An Empirical Study of Mobile Network Behavior and Application Performance in the Wild

Shiwei Zhang[1,2] Weichao Li[1,2] Daoyuan Wu[3] Bo Jin[1,2]
Rocky Chang[4] Debin Gao[3] Yi Wang[1,2] Ricky Mok[5]

[1]Southern University of Science and Technology, China
[2]Peng Cheng Laboratory, China
[3]Singapore Management University, Singapore
[4]The Hong Kong Polytechnic University, Hong Kong
[5]CAIDA / University of California, San Diego, USA

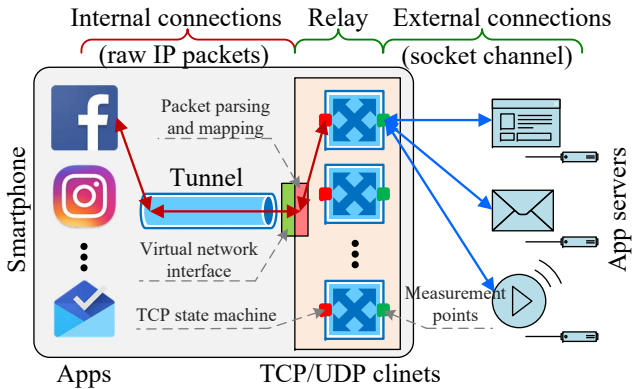
ACM/IEEE IWQoS, June 24, 2019

Overview

- Analyze a **two-year long** dataset obtained by a **mobile crowdsourcing** app.
- Characterize the **performance** of different protocols, DNS deployments, IP anycast, etc. **in the wild**.
- Propose a **performance degradation detection** method based on Apriori algorithm, tailored for **imbalaced** and **sparse** datasets.

Data Collection*

- VPN-based
  - Real user traffic
  - No “root” is needed
- Crowdsourcing
- Per-app measurement



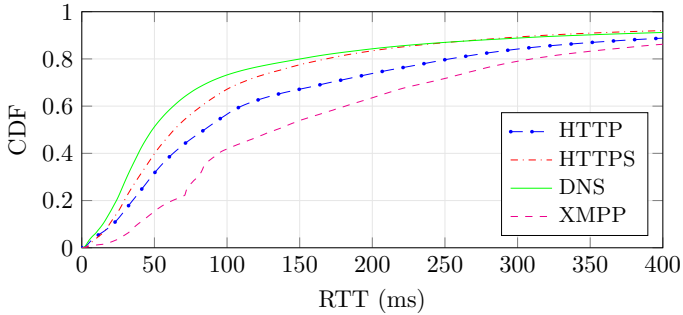
*MopEye: Opportunistic Monitoring of Per-app Mobile Network Performance, USENIX ATC'17

Dataset

- 20M records from 11k users in 173 countries.
- Interesting findings:
  - only **5.94%** of WiFi measurements were observed to have >300Mbps PHY rates.
  - more than **one third** of the 653 ISPs have **no 4G** measurements observed, mainly in Africa and Asia.

Protocols

XMPP (mainly used by IM and VoIP) traffics experience longer latency than HTTP(s).



DNS Redirection and IP Anycast* performance

- Using DNS servers located in a different country could increase the median latency by about 50%.
- Domains that already deployed DNS redirections could still benefit from IP Anycast.

*Identified using iGreedy: https://anycast.telecom-paristech.fr/dataset/

Advertisement Servers*

- For some apps more than 50% of traffic goes to advertisement providers.
- However, these traffic often has longer RTTs than the apps' own API servers.
- Companies with decent CDN deployments could improve the loading time by caching ads themselves.

*Identified using EasyList: https://easylist.to/

Why yet another performance degradation detection algorithm

- **Imbalanced**: For example, 83.5% of the 16,868 HSPAP measurements for ISP Mobilis are from one user. If those measurements are excluded, the median RTT can decrease from 332ms to 219ms.
  - normal association rules method bias to the performance of the dominating user.
- **Sparse**: Although the total number of observations is huge, records for each combination of features can be very small.
  - impossible to model the normal performance for all combinations of features separately.
- **Large**: We need a scalable method to process the increasingly large data.

Our Method

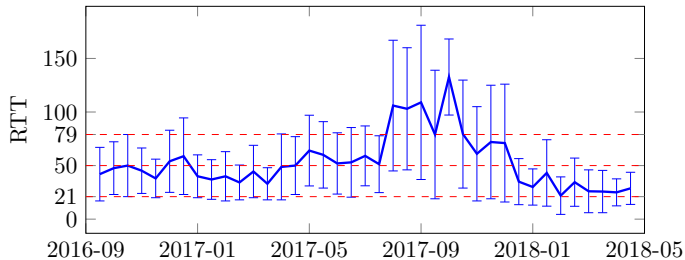
1. Find feature sets with enough records using the Apriori algorithm.
2. Filter the candidate feature set using using some huristic rules.
3. Identify performance degradation events by comparing the meadian RTT of similar feature sets.
4. Use hypothesis test to verify influence of the anormaly feature we found.

Evaluation

1. Low false positive rate in random data
   - We randomly shuffle the RTT of the records.
   - We mathmatically proved that the probability of our methods flagging an anomaly is very small in our configuration.

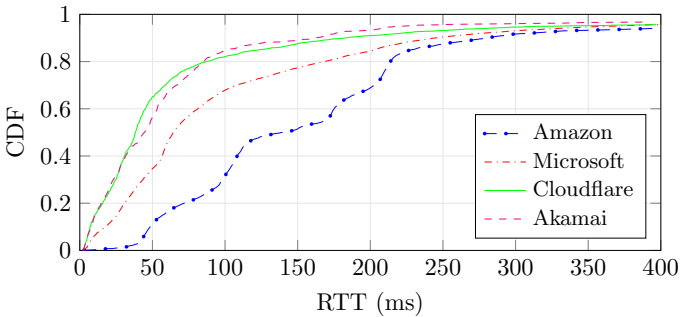
Evaluation

2. Real world case of Google Germany



Evaluation

3. Real world case of Microsoft Office Mobile



Conclusion

- Though IEEE 802.11ac equipment has become the mainstream in the market, only a small portion (6%) of Wifi connections exceed PHY rates of 300Mbps.
- More than one third of ISPs we observed do not deploy 4G networks.
- Many users use external DNS resolvers. IP Anycast may improve the mobile app performance in this case.
- XMPP traffic experiences longer RTT than HTTPS, which suggests that IM and VoIP services still have room for improvement.
- Advertisements servers often have longer latency than application servers.

Future Works

- 5G deployment and performance
- Actively measure the server when unexpected high RTTs are observed.

Thank you!

Shiwei Zhang (zhangsw@mail.sustech.edu.cn)