Collaboration - ITU

Coming Up: https://aiforgood.itu.int/event/itu-ai-ml-in-5g-grand-challenge-finale-2/

Collaboration Options

1. Hosts
    a. Propose a problem achieve better solution
2. Advisory for Problem(s)
    a. Relevance, ML-Technology, etc.
3. Evaluation Panel
    a. Solution evaluation.
4. Competitor
    a. Research-group/Researchers participating.
5. Webinars/Talks
    a. Relevant to the problems.
    b. Present or attend.
6. Engage in Programming Challenges
    a. BuildAThon
    b. Mentoring in this challenge.

Contact Vishnu Ram: Vishnu Ram OV <vishnu.n@ieee.org>

ITU-T Secretariat Contacts.

Problems, Hosts, Competitors and Datasets

Host-Based Categorization

Sl. no	Host	2020	2021
1	Federal University of Pará (UFPA), Brazil	beam-selection: Machine learning applied to the physical layer of millimeter-wave MIMO systems	Reinforcement learning: scheduling and resource allocation
2	UPF, Spain	Improving the capacity of IEEE 802.11 WLANs through machine learning	1.Finding groups of BSSs 2.Federated Learning for Spatial Reuse in WLANs
3	BNNUPC, Spain	Graph Neural Networking (GNN) Challenge 2020	GNN Challenge. 2021: Creating a Scalable Network Digital Twin
4	North Carolina State University, US	channel estimation: Machine learning applied to the physical layer of millimeter-wave MIMO systems	Localization: Multidevice localization with mmWave signals in a factory environment
5	NEC	Network state estimation by analysing raw video data	Location estimation using RSSI of wireless LAN
6	KDDI	Analysis on route information failure in IP core networks by NFV-based test environment	Network failure detection and root cause analysis in 5GC by NFV-based test environment
7	Turkcell	Using weather info for radio link failure (RLF) prediction	Radio Link Failure Prediction
8	St. Petersburg State Univ. of Telecom.	Traffic recognition and Long-term traffic forecasting based on AI algorithms and metadata	Forecasting Model for Service Allocation Network Using Traffic Recognition
9	China Unicom	1.Fault localization of loop network devices based on MEC platform 2.Energy-saving prediction of base station cells in mobile comm. network 3.Core network KPI index anomaly detection 4.Alarm and prevention for public health emergency based on telecom data 5.Configuration knowledge graph construction of loop network devices based on MEC architecture	Network anomaly detection based on logs
10	China Mobile	1.Network topology optimization 2.Out of service (OoS) alarm prediction of 4/5G network base station	Cross Layer user experience optimization – Radio link performance prediction
11	NIST		WALDO (Wireless Artificial intelligence Location DetectiOn): sensing using mmWave communications and ML

12	Xilinx		Hardware-Efficient Modulation Classification with RadioML
13	Univ. of Alabama		RF-Sensor Based Human Activity Recognition
14	ITU: FG-AN		Network resource allocation for emergency management based on Closed loop analysis
15	ZTE		Delivery route optimization

Problem-Domain based Categorization

The categories of the problems are as follows:
RAN: Radio Access N/W

PHY: Physical Layer

NFV: Network Functions Virtualization

WLAN: Wireless LAN(802.11)

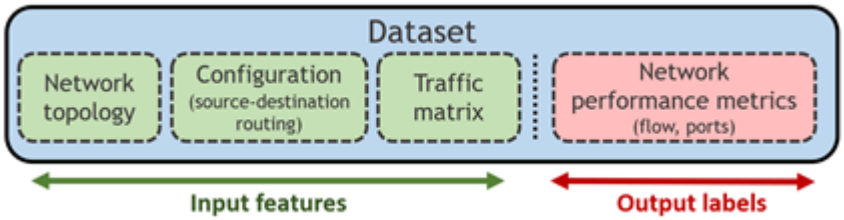
TE: Traffic Engineering

AR: Artificial Reality

Sl. No	Category of the Problems	Problems	Host	#of Competitors	Dataset
1	RAN	ITU-ML5G-PS-016: Location estimation using RSSI of wireless LAN	RISING Japan	38	The training data includes the location of AP, RSSI information within the coverage of the AP and its corresponding location. Link is not available.
2	NFV	Network failure detection and root cause analysis in 5GC by NFV-based test environment. Link	KDDI Japan	35	Did not find the public link. The data sets used for this challenge were created in the NFV-based test environment simulated for 5GC. In this sense, they are synthetic data, but as similar as the real data, resulting from our NFV-based test environment.
3		Federated Learning for Spatial Reuse in a multi-BSS (Basic Service Set) scenario Link	UPF	22	A dataset generated with the Komondor simulator is provided to train ML models. Link
4	RAN	Radio Link Failure Prediction Link	Turkcell	35	Zip file contains the following tab separated files (tsv): • **distances.tsv**: pair-wise distances • **met-forecast.tsv**: meteorology 5-day forecasts • **met-real.tsv**: meteorology historic realizations • **met-stations.tsv**: meteorology station information • **rl-kpis.tsv**: radio link KPIs and configuration parameters • **rl-sites.tsv**: radio link site information GitHub: Link Training Dataset: Link Exemplary Colab Project by Turkcell: Link
5		Build-a-thon (PoC) Network resource allocation for emergency management based on closed-loop analysis Link	ITU Focus Group on Autonomous - Networks	26	Powerpoint slides for the progress of teams during the challenge Link

6	TE	ITU-ML5G-PS-011: Combinatorial Optimization Challenge:- Delivery route optimization Link	ZTE	19	The dataset is provided in the Dataset module.
					Each sample in the dataset is composed of 2 elements:
					
					1. A json object that describes the graph of the transportation network, see the below example:
					{
					nodes:[0,1,2] ,# id list of the transportation nodes
					edges:[
					(0,1):{ # (starting point, end point)
					lanes: [0,1,2], # id list of the lanes, the length of the array is the number of the lanes between the two adjacent nodes
					lane_weights:[10, 10, 20] # The weight that each lane can carry
					}
					]
					}
					2. A json object that describes the starting point, endpoint, weight information of each delivery, like:[{“a_end”: 138, “z_end”: 206, “weight”: 1}, {“a_end”: 74, “z_end”: 156, “weight”: 1}]

7	TE	Forecasting Model for Service Allocation Network Using Traffic Recognition Link	The Bonch-Bruevich Saint Petersburg State University of Telecommunications	16	The training data in pcap format can be downloaded here.

8	PHY	RF-Sensor Based Human Activity Recognition Link	The University of Alabama	24	Github with Data Download: https://github.com/ci4r/CI4R-Activity-Recognition-datasets Watch the webinar here: https://aiforgood.itu.int/events/understanding-how-people-move-using-modern-civilian-radar/
9		ML5G-PHY-Localization: Multidevice localization with mmWave signals in a factory environment Link	North Carolina State University	17	The training data set consists of a collection of channels associated to the links between devices and access points in a factory environment and also their associated positions in the room. The factory environment has been simulated by ray tracing, with 12 access points (AP) located at the ceiling. The devices are randomly placed at the room and assigned to the AP with highest gain during the sector level sweep. For simplification, we assume perfect synchronization, this means that the first path arrives to the AP at time 0. The beam-training is measured. Then, with probabilities 50%, 25% and 25%, 0, 1 or 2 other devices interfere with the beam-training through random 4-QAM OFDM symbols using 512 frequency carriers. Two datasets are provided, namely “train” and “test”. The dataset “train” consists on all the information required for the training of your algorithm, this is UE location and the AP it is associated to, while the dataset “test” lacks the UE localization. Each participant can get access to each of the 12000 samples measurement through the provided executable file by passing as flags the name of the dataset and the indices [0-11999] of the samples to generate. Example: “gen_channel.exe train 0 1 2” generates the first 3 samples of the “train” dataset. We also provide Python code to more easily understand the dataset through visualization. The provided Python code generates the measurements regarding the sample indexes 0, 4 and 8, which happen to have 0, 1 and 2 interferers and then prints the information regarding these samples interferer and plots the power spectrums in time, direction of departure and direction of arrival. These files can either be downloaded from the dataset tab once the team is registered or from the NCSU webpage of the problem.
10	AR	ML5G-PHY-Reinforcement learning: scheduling and resource allocation Link	UFPA	37	The dataset is composed of “traces” from simulations executed with Unreal Engine + AirSim, and will be available on June 30 at this folder
11	WLAN	WALDO (Wireless Artificial intelligence Location DetectiOn): sensing using mmWave communications and ML. Link	National Institute of Standards and Technology	31	For software and dataset resources, please refer to https://github.com/usnistgov/PS-002-WALDO. The challenge dataset relies on two open-source software: 1. IEEE 802.11ay packet generator available at: https://github.com/usnistgov/PS-002-WALDO. 2. NIST Q-D channel realization software: https://github.com/wigig-tools/qd-realization used to generate the millimeter wireless channel. **Training Dataset** The training datasets consist of a collection of received packets at different SNR. The IEEE 802.11ay preamble is filtered through the NIST Q-D channel realizations, which model the millimeter wave wireless signal propagation in a room with multiple people moving.
12	RAN	Lightning-Fast Modulation Classification with Hardware-Efficient Neural Networks Link	Xilinx	68	This challenge uses the RadioML 2018.01A dataset provided by DeepSig. You will be required to provide your contact information to download the dataset. Please note that this dataset is released by DeepSig under the Creative Commons Attribution - NonCommercial - ShareAlike 4.0 License a nd you must abide by the terms of this license to participate in our challenge.\ **Train/test split** Please note that we use a fixed train/test split on the dataset for a fair comparison, and you must abide by this in order to provide a valid submission. This split is given by the fixed random seed and other parameters in the Jupyter notebook we provide in the sandbox.
13		Network anomaly detection based on logs Link	China Unicom	45	The data is the log data of network equipment, including time stamp, log information, etc. The log is in text format. The data is divided into training set and test set. The training set contains tag information, that is, with exception and without exception; it is a Boolean tag representing a true or false result. The test set comes from the same network equipment as the training set. The test set does not contain tag information. After modelling, the contestants predict the test set and get its tag. The data could not be found.
14		Graph Neural Networking Challenge 2021 - Creating a Scalable Network Digital Twin Link	Barcelona Neural Networking Center	47	We provide a dataset generated with the OMNet++ network simulator, which is a discrete event packet-level network simulator. The dataset contains samples simulated in several topologies and includes hundreds of routing configurations and traffic matrices. Each sample is labelled with network performance metrics obtained by the simulator: per-flow performance statistics (mean per-packet delay, jitter and loss), and port statistics (e.g. queue utilization, size).  Data is divided in three different sets for training, validation and test. As the challenge is focused on scalability, the validation dataset contains samples of networks considerably larger (51-300 nodes) than those of the training dataset (25-50 nodes). Likewise, the test dataset will be released at the end of the challenge, just before the evaluation phase starts, and it will contain samples following the same distribution as in the validation dataset. Please, find a detailed description of the datasets and the links to download them here: https://bnn.upc.edu/challenge/gnnet2021/dataset

Thoth as a Competitor

1. NFV Related Problems
2. Start with 2021 Problems.
    a. KDDI
    b. China Unicom*

Thoth/Anuket/LFN as a Host - ?

LFN.

Tentative Problems :

1. Loss Characterization
2. Synthetic Observability (logs and metrics) Data Generation GANs.