

Machine Learning for Mobile Communication Systems

Problem for the Final Project

Statement

The dataset D stored in the file at this [link](#) reports the activity of one LTE cell during approximately 1 month of operation.

The dataset contains the following information:

- *overalltime*: in seconds, starting from the initial date of the measurements (29/06/2016)
- *overallusers*: number of transmitting users
- *overallrbdw*: total number of resource blocks in downlink
- *overallrbdwmean*: mean number of resource blocks in downlink
- *overallrbdwstd*: standard deviation of resource blocks in downlink
- *overallratedw*: rate (bit/s) in downlink
- *overallratedwmean*: mean rate (bit/s) in downlink
- *overallratedwstd*: standard deviation of rate in downlink
- *overallmsgdw*: number of messages in downlink
- *overallretxdw*: number of retransmitted packets in downlink
- *overallrbup*: total number of resource blocks in uplink
- *overallrbupmean*: mean number of resource blocks in uplink
- *overallrbupstd*: standard deviation of resource blocks in uplink
- *overallrateup*: rate (bit/s) in uplink
- *overallrateupmean*: mean rate (bit/s) in uplink
- *overallrateupstd*: standard deviation rate (bit/s) in downlink
- *overallmsgup*: number of messages in uplink
- *overallretxup*: number of retransmitted packets in uplink

Problem: Given D , build one model based on reinforcement learning able to automatically identify anomalous points due to flash crowds. Select the most appropriate state variables, actions and reward to get the highest performance in term of accuracy and F-score. The cumulative reward of the proposed model over the number of episodes must be shown to discuss algorithm convergence.

A-priori information: every Sunday from 9am to 16pm, a popular flea market took place in the monitored zone attracting a high number (anomalous) number of people. This portion of the dataset may be used as ground-truth for the anomalous points.

Instructions to deliver your project

What: a written report (paper style, two columns, pdf) including problem statement, proposed solution/method, achieved results and discussion. The generated code should be also delivered.

How: by email

When: 1 July 2022

Projects should be individual. A group (max 2 people) may be accepted in special cases, upon request.

Project delivery, questions and comments: paolo.dini@cttc.es