

Unsupervised Flow Characterization for Real-Time Network Digital Twins

Master's Thesis

Project

The development of Network Digital Twins for future wireless communication systems (like 6G Radio Access Networks) requires precise modeling and understanding of network traffic. However, modern network analysis faces major challenges: traditional methods like Deep Packet Inspection (DPI) and port-based classification fail due to the widespread adoption of heavy encryption (e.g., Layer 4 upwards, VPNs) and dynamic port allocations.

In this thesis, you will work with real packet-level measurement data (PCAP) to characterize network flows purely by their statistical properties (such as packet sizes, inter-arrival times, and flow duration). You will implement unsupervised machine learning algorithms (clustering) to automatically group similar traffic behaviors without relying on labeled application data. A key focus of the thesis will be evaluating these models not just for clustering accuracy, but for computational complexity, ensuring the chosen approach is lightweight enough to run in real-time on a digital twin base station.

Tasks

1. Work with a large real-world packet-level dataset to extract statistical flow features.
2. Implement baseline and advanced unsupervised clustering approaches (e.g., k-means, Autoencoders, etc.) for flow characterization.
3. Compare different approaches with respect to clustering quality, adaptability to noise, and computational complexity (real-time capability).

Requirements

- ✓ Basic programming knowledge in Python (data processing and ML libraries)
- ✓ Basic knowledge of machine learning concepts (especially unsupervised learning)
- ✓ Interest in computer networks and statistical traffic analysis
- ✓ Interest in addressing problems associated with real-world data

Institute

Communications
Engineering
Lab

Hertzstr. 16
Gebäude 06.45
76187 Karlsruhe
www.cel.kit.edu

Contact

M.Sc.
Enes Köktas

Room 209
enes.koektas@kit.edu