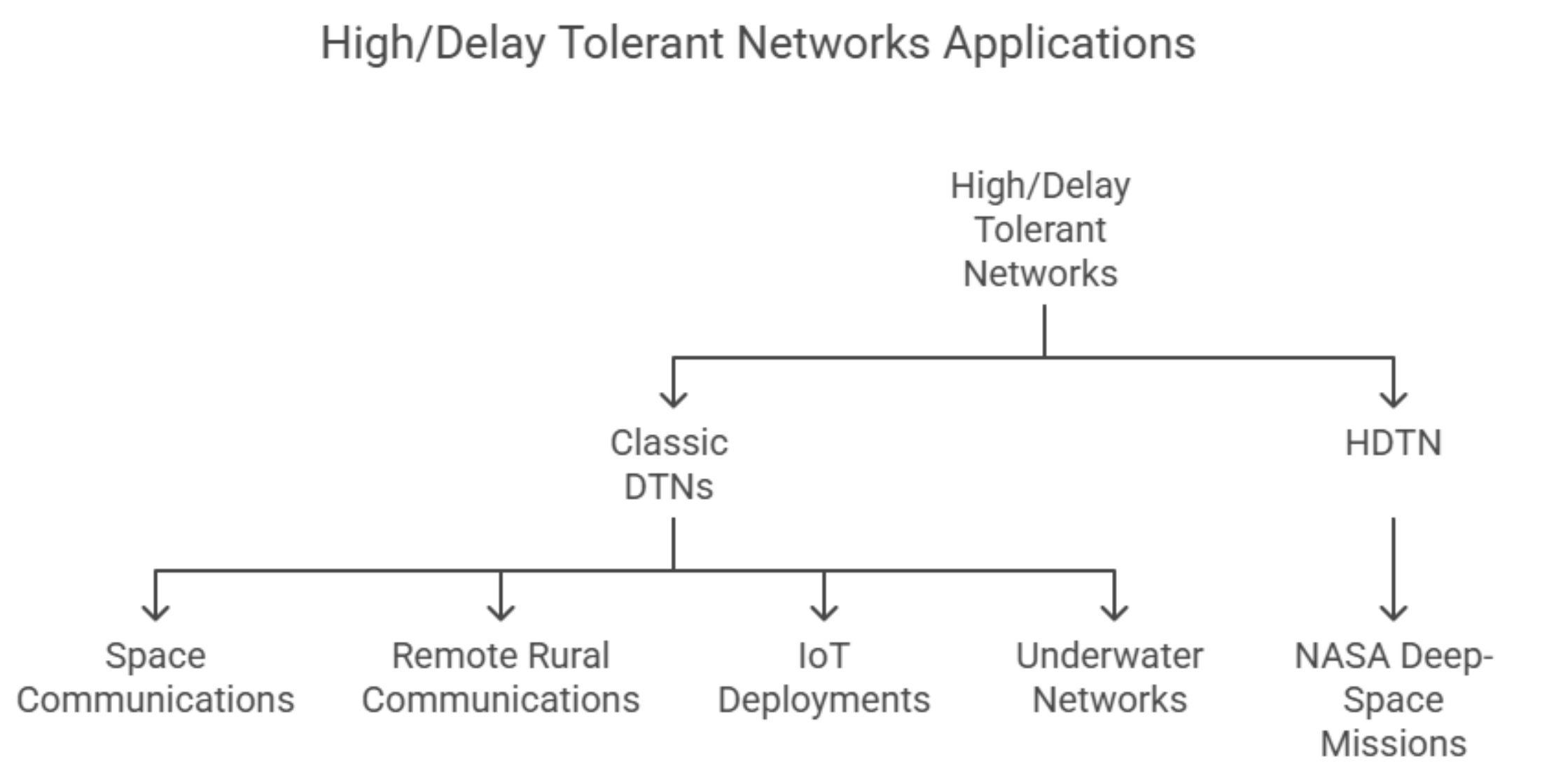
Reinforcement Learning for Autonomous Routing in High-Delay Tolerant Networks

Nashithuddin Mohammed[1]*, Sheetal Kumar[2]*, Mohammad Abdus Salam[3]
[1,2]*Graduate Student*, [3]*Professor*

Abstract

High Delay Tolerant Networks (HDTNs) are designed to support communication in environments characterized by intermittent connectivity, high latency, and frequent disruptions. Conventional routing protocols often lack the adaptability required to handle dynamic topologies and unpredictable link availability, limiting overall network performance.



High/Delay Tolerant Networks Applications

High/Delay Tolerant Networks
Classic DTNs
HDTN
Space Communications
Remote Rural Communications
IoT Deployments
Underwater Networks
NASA Deep-Space Missions

This work explores Reinforcement Learning (RL) for autonomous routing in HDTNs. RL agents learn optimal forwarding strategies through interaction and delayed rewards, improving delivery, reducing overhead, and adapting to network changes.

Deep RL techniques use traffic history, mobility patterns, and buffer states to make intelligent routing decisions under uncertainty. Applications in space communications, military operations, and disaster recovery demonstrate the advantages of RL over conventional methods, enhancing throughput, reducing packet loss, and increasing resilience.

These findings highlight the potential of learning-enabled routing to improve the efficiency and reliability of next-generation HDTNs.

Problem Statement

The deployment of large-scale HDTNs is hindered by the limitations of existing routing protocols:

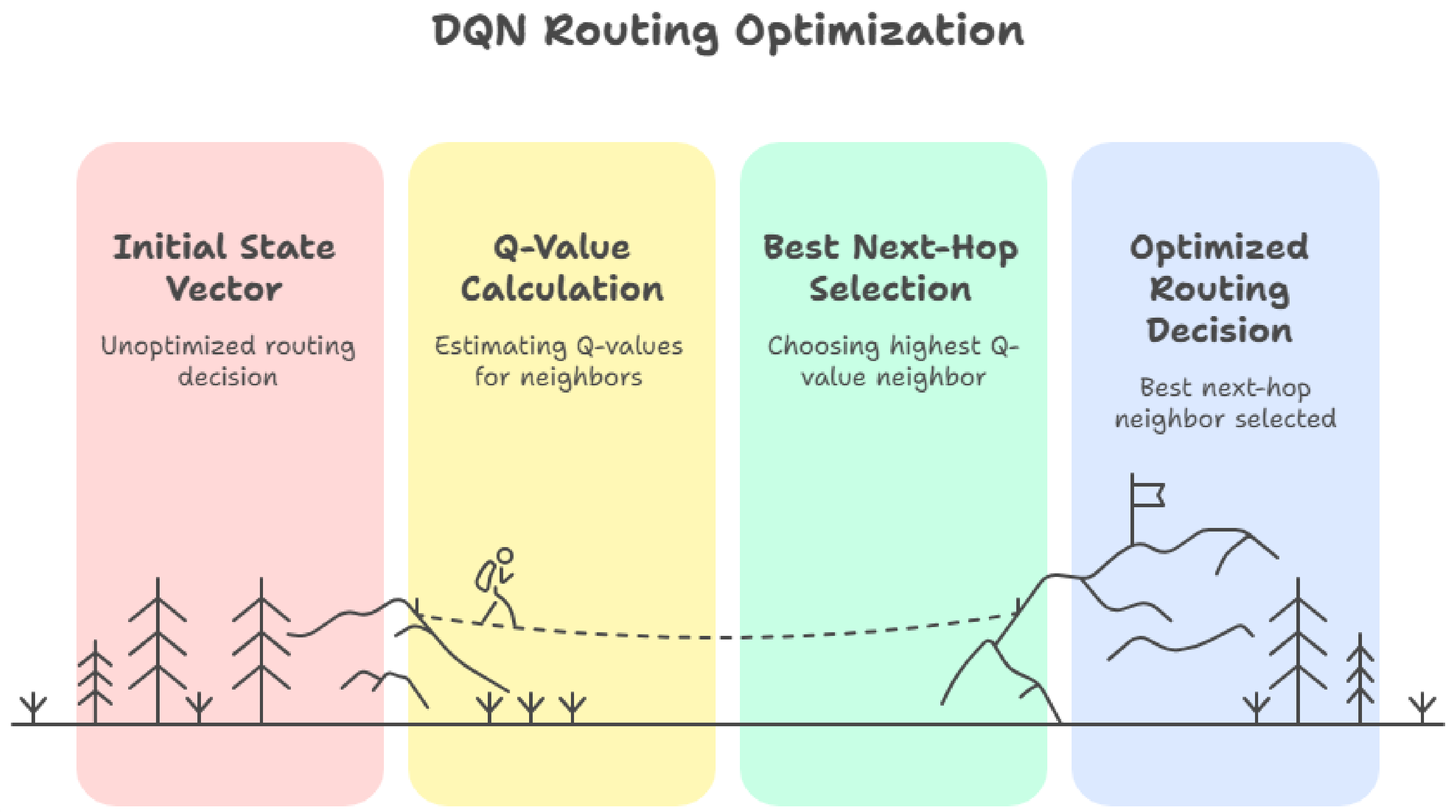
Lack of Adaptability: Protocols like Contact Graph Routing (CGR) rely on a pre-computed, static contact plan and struggle with probabilistic contacts, unplanned disruptions, or dynamic topology changes. They cannot learn from or adapt to network evolution.

Scalability Bottlenecks: Contact Graph Routing (CGR)'s computational complexity grows significantly with the number of contacts, making it infeasible for resource-constrained nodes in large networks (e.g., mega-constellations).

Handling Uncertainty: Traditional methods assume perfect knowledge of future contacts. In-reality, contacts can be uncertain in timing, duration, or data rate due to orbital prediction errors or environmental factors.
There is a critical need for routing schemes that are adaptive, scalable, and resilient to uncertainty.

Background and Method of Approach

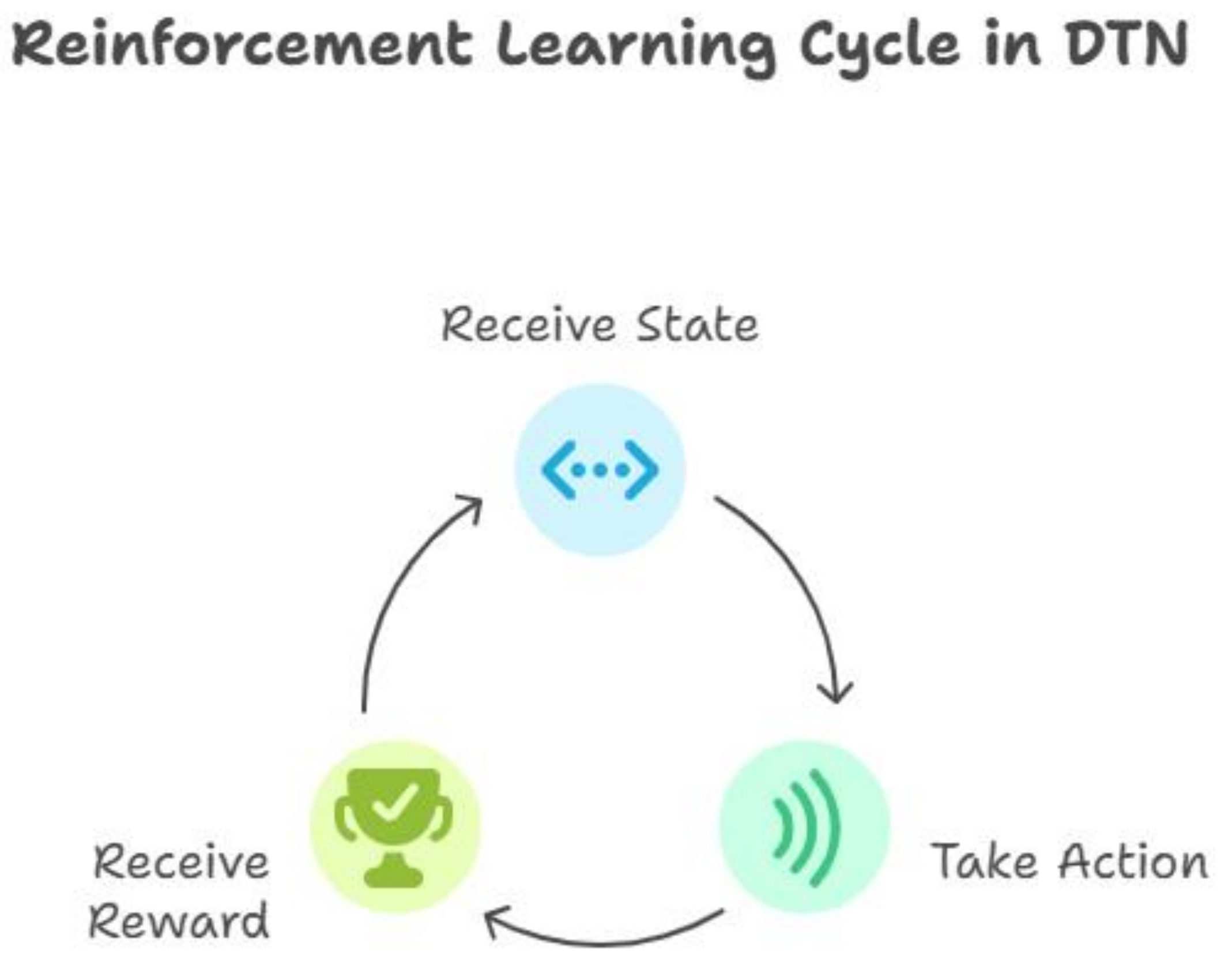
- **Delay-Tolerant Networking (DTN):** An architecture designed for environments lacking continuous end-to-end connectivity, operating on a store-carry-and-forward principle. Contacts are classified as opportunistic, probabilistic, or scheduled.
- **Contact Graph Routing (CGR):** The state-of-the-art for space DTNs. CGR uses a scheduled contact plan to compute time-aware routes via an adapted Dijkstra's algorithm. While highly effective in predictable environments, its scalability and adaptability are limited.
- **Machine Learning in DTN Routing:** Prior research has explored ML for routing, primarily using supervised learning (e.g., Decision Trees, Neural Networks) to classify the best next hop. However, these approaches often fail to exploit the graph structure and do not focus on computational stability.
- **Reinforcement Learning for Routing:** RL formulates routing as a sequential decision-making problem:
    - **Q-Routing & Standard Q-Learning:** Foundational algorithms where nodes maintain Q-tables estimating delivery cost via neighbors. Successfully applied in VANETs and underwater networks.
    - **Deep Q-Learning (DQN):** Uses deep neural networks to approximate the Q-function, enabling generalization in large state spaces. Proposed for complex environments like lunar networks.
    - **Advanced Reinforcement Learning:** Techniques like Double Q-Learning reduce overestimation bias, and Multi-Agent RL frameworks allow distributed nodes to collaboratively optimize global network performance.



DQN Routing Optimization

The core of RL-based routing involves an agent at each node learning an optimal policy.

- **State Space (s):** Defines the agent's perception of the environment. Can include:
    - Local buffer occupancy and bundle priorities.
    - History of contacts with neighboring nodes.
    - Current time and predicted node mobility.
    - Neighbor discovery information.
- **Action Space (a):** The set of possible decisions at each step. The primary action is selecting a next-hop neighbor to forward a bundle to. The action to store and carry the bundle is also implicit.
- **Reward Function (r):** A scalar feedback signal that the agent aims to maximize. In HDTNs, rewards are often delayed.
    - **Positive Reward:** Granted upon confirmed end-to-end bundle delivery.
    - **Negative Reward/Penalty:** Applied for bundle drops, excessive latency, or resource consumption.



Reinforcement Learning Cycle in DTN

- **Learning Algorithm:** The agent uses an RL algorithm (e.g., DQN)to iteratively update its policy. It explores the environment (trying different actions) and exploits learned knowledge to maximize cumulative future reward, ultimately learning which next-hop leads to the most successful and efficient delivery.

Results

- **Improved Delivery Metrics:** RL approaches demonstrate a capacity to achieve higher delivery ratios and lower latency compared to probabilistic and flooding-based protocols (e.g., Prophet, Epidemic)in scenarios with unpredictable mobility.
- **Enhanced Adaptability:** In simulations of uncertain contact plans, RL agents successfully learned to route around failed or delayed contacts, outperforming static CGR which relies on perfect a priori knowledge.
- **Overhead Reduction:** By learning efficient paths, RL reduces unnecessary bundle replication, a common drawback of flooding-based methods, thereby conserving precious bandwidth and node storage.
- **Resilience:** RL-based routers show improved performance in the face of network disruptions and node failures, as they can continuously adapt their policies without human intervention.
- **Trade-off:** While a pure RL approach may not always match the optimal delivery time of a perfectly planned CGR, it offers superior performance in realistic, uncertain conditions and provides a stable, low-latency computational footprint once trained.

Conclusion

Reinforcement Learning presents a powerful and promising framework for overcoming the fundamental routing challenges in High Delay Tolerant Networks. By enabling nodes to autonomously learn and adapt to dynamic, disruptive, and delayed environments, RL moves beyond the limitations of static algorithms like CGR. While challenges remain—including training time, convergence guarantees, and integration into flight-grade software—the potential for RL to enhance the throughput, resilience, and autonomy of HDTNs is clear. Future work should focus on hybrid CGR-RL systems, transfer learning, and the standardization of learning-based modules to pave the way for the next generation of intelligent space and terrestrial internetworks.

References

[1] Yesuf, Fuad Yimer, and M. Prathap. "CARL-DTN: context adaptive reinforcement learning based routing algorithm in delay tolerant network." arXiv:2105.00544 (2021). https://arxiv.org/abs/2105.00544

[2] Fraire, Juan A., Olivier De Jonckère, and Scott C. Burleigh. "Routing in the space internet: A contact graph routing tutorial." Journal of Network and Computer Applications 174 (2021): 102884. https://doi.org/10.1016/j.jnca.2020.102884

[3] Dudukovich, Rachel, and Christos Papachristou. "Delay tolerant network routing as a machine learning classification problem." 2018 NASA/ESA Conference on Adaptive Hardware and Systems (AHS). IEEE, 2018. https://doi.org/10.1109/ahs.2018.8541460

Real Time Communications
Conference and Expo at Illinois Tech
IEEE International Conference
IEEE RTC 2025 - October 7-8, 2025 - IIT, Chicago, IL, USA