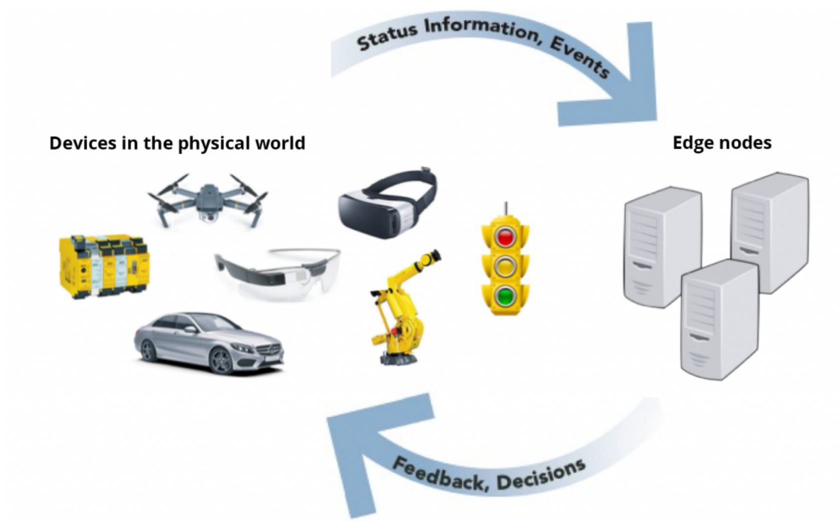
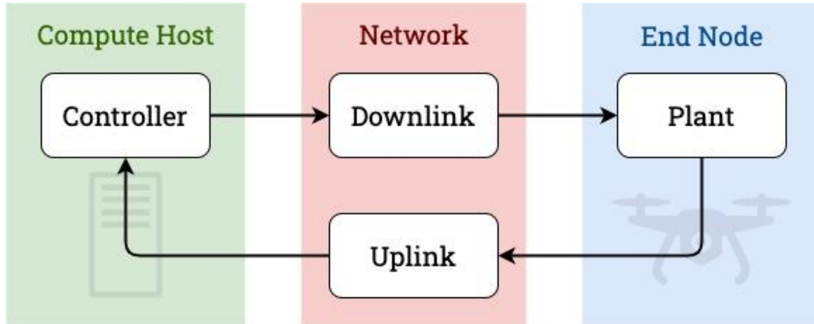
Predicting 5G Network Responsiveness with Deep Learning

Samie Mostafavi
ssmos@kth.se



Latency-critical applications are emerging more than ever with the advent of fast and accessible computing infrastructures. Examples of latency-critical applications include cyber-physical systems such as remote-controlled robotics systems and human-in-the-loop applications such as cloud gaming, augmented reality, and virtual reality. In addition to the bandwidth requirement in these applications, the end nodes require a timely response from the server for smooth operation. There is an increasing need for these applications to be able to run wirelessly. Due to the unpredictable nature of the wireless medium, a wireless link (e.g., 5G) cannot limit the communication delay. Shadowing, fading, and interference are stochastic phenomenons that cause transient high error rates, hence higher latencies.

To maintain the desired quality of service, either the wireless network must be tuned to guarantee a certain delay level, or the application must adapt according to the delay conditions. In both cases, a model that can predict the delay of the network is needed. For instance, using that model, a remote-controlled robot can avoid high latency areas (probably because of poor radio coverage) in path planning.



The end-to-end latency measurements could express the responsiveness of the communication network. End-to-end latency denotes the time duration a packet travels from the end node to the server via uplink and back to the end node through the downlink. We aim to model this notion as a random variable and compute its probability density conditioned on related factors, such as signal strength indicators or the terminal's location.
Analytical modeling is shown to be intractable in finding the relation between the end-to-end delay probability and the related factors. Therefore, data-driven methods must be utilized. In this project, a deep learning-based probability prediction scheme will be devised for predicting the responsiveness of a software-defined 5G network from the related parameters, such as the location or signal strength indicators.

Project Goals:

1. A survey on networked systems end-to-end delay prediction works.
2. Propose an approach to predict the delay density prediction from the network state.
3. Implement and validate the proposed scheme on the software-defined 5G network.

Implementation Tasks:

1. Run [Openairinterface5G] network on the ExPECA testbed's software defined radios.
2. Develop a software that collects the end-to-end delays to form the dataset for training the machine learning model.
3. Implement the machine learning application based on your proposed approach that can predict the end-to-end delay probabilities from the network state.
4. Collect end-to-end delay measurements and train the model.
5. Evaluate the model.

Required Skills:

Required:
Linux, Docker, Python

References and Background Material:

[1] Mohammed S. Elbamby, Cristina Perfecto, Chen-Feng Liu, Jihong Park, Sumudu Samarakoon, Xianfu Chen, and Mehdi Bennis. 2019. Wireless Edge Computing With Latency and Reliability Guarantees. *Proc. IEEE* 107, 8 (August 2019), 1717–1737. DOI:https://doi.org/10.1109/JPROC.2019.2917084

[2] Majid Raeis, Ali Tizghadam, and Alberto Leon-Garcia. 2020. Probabilistic Bounds on the End-to-End Delay of Service Function Chains using Deep MDN. *arXiv:2006.16368 [cs]* (June 2020). Retrieved September 13, 2021 from http://arxiv.org/abs/2006.16368

[3] Jonas Rothfuss, Fabio Ferreira, Simon Boehm, Simon Walther, Maxim Ulrich, Tamim Asfour, and Andreas Krause. 2020. Noise Regularization for Conditional Density Estimation. *arXiv:1907.08982 [cs, stat]* (February 2020). Retrieved October 3, 2021 from http://arxiv.org/abs/1907.08982

[4] https://expeca.proj.kth.se/