

Challenges in RF Data Acquisition, Management, Dissemination

Session 3: Mobile and Wireless Measurement Challenges

Kaushik Chowdhury

Professor and Associate Director for Institute for the Wireless IoT
Northeastern University

Community Pain Points [for the Wireless Community]

- No representative datasets for the wireless community
 - MNIST for image processing community
- Growing interest in ML/AI, hampered by lack of diverse datasets
 - No clear categorization of technology options that intuitively leads to discovery of resources
 - Need for a centralized “focal point” for the wireless community
- Lack of tools to create datasets on wireless testbeds
 - External testbeds are complex, need time investment
 - Privacy concerns
- Lack of tools to share, disseminate datasets
 - Lack of community leadership for standardized metadata representation and sharing
 - Lack of education and awareness for wireless engineers

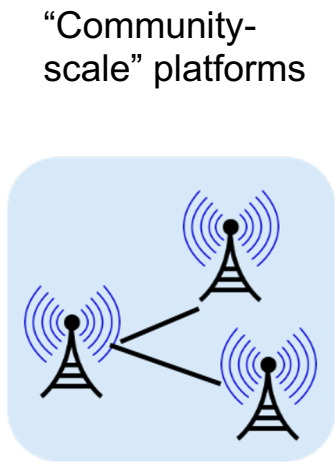
Sources of Wireless Datasets



User Testbeds & Simulation



Colosseum



NSF PAWR Platforms

New Opportunities

A large orange bracket spans across the bottom of the 'Colosseum' and 'NSF PAWR Platforms' sections, with the text 'New Opportunities' centered below it.

NSF PAWR Platforms – Rich Sources of Datasets

NSF PAWR PLATFORMS

POWDER

Salt Lake City, UT



Software defined networks and mMIMO

COSMOS

West Harlem, NY



mmWave, full duplex and backhaul research

AERPAW

Raleigh, NC



Unmanned aerial vehicles, mmWave and mobility

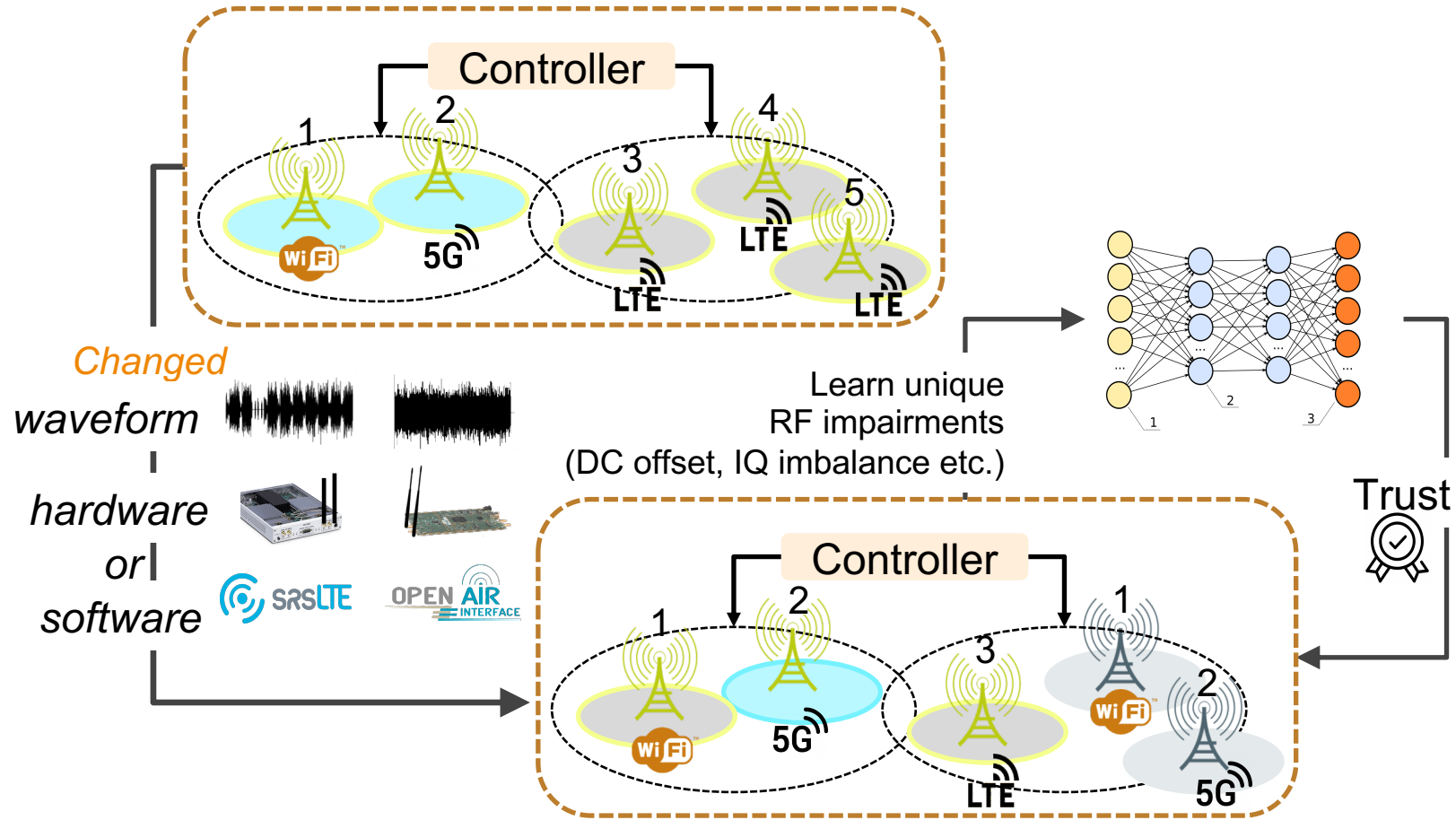
COLOSSEUM

Boston, MA



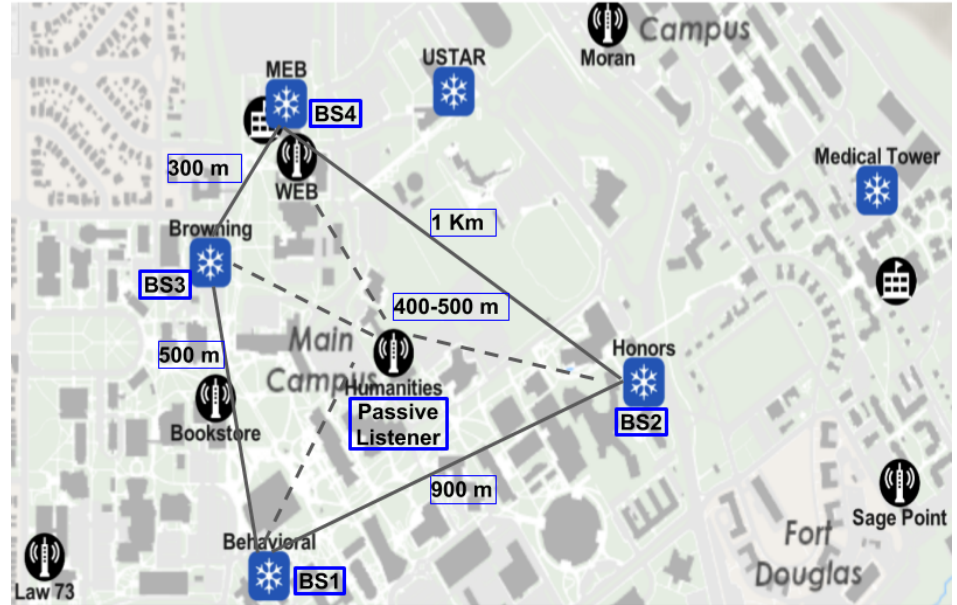
Spectrum sharing, mMIMO, mobility, IoT, mesh networks

Example Problem: RF Fingerprinting



Example Measurement and Dataset Generation

POWDER PAWR platform University of Utah campus, in Salt Lake City, spanning 6 km², composed of SDRs and open-source software stacks.



Example Measurements & Dataset



[HOME](#) [ABOUT US](#) [PEOPLE](#) [PUBLICATIONS](#) [OUTREACH](#) [DOWNLOADS](#) [CONTACT US](#)

[Home](#)

POWDER RF Fingerprinting Dataset



Datasets for RF Fingerprinting on the POWDER Platform

Download Datasets:

Please use below link to download the dataset:

Dataset#1: [Raw IQ samples of over-the-air transmissions from 4 Base Stations deployed in the POWDER platform, in Salt Lake City, Utah, USA.](#)

These datasets were used for the paper "Trust in 5G Open RANs through Machine Learning: RF Fingerprinting on the POWDER PAWR Platform", IEEE GLOBECOM 2020. Any use of this dataset, which results in an academic publication or other publication that includes a bibliography, should contain a citation to our paper. Here is the reference for the work:

Conference version: [PDF](#)

G. Reus-Muns, D. Jaisinghani, K. Sankhe and K. R. Chowdhury, "Trust in 5G Open RANs through Machine Learning: RF Fingerprinting on the POWDER PAWR Platform," IEEE Globecom, 7-11 December 2020, Taipei, Taiwan.

[Copy Bibtext](#)

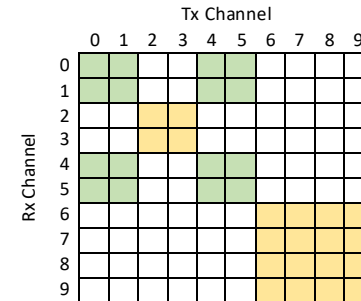
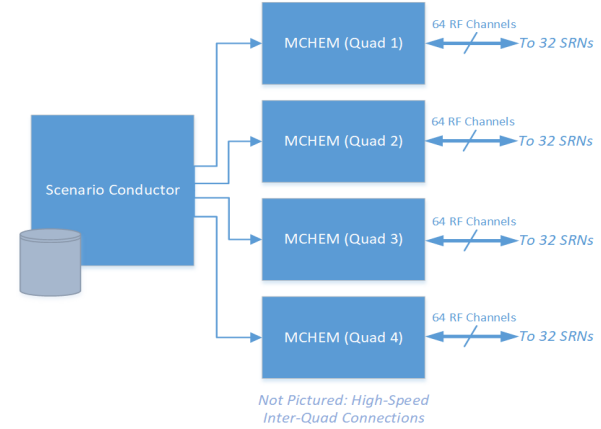
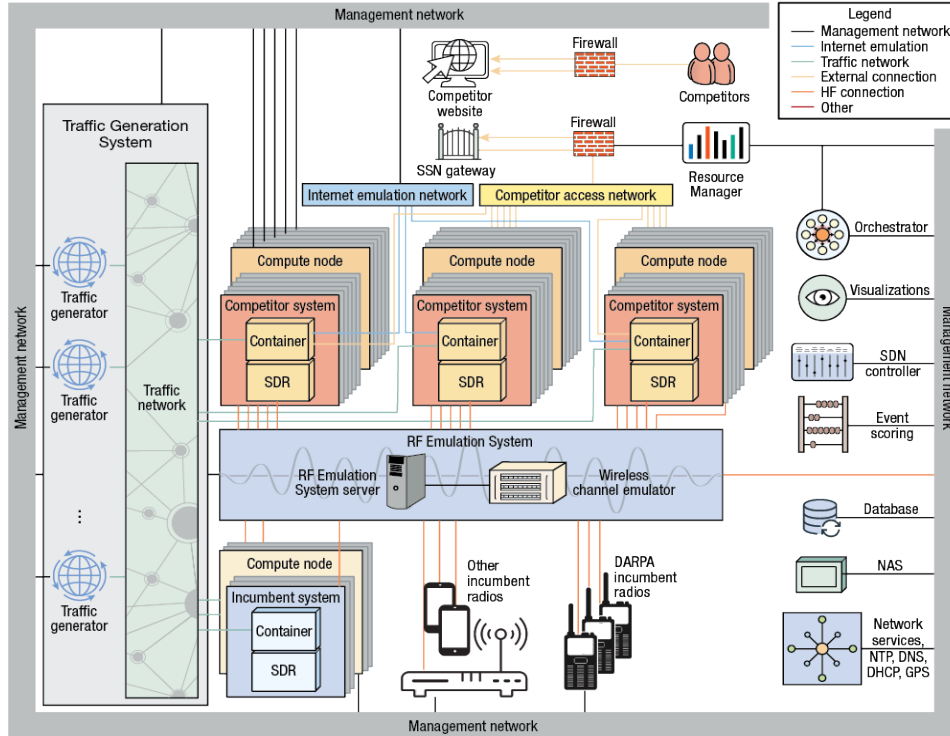
<https://genesys-lab.org/powder>

Colosseum



- 900 TB of Network Attached Storage (NAS)
- 171 high-performance servers
- 256 USRP X310s (128 as communications devices, 128 as part of the channel emulator)
- 18 10G switches
- 19 clock distribution systems
- 25.6 GHz total instantaneous bandwidth
- 52 TB/s of digital RF data
- 320 FPGAs
- Hundreds of high-speed optical connections
- Software-based traffic generation solutions hosted on a pool of dedicated servers
- Full-mesh networking capability
- 21 racks of radios, FPGAs, servers and support equipment

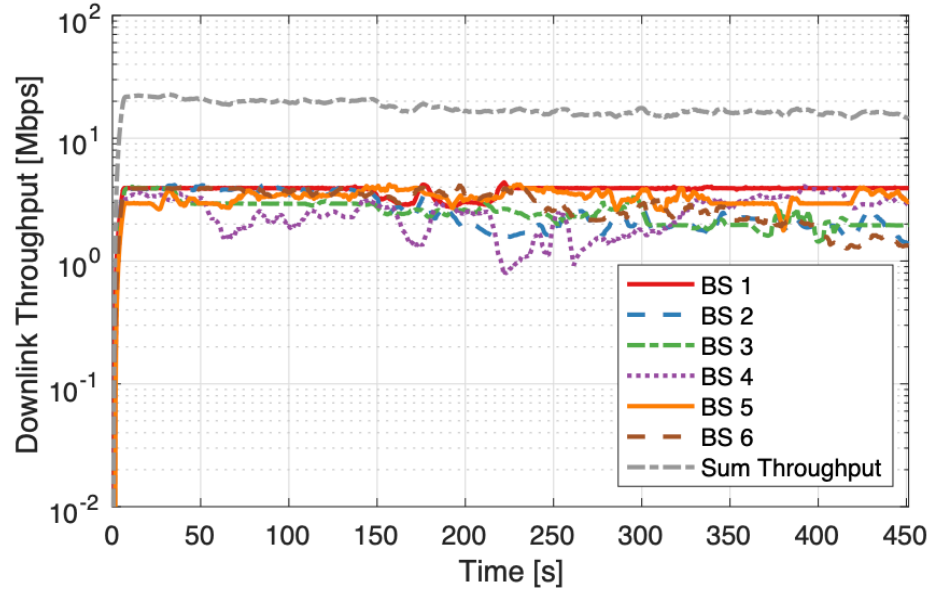
Example Experiment Design



- Legend
- Green: Reservation A - Channels 0,1,4,5
 - Yellow: Reservation B - Channels 2,3,6,7,8,9
 - White: Connectivity Disabled

Example Measurements & Dataset

- Cellular network w/ srsLTE: 6 interfering base stations w/ 24 users
- Downlink video streaming
- Pedestrian user mobility
- Real-world scenario with base station locations in Boston Public Garden



Downlink throughput

The screenshot shows a software interface titled "4G/5G Networks Software". It features a white text box overlaid on a map background. The text box contains the following information:

- srsLTE**
- 4G/5G Cellular Networking
- Image name: `srsLte-20-04`
- Credentials: `root/ChangeMe`
- Image repository: <https://github.com/colosseum-wiot/colosseum-srslte-20-04>

Below this, it states: "srsLTE is a free and open-source LTE software suite developed by SRS. It includes the following features:"

- srsUE - a complete SDR LTE UE application featuring all layers from PHY to IP
- srsENB - a complete SDR LTE eNodeB application
- srsEPC - a light-weight LTE core network implementation with MME, HSS and S/P-GW
- a highly modular set of common libraries for PHY, MAC, RLC, PDCP, RRC, NAS, STAP and GW layers.

At the bottom, it says: "See the srsLTE GitHub repository for more details."

Base station locations

Open Challenges: Technical

Create representative community dataset(s)

- MNIST for wireless
- One central stop for the community

Accelerate 6G research

- Real-time and offline dataset generation, processing tools

Influence design of future experimental testbeds

- Show feasibility of native support/APIs for community testbeds

Open Challenges: Education and Workforce Dev

Train “tomorrow’s” wireless engineer

- Train wireless/networking workforce on how to build, store, share datasets
- Impact broad curriculum changes

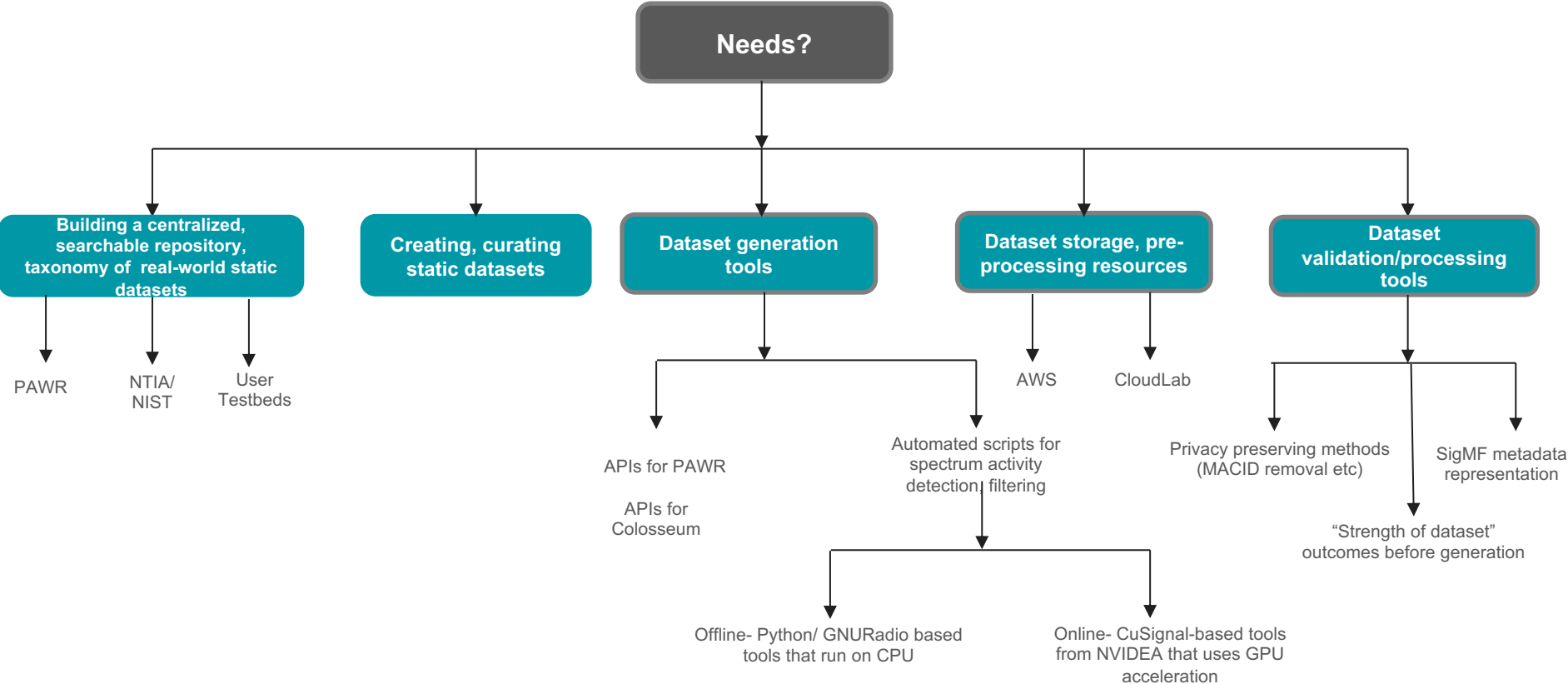
Bridge communities

- Provide resources, competitions for wireless engineers to build collaborations with ML/AI experts
- Engage with industry as they roll out new tools

Democratize access to wireless datasets

- Enable anyone to create datasets with minimal expert knowledge
- New hands on activities K-12 and beyond

Community Needs



Thank You