Throughput Prediction in Cellular Networks

Darijo Raca

Supervisors: Prof. Cormac Sreenan and Dr. Ahmed Zahran

Introduction

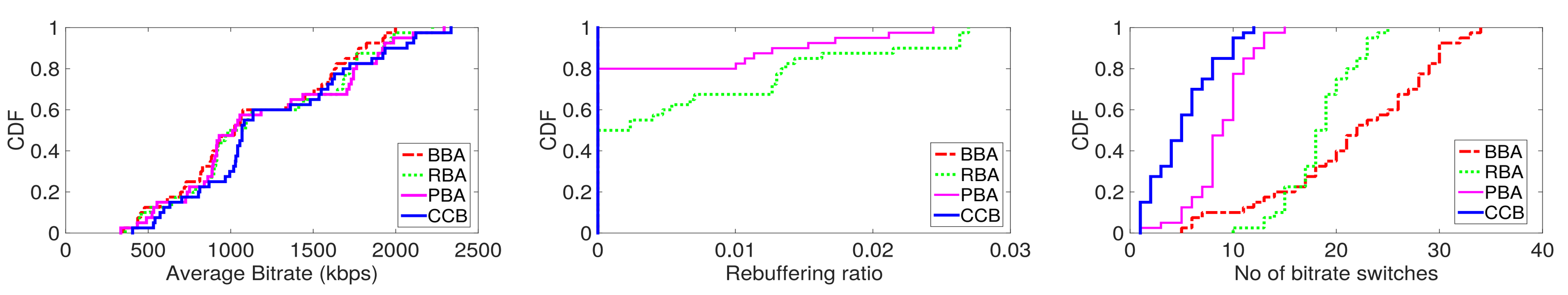
- Mobile data traffic grows rapidly (63% increase in 2016),
- Mobile video has the largest share in total mobile traffic (currently 60%, expected to rise to 78% by the 2021)
- Majority of video content providers use adaptive-bit-rate (ABR) streaming techniques for delivering content to the end users

Motivation

- State-of-art ABR algorithms use different approaches for estimating available bandwidth: historical throughput, buffer occupancy, and traffic shaping
- There is significant discrepancy between inferred and actual bandwidth in cellular networks

Recent studies show possibility of accurate throughput prediction on small and medium time scales. Some of adaptation algorithms leverage possibility of having a semi-accurate throughput prediction:

- PBA shows that overall user experience can be increased if the algorithm has a hint about future throughput values
- Similarly, the CrystalBall algorithm takes prediction errors into account when deciding for the next chunk quality



Challenges

- Data rate in cellular networks can fluctuate by an order of magnitude over a span of a few seconds
- These fluctuations can be a consequence of various factors: rapid changes in the underlying radio channel conditions and system load (devices leaving and entering the network); burstiness caused by frame scheduling algorithms; interaction between feedback loops at the different time scales (radio retransmission, TCP)

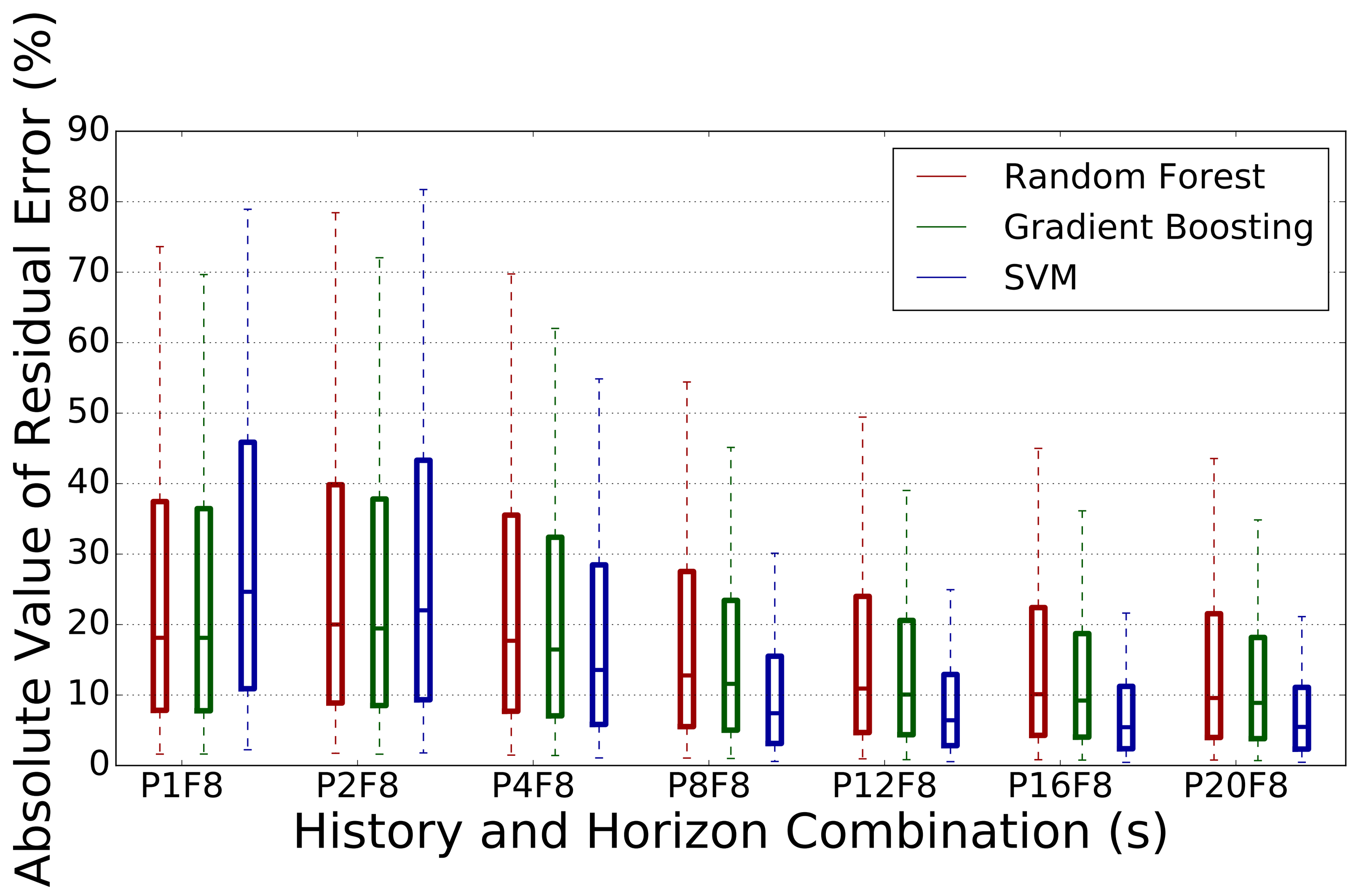
Methodology

- Conducting experiments in a real LTE cellular network (major US mobile operator) - continuous download of a large/small file
- Different mobility models: static, pedestrian, driving
- Collecting device based KPIs (e.g., CQI, RSRP, RSRQ, SINR) and analogous network based KPIs (e.g., cell load, the average throughput of users)
- Leveraging different machine learning algorithms for throughput prediction
- Feature engineering (how to summarize KPIs)
- Analysing impact of the device and/or network based KPIs on throughput prediction accuracy
- **Evaluation metrics:** absolute value of residual error and coefficient of determination

Preliminary Results

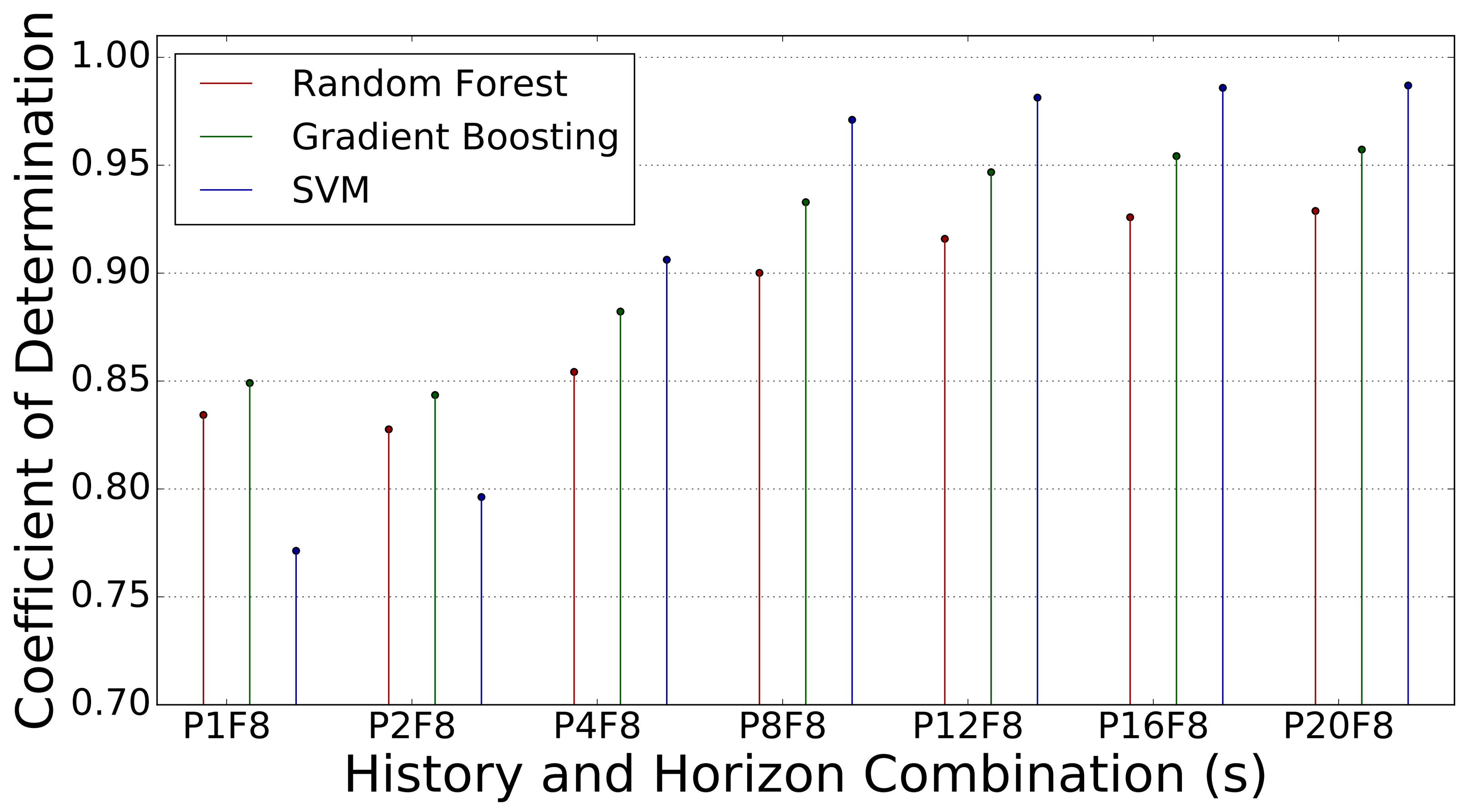
Following results are based only on device KPIs. For the future work, we plan to incorporate information obtained from the network (i.e., from the cell)

Boxplot for Residual Errors



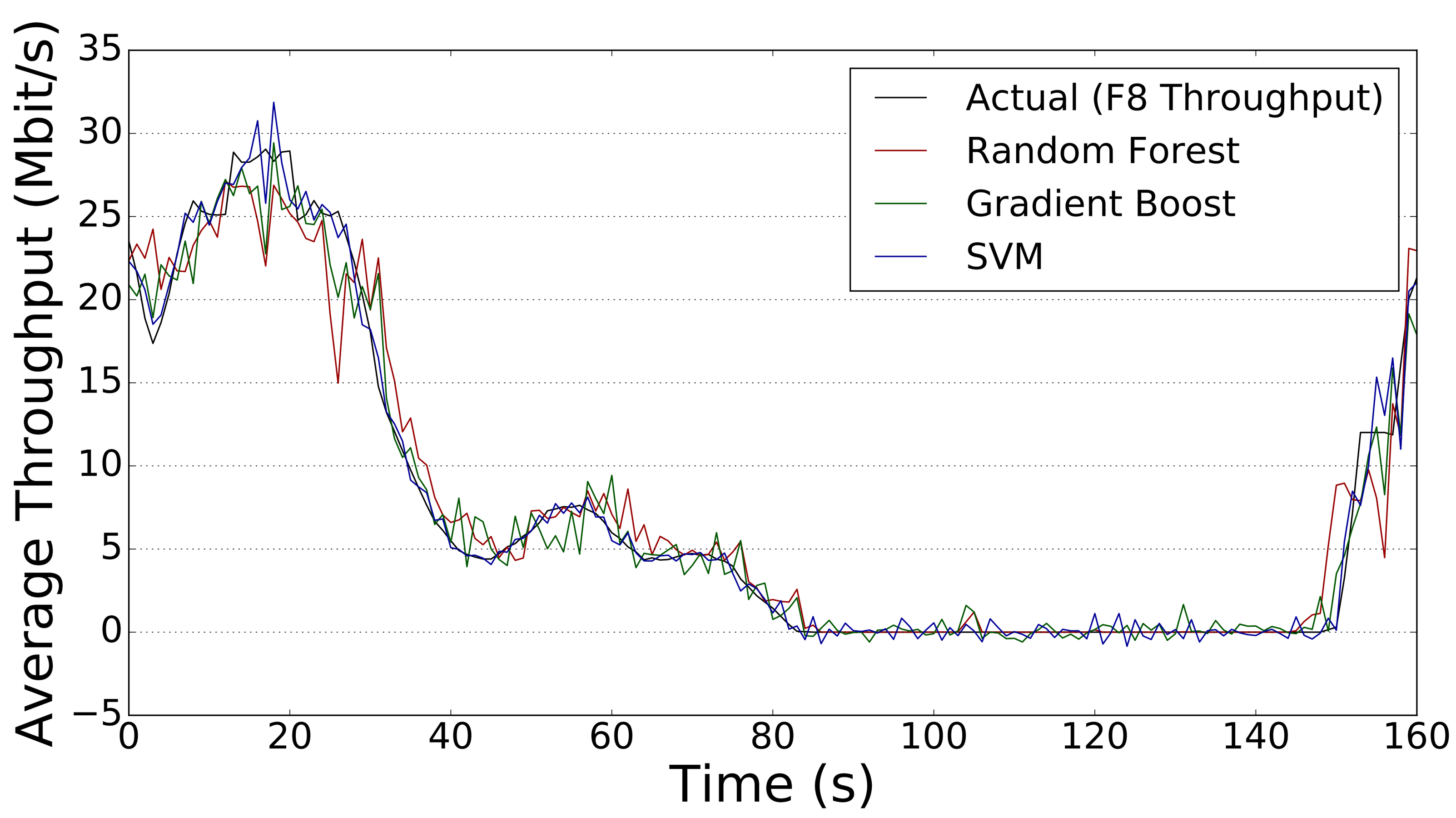
- **Increasing history information helps in lowering overall error**
- **90th percentile of the absolute residual error for the eight second prediction window is less than 22%**

Comparison of CoD



- **With larger history information all ML algorithms achieve high R^2 value ($>>$ 0.9). For larger history, 98% of the variability observed in throughput values can be explained by regression model based on SVM**

Time Series Plot



- **Regression model with SVM follows actual throughput more closely than other two ML schemes**