

Who slowed down my app?

A Measurement Study of Response Times of Android Apps in India

BTP PRESENTATION (SEMESTER-2)

8 CREDITS

Abhivandan Pandey(2017007)
Prince Sachdeva(2017080)

Advisors
Dr. Mukulika Maity
Dr. Arani Bhattacharya

Motivation

App Quality of Experience

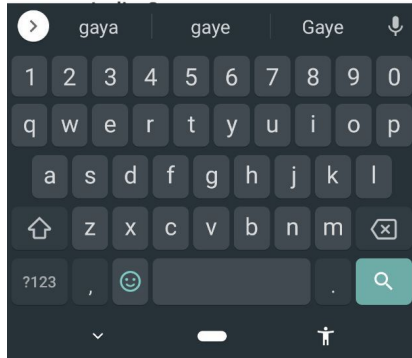
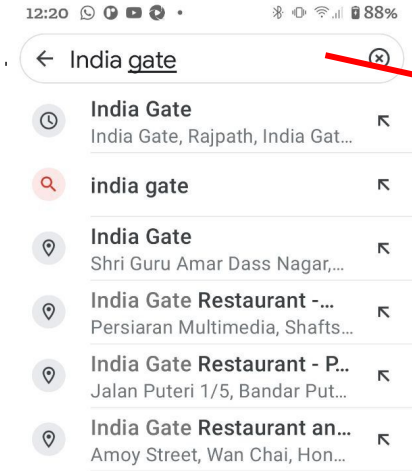
- QoE refers to the metrics end users use to judge the quality of services they receive
- Users would like apps to respond quickly to their requests, consume less mobile data and consume less energy to ensure sufficient battery life



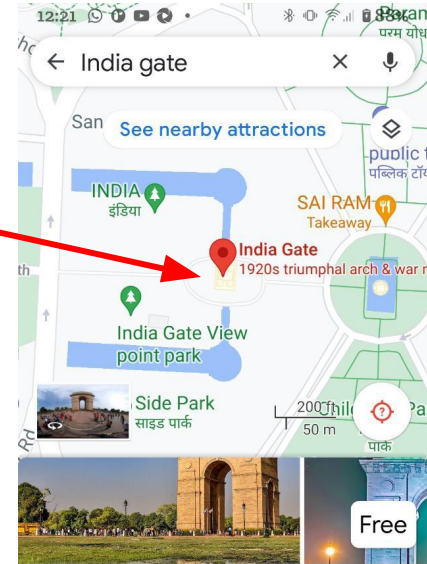
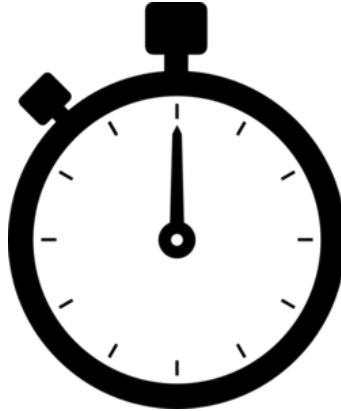
App Response Time

- App response time is defined as the time needed to reflect UI changes corresponding to a user's action.
- Amongst all the other QoE metrics, the app response time is the most direct way for end users to judge app performance.
- Response time of an app is influenced by several factors like phone hardware, network conditions, type of network.

Searching a location



Red Balloon
appears



India Gate

4.6 ★★★★★ (202,033)

Monument

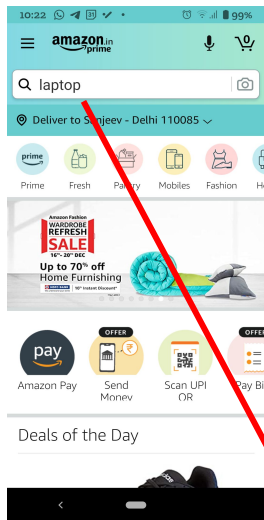
Temporarily closed

△ Hours or services may differ

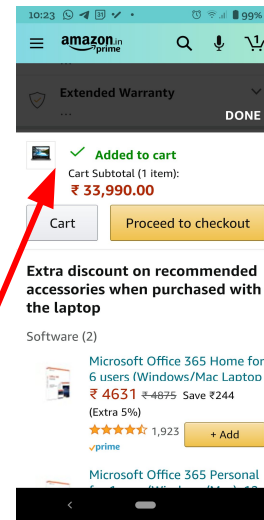
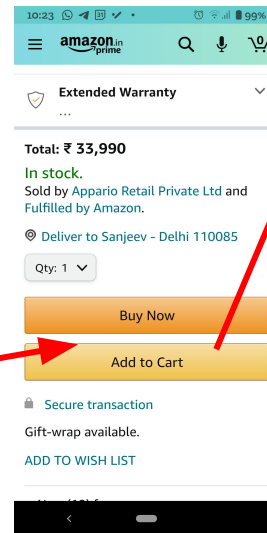
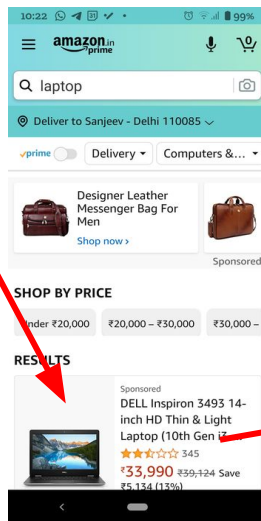
Directions

Call

Directory



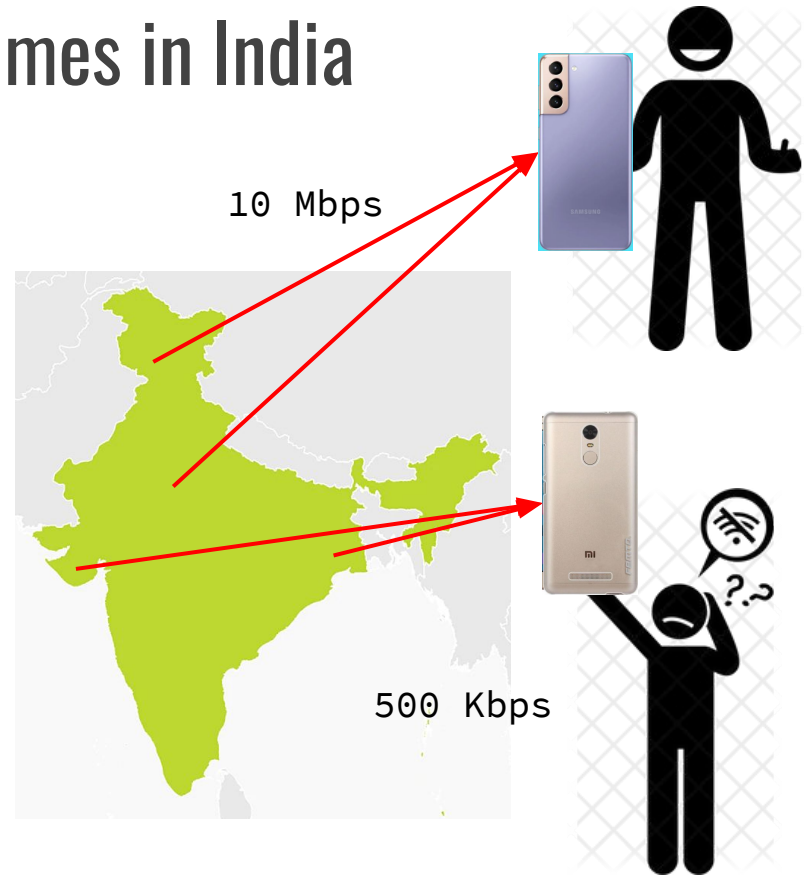
Searching
a product



Adding to
cart

Need for Study of App Response Times in India

- Prior studies in US have shown that smartphone hardware affects app response time the most
- India has non-uniform network coverage, with many areas still lacking adequate internet facilities
- Also there is a huge divergence in the smartphone used by the Indian population



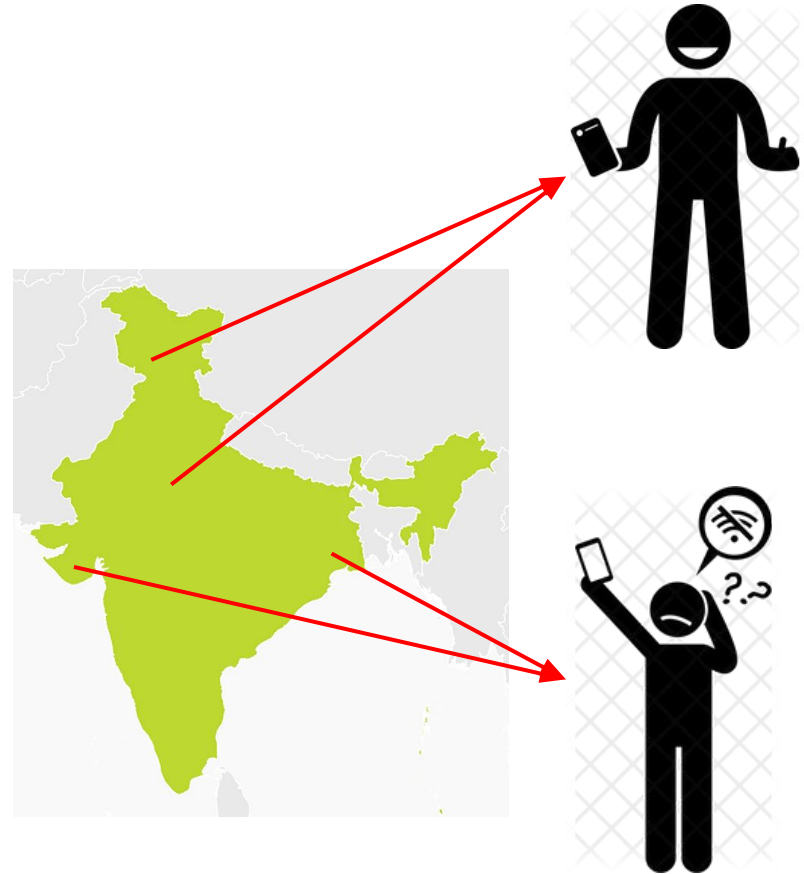
Need for Study of App Response Times in India

— — —

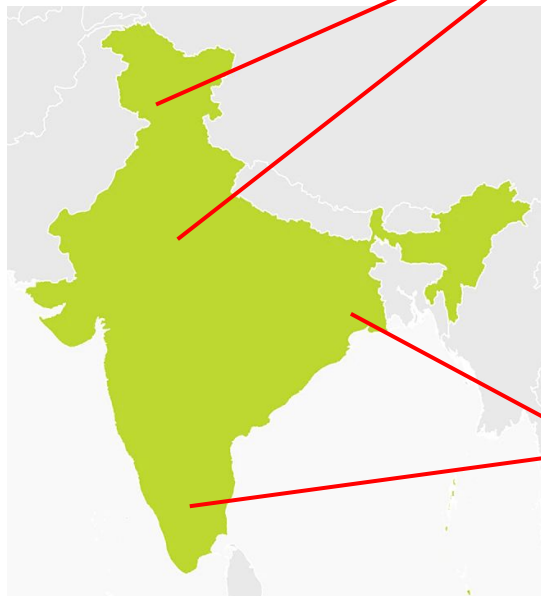
- So there is a need for a study to know which factors influence the app response time in developing countries like India.
- This will also help app developers diagnose performance bottlenecks and failures experienced by their apps in the wild.

Field Experiments

- This data must reflect end user experience under real network conditions.
- So we need to conduct field experiments to gather relevant data and make appropriate conclusions.



Motivation



VS



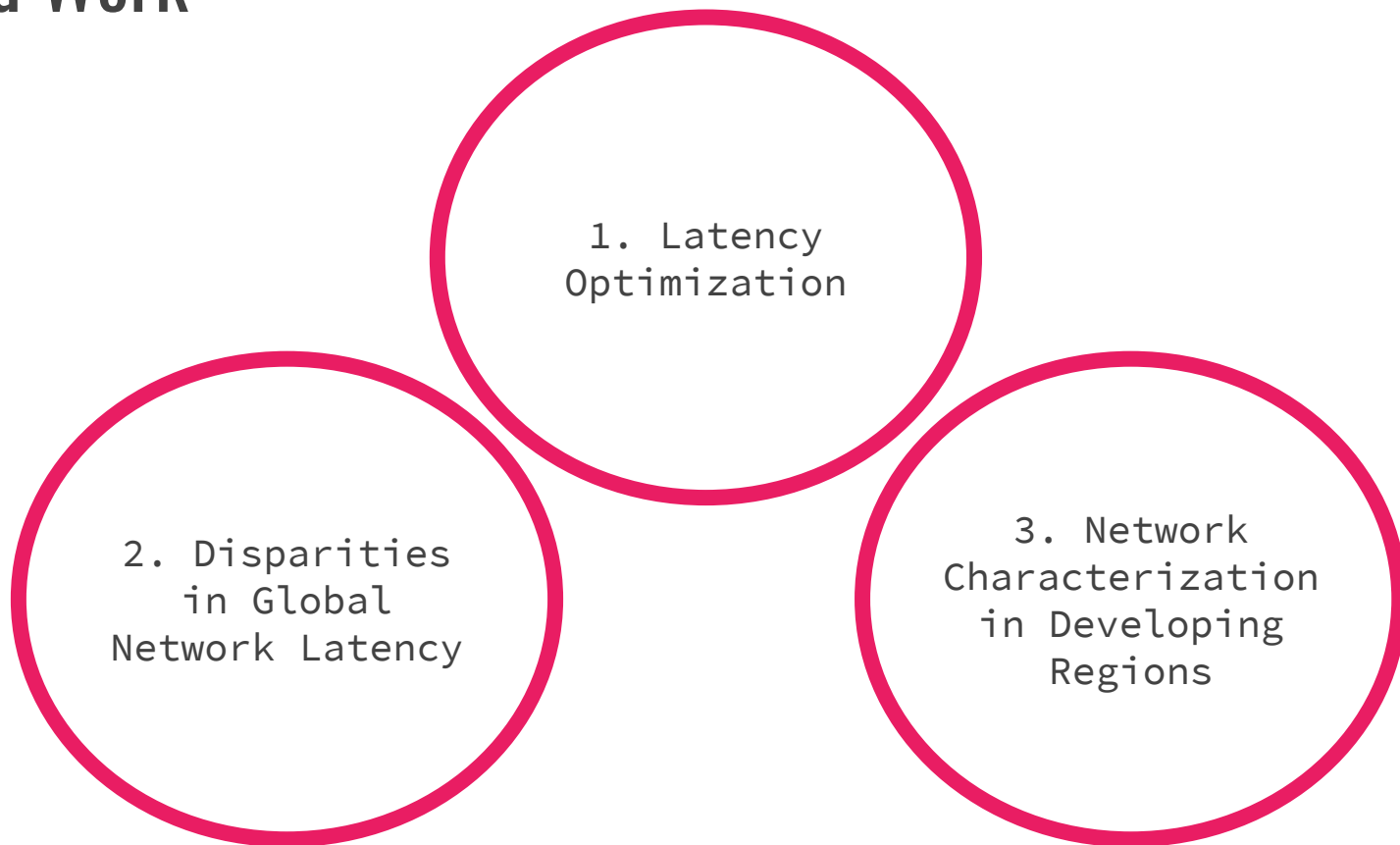
Problem Statement

To understand which of the factors affect the app response times the most:

- type of smartphone used,
- location of the users,
- the version of the apps,
- nature of backhaul network,
- network type

Related Work

Related Work



1. Latency Optimization

**Aim to
measure and/or
optimize the
latency of
mobile web and
apps**

AppInsight[1]
identified
the critical
path in the
execution of
mobile apps and
web

Yun et al[2]
showed that mobile
apps tend to have
lower latency than
web apps in most
cases

QoE Doctor [3]
and PerfProb[4]
also did not
crowdsource data
collection

UI-automator provided by
the Android framework which
requires modification of
the APK

2. Disparities in Global Network Latency

M-Lab [5] performed active tests to identify regional variations in Internet latency across the world

Studies the disparities in network latency across time and region

Bozkurt et al[6] identified the **geographical** and protocol level differences that lead to large diversities in the end-user Internet latency

Vaibhav et al[7] identified the disparities in last-mile latency depending on the **time of day**, ISP used, and **geographical location**.

3. Network Characterization in Developing Regions

Rod  rick et al[8]
and Zaki et al[9]
looked at the web
latencies within
Africa and Ghana
resp.

**Deals with the
performance of
smartphones and
networks in
developing
countries**

Sharma et al.
[11]
studied the
latency observed
by Indian
cellular networks
before the
deployment of
LTE.

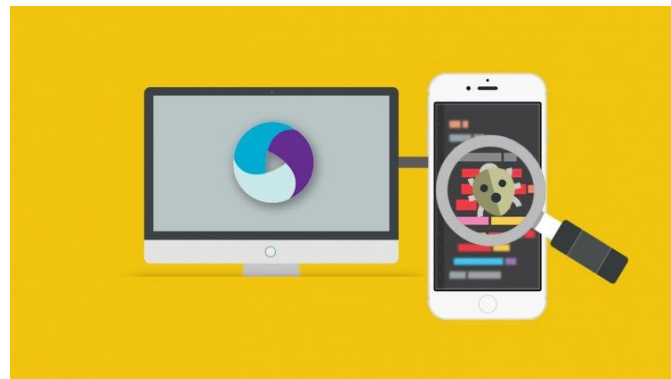
Ravi et al.[10]
characterized the
web pages visited by
the users and
identified memory as
the bottleneck.

Approach





Approach

— — —





Crowdsource our desktop application, EvalApp, to collect data in an automated fashion and perform causal analysis







List of Apps and their features for testing performance

App Type	Application	Actions
Entertainment	 Youtube	<ul style="list-style-type: none">• Search a channel• Play a video
	 Hotstar	<ul style="list-style-type: none">• Play a video• Opening trending page
Social	 LinkedIn	<ul style="list-style-type: none">• View your profile• Check your connections• Search person, (Bill Gates)
	 Facebook	<ul style="list-style-type: none">• Search person• Post in a Facebook group

List of Apps and their features for testing performance

Category	Application	Testing features
News	 Google News	<ul style="list-style-type: none">• Search news about Delhi• Open the headlines page
	 Dailyhunt	<ul style="list-style-type: none">• Search news about sports• Open live tv page
Shopping	 Amazon	<ul style="list-style-type: none">• Searching a product• Opening the product profile• Adding the product to cart• Go to cart page• Removing the product from the cart
	 Flipkart	

List of Apps and their features for testing performance

Category	Application	Testing features
Payment	 PayTM	<ul style="list-style-type: none">• Pay Re. 1/- from the wallet.
Navigation	 Google Maps	<ul style="list-style-type: none">• Search a location
Messaging	 WhatsApp	<ul style="list-style-type: none">• Send a hard-coded message ("Hi, this is an automated test") to the respective group.
	 Telegram	

Approach

EvalApp, an automated framework that uses Appium to conduct experiments on mobile applications.



Evaluation Of Apps

E-mail

Area Pin Code

Check which apps to run:

Non-payment apps:

☒ Amazon

☒ LinkedIn

☒ YouTube

☒ Google News

☒ Flipkart

☒ Facebook

☒ Hotstar

☒ DailyHunt

☒ Telegram

☒ WhatsApp

Payment apps:

☒ Amazon Pay UPI

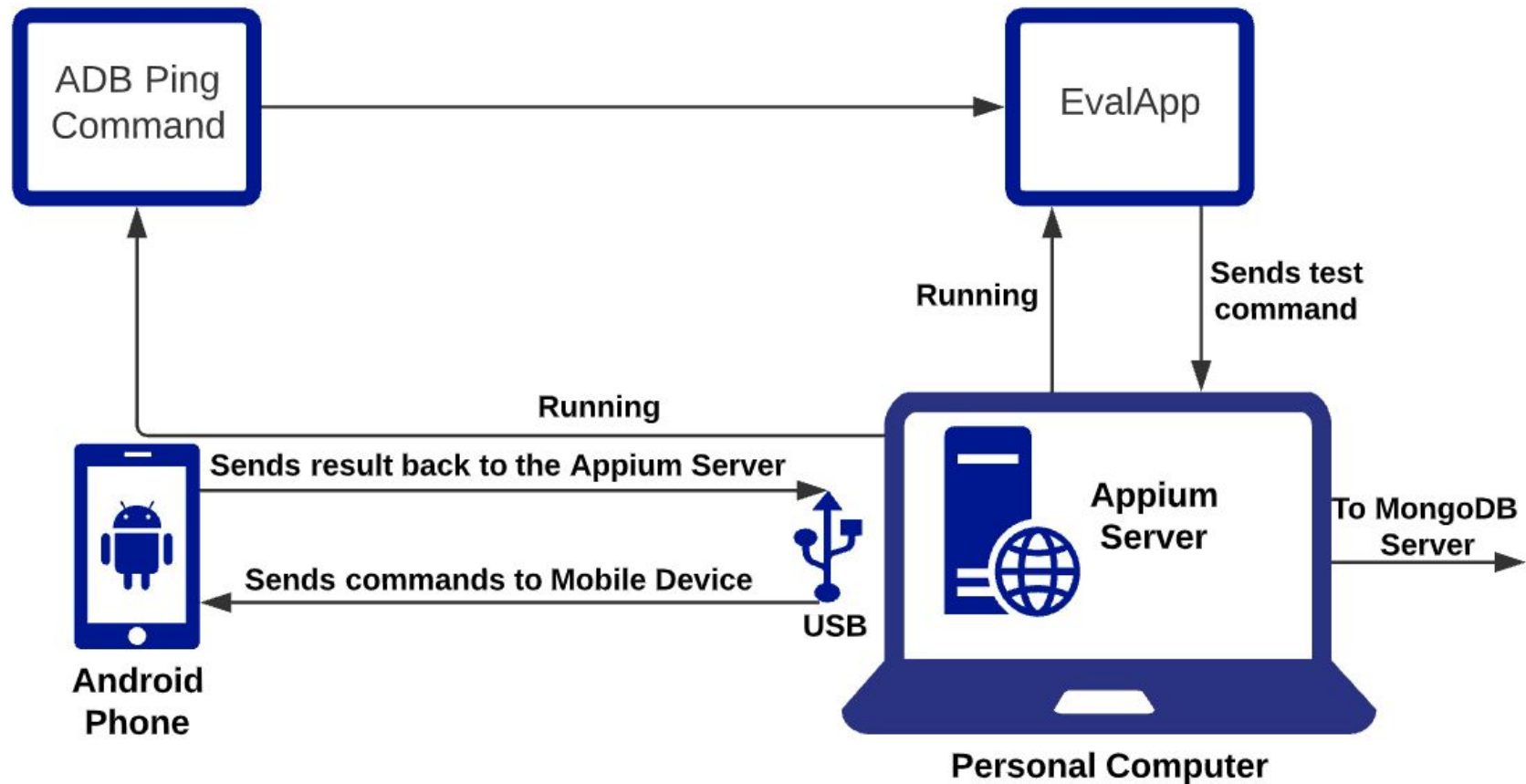
☒ Paytm

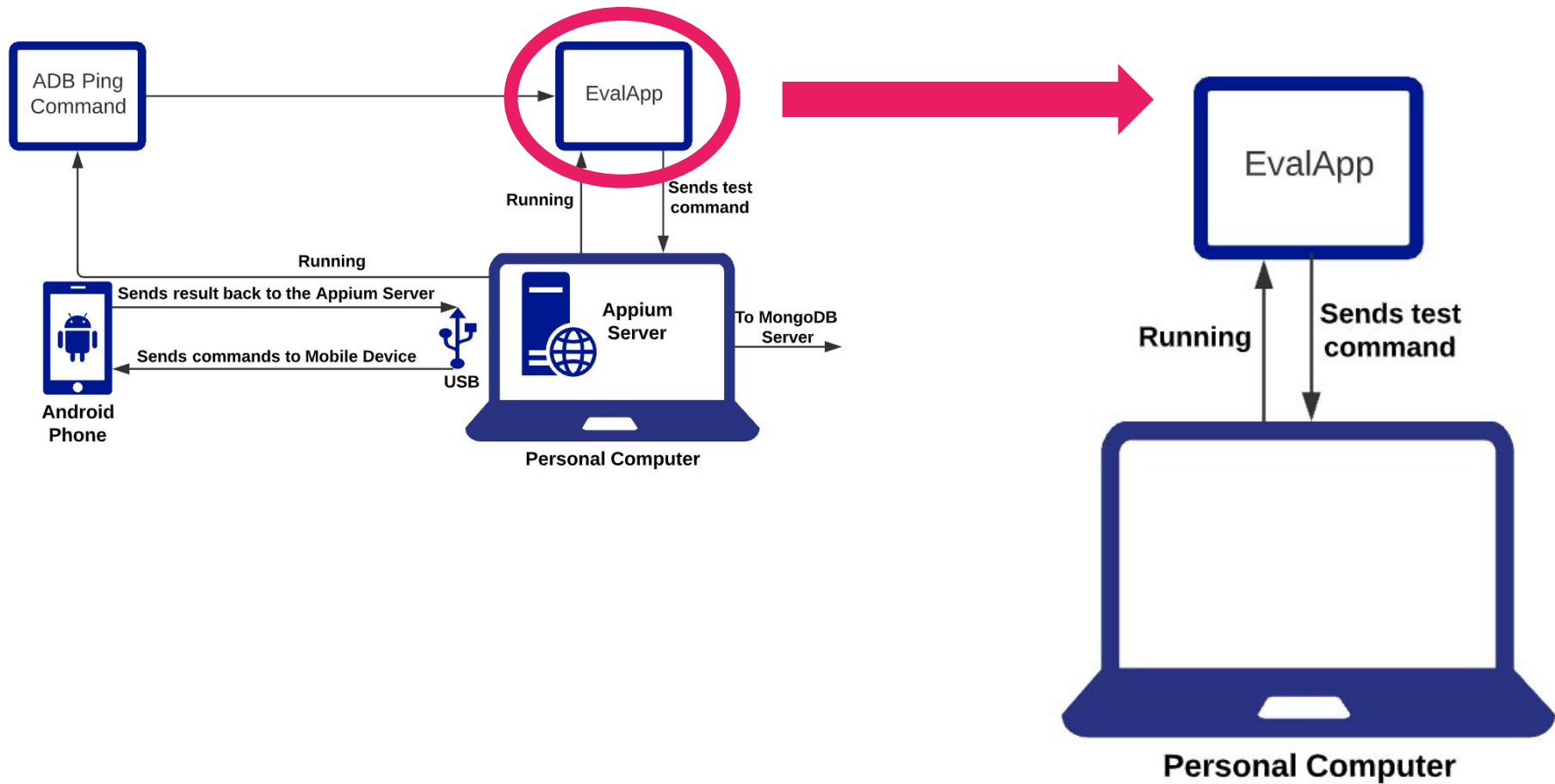
☒ Mobikwik

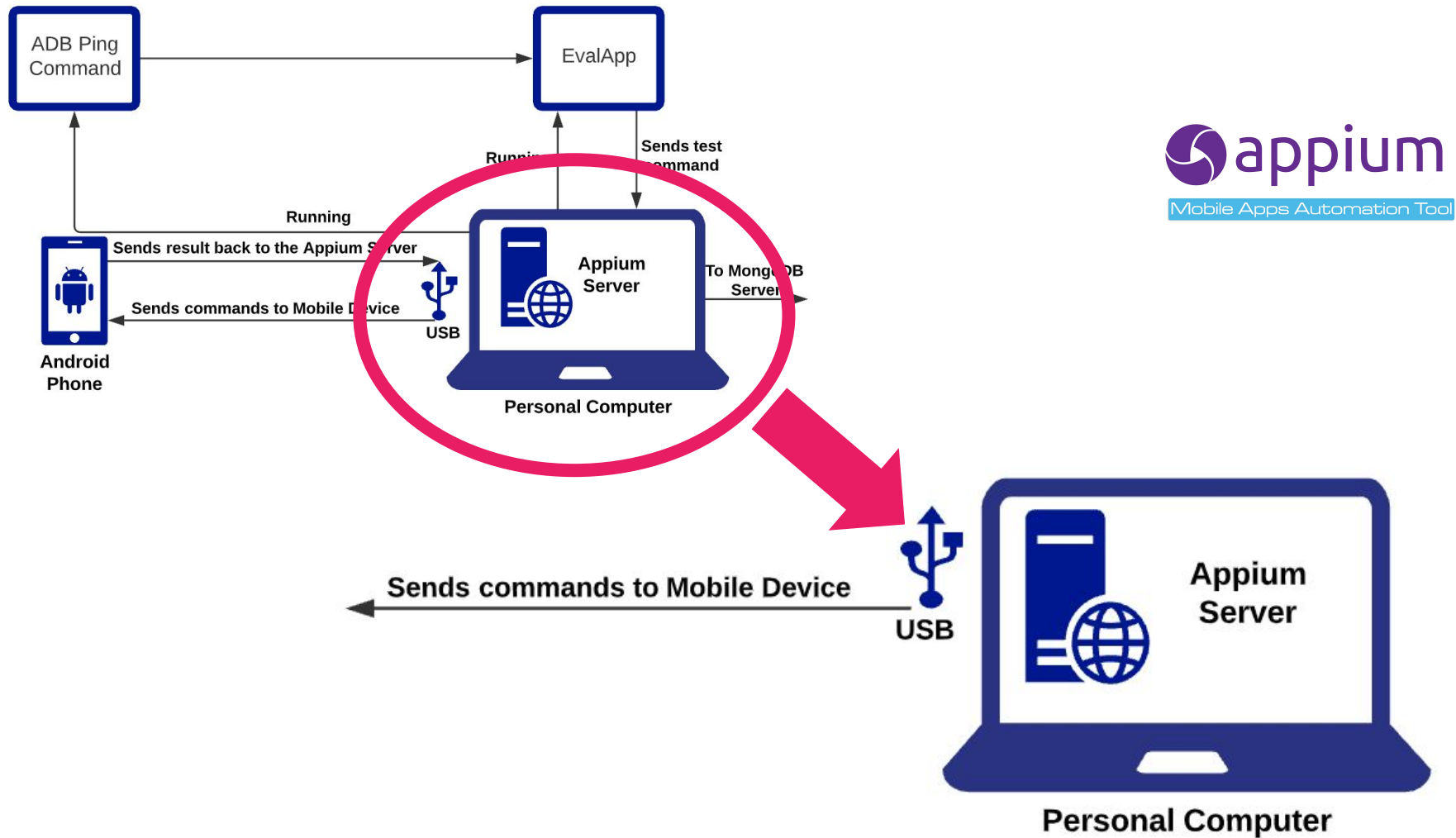
ABOUT

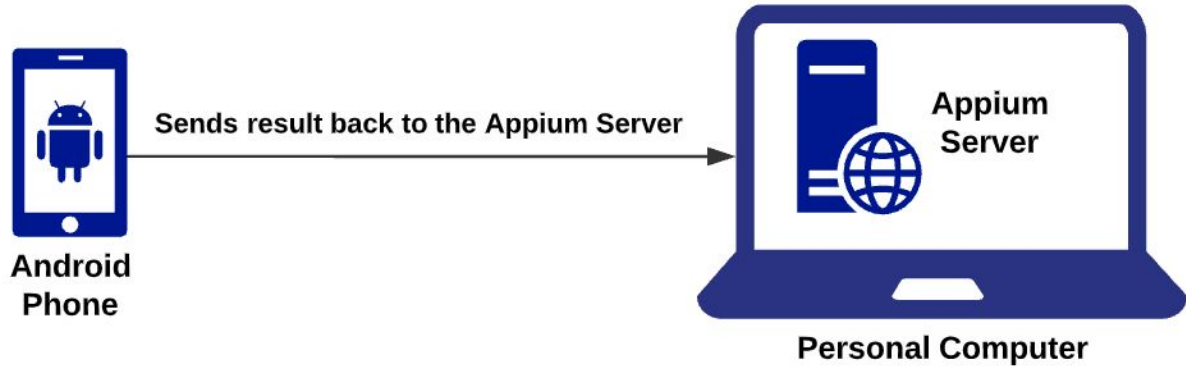
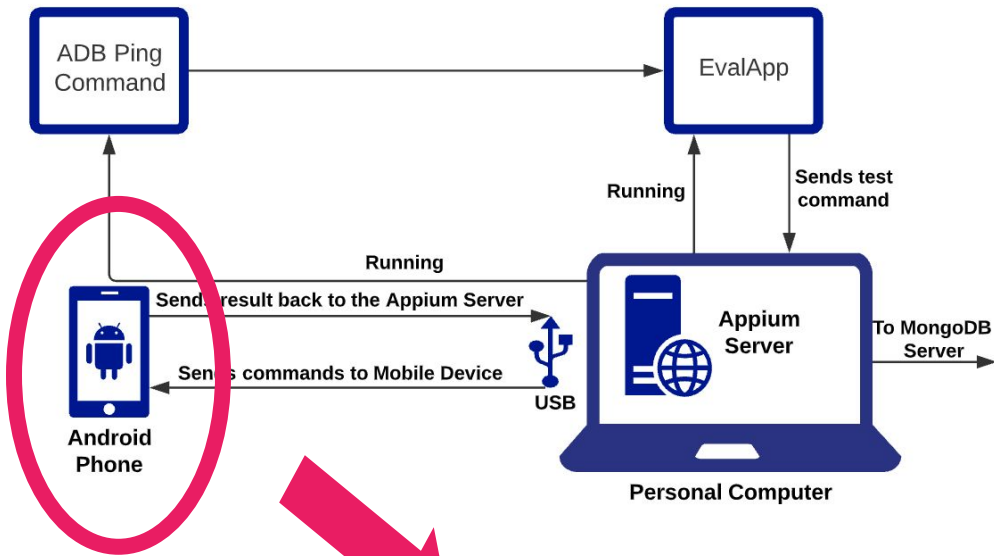
PROCEED

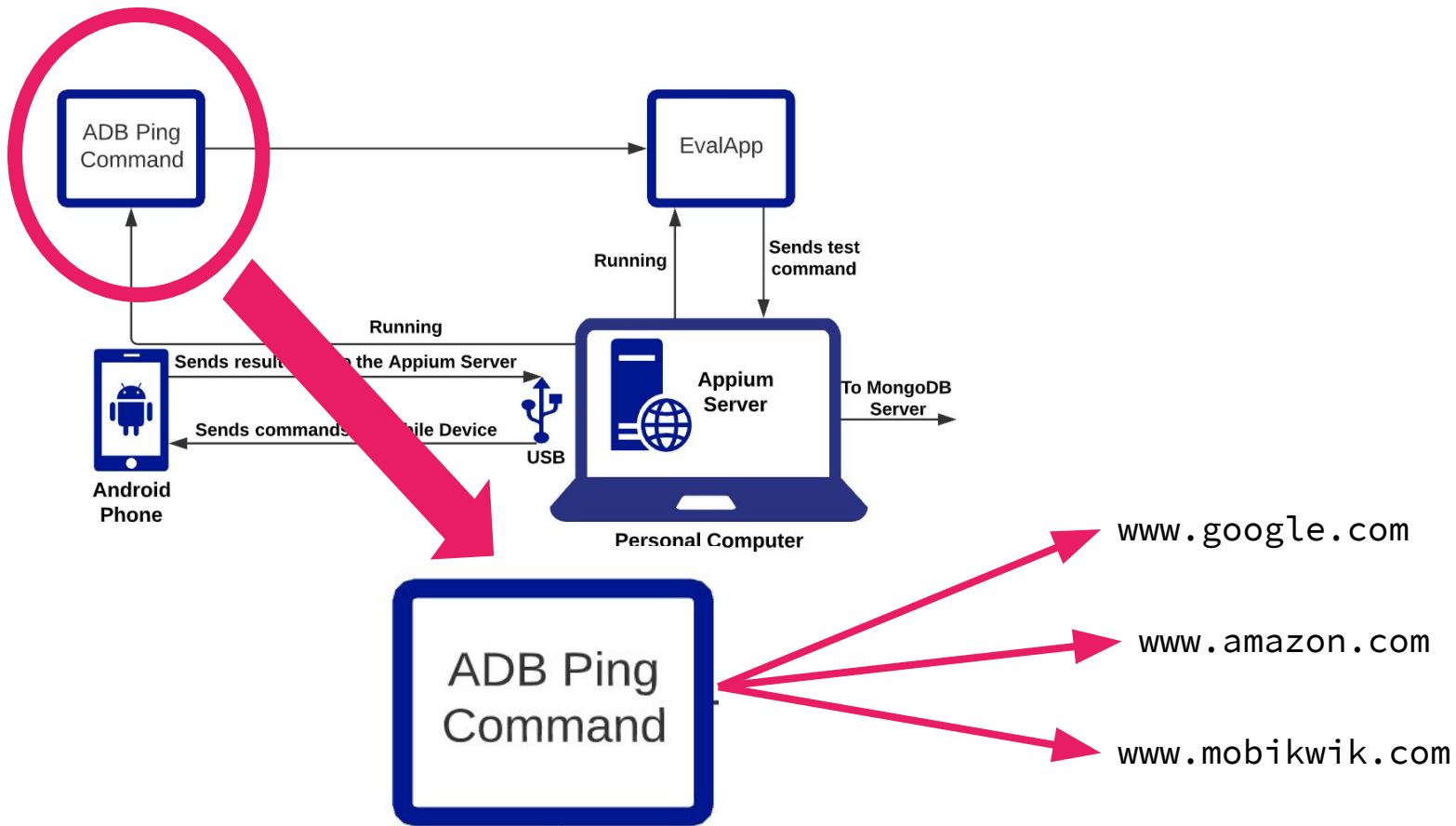
Workflow of EvalApp

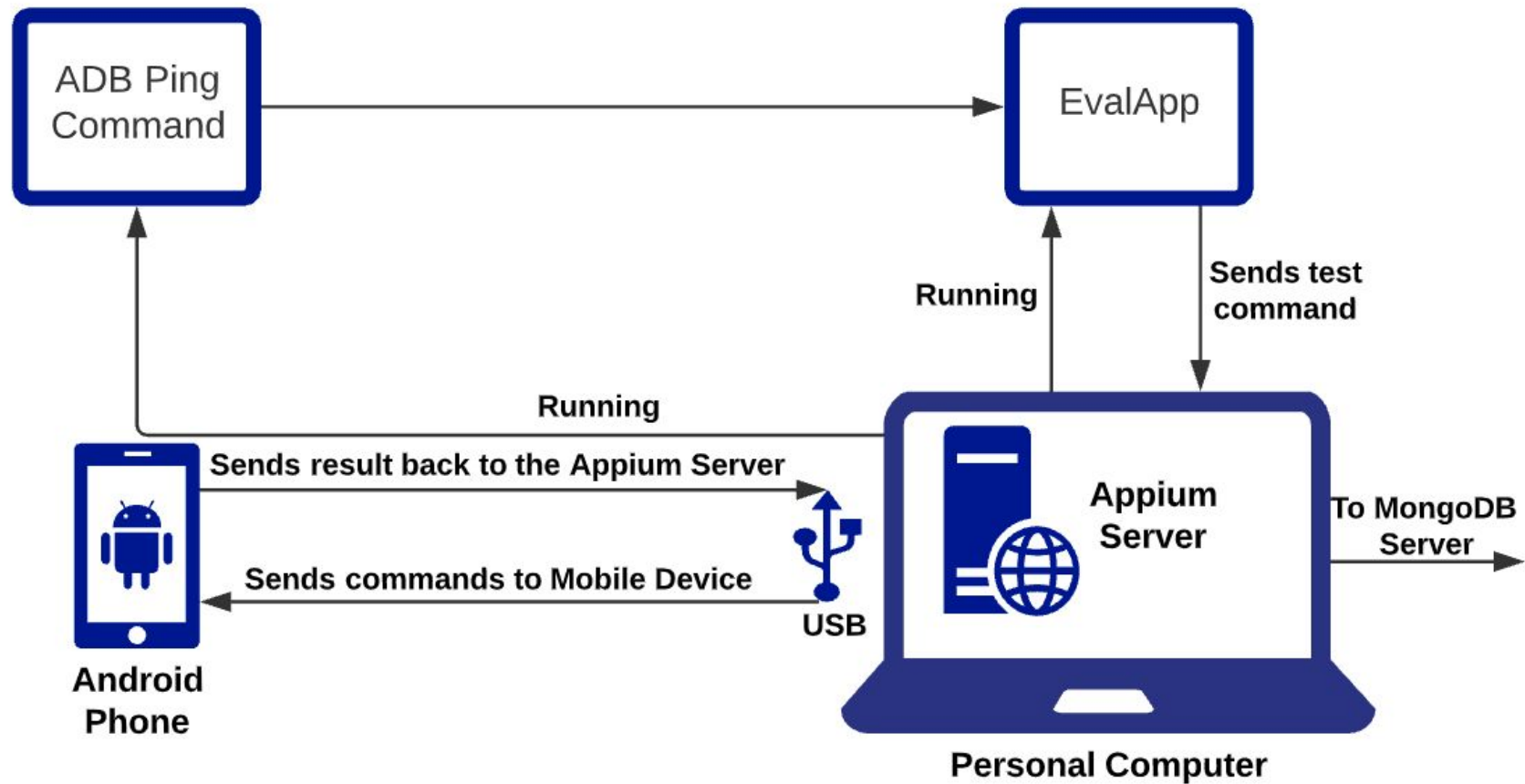













Field Experiment : Data Collection

Challenges Faced

- Requested about 100+ students to volunteer but only 41  finally performed the experiments.
- Conducted regular meetings to resolve the issues being faced by the volunteers.
- Some of the issues were:
 - Phone models like Realme, Xiaomi, OnePlus had different set of USB debugging settings
 - Installation issues in different Linux distributions

Challenges Faced

— — —

- Some of the reasons because of which the volunteers were not able to perform the experiments:
 - Faulty USB cables
 - Not having enough space in phone to install the apps
 - Our framework is not compatible with MacOS

41

Users

21

Locations

8

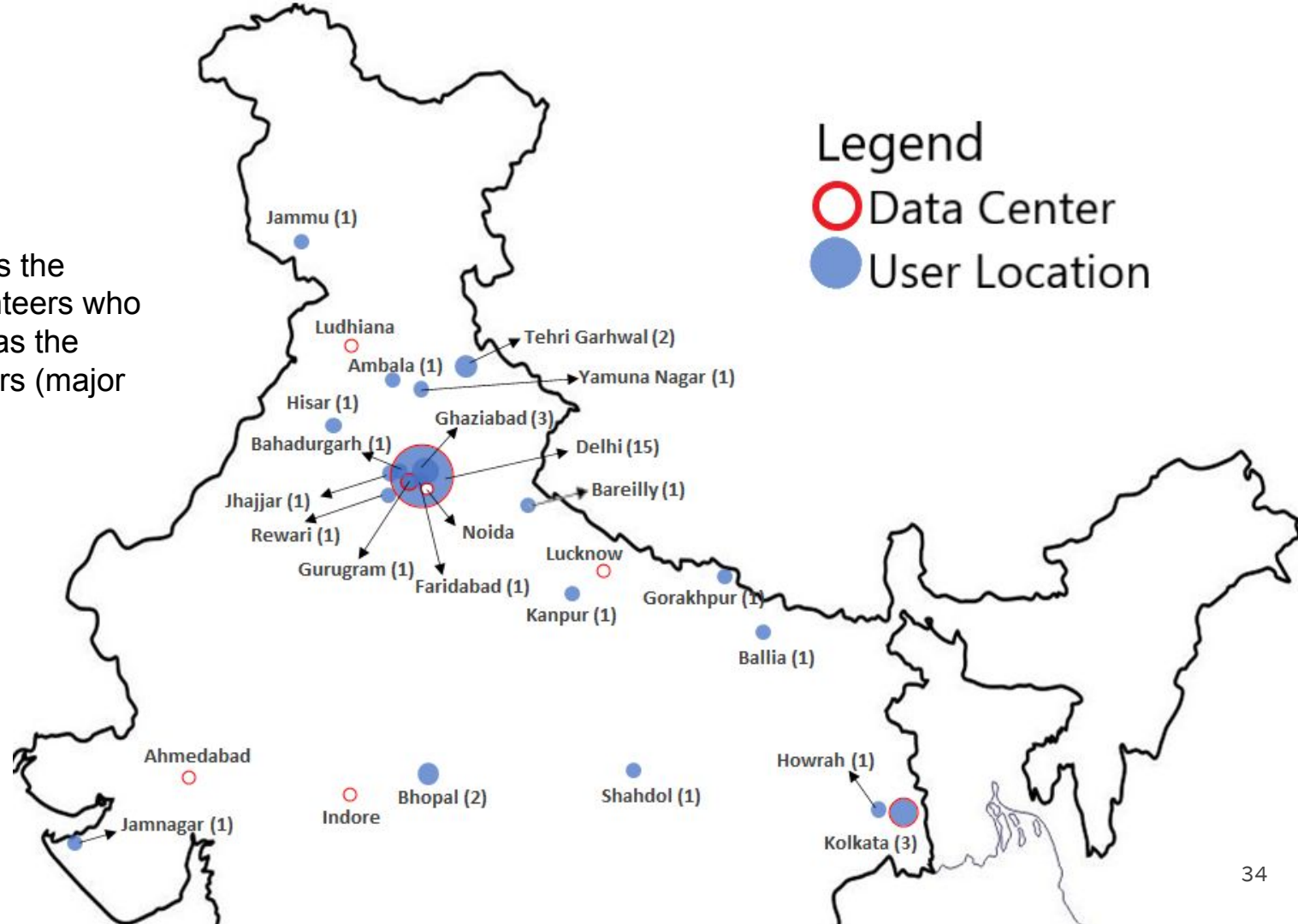
Data Centers

16000

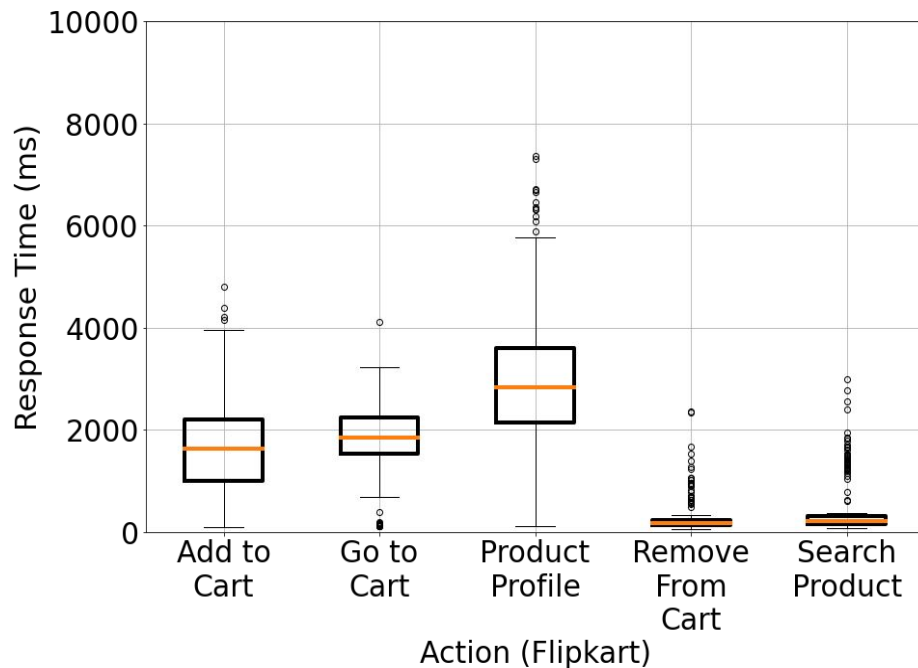
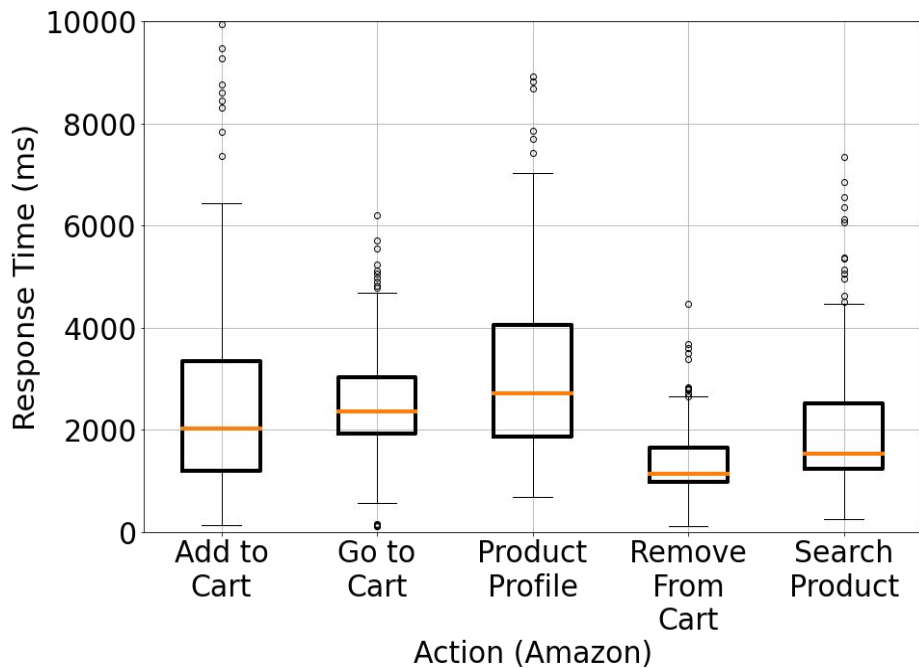
Data Samples

— — —

This figure represents the locations of the volunteers who run EvalApp as well as the cities with data centers (major cities).



Category: Shopping



Category: Shopping

Amazon

- Median Response Times taken by different actions are:

- Add to Cart: 2023.0 ms
- Go to Cart: 2365.0 ms
- Product Profile: 2717.5 ms
- Remove from Cart: 1126.0 ms
- Search Product: 1535.0 ms

Action	UI Response
Add to cart	Cart value increased
Go to cart	Cart displayed
Product profile	Product rating stars shown

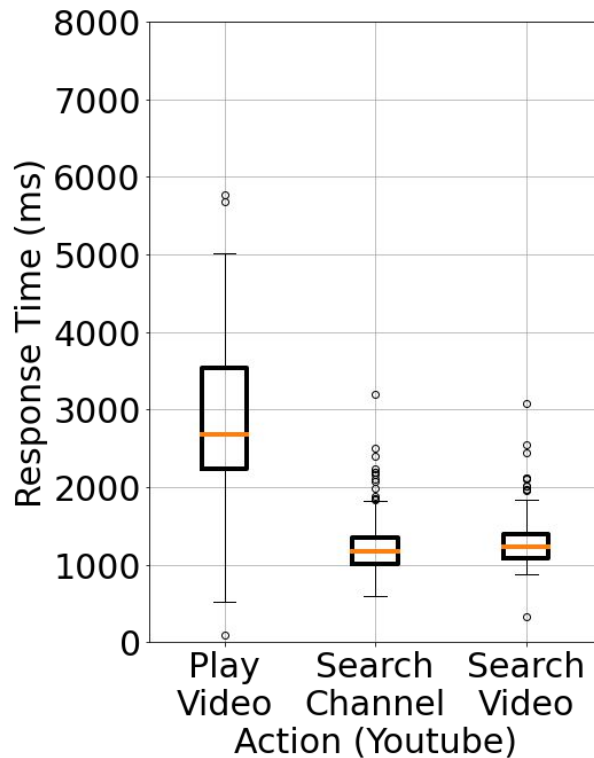
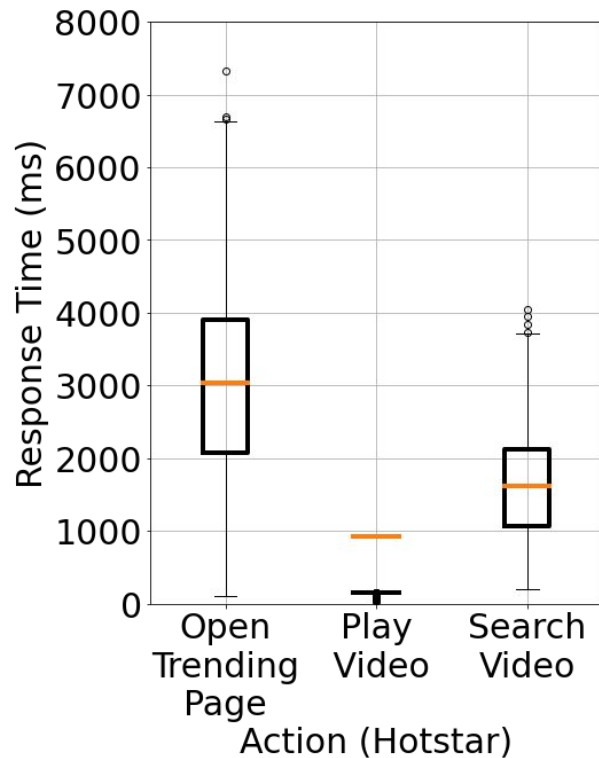
Flipkart

- Median Response Times taken by different actions are:

- Add to Cart: 1627.0 ms
- Go to Cart: 1854.0 ms
- Product Profile: 2835.0 ms
- Remove from Cart: 173.5 ms
- Search Product: 214.5 ms

Action	UI Response
Remove from cart	Cart value decreased
Search Product	Search results displayed

Category: Entertainment



Category: Entertainment

Hotstar:

- Median Response Times taken by different actions are:
 - Open trending page: 3031.0 ms
 - Play video: 927.0 ms
 - Search video: 1612.0 ms

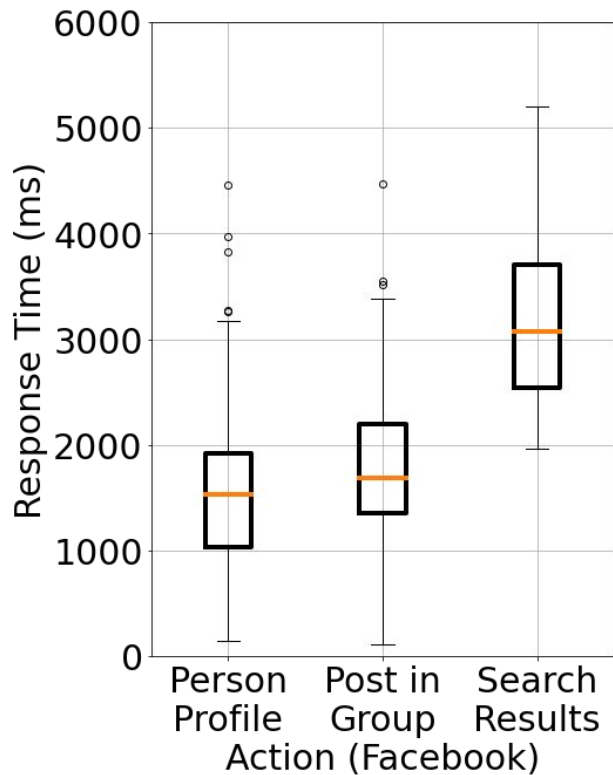
YouTube:

- Median Response Times taken by different actions are:
 - Search channel: 1169.5 ms
 - Play video: 2673.0 ms
 - Search video: 1223.0 ms

Action	UI Response
Open trending page	Video player box appeared
Play video	Video title displayed
Search video/channel	Search results displayed

Category: Social Media

-- -- --



Category: Social Media

Facebook:

- Median Response Times taken by different actions are:
 - Searching a person:
 - Search person: 3075.0 ms
 - Person Profile: 1534.5 ms
 - Post in a group: 1686.0 ms

LinkedIn:

- Median Response Times taken by different actions are:
 - Searching a person:
 - Search Results: 652.5 ms
 - Profile Page: 833.0 ms
 - Check connections: 1395.0 ms

Action	UI Response
Search Person	Search result displays
Person profile/Profile Page	Person profile picture appears
Post in a group	Post like button appears
Check connections	List of connections appears

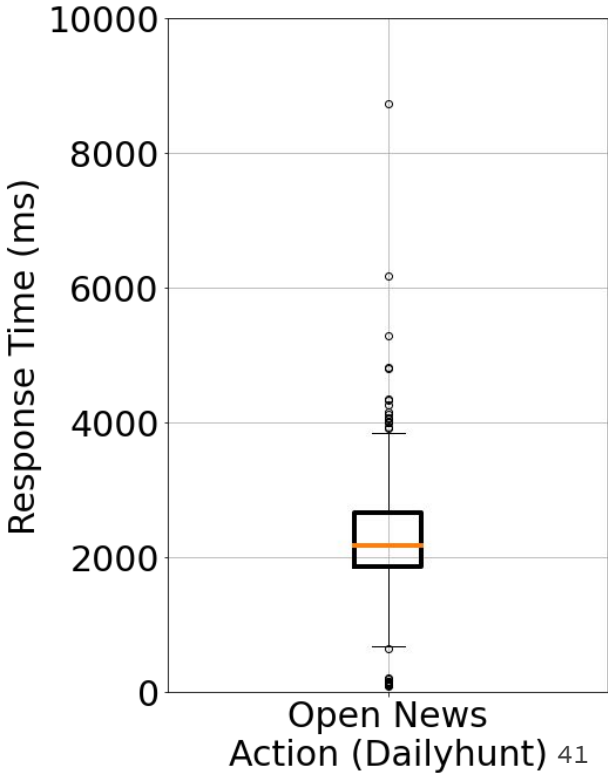
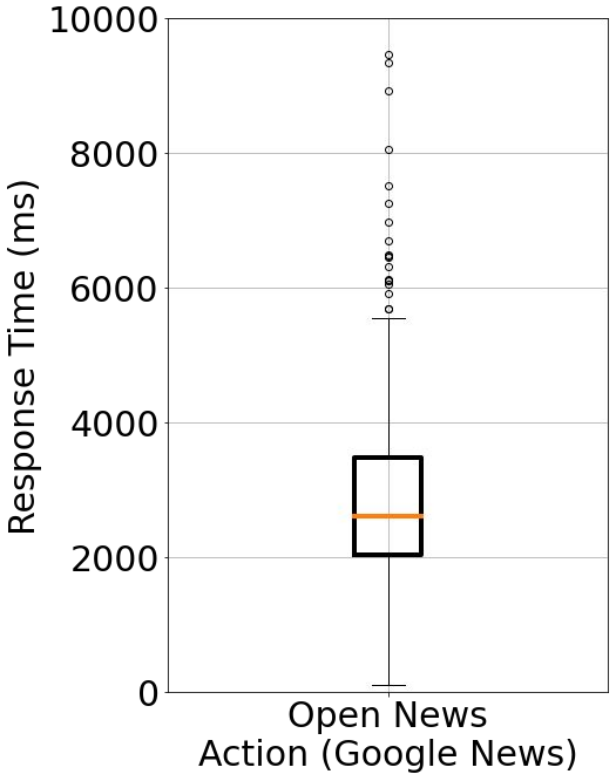
Category: News

Median Response Time
taken for opening a
news page by:

Google News: 2609.5 ms

DailyHunt: 2167.0 ms

Action	UI Response
Open News	News title displays



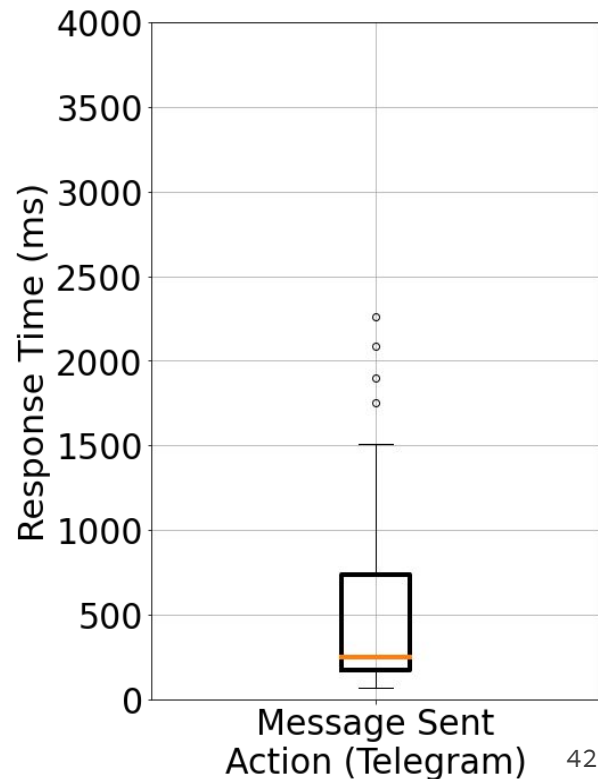
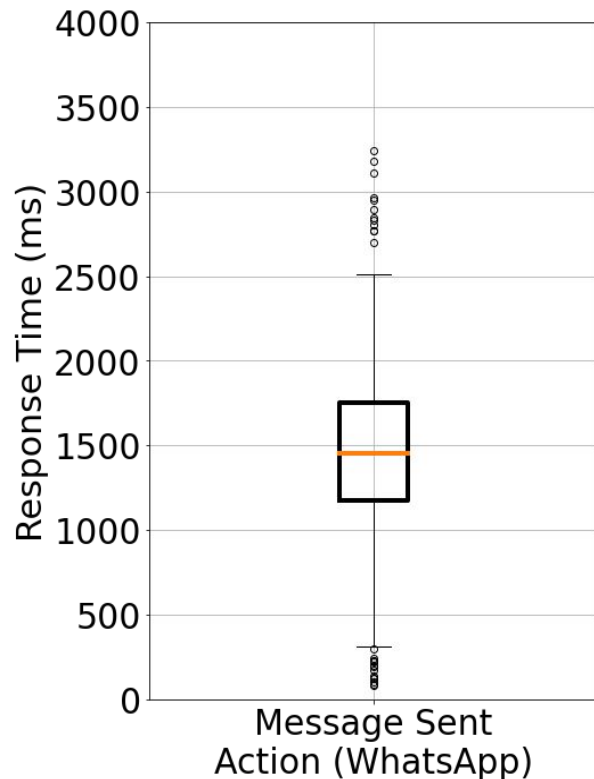
Category: Messaging

Median Response Time
taken for sending a
message by:

WhatsApp: 1454.5 ms

Telegram: 251.5 ms

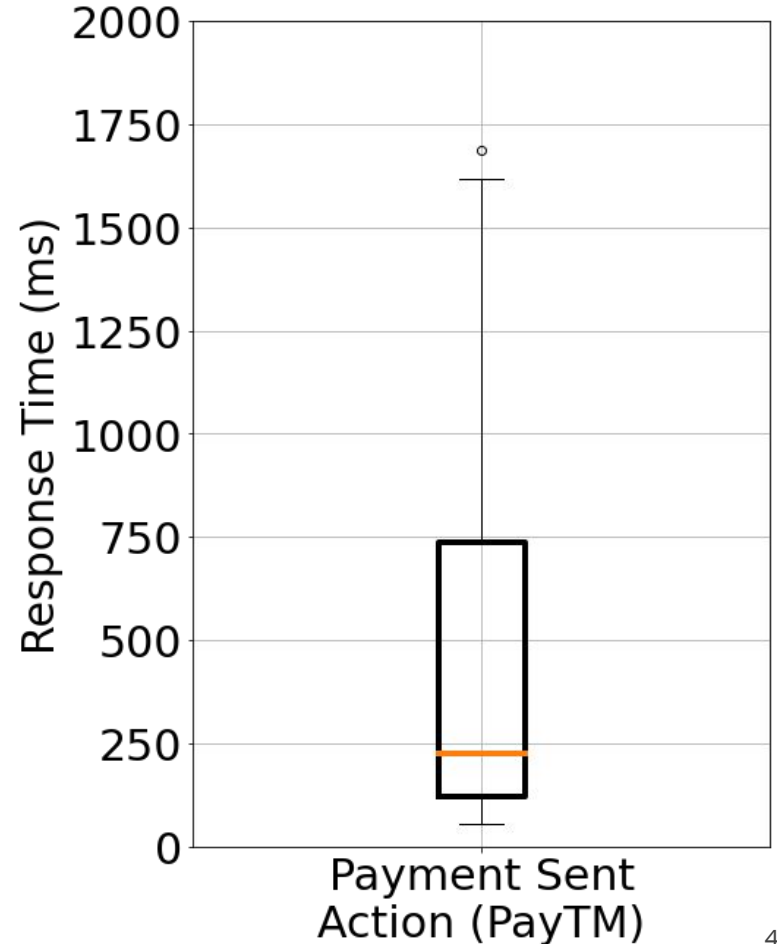
Action	UI Response
Send message	Single tick appears



Category: Payment

Median Response Time taken for sending money through wallet by PayTm is 227.0 ms

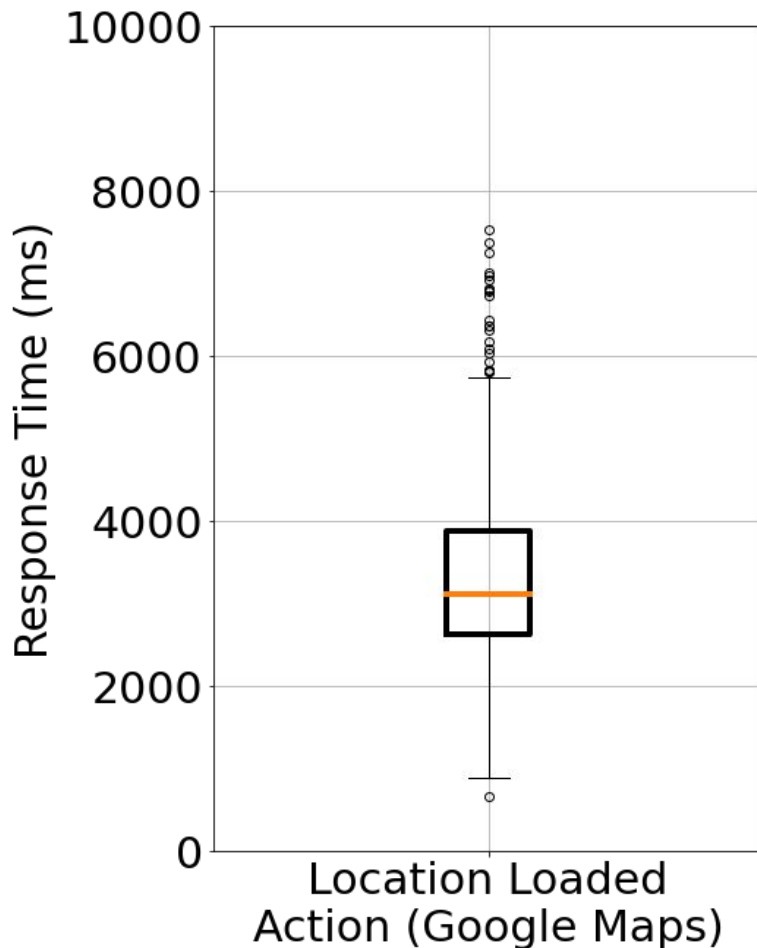
Action	UI Response
Send money	Payment successful icon is shown



Category: Navigation

- Median Response Time taken for loading a location in Google Maps is 3110.0 ms

Action	UI Response
Search a location	Street thumbnail displays



Features Collected

— — —

Feature Name	Range of Values
Android Version	7 to 11
RAM of mobile device	2 GB - 8 GB
Distance from datacenter	0 KM - 397 KM (normalized to 0 - 1)
Time of Day	Morning (0) and Evening (1)
Type of Network	WiFi (0), Mobile Data (1) and Mobile Hotspot (2)
RTT of pings to google.com	4 ms to 2000 ms (normalized to 0 - 1)

Features Collected

Which of these
features is
the most
important?

Feature Name	
Android Version	
RAM of mobile device	
Distance from datacenter	
Time of Day	Morning (0) and Evening (1)
Type of Network	WiFi (0), Mobile Data (1) and Mobile Hotspot (2)
RTT of pings to google.com	4 ms to 2000 ms (normalized to 0 - 1)

Field Experiment : Causal Analysis

Causal Analysis

— — —

- Computing correlation between a feature and the response time values was not possible because the other features were not constant.
- Also if we try to keep all the other features constant then size of data becomes very small and thus insufficient to make correct conclusions.
- So we moved to more sophisticated causal analysis methods.

Method

— — —

1. K-Means Clustering

- K-means algorithm is an iterative algorithm that tries to partition the dataset into K distinct non-overlapping clusters.
- Silhouette analysis is used to determine the optimal value of K i.e. number of clusters.
- We then create clusters of app response times using K-means algorithm.

Method

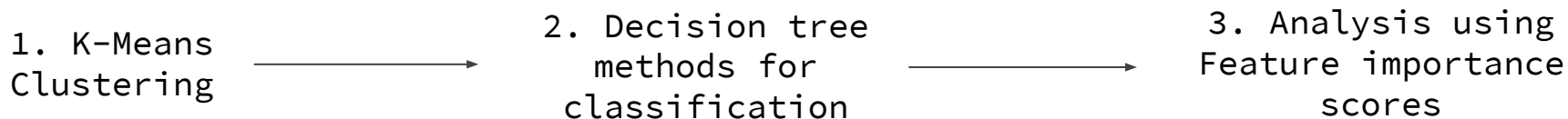
— — —



- Next, we used decision trees to classify the data samples into these clusters with the help of all the six features.

Method

— — —



- After classification, we found out the importance of all the features to identify the most influential feature for every application.

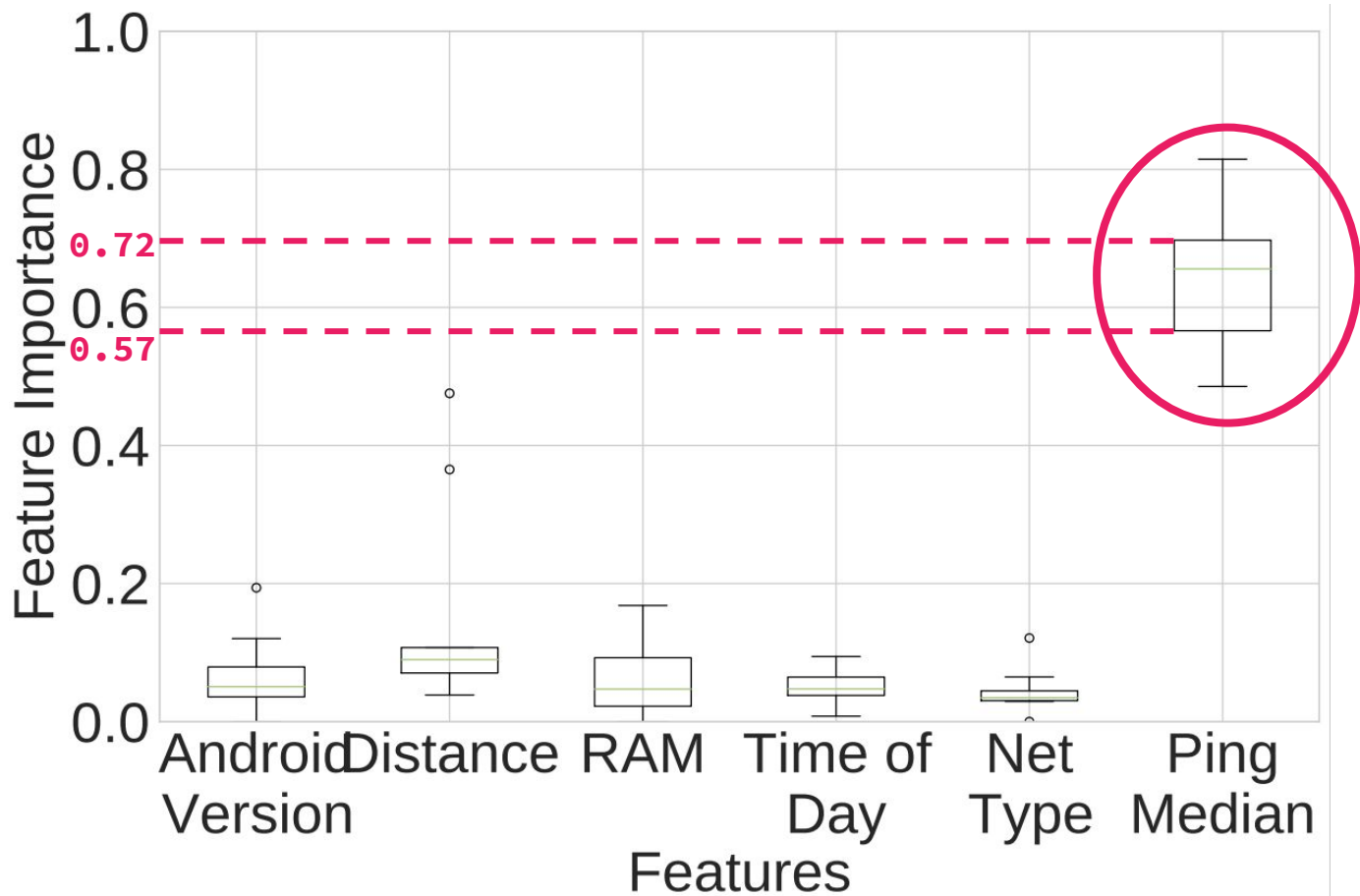
Method

— — —

- K-means algorithm is an iterative algorithm that tries to partition the dataset into K distinct non-overlapping clusters.
- Silhouette analysis is used to determine the optimal value of K i.e. number of clusters.
- We then create clusters of app response times using K-means algorithm.
- Now, we used decision trees to classify the data samples into these clusters with the help of all the six features.
- After classification, we found out the importance of all the features to identify the most influential feature for every applications.

Results

Including
Ping Median

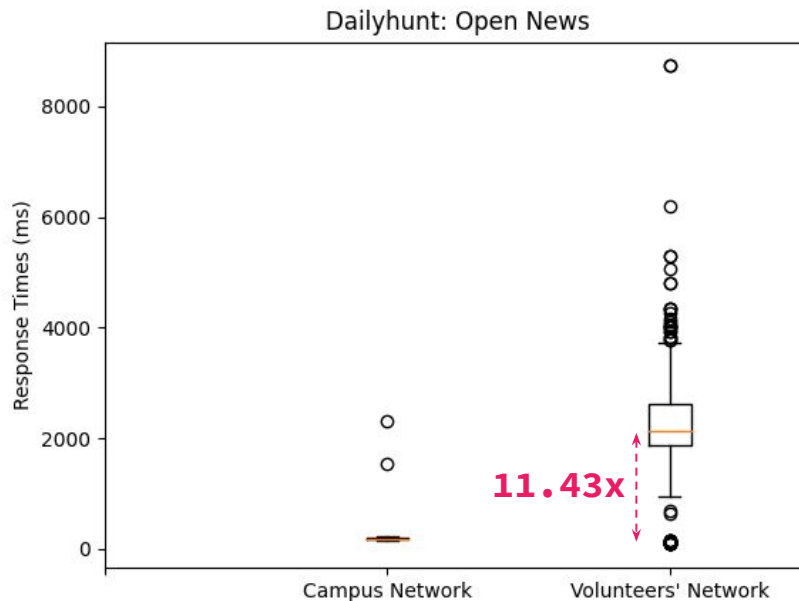
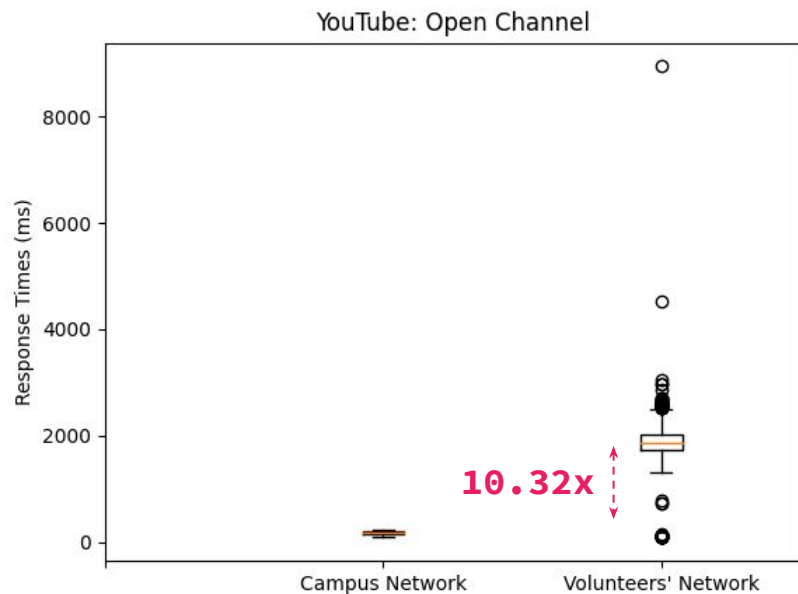


Confirming that Network is the Primary Bottleneck

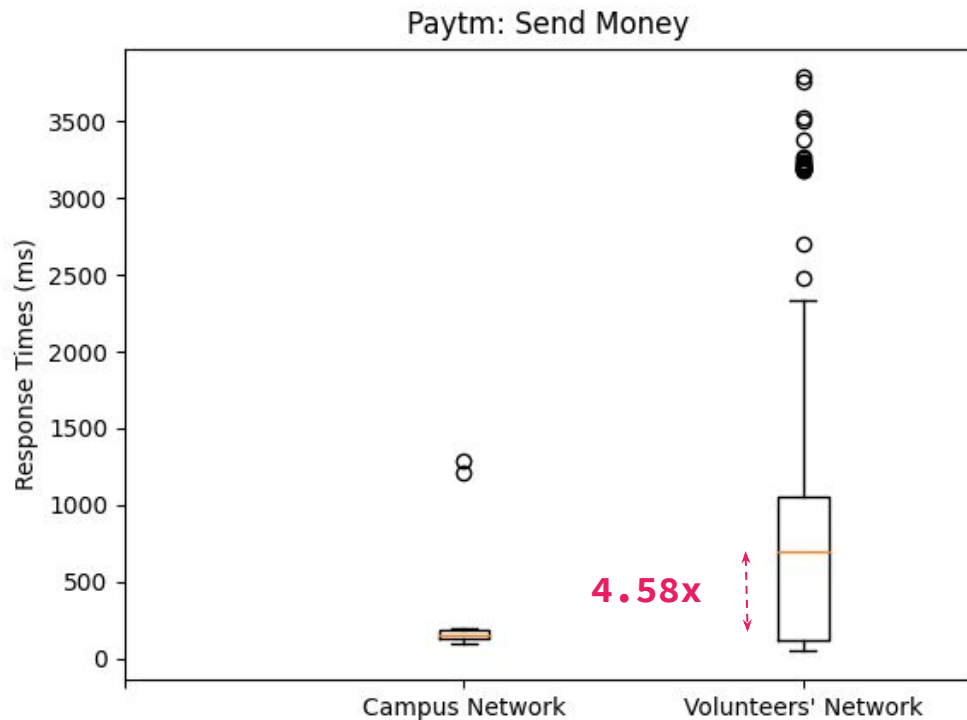
— — —

- Performed a controlled experiment in our campus using a Samsung A50S smartphone (RAM 4GB) during a period of low congestion.
- Median network latency observed with NKN was 9.1ms compared to 67.6ms for commercial home networks used by volunteers.
- We observed a significant reduction in response times for most of the actions.

Confirming that Network is the Primary Bottleneck

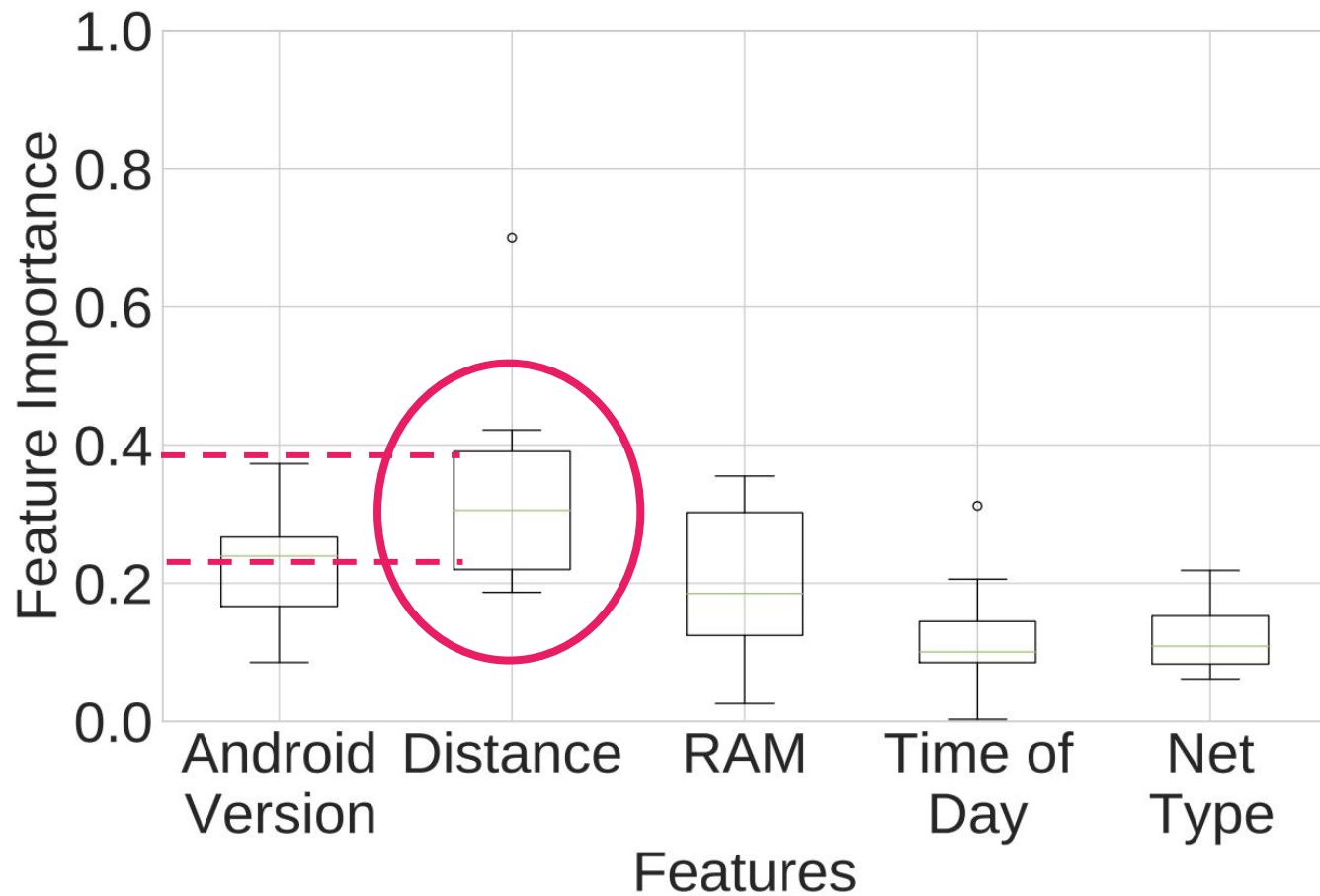


Confirming that Network is the Primary Bottleneck



Results

Without
Ping Median



Results

We list down the app-wise important features apart from the ping,

Feature	Apps
Distance	Youtube, LinkedIn, Hotstar, Flipkart, Amazon, Telegram, Whatsapp, Google news, PayTM, Dailyhunt, Google Maps
RAM	Facebook, Google Maps
Android version	Facebook, Dailyhunt

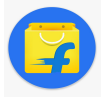
Controlled Experiment : Data

Why is it required?

— — —

- Users expect better performance with an updated version of their app.
- But this does not happen always.
- We cannot perform field experiment for this kind of analysis because it will be inconvenient for the volunteers to try out different app versions of the same app repeatedly.
- Also we needed same network conditions across the app versions which will be difficult for us to ensure in a field experiment

Performed on 4 Apps with 3 versions each



1. Flipkart



2. Google Maps



3. YouTube



4. Telegram

Performed on 2 Android OS



Android 5



Android 8

Device Details

Lenovo A6020a40
2 GB RAM

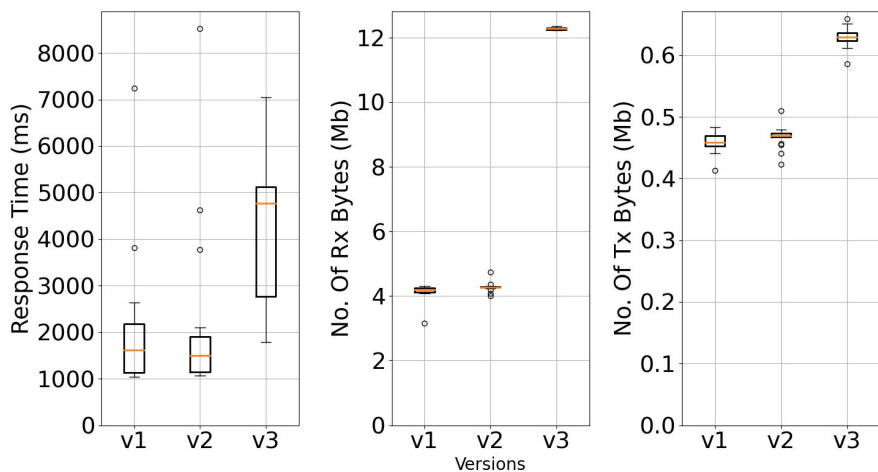
Features Collected

Response Times
Rx Bytes
Tx Bytes

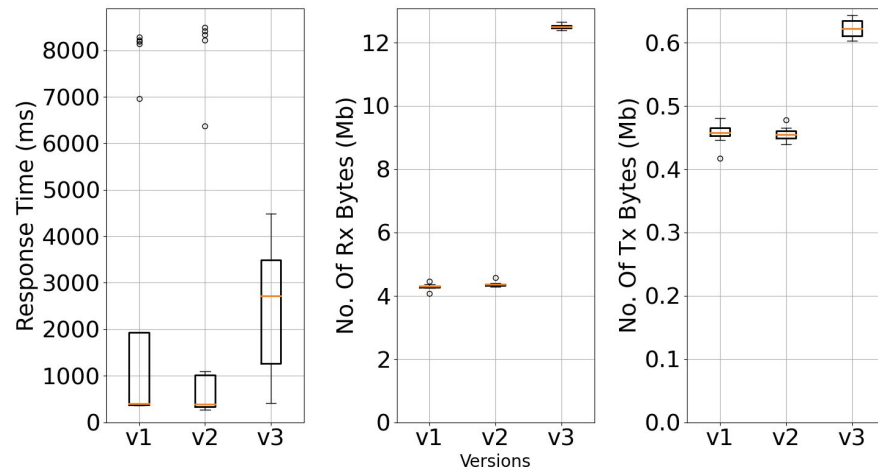
Controlled Experiment : Analysis

Flipkart with versions: v6.10, v6.15, v7.15

Test Name: Open product profile



Experiment Run on Android 5

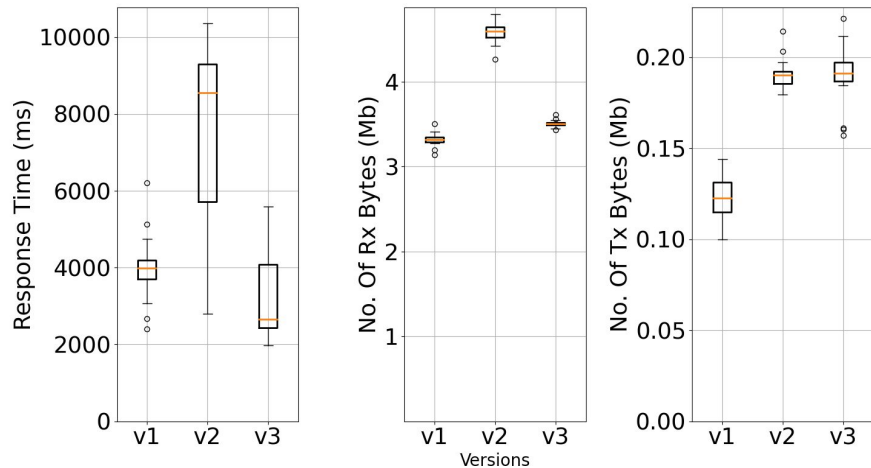


Experiment Run on Android 8

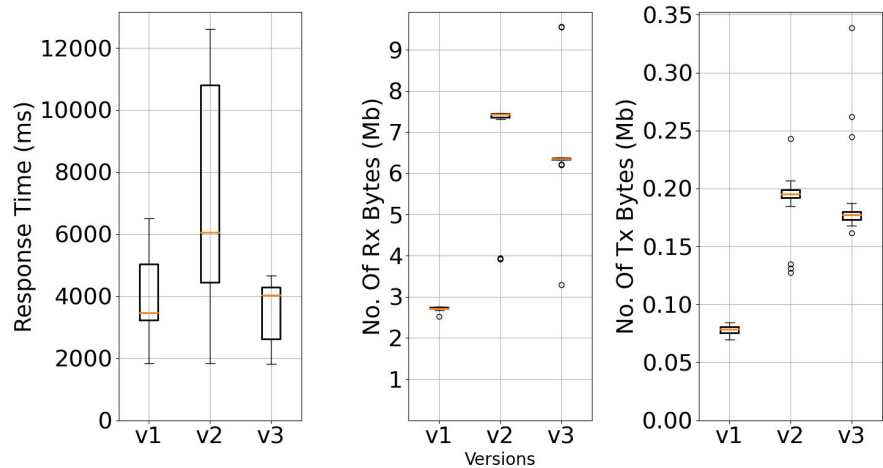
v1: v6.10, v2: v6.15, v3: v7.15

Google Maps with versions: v9.67.1, v10.8.1, v10.32.2

Test name: Search a location



Experiment Run on Android 5

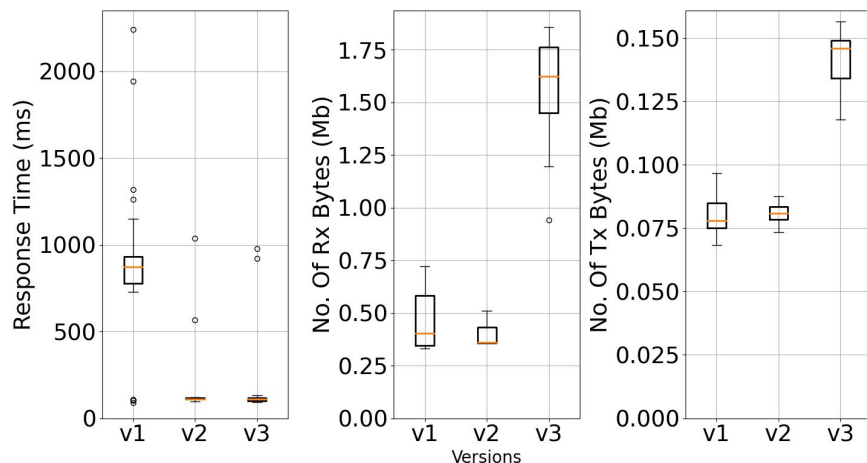


Experiment Run on Android 8

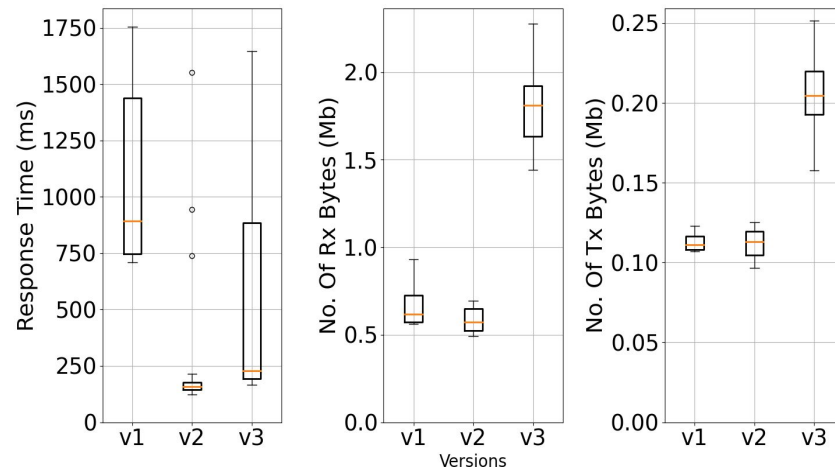
v1: v9.67.1, v2: v10.8.1, v3: v10.32.2

Telegram with versions: v5.15, v6.3.0, v7.6.0

Test name: Send a message



Experiment Run on Android 5

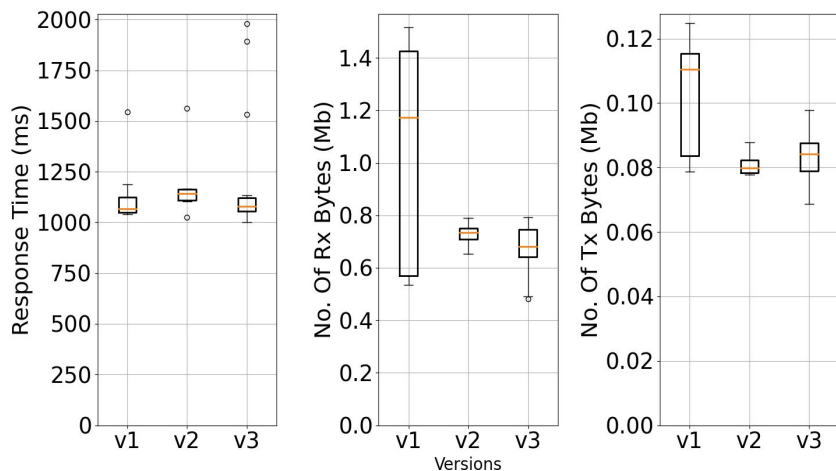


Experiment Run on Android 8

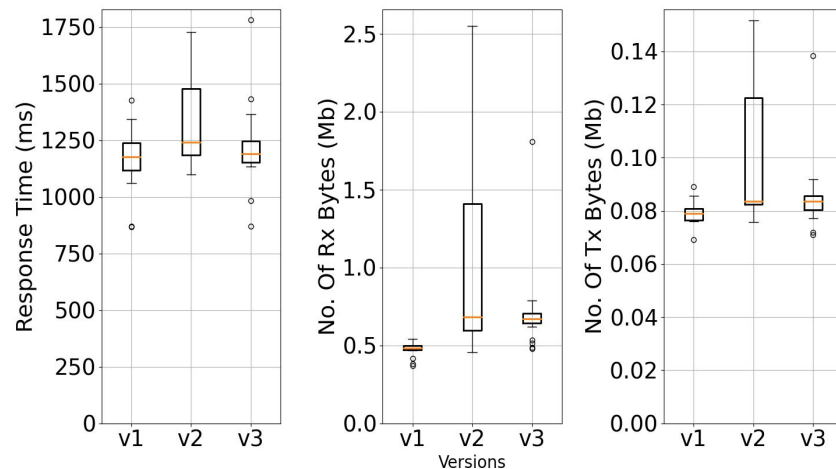
v1: v5.15, v2: v6.3.0, v3: v7.6.0

YouTube with versions: v14.43.55, v15.50.35, v16.02.32

Test Name: Search a channel



Experiment Run on Android 5



Experiment Run on Android 8

v1: v14.43.55, v2: v15.50.35, v3: v16.02.32

Conclusion

Conclusion

- Field Experiment analysis:
 - Network is the most contributing factor in terms of the app response time.
 - Next major factor is the distance from the major cities where the data centers are situated.
 - So our recommendation to the app developers is that they modify their apps to reduce data consumption.

Conclusion

- Controlled Experiment analysis:
 - The response time of apps do not always improve with version upgrades.
 - Also in most cases there exists a relationship between the response times and the amount of data usage across app versions

**We have submitted our
work as a short paper in
IMC-2021**

Future Work

Future Work

— — —

- Conducting this study across different demography and other regions including remote regions to remove potential bias is an immediate future work.
- A detailed analysis to figure out why network affects the app latencies the most is also a potential future work.

Conclusion

- Field Experiment analysis:
 - Network is the most contributing factor in terms of the app response time.
 - Next major factor is the distance from the major cities where the data centers are situated.
 - So our recommendation to the app developers is that they modify their apps to reduce data consumption.

Any questions?

References

— — —

- [1] Ravindranath, Lenin, et al. "Appinsight: Mobile app performance monitoring in the wild." 10th {USENIX} Symposium on Operating Systems Design and Implementation ({OSDI} 12). 2012.
- [2] Ma, Yun, et al. "A tale of two fashions: An empirical study on the performance of native apps and web apps on android." IEEE Transactions on Mobile Computing 17.5 (2017): 990-1003.
- [3] Chen, Qi Alfred, et al. "Qoe doctor: Diagnosing mobile app qoe with automated ui control and cross-layer analysis." Proceedings of the 2014 Conference on Internet Measurement Conference. 2014.
- [4] Hong, David Ke, et al. "Perfprobe: A systematic, cross-layer performance diagnosis framework for mobile platforms." 2019 IEEE/ACM 6th International Conference on Mobile Software Engineering and Systems (MOBILESoft). IEEE, 2019.
- [5] Høiland-Jørgensen, Toke, et al. "Measuring latency variation in the internet." Proceedings of the 12th International on Conference on emerging Networking EXperiments and Technologies. 2016.
- [6] Bozkurt, Ilker Nadi, et al. "Why is the internet so slow?!" International Conference on Passive and Active Network Measurement. Springer, Cham, 2017.

References

— — —

- [7] Bajpai, Vaibhav, Steffie Jacob Eravuchira, and Jürgen Schönwälder. "Dissecting last-mile latency characteristics." ACM SIGCOMM Computer Communication Review 47.5 (2017): 25-34.
- [8] Fanou, Rodérick, et al. "Exploring and analysing the African web ecosystem." ACM Transactions on the Web (TWEB) 12.4 (2018): 1-26.
- [9] Zaki, Yasir, et al. "Dissecting web latency in ghana." Proceedings of the 2014 Conference on Internet Measurement Conference. 2014.
- [10] Naseer, Usama, Theophilus A. Benson, and Ravi Netravali. "WebMedic: Disentangling the Memory-Functionality Tension for the Next Billion Mobile Web Users." Proceedings of the 22nd International Workshop on Mobile Computing Systems and Applications. 2021.
- [11] Sharma, Asheesh, et al. "Revisiting the state of cellular data connectivity in india." Proceedings of the 2015 Annual Symposium on Computing for Development. 2015.