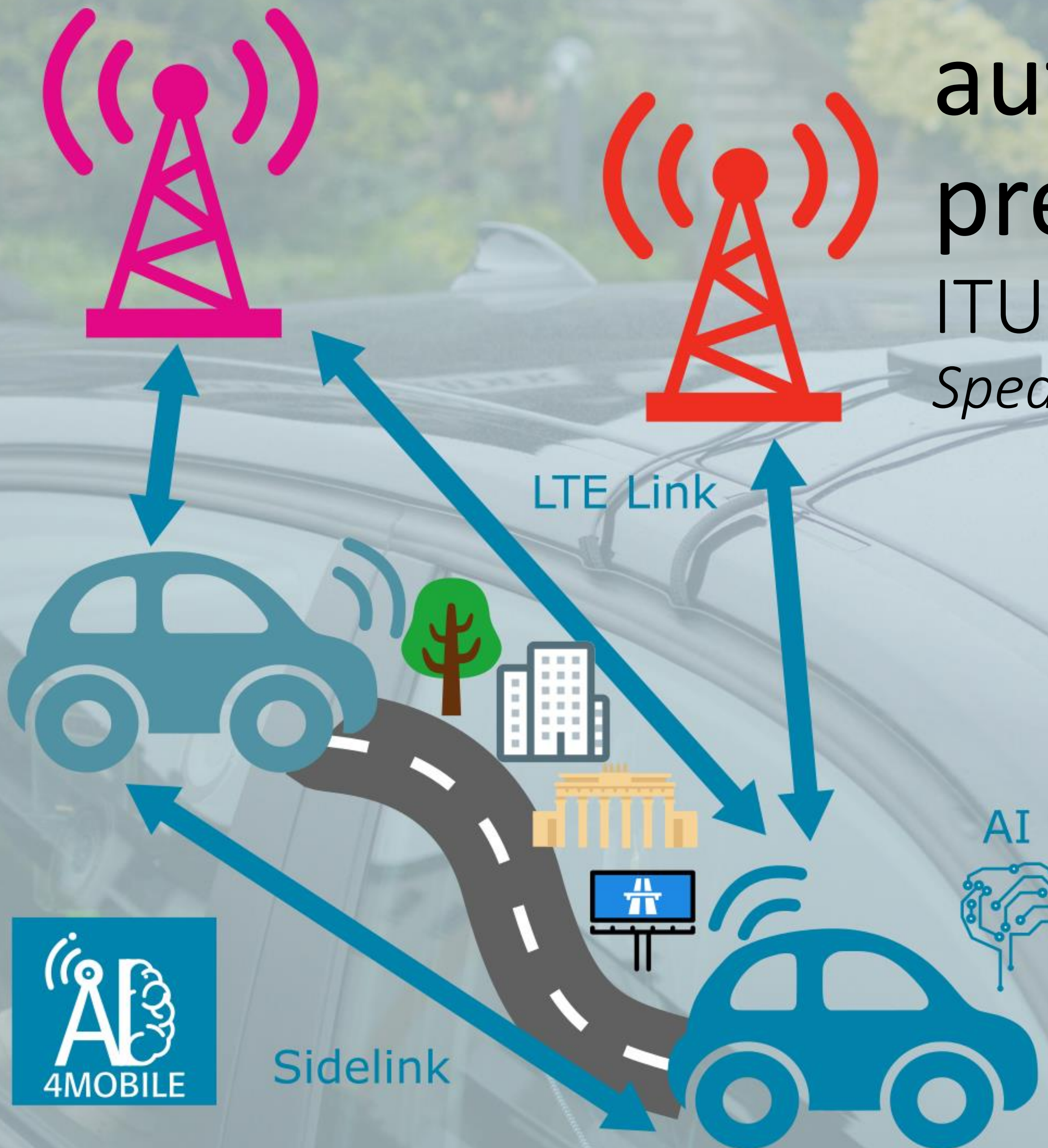
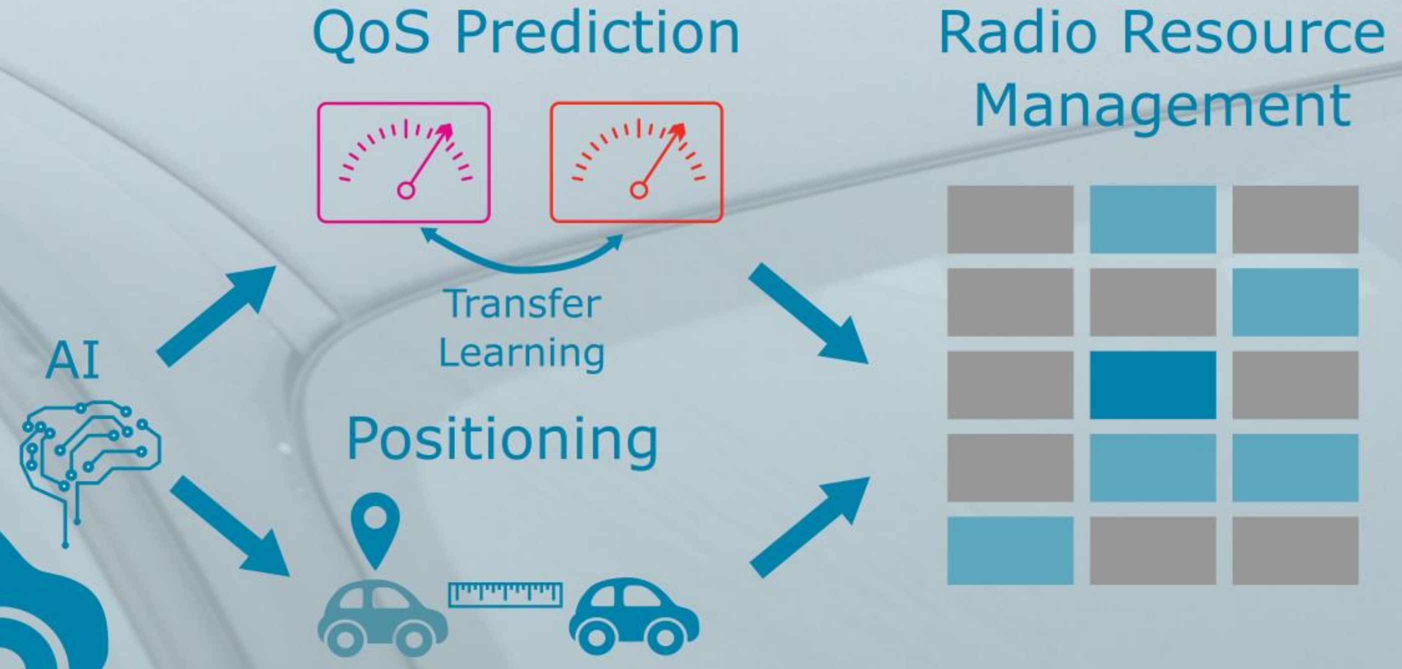
Multi-environment automotive QoS prediction using AI/ML

ITU AI/ML in 5G Challenge in 2023

Speaker: Rodrigo Hernangómez
(Fraunhofer Heinrich Hertz Institute)





GEFÖRDERT VOM
Bundesministerium
für Bildung
und Forschung

AI4Mobile

AI-AIDED WIRELESS SYSTEMS FOR MOBILITY IN INDUSTRY AND TRAFFIC

- Research project funded by the German Ministry for Education and Research
- Consortium of 3 multinationals, 3 SMEs and 3 research institutions
  - Duration: 2020-2023
- **Research goal** – AI/ML for predictive Quality of Service (QoS) at high mobility
  - Twofold meaning of high mobility:
    - Traffic: Connected cars
    - Industry: Automated Guided Vehicles
- More info at ai4mobile.org

GEFÖRDERT VOM
Bundesministerium
für Bildung
und Forschung

AI4Mobile

DATASETS

- Goal: Test ML algorithms for predictive QoS
- Extensive measurement campaigns available at IEEE Dataport
- Broad scope: Cellular/Sidelink/Campus network; Automotive/Industry

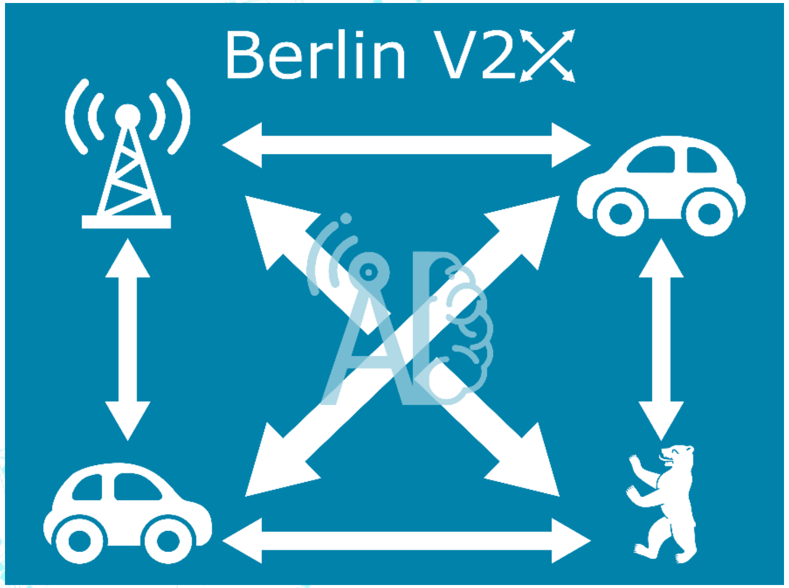
Datasets				
	Automotive	Highway	Sensor data	Cellular
	Industry	Urban	Campus-Net	Sidelink
Berlin V2X	Berlin Vehicle to Everything			
iV2I+	industrial Vehicle to Infrastructure + Sensor			
iV2V	industrial Vehicle to Vehicle			

4MOBILE

Multi-environment automotive QoS prediction with AI

OUTLINE

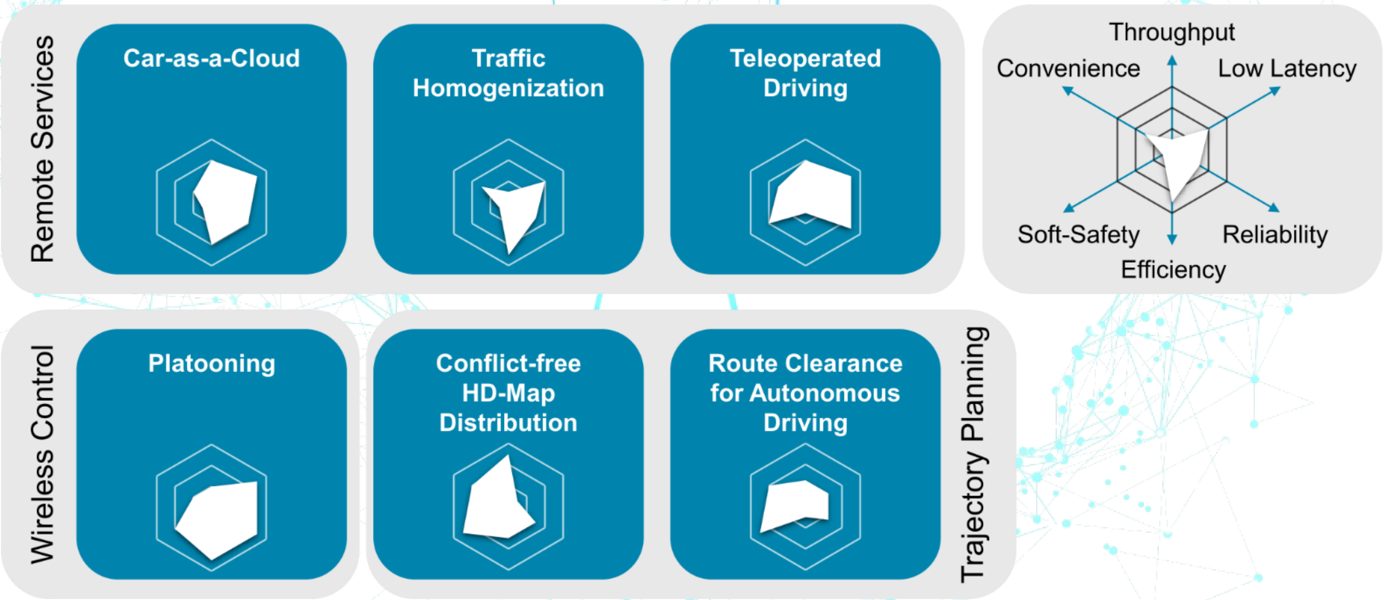
- Motivation & Challenges
- Dataset
- Problem Statement



Multi-environment automotive QoS prediction with AI

Motivation

- Vehicle-to-everything (V2X) communication at the core of new car services
- High demands on Quality of Service (QoS) and proactive resource allocation
  - Leverage boom of ML in communication networks for QoS prediction



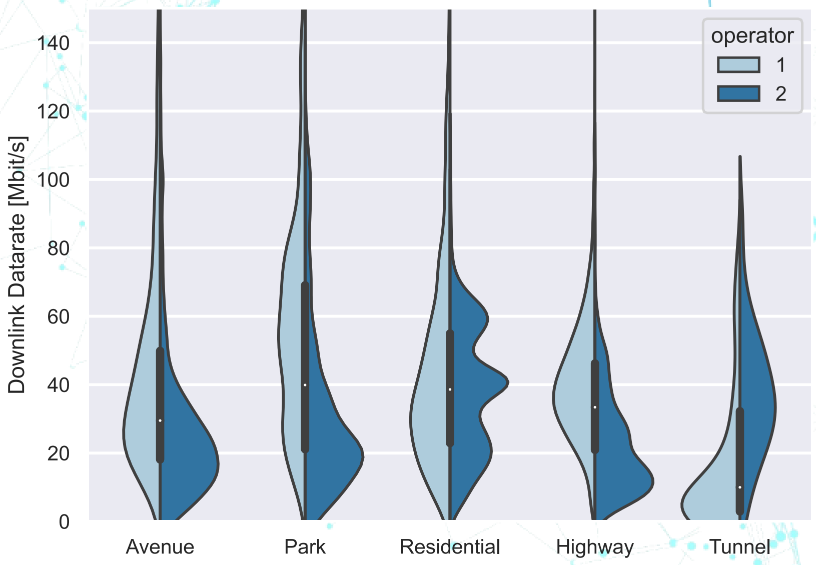
(Külzer *et al.*, 2021)

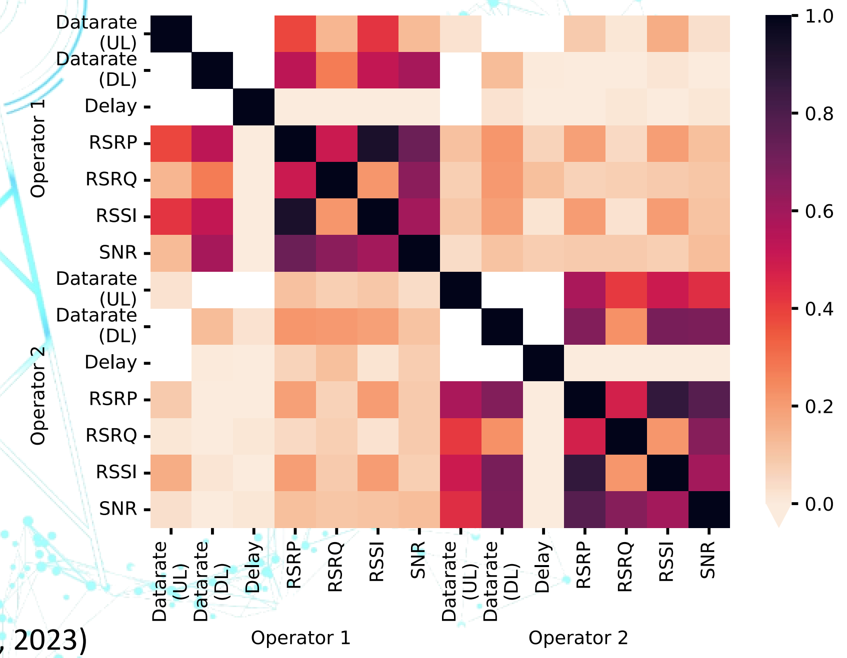
GEFÖRDERT VOM Bundesministerium für Bildung und Forschung

Multi-environment automotive QoS prediction with AI

Challenges for AI/ML

- Drastic changes to the radio environment → Data distribution drifts (no i.i.d.)
- Generalization across different entities (vehicles/operators) or scenarios
  - Domain adaptation techniques might be helpful





(Hernangómez *et al.*, 2023)

GEFÖRDERT VOM
Bundesministerium für Bildung und Forschung

GEFÖRDERT VOM
Bundesministerium für Bildung und Forschung

BerlinV2X Dataset

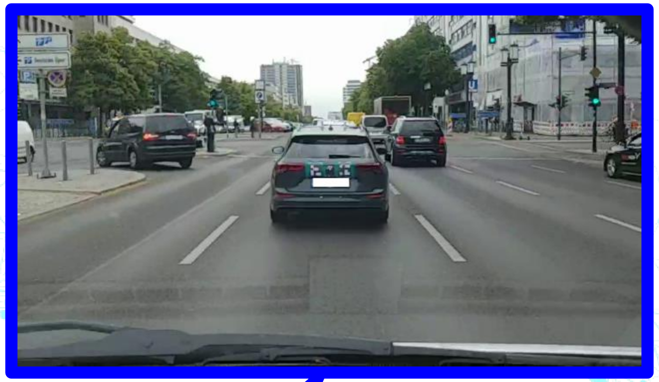
Measurement Campaign

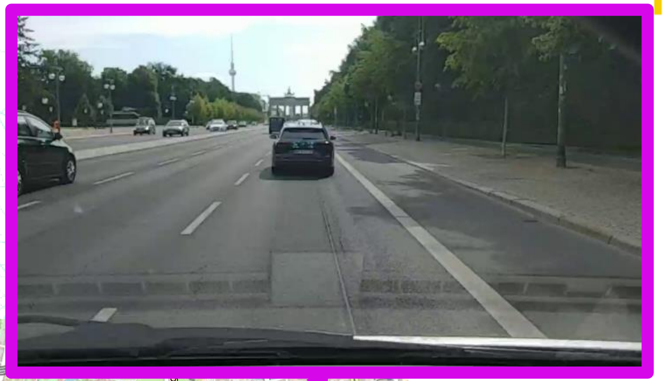
- 2 commercial LTE mobile network operators
  - LTE frequency bands 700-2700 MHz
- V2V sidelink (3GPP Rel. 14, PC5 mode 4)
  - Out of the challenge's scope
- Simultaneous measurements from up to 4 vehicles
  - 2 vehicles per operator
  - Different drive modes: platoon vs. pair driving
  - 3 throughput profiles for diverse QoS measurements
- 17 rounds over 3 days
  - 1 round = 17 km across West Berlin
  - 45-60 minutes on a weekday
  - Diverse urban areas

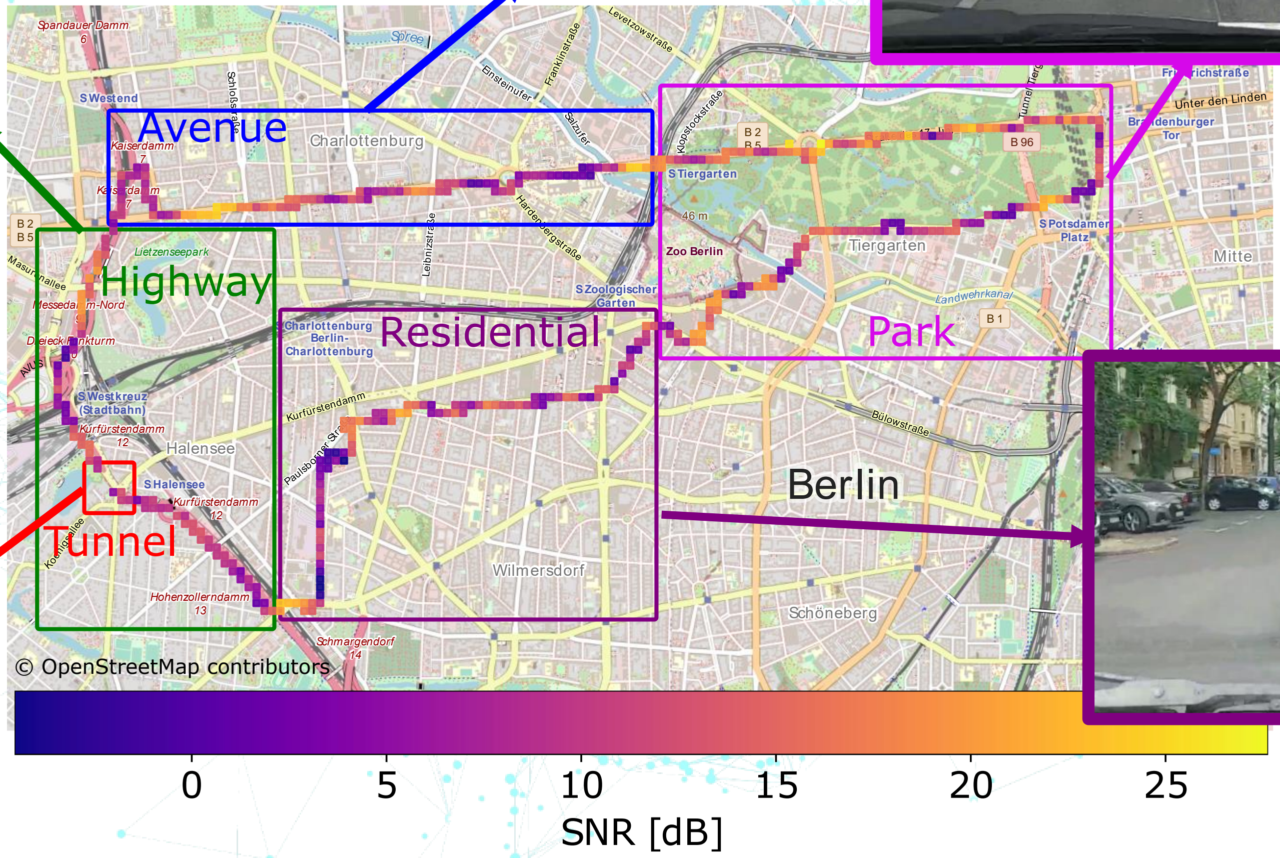
GEFÖRDERT VOM
Bundesministerium
für Bildung
und Forschung

BerlinV2X Dataset

URBAN AREAS







4MOBILE

GEFÖRDERT VOM
Bundesministerium für Bildung und Forschung

BerlinV2X Dataset

Measurement methodology

- Full LTE stack captured with MobileInsight (MI)
  - (Li *et al.*, 2016)
- Traffic exchange with a server located at HHI
  - Datarate and jitter measurements with iperf
  - Ping-based delay measurements
  - Target datarate according to QoS measurement
- GPS localization
- Side information via APIs
  - Weather (DarkSky)
  - Traffic conditions (HERE)

Throughput profiles	QoS measurement	Target datarate
Low throughput	UL/DL delay	400 kbps
High throughput	UL datarate	75 Mbps
High throughput	DL datarate	350 Mbps

BerlinV2X Dataset

Captured Data

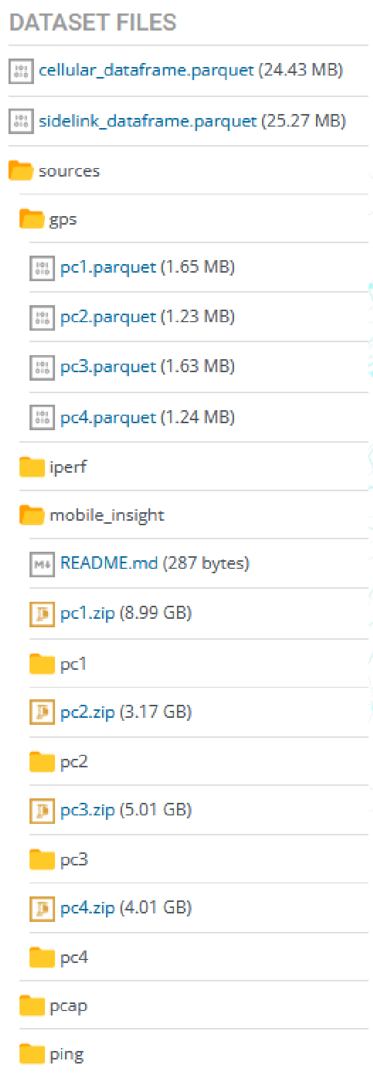
Data category	Source	Tool	Sampling interval	Features
LTE stack	In-vehicle device	MobileInsight	10 ms	PHY: SNR, RSRP, RSRQ, RSSI
			20 ms	PDSCH/PUSCH: RBs, TB Size, DL MCS, UL Tx Power
			Event-based	RRC: Cell Identity, DL/UL frequency, DL/UL bandwidth
Quality of Service	In-vehicle device	ping	1 s	Delay
		iperf	1 s	DL Datarate, Jitter
	Server	iperf	1 s	UL Datarate, Jitter
Position	GPS	NA	1 s	Latitude, Longitude, Altitude, Velocity, Heading
Side information	Internet database	HERE API	5 min	Traffic Jam Factor, Traffic Street Name, Traffic Distance
		DarkSky	1 hour	Cloud cover, Humidity, Precipitation Intensity & Probability, Temperature, Pressure, Wind Speed
Metadata	NA	NA	NA	Scenario, operator, drive type, target datarate, direction

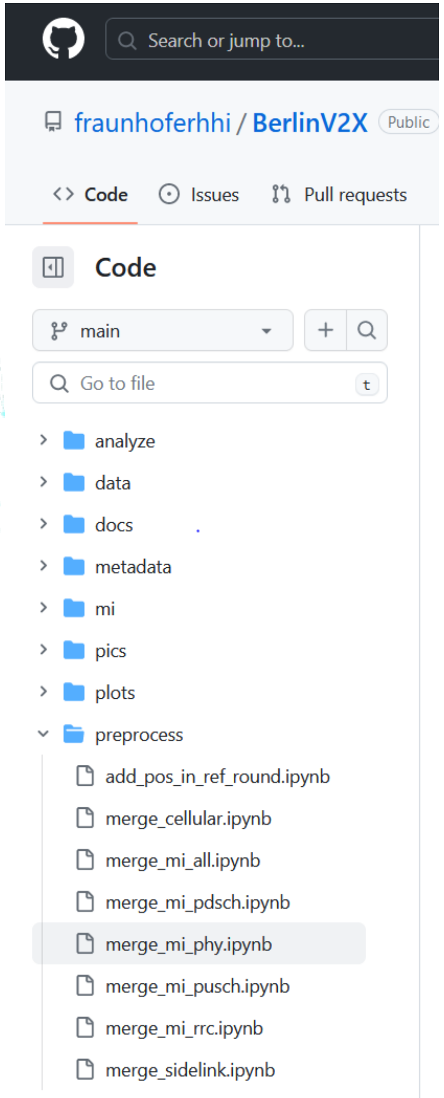
GEFÖRDERT VOM
Bundesministerium für Bildung und Forschung

BerlinV2X Dataset

Preprocessing

- Raw measurements as [parquet](#) files in [IEEE Dataport](#) (sources)
- Preprocessing code in [GitHub](#) under [preprocess](#)
  - Downsample to 1s
  - Extract relevant LTE params from MI (`merge_mi_*.ipynb`).
    - Merge on device, cell, and timestamp (`merge_mi_all.ipynb`)
    - Concatenate values for primary and secondary cells (carrier aggregation)
  - Merge all data on timestamp and device (`merge_cellular.ipynb`)
    - Select QoS data from server/device
    - Label with measurement metadata





BerlinV2X Dataset

PREPROCESSING

- **Excerpt of** `cellular_dataframe.parquet`

timestamp	device	PCell_RSSI _max	PCell_Cell _Identity	SCell_RSSI _max	SCell_Cell _Identity	...	ping_ms	datarate	Latitude	Longitude	temperature	Traffic Jam Factor	measured_ qos	drive_mode	target_ datarate	direction	operator	area
2021-06-22 14:34:10	pc4	-47.118750	29127680.0	-52.660625	29127683.0	...	2238.0	37700000.0	52.514862	13.322625	21.57	2.53887	datarate	platoon	350000000	downlink	1	Avenue
2021-06-22 14:34:10	pc3	-58.760000	51447562.0	-69.550000	51447567.0	...	47.7	403000.0	52.515300	13.323007	21.62	2.53887	delay	platoon	400000	downlink	2	Avenue
2021-06-22 14:34:11	pc3	-61.433125	51447562.0	-75.035625	51447567.0	...	41.9	403000.0	52.515213	13.322935	21.62	2.53887	delay	platoon	400000	downlink	2	Avenue
2021-06-22 14:34:11	pc2	-93.064375	51447562.0	-92.653750	NaN	...	35.9	403000.0	52.514997	13.322730	21.57	2.53887	delay	platoon	400000	downlink	2	Avenue
2021-06-22 14:34:12	pc2	-92.622500	51447562.0	-92.585625	NaN	...	44.9	403000.0	52.514923	13.322672	21.57	2.53887	delay	platoon	400000	downlink	2	Avenue
2021-06-22 14:34:12	pc3	-62.138125	51447562.0	-74.110000	51447567.0	...	40.7	394000.0	52.515130	13.322865	21.62	2.53887	delay	platoon	400000	downlink	2	Avenue
2021-06-22 14:34:13	pc3	-60.440625	51447562.0	-74.069375	51447567.0	...	38.8	413000.0	52.515048	13.322798	21.62	2.53887	delay	platoon	400000	downlink	2	Avenue
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...

GEFÖRDERT VOM
Bundesministerium für Bildung und Forschung

GEFÖRDERT VOM
Bundesministerium
für Bildung
und Forschung

BerlinV2X Dataset

RESOURCES

- Data on **IEEE Dataport**: https://ieee-dataport.org/open-access/berlin-v2x
- Code on **GitHub**: https://github.com/fraunhoferhhi/BerlinV2X
- Documentation
  - Readme on **GitHub** and **IEEE Dataport**
  - Hernangómez, R. *et al.* (2023) 'Berlin V2X: A Machine Learning Dataset from Multiple Vehicles and Radio Access Technologies', in *2023 IEEE 97th Vehicular Technology Conference (VTC2023-Spring). 2023 IEEE 97th Vehicular Technology Conference (VTC2023-Spring)*, Florence, Italy. Preprint available at: https://doi.org/10.48550/arXiv.2212.10343.

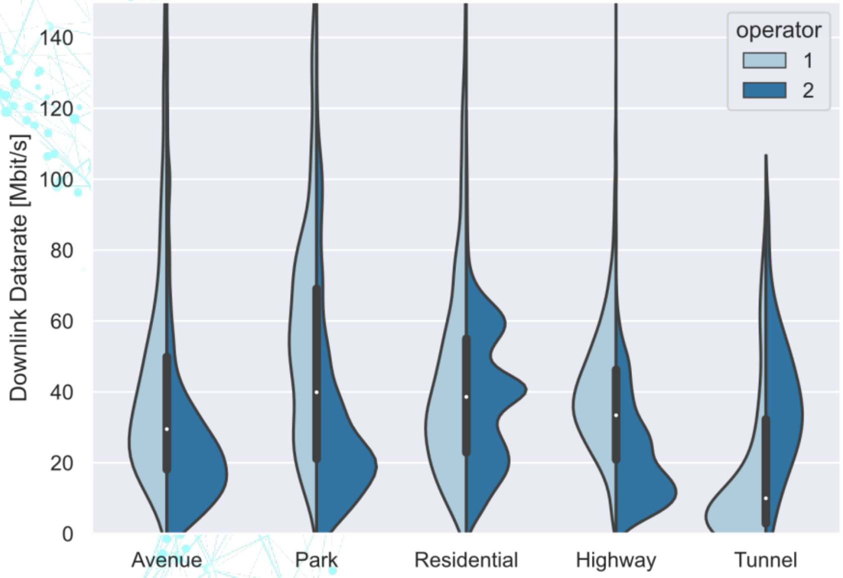
IEEEDataPort™ GitHub VTC2023-Spring FLORENCE Connecting the Mobile World arXiv

AI 4MOBILE

ITU Problem Statement

QoS Prediction across Environments

- Automotive communication is multi-environmental in essence
  - Underlying data distributions differ across vehicles, operators, areas
  - i.i.d. assumptions are systematically violated
- Train/test data split along such environments
  - Random split is often too indulgent (Palaios *et al.*, 2023)



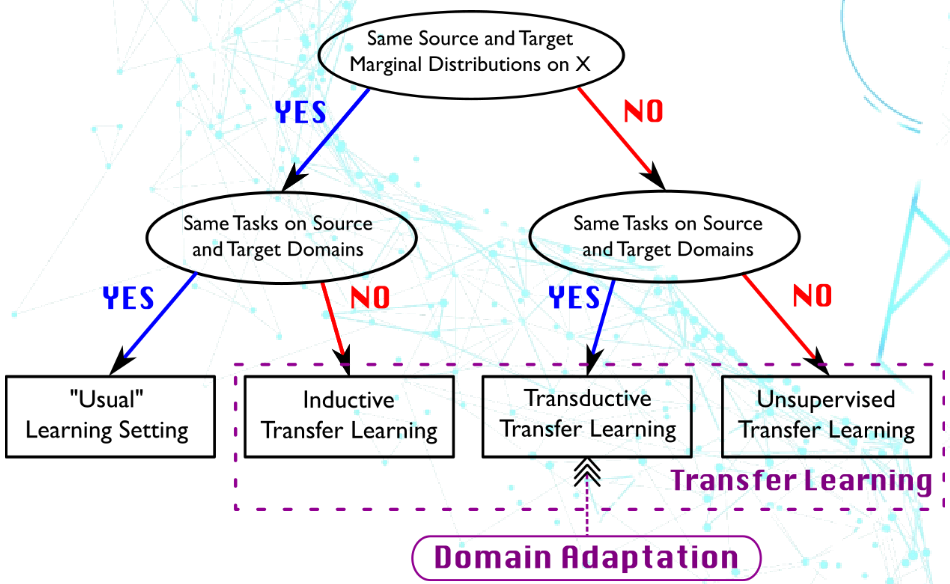
(Hernangómez *et al.*, 2023)

GEFÖRDERT VOM Bundesministerium für Bildung und Forschung

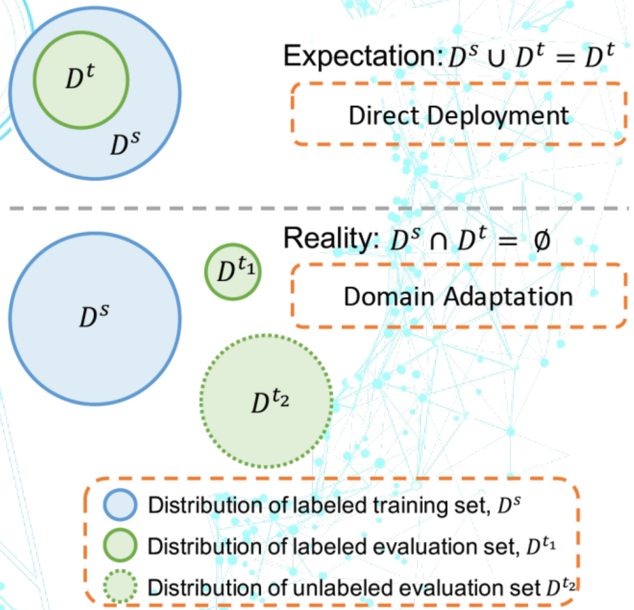
ITU Problem Statement

Domain Adaptation

- Discussed as domain adaptation / transfer learning / concept drift
- Automotive QoS prediction hardly fits the „usual“ learning setting



(Redko *et al.*, 2019)

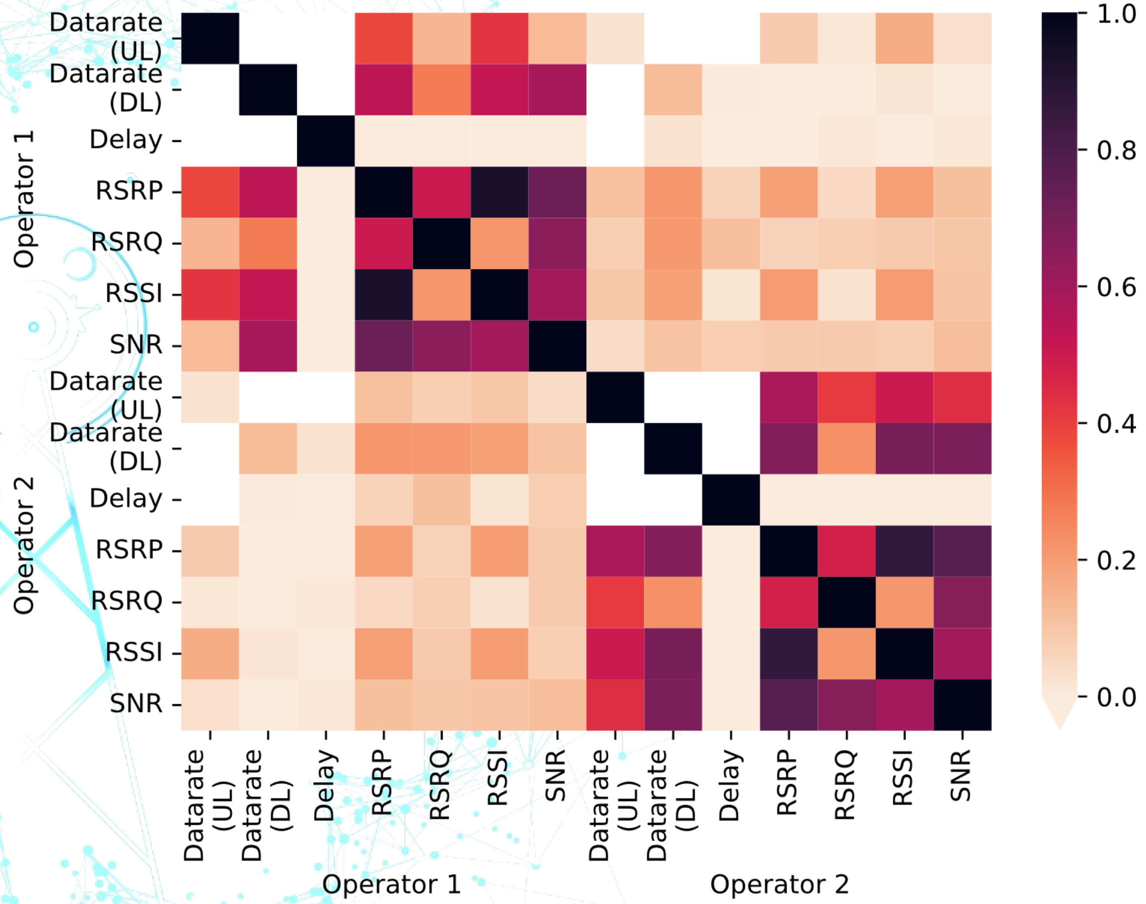


(Xu *et al.*, 2019)

ITU Problem Statement

Open Questions

- Which QoS parameter?
  - Focus on datarate (uplink or downlink)
- Which input features?
- How to split train/test data?
  - Areas
  - Vehicles
  - Operators
  - Uplink/downlink
- Which algorithm?
  - Base ML regressor or Neural Network
  - Fine-tuning techniques: DA/TL



ITU Problem Statement

Reference Example

- Available on GitHub
  - Focused on downlink datarate as QoS with train/test split along operators
- Improvements
  - You can define a different pQoS problem
  - Feature selection is up to you
    - The less features, the better
  - You may choose a different train/test split (along discussed environments)
  - You are free to choose your ML/DL algorithms
  - You may apply (unsupervised) domain adaptation and transfer learning techniques
  - You can create your own preprocessing pipeline
    - Upsample GPS and datarate to ms-range instead of LTE downsampling
    - Include other features from the LTE stack

GEFÖRDERT VOM Bundesministerium für Bildung und Forschung

ITU Problem Statement

Evaluation

- Weighted score
  1. Coefficient of determination R^2
     - Available for Scikit-learn, PyTorch, and Tensorflow:

$$R^2(y, \hat{y}) = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$$

  2. Number of used features:

$$R_{features} = (useful\ features\ - used\ features) = 84 - used\ features$$

  3. Problem setup (qualitative)
     - The choices on predicted QoS, train and test datasets, etc.
- Submission on ITU's challenge platform
  - Scores.csv
  - Code deliverable
  - Brief report (2-5 pages)
  - (Optional) model and weights

team_id	predicted_qos	train_set	test_set	r2_score	used_features
REFERENCE	downlink datarate	operator 1	operator 2	0.806633	39

GEFÖRDERT VOM
Bundesministerium für Bildung und Forschung

GEFÖRDERT VOM
Bundesministerium
für Bildung
und Forschung

ITU Problem Statement

TIMELINE

- Competition Phase
  - Registration from **23 May 2023** to **31 August 2023**
  - Submission deadline **8 September 2023**
  - Evaluation of solutions: **31 October 2023**

- Evaluation Phase
  - **November 2023 –** Judges Panel evaluates the best solutions from the Competition Phase
  - **28 – 30 November 2023** – Best solutions pitch in a 3-day event to determine the finalists
  - **13 December 2023 –** Grand Challenge Finale

- Winning Prize: 1000 CHF + certificates

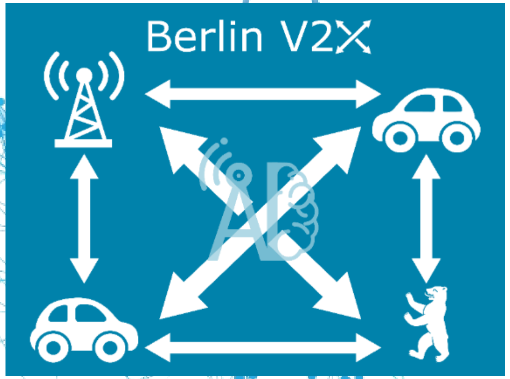
AI 4MOBILE

References

- Hernangómez, R. *et al.* (2023) 'Berlin V2X: A Machine Learning Dataset from Multiple Vehicles and Radio Access Technologies', in *2023 IEEE 97th Vehicular Technology Conference (VTC2023-Spring)*. *2023 IEEE 97th Vehicular Technology Conference (VTC2023-Spring)*, Florence, Italy. Available at: https://doi.org/10.48550/arXiv.2212.10343.
- Külzer, D.F. *et al.* (2021) 'AI4Mobile: Use Cases and Challenges of AI-based QoS Prediction for High-Mobility Scenarios', in *2021 IEEE 93rd Vehicular Technology Conference (VTC2021-Spring)*. *2021 IEEE 93rd Vehicular Technology Conference (VTC2021-Spring)*, Helsinki, Finland, pp. 1–7. Available at: https://doi.org/10.1109/VTC2021-Spring51267.2021.9449059.
- Li, Y. *et al.* (2016) 'MobileInsight: extracting and analyzing cellular network information on smartphones', in *Proceedings of the 22nd Annual International Conference on Mobile Computing and Networking*. *MobiCom'16: The 22nd Annual International Conference on Mobile Computing and Networking*, New York City New York: ACM, pp. 202–215. Available at: https://doi.org/10.1145/2973750.2973751.
- Palaios, A. *et al.* (2023) 'The Story of QoS Prediction in Vehicular Communication: From Radio Environment Statistics to Network-Access Throughput Prediction'. arXiv. Available at: https://doi.org/10.48550/arXiv.2302.11966.
- Redko, I. *et al.* (2019) *Advances in Domain Adaptation Theory*. Elsevier. Available at: https://doi.org/10.1016/C2016-0-05108-2.
- Xu, X. *et al.* (2019) 'd-SNE: Domain Adaptation Using Stochastic Neighborhood Embedding', in. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2497–2506. Available at: http://openaccess.thecvf.com/content_CVPR_2019/html/Xu_d-SNE_Domain_Adaptation_Using_Stochastic_Neighborhood_Embedding_CVPR_2019_paper.html (Accessed: 4 May 2020).

THANK YOU.

Questions?



Contact: ml5g-challenge@ai4mobile.org

GEFÖRDERT VOM
Bundesministerium
für Bildung
und Forschung