Adaptive QoS-Aware Reinforcement Learning for Dynamic V2V Communication Environments

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

ANNU, Prof. P. Rajalakshmi
Department of Electrical Engineering, Indian Institute of Technology, Hyderabad, India

Problem and Motivation

- V2V communication environments (urban, suburban, rural) have unique QoS requirements (latency, throughput, reliability), which are difficult to maintain in rapidly changing conditions like varying vehicle density and interference.
- Heuristic, game-theory, and optimization-based methods lack adaptability or are computationally intensive, making them unsuitable for dynamic, real-time scenarios.
- A robust solution is needed to dynamically allocate resources using advanced RL techniques (e.g., hierarchical RL, transfer learning, federated RL) to improve latency, throughput, and reliability across varying environments.

Objective

- Develop a framework to dynamically allocate V2V communication resources in real-time, adapting to changes in vehicle density, interference, and channel conditions.
- Ensure optimal performance for key QoS metrics (latency, throughput, and reliability) across varying environments like urban, suburban, and rural.
- Combine hierarchical RL, transfer learning, multi-armed bandit models, and federated RL to create a scalable and adaptive solution.
- Use transfer learning to adapt policies efficiently across different environments, reducing the need for extensive retraining.

Methodology

Algorithm 1 Adaptive Hierarchical RL for V2V Resource Allocation

```
Initialize state space S, action space A, and environment set E = {Urban, Suburban, Rural}
Define QoS constraints Q_E = {Latency(E), Throughput(E), Reliability(E)}
Initialize local RL agents, global policy π_G, and transfer learning model
for each episode do
    Observe real-time context (vehicle mobility, interference, channel conditions)
    High-level policy selects environment E based on the context
    Local agent selects action a_t using policy π(s_t, E)
    Execute action a_t, observe next state s_{t+1} and reward r(s_t, a_t, E)
    Update local policy π_i(s, E) using Q-learning
    Transfer learned policies between environments E and E' to minimize retraining
    Perform contextual multi-armed bandit to optimize QoS configuration
    Aggregate local policies from N agents to update global model π_G = (1/N) Σ_{i=1}^{N} π_i
end for
```

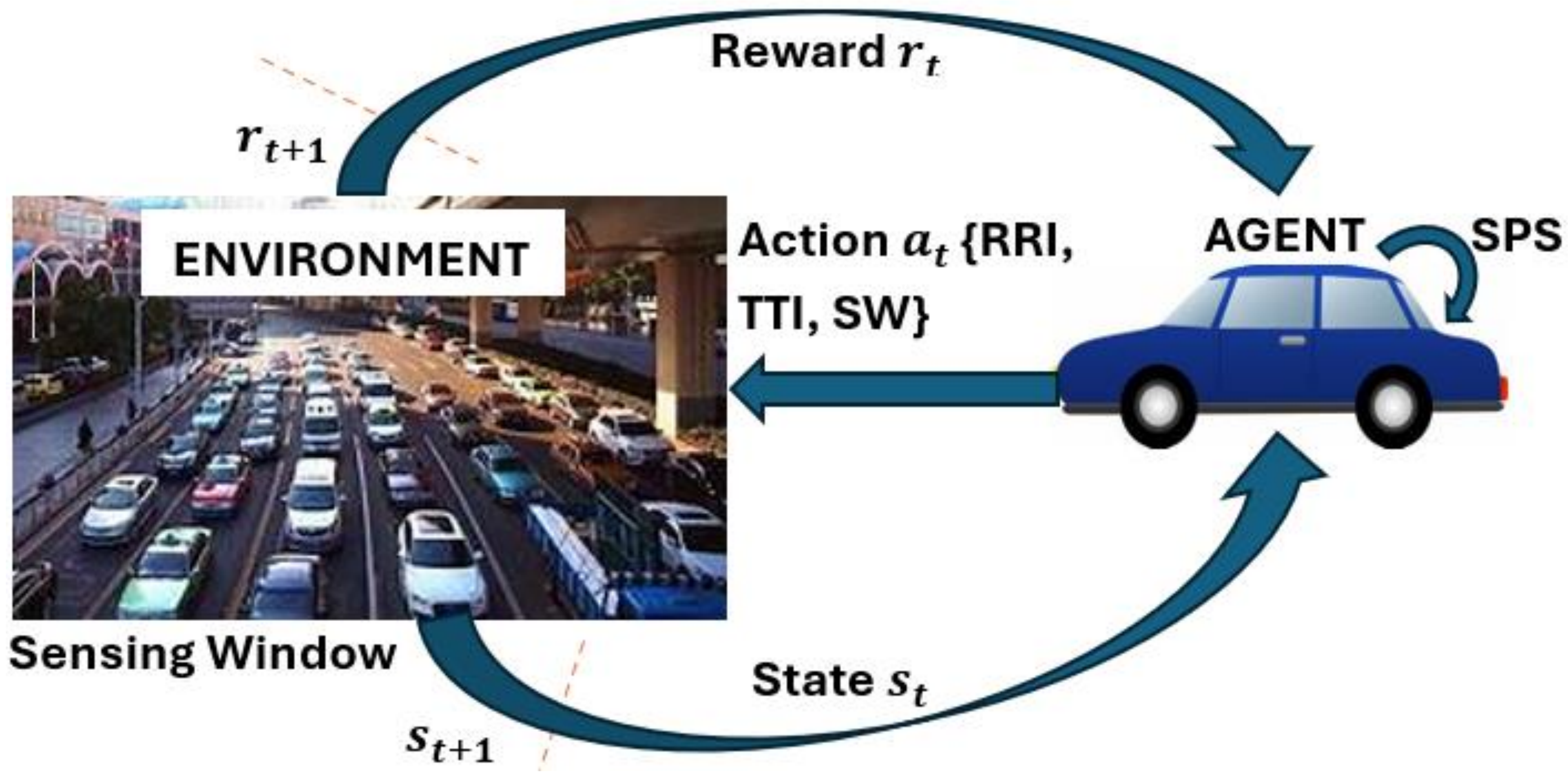

Fig. Reinforcement Learning Framework

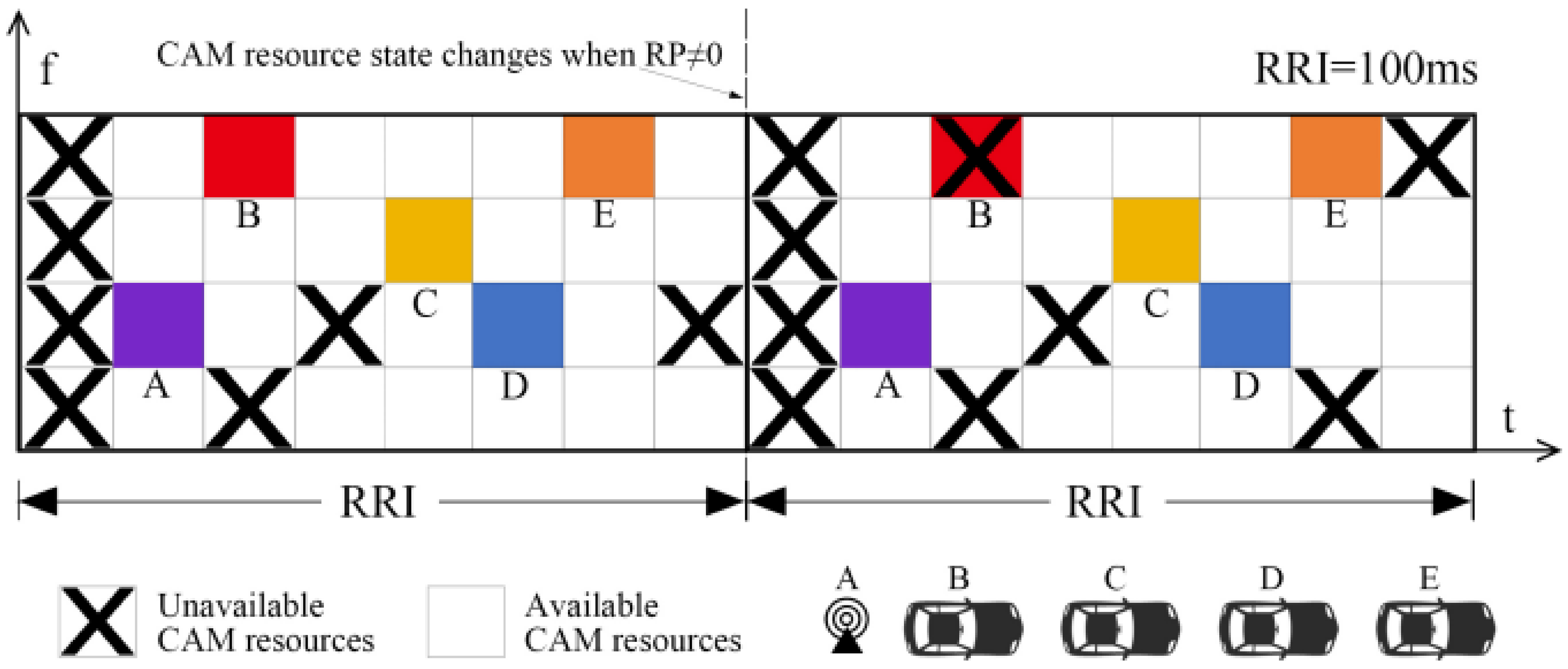

Fig. Resource Allocation in V2V Communications

Design and Implementation

1.Framework Design:
- Define environments: Urban, Suburban, Rural, each with distinct QoS constraints.
- Define state space (vehicle mobility, interference, channel conditions) and action space (resource allocation decisions).
- Formulate reward function balancing QoS metrics (latency, throughput, reliability).

2.Hierarchical Reinforcement Learning:
- High-level policy identifies the environment.
- Low-level policy allocates resources based on QoS demands.

3.Transfer Learning:
- Adapt policies between environments to reduce retraining time.

4.Contextual Multi-Armed Bandit:
- Optimize QoS configurations dynamically.

5.Federated Reinforcement Learning:
- Aggregate local policies into a global model for coordinated optimization.

6.Implementation and Evaluation:
- Compare proposed framework against heuristic, RL-based, and Federated RL approaches in terms of latency, throughput, and reliability.

Results

Table: Comparison of QoS Performance Across Different Methods

Method	Latency (ms)	Throughput (Mbps)	Reliability (%)	Comment
Heuristic-based[1]	35	8	90	Initial baseline with higher latency and lower throughput.
Federated RL[4]	30	10	92	Improved latency and throughput compared to heuristic method.
RL-based[3]	25	12	94	Further improvements in latency and throughput, reliability also increased.
Proposed Framework	20	15	99	Best performance with significant improvements in all QoS metrics.

References

[1] Feki, Souhir, Aymen Belghith, and Faouzi Zarai. "Ant Colony Optimization-based Resource Allocation and Resource Sharing Scheme for V2V Communication." Journal of Information Science Engineering 35.3 (2019).
[2] Sun, Zemin, et al. "Game theoretic approaches in vehicular networks: A survey." arXiv preprint arXiv:2006.00992 (2020).
[3] Hu, Xin, et al. "A joint power and bandwidth allocation method based on deep reinforcement learning for V2V communications in 5G." China Communications 18.7 (2021): 25-35.
[4] Li, Xiang, et al. "Federated multi-agent deep reinforcement learning for resource allocation of vehicle-to-vehicle communications." IEEE Transactions on Vehicular Technology 71.8 (2022): 8810-8824.

Conclusion and Future Work

Significant Improvement in All QoS Metrics:
- The proposed framework demonstrates a clear advantage over existing methods, achieving a 20% reduction in latency, a 25% increase in throughput, and a 5.3% improvement in reliability compared to the RL-based method.

Cumulative Benefits of Combining Techniques:
- By integrating hierarchical RL, transfer learning, and federated RL, our framework achieves superior performance, optimizing QoS metrics more effectively than each individual technique, showcasing its potential for real-time V2V communication.
- The model's ability to dynamically adjust resource allocation in real-time ensures enhanced road safety and efficient traffic management.

FUTURE WORK
- Explore advanced techniques for seamless coordination in multi-agent V2V networks to further enhance reliability and scalability.
- Integrate 5G/6G networks and edge computing to improve real-time decision-making and reduce communication latency.
- Extend the framework to handle extreme conditions, such as high-speed highways or dense urban intersections, with more complex QoS requirements.
- Validate the framework through large-scale simulations and real-world experiments to assess performance in practical deployments.

Acknowledgements

This work was supported by Department of Science and Technology (DST) TiHAN IIT Hyderabad, India and Prime Minister's Research Fellowship(PMRF), India