Network Optimization with GNNs and Deep Reinforcement Learning

Workshop on Explainability and Applicability of Graph Neural Networks
06/09/23 - 08/09/23 Kassel, Germany

Paul Almasan

Joint work with José Suárez-Varela, Krzysztof Rusek, Prof. Pere Barlet-Ros and Prof. Albert Cabellos-Aparicio

Telefónica

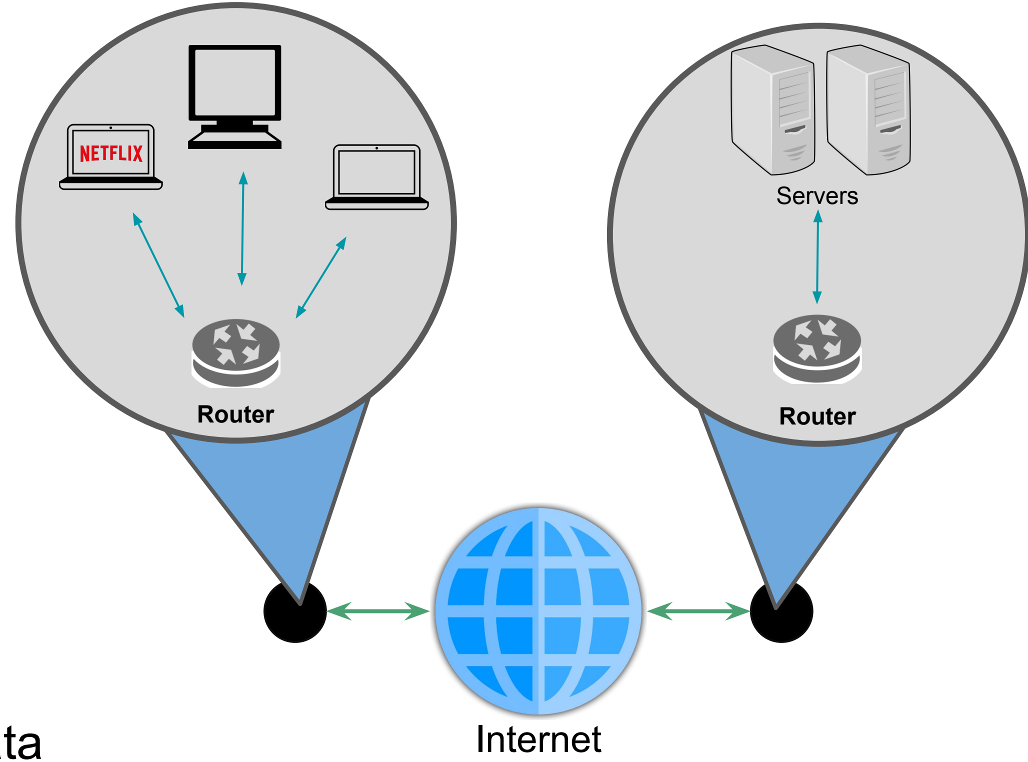
How can we leverage GNNs for network optimization?

Outline

- **What** are computer networks and **why** should we manage them efficiently?
- **How** can we efficiently manage computer networks?
- DRL meets GNNs: Routing Optimization in Optical Networks

What are computer networks and why should we manage them efficiently?

A collection of computers, servers, and other devices linked together for the purpose of **sharing data and resources**



Key Components:

1. **Nodes**: Devices such as computers, servers, routers, and switches
2. **Links**: Physical cables or wireless connections used for data transmission
3. **Protocols**: Set of rules governing data exchange and communication

Computer networks experienced a considerable **growth in novel networked applications, network traffic and connected devices**[1] in the last years

- Stringent network requirements

- High throughput

- Adapt to dynamic topology

- Low deterministic latency

[1]Cisco Annual Internet Report (2018–2023) White Paper

Computer networks experienced a considerable **growth in novel networked applications, network traffic and connected devices**[1] in the last years

- Stringent network requirements

- High throughput

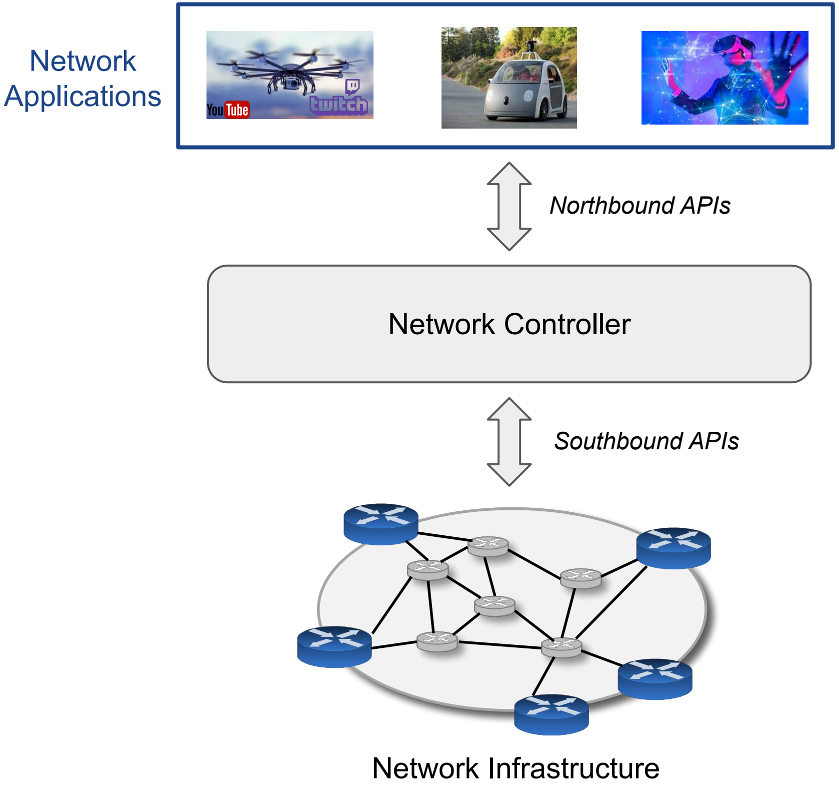
- Adapt to dynamic topology

- Low deterministic latency

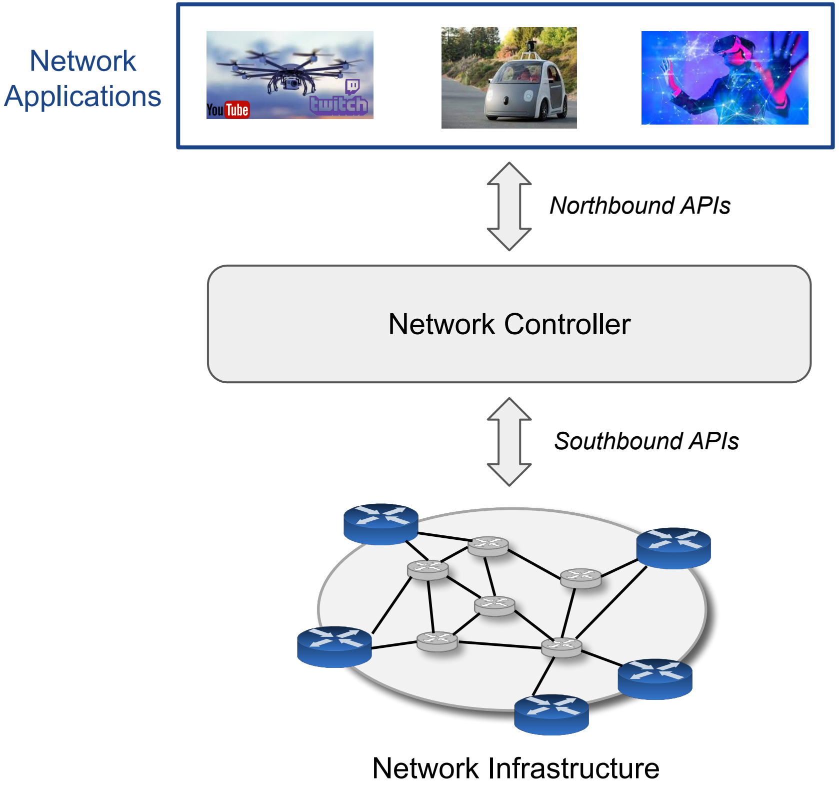
Computer networks are becoming more complex and costly to manage

[1]Cisco Annual Internet Report (2018–2023) White Paper

Why is network management important?



Why is network management important?



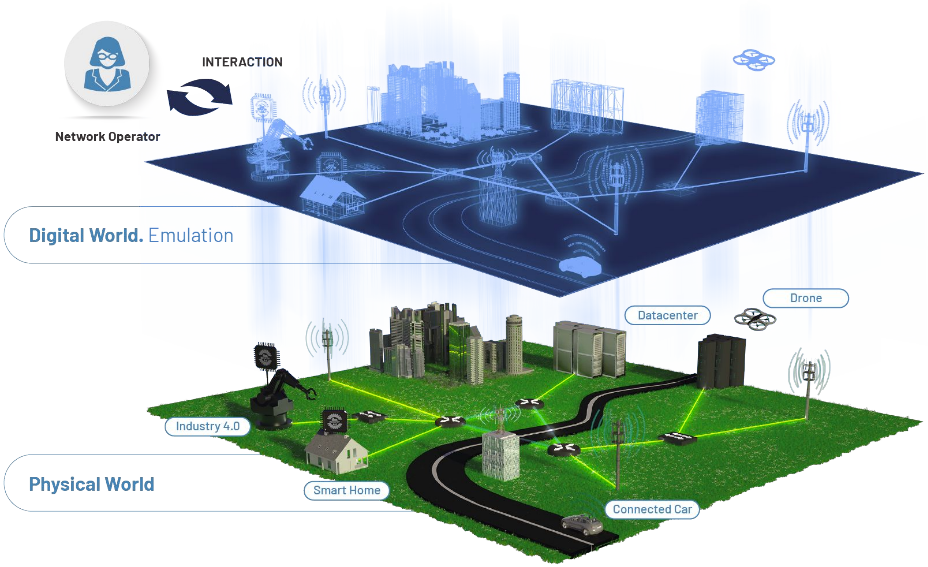
Different network applications with **heterogeneous network requirements**

- ✓ Real-time
- ✓ Ultra-low latency
- ✓ High throughput
- ✓ Reliability

Network management is **costly**

<u>How</u> can we efficiently manage computer networks?

The **Network Digital Twin**[1,2] **(NDT)** paradigm emerged as a **key enabler for efficient control and management** of modern networks



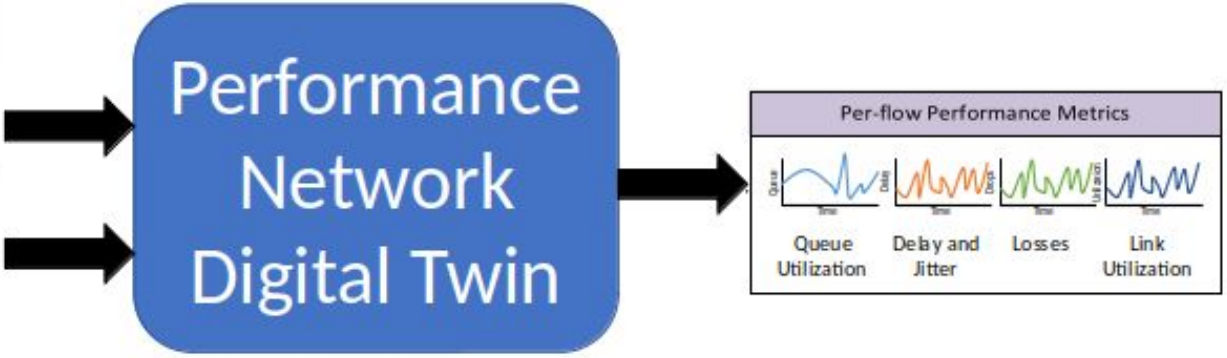
- ✔ Digital ML-based network model
- ✔ Mimics real network's behaviour
- ✔ Accurate estimated performance metrics
- ✔ Fast inference

[1]Paul Almasan, et al. "Network Digital Twin: Context, Enabling Technologies and Opportunities," in IEEE Communications Magazine, doi: 10.1109/MCOM.001.2200012.
[2]Wu, Y., Zhang, K., & Zhang, Y. (2021). Digital twin networks: A survey. IEEE Internet of Things Journal, 8(18), 13789-13804.

Performance Network Digital Twin

Network Configuration
- Topology
- Link Capacity
- Routing

Traffic Load



Rusek, K., Suárez-Varela, J., Almasan, et. al. (2020). Routenet: Leveraging graph neural networks for network modeling and optimization in sdn. *IEEE Journal on Selected Areas in Communications*, *38*(10), 2260-2270.
Ferriol-Galmés, M., Paillisse, J., Suárez-Varela, J., Rusek, K., Xiao, S., Shi, X., ... & Cabellos-Aparicio, A. (2023). RouteNet-Fermi: Network Modeling With Graph Neural Networks. *IEEE/ACM Transactions on Networking*.

There are relevant academic, industrial and standardization efforts put to make the NDT become a reality

- Research[1,2,3]

- Industry

Network Digital Twin

Use a virtual replica of your network to make your i NOKIA more efficient and productive

Press release MWC 2023

Telefónica develops a digital twin to improve the production of industrial parts

ericsson.com Products and solutions Discover Transforming enterprises Future technologies

HOME › REPORTS AND PAPERS › ... › ERICSSON TECHNOLOGY REVIEW ARTICLES › NETWORK DIGITA

Network digital twins – outlook and opportunities

- Standards[4,5]

Workgroup: Internet Research Task Force
Internet-Draft: draft-irtf-nmrg-network-digital-twin-arch-02
Published: 24 October 2022
Intended Status: Informational
Expires: 27 April 2023
Authors:
C. Zhou, *China Mobile*; H. Yang, *China Mobile*; X. Duan, *China Mobile*; D. Lopez, *Telefonica I+D*; A. Pastor, *Telefonica I+D*; Q. Wu, *Huawei*; M. Boucadair, *Orange*; C. Jacquenet, *Orange*

IETF

Digital Twin Network: Concepts and Reference Architecture

ITU-T Y.3090 (02/2022)

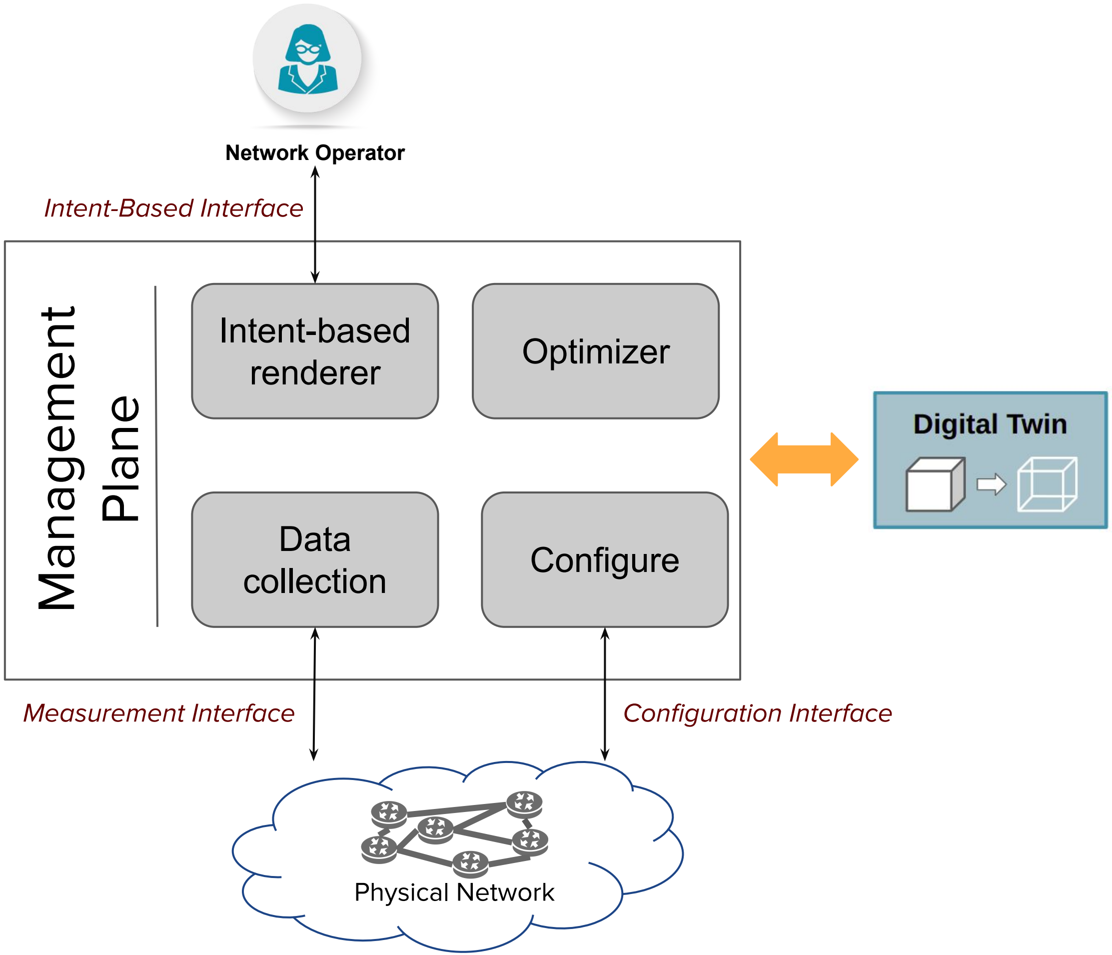
ITU International Telecommunication Union

Digital twin network – Requirements and architecture

Recommendation ITU-T Y.3090 describes the requirements and architecture of a digital twin network (DTN). A digital twin network is a virtual representation of a physical network. It is useful for analysing, diagnosing, emulating, and controlling the physical network based on data, model, and interface to achieve a real-time interactive mapping between a physical network and a digital twin network.

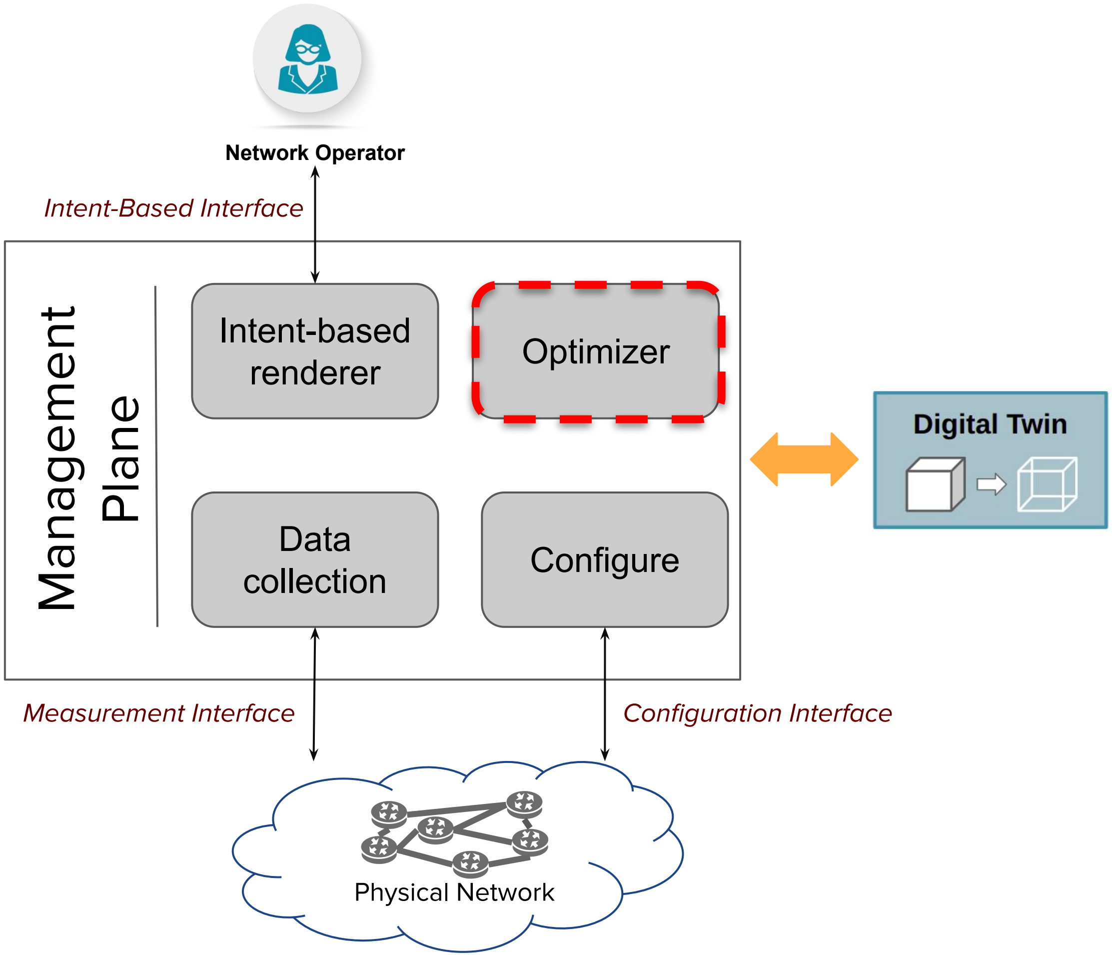
[1]Wu, Y., Zhang, K., & Zhang, Y. (2021). Digital twin networks: A survey. IEEE Internet of Things Journal, 8(18), 13789-13804.
[2]Nguyen, H. X., Trestian, R., To, D., & Tatipamula, M. (2021). Digital twin for 5G and beyond. IEEE Communications Magazine, 59(2), 10-15.
[3]Ferriol-Galmés, M., et al. (2022). Building a digital twin for network optimization using graph neural networks. Computer Networks, 217, 109329.
[4]C. Zhou et al., “Digital Twin Network: Concepts and Reference Architecture,” IETF, Internet-Draft, 2022.
[5]ITU-T, “Digital twin network: Requirements and architecture,” Recommendation ITU-T Y.3090, 2022.

Network Digital Twin general architecture[1]



[1]https://datatracker.ietf.org/doc/html/draft-paillisse-nmrg-performance-digital-twin-01

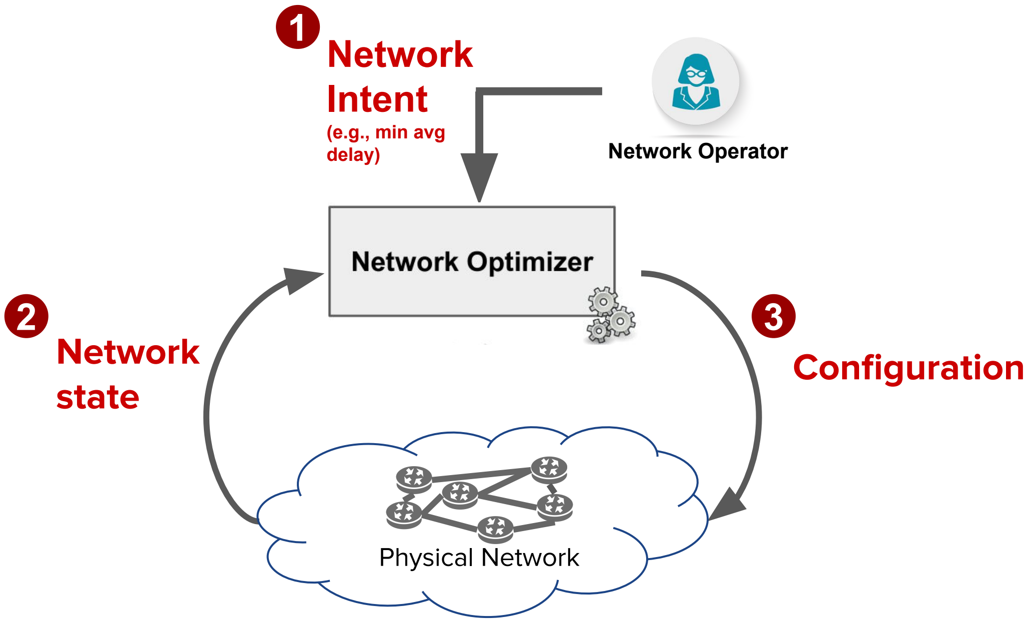
Network Digital Twin general architecture[1]



[1]https://datatracker.ietf.org/doc/html/draft-paillisse-nmrg-performance-digital-twin-01

Network optimization consists on **using effectively the network resources**

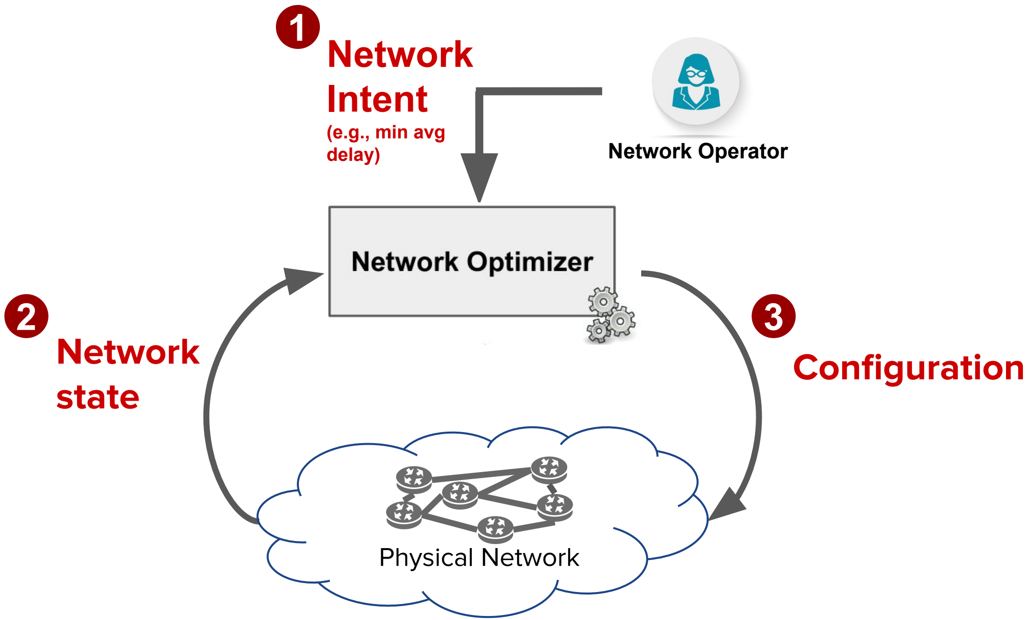
- ✔ Better utilization of limited resources
- ✔ Reduces operational costs
- ✔ Increase QoE



Network optimization consists on **using effectively the network resources**

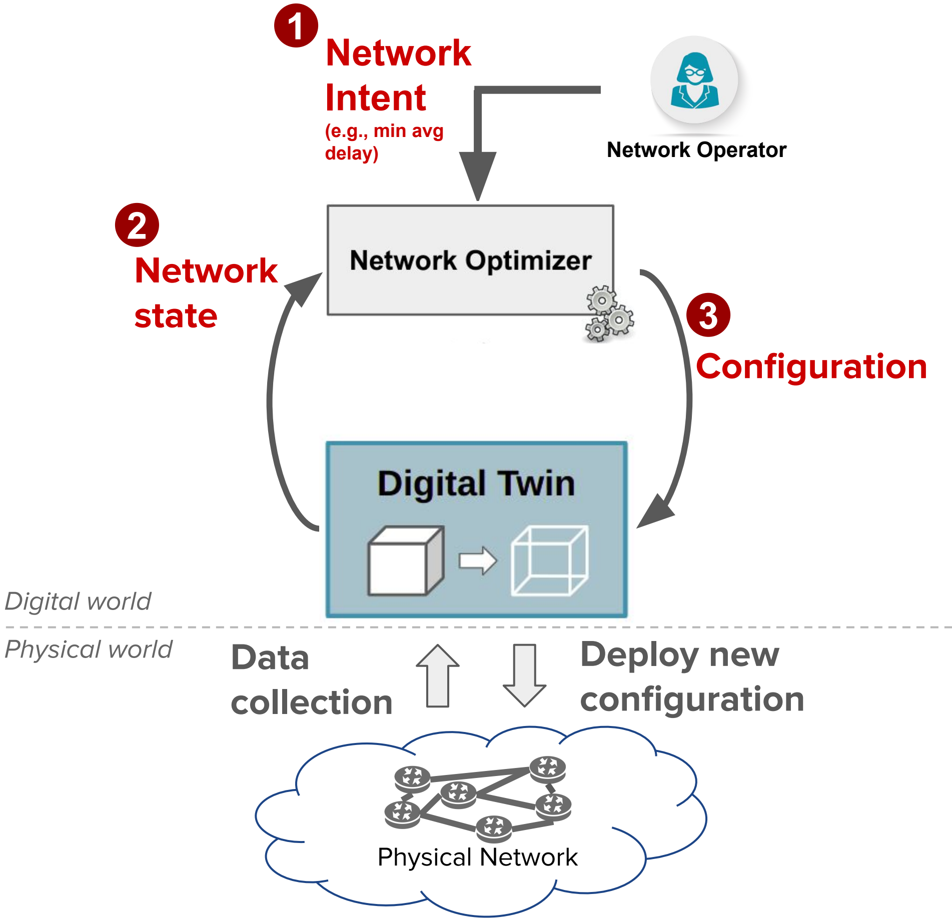
- ✔ Better utilization of limited resources
- ✔ Reduces operational costs
- ✔ Increase QoE

- ✖ Break the network
- ✖ Costly process
- ✖ Online optimization



NDTs enable **efficient real-time network optimization**[1,2]

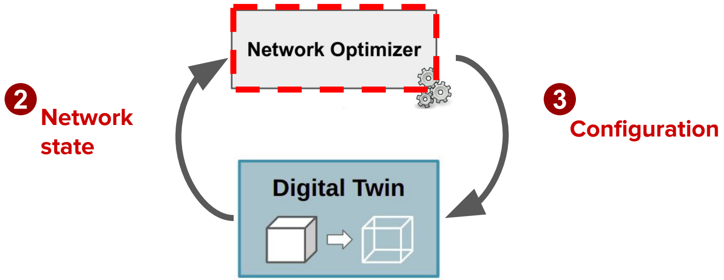
- ✔ **Test new configurations** without compromising the physical network
- ✔ **High quality** configurations
- ✔ **Online** optimization



[1]Paul Almasan, et al. "Network Digital Twin: Context, Enabling Technologies and Opportunities," in IEEE Communications Magazine), doi: 10.1109/MCOM.001.2200012.
[2]Ferriol-Galmés, M., et al. (2022). Building a digital twin for network optimization using graph neural networks. Computer Networks, 217, 109329.

How can we implement the Network Optimizer?

Method*	Execution cost	Performance
Heuristics[1,2]	Low	Low
Mathematical Optimizers[3,4] (e.g., CP, ILP)	High	High
SoA Machine Learning[5,6]	High (training)	High



*Existing methods in 2019-2021

[1]Fortz, B., & Thorup, M. (2000, March). Internet traffic engineering by optimizing OSPF weights. In Proceedings IEEE INFOCOM 2000. (Vol. 2, pp. 519-528).

[2]Wang, N., Ho, K. H., Pavlou, G., & Howarth, M. (2008). An overview of routing optimization for internet traffic engineering. IEEE Communications Surveys & Tutorials, 10(1), 36-56.

[3]Gong, L., et. al. (2013). Efficient resource allocation for all-optical multicasting over spectrum-sliced elastic optical networks. Journal of Optical Communications and Networking, 5(8), 836-847.

[4]Hartert, R., Vissicchio, et. al. (2015). A declarative and expressive approach to control forwarding paths in carrier-grade networks. ACM SIGCOMM computer communication review, 45(4), 15-28.

[5]Chen, X., et. al. (2018). Deep-RMSA: A deep-reinforcement-learning routing, modulation and spectrum assignment agent for elastic optical networks. In Optical Fiber Communication Conference (pp. W4F-2).

[6]Suárez-Varela, et. al. (2019). Routing in optical transport networks with deep reinforcement learning. Journal of Optical Communications and Networking, 11(11), 547-558.

SoA Machine Learning

We were failing to learn in computer networks

- In some cases **worse than simple well-known heuristics**
- **Ad-hoc solutions tailored to specific problems,** in some cases transforming the problem to prevent learning graph

Poor performance when evaluated **on data different than in training (does not generalize)**

- Require **re-training the ML model when there is a change in the network** (e.g., link failure)

SoA Machine Learning

We were failing to learn in computer networks

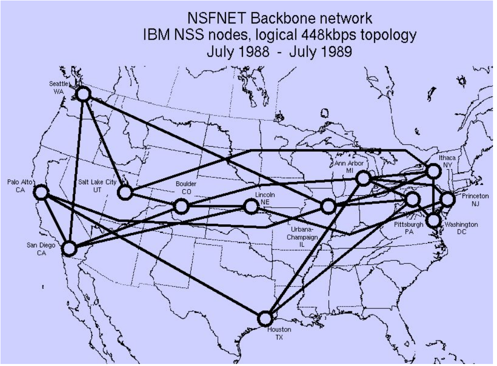
- In some cases **worse than simple well-known heuristics**
- **Ad-hoc solutions tailored to specific problems,** in some cases transforming the problem to prevent learning graph

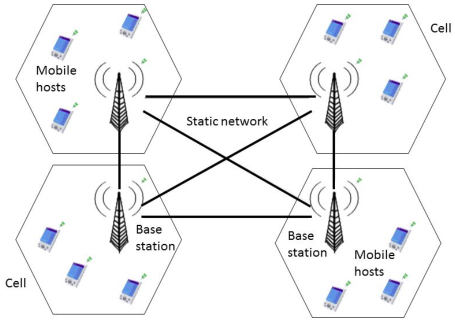
Poor performance when evaluated **on data different than in training (does not generalize)**

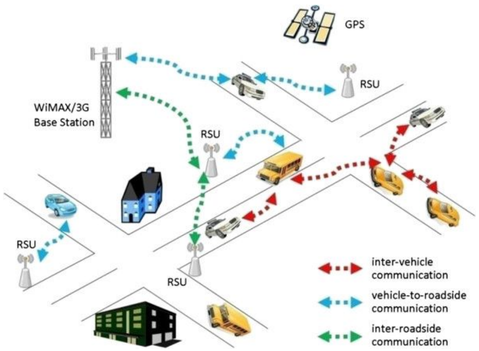
- Require **re-training the ML model when there is a change in the network** (e.g., link failure)

The main reason for this is that standard Neural Networks <u>are not suited</u> to learn information structured as a graph

Networks are fundamentally represented as graphs

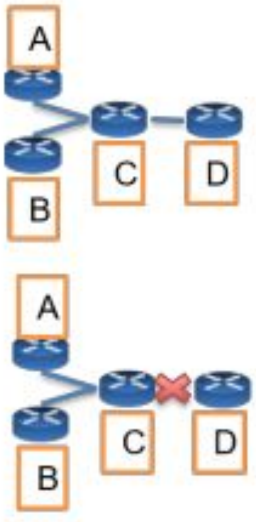






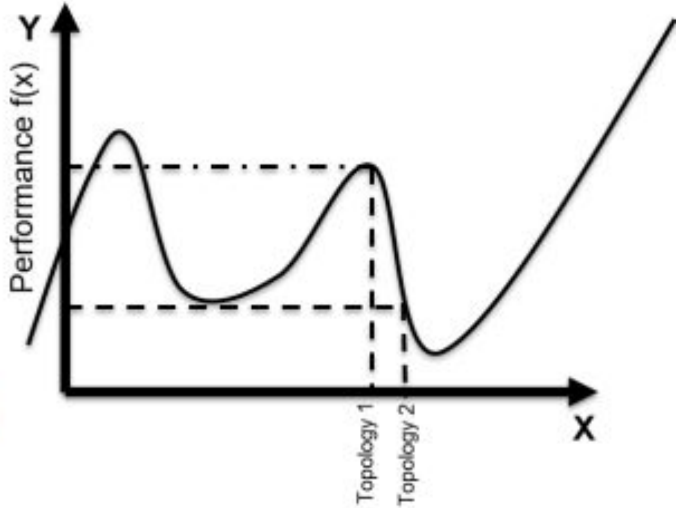
Limitations of traditional NNs

- Networks are variable in size (number of links and nodes)
- Information is relational
- Modeling networks with traditional NNs is very hard



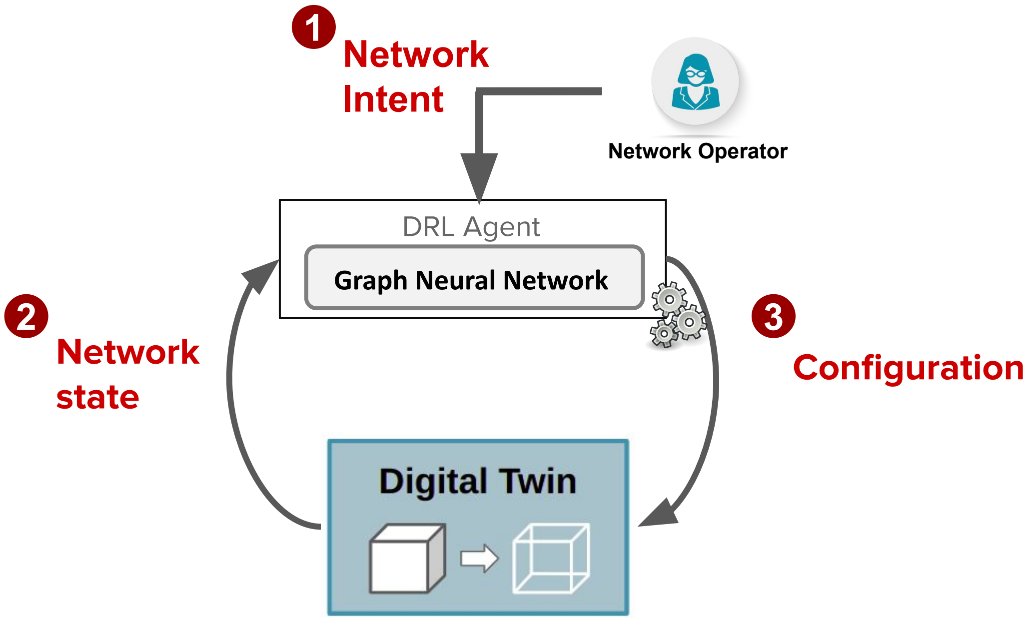
	A	B	C	D
A	1	0	1	0
B	0	1	1	0
C	1	1	1	1
D	0	0	1	1

	A	B	C	D
A	1	0	1	0
B	0	1	1	0
C	1	1	1	0
D	0	0	0	1



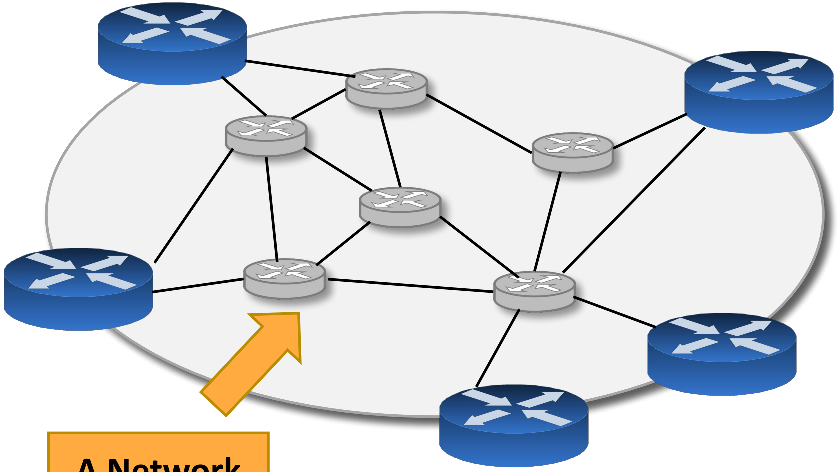
Similarity

We proposed a **Deep Reinforcement Learning + Graph Neural Networks architecture for routing optimization**[1]



[1] Paul Almasan, et al. "Deep reinforcement learning meets graph neural networks: Exploring a routing optimization use case." Computer Communications, 196, pp. 184-194, 2022.

Routing Optimization



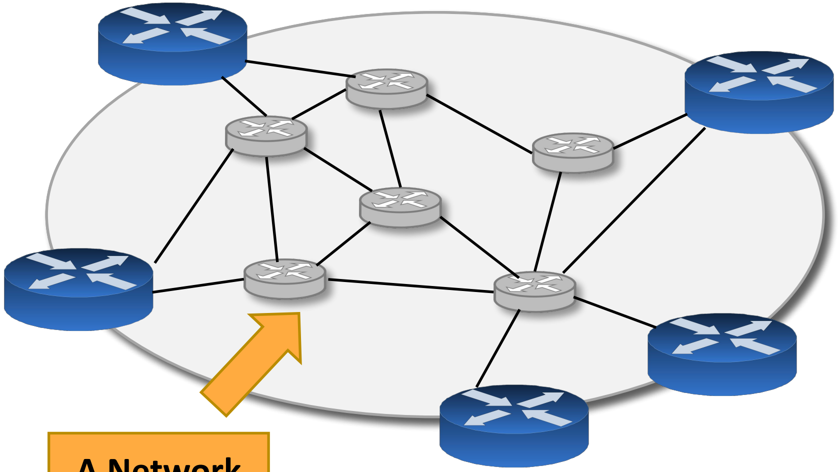
A Network Topology

Routing Optimization

A Traffic Matrix: Source to Destination

	A	B	C	D	E
A	0	2	3	2	9
B	1	0	2	1	6
C	2	4	0	1	1
D	1	6	2	0	5
E	9	2	3	1	0

Traffic Matrix (TM)



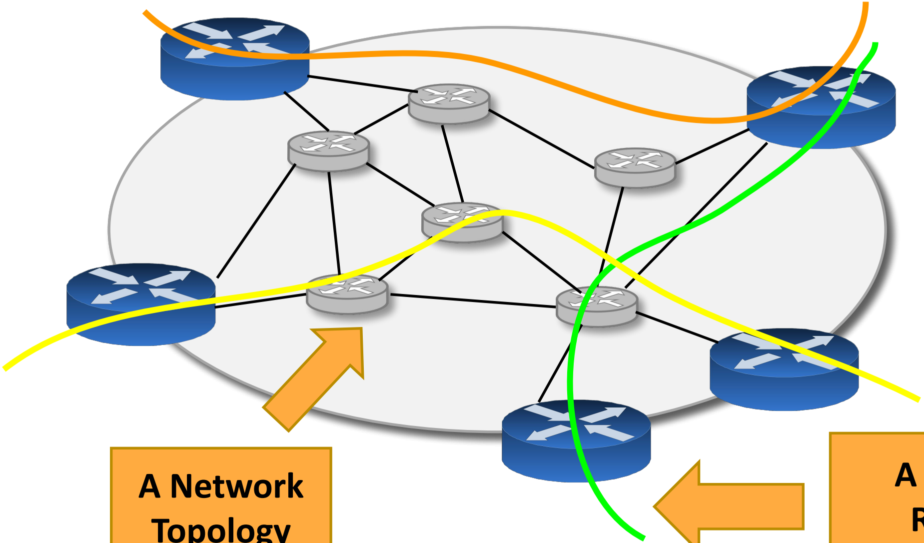
A Network Topology

Routing Optimization

A Traffic Matrix: Source to Destination

	A	B	C	D	E
A	0	2	3	2	9
B	1	0	2	1	6
C	2	4	0	1	1
D	1	6	2	0	5
E	9	2	3	1	0

Traffic Matrix (TM)



A Network Topology

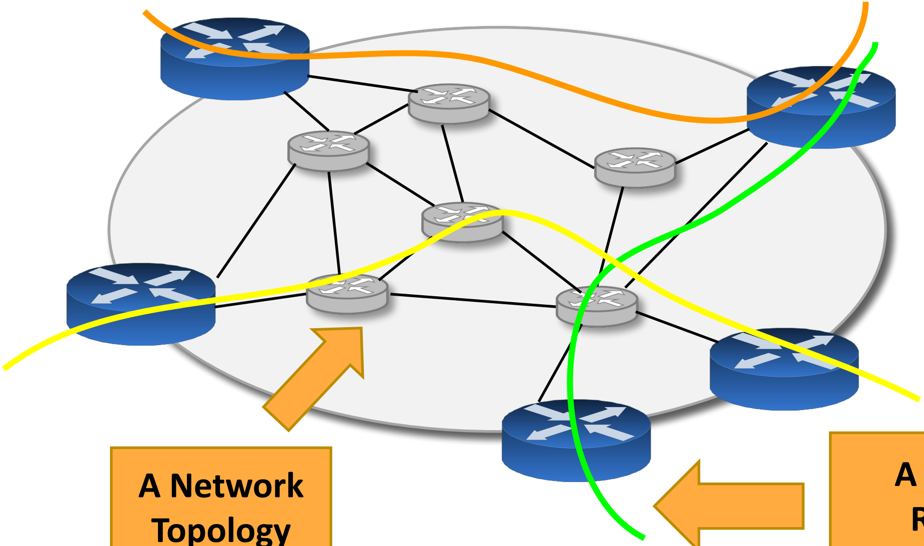
A src-dest Routing Configuration

Routing Optimization

A Traffic Matrix: Source to Destination

	A	B	C	D	E
A	0	2	3	2	9
B	1	0	2	1	6
C	2	4	0	1	1
D	1	6	2	0	5
E	9	2	3	1	0

Traffic Matrix (TM)



A Network Topology

A src-dest Routing Configuration

Which is the best routing configuration that satisfies some constraint? E.g., minimize link congestion

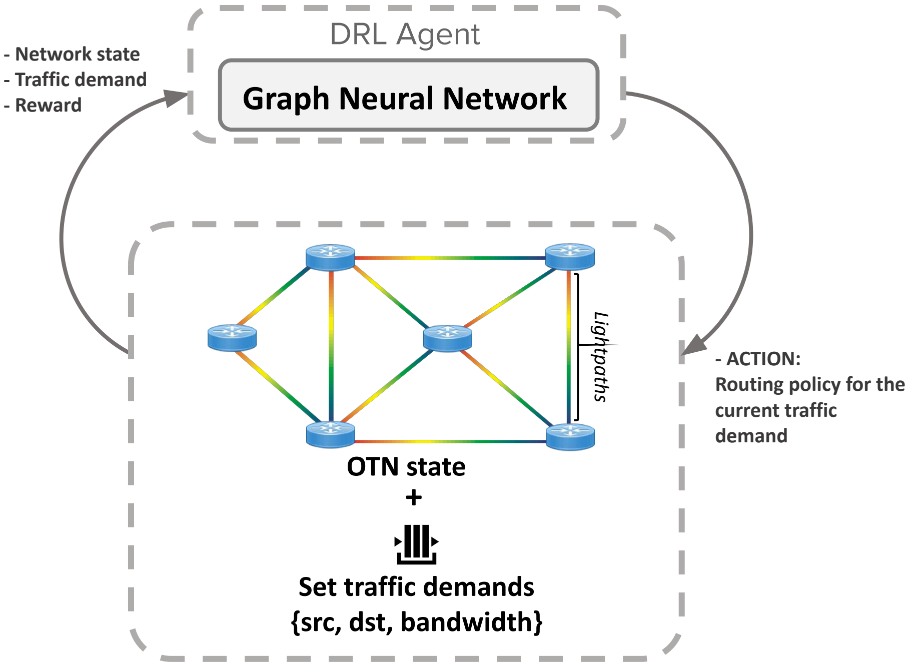
DRL meets GNNs: Routing Optimization in Optical Networks

Routing optimization scenario[1]

1. **Network state** composed by a network topology with link capacities

2. Set of **network traffic demands**

3. **An agent that allocates** the incoming **traffic demands** on the network state



Objective → Maximize the number of traffic demands allocated

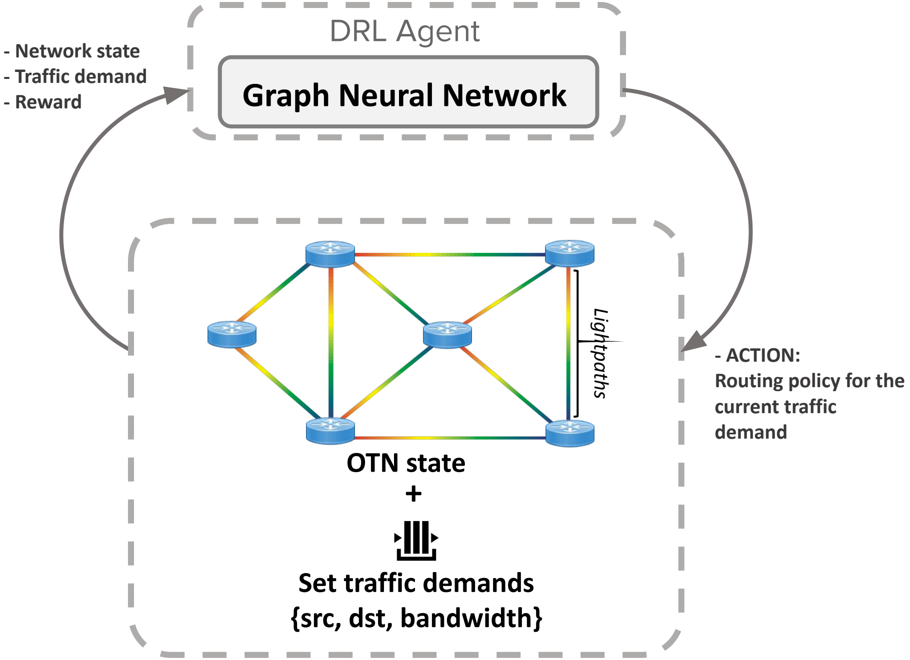
[1]Paul Almasan, et al. "Deep reinforcement learning meets graph neural networks: Exploring a routing optimization use case." Computer Communications, 196, pp. 184-194, 2022.

Proposed solution

DRL Agent implements the **Deep Q-Network**[1] (DQN) learning method

Integrate GNNs into DRL agents and we design a **problem specific action space**

We used a **NDT implemented in python with the minimum functionalities** needed for this optimization problem

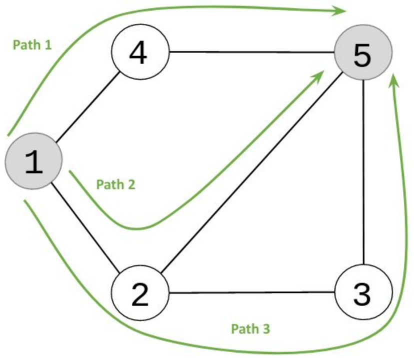


[1]Mnih, V., Kavukcuoglu, et al. (2015). Human-level control through deep reinforcement learning. Nature, 518(7540), 529-533.

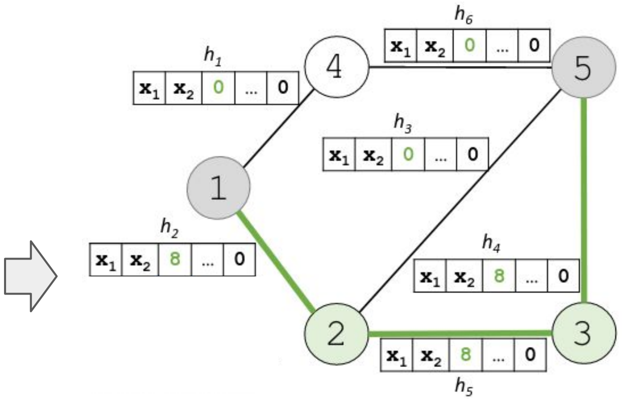
Action representation

Introduced within the network state

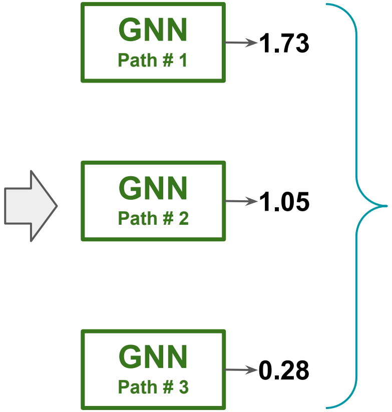
Choose one path to allocate from the pre-defined set of ***K* paths** for each traffic demand



Traffic demand: {1, 5, 8}

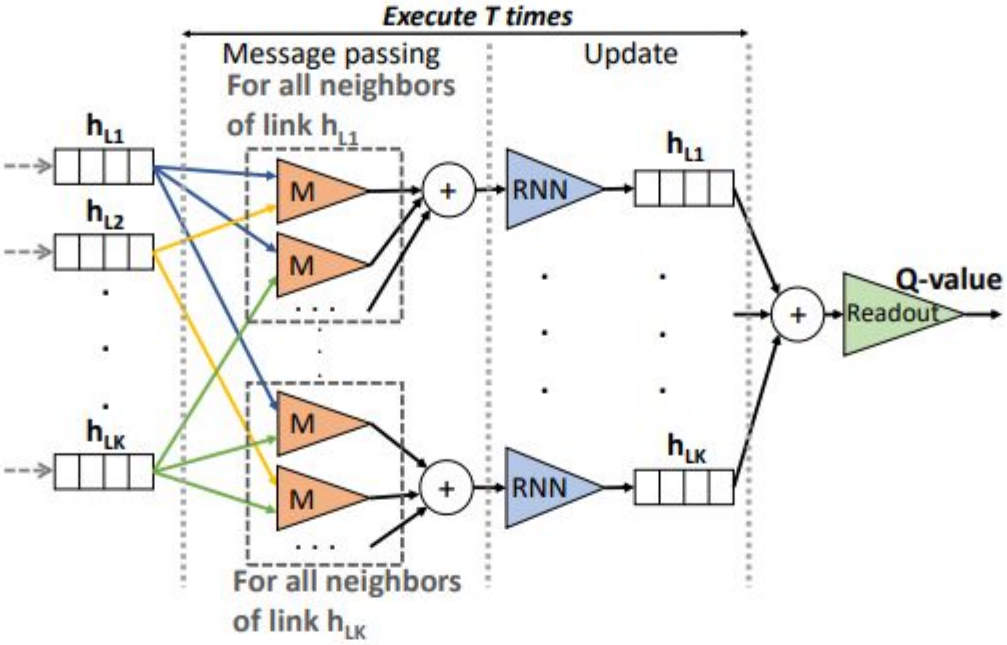


Notation	Description
x_1	Link available capacity
x_2	Link Betweenness
x_3	Action vector (bandwidth allocated)
$x_4 - x_N$	Zero padding



DRL agent samples action and applies it to network environment

GNN implementation



Algorithm 1 Message Passing

Input : $\mathbf{x}_l$
Output : $\mathbf{h}_l^T, q$

for each $l \in \mathcal{L}$ **do**
 $h_l^0 \leftarrow [\mathbf{x}_l, 0 \ldots, 0]$
for $t = 1$ **to** T **do**
 for each $l \in \mathcal{L}$ **do**
  $M_l^{t+1} = \sum_{i \in N(l)} m\left(h_l^t, h_i^t\right)$
  $h_l^{t+1} = u\left(h_l^t, M_l^{t+1}\right)$
$rdt \leftarrow \sum_{l \in \mathcal{L}} h_l$
$q \leftarrow R(rdt)$

---

[1]Paul Almasan, et al. "Deep reinforcement learning meets graph neural networks: Exploring a routing optimization use case." Computer Communications, 196, pp. 184-194, 2022.

Experimental Results

Baselines

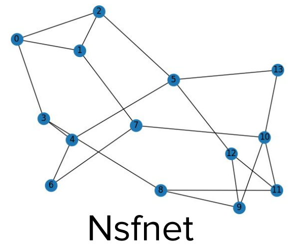
Action space for all baselines is limited to **K=4 paths**[1]

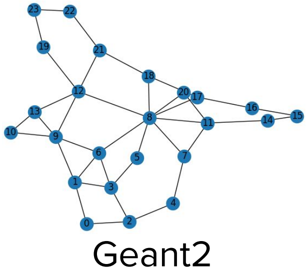
- **SoA DRL[2]:** The DRL agent implements a standard Fully Connected NN and they use an **elaborated representation of the network state** into a matrix
- **Theoretical Fluid: Traffic demands are be split into the K paths proportionally** to the available capacity. This routing policy is aimed at avoiding congestion on links
- **Load Balancing (LB): Selects uniformly one path** among the K candidate shortest paths to allocate the traffic demand

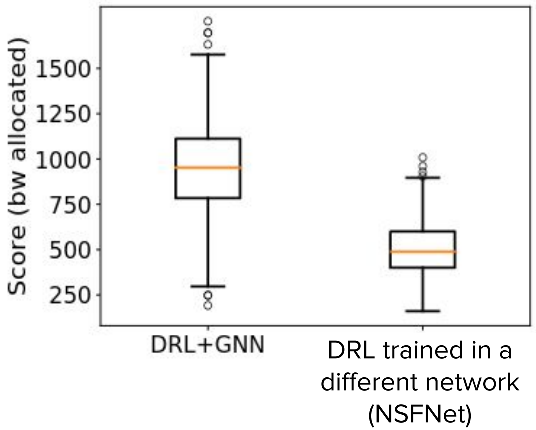
[1]Paul Almasan, et al. "Deep reinforcement learning meets graph neural networks: Exploring a routing optimization use case." Computer Communications, 196, pp. 184-194, 2022.
[2]Suárez-Varela, J., Mestres, A., Yu, J., Kuang, L., Feng, H., et al., (2019). Routing in optical transport networks with deep reinforcement learning. Journal of Optical Communications and Networking, 11(11), 547-558.

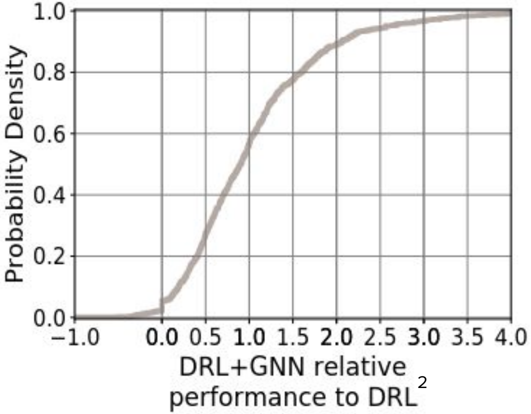
Evaluation I: Generalization

We **trained on the Nsfnet[1] topology** a SoA DRL agent[2] and our DRL+GNN architecture and **evaluated on the Geant2[3] topology**





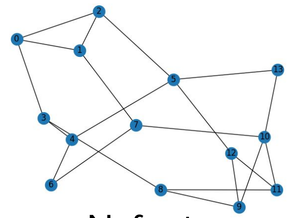




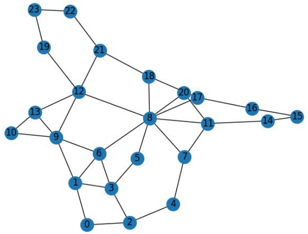
[1]Hei, X., Zhang, J., Bensaou, B., & Cheung, C. C. (2004). Wavelength converter placement in least-load-routing-based optical networks using genetic algorithms. Journal of Optical Networking, 3(5), 363-378.
[2]Suárez-Varela, J., Mestres, A., Yu, J., Kuang, L., Feng, H., et al., (2019). Routing in optical transport networks with deep reinforcement learning. Journal of Optical Communications and Networking, 11(11), 547-558.
[3]Barreto, F., Wille, E. C., & Nacamura Jr, L. (2012). Fast emergency paths schema to overcome transient link failures in ospf routing. arXiv preprint arXiv:1204.2465.

Evaluation I: Generalization

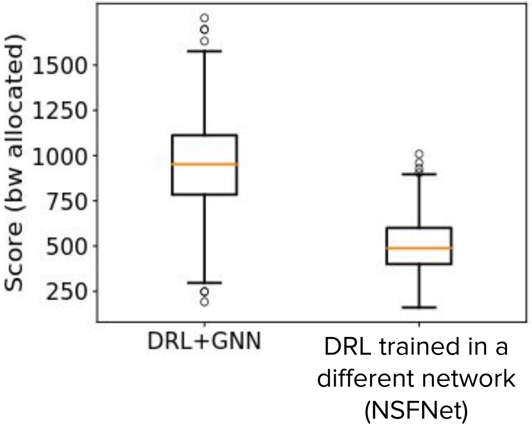
We **trained on the Nsfnet[1] topology** a SoA DRL agent[2] and our DRL+GNN architecture and **evaluated on the Geant2[3] topology**

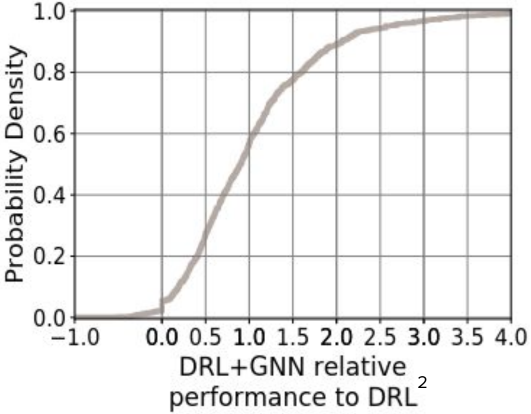


Nsfnet



Geant2





Our DRL+GNN agent is able to **generalize to the unseen Geant2 topology**

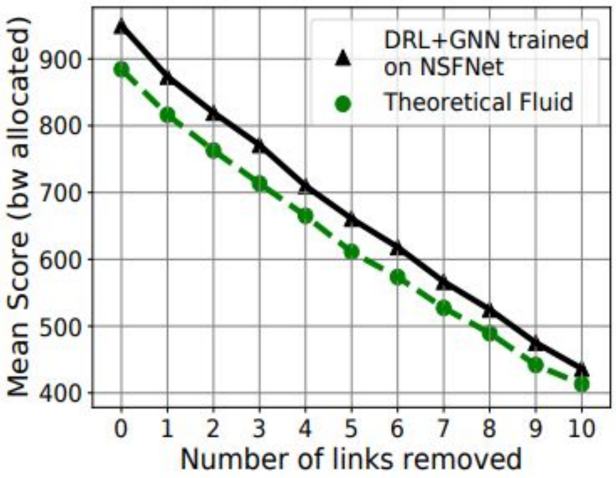
[1]Hei, X., Zhang, J., Bensaou, B., & Cheung, C. C. (2004). Wavelength converter placement in least-load-routing-based optical networks using genetic algorithms. Journal of Optical Networking, 3(5), 363-378.
[2]Suárez-Varela, J., Mestres, A., Yu, J., Kuang, L., Feng, H., et al., (2019). Routing in optical transport networks with deep reinforcement learning. Journal of Optical Communications and Networking, 11(11), 547-558.
[3]Barreto, F., Wille, E. C., & Nacamura Jr, L. (2012). Fast emergency paths schema to overcome transient link failures in ospf routing. arXiv preprint arXiv:1204.2465.

Evaluation II: Link failure

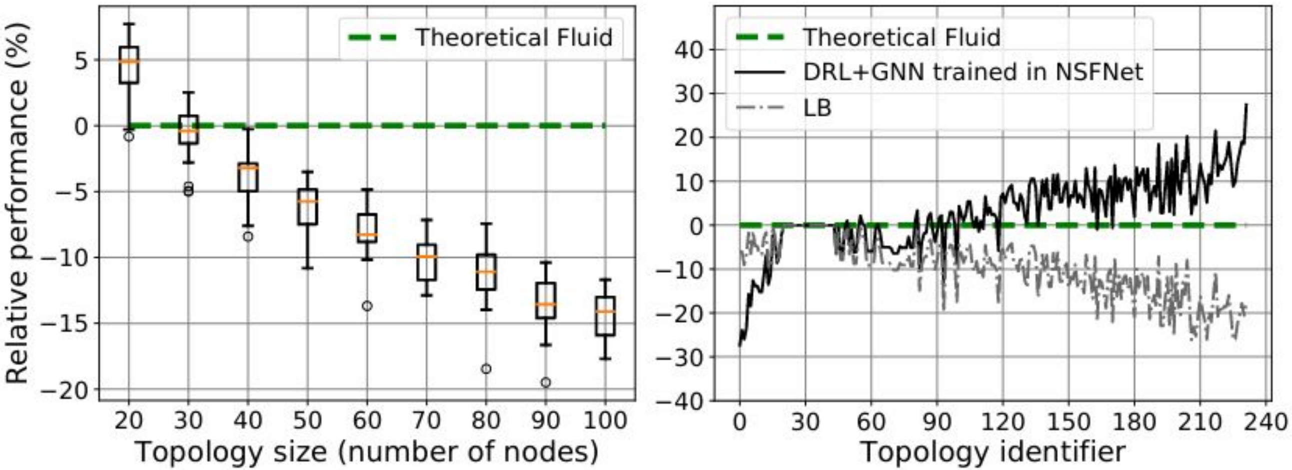
Changes in network connectivity are unpredictable and they have a **significant impact** in protocol convergence and **network performance**

We considered a range of scenarios that can experience **up to 10 link failures**

- Links are randomly removed from the Geant2 topology



Evaluation III

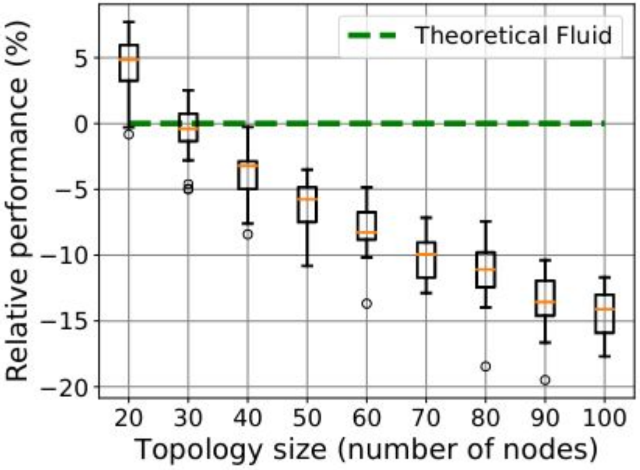


(a) 180 synthetic topologies (20 different per topology size)

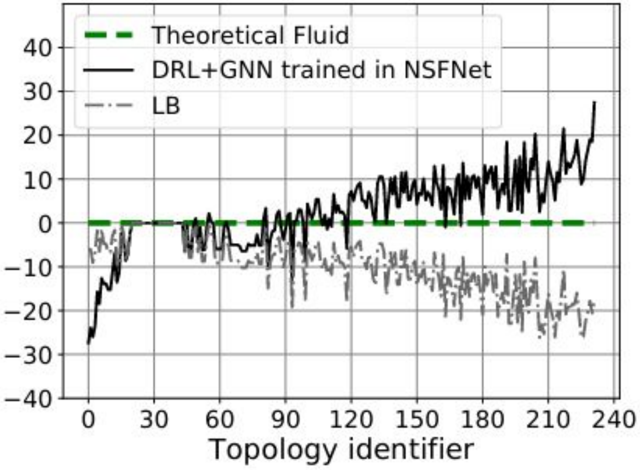
(b) 232 real-world topologies[1]

[1]Paul Almasan, et al. "Deep reinforcement learning meets graph neural networks: Exploring a routing optimization use case." Computer Communications, 196, pp. 184-194, 2022.

Evaluation III



(a) 180 synthetic topologies (20 different per topology size)



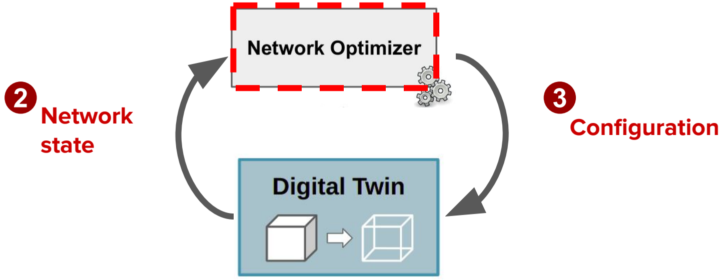
(b) 232 real-world topologies[1]

The DRL+GNN architecture is **robust to operate in real-world topologies** that largely differ from the topologies seen during training

[1]Paul Almasan, et al. "Deep reinforcement learning meets graph neural networks: Exploring a routing optimization use case." Computer Communications, 196, pp. 184-194, 2022.

Conclusions

Method	Execution cost	Performance
Heuristics[1,2]	Low	Low
Mathematical Optimizers[3,4] **(e.g., CP, ILP)**	High	High
SoA Machine Learning[5,6]	High (training)	High
Our DRL+GNN solution	Low (generalization)	High



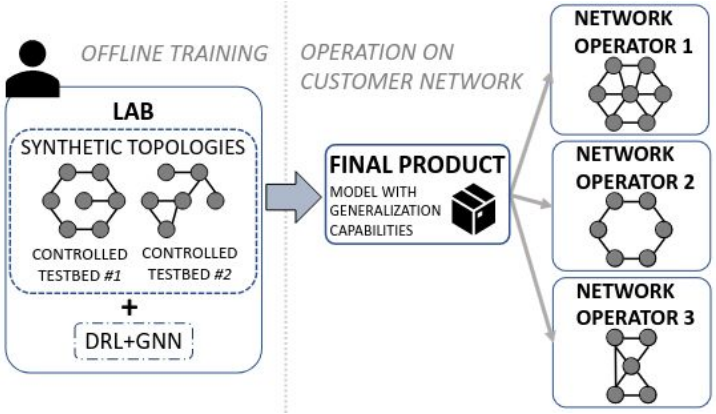
[1]Fortz, B., & Thorup, M. (2000, March). Internet traffic engineering by optimizing OSPF weights. In Proceedings IEEE INFOCOM 2000. (Vol. 2, pp. 519-528).
[2]Wang, N., Ho, K. H., Pavlou, G., & Howarth, M. (2008). An overview of routing optimization for internet traffic engineering. IEEE Communications Surveys & Tutorials, 10(1), 36-56.
[3]Gong, L., et. al. (2013). Efficient resource allocation for all-optical multicasting over spectrum-sliced elastic optical networks. Journal of Optical Communications and Networking, 5(8), 836-847.
[4]Hartert, R., Vissicchio, et. al. (2015). A declarative and expressive approach to control forwarding paths in carrier-grade networks. ACM SIGCOMM computer communication review, 45(4), 15-28.
[5]Chen, X., et. al. (2018). Deep-RMSA: A deep-reinforcement-learning routing, modulation and spectrum assignment agent for elastic optical networks. In Optical Fiber Communication Conference (pp. W4F-2).
[6]Suárez-Varela, et. al. (2019). Routing in optical transport networks with deep reinforcement learning. Journal of Optical Communications and Networking, 11(11), 547-558.

Summary

NDTs enable **the development of more efficient network control and management tools** in modern networks

The DRL+GNN represents a **first step towards ML-based optimizers that generalize to other topologies**

- High optimization performance and fast inference
- Small execution cost (no re-training)



Source code and datasets are publicly available[1]

[1]https://github.com/knowledgedefinednetworking/DRL-GNN

Thank you!

Please reach out if you want to know more about applications of GNNs in mobile networks

@PaulAlmasan Paul Almasan

Telefónica