Al-Najah National University
Computer Network and Information Security
Network Design

Analyzing Throughput and Stability in Cellular Networks

Instructor:Dr. Amjad Hwash

Dima Bshara
Mohammed Adnan

1 Objective

The purpose of this paper is to understand and identify the prime factors of cellular network performance that affect end-user experience. So they use a dataset collected using net radar, a platform that measures cellular network performance crowdsourced from mobile end devices. Using this dataset they will develop a methodology (a classifier using a machine learning approach) for understanding cellular network performance.

Also, they examine the main characteristics of cellular networks related to throughput from the perspective of mobile user activity, MNO, smartphone models, link stability, location, and time of day. They also perform a network-wide correlation and statistical analysis to obtain a basic understanding of the influence of individual factors.

2 Terms and Activities

2.1 Data collection

The essence of this paper is to investigate factors the throughput and stability of cellular networks based on a longitudinal dataset collected using the (Netradar) measurement platform, which is easy to gather from mobile devices operating systems. After collection data, they reached:

To confirm that the time of a day and the location **affect the throughput**: in cities, the throughput drops during peak hours as congestion increases. They also observe that switches happen from legacy radio technologies to more advanced ones during the day. Also, they show that it is possible to **predict** the probability of sudden drops in bit rate in a cellular network with 90accuracy and 78% of kappa value, just relying on simply accessible information such as device model, location, network technology type, time of the day, and latency. Then they trained a model that can predict the average TCP download speed based on the first 5 seconds of the median bit rate value of throughput.

2.2 Performance Metrics.

Usually, it assumes that newly released devices come with better performance than the old ones. After that, they used a device's year of release to confirm this hypothesis for throughput. So, they discovered that our belief was wrong because after they examined how much the release of the new device model improved the performance of TCP downlink speed, they found that some old devices have better downlink speed than the most recent ones. For example, model SM-G388F, released in 2015, has a median downlink speed of about 14 MBps, while the older model GT-I9305, released in 2012, recorded a median downlink speed of 22 MBps. They also notice that there high range of variation between different device brands accessing the same radio technology. Also, they observe that even for the same brand, it is not always the newest device model whose TCP download performance is best.

2.3 Statistical Methodology

They next investigate the impact of location and time of day on cellular network performance as shown in the geographical distribution of throughput values for three MNO's in Finland. It shows that there is a high variation in throughput by MNO for different locations in LTE networks. In the Uusimaa region, DNA has the highest median throughput of 28.31 Mbps, compared to Elisa and TeliaSonera that have 21.28 and 19.6 Mbps, respectively. In other areas such as Pirkanmaa (Tampere), Elisa and TeliaSonera show 12.5% and 8,4% median throughput improvement over DNA, respectively.

2.4 Stability and its modeling.

Stability:
They classified TCP throughput stability based on sudden dropouts only using device model, location, network technology type, time of the day, and latency. It allows inferring network stability using the "minimal information" (e.x without measuring download and upload rates), which turns reduces measurement overhead. They also trained a model that can predict the average TCP throughput based on the throughput of the first five seconds (mostly TCP slow-start phase).

Modeling: They evaluate three prediction models. Namely, bagging-based classification, random forest, and conditional inference tree using different features(TCP downlink speed, latency, radio technology, time of the day, location, operator network, signal strength, base station information, OS platform, device information, and battery level). They also consider the technology switch happening during the measurement session. So, they take the radio technology that accesses at the beginning and the end of the measurement into account. Finally, they consider how many times does the radio technology change during the test.

3 design phase characteristics

1. Data collection is the first phase (Analyze requirement) because it analysis and collect the data from different resource
2. performance metric is in Design phase
3. Statistical Methodology is in Arch
4. Stability Arch
5. Modeling Design

4 Conclusion

In the end, We conclude that the throughput observed by a cellular network user depends on various factors such as device model, location, and time of the day where metropolitan areas during peak hours showed more drops in throughput. By examining the data received by the client every 50 ms, they find that about 30% of the total measurements experience sudden drops to zero bit rate for at least 200 ms. TCP is used to deliver data reliably, and its performance is sensitive to losses, latency, and jitter. So that a sudden dropout could create noticeable performance degradations, and real-time applications using other transports would be affected. They also trained a model that can predict the average TCP throughput based on the throughput of the first five seconds (mostly TCP slow-start phase).