Annalisa Navarro

Autonomous Orchestration of Next-Generation Networks Using Reinforcement Learning

Tutor: Prof. Roberto Canonico

Cycle: XXXVIII

Year: III

Candidate's information

MSc degree: Computer Engineering

DIETI Research group/laboratory: Comics/ArcLab

PhD start date – end date: 01/11/2022 – 31/10/2025

Scholarship type: UNINA

Period abroad: 15/05/2024 – 15/10/2024 at Deutsche Telekom Chair of Communication Networks (Dresden University of Technology, Germany)

Summary of study activities (1/2)

Ad hoc courses

- Using Deep Learning Properly
- Scientific Programming and Visualization with Python
- Statistical Data Analysis for Science and Engineering Research
- Virtualization Technologies and Their Applications
- Big Data Architecture and Analytics
- Artificial Intelligence and Natural Language Processing
- Hands-on Network Intrusion Detection via Machine Learning and Deep Learning
- Strategic Orientation for STEM Research & Writing

PhD Schools

- RESTART Tech Camp on 5G and Open RAN
- 20th TAROT Summer School on Software Testing, Verification & Validation

56 seminars

Summary of study activities (2/2)

Conferences attended

- IEEE International Conference on Machine Learning for Communication and Networking - 2025
- IEEE International Conference on Distributed Computing Systems - 2024
- IEEE Conference on Network Function Virtualization and Software Defined Networks - 2023
- IEEE International Conference on Network Softwarization - 2023
- IFIP Networking Conference - 2023
- IEEE Latin-American Conference on Communications - 2022

PhD Year	Courses	Seminars	Research	Tutoring
1st	23.0	13.4	26.2	0.32
2nd	12.0	7.6	44.0	2.36
3rd	4.0	4.8	65.0	0.16
Total	39.0	25.8	135.2	2.84

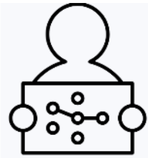
Research area

- **Automated Management and Control of Modern Communication Networks.**
- Main challenges in this context are:


performance


security


reliability


explainability

Case studies: Software Defined Wide Area Networks (***SD-WAN***), Open Radio Access Networks (***O-RAN***), 5G Core Network (***5GC***), Cyber Physical Systems (***CPS***), Low Power Wide Area Networks (***LP-WAN***)

Research products

[P1]	A. Botta, R. Canonico, A. Navarro, G. Stanco, G. Ventre, *Adaptive overlay selection at the SD-WAN edges: A reinforcement learning approach with networked agents***,** **Computer Networks,** vol. 243, p. 110310, Apr. 2024, DOI: 10.1016/J.COMNET.2024.110310.
[P2]	G. Stanco, A. Navarro, F. Frattini, G. Ventre, A. Botta, *A comprehensive survey on the security of low power wide area networks for the Internet of Things*, **ICT Express,** vol. 10, no. 3, pp. 519-552, 2024, DOI: 10.1016/j.icte.2024.03.003.
[P3]	R. Canonico, G. Esposito, A. Navarro, S.P. Romano, G. Sperlì, A. Vignali *An Anomaly-based Approach for Cyber-Physical Threat Detection using Network and Sensor Data,* **Computer Communications,** vol. 234, pp. 108087, 2025, DOI: 10.1016/j.comcom.2025.108087.
[P4]	R. Canonico, G. Esposito, A. Navarro, S.P. Romano, G. Sperlì, A. Vignali, *Empowered Cyber-Physical Systems Security using both Network and Physical Data,* **Computers & Security**, vol. 152, pp. 104382, 2025, DOI: j.cose.2025.104382.
[P5]	R. Canonico, F. Lista, A. Navarro, G. Sperlì and A. Vignali, *Threat Detection in reconfigurable Cyber Physical Systems through Spatio-Temporal Anomaly Detection using Graph Attention Network,* **Computers & Security,** vol. 156, pp. 104509, 2025, DOI: 10.1016/j.cose.2025.104509.
[P6]	R. Botta, A. Canonico, A. Navarro, S. Ruggiero, G. Ventre, *AI-enabled SD-WAN: the case of Reinforcement Learning,* **IEEE Latin-American Conference on Communications (LATINCOM),** Rio de Janeiro, Brazil, November 30 – December 2, pp. 1-6, DOI: 10.1109/LATINCOM56090.2022.10000667
[P7]	A. Botta, R. Canonico, A. Navarro, G. Stanco, G. Ventre, *Scalable Reinforcement Learning for Dynamic Overlay Selection in SD-WANs,* **IFIP Networking Conference,** Barcelona, Spain, June 12-15, 2023, pp. 1-9, DOI: 10.23919/IFIPNetworking57963.2023.10186399

Research products

[P8]	A. Navarro, R. Canonico, A. Botta, *Software Defined Wide Area Networks: Current Challenges and Future Perspectives,* **IEEE 9th International Conference on Network Softwarization (NetSoft),** Madrid, Spain, June 19-23, pp. 350-353, DOI: 10.1109/NetSoft57336.2023.10175458
[P9]	Botta, R. Canonico, A. Navarro, G. Stanco, G. Ventre, *Towards a Highly-Available SD-WAN: Rapid Failover based on BFD Protocol,* **IEEE Conference on Network Functions Virtualization and Software-Defined Networking (NFV-SDN),** Dresden, Germany, Nov. 7-9, 2023, pp. 153-158, DOI: 10.1109/NFV-SDN59219.2023.10329617
[P10]	Navarro, A. Botta, A. Canonico, R. Stanco, G. Ventre, G. Buonocunto, A. Fresa, A. Gentile, V. Scommegna, L. Vicario, E. *Edge to Cloud Network Function Offloading in the ADAPTO Framework,* **Advanced Information Networking and Applications (proceedings of AINA 2024),** Apr. 9, 2024, pp. 69-78, Springer Nature Switzerland, DOI: 10.1007/978-3-031-57931-8_7.
[P11]	Botta, R. Canonico and A. Navarro, *Explainable Reinforcement Learning for Network Management via Surrogate Model,* **2024 IEEE 44th International Conference on Distributed Computing Systems Workshops (ICDCSW),** *Jersey City, NJ, USA, 2024, pp. 35-40, doi: 10.1109/ICDCSW63686.2024.00012.*
[P12]	A. Botta, R. Canonico, A. Navarro, G. Stanco, G. Ventre, A. Buonocunto, A. Fresa, V. Gentile, L. Scommegna, E. Vicario, E. Mingozzi, A. Virdis, M. Cucurachi, *ADAPTO: Scaling and Offloading Cloud-Native Network Functions in Future Mobile Networks,* **Advanced Information Networking and Applications (proceedings of AINA 2025),** Barcelona, Spain, Apr. 2025, pp. 337-346, Springer, DOI: 10.1007/978-3-031-87778-0_33.
[P13]	A. Navarro, A. Botta, R. Canonico, Y. Wang, F. H. P. Fitzek and G. T. Nguyen, *2OffRAN: Offline Off-Policy Reinforcement Learning for Safe Handover in O-RAN,* **IEEE International Conference on Machine Learning for Communication and Networking** Barcelona, Spain, May. 2025, pp. 1-6, IEEE, DOI: 10.1109/ICMLCN64995.2025.11140172.
[P14]	Zinno, S., Navarro, A., Rotbei, S., Pasquino, N., Botta, A., & Ventre, G., *A lightweight deep learning approach for latency prediction in 5G and beyond,* **1st Workshop on Integrated Wireless Networking and Computing (IWNC),** IEEE Conference on Network and Service Management (CNSM) Bologna, Italy, October. 2025, pp. 27-31, IEEE, [YET TO APPEAR]

Prizes and awards

13th TMA PhD School Best Poster Award

- «2OffRAN: Offline Off-Policy Reinforcement Learning for Safe Handover in O-RAN» Annalisa Navarro, Roberto Canonico, Alessio Botta

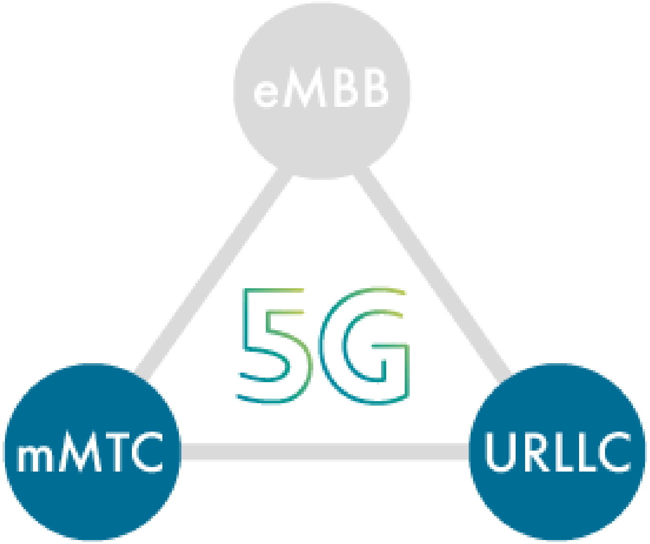
IEEE ComSoc CSIM TC Conference Best Paper Award 2023

- A. Botta, R. Canonico, A. Navarro, G. Stanco and G. Ventre, "Towards a Highly-Available SD-WAN: Rapid Failover based on BFD Protocol," 2023 IEEE Conference on Network Function Virtualization and Software Defined Networks (NFV-SDN), Dresden, Germany, 2023, pp. 153-158, doi: 10.1109/NFV-SDN59219.2023.10329617.

Travel Grants

- 2023 IEEE 9th International Conference on Network Softwarization (NetSoft)
- 2024 IEEE 44th International Conference on Distributed Computing Systems (ICDCS)
- 2025 Network Traffic Measurement and Analysis Conference (TMA)

Problem



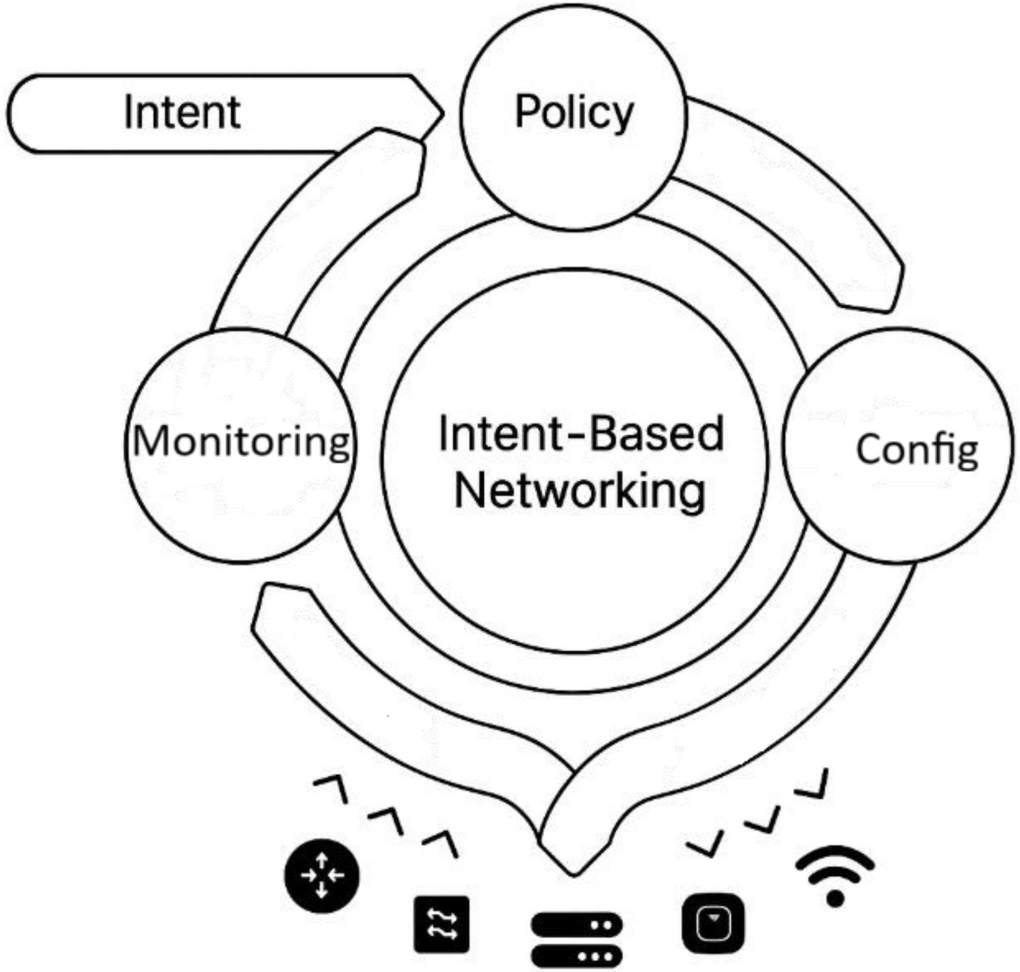
- **Next Generation Networks (5G/6G)** aim to support URLLC, massive IoT, and immersive applications with **stringent requirements**.
  - high throughput, ultra-low latency, scalability, and flexibility…

Traditional network management and control systems (manual, static) lack responsiveness and are error-prone, making them inadequate for 5G/6G networks.

- **Autonomous network orchestration** is the automated coordination of network resources to meet objectives under dynamic network conditions.

State of the Art

- To enable **autonomous orchestration**, networks are being redesigned for programmability and virtualization through **SDN** and **NFV**.
- Machine Learning, particularly **Reinforcement Learning** (RL), provides a promising framework to realize **network intents.**
- However, **its deployment in operational telecom networks remains limited due to complexity, risky exploration, and lack of transparency.**



Contribution

Automatically orchestrate next-gen telecom networks via Reinforcement Learning while guaranteeing 1) scalability, 2) explainability, 3) reliability.

Scalability

Challenge: RL complexity gets unfeasible when network size grows (e.g. number of sites, paths…)

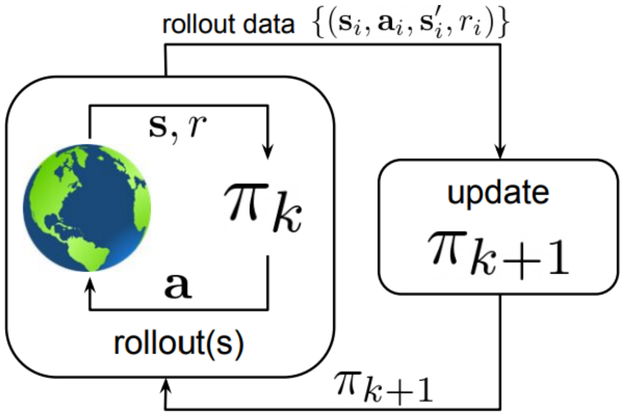
Explainability

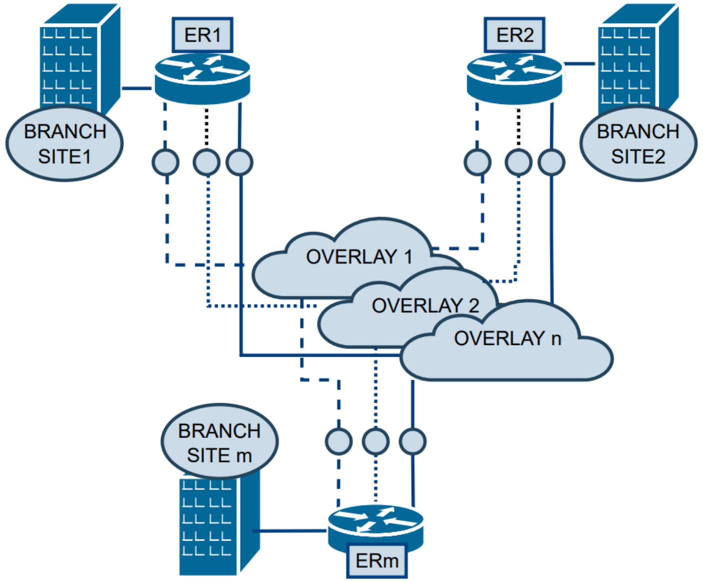
Challenge: RL is seen as black box: need to explain RL actions to foster network admins’ trust and adoption.

Reliability

Challenge: RL learns a policy while interacting with the network; it is critical to avoid network operations’ disruption.

Online RL for autonomous orchestration





The ***environment*** is the entire WAN, but only some parameters can be measured (i.e. site-to-site delay).

The ***reward*** is positive when latency remains below the threshold set via the network policy and negative otherwise.

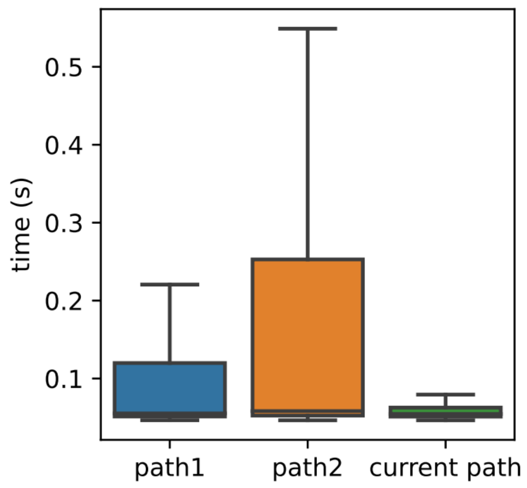
An ***action*** corresponds to instructing the ER to select one of the available overlays to forward traffic.

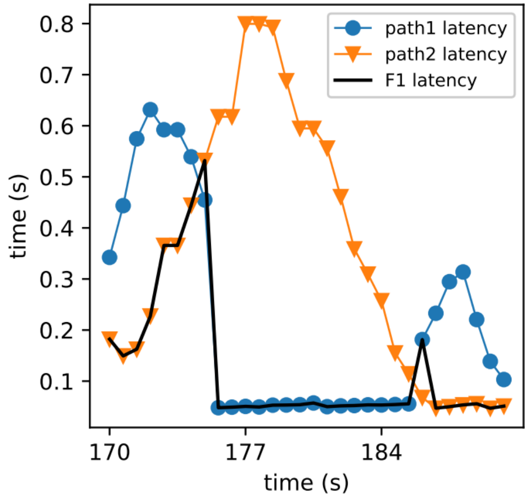
The ***state*** depends on the overlay used for each couple of sites and the condition of each overlay.

[P6] Online RL for autonomous orchestration: **results**

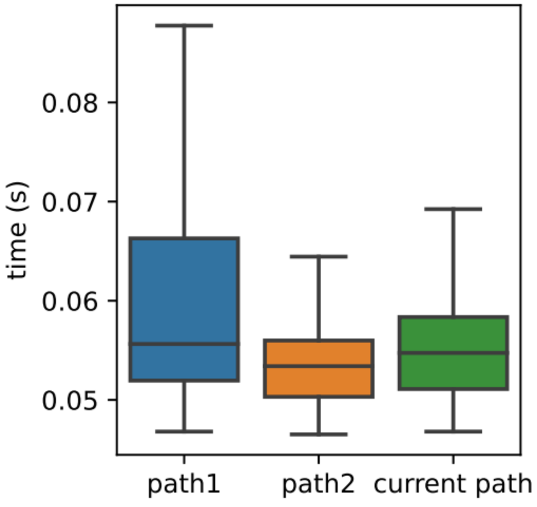
SD-WAN

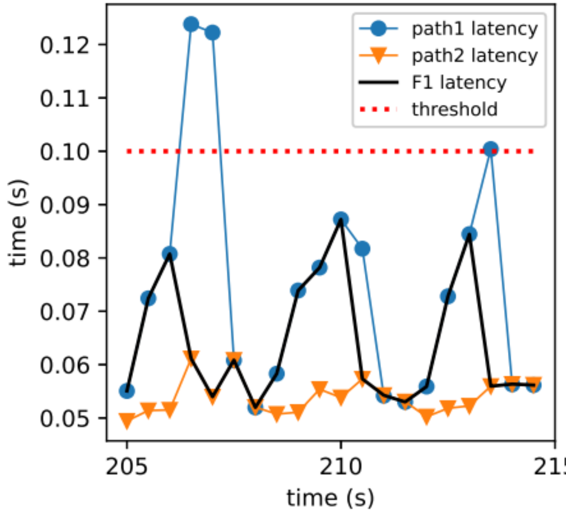
- **Network Intent based on latency**





- **Network Intent based on latency and cost**





- The SD-WAN flow manages to get lower latency values with respect to the delay measured on the individual alternative paths.
- The SD-WAN flow is routed to the path that is experiencing the lowest delay.

- The SD-WAN flow manages to get lower latency values with respect to the best-effort path.
- The SD-WAN flow preferentially uses the best-effort, less expensive path.
- The SD-WAN flow uses path 2 only when path 1 delay increases.

iteePhD information technology electrical engineering

Contribution

Automatically orchestrate next-gen telecom networks via Reinforcement Learning while guaranteeing **1) scalability**, 2) explainability, 3) reliability.

Scalability

Challenge: RL complexity gets unfeasible when network size grows (e.g. number of sites, paths…)

Explainability

Challenge: RL is seen as black box: need to explain RL actions to foster network admins' trust and adoption.

Reliability

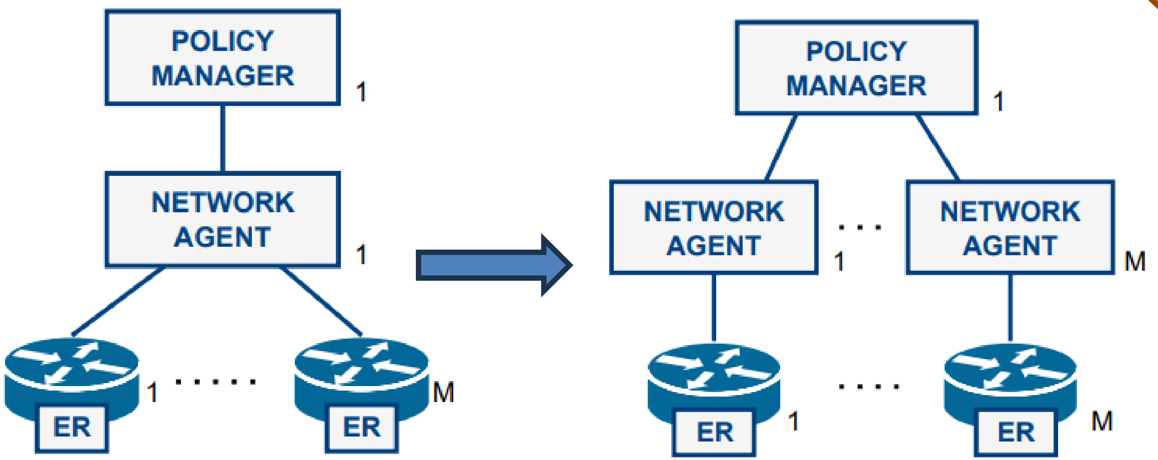
Challenge: RL learns a policy while interacting with the network; it is critical to avoid network operations' disruption.

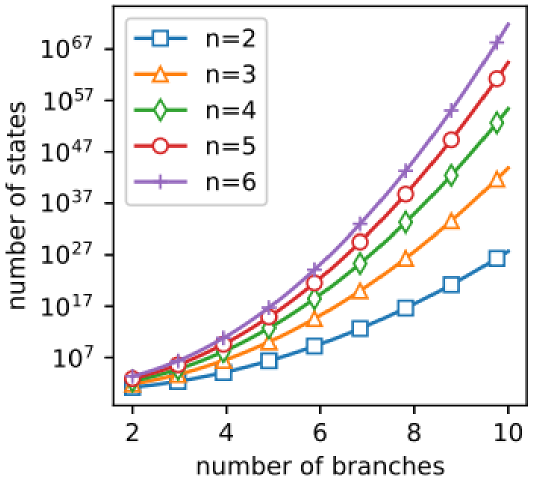
[P7] MARL for **scalable** orchestration

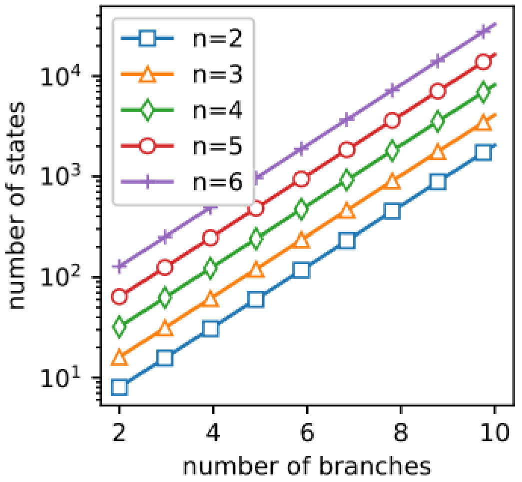
In complex SD-WAN scenarios with multiple sites and overlays (e.g. Enel) the number of actions and states of Single Agent RL explodes.

To overcome this issue, the overlay selection problem is addressed with a Multi Agent Reinforcement Learning (MARL) approach.

Complexity gets unfeasible with SARL, while MARL keeps complexity manageable.







SD-WAN

[P1] MARL-NA for scalable **global** orchestration

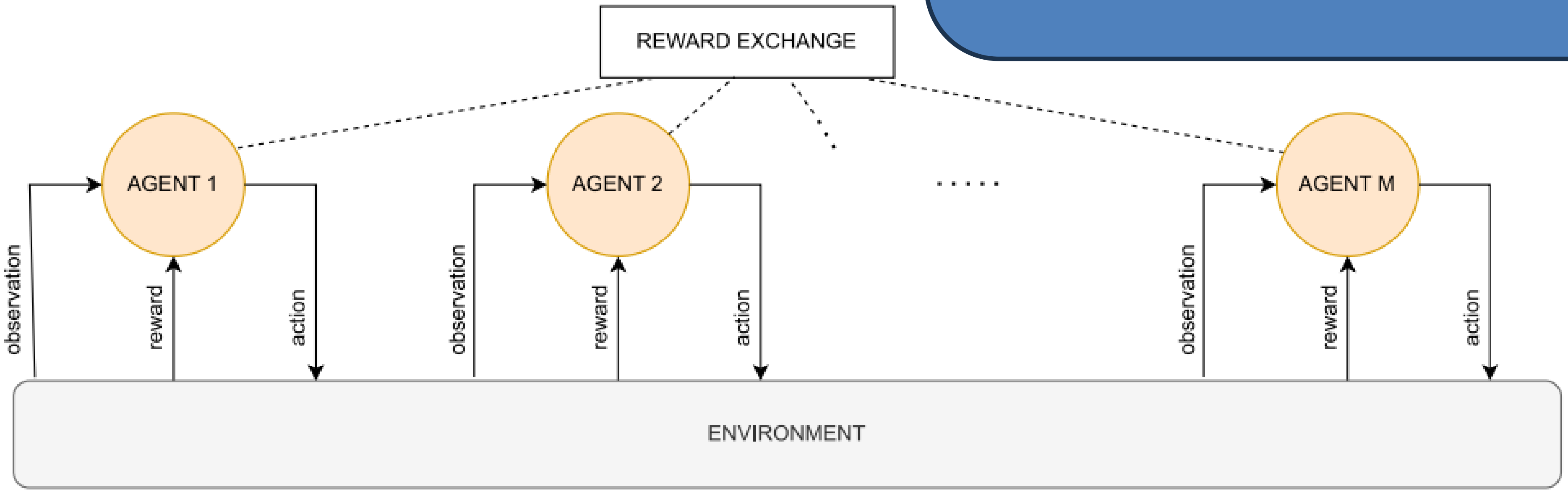
Fully Decentralized MARL

- Agents act independently
- Each agent receives its own reward

MARL with Networked Agents (MARL-NA)

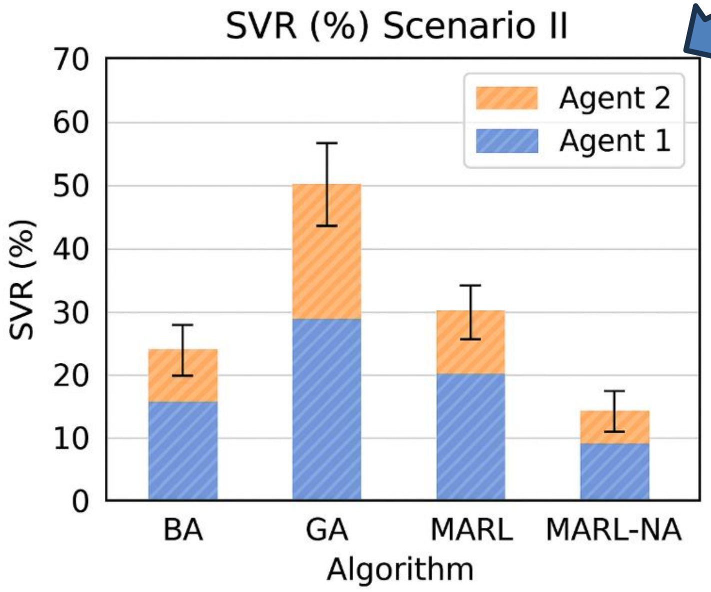
- Each agent receives its local reward and rewards from neighbors
- Each agent computes an adjusted reward:

$$r = r_i \cdot (1 - \lambda) + \lambda \cdot \sum_j NR_{ij}$$



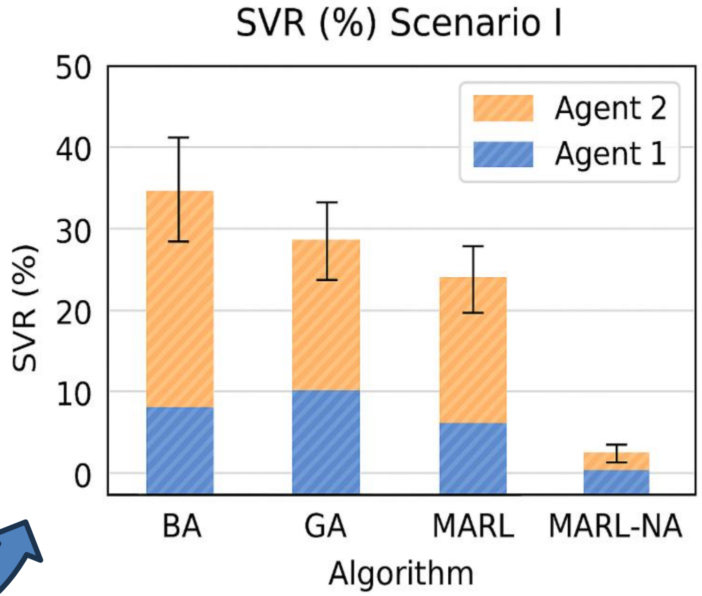
A. Botta, R. Canonico, A. Navarro, G. Stanco, G. Ventre, *Adaptive overlay selection at the SD-WAN edges: A reinforcement learning approach with networked agents*, **Elsevier Computer Networks**

MARL-NA for **scalable global orchestration**: results



***Network Intent*:**
A1, A2: *"Traffic from site1 to site3 and from site2 to site3 must experience a maximum delay of 300 ms"*

***Network Intent*:**
A1: *"Traffic from site1 to site3 must experience a maximum delay of 400 ms";*
A2: *"Traffic from site2 to site3 must experience a maximum delay of 200 ms"*



SVR: SLA violation ratio; BA: Blind Algorithm; GA: Greedy Algorithm

Contribution

Automatically orchestrate next-gen telecom networks via Reinforcement Learning while guaranteeing 1) scalability, **2) explainability**, 3) reliability.

Scalability

Challenge: RL complexity gets unfeasible when network size grows (e.g. number of sites, paths…)

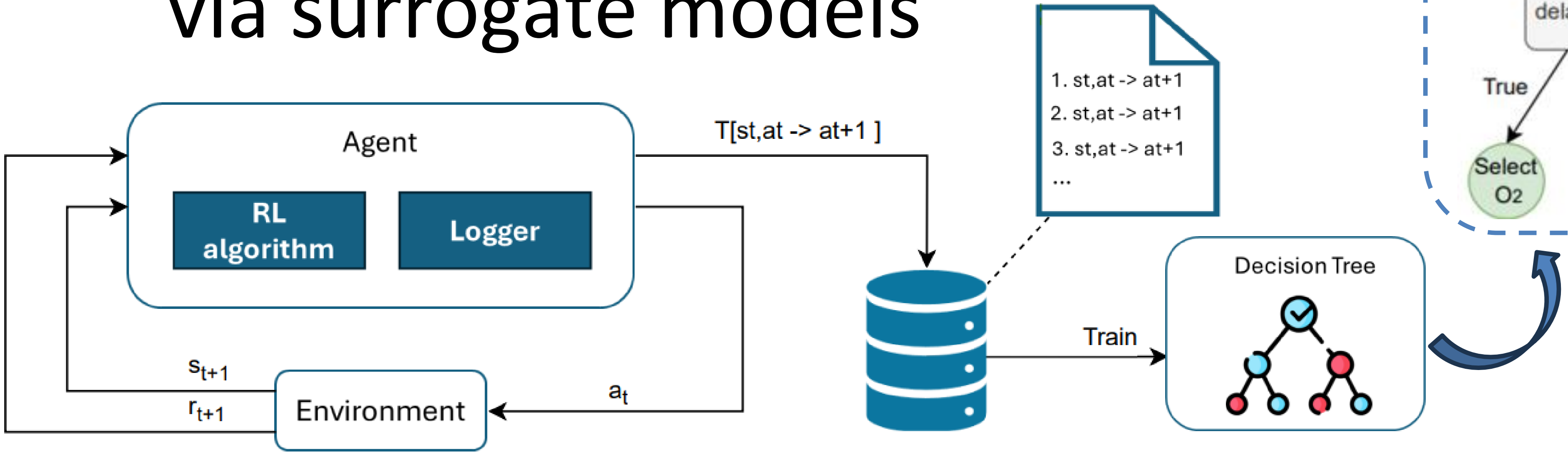
Explainability

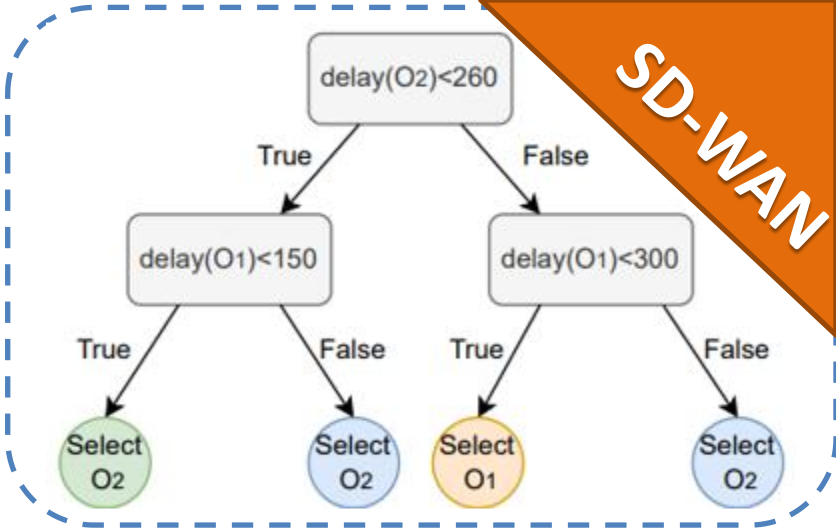
Challenge: RL is seen as black box: need to explain RL actions to foster network admins' trust and adoption.

Reliability

Challenge: RL learns a policy while interacting with the network; it is critical to avoid network operations' disruption.

[P11] **Explainable** orchestration via surrogate models





- Build an interpretable model from a black-box complex RL agent, through imitation.
- **Decision Trees (DTs)** can be used to create a simplified, interpretable representation.
- Once the RL agent is deployed, the state-action transitions T are logged.
- These data samples are then used to train a DT that learns the decision taken by the RL agent based on the previous decision and the current state.
- The DT is used to explain the RL model behavior.

Contribution

Automatically orchestrate next-gen telecom networks via Reinforcement Learning while guaranteeing 1) scalability, 2) explainability, **3) reliability**.

Scalability

Challenge: RL complexity gets unfeasible when network size grows (e.g. number of sites, paths…)

Explainability

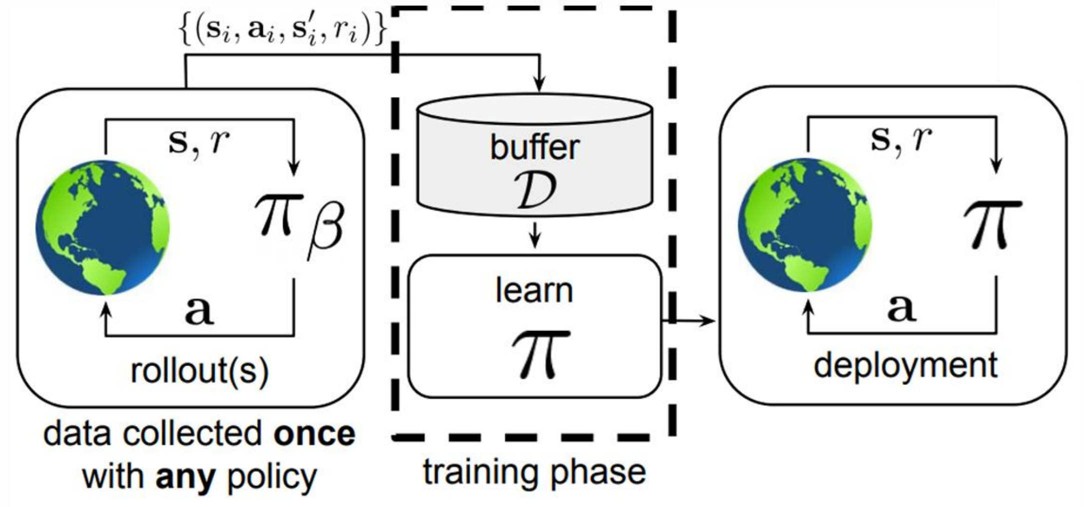
Challenge: RL is seen as black box: need to explain RL actions to foster network admins' trust and adoption.

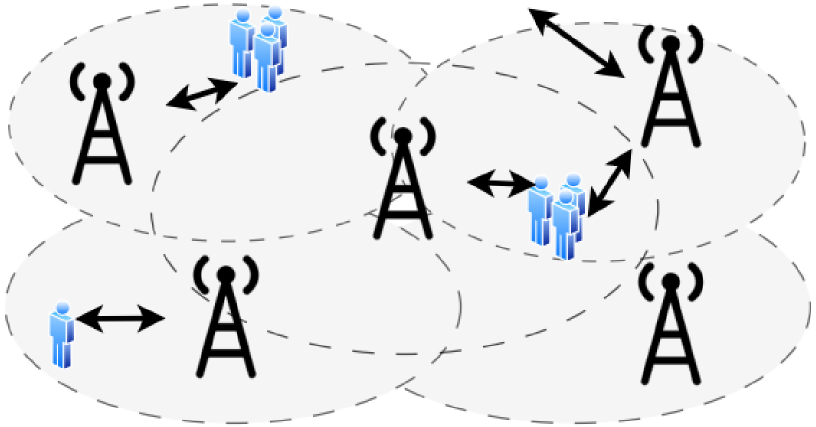
Reliability

Challenge: RL learns a policy while interacting with the network; it is critical to avoid network operations' disruption.

Offline RL for autonomous **reliable** orchestration

O-RAN





The ***environment*** is the entire **O-RAN** system, including all base stations and user equipments.

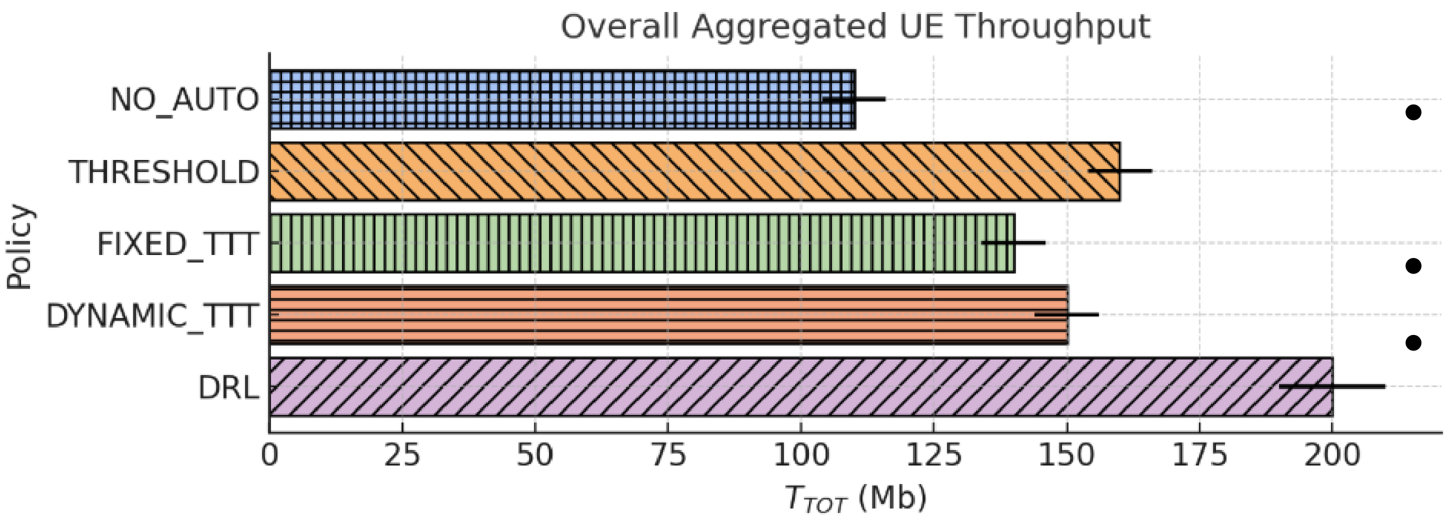
The ***reward*** is based on the number of UEs achieving the target throughput established with the policy.

An ***action*** corresponds to assigning each UE to a serving base station at the next time step.

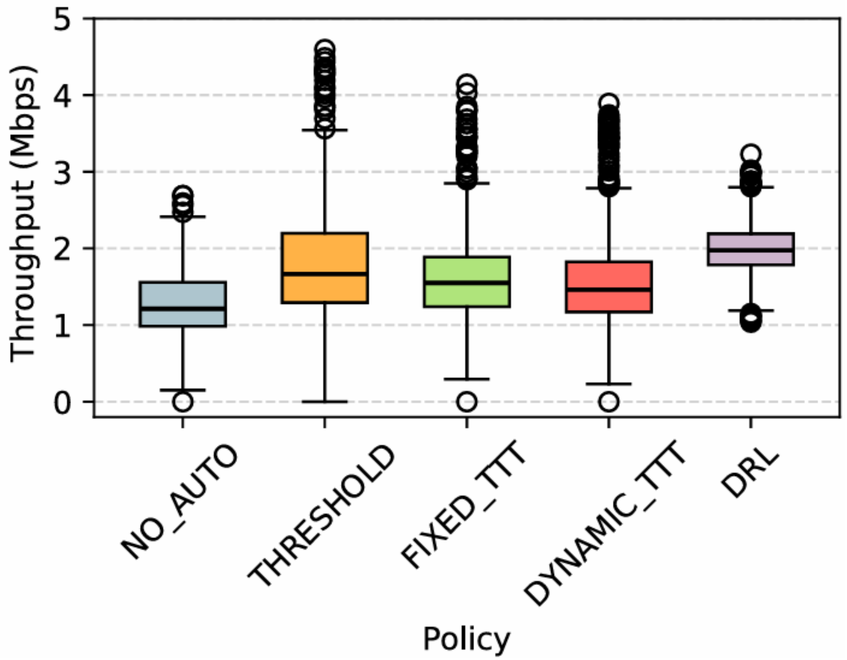
The ***state*** depends on the *current serving BS of each UE*, the *SINR* values from serving and neighboring BSs, the *number of active UEs per BS*, and the *throughput achieved by each UE*.

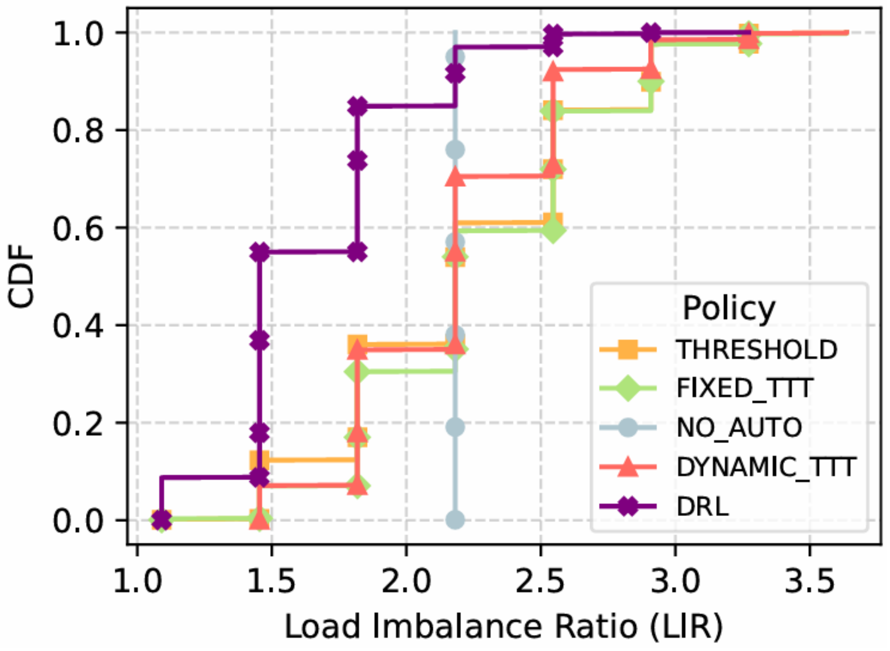
[P13] Offline Off-Policy RL: results

O-RAN



- DRL achieves higher and more stable throughput compared to static policies.
- DRL lowers the load imbalance.
- The offline off-policy approach avoids disrupting the normal network operation.





itee PhD
information technology
electrical engineering

Conclusion

- Traditional network management and control are inadequate for **5G and beyond.**
- **Reinforcement Learning (RL)** is a key enabler of **autonomous network orchestration**.
- Identified and addressed major challenges limiting RL adoption in real networks: **scalability, explainability, and reliability**.
  - Proposed **decentralized cooperative MARL** to improve scalability and enable global policy compliance.
  - Introduced **explainable RL** via surrogate models to improve transparency and operator trust.
  - Designed an **offline RL framework** to enhance reliability, building a safe-to-deploy agent.

This thesis contributes to bridging the gap between RL research and its practical adoption in real next-generation networks, enabling autonomous network orchestration.

Thank you for your attention.