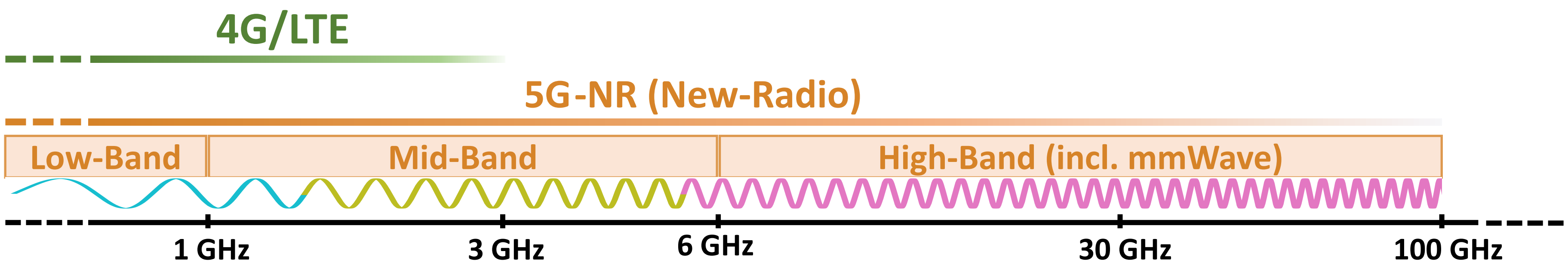
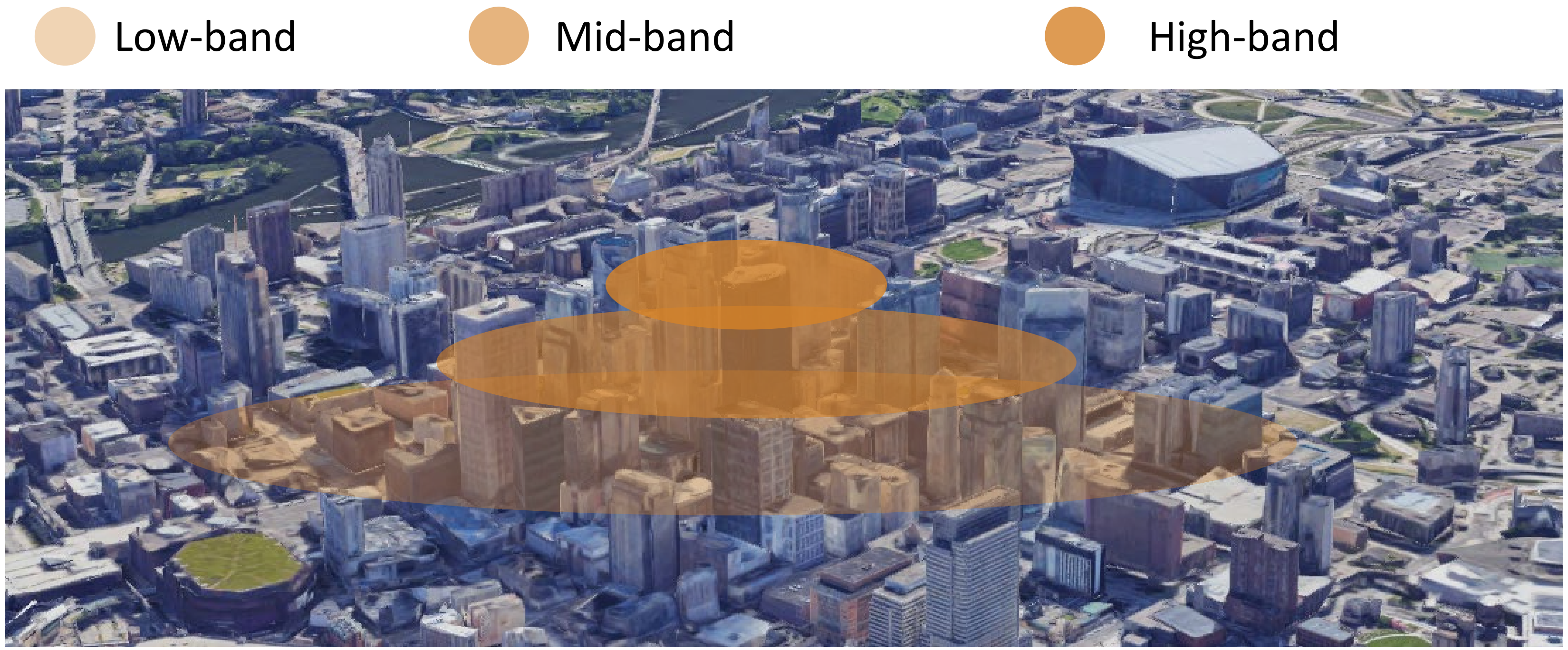
5GNN: Extrapolating 5G Measurements through GNNs

Wei Ye, Xinyue Hu, Tian Liu, Ruoyu Sun,
Yanhua Li, **Zhi-Li Zhang**

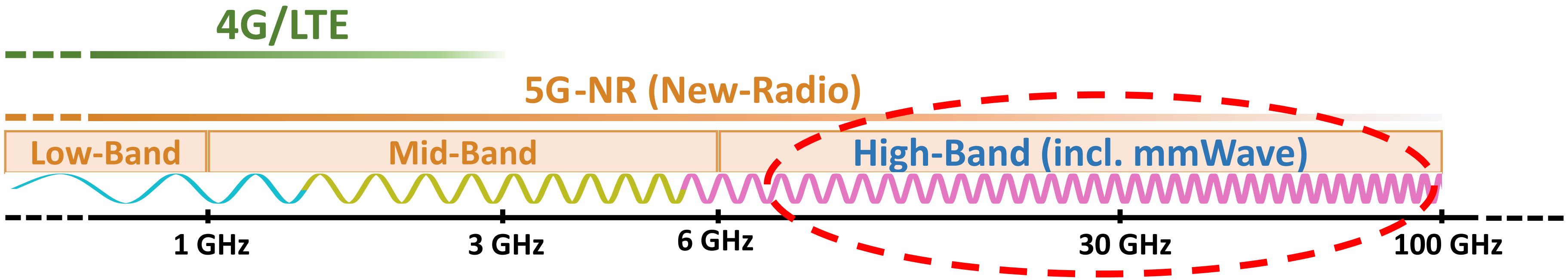
research supported in part by

5G Band Is Diverse





With Ultra High BW, 5G mmWave Can Enrich Human Experience

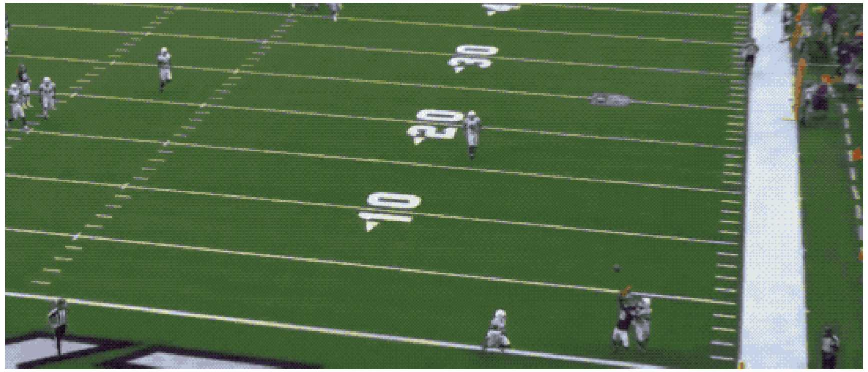


With theoretical throughput up to 20 Gbps!

Key to support apps requiring ultra-high bandwidth

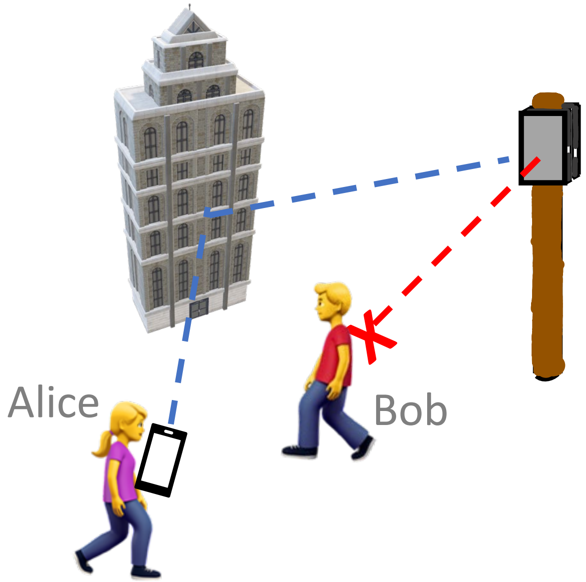
AR/VR Services

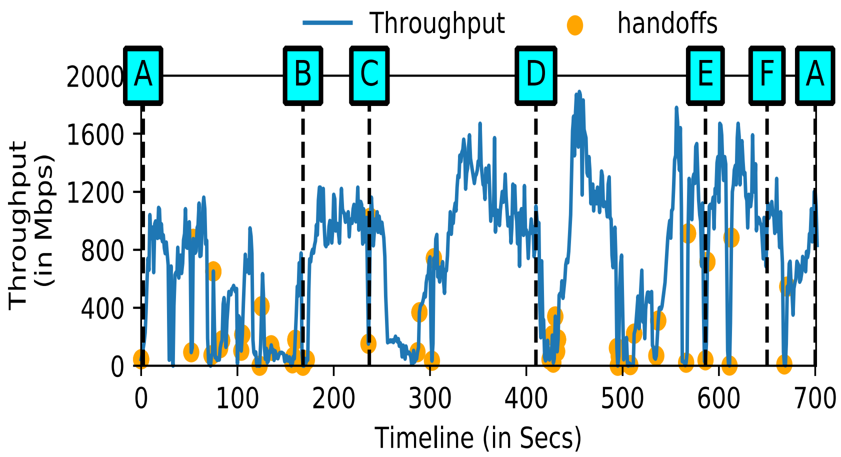
Source: Intel True View

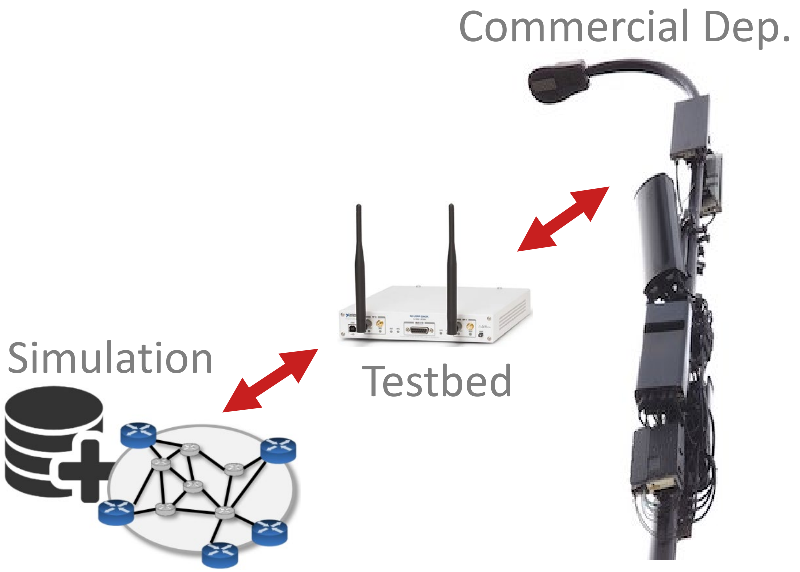


Volumetric Content Delivery

But mmWave 5G Also Poses New Challenges







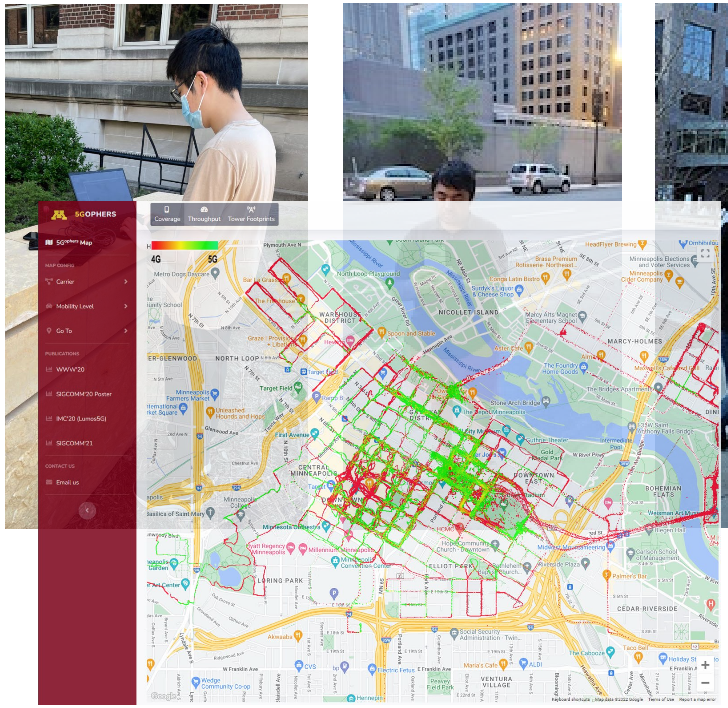
Multi-path signals due to **reflection** and **blockage**

Wild and frequent **fluctuations** in throughput

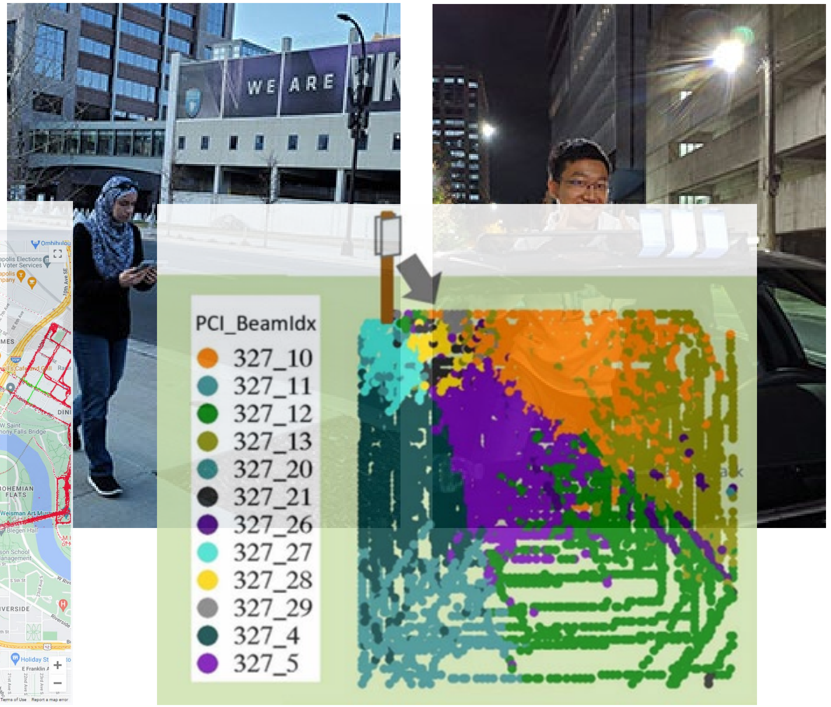
Gaps among **digital world**, **local testbed**, and **commercial deployment**

- **Real-world measurements are crucial in understanding 5G network performance!**

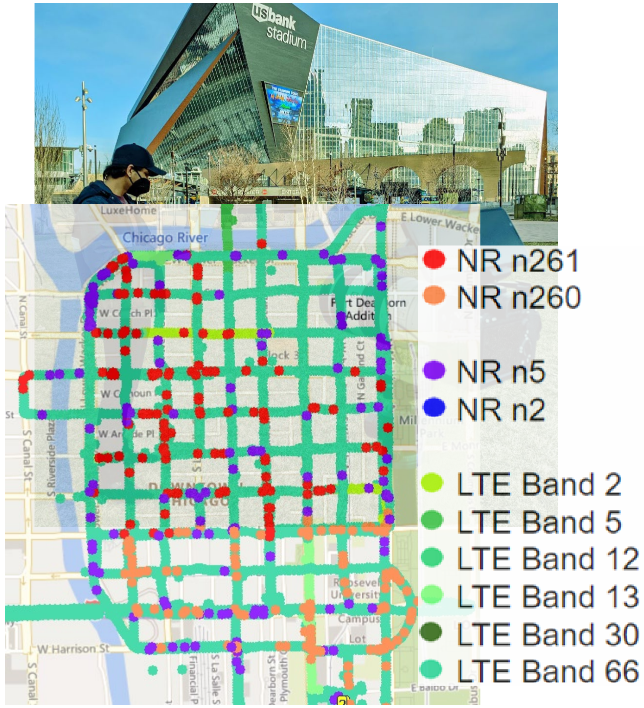
Measurements are Critical to Understanding Commercial 5G Performance



5Gopher: Mapping throughput



5G Beam coverage



5G Band study

- In-depth **"in-the-field" measurements** are critical to understanding the commercial 5G network performance.

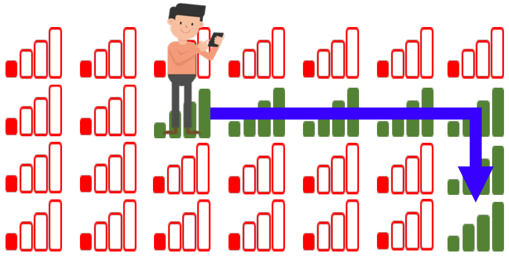
Measurements of Commercial 5G Networks: Time-Consuming and Costly

However,

(1) Measurement process is **laborious** and **costly**

(2) Data can only be collected in **limited areas** of (potentially large) space

(3) The limited coverage data will **introduce biases** e.g., to ML models



Can we utilize relatively few measurement data points to "construct" a more complete 5G coverage and performance map?

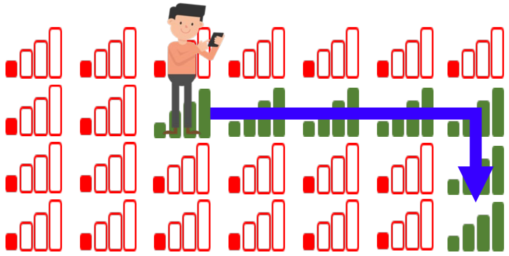
Measurements of Commercial 5G Networks: Time-Consuming and Costly

However,

(1) Measurement process is **laborious** and **costly**

(2) Data can only be collected in **limited areas** of (potentially large) space

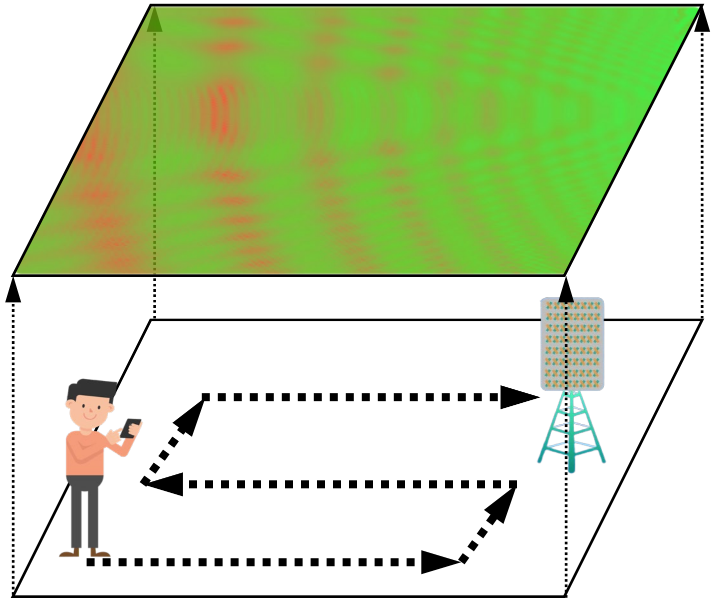
(3) The limited coverage data will **introduce biases** e.g., to ML models



5G Measurement Extrapolation Problem

5G Measurement Extrapolation Problem and GNNs

Given a collection of 5G "signal" measurements collected in the **limited number of locations** in a target 5G coverage area:

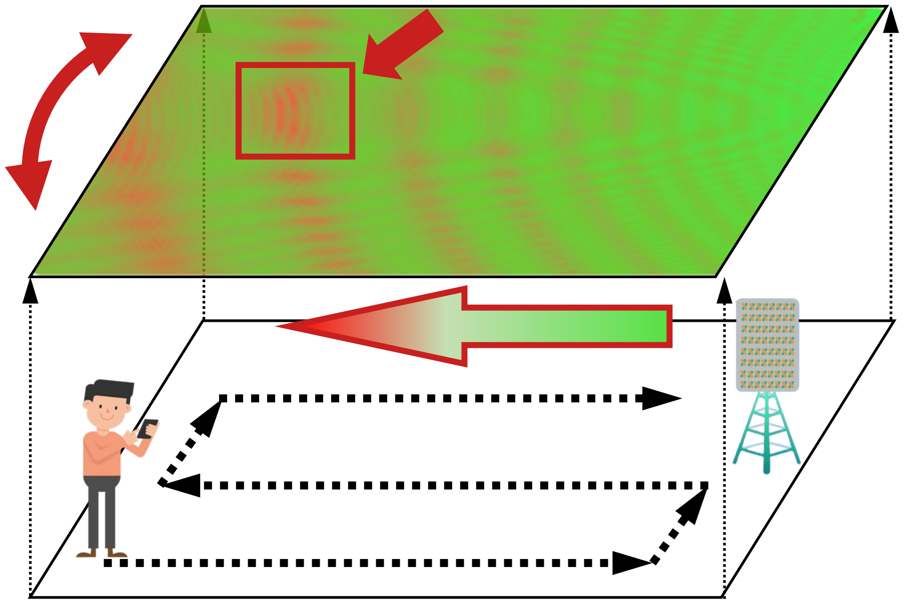


Can we infer or **extrapolate 5G "signals" at other locations** within the area that we do not have samples?

GNN is good at capturing spatial correlations!

5G Measurement Extrapolation Problem and GNNs … *But!*

The state-of-art GNNs cannot fully capture ***physical characteristics*** *of 5G radio frequencies!*



Local Properties:

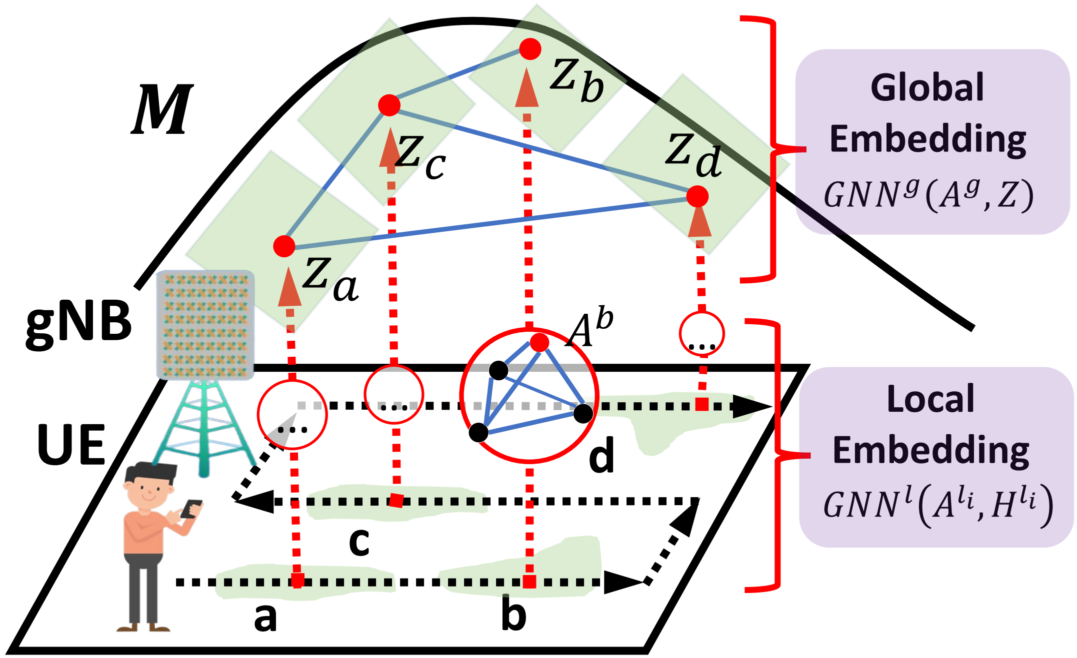
signal waveform variations, noise, and interference

Global Properties:

Radio signal propagation, attenuation, fading, and path losses

Our Solution → ***5GNN***

Proposed Method: 5GNN



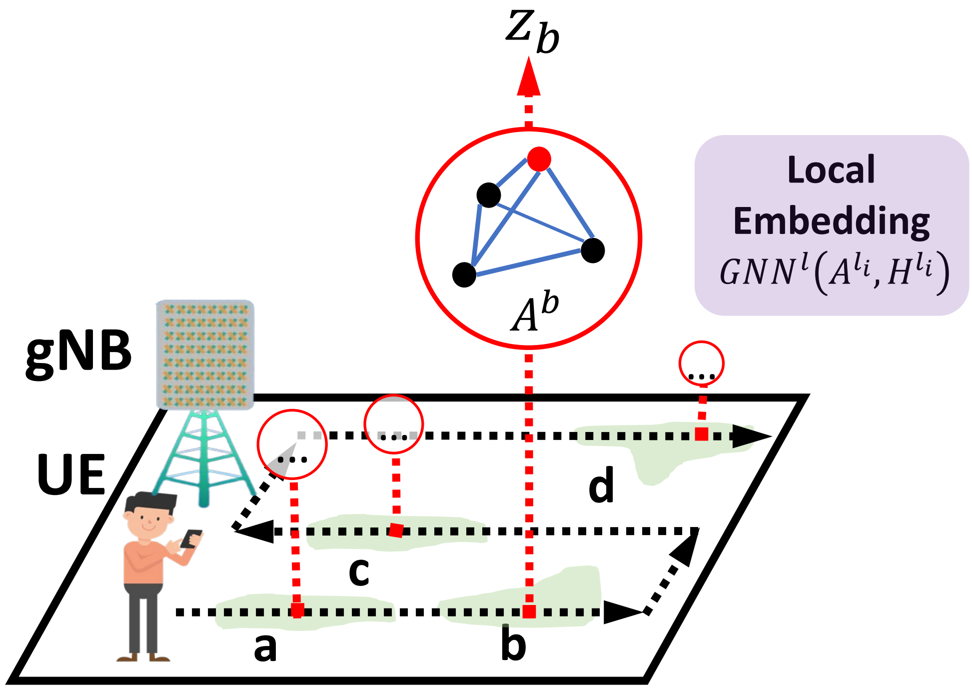
Radio Physical Characteristics

Radio signal propagation, attenuation, fading, and path losses

Local signal waveform variations, noise, and interference

5GNN is a tower information-free, physical-inspired, and graph-based learning approach.

Proposed Method: 5GNN



We aim to **learn a local ("smooth") embedding** of data points lying within a local neighborhood.

Radio signal's local characteristics:

Local signal waveform variations, noise, and interference

Stage 1: Local embedding

$$z_i = F(GNN^l(A^{l_i}, H^{l_i}))$$

$F(.)$ flattens the matrix into vector

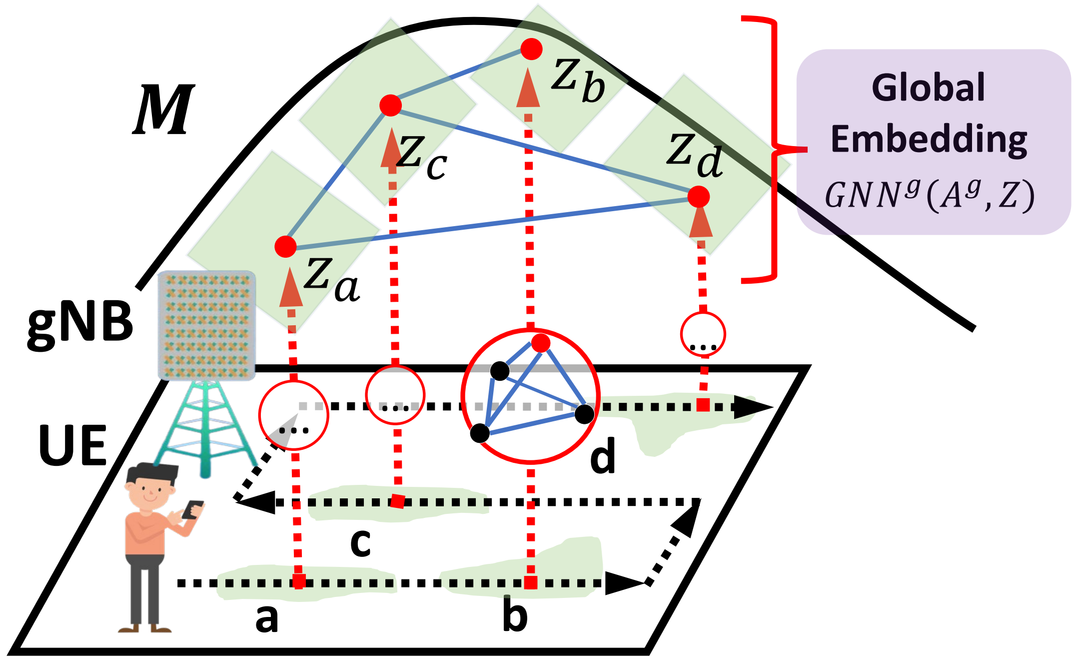
GNN^l is applied on the local graph and can be **any off-the-shelf baseline GNN model.**

For each data point i, sample the k-nearest neighboring $(j, k \in Nbr(i))$ into a local graph A^{l_i}.

$$A^{l_i}_{jk} = \exp(-\frac{d(c_j, c_k)}{2\sigma^2})$$

$$H^{l_i} = \begin{bmatrix} c_i & 0 & 0 \\ c_j & X_j & y_j \\ \dots & \dots & \dots \end{bmatrix} \begin{array}{l} \leftarrow \text{center node} \\ \leftarrow \text{neighbor node} \\ \\ \end{array}$$

Proposed Method: 5GNN



We aim to **approximate the (global) 5G signal manifold** by patching together the local charts in an appropriate manner.

Radio signal's global characteristics:

Radio signal propagation, attenuation, fading, and path losses.

Stage 2: Global embedding

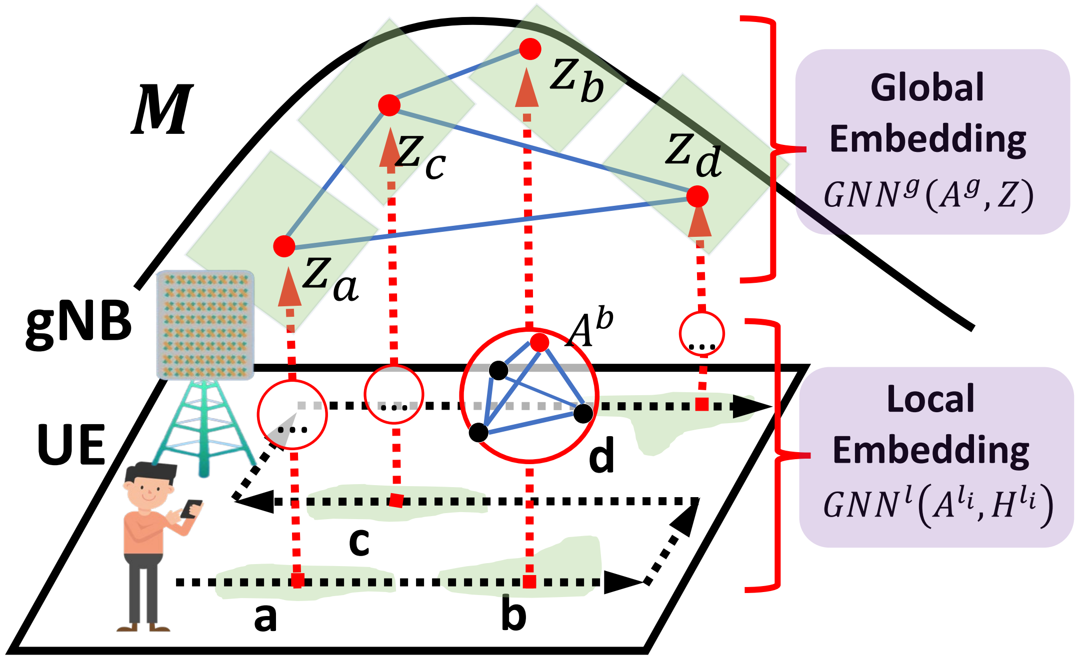
$$\hat{Z} = GNN^g(A^g, Z)$$

GNN^g is applied on the global graph and can be **any off-the-shelf baseline GNN model**.

We build a global kNN graph A^g over all the local charts.

$Z = [z_0, z_1, \dots, z_N]^T$ is the feature matrix

Proposed Method: 5GNN



Stage 3: Joint Training

$$\arg\min_{\theta} \; Loss(\boldsymbol{y}_{pred}, \boldsymbol{y}_{true})$$

$$\boldsymbol{y}_{pred} = MLP([\hat{Z}||Z])$$

Stage 2: Global embedding

Try to approximate the (global) 5G signal manifold.

$$\hat{Z} = GNN^l(A^g, Z)$$

$$Z^g = kNN(\{\boldsymbol{p} \text{ in } \boldsymbol{Batch}\}) \qquad Z = [z_0, z_1, \ldots, z_N]^T$$

Stage 1: Local embedding

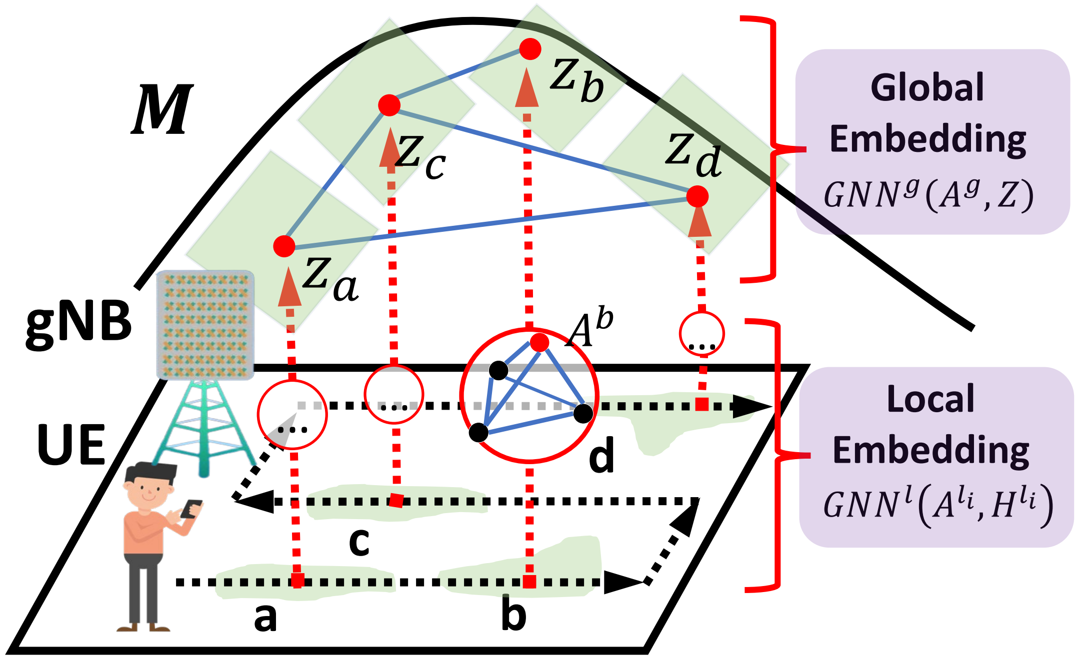
Try to learn a local ("smooth") embedding.

$$z_i = F(GNN^l(A^{l_i}, H^{l_i}))$$

$$A_{jk}^{l_i} = \exp\left(-\frac{d(c_j, c_k)}{2\sigma^2}\right) \qquad H^{l_i} = \begin{bmatrix} c_i & 0 & 0 \\ c_j & X_j & y_j \\ \ldots & \ldots & \ldots \end{bmatrix}$$

Back propagation

Proposed Method: 5GNN



Any existing GNN models can be used for local and global embeddings

Stage 3: Joint Training

$$\arg\min_{\theta} \ Loss(\boldsymbol{y}_{pred}, \boldsymbol{y}_{true})$$

$$\boldsymbol{y}_{pred} = MLP([\hat{Z}||Z])$$

Stage 2: Global embedding

Try to approximate the (global) 5G signal manifold.

$$\hat{Z} = GNN^l(A^g, Z)$$

$$Z^g = kNN(\{\boldsymbol{p} \text{ in } \boldsymbol{Batch}\}) \qquad Z = [z_0, z_1, \dots, z_N]^T$$

Stage 1: Local embedding

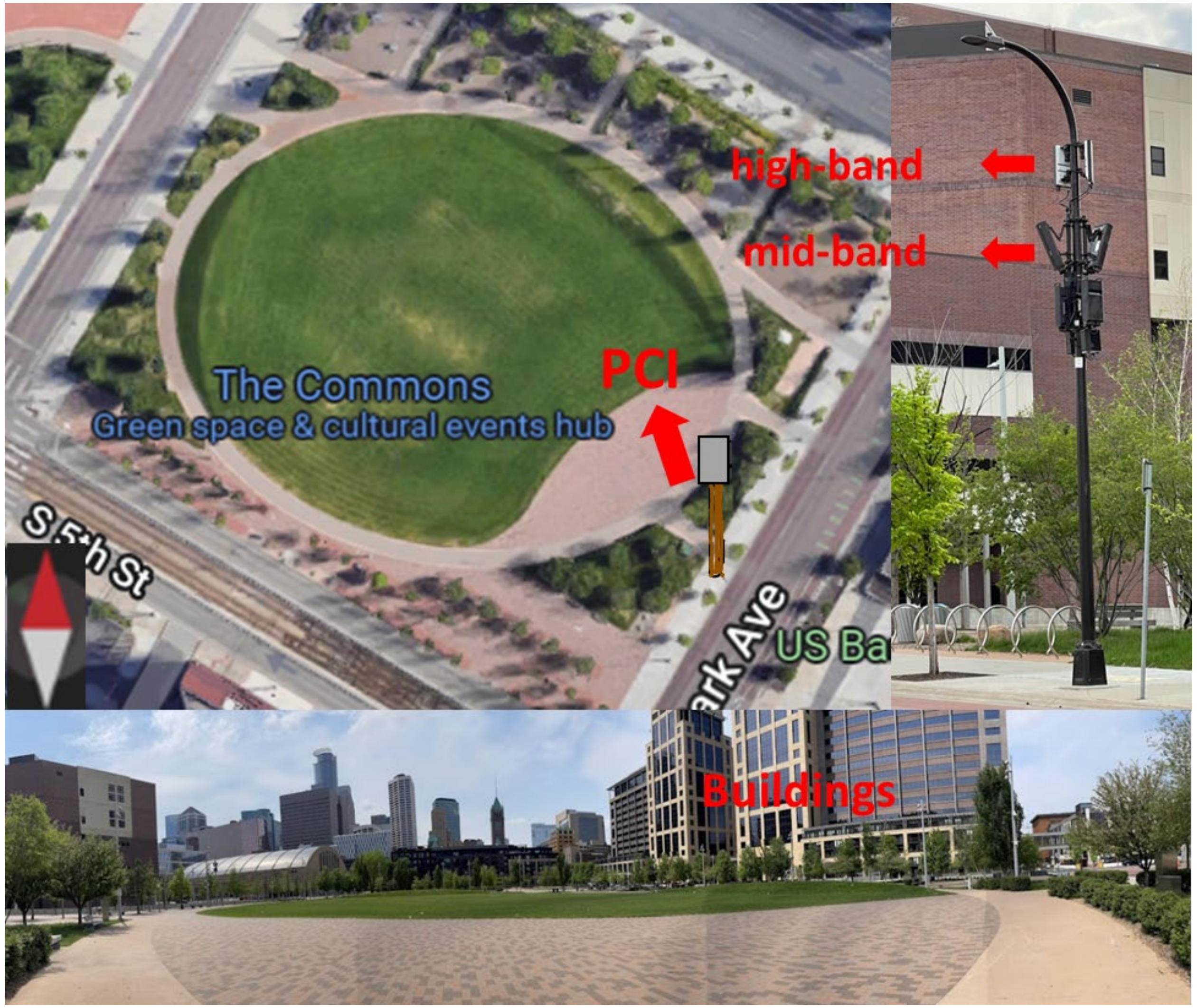
Try to learn a local ("smooth") embedding.

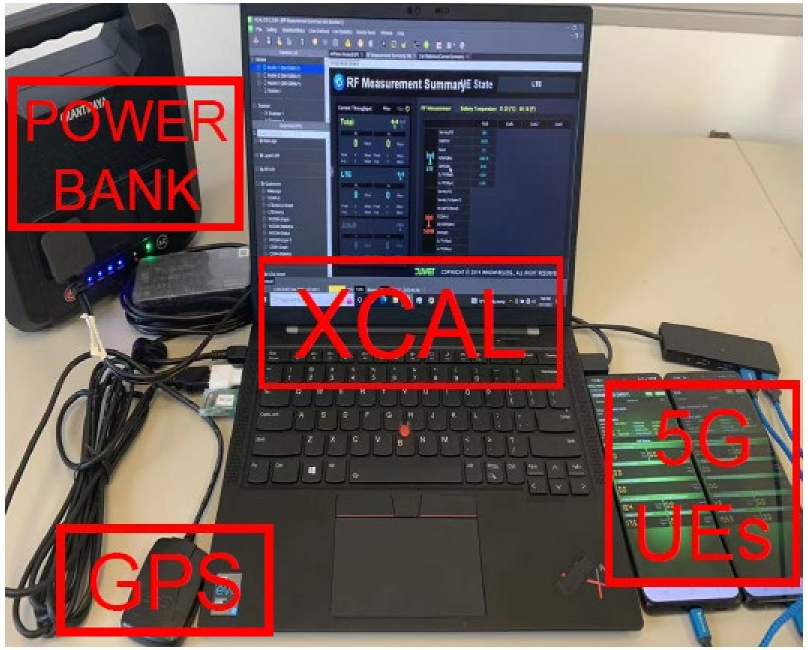
$$z_i = F(GNN^l(A^{l_i}, H^{l_i}))$$

$$A^{l_i}_{jk} = \exp(-\frac{d(c_j, c_k)}{2\sigma^2}) \qquad H^{l_i} = \begin{bmatrix} c_i & 0 & 0 \\ c_j & X_j & y_j \\ \dots & \dots & \dots \end{bmatrix}$$

Back propagation

Evaluation: Measurement Campaigns and Data Collections



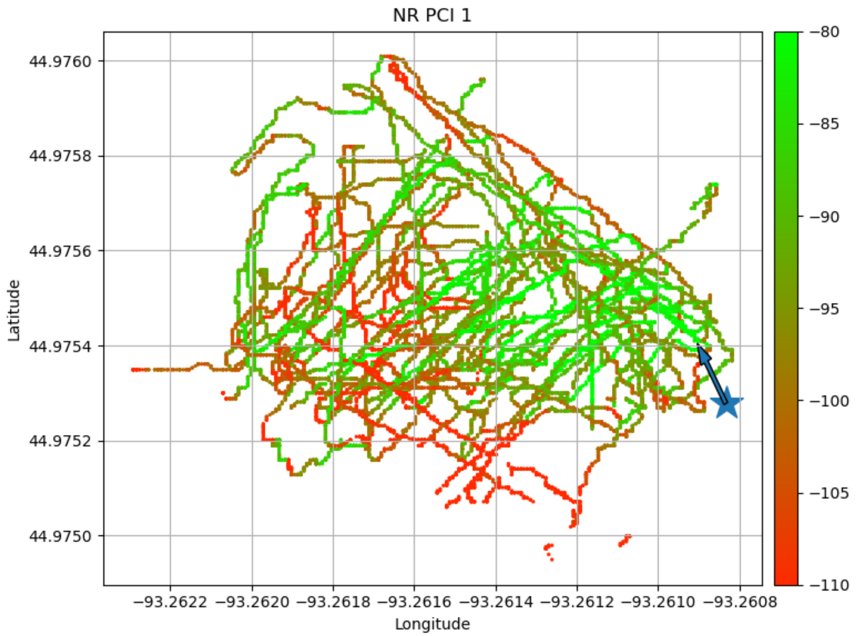


We conducted the **comprehensive measurement campaigns** with the professional tools at the public park covered by diverse 4G/5G bands.

Datasets

Summary of Collected Dataset

Total area covered	8000 m^2
Technologies	4G-lowBand/midBand; 5G-mmWave
Data samples	Total 200k+ with 100ms sampling rate



Data Fields of Signal Imputation Task

c	Longitude; Latitude Angle: the orientation angle of UE
y	BRSRP: filtered beam reference signal received power [dBm]

Data Fields of Channel Quality Regression Task

c	Longitude; Latitude Angle: the orientation angle of UE
X	ss-RSRP: synchronize signal reference signal received power [dBm] ss-RSRQ: synchronize signal reference signal received quality [dB] csi-RSRP [SSBRI]: CSI reference signal received power [dBm] csi-RSRP [CRI]: CSI reference signal received power [dBm] SINR: signal to interference & noise ratio [dB] Pathloss: reduction in power density as signal propagates [dB] BLER: block error rate [%]
y	CQI: channel quality indicator

We use the **collected commercial 5G data** and also **DeepMIMO simulated data** for evaluation.

Evaluation Setup

- We compare 5GNN with other **two** state-of-the-art learning **paradigms** (*PE-GNN [NeurIPS-ws'22]* and *Kriging-GNN [AAAI'20]*) for "geography inference" problems.

- We consider **three** representative **baseline GNNs** (*GCN [ICLR'17]*, *GraphSAGE [NeurIPS'17]*, and *GIN [ICLR'19]*) to combine with above learning paradigms.

- We also consider **one** widely-used classical statistics-based method --- Universal Kriging (UK)

Metrics for Comparisons: RSME and MAE

Results of Signal Imputation Task: GCN as an example

Note: P1 denotes PE-GNN paradigm; P2 is Kriging-GNN paradigm.

Datasets	Metrics	UK	GCN		
			P1	P2	5GNN
DeepMIMO_Mid	RMSE	0.0465	0.0584	0.0451	**0.0440**
	MAE	0.0342	0.0444	0.0334	**0.0320**
DeepMIMO_High	RMSE	0.0767	0.0722	0.0720	**0.0701**
	MAE	0.0584	0.0552	0.0550	**0.0535**
4G_Signal_Low	RMSE	0.1840	0.1216	0.1129	**0.1076**
	MAE	0.1430	0.0955	0.0864	**0.0784**
4G_Signal_Mid	RMSE	0.0899	0.0943	0.0849	**0.0806**
	MAE	0.0684	0.0740	0.0662	**0.0616**
5G_Signal_High	RMSE	0.1588	0.1598	0.1453	**0.1366**
	MAE	0.1201	0.1297	0.1156	**0.1015**

- *5GNN* is **consistently** superior other graph-based learning **Paradigms**.
- It reduces errors up to **12.8%** on the imputation task.

Results of Signal Imputation Task: GCN as an example

Note: P1 denotes PE-GNN paradigm; P2 is Kriging-GNN paradigm.

Datasets	Metrics	UK	GCN P1	GCN P2	GCN 5GNN
DeepMIMO_Mid	RMSE	0.0465	0.0584	0.0451	**0.0440**
	MAE	0.0342	0.0444	0.0334	**0.0320**
DeepMIMO_High	RMSE	0.0767	0.0722	0.0720	**0.0701**
	MAE	0.0584	0.0552	0.0550	**0.0535**
4G_Signal_Low	RMSE	0.1840	0.1216	0.1129	**0.1076**
	MAE	0.1430	0.0955	0.0864	**0.0784**
4G_Signal_Mid	RMSE	0.0899	0.0943	0.0849	**0.0806**
	MAE	0.0684	0.0740	0.0662	**0.0616**
5G_Signal_High	RMSE	0.1588	0.1598	0.1453	**0.1366**
	MAE	0.1201	0.1297	0.1156	**0.1015**

- *5GNN* also outperforms the Kriging methods and reduces up to **25.3%** error rates on the **5G high band**.

Results of Channel Quality Regression Task: GCN as an example

Note: P1 denotes PE-GNN paradigm; P2 is Kriging-GNN paradigm.

Datasets	Metrics	UK	GCN P1	GCN P2	GCN 5GNN
4G_CQI_Low	RMSE	0.1840	0.1859	0.1710	**0.1611**
	MAE	0.1430	0.1518	0.1365	**0.1259**
4G_CQI_Mid	RMSE	0.1435	0.1387	**0.1328**	**0.1328**
	MAE	0.1092	0.1113	0.1053	**0.1017**
5G_CQI_High	RMSE	0.1926	0.1818	0.1724	**0.1643**
	MAE	0.1516	0.1523	0.1417	**0.1287**

- *5GNN* reduces errors up to **9.2%** on the regression task.
- *5GNN* can efficiently capture signals' **local and global spatial correlations**.

5GNN using different GNN baselines

Signal Imputation Task

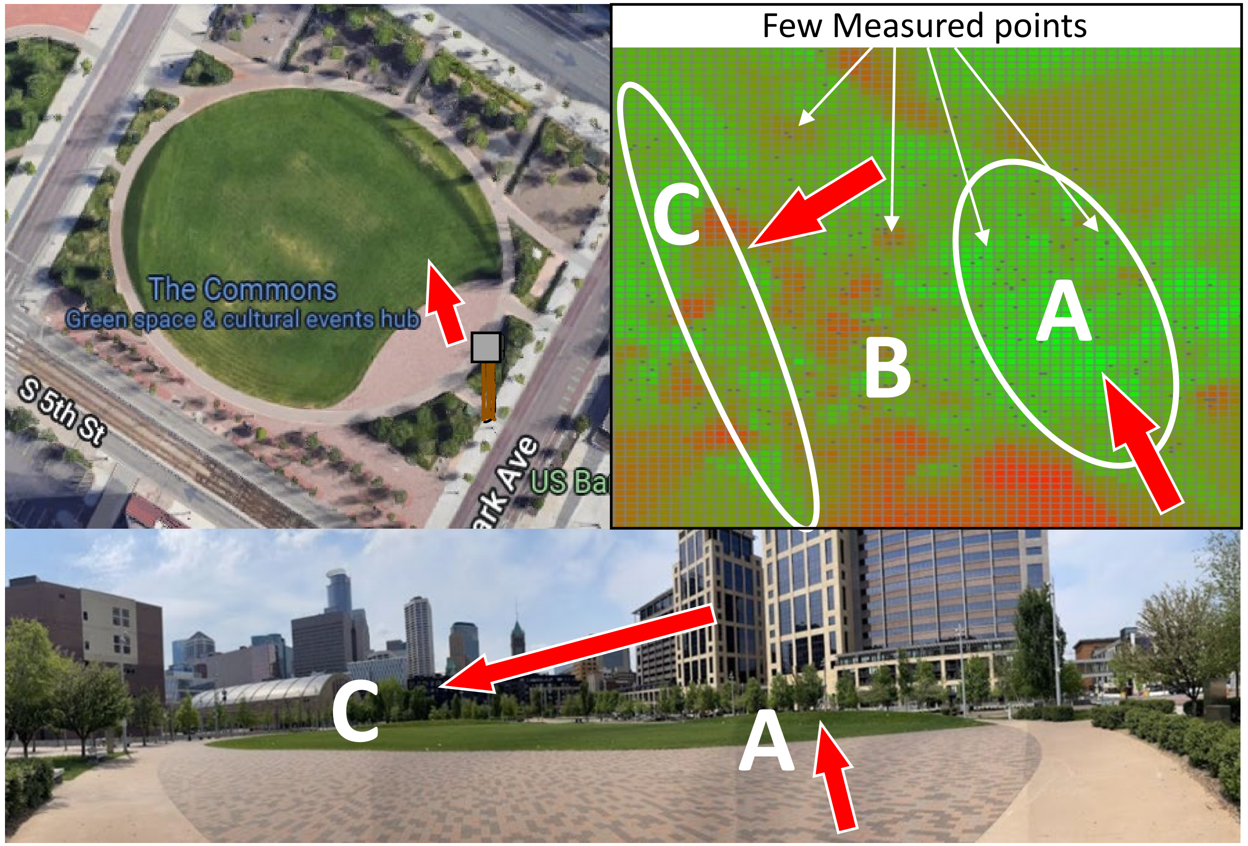
Datasets	Metrics	UK	GCN			GraphSAGE			GIN		
			P1	P2	5GNN	P1	P2	5GNN	P1	P2	5GNN
DeepMIMO_Mid	RMSE	0.0465	0.0584	0.0451	0.0440	0.0559	0.0438	0.0437	0.0531	0.0458	**0.0436**
	MAE	0.0342	0.0444	0.0334	0.0320	0.0423	**0.0315**	**0.0315**	0.0410	0.0344	0.0316
DeepMIMO_High	RMSE	0.0767	0.0722	0.0720	0.0701	0.0723	0.0696	**0.0690**	0.0722	0.0719	0.0703
	MAE	0.0584	0.0552	0.0550	0.0535	0.0552	0.0524	**0.0519**	0.0555	0.0545	0.0537
4G_Signal_Low	RMSE	0.1250	0.1216	0.1129	0.1076	0.1039	**0.1018**	0.1020	0.1223	0.1130	0.1071
	MAE	0.0915	0.0955	0.0864	0.0784	0.0771	0.0758	**0.0745**	0.0964	0.0874	0.0787
4G_Signal_Mid	RMSE	0.0899	0.0943	0.0849	0.0806	0.0887	0.0796	**0.0795**	0.0939	0.0851	0.0816
	MAE	0.0684	0.0740	0.0662	0.0616	0.0688	**0.0603**	**0.0603**	0.0735	0.0663	0.0622
5G_Signal_High	RMSE	0.1588	0.1598	0.1453	0.1366	0.1218	0.1194	**0.1187**	0.1574	0.1457	0.1361
	MAE	0.1201	0.1297	0.1156	0.1015	0.0889	0.0863	**0.0855**	0.1260	0.1157	0.1009

Channel Quality Regression

Datasets	Metrics	UK	GCN			GraphSAGE			GIN		
			P1	P2	5GNN	P1	P2	5GNN	P1	P2	5GNN
4G_CQI_Low	RMSE	0.1840	0.1813	0.1710	0.1611	0.1754	**0.1594**	0.1605	0.1809	0.1712	0.1602
	MAE	0.1430	0.1460	0.1365	0.1259	0.1416	**0.1234**	0.1252	0.1472	0.1370	0.1247
4G_CQI_Mid	RMSE	0.1435	0.1336	0.1328	0.1328	0.1310	0.1278	**0.1276**	0.1345	0.1364	0.1329
	MAE	0.1092	0.1042	0.1053	0.1017	0.1029	0.0982	**0.0972**	0.1045	0.1076	0.1015
5G_CQI_High	RMSE	0.1926	0.1730	0.1724	0.1643	0.1751	0.1645	**0.1629**	0.1748	0.1726	0.1638
	MAE	0.1516	0.1455	0.1417	0.1287	0.1469	**0.1279**	**0.1279**	0.1493	0.1422	0.1292

- *5GNN* consistently obtains superior or compatible results under different baseline GNNs.

Visualization of Radio Maps



5GNN can efficiently **generate the radio map** based on the **few** measured data points, thus assisting future measurements.

Summary

- Put forth the **5G measurement extrapolation problem**
  - Demonstrate 5G measurement is low efficient.
  - Argue for the need to account for **local** and **global** dependencies in extrapolation problem.

- Propose ***5GNN***
  - A tower information-free, physical-inspired, and graph-based learning paradigm.
  - *5GNN* reduces the error rates (up to) **12.8%** and **9.2%** on the signal imputation task and channel quality regression task.

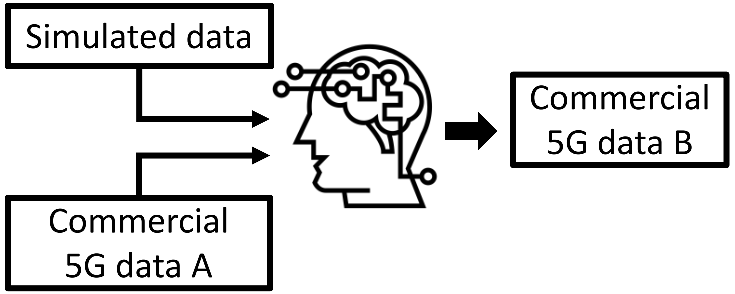
- Conduct comprehensive **Experiments**
  - Collect the commercial 5G network data.
  - For dataset access, codes and more info, visit us @

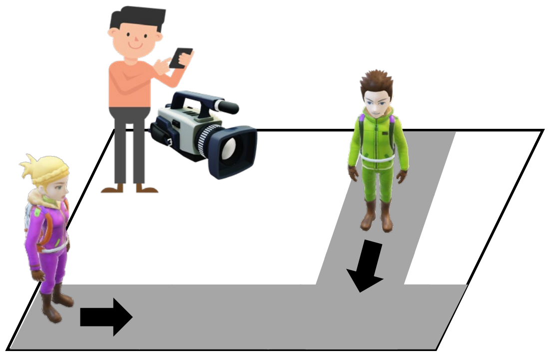
https://github.com/StrongWeiUMN/5GNN

Thank you

Questions

Ongoing and Future Works





Machine learning model generalization ability

Real-time environmental perception

AI-assisted measurement route recommendation

Back up

Results of Signal Imputation Task

Datasets	Metrics	UK	GCN			GraphSAGE			GIN		
			P1	P2	5GNN	P1	P2	5GNN	P1	P2	5GNN
DeepMIMO_Mid	RMSE	0.0465	0.0584	0.0451	0.0440	0.0559	0.0438	0.0437	0.0531	0.0458	**0.0436**
	MAE	0.0342	0.0444	0.0334	0.0320	0.0423	**0.0315**	**0.0315**	0.0410	0.0344	0.0316
DeepMIMO_High	RMSE	0.0767	0.0722	0.0720	0.0701	0.0723	0.0696	**0.0690**	0.0722	0.0719	0.0703
	MAE	0.0584	0.0552	0.0550	0.0535	0.0552	0.0524	**0.0519**	0.0555	0.0545	0.0537
4G_Signal_Low	RMSE	0.1250	0.1216	0.1129	0.1076	0.1039	**0.1018**	0.1020	0.1223	0.1130	0.1071
	MAE	0.0915	0.0955	0.0864	0.0784	0.0771	0.0758	**0.0745**	0.0964	0.0874	0.0787
4G_Signal_Mid	RMSE	0.0899	0.0943	0.0849	0.0806	0.0887	0.0796	**0.0795**	0.0939	0.0851	0.0816
	MAE	0.0684	0.0740	0.0662	0.0616	0.0688	**0.0603**	**0.0603**	0.0735	0.0663	0.0622
5G_Signal_High	RMSE	0.1588	0.1598	0.1453	0.1366	0.1218	0.1194	**0.1187**	0.1574	0.1457	0.1361
	MAE	0.1201	0.1297	0.1156	0.1015	0.0889	0.0863	**0.0855**	0.1260	0.1157	0.1009

Note: P1 denotes PE-GNN paradigm; P2 is Kriging-GNN paradigm.

- *5GNN* is **consistently** superior other graph-based learning paradigms.
- It reduces errors up to **12.8%** on the imputation task.

Results of Signal Imputation Task

Datasets	Metrics	UK	GCN			GraphSAGE			GIN		
			P1	P2	5GNN	P1	P2	5GNN	P1	P2	5GNN
DeepMIMO_Mid	RMSE	0.0465	0.0584	0.0451	0.0440	0.0559	0.0438	0.0437	0.0531	0.0458	**0.0436**
	MAE	0.0342	0.0444	0.0334	0.0320	0.0423	**0.0315**	**0.0315**	0.0410	0.0344	0.0316
DeepMIMO_High	RMSE	0.0767	0.0722	0.0720	0.0701	0.0723	0.0696	**0.0690**	0.0722	0.0719	0.0703
	MAE	0.0584	0.0552	0.0550	0.0535	0.0552	0.0524	**0.0519**	0.0555	0.0545	0.0537
4G_Signal_Low	RMSE	0.1250	0.1216	0.1129	0.1076	0.1039	**0.1018**	0.1020	0.1223	0.1130	0.1071
	MAE	0.0915	0.0955	0.0864	0.0784	0.0771	0.0758	**0.0745**	0.0964	0.0874	0.0787
4G_Signal_Mid	RMSE	0.0899	0.0943	0.0849	0.0806	0.0887	0.0796	**0.0795**	0.0939	0.0851	0.0816
	MAE	0.0684	0.0740	0.0662	0.0616	0.0688	**0.0603**	**0.0603**	0.0735	0.0663	0.0622
5G_Signal_High	RMSE	0.1588	0.1598	0.1453	0.1366	0.1218	0.1194	**0.1187**	0.1574	0.1457	0.1361
	MAE	0.1201	0.1297	0.1156	0.1015	0.0889	0.0863	**0.0855**	0.1260	0.1157	0.1009

Note: P1 denotes PE-GNN paradigm; P2 is Kriging-GNN paradigm.

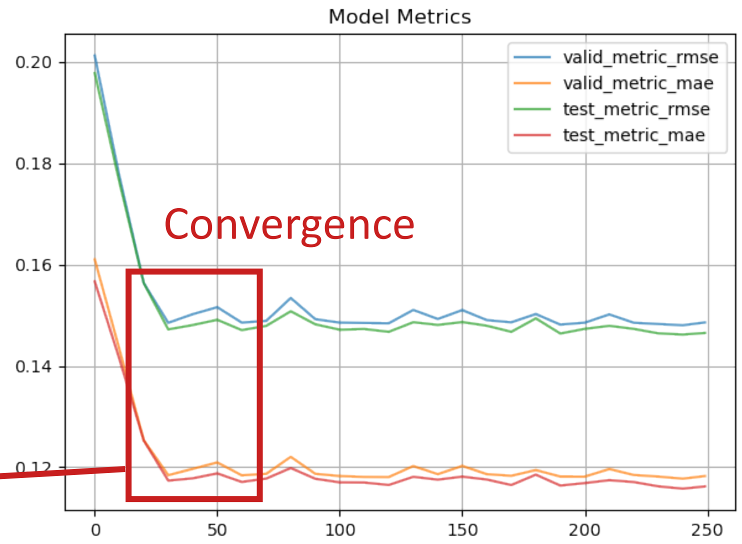
- *5GNN* also outperforms the Kriging methods and reduces up to **25.3%** error rates on the **5G high band**.

Discussions on Running Time

We use 64 vCPU cores for universal kriging, while 8 vCPU cores with 1 GPU for other GCN-based methods. Those machines **cost the same price** on AWS as a fair comparison.

	Training Time		Inference Time
UK	182.4 s		0.135 ms/sample
GCN-P1	0.39 s/epoch	45 epochs	0.006 ms/sample
GCN-P2	0.18 s/epoch	42 epochs	0.178 ms/sample
GCN-5GNN	0.29 s/epoch	46 epochs	0.183 ms/sample

Note: P1 denotes PE-GNN paradigm; P2 is Kriging-GNN paradigm.



- Different learning paradigms have **different characteristics** of training/inference time.
- *5GNN* is competitive with other methods and can **meet the requirements** for real-world measurement, considering each measurement run usually takes tens of minutes.

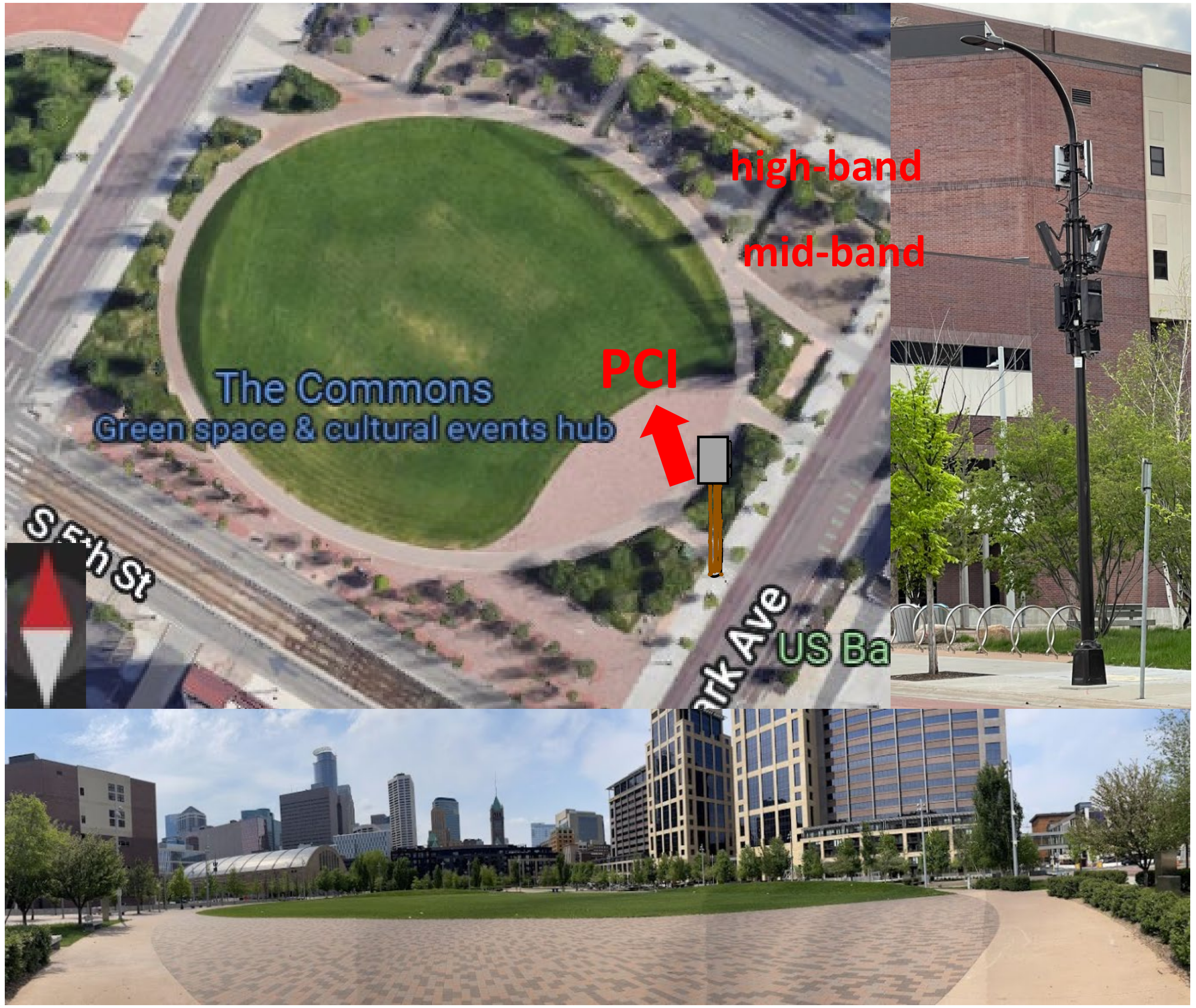
Old Slides

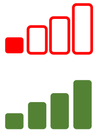
5GNN: Extrapolating 5G Measurements through GNNs

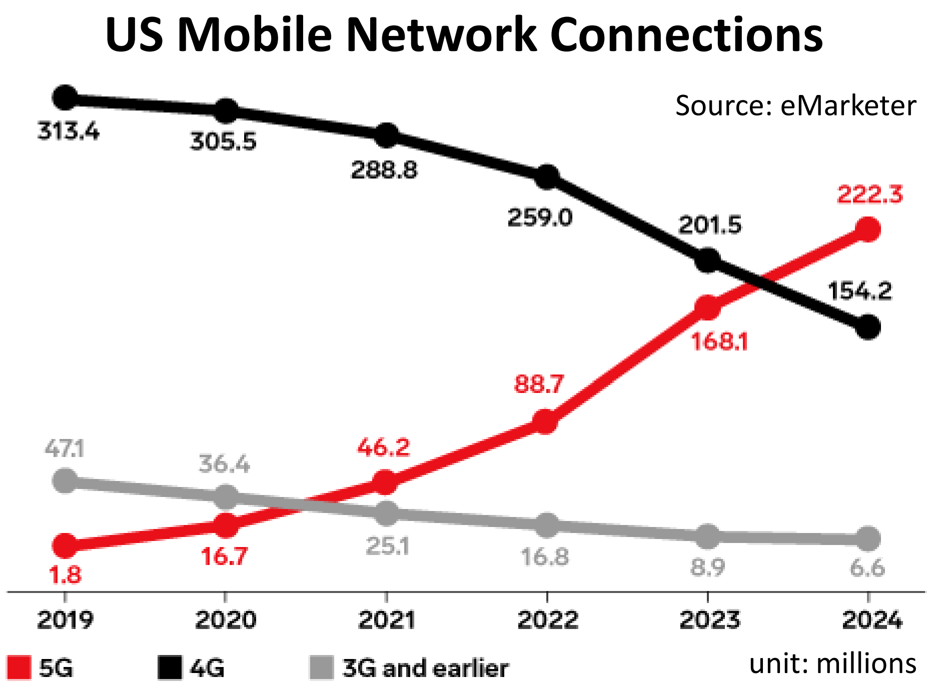
https://github.com/StrongWeiUMN/5GNN

Wei Ye, Xinyue Hu, Tian Liu, Ruoyu Sun,
Yanhua Li, **Zhi-Li Zhang**

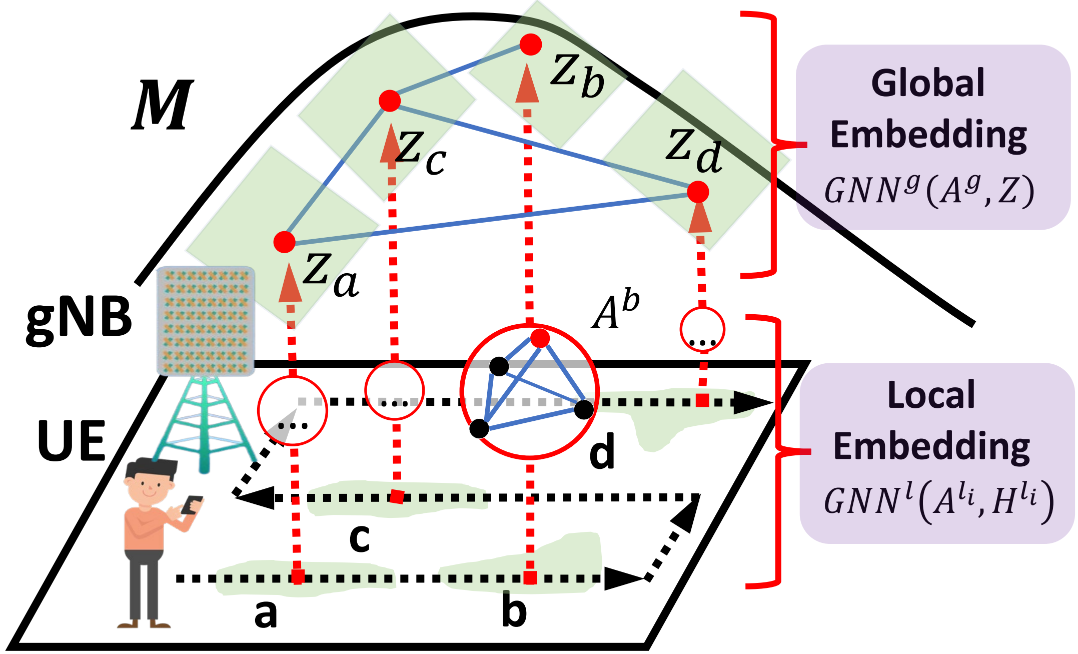
UNIVERSITY OF MINNESOTA

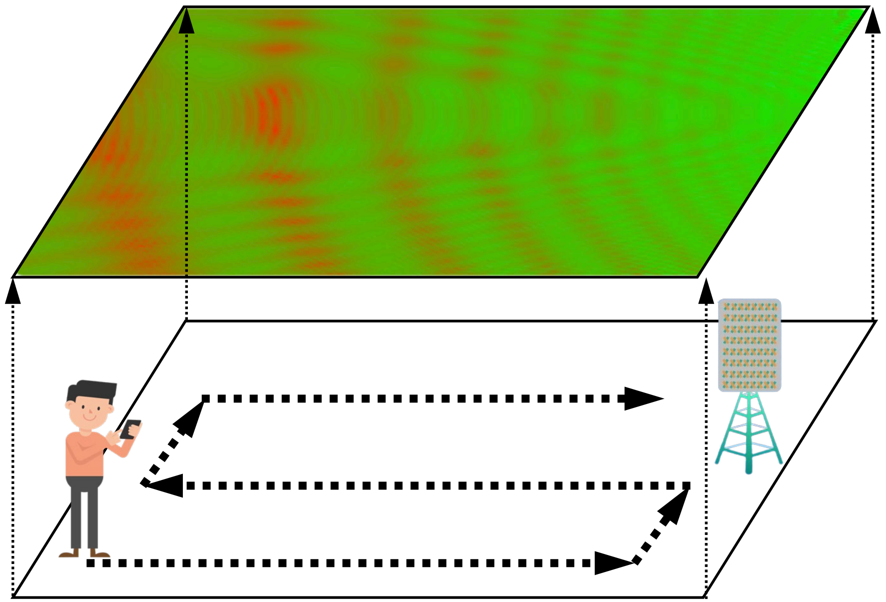

high-band
mid-band
PCI
The Commons
Green space & cultural events hub
S 5th St
ark Ave
US Ba




US Mobile Network Connections
Source: eMarketer
313.4
305.5
288.8
259.0
201.5
154.2
222.3
168.1
88.7
46.2
16.7
1.8
47.1
36.4
25.1
16.8
8.9
6.6
2019
2020
2021
2022
2023
2024
5G
4G
3G and earlier
unit: millions

Lumos5G achieves an overall weighted average F1 score of up to 0.96 (with 3 classes) and 1.37× to 4.84× reduction in throughput prediction error compared to existing approaches.





Summary

Address the **5G measurement extrapolation problem**.

Argue for the need to account for both **local** and **global** dependencies in 5G signal and feature maps.

Propose ***5GNN*** - a tower information-free, physical-inspired, and graph-based learning paradigm.

5GNN reduces the error rates (up to) **12.8%** and **9.2%** on the signal imputation task and channel quality regression task, compared to other state-of-art methods on both synthetic and real-world datasets.

Conduct field experiments to collect the **commercial 5G network data** for this study.

For dataset access and more info, visit us @

https://github.com/StrongWeiUMN/5GNN