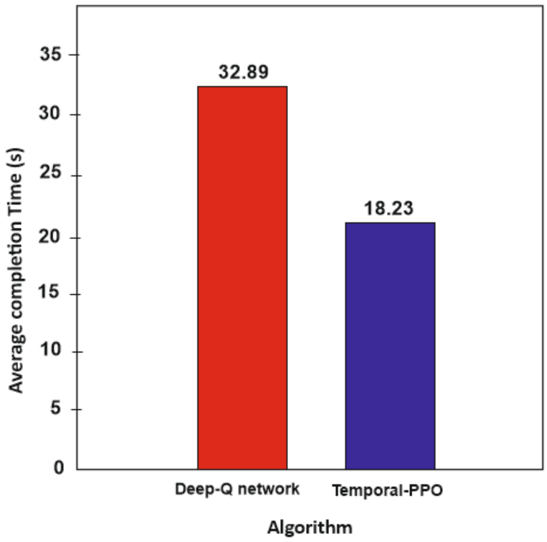


Figure 2: Average Completion Time on each model.

Significance:

This study holds significant implications for the advancement of autonomous technology and its broader impact on society:

1. Enhanced Efficiency: By optimizing energy usage and reducing computational overhead, our approach can improve the overall efficiency of autonomous system, leading to cost savings and environmental benefits.
2. Improved Safety: Real-time information sharing among autonomous system can enhance safety by enabling proactive responses to changing road conditions and potential hazards.
3. Scalability: The developed framework can be scaled to accommodate a growing network of AVs, paving the way for widespread adoption and integration into smart transportation and connected systems.
4. Technological Innovation: Our research contributes to the ongoing innovation in ground autonomous vehicle technology, driving forward the development of smarter and more sustainable transportation solutions. Overall, this study represents a crucial step towards realizing the full potential of ground autonomous vehicles and their role in shaping the future of transportation.