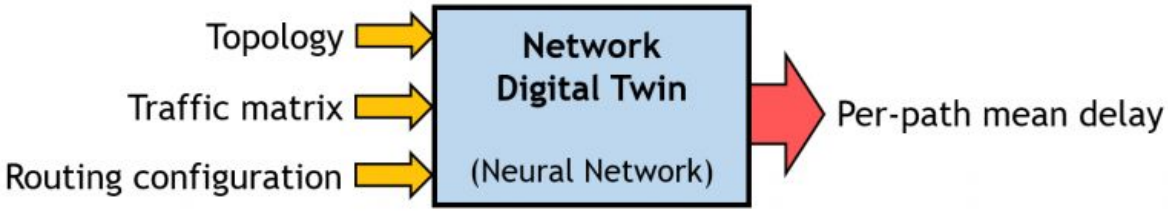
ITU-ML5G-PS-001: Graph Neural Networking Challenge 2021 - Creating a Scalable Network Digital Twin

Miquel Farreras Casamort[1], Paola Soto Arenas[2], Lluís Fàbrega Soler[1], Pere Vilà Talleda[1]
GAIN TEAM

[1]Institute of Informatics and Applications, Universitat de Girona, Girona, Spain
[2]IDLab, University of Antwerp, in collaboration with imec, Antwerpen, Belgium

Index

- Description of our best approach
- Problems and possible improvements
- Conclusions and acknowledgments
- Q&A

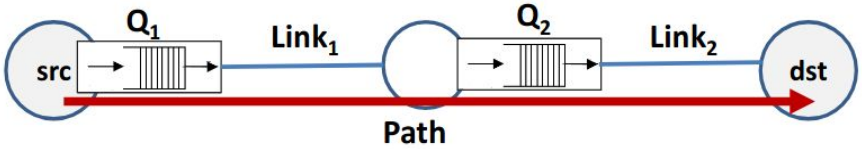


imec Universiteit Antwerpen BCDS Universitat de Girona

Description of our best approach

Solution based on RouteNet Tensorflow

1st step: Apply the queue occupancy approach



$Delay_{L1} = Occupancy_{Q1} * Size_{Q1} / Cap_{L1}$

$Delay_{Path} = \sum_{k=1}^{N\ links} Delay_{L_k}$

occupancy = avgPortOccupancy/queueSize

occupancy$_{Q1}$ in 0-1 values range

avgPortOccupancy in number of packets

queueSize ($Size_{Q1}$) in number of packets

capacity (Cap_{L1}) = (bits/time unit)/avgPktSize → in packets/time unit

Result on evaluation platform: 30.40 % MAPE

imec Universiteit Antwerpen BCDS Universitat de Girona

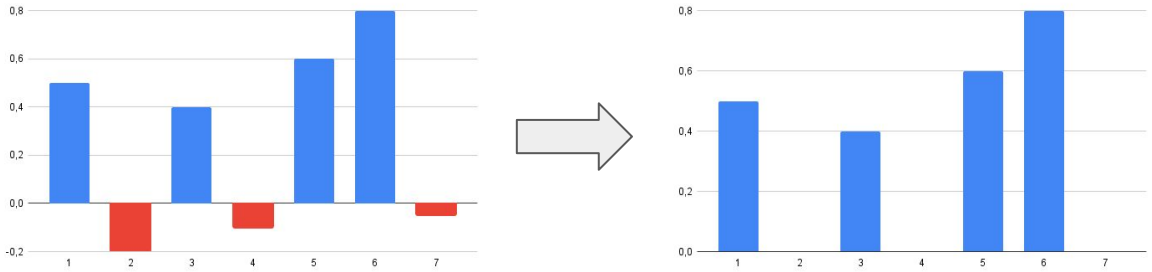
Description of our best approach

2nd step: Limit occupancy results

Few results predicted a negative occupancy

Solution: negative predicted occupancy rounded to zero

Result on evaluation platform: 30.18 % MAPE



Description of our best approach

3rd step: Path length feature

Setting 1: Longer paths
Setting 2: Higher capacity links
Setting 3: Mix of both features

Worse results for artificial longer paths (Setting 1)

Solution: new path length attribute for each path

TABLE I
MAPE ORIGINAL ROUTENET VS OCCUPANCY IMPLEMENTATION

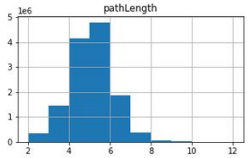
	Full validation	**Setting 1**	**Setting 2**	**Setting 3**
Original model	187,28	79,14	253,07	247,21
Our occupancy approach	28,82	63,58	21,02	21,54
		50 nodes Setting 1	**50 nodes Setting 2**	**50 nodes Setting 3**
Original model		50,15	68,48	44,66
Our occupancy approach		18,86	21,04	17,18
		300 nodes Setting 1	**300 nodes Setting 2**	**300 nodes Setting 3**
Original model		92,97	345,13	368,01
Our occupancy approach		96,41	31,56	24,26

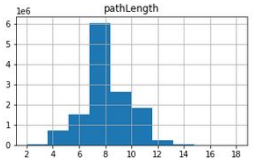
Description of our best approach

4th step: Normalize predictor variables

Data distributions in different value ranges in train and validation datasets:

- bandwidth (link property)
- AvgBw (path property)
- pathLength





Solution: min-max normalization as a preprocessing in the transformation function (as suggested on the challenge GitHub)

Result on evaluation platform (3 and 4 steps together): 28.73 % MAPE

imec Universiteit Antwerpen BCDS Universitat de Girona

Description of our best approach

Parameters

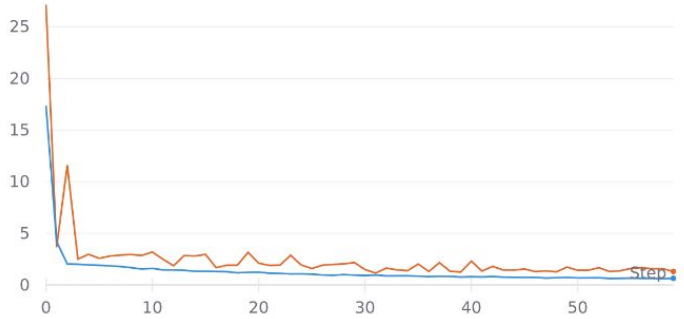
No modifications to loss functions, hyper-parameters, or Routenet architecture

Number of epochs reduced from 100 to 60 for faster testing

Features used

Path features	
Traffic	Average bandwidth through the path
PathLength	Number of nodes in path
Link features	
Capacity	Link bandwidth (bits/time unit)

Fig. 1. MAPE on train (blue) and validation (orange) datasets



Other things we tried

New feature, offeredTraffic of a link
Sum of the average traffic of all the paths passing through a link divided by the capacity of that link, resulting in a scalar variable assigned to the link state.

Include pktsGen feature
Include the number of packets generated per second, max-min normalized, as a path feature. Higher correlation to delay was later seen in the data analysis.

Average packet size (APS) for each link for occupancy post processing
Average traffic in bits per second for each path, divided by the packets generated.
Calculate all the APS for each link forming a path. Do for all paths.
For each link, obtain the APS of all the paths that go through each link.

Early stopping

Low occupancy MAPE, high delay MAPE?
Low validation 1.317% MAPE for the occupancy prediction, 47% MAPE for the delay prediction

Conclusions and acknowledgements

First experience in GNN for some participants → satisfied with the result

We plan to work further to improve the results now that the test dataset with labels is released.

This competition has helped us to gain experience in tuning and developing GNNs, a field that has a steep learning curve.

The GAIN team wants to thank the contributions and insights of Dr. Miguel Camelo and the University of Antwerp having given access to us to their GPU platform.

We are looking forward to the next challenge

Any questions?