A Graph Convolutional Neural Network approach for throughput prediction in next-generation WLANs

ITU-ML5G-PS-013: Improving the capacity of IEEE 802.11 WLANs through Machine Learning

Team ATARI: Paola Soto[1,2], David Góez[2], Natalia Gaviria[2], Miguel Camelo[1]

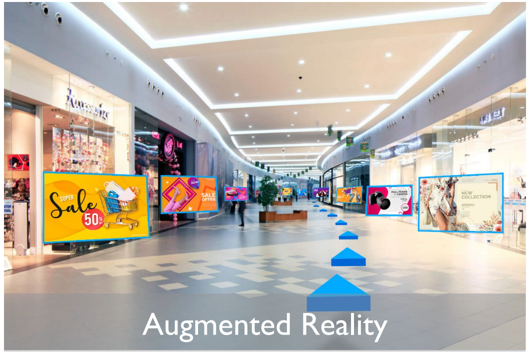
1 University of Antwerp - imec, IDLab, Department of Mathematics and Computer Science, Antwerp, Belgium

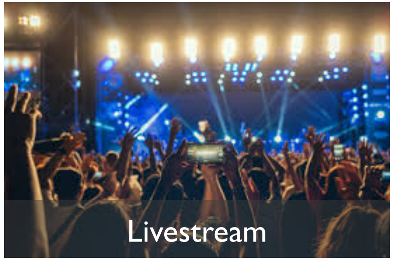
2 Universidad de Antioquia, Department of Telecommunications Engineering, Medellín, Colombia

Email: paola.soto-arenas@uantwerpen.be

Context

Million connected devices with higher demands in bandwidth





- New services are becoming a reality
  - Higher Bandwidth
  - High availability for high number of users
- Wi-Fi will remain an important technology to support such services
  - Wi-Fi hotspots will fourfold increase by 2023[1]

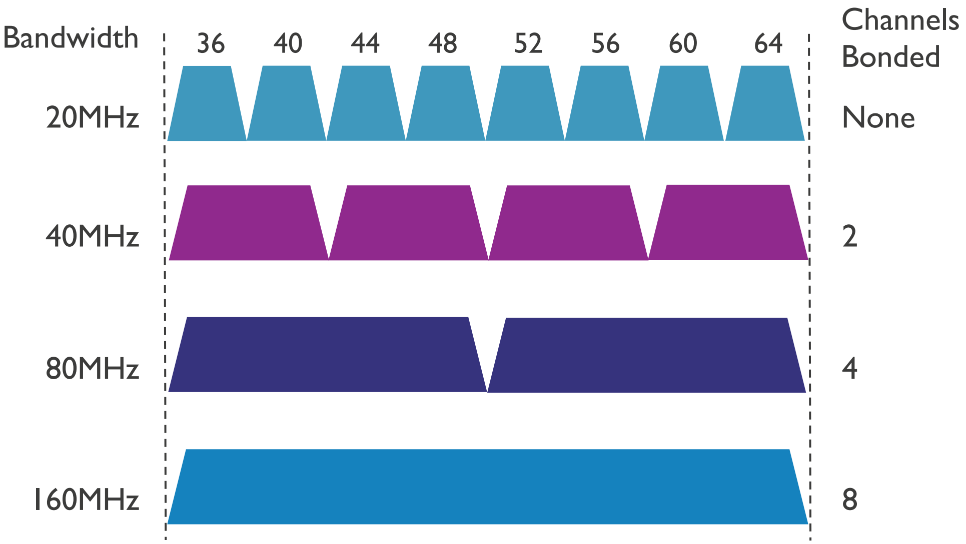
imec IDLab INTERNET & DATA LAB UNIVERSITEIT GENT Universiteit Antwerpen UNIVERSIDAD DE ANTIOQUIA

1. Cisco annual internet report (2018–2023) white paper. Available on: http://www.cisco.com/go/vni

Channel Bonding

Provide a higher channel capacity

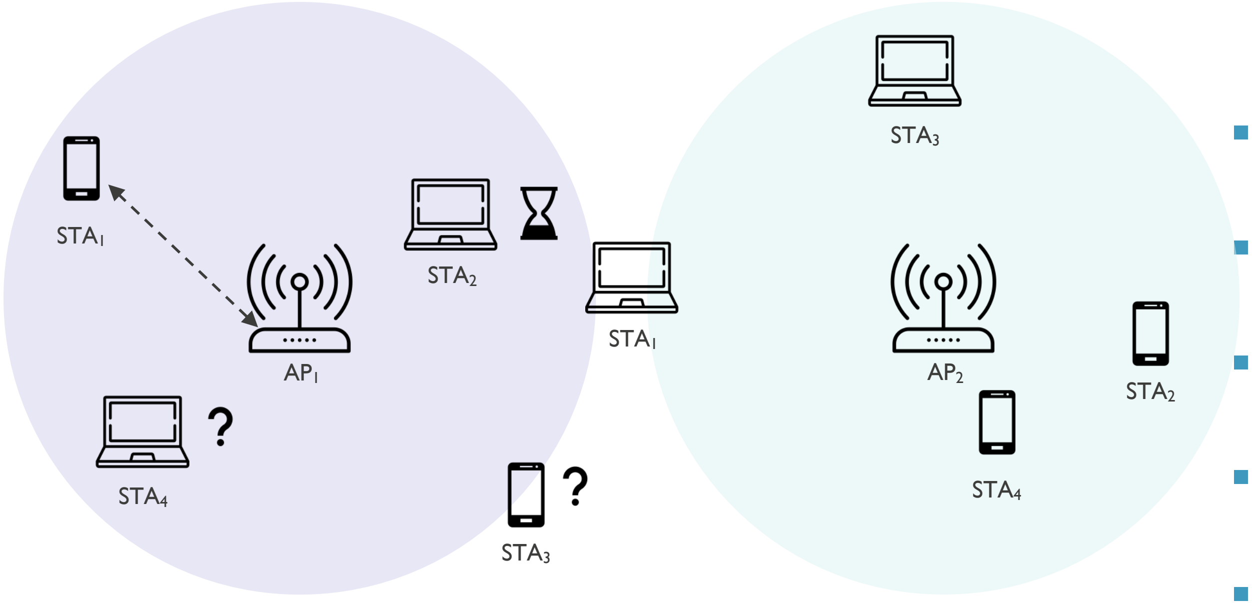
- Introduced in 802.11n[2] (2009).
- Two adjacent channels are bonded.
  - Recent versions (802.11ac/ax) allow more channels to be bonded.



2. IEEE, "Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment 5 : Enhancements for Higher Throughput", Standard, 2009.

Channel Bonding Problems

OBSS interactions



- Stochastics interactions based on the CSMA algorithm.
- Transmission over the full/sub set of channels.
- High data rate = high packet error.
- Inter STA – Inter WLAN interference.
- Hidden or exposed nodes.
- Starvation.

Previous Solutions

- Analytical models[3]
- Markov Chains[4]
- Simulations

- Assumptions that do not hold.
- Complexity increases with the number of devices.

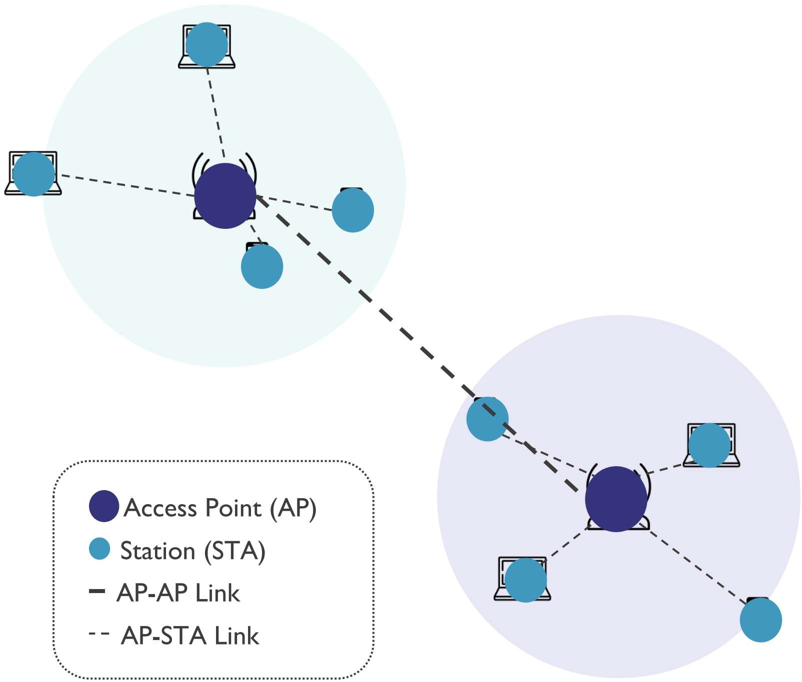
- Machine Learning / Artificial Intelligence[5]
  - Reinforcement Learning to select CB Policy[6]

3. L. Deek et al, “Intelligent channel bonding in 802.11n wlans,” IEEE Transactions on Mobile Computing, vol. 13, no. 6, pp. 1242–1255, 2013
4. S. Barrachina-Muñoz, F. Wilhelmi, and B. Bellalta, “Dynamic channel bonding in spatially distributed high-density wlans,” IEEE Transactions on Mobile Computing, vol. 19, no. 4, pp. 821–835, 2019.
5. ITU-T Y.3170-series - Machine learning in future networks including IMT-2020: use cases,” ITU-T, Specification, 2019. [Online]. Available: https://www.itu.int/rec/T-REC-Y.Sup55-201910-I/en4
6. Y. Luo and K.-W. Chin, “Learning to bond in dense wlans with random traffic demands,” IEEE Transactions on Vehicular Technology, vol. 69, no. 10, pp. 11 868–11 879, 2020.

Proposed Solution: Graph Neural Networks – GNN[7]

DL do not exploit graph-like structured data

- A graph is 2-tuple G=(V, E) where V are the nodes and E are the links.
- Each node has its own set of features.
  - Node type
  - SINR
  - Channels used
  - Airtime
- Each link has its own set of features
  - Distance
  - RSSI
  - Interference

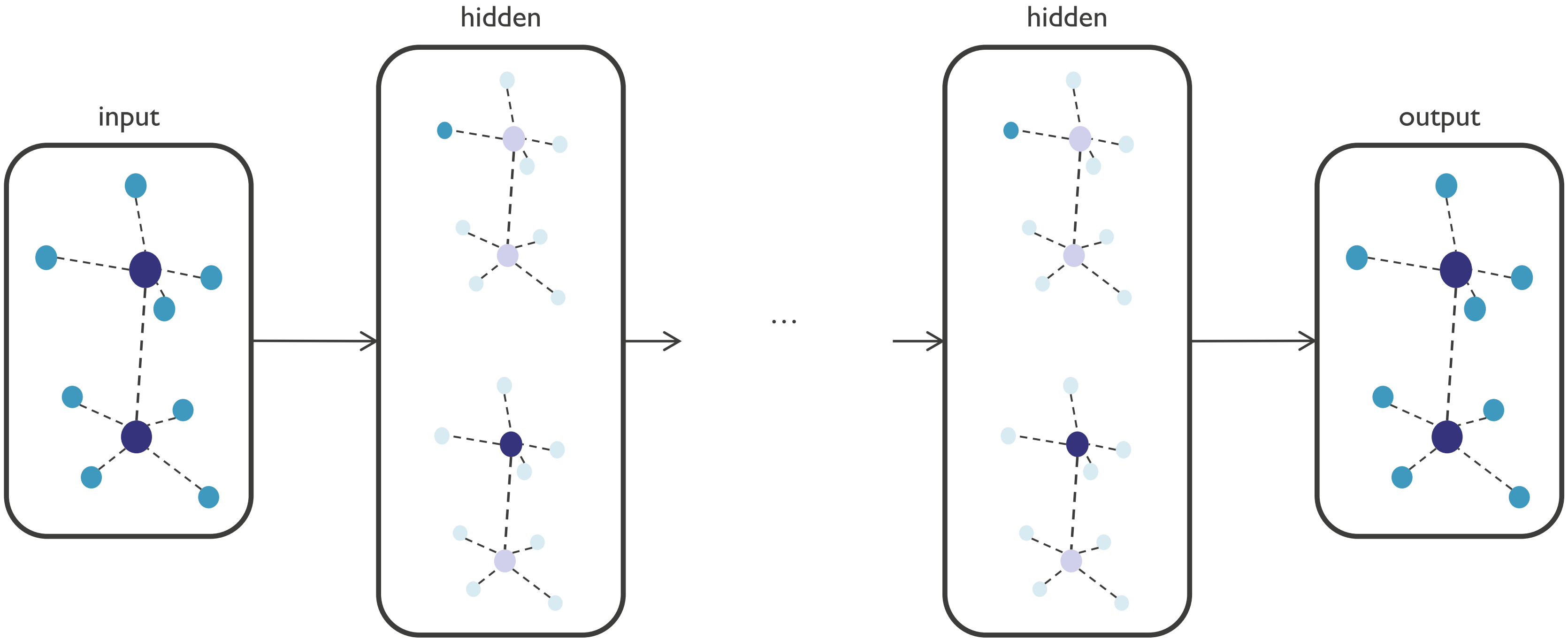


imec IDLab INTERNET & DATA LAB UNIVERSITEIT GENT Universiteit Antwerpen UNIVERSIDAD DE ANTIOQUIA

7. F. Scarselli, et al, "The graph neural network model," IEEE Transactions on Neural Networks, vol. 20, no. 1, pp. 61–80, 2008.

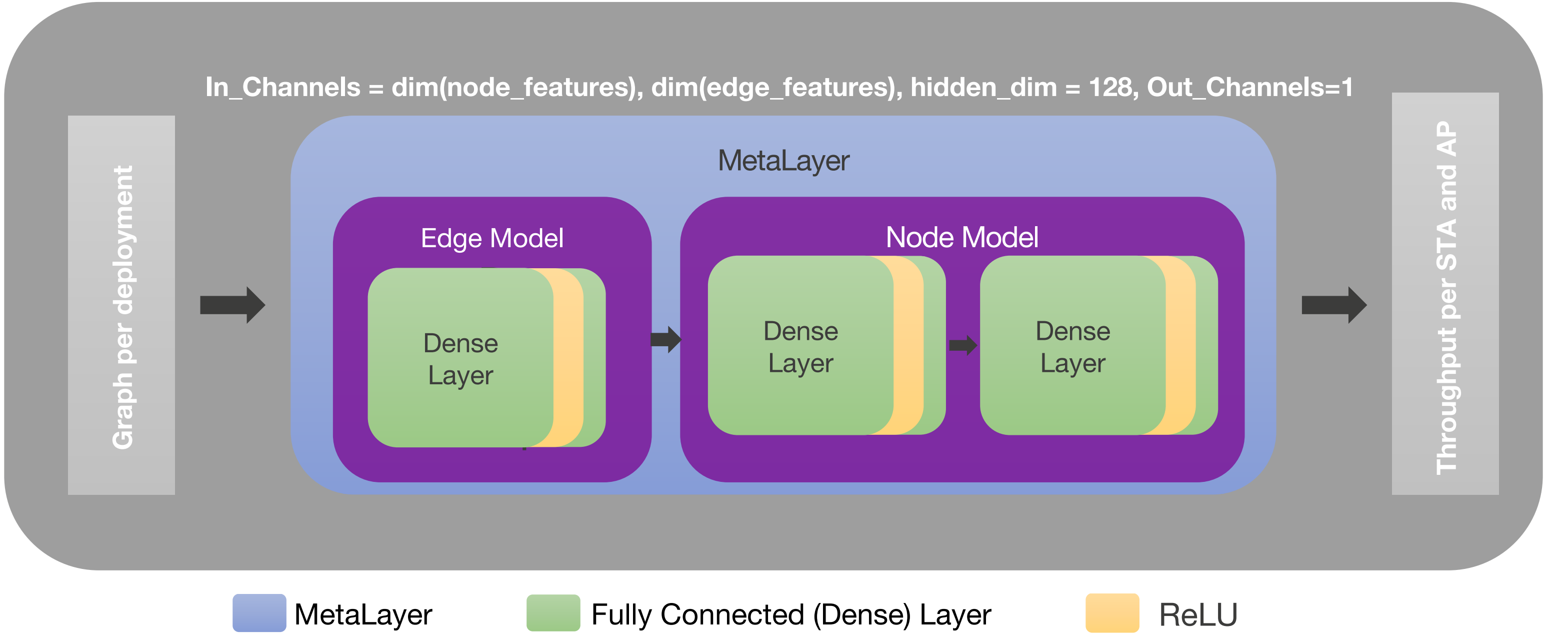
Graph processing

Follows a traditional multilayer approach of (Deep) Neural Networks



Model - MetaNet

Each layer is designed to find patterns in the graph-like structured data[8]



imec IDLab INTERNET & DATA LAB UNIVERSITEIT GENT Universiteit Antwerpen UNIVERSIDAD DE ANTIOQUIA

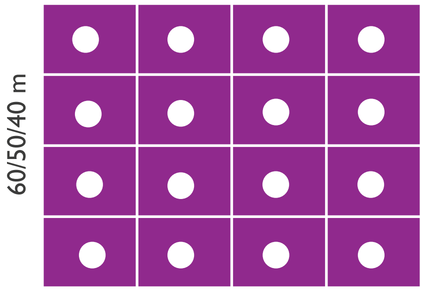
8. P. W. Battaglia et al., "Relational inductive biases, deep learning, and graph networks,", arXiv preprint arXiv:1806.01261, 2018

Training Dataset

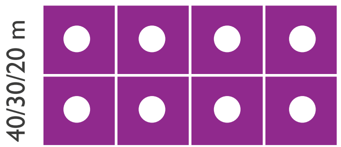
Characteristics

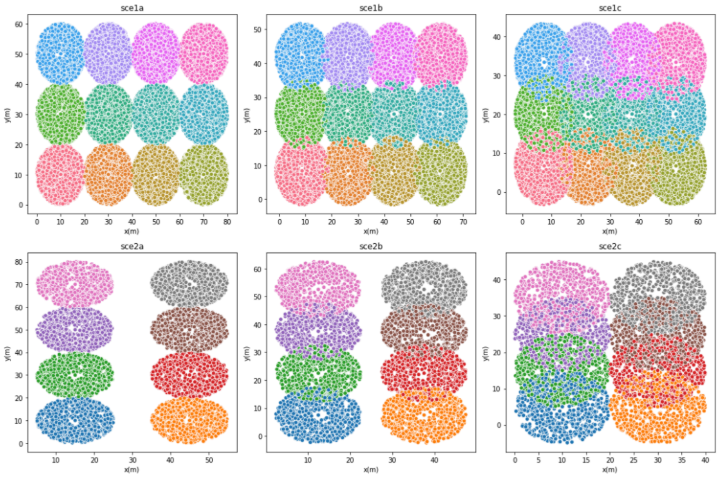
- Two scenarios with modifications on user density per m^2
- 3 map sizes per scenario.
- 100 deployments per scenario.
- Each scenario introduce more interference than the other.
- Different channel configurations.
- 80% Training – 480 Graphs.
- 20% Validation – 120 Graphs.

Scenario 1: 12APs 10-20 STAs



Scenario 2: 8APs 5-10 STAs

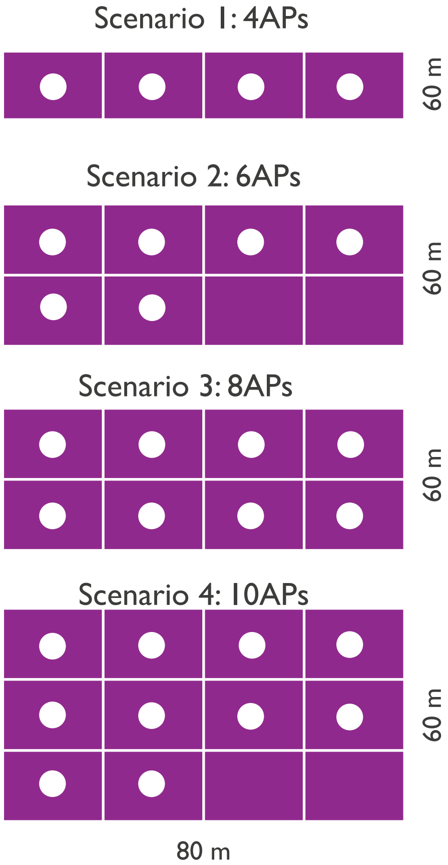




Test Dataset

Characteristics

- Four testing scenarios.
- Different spatial distribution than training.
- 50 random deployments per scenario.

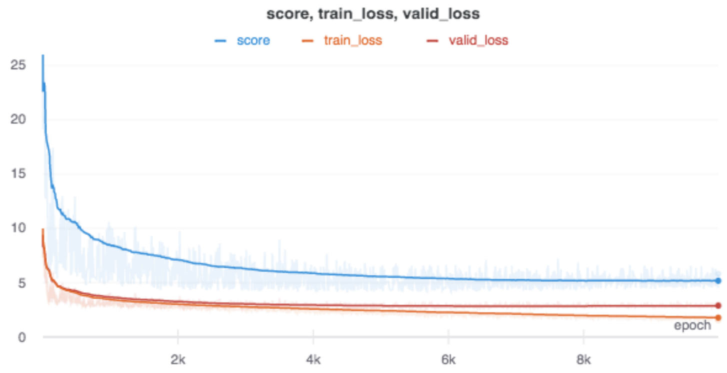


Results

Training

- Training
  - Masked Loss
    - Training loss is designed to learn the STAs' throughput.
    - Throughput AP is computed from predicted STAs throughput.
  - Adam optimizer
  - 10000 epochs
  - 3h of training time.

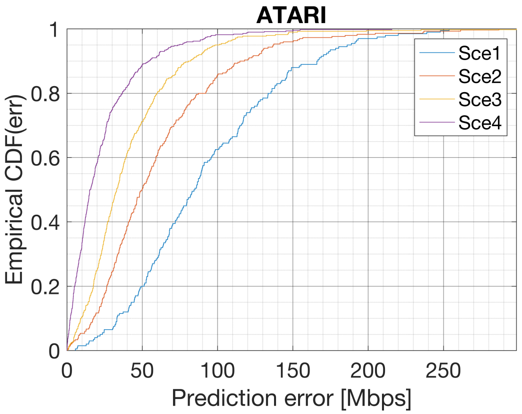
Metric	Loss Value
Train loss (RMSE)	2.057
Validation loss (RMSE)	2.652
Masked Loss (score)	4.197



Results

Testing

- Model is overfitting.
  - Novel architecture for this type of problem.
  - The more similar the test scenario to train scenario, the better the prediction.
  - The amount of data needed to train is big.
- Techniques to avoid overfitting
  - Perform data augmentation.
  - Add regularizers, e.g., dropout.



Scenario	RMSE*
1	26.97
2	18.94
3	12.76
4	8.73

*Over all devices

imec IDLab UNIVERSITEIT GENT Universiteit Antwerpen UNIVERSIDAD DE ANTIOQUIA

Improvements

- Change MetaLayers by Graph Convolutional Layers.
- Implemented a Feed Forward model and a Convolutional model as baseline.
- Next Publication.

Scenario	GCN	FNN	CNN
1	9.384	16.250	16.403
2	8.247	15.562	15.456
3	6.822	11.774	11.485
4	6.481	12.025	11.565

*RMSE over all devices

imec

embracing a better life