MYSPEEDTEST: ACTIVE AND PASSIVE MEASUREMENTS OF CELLULAR DATA NETWORKS

A Thesis
Presented to
The Academic Faculty

by

Sachit Muckaden

In Partial Fulfillment
of the Requirements for the Degree
Master of Science in the
School of Computer Science

Georgia Institute of Technology
May 2013

MYSPEEDTEST: ACTIVE AND PASSIVE MEASUREMENTS OF CELLULAR DATA NETWORKS

Approved by:

Professor Nick Feamster, Advisor
School of Computer Science
Georgia Institute of Technology

Professor Patrick Traynor
School of Computer Science
Georgia Institute of Technology

Professor Raghupathy Sivakumar
Department of Electrical and Computer Engineering
Georgia Institute of Technology

Date Approved: 4 April 2013

To my parents and brother,

Joseph, Maryann and Aditya Muckaden,

the only reason I'm even capable of being at an institution as amazing as Georgia Tech.

ACKNOWLEDGEMENTS

Working with Dr. Nick Feamster has been the best academic experience of my life. Right from the Computer Networks class in the my first semester at Georgia Tech to being my advisor and guide for the MySpeedTest project, I have been in complete awe of Nick. Despite being extremely busy, since he's one of the most respected researchers in the Internet measurement community, he has always found the time to guide me and the project. Not just technically but in terms of manpower, direction and partnerships. I feel like every meeting with him, even if it last 10 minutes, gives me the opportunity to learn something new and grow not just as a student, but as a researcher. Despite my being an inexperienced Masters student, he has given me independent reign over several decisions and it means a lot to be entrusted with such responsibility. I have thoroughly enjoyed my work and cannot wait to continue working with him as a doctoral student.

I would like to thank Dr. Patrick Traynor and Dr. Raghupathy Sivakumar for agreeing to be on my committee. Dr. Traynor is an expert on cellular technology, and his being on my committee is an immense honor. Dr. Sivakumar is an expert on wireless technologies and I'm extremely glad that I got the opportunity to interact with someone as brilliant and well respected as him. I hope that I will have the opportunity to work with him closer in the future.

My deepest gratitude to Abhishek Jain and Young Kim, the original pioneers of the project. They began working on MySpeedTest before I was involved. Abhishek is the best coder I have ever met and I have learned a lot from him. Despite graduating from Georgia Tech he has always been available to me for any help with the project or my work in general. Young continues to be at Georgia Tech and I hope that he will take MySpeedTest to new heights in the period that I am away.

I would like to thank Dominic Hamon, Meredith Whitaker and Tiziana Refice for always being willing to help us and partnering with us in getting MySpeedTest on the Measurement Lab servers. They have been constantly supportive and have given us great insight in to working with mobile measurements and collecting and publishing data.

Thank you to Marshini Chetty and the Research ICT Africa team for believing in MySpeedTest and deploying as part of their research study in South Africa. This partnership gives us the opportunity to learn so much about cellular technology in developing areas and is an important step forward for MySpeedTest.

A very special thank you to Srikanth Sundaresan. Srikanth is a PhD student in the Networking and Telecommunications Lab. He has been guiding me ever since I been a student at Georgia Tech and before he remembers knowing me. Every time I need an opinion with work, or advice about life in general he has always been there for me. Even when he has paper deadlines to meet he will still find the time to give me a few pointers. I have greatly benefitted from his knowledge and experience. Thanks to all the other members of the GTNoise group - Joon, Sam, Bilal, Shuang, Abhinav, Sarthak, Swati, Shahbaz, Aman, Alfred and Mi Seon for being excellent lab-mates and always being willing to help. The environment in the lab has been fantastic and really conducive to learning.

Being a student at Georgia Tech can be a roller coaster ride, with several ups and downs. My girlfriend Sacchi has always been there for me to help me through the lows. Despite the 10 plus hour time difference she has stood by me and listened to hours of my complaining. My friends Melonia, Nitya, Gary, Salil and Abhiram too have helped through this ride. Sharing in all my accomplishments and failures whilst at Georgia Tech. I owe them more than I can express. Thanks to Aishwarya Babu for painstakingly helping with this thesis.

Lastly, I would like to thank my parents and brother - Joseph, MaryAnn and Aditya Muckaden. The pillars of my life and the reason that I am who I am. They have believed in me and supported me all my life. Their love and support has made this whole journey possible.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

As the number of mobile users and applications across the globe increases rapidly, there is an increasing need for application developers, cellular service providers and users to understand data network performance as seen from cellular devices. Cellular technology is advancing very rapidly; in a few years service providers have moved from 2G to 3G and now 4G services. Within the US itself each service provider offers users different service options based on the underlying technology. It is important while trying to understand cellular data network performance that these technologies behave differently and hence performance may vary significantly across them. In addition, unlike wire-line broadband Internet measurements, cellular data networks could be affected by several different factors. The battery level of the phone, the current operating system, the radio firmware and many other parameters could be a factor in the performance observed by users.

In order to study these effects, MySpeedTest was designed. MySpeedTest is an Android application which has been taking longitudinal measurements of cellular data networks. The application is on the Google Play Store and has been used by over 6000 users from across the globe over the last 5 months, resulting in over 3.5 million measurements. In the US alone it has 656 users, across all service providers and network technologies. The tool works in the background periodically and takes active and passive measurements of the cellular data network. In this thesis is presented the design of MySpeedTest and the performance results sliced up by the network technologies, for devices in the US. Statistical analysis is used to determine which of several metadata parameters associated with each measurement has an effect on the observed performance. The findings provide insights into the performance seen by

users in the US and the significant amount of variation over different cellular network technologies.

CHAPTER I

INTRODUCTION

There are over 6 billion cell phones in active use around the world today. Of these, over 1 billion use smart phones according to [23]. This number is likely to increase with users relying on cellular data networks not just for day-to-day activities like checking email and browsing the Internet but critical activities like mobile banking. The Google Play Store [15] alone had 600,000 applications with over 20 billion downloads as of 2012 [1]. As these numbers increase, it is essential that service providers and application developers better understand the performance that users see from their devices. In addition it is important to understand the effect that these performance numbers have on the applications that the user uses frequently. This information will enable application developers and service providers provide better quality applications and services in a highly competitive environment. Also, users can make smarter decisions about cellular devices, providers and applications based on the longitudinal performance seen in they're areas.

1.1 Problem Statement

Benchmarking cellular data network performance is not a straightforward task. It isn't as simple as conducting one-time speed tests since the context of the measurements are continuously changing. In addition, data caps imposed on users make frequent speed tests, which have high data usage costs, implausible and annoying to users. It is important to figure out the right set of tests which accurately characterize data network performance without being data heavy. Unlike with wire-line broadband measurements like [35] writing applications for cellular devices means that you cannot use standard throughput measuring tools like [16, 18, 21]. Hence, writing accurate measurements is also an important challenge. Also, it is important to capture the entire context of the measurement. Locations, device energy levels, network signal strength, service provider are few of the factors that could potentially affect the performance. In order to be able to accurately determine what affects the performance, gathering all the pertinent information is necessary.

Service providers do not have a clear picture on the kind of performance that users of their services are seeing. In order to establish a baseline performance level ”one-time” tests are not sufficient. Continual measurements are essential to determine on an average case what kind of performance is observed by the users. This is especially important as the cellular service market gets more competitive. It is even more unclear as to what parameters affect the performance seen from cell phones. If these can be determined then service providers are better informed as to how to improve services. For example, if it can be determined that one type of cell tower induces higher network latency then service providers can make investments to improve performance. This is also true for application developers. If developers are better informed of what conditions affect the network performance they can be most aware of the performance that the applications can expect and tailor their applications to function better. For example, knowing that in peak hours, the time to acquire a dedicated channel for data transmission is higher implies that applications like video calling applications can factor this in while determining the bit rate at which to send multimedia packets. Another important aspect is to determine how the performance affects users and the applications that they have installed. A large percentage of applications on cellular devices use the network. With streaming applications like Pandora and Netflix gaining prominence, it is interesting to gauge the extent to which performance affects the applications that users use often. For example, does higher latency or high packet delay variation mean that users do not use Netflix as frequently as users with lower latency? This will give keen insight into the user behavior towards applications that use the network and which performance numbers determine the users' favor.

1.2 MySpeedTest: A solution

To solve most of these problems MySpeedTest was built. It is an Android application that is on the Google Play Store. On installing the application the user is presented with the Terms and Conditions and the Privacy Policy. Once the user accepts the terms and conditions the application is installed on the users device. It then proceeds to take measurements in the background every 15 minutes.

These measurements include -
Active Probes that measure network parameters like latency, speed and loss rates.
Passive Measurements that measure the applications installed on the users device and the network usage in bytes of each of the applications.
Metadata which gets the context in which the particular measurement was taken. This is achieved by querying the Android API for various metadata parameters.

MySpeedTest introduces a unique set of tests and test scheduling framework. The details of the tests and scheduling framework can be found in Chapter 2. On completing the test, the data is transmitted to Georgia Tech servers for storage. The data storage details also can be found in Chapter 2.

1.2.1 Design Choices

1.2.1.1 iOS versus Android versus Both

MySpeedTest is an Android only application. The most important requirement for MySpeedTest is to be able to take periodic background measurements. Unlike Android OS, the iOS platform does not provide a simple method to achieve this. Thereby making the design of the application for iOS far more complex. In addition, according to the Gartner report for Quarter 2 made in August, 64.1% were powered by Android and 18.1% were powered by iOS. Also, the Android operating system is supported by many different cellular devices vendors across the world, allowing for measurements from a variety of devices with different features. This increases the richness of the data set. For these reasons, MySpeedTest was chosen to be an Android only application.

1.2.1.2 Rooted devices versus Play Store application

Traditionally there have been 2 approaches to measurements from Android phones because of the nature of the Android subsystem. In the first approach researchers write applications for the Google Play Store which are allow for a large number of users from all across the world. The limitation of this approach is that the data that may be collected and the tools that may be used by an application installed, is limited. This limitation is imposed by the user not being a "root access user". Android consists of a kernel based on Linux 2.6 and from Android 4.0, version 3.x which requires root access to perform certain privileged tasks. Therefore this approach does not allow for researchers to gather packet level details during measurements.

The other approach is to use a small number (usually 5 - 20) of phones which have been "rooted", that is the user of the phone has root access to the device kernel. This enables far more detailed measurements since packet level details may be collected. Also tools such as iperf [16] maybe cross compiled for ARM and run from the Android device. The limitation with this approach is that the environment in which measurements are gathered is very limited. The devices are generally handled by members of the same research group or institute. The number of devices and measurements is also limited.

MySpeedTest uses the first approach. Previous studies that employ this method are [29, 32, 33]. The tests designed were extremely simple and involved low data costs. They are run from within an application without requiring packet level details but still being indicative enough to understand network performance. It has thus been possible to gather many millions measurements from 165 countries across the world. The large number of data points is essential to be able to make inferences regarding performance across different service providers and in different locations.

1.2.2 Unique Contributions of MySpeedTest

Junxian et al. developed a longitudinal cellular data network measurement tool Mobiperf [30]. Besides Mobiperf, Ookla's speedtest.net [22] has a mobile client that is fairly popular. The FCC too has been working with SamKnows to develop a mobile application for large scale deployment across the US. However MySpeedTest has several unique features that are tailored to answer the questions presented in the Problem Statement above. The following sections elaborate on these contributions.

1.2.2.1 Active Measurements

[30] focused on TCP throughput and TCP round trip latency as the primary indicators of network performance for large scale longitudinal studies of cellular data networks. Newer tests have added HTTP performance measurements and traceroutes. The application being developed for the FCC also takes similar measurements. While these are interesting indicators of performance MySpeedTest uses a different set of tests aimed at establishing a baseline performance level. The active tests used by MySpeedTest are - TCP single threaded throughput (similar to [30]), pings to geographically distributed servers and popular content providers, IP packet delay variation over UDP, UDP loss rates and warm up tests aimed at measuring the time it takes to acquire a dedicated channel for data transmission. The full details of the tests can be found in Chapter 2. The aim with the tests was to minimize data usage while still getting useful information about the network.

The TCP throughput is always an important indicator of performance and often the metric that users are most interested in. Service providers often make speed guarantees and this metric helps users decide if they are getting what they paid for. This test however is similar to those implemented by previous works.

The longitudinal ping tests to the two most popular content websites - Facebook and Google are unique to MySpeedTest. This is indicative of the placement of content and how it varies across locations and service providers. It also establishes the baseline achievable performance to these websites. Xu et al. presented data characterizing cellular data network infrastructure in [36]. It was found that routing in cellular data networks is very restrictive because there are very few gateway GPRS support nodes (GGSN's) that act as the gateway between GPRS networks and external IP networks. It was suggested that in order to reduce the latency by upto 50% content should be placed close to the GGSN's. As a follow up to this research, MySpeedTest sends ping packets to the 2 most popular content websites are used to determine the relative placement of content by different service providers in different countries. Pings are also sent to 3 other servers - two in the US and one in Europe.

The 2 UDP based tests are extremely light-weight in terms of data usage and are designed in accordance with RFC 2680 [24] and RFC 3393 [27]. UDP based loss gives a good idea about packet delivery rates in cellular networks. As explained in Section 1.2.2 the choice to not use rooted phones means that if a TCP based test was used instead, the retransmission information would not be available. There has been work on retransmissions over TCP using a set of 20 rooted phones in [30]. However in order to get a better sense of how many packets are dropped on the cellular network especially in the case of low signal strength on a continual it was important to use a protocol that does not do automatic retransmission. Hence MySpeedTest uses a UDP loss based test. With streaming applications like Pandora and Netflix becoming popular in the US, IP packet delay variation, often refered to as jitter is becoming an extremely important metric. With high queue latency in networks today, variable amounts of delay have been introduced. This could be especially true in cellular network where researchers have found notoriously large buffers. MySpeedTest is the first measurement application for cellular devices that conducts active probes of the IP packet delay variation.

Previous studies [25, 34] have studied the time that it takes to acquire and release a dedicated channel in cellular data networks. It is an extremely important metric because it directly relates to the energy efficiency of the device. Researchers have argued that the current methods used in cellular systems is inefficient in terms of energy usage due to the time it takes to release the dedicated channel. The argument however is that the time taken to acquire the dedicated might not be constant depending on the network technology being used. The state diagram from channel usage is different for 3G and 4G based technologies. However, within the different types of 3G technologies itself there maybe warm up times. The warm up time could be affected by a number of parameters including the type of device, the location of the device, the time of day etc, which is important to investigate. MySpeedTest does the first longitudinal measurement of the warm up time to infer the parameters which affect this metric and suggest improvements.

1.2.2.2 Metadata - In context measurements

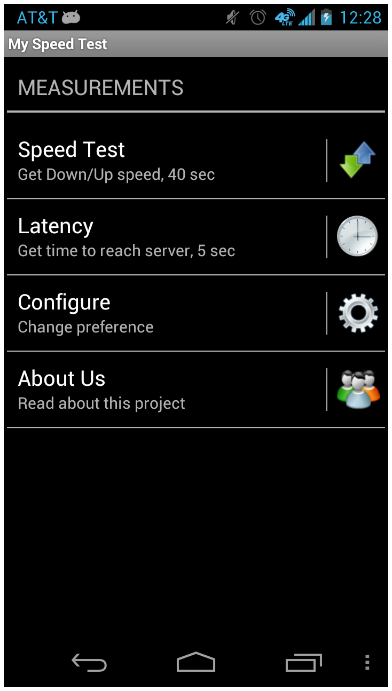
Previous tools for measuring cellular data network performance have neglected to collect metadata that entirely sums up the context of the measurement. Cellular data networks measurements vary greatly from fixed line measurements because there are so many more parameters that could affect performance metrics. Measurement tools like [30] have collected metadata pertaining to the network such as the signal strength. This thesis argues that several other parameters could affect performance. The Android API is very rich in the data that can be queried from the device. For a full list of metadata parameters gathered please refer to Chapter 2. MySpeedTest collects far more metadata parameters than any other tool. The argument is that the by clearly defining the context of the measurement, better conclusions can be made. For example, it is possible that low battery level may cause the performance to degrade. This information will be highly valuable to application developers who may tailor applications to minimize network usage while the performance is known to be bad. In addition, gathering the specific network type enables slicing up the data and performance by the network type. This thesis argues that for a fixed set of conditions different network types perform significantly differently. In the past performance results have been presented without considering the specific network type of the measurement. Chapter 3 presents an in depth discussion of the network types prevalent in the US. This thesis present all results divided by network type for a more accurate understanding of the performance.

1.2.2.3 Passive Measurements

Perhaps the most important question to be answered is if the performance observed affects user behavior towards applications at all. If this is so then what metrics affect user behavior the most and how. This information would enable service providers to narrow down on the key metrics that they need to improve in order to ensure user satisfaction. Studies such as [26] have studied application usage in the past. MySpeedTest, measures the network usage of each of the applications. These values are noted every 15 minutes, along with the active measurements. This gives a unique view into the user behavior as affected by the network performance. It enables pinpointing exactly which performance metrics largely affect user behaviour. This thesis finds that throughput isn't the only factor which affects user behavior. In fact, latency and IPDV play an equally important role. Chapter 3 highlights details of the passive measurements.

1.2.2.4 User Interface

There are close to a million applications [1] on the Google Play Store. An important part of getting a large number of users for an application across the globe is to have a user interface that is unique and easy to use. MySpeedTest is proof of this concept. The user interface of MySpeedTest went through several iterations. The unique feature offered by MySpeedTest is a simple user interface and a graph plotting engine. It uses AChartEngine which is an freely available graphing software for Android. The active measurements are stored in an SQLLite database which is part of the Android subsystem. This enables plotting of graphs using longitudinal data. Which makes it possible for users to constantly monitor the performance the corresponding device is seeing. Previous Internet broadband studies as well as cellular data network studies have made this available to users using a website url and authentication. Storing data efficiently on the device and plotting graphs on the click of the button make MySpeedTest far more appealing to users.



MEASUREMENTS

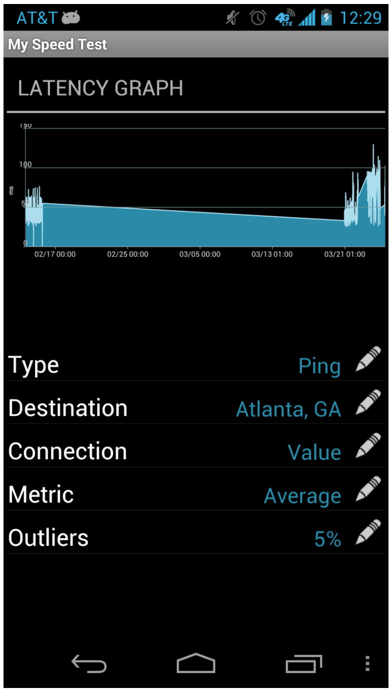
Speed Test
Get Down/Up speed, 40 sec

Latency
Get time to reach server, 5 sec

Configure
Change preference

About Us
Read about this project

Figure 1: The main test screen of MySpeedTest



Type Ping

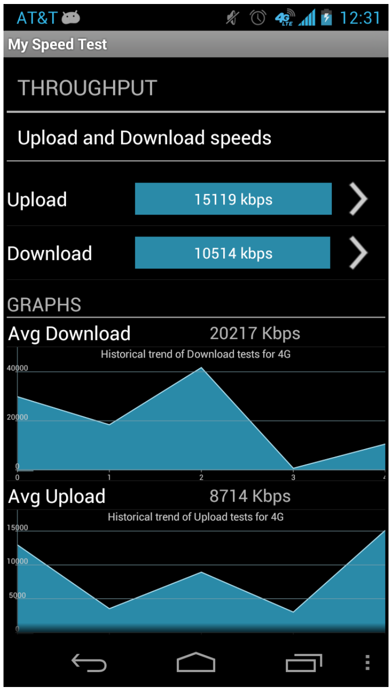
Destination Atlanta, GA

Connection Value

Metric Average

Outliers 5%

Figure 2: Graph displaying the latency to one server over the last month



My Speed Test

THROUGHPUT

Upload and Download speeds

Upload 15119 kbps

Download 10514 kbps

GRAPHS

Avg Download 20217 Kbps

Historical trend of Download tests for 4G

Avg Upload 8714 Kbps

Historical trend of Upload tests for 4G

Figure 3: Graph displaying results for last 5 throughput measurements

CHAPTER II

MYSPEEDTEST: IMPLEMENTATION DETAILS

2.1 Test details

MySpeedTest is available for download in the Google Play Store. Once a user installs the application and agrees to the Terms and Conditions, MySpeedTest conducts active network probes, takes passive measurements and collects metadata periodically in the background. Below is described each of the tests.

2.1.1 Active probes

Active measurements give insight into the network performance conditions. Most of these tests, except for the speed test are conducted periodically in the background. The periodicity of the tests varies with each test depending primarily on the data usage of each test. The active tests measure against a server located at Georgia Tech. The server is named ruggles.

2.1.1.1 Speed Test

The speed test measures single-threaded TCP throughput in both the uplink and downlink directions separately. In order to capture the variation of speeds prevalent on a cellular network and ensure that TCP slow start does not affect the measurement, these tests last 20 seconds each.

In the uplink test, the client first creates a single threaded TCP connection to the ruggles server. Once the handshake is complete the measurement is ready to begin. A random string of size equal to the maximum size that can fit in the payload without fragmentation is generated. For 20 seconds, packets are generated and sent to the server using this string. The rate of sending packets is limited by 2 factors - the network and the processor speed of the device. These are both characteristics that we intend to measure. It should be noted that this test does not measure the bandwidth of the link, but rather the throughput achievable from the device over TCP. The server maintains a count of the amount of data received. On receiving the FIN packet, the server notes the time taken to receive the data and computes the throughput. The calculated value is then sent over to the client for display. The calculation is done at the receiver to overcome the effects of buffering, in order to get a more accurate measurement.

The downlink test is very similar to the uplink test. In this case the string is generated at the server and the calculation is done at the client end. The duration of the test is the same.

The speed tests are very high on data usage. Especially on faster networks today that send data at a very high rate. One optimization could be to tailor the duration of the test to the type of network being used or use fixed size data instead of fixed duration. The downside with fixed sized data is that on slower networks this could cause the test to be very long. This will be an important consideration in further iterations of the tool. Because the test is so data intensive; in the case of LTE networks could go up to 7.5 MB; and data is very expensive in certain countries, this test is never conducted in the background. It is only conducted if initiated by the user. This limits the richness of the data that can be collected with speed tests. However, since majority of the research in the past has been with speed tests and a few hundred measurements are still conducted everyday by users of MySpeedTest, this design was not changed.

2.1.1.2 Latency Tests

This test sends a set of 5 ICMP ping packets to 5 different servers. There are 2 sets of servers used - content servers and BISMark servers set up by the GTNoise group at Georgia Tech. The servers are - Facebook, Google, Atlanta, Oakland and Napoli, Italy. Since ping packets are only 64 bytes, we can be sure that they do not cause congestion. Hence, for the sake of quick response time when the user runs the test, the set of pings to the servers are conducted in parallel using 5 different threads. The min, max, average and standard deviation is recorded for each server.

We also send pings to the first hop "pingable" hop to measure last mile latency. This is achieved by setting the Time to Live (TTL) field as 1. It was observed from our initial measurements that very often the first hop is not pingable. Hence successive hops are pinged till the first pingable hop is identified. Sundaresan et al. in [35] found that last-mile latency was often a dominating factor in determining end-user performance in home networks. This can be even more pronounced in cellular networks due to the nature of the last hop wireless link.

In the next iteration of the application, the list of servers to ping will be a dynamic list retrieved from the ruggles server at Georgia Tech before each latency test. That will enable determining the set of servers on this list based on the most interesting questions to be answered at the current time. Also, the number of pings to each server will be increased from 5 to at least 10 or 15. This is because through the results we have determined that 5 pings show very high variability and hence to find a more accurate reading, the number of measurements should be increased.

This test is done periodically in the background every 15 minutes in addition to when the user manually begins a test.

2.1.1.3 IP Packet Delay Variation Test

According to RFC 3393 [27], IP Packet delay variation is defined for 2 packets from the server to the client, selected by some function F, as the difference in the one-way delay between the server and client at time T2 and the one-way delay between server and client at time T1. Where T1 is time when the server sent the first of the 2 packets

```
  I1   P(i)        P(j)                    P(k)                        I2

MP1  |---------------------------------------------------------------------|
         |\         |\                       |\
         | \        | \                      | \
         |  \       |  \                     |  \
         |   \      |   \                    |   \
         |dTi \     |dTj \                   |dTk \
         |<--->v    |<--->v                  |<--->v

MP2  |---------------------------------------------------------------------|

 I1           P(i)         P(j)                   P(k)                 I2
```

Figure 4: Figure depicts how IPDV works. IPDV = dTj - dTi

and T2 is the time that the server sent the second packet.

Multiple such readings are recorded per measurement. An interval of 10 packets is chosen for each reading. The selection function F is such that it randomly chooses A stream of 64 byte UDP packets is generated at the server. The payload contains the timestamp at the server as well as a the sequence number of the packet. The stream is generated using a Poisson Sampling methodology. The reason for using Poisson's sampling is that it results in an exponential distribution of samples for each interval which makes the arrival of samples "unpredictable" and the sampling is asymptotically unbiased even if the sampling affects the network state [31]. In each interval, the packets sampled is varied to introduce as much unpredictability as possible. The server sends a stream of 500 packets, hence 50 readings are taken at the client. All these readings are sent to the server for storage. The data rate of sending packets is very low (less than 1 Kbps) so that the test does not induce congestion and contribute to variation in inter packet delay.

Since the test involves sending 500 packets, each test uses 32 Kbytes of data. Therefore this test is conducted only once every 2 hours. In the current iteration of MySpeedTest, a front end has not been developed for this test. It only runs in the background on a periodic basis.

2.1.1.4 Loss Test

The loss test is designed according RFC 2680 - One-Way Loss Test [24]. The loss is measured in the direction from server to client. The RFC defines a loss event to occur if the delay between sending the packet at the server and receiving it at the client is above a fixed threshold. We set this threshold at 25 seconds.

According to the RFC, the stream for this test should also be generated by a Poisson Sampling methodology. The sample is defined according to a Poisson process to avoid self-synchronization and to make the sample as statistically unbiased as possible. Hence the Poisson's process is used to generate the delay measurement packets. The packets will not reach the client according to the Poisson's distribution because of the variable delay induced by the network.

MySpeedTest uses the same packets in the IPDV test for the loss test to further reduce the data usage. The packets contain timestamps from when they are sent out at the server. The timestamp is noted when the packets arrive at the client. The difference between the 2 timestamps are recorded. If this exceeds the threshold then the packet is marked as lost. If the packet does not arrive for threshold number of seconds then it is also marked as lost. There is the issue of synchronization of clocks on the server and the client. This is why the threshold is set at a high value to minimize the effect of skew and drift in the clocks.

Since the test is run in conjunction with the IPDV test, it also runs once every 2 hours. It runs only in the background and does not currently have a front end.

2.1.1.5 Warmup Measurement

This measurement is used to find out how much time it takes to acquire a dedicated channel for data communication. The test sends out a sequence of 40 ping packets to the ruggles server; one packet every 200 milliseconds. The round trip time is noted for each packet. This test is the first one conducted on every periodic set of measurements. This is because the other tests need to be conducted when the device already has a dedicated channel. Once the dedicated channel is achieved, the round trip times decrease considerably. As will be showed in Chapter 4, the round trip time drops by almost 80 percent.

As mentioned this test is conducted every 15 minutes in the background and is the first of the periodic set of tests conducted.

2.1.2 Passive Measurements

MySpeedTest records data pertaining to data usage of applications installed on the phone. This data is made available through a simple procedure call to the Android API. The following data is collected-

The total number of bytes sent and received by the device since it was last powered on.

For each application currently installed on the device:

- Package name of the application. For example com.android.browser.
- Total number of bytes transmitted since the device was last powered on.
- The number of bytes received since the device was last powered on.
- Is the application running at the time of the measurement.

These passive measurements are recorded every 15 minutes. The challenge is to maintain a running count of the data usage of an application even after the device is restarted. We use a simple method to get a fairly accurate measurement even if the device is restarted within the 15 minute time period between successive tests. This is done by checking if the data usage of the application is less than it was at the time of the previous measurement. If this is true then we add this new value to the maintained running count.

2.1.3 Metadata collection

Metadata is collected along with every measurement in order to establish the complete context of the measurement. This metadata is collected by querying the Android API. The following metadata is collected.

1. local time
2. Is Manual - Was the test started by the user or was it a scheduled periodic test.
3. Network Information -
    - Network Type: Mobile 3G/ Mobile 2G/ Wifi
    - Cell Tower Id: Each cell tower has a unique ID
    - Hashed deviceid: The IMEI number of the device. This value is hashed on the device itself to ensure privacy
    - Signal Strength: The network signal strength. On Android devices this is measurement on a 32 point scale (0-31)
    - Cell Type : GSM or CDMA
    - Connection Type : Mobile 2G / 3G / 4G
    - Latitute and Longitude of the Base Station
    - Network Operator Id
    - CDMA System ID
    - CDMA Network ID
    - Data State: CONNECTED / CONNECTING / DISCONNECTED / SUSPENDED
    - Data Activity: DATA-ACTIVITY-IN / DATA-ACTIVITY-OUT/ DATA-ACTIVITY-NONE / DATA-ACTIVITY-DORMANT

4. Battery -
    - Battery level: The battery level is measured on a scale of 100
    - Battery technology: Li-Ion (for example)
    - Is Plugged: Is the device currently plugged into a charger
    - Health: The health level of the battery
    - Voltage: Current voltage level of the battery
    - Temperature: The current temperature of the battery
5. Device Specific Info -
    - Billing cycle / Data cap : Provided by user.
    - Phone brand and manufacturer
    - Android version
    - Radio firmware version
    - Country
    - Hashed phone number
    - Software Version number for the device, for example, the IMEI/SV for GSM phones
    - Service provider
    - SIM card state and serial number

2.2 A note on test scheduling

For the background tests, if the device is not being currently used, acquiring a dedicated channel increases the power consumption from around 700mW to around 1500mW according to Feng Qian et al. in [34]. In order to ensure that all the tests mentioned above do not wake up the radio when the device is on low battery, the battery level and radio state is checked before beginning every test. If the battery level is low (less than 10 but greater than 5 out of 100) we will not perform the test unless the device radio is already in the connected state. If the battery is very low (less than 5) we will not perform tests at all.

2.3 Data collection and storage backend

The data is collected and stored on the ruggles server at Georgia Tech. After each measurement the data is compiled into JSON format and put into a queue for sending to the server. The data communication is achieved via an HTTP POST request. If for any reason the data communication to the server fails, the message remains in the queue for sending at a later time. In order to ensure the radio is not unneccesarily woken up, the communication is only re-tried when the next measurement takes place.

The ruggles server runs a python-django backend service on an apache server. A postgresql relational database is designed using models in django. Django manages receiving the data, parsing for errors and inserting into the database. Each request from the client is treated as a single transaction in the database. This enables very quick and efficient inserts into the database.

In addition to the database maintained on the ruggles server, each device also maintains a database using SQLLite. The measurement results from the speed and latency tests described in section 2.1.1.1 and 2.1.1.2 are stored in this database. The database on the device enables quick querying and plotting of graphs for the users to observe the longitudinal trends on their data network.

2.4 Challenges

During the implementation of MySpeedTest several challenges were faced. The most important are listed below. Some of these continue to be issues and we are in the process of working on partnerships to overcome these challenges.

2.4.1 Absense of standard throughput measurement tools

Most standard throughput and bandwidth measurement tools like those described in [18, 16, 21] require root access. In addition cross compiling these tools for ARM processors is a complex tasks. The busybox toolkit for ARM processors has a list of tools cross compiled for use on Android device. However, installing the toolbox requires root access.

Hence it was necessary to implement a speed testing tool. The implementation of the tool is described in Section 2.1.1.1. This however, is an unsophisticated method that roughly approximates the throughput of the link at the current time. While, this is interesting enough, it would be beneficial to be able to use a standard tool that has gained acceptance. In addition, validating measurements becomes another even more important task. The best method for validation of speed measurements was to compare the results with similar such tools that do speed measurements such as [30, 22]. While this method is generally accepted by other researchers in the space, we intend to make more effort on establishing a standard for speed test measurements.

2.4.2 Location of the measurement server

Currently the only measurement server is the ruggles server located at Georgia Tech in Atlanta. Because the measurement uses TCP, high round trip times can greatly affect the measured throughput. For devices in the US, this round trip time is off the order of 60 milliseconds on an average case for UMTS or EVDO networks, which are the most prevalent. This is tolerable and gives fairly accurate readings. However, for devices in other countries which are geographically further away this is a major concern. For example in South Africa where MySpeedTest has almost 200 users, the round trip time ranges between 300 - 400 milliseconds on an HSPA network. Assuming a window size of 65535 and a latency of 300 milliseconds, the maximum achievable throughput over TCP is 1747 kbps. This is much lower than the actual throughput.

For this reason, this thesis does not present results for speed tests. We are currently partnering with Google and the Measurement Lab team [17] to use they're globally distributed measurement servers. This is reduce round trip considerably and will yield far more accurate measurements.

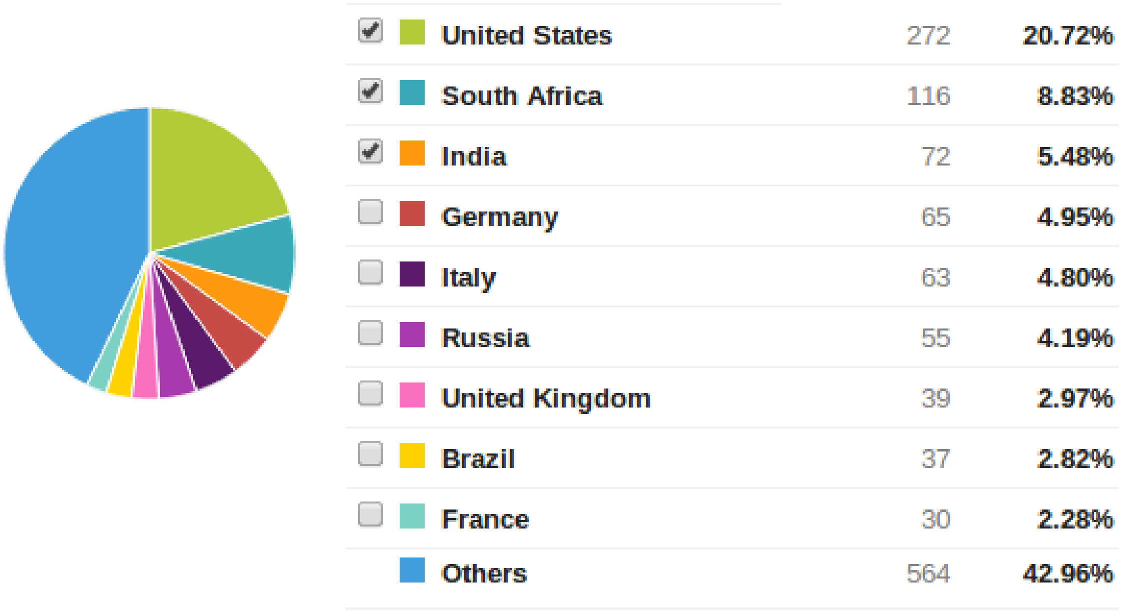
2.4.3 User privacy

User privacy is a very important issue with measurement studies that involve mobile devices. The FCC is also involved in policy documents to identify what information collected from mobile devices is potentially risky. A standards document for mobile data collection is being worked on by policymakers from across the US.

There are 2 parameters that we collect that we deemed as possibly personally identifiable information (PII). These were the IMEI number of the device and the phone number. Hence both these values are hashed using a widely accepted encryption algorithm on the device before being transmitted to the server. Therefore this information is not available to us either.

Researchers in Internet measurements have argued the cause of making data sets publicly available for other researchers to study. While the data collected by MySpeedTest will be made publicly accessible there exist several privacy issues at the current time. Firstly, a combination of metadata parameters that we collect could serve to identify a user. For example a combination of the device type and location. In addition, not being able to identify a user, but the users locations could be potentially dangerous. For these reasons we do not yet collect fine grained GPS locations. Before the data can be released it will need to be made adequately anonymous and aggregated.

Secondly, the application usage data is suggestive of user behavior. This data is extremely sensitive. We have worked with the institutional review board at Georgia Tech to ensure that the data that the passive measurement data that we will collect



☑	■	**United States**	272	**20.72%**
☑	■	**South Africa**	116	**8.83%**
☑	■	**India**	72	**5.48%**
☐	■	**Germany**	65	**4.95%**
☐	■	**Italy**	63	**4.80%**
☐	■	**Russia**	55	**4.19%**
☐	■	**United Kingdom**	39	**2.97%**
☐	■	**Brazil**	37	**2.82%**
☐	■	**France**	30	**2.28%**
	■	**Others**	564	**42.96%**

Figure 5: Distribution of active installs of MySpeedTest by country

does not constitute more than minimum risk to the participants of the study. This thesis will not present any information with regards to the passive measurements. Also, this information will not be made publicly available.

2.5 Deployment details

This section highlights some of the important deployment statistics of MySpeedTest. The diversity in the devices, countries, carriers and network types contributes to the importance of the measurement information.

- The tool is deployed in 138 different countries. The Figure 4 below is taken from [14] and shows the distribution of countries. The Play Store does not show Iran which also has a large number of users of the application. Our recent partnership with Research ICT Africa is the reason for the large deployment in South Africa.

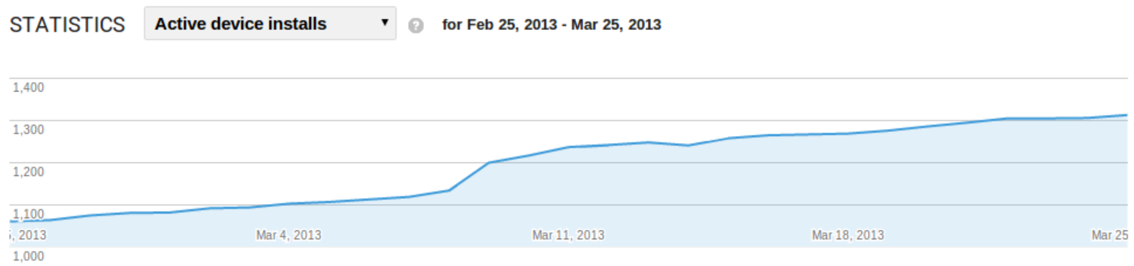


Figure 6: Active installations of MySpeedTest from Feb 25, 2013 to March 25, 2013

- At the time of writing this thesis, there were over 6000 total installs of the application. Figure 5 is taken from [14] and shows the number of active installs over the last month.

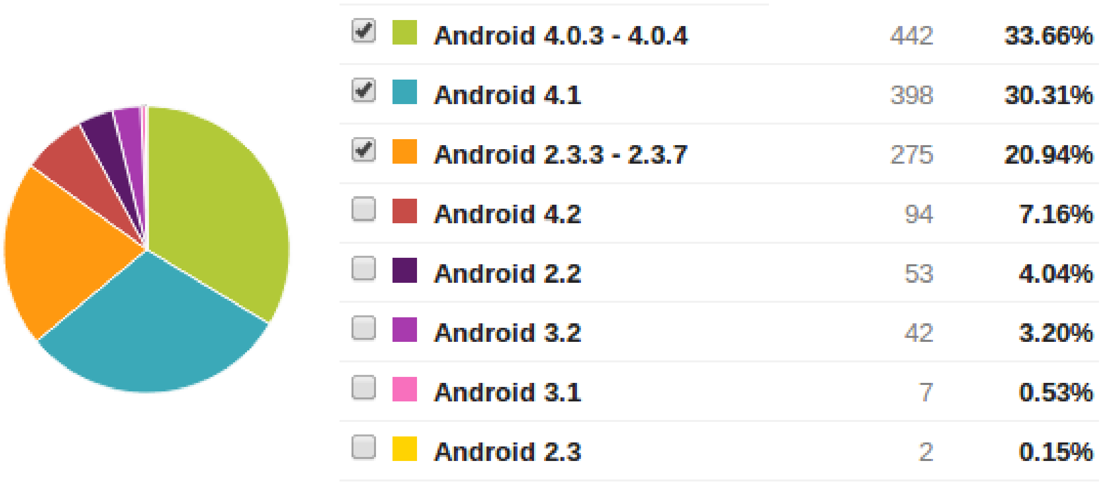
- The tool is deployed on 831 different types of devices running the Android operating system.

- The application spans all Android version from 2.2 - 4.2. Figure 8 is taken from [14] and illustrates the distribution by Android version for currently active installs.

- The application has users over the 4 major carriers in the US. Figure 7 is taken from [14] and shows the distribution by carrier of active installs of the application. Table 1 below shows the number of measurements per carrier for the 4 major US carriers - Verizon Wireless, AT&T, Sprint and T-Mobile-US.

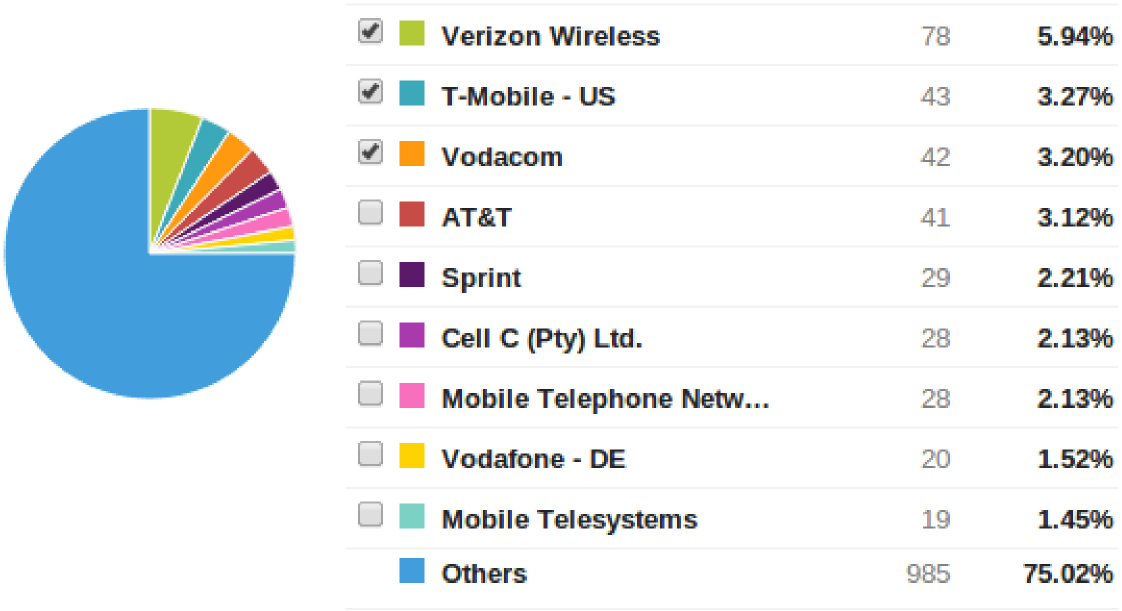
- Table 2 belows shows the number of measurements for devices in the US by the different network types. This is key in the analsis presented Chapter 5 of this thesis.

- At the time of writing this thesis there were 3.47 million measurements totally, of which 0.6 million are for devices in the US.



☑	🟩	**Android 4.0.3 - 4.0.4**	442	**33.66%**
☑	🟦	**Android 4.1**	398	**30.31%**
☑	🟧	**Android 2.3.3 - 2.3.7**	275	**20.94%**
☐	🟥	**Android 4.2**	94	**7.16%**
☐	🟪	**Android 2.2**	53	**4.04%**
☐	🟪	**Android 3.2**	42	**3.20%**
☐	🟥	**Android 3.1**	7	**0.53%**
☐	🟨	**Android 2.3**	2	**0.15%**

Figure 7: Distribution of active installs of MySpeedTest by Android Version



☑	🟩	**Verizon Wireless**	78	**5.94%**
☑	🟦	**T-Mobile - US**	43	**3.27%**
☑	🟧	**Vodacom**	42	**3.20%**
☐	🟥	**AT&T**	41	**3.12%**
☐	🟪	**Sprint**	29	**2.21%**
☐	🟪	**Cell C (Pty) Ltd.**	28	**2.13%**
☐	🟥	**Mobile Telephone Netw...**	28	**2.13%**
☐	🟨	**Vodafone - DE**	20	**1.52%**
☐	🟩	**Mobile Telesystems**	19	**1.45%**
	🟦	**Others**	985	**75.02%**

Figure 8: Distribution of active installs of MySpeedTest by carrier

Table 1: Shows the number of measurements per carrier for the top 4 carriers in the US

Carrier	Number of measurements
AT&T	158591
T-Mobile	149341
Verizon Wireless	146278
Sprint	119087

Table 2: Shows the number of measurements per network type for devices in the US

Network Type	Number of measurements
LTE	189174
EVDO_A	147184
UMTS	112337
HSDPA	54010
HSPA	42996
EDGE	23107
1xRTT	12030
GPRS	4634
WIMAX	4556
EVDO_0	1909

CHAPTER III

NETWORK TECHNOLOGIES: AN OVERVIEW

There are 2 major standards for deployment of cellular technologies - GSM and CDMA. Within each of these standards there are different technologies being implemented. This section provides a very high level overview of these technologies in order to help the reader get a better understand of the results presented in the next chapter.

A cellular network is characterized by the division of the mobile service area into regular shaped cells, generally hexagonal. Each of these cells is assigned a set of frequencies and a base station. Figure 9 below represents the structure of a cellular network.

The first generation (1G) of mobile systems refers to the analog cellular technology. 1G, based on circuit switching, supported voice calls and data, which only in the form of analog signals, could be exchanged between phones. This system was outmoded by the digital system because it was incapable of supporting the required wide spectrum and had limitations in terms of security. [2]

The succeeding generations are second generation (2G) digital cellular networks, third generation (3G) broadband data services and fourth generation (4G) native-IP networks.

The second generation technology comprises two competing standards, GSM and CDMA. These two circuit switched standards evolved into their 2.5/3G counterparts which work on packet switching. [3]

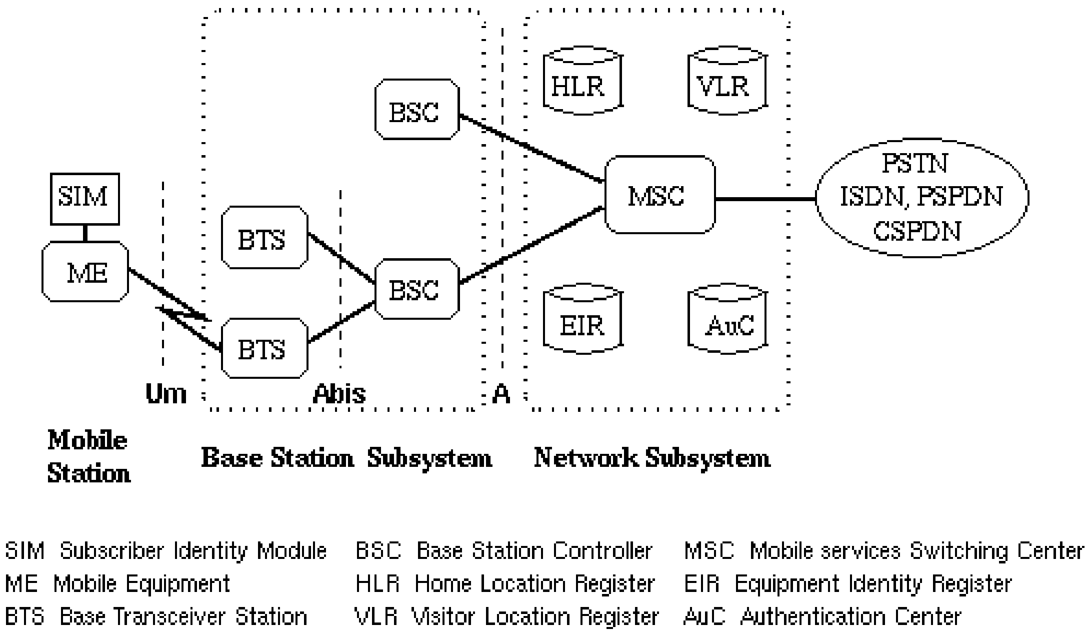


Figure 9: An overview of cellular network infrastructure

3.0.1 GSM

GSM (Global System for Mobile Communications, originally Groupe Spcial Mobile) is the European standard for digital mobile communications which is now the most widely used. It uses a combination of TDMA(Time Division Multiple Access) and FDMA(Frequency Division Multiple Access) technologies for delivering the service. Figure 10 shows the GSM network infrastructure. [4, 5]

The ability to switch handsets, good sound quality and encryption are some of the aspects that make the GSM standard most popular worldwide.

A key feature of GSM is the Subscriber Identity Module, or SIM card. The SIM is a detachable smart card which stores the users subscription information. It is also equipped with capability to store a fixed number of contact numbers. This allows the user to retain information after switching handsets. Most important the user may change operators while retaining the handset simply by changing the SIM. This is probably the most compelling reason for the popularity of GSM worldwide. It enables users moving countries to retain the same device but switch over to a different carrier

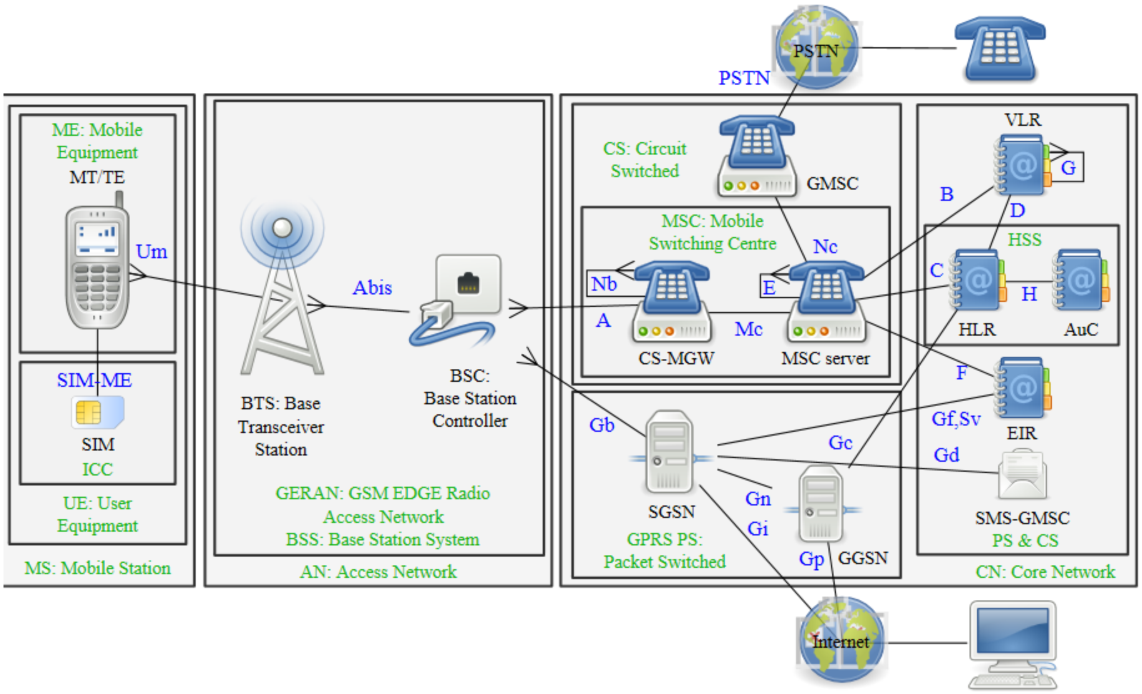


Figure 10: GSM network infrastructure

provided the device is compatible with frequency band being used by the carriers. The low data transfer speeds(9.6 kbps) of the 2G GSM technology and the increasing demand for mobile web applications led to the development of the 2.5G, a step halfway between 2G and 3G. The move was made from circuit switching to GPRS (General Packet Radio Service), a 2.5G technology, which is based on packet switching. It is able to achieve throughput rates of up to 40 kbps. This technology allows people to access the web from almost anywhere on their mobile. [5]

EDGE(Enhanced Data rates for GSM Evolution), also known as EGPRS(Enhanced GPRS), is an upgrade on the existing GSM technology with a capacity three times that of GSM/GPRS. [6]

UMTS(Universal Mobile Telecommunications System) is a 3G(third generation) technology based on the GSM standard. This technology is capable of providing high data rates and thus delivering high speed Internet services. UMTS uses W-CDMA(Wideband Code Division Multiple Access) for radio access. This is different from the set of

Table 3: Summary of GSM standards

	2G	2.5G	3G	3.5G	4G
GSM Standards	GSM	GPRS,EDGE	UMTS	HSDPA,HSUPA,HSPA	HSPA+,LTE

CDMA standards. Users of deployed in UMTS networks can expect data rates of up to 384 kbits/sec. [7]

HSDPA(High-Speed Downlink Packet Access), an upgrade to the UMTS networks, was developed to cater to the need for improved data download services. HSUPA(High-Speed Uplink Packet Access) provides enhanced upload speeds. HSDPA added a new transport layer channel, High-Speed Downlink Shared Channel (HS-DSCH), to UMTS. Besides improving data rates, HSDPA also decreases latency and so the round trip time for applications. HSDPA deployments are supposed to offer speeds upto 7.2Mbps but usually gives speed of about about 3 - 4 Mbps. [8]

HSDPA and HSUPA are together referred to as HSPA(High-Speed Packet Access) a technology standardized by 3GPP (Third Generation Partnership project, a group of six telecommunications standard development organizations).

HSPA+ or evolved HSPA is an improvement of the HSPA network. Theoretically, it is supposed to offer speeds up 168 Mbits/sec in the downlink direction and 22 Mbits/sec in the uplink direction. The actually speed achieved by the user are typically lower. These speeds are possible due to the use of a multiple-antenna technique known as MIMO (for ”multiple-input and multiple-output”) and higher order modulation (64QAM) or combining multiple cells into one with a technique known as Dual-Cell HSDPA. HSPA+ has been dubbed as cost effective because it upgrades the existing 3G network and provides a method for telecom operators to migrate towards 4G speeds without deploying a new radio interface. [9]

Table 3 summarizes the different GSM based technologies.

Table 4: Summary of CDMA standards

	2G	2.5G	3G
CDMA Standards	CDMA(cdmaOne)	CDMA2000 1x (1xRTT)	CDMA2000 1xEV-DO

3.0.2 CDMA

CDMA(Code Division Multiple Access) is a wireless standard based on the channel access method CDMA which uses spread spectrum technology. Several users can share the same frequencies and be active at all times. CDMA mobile phones are most popular in North America but not so much in many other countries, the reason being the inability to switch handsets and carriers.

CDMA (cdmaOne or IS-95) is the 2G competitor of the 2G GSM standard and is based on circuit switching. CDMA2000 1x (1xRTT), the 2.5G counterpart and its successors work on packet switching. he 1X standard supports packet data speeds of up to 153 kbit/s with real world data transmission averaging 80100 kbit/s in most commercial applications. The 1X standard supports packet data speeds of up to 153 kbit/s with real world data transmission averaging 80100 kbit/s in most commercial applications.

The 3G version of CDMA technology is the CDMA2000 1xEV-DO (EVDO) (Evolution Data Optimized). It uses code division multiple access (CDMA) as well as time division multiple access (TDMA) to maximize both individual user's throughput and the overall system throughput. [10]

Table 4 summarizes CDMA standards.

3.0.3 LTE

LTE(Long Term Evolution) also referred to as 4G LTE is a high-speed wireless communication standard based on GPRS/EDGE and UMTS/HSPA. It provides higher

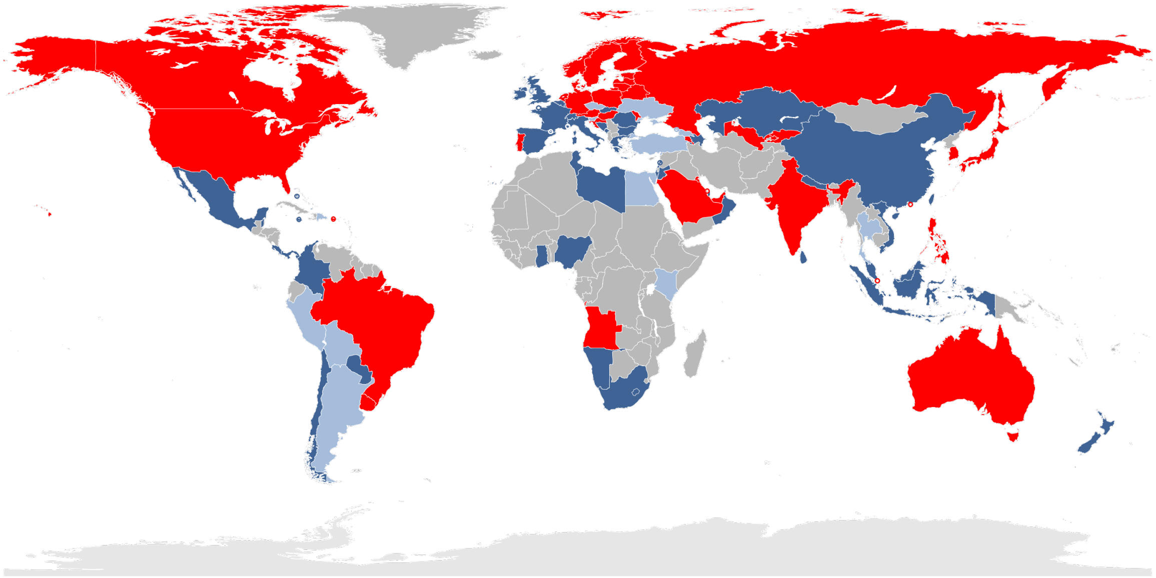


Figure 11: Global deployment of LTE networks

data rates and lower latency. LTE is the first global mobile phone standard since carriers using both GSM and CDMA standards have launched LTE networks across the globe. Figure 11 is taken from [11] and shows the global deployment of LTE networks. LTE is supposed to offer peak download rates up to 299.6 Mbit/s and upload rates up to 75.4 Mbit/s depending on the user equipment.

CHAPTER IV

RESULTS

This section presents early analysis of the data accumulated by MySpeedTest. The chapter is divided into sections presenting findings from the latency, warm up and IPDV tests. So far analysis has only been for devices in the US. Hence, the graphs in the following sections represent devices only in the US.

An attempt is made to classify measurements by the different metadata parameters. But the focus is on making sure that different network types are always represented differently. As the following sections will show, the different network types perform very differently in the same set of conditions. Hence we have attempted at each stage to treat each network type separately. This differentiates the analysis presented in this section from those conducted in [30]. This is enabled because of the metadata collected by MySpeedTest. For details of the metadata collected refer to Section 2.1.3.

In order to better understand the results, there is a need to understand the 4 major carriers in the US and the network technologies offered as part of their services. Table 5 summarizes this.

Table 5: Summary of the services offered by 4 major US carriers

Carrier Name	Standard	2G	3G	4G
AT&T	GSM	GPRS,EDGE	UMTS,HSDPA,HSPA	LTE
T-Mobile	GSM	GPRS,EDGE	UMTS,HSDPA,HSPA	HSPA+
Verizon-Wireless	CDMA	1XRTT	EVDO_A,EVDO_0	LTE
Sprint	CDMA	1XRTT	EVDO_A,EVDO_0	LTE

4.1 Latency test

This section summarizes results from the latency tests. The details of the latency test can be found in Section 2.1.1.2. As described, the test conducts 5 pings to 5 different servers. The minimum, maximum, average and standard deviation over those 5 pings is recorded.

4.1.1 Network Types

Figure 12 highlights the most crucial point of this section. It plots the median values of latency measurements for all network types for network technologies in the US, to facebook.com. As can be clearly seen from the figure, the values greatly differ depending on the specific network type. LTE over GSM seems to have the best performance while 1xRTT, which is the 2G technology over CDMA, not only has the highest median latency, but also the most variability.

Figure 13 shows the same data for the different GSM technologies. What is very interesting to note here is that UMTS seems to have lower ping times when compared to HSDPA and HSPA. As mentioned in the Chapter 3 this contrary to the expectation. HSDPA is a high speed improvement on UMTS which is expected to have lower latency, in addition to higher speed. Although, the figure clearly shows that in US networks, UMTS seems to be have lower latency. In fact, there UMTS performs almost 25 percent better than HSDPA in terms of latency to facebook.

Figure 14 looks at the different CDMA technologies. CDMA technologies perform as expected. The performance improvement from 1xRTT to EVDO in terms of round trip latency is very significant - almost 270 milliseconds. LTE shows better performance than EVDO as well.

Figure 15 shows the different 3G technologies. This view provides an interesting comparison between CDMA technology EVDO and the popular GSM 3.5G technology -

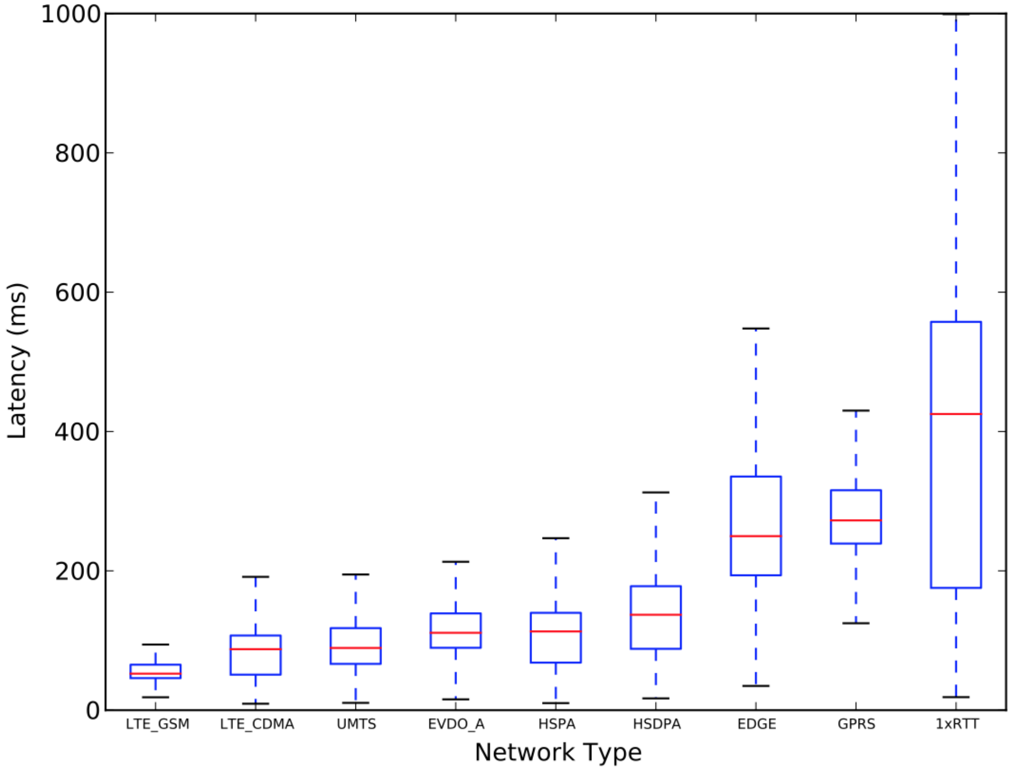


Figure 12: Boxplot of minimum ping times to www.facebook.com for all network types

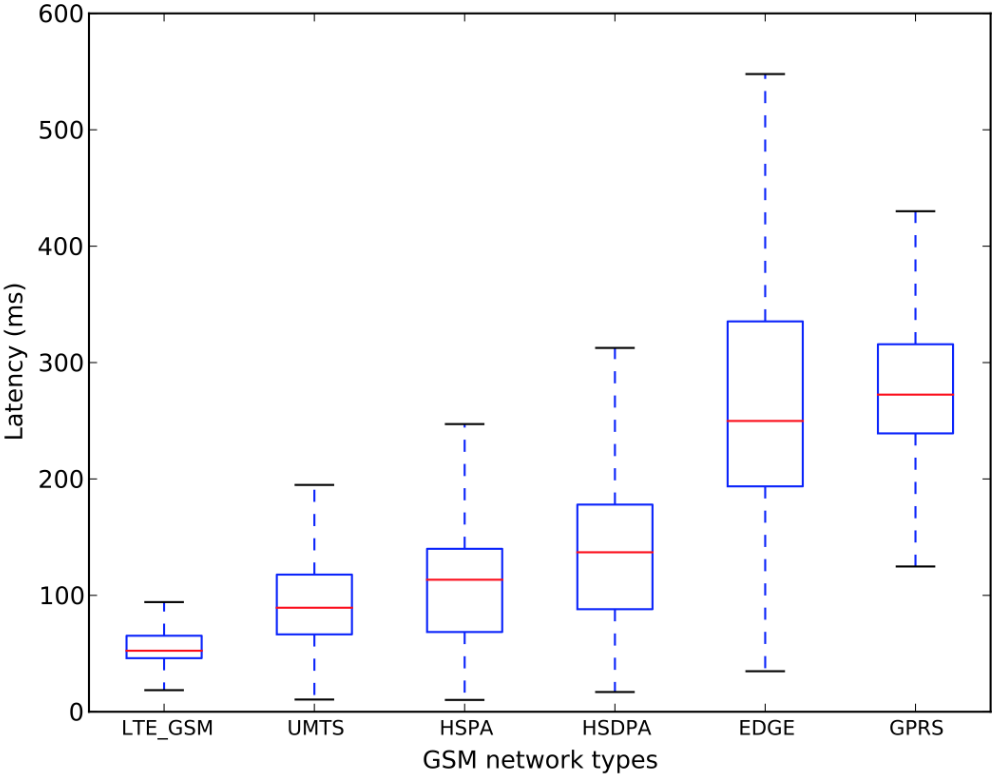


Figure 13: Boxplot of minimum ping times to www.facebook.com for different GSM networks

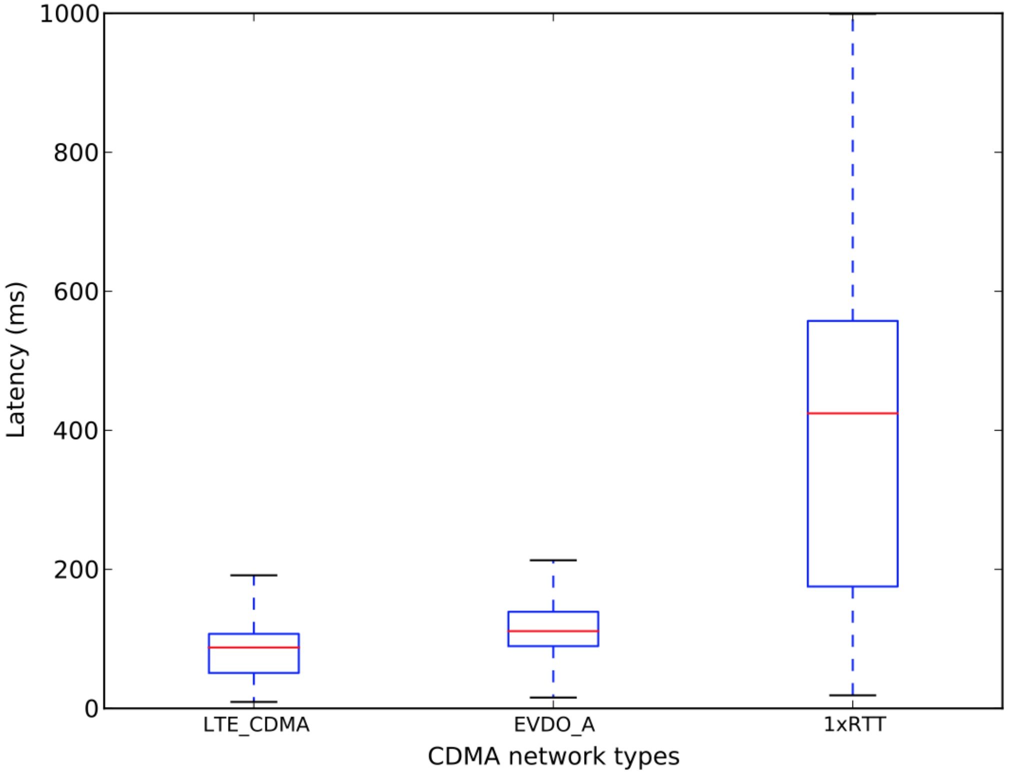


Figure 14: Boxplot of minimum ping times to www.facebook.com for different CDMA networks

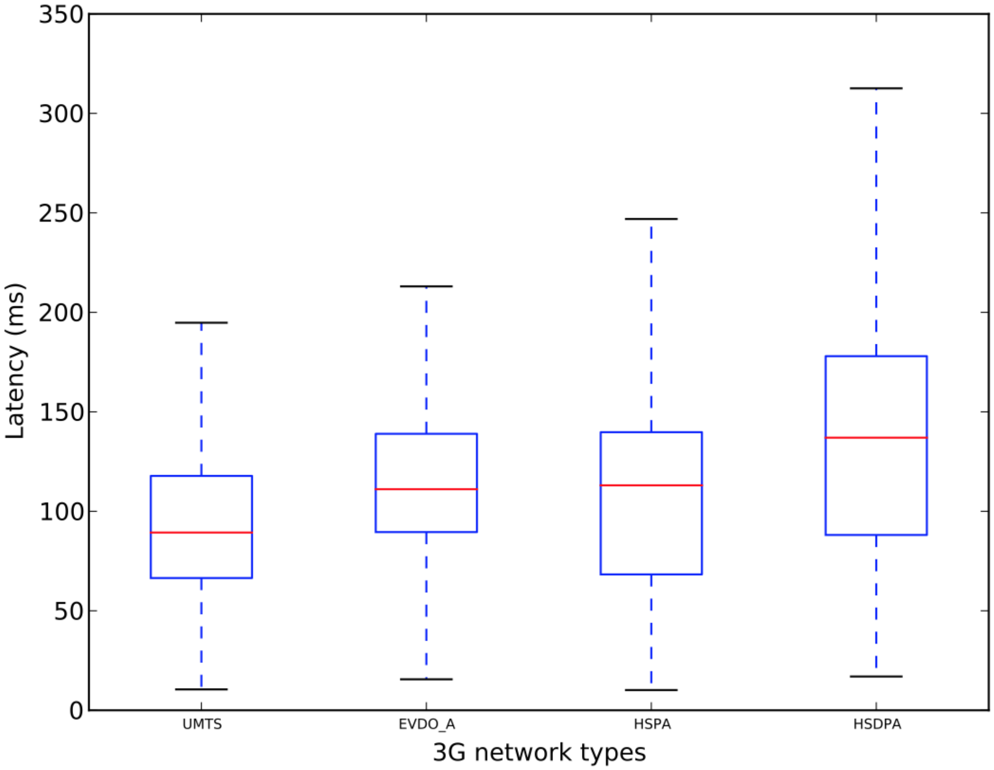


Figure 15: Boxplot of minimum ping times to www.facebook.com for different 3G networks

HSDPA. While surprisingly UMTS still outperforms both of these, EVDO performs on par HSPA but definitely better than the more popular HSDPA.

Figure 16 shows a very interesting finding - LTE over GSM performs almost twice as better as LTE over CDMA. The number of measurements for each of these technologies is similar and hence statistical anomalies are not responsible for this difference. The difference might lie in the service providers that offer these services. This is an area that we will explore in following subsections.

4.1.2 Data activity

One of the interesting metadata parameters that we collect is the data activity before the start of the measurements. The different data activities on Android devices are

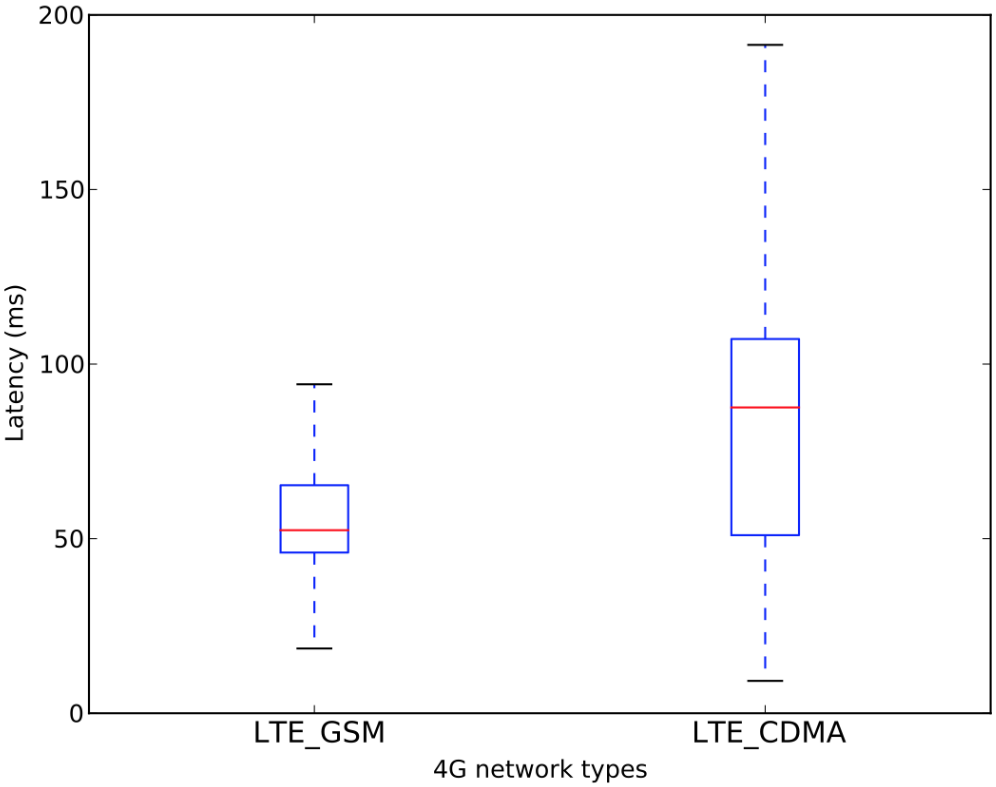


Figure 16: Boxplot of minimum ping times to www.facebook.com for different 4G networks

Table 6: Details of Data Activity States

Data Activity State	Details
DATA_ACTIVITY_IN	Currently receiving IP PPP traffic.
DATA _ACTIVITY _INOUT	Currently both sending and receiving IP PPP traffic.
DATA _ACTIVITY _NONE	No traffic.
DATA _ACTIVITY _OUT	Currently sending IP PPP traffic.
DATA _CONNECTED	Connected but not currently sending or receiving traffic.
DATA _CONNECTING	Currently setting up a data connection.
DATA _DISCONNECTED	Disconnected.
DATA _SUSPENDED	Suspended.

summarized in Table 6. The list is taken from [12]

Figure 17 shows the latency to facebook depending on the data activity state. We have restricted the plots to DORMANT, INOUT and NONE because we have the most number of measurements for these states. As can be seen from the figure, the data activity state does not play a significant role in the latency measurements. This is because of the way the measurements are conducted. The warm up measurement ensures that the radio has a dedicated channel and hence the data state is INOUT at the time of the latency tests.

4.1.3 Android Version

Android Version of the device generally indicate several performance improvements. These are not restricted to user and developer features, but several platform improvements as well [13]. Figure 7 details the deployment of MySpeedTest on the different Android versions. In this section, the results for latency measurements to facebook for 3 different Android versions are presented.

1. Android version 2.3.6 - Gingerbread: Gingerbread upgraded the Linux kernel to version 2.6.35. The networking APIs were improved over previous versions

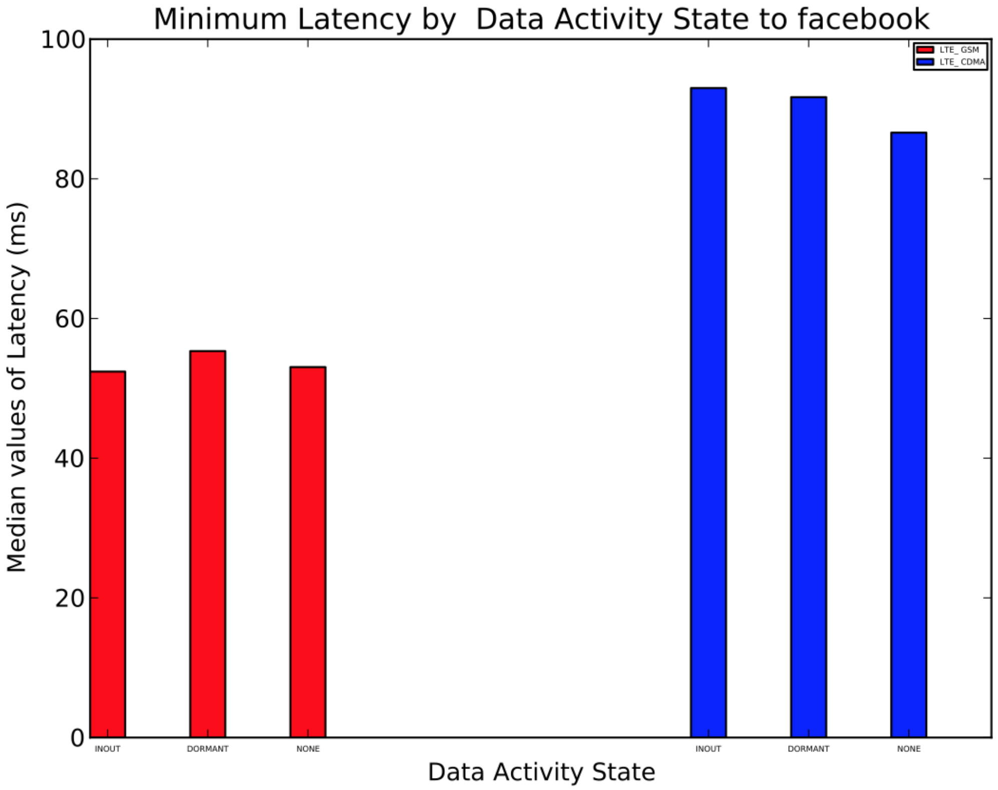


Figure 17: Latency to www.facebook.com based on the data activity state

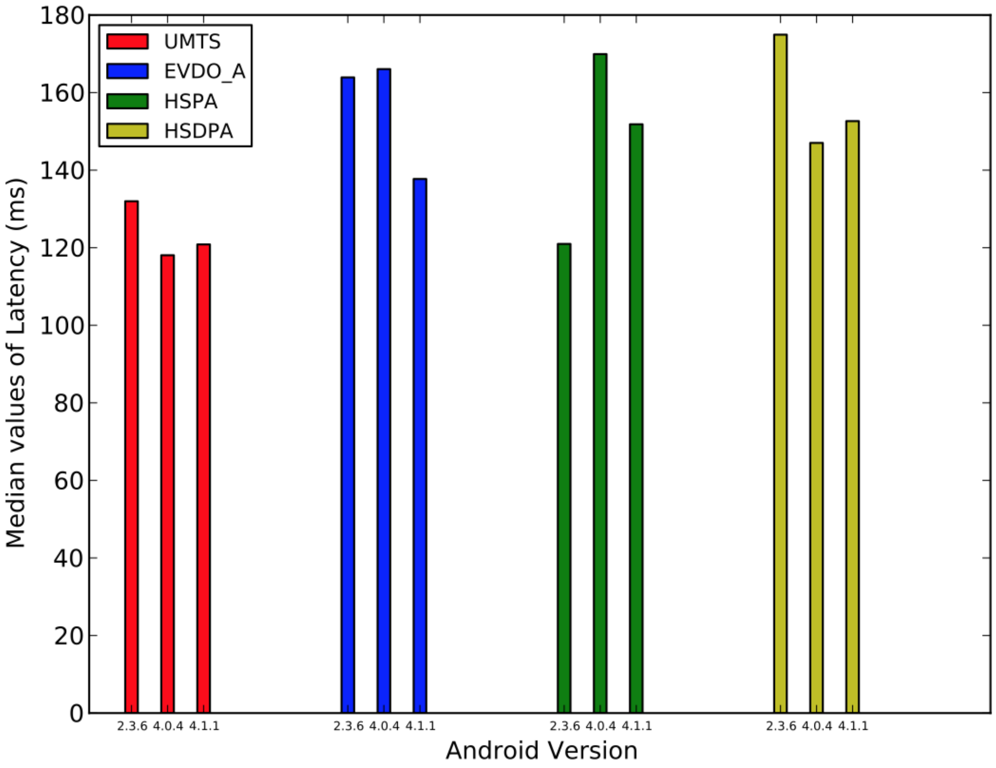


Figure 18: Latency to www.facebook.com based on the Android version for different 3G networks

and support for Near Field Communications was added.

2. Android version 4.0.4 - Ice-cream Sandwich: The Ice cream Sandwich OS is the first to use a Linux 3.x kernel.
3. Android version 4.1.1 - Jelly Bean: The Jelly Bean OS was the newest version of Android at the time of launching MySpeedTest. It was incremental update to the Ice cream Sandwich and about 35 percent of MySpeedTest users use this operating system.

Figure 18 shows the latency for the different 3G technologies by the Android Version. It seems like the Android version and in turn the Linux kernel version does not show a

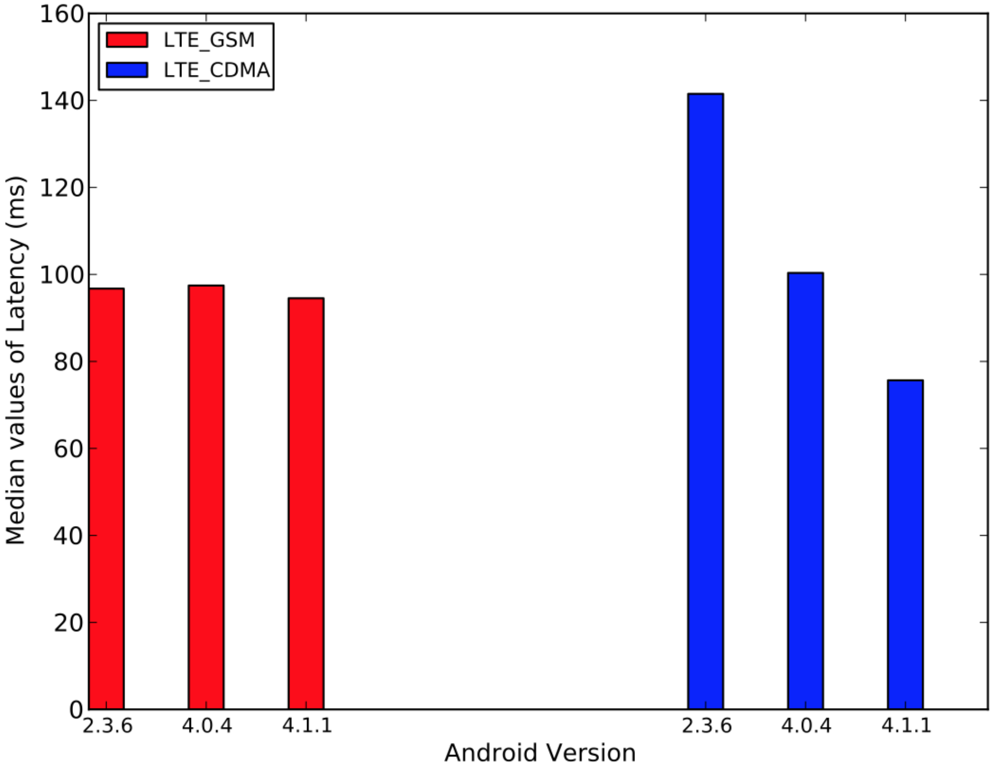


Figure 19: Latency to www.facebook.com based on the Android version for different 4G networks

clear pattern for the various 3G technologies. Version 4.1.1 seems to perform the best on average. The later Android versions have a larger deployment of 4G technologies. The results for the latency test for the different 4G technologies, split up by the Android version are presented in Figure 19. For the CDMA, there is very obvious pattern based on the Android version. Version 4.1.1 shows great improvement over version 2.3.6. This is less pronounced for the GSM technologies. This is indicative that the newer devices which support Android 4.1.1 over CDMA have caused the improvement.

4.1.4 Service Providers

This section describes the performance over the 4 major US carriers. The figures show the round trip time to facebook, but the relative performance remains the same over for the other server destinations as well. For the sake of brevity these figures are not shown.

Figure 20 shows the performance of the service providers for 3G services offered. As highlighted in Table 5 AT&T and T-Mobile offer GSM services, while Verizon Wireless and Sprint offer CDMA services. For GSM services, T-Mobile consistently outperforms AT&T for all the different 3G network types. The round trip latency is almost double in the case of AT&T. This difference is true even to the other destinations; including Oakland, California and Napoli, Italy. This indicates that it is not the peering relationship that T-Mobile shares with facebook and google that leads to this difference. For CDMA services, Verizon seems to provide faster round trip times on 3G than Sprint.

Figure 21 shows the performance for 4G services. An important thing to note is that T-Mobile offers 4G services on HSPA+ networks. Unfortunately, the Android API does not provide information about the network type for HSPA+ networks. It might be possible to infer this based on the latencies achieved but this thesis refrains from making these inferences. For the remaining service providers, it can be seen that Verizon Wireless seems to have much higher latencies than the other service providers. In fact Sprint seems to have lower latencies than AT&T which runs over GSM. Therefore, the difference seen for the 2 standards in Figure 16 is contributed to by the service providers.

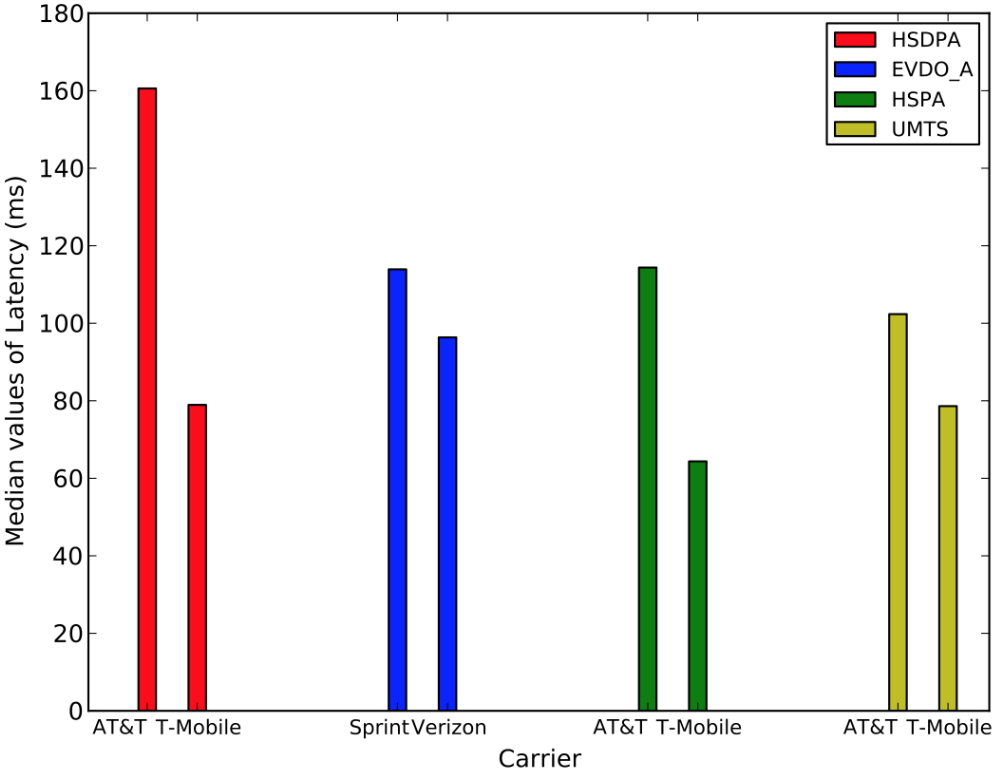


Figure 20: Latency to www.facebook.com based on the Carrier for different 3G networks

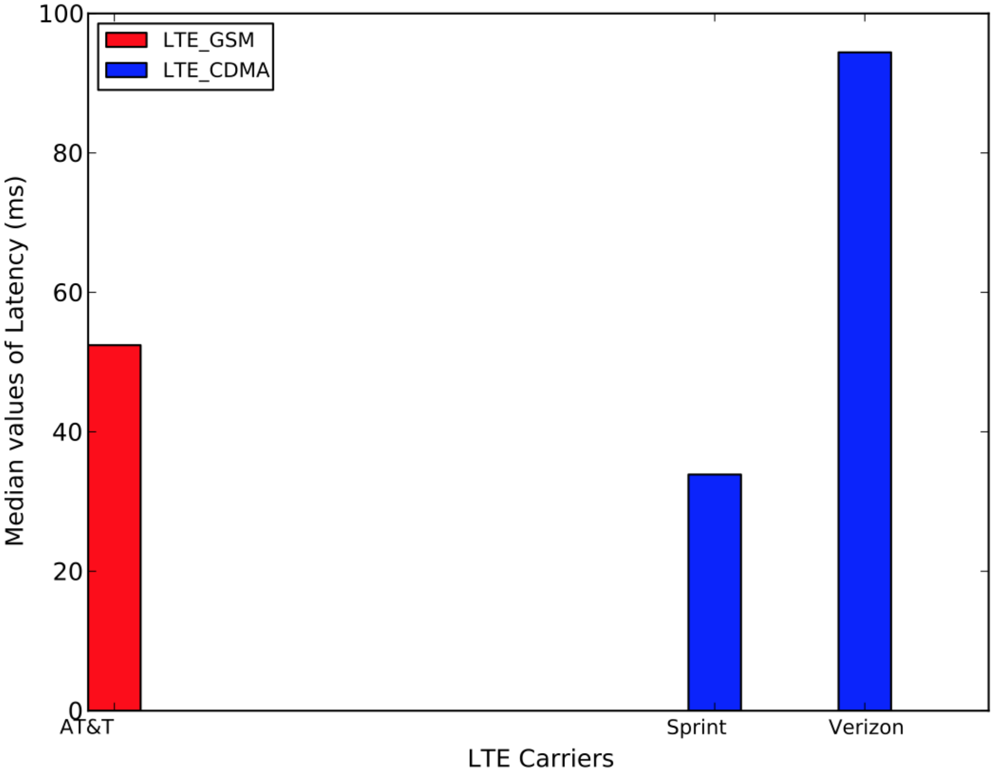


Figure 21: Latency to www.facebook.com based on the Carrier for different 4G networks

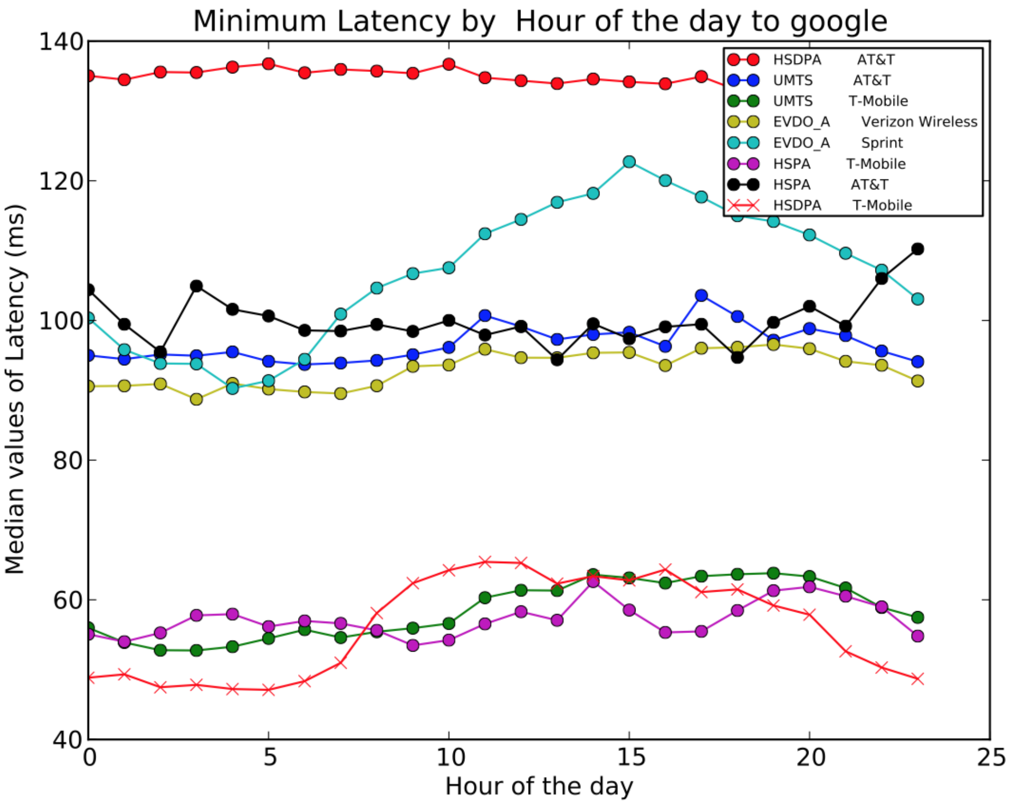


Figure 22: Time of day variation of latency to www.google.com for 3G networks

4.1.5 Time of day

This section characterizes the change in latency over time of day for the different carriers and network types. In cellular networks time of day effects could be caused by congestion due to increased number of users during peak hours.

Figure 22 shows the performance over the hour of the day for 3G services offered by the carriers. T-Mobile and Sprint networks clearly show an increase in latency during peak hours. From about midnight to 5 am, the latency decreases by almost 50 percent for T-Mobile over HSDPA, and 33 percent for Sprint over EVDO_A.

Figure 23 shows performance for 4G networks. AT&T's LTE services do show a slight time of day trend. This isn't as pronounced as in the 3G case though.

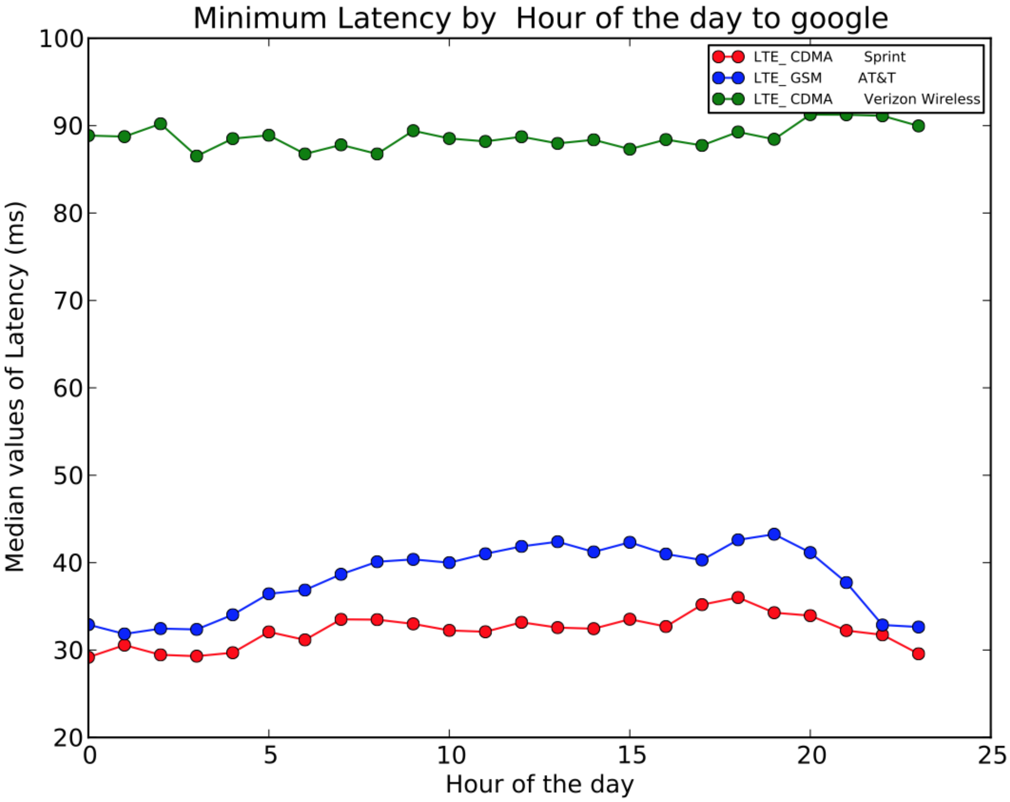


Figure 23: Time of day variation of latency to www.google.com for 4G networks

4.1.6 Signal Strength

The Android API represents signal strength at a more fine grained level than the bars presented to the user on the device. The signal strength is measured on a 32 point scale from 0 to 31. Unfortunately, for CDMA technologies, this value is rarely available from the device. Hence this section only presents data for GSM technologies.

Figure 24 shows the latency for different values of signal strength for the different 3G technologies. As can be seen from the figure, the latency does not show a fixed pattern for different values of signal strength. This is contrary to what might be expected since signal strength greatly affects phone calls from devices.

Figure 25 re-iterates this fact by showing the same effect after splitting the measurements by the different service providers.

This concludes the section on latency results. There are several more metadata parameters to explore and more in-depth statistical analysis to be done. These details can be found in the future works section.

4.2 IP packet delay variation

This section highlights the packet delay variation over the different network types. The IPDV test takes 50 records per measurement. For the analysis, the median value of IPDV is selected over the 50 records. Figure 26 shows boxplots for IPDV values with whiskers for the minimum and maximum values.

The first thing to notice is that in a set of 50 records, the median value of IPDV for all the carriers and network types is well below 25 milliseconds. In fact, AT&T-HSDPA sees a median of about 3 milliseconds. AT&T performs better than all the other carriers in terms of the packet delay variation.

For the 4G technologies, AT&T and Verizon show similar performance. While Sprint has a median value of IPDV comparable to that of AT&T and Verizon Wireless, the variability for Sprint seems high.

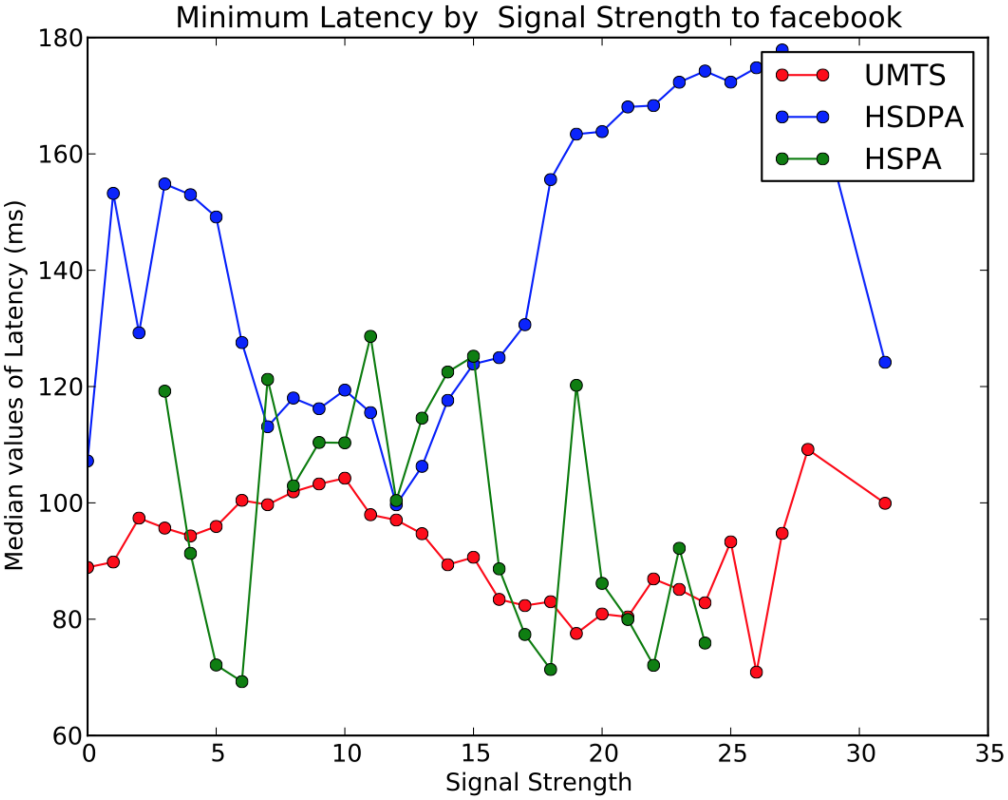


Figure 24: Variation of latency to www.facebook.com with network signal strength for 3G networks

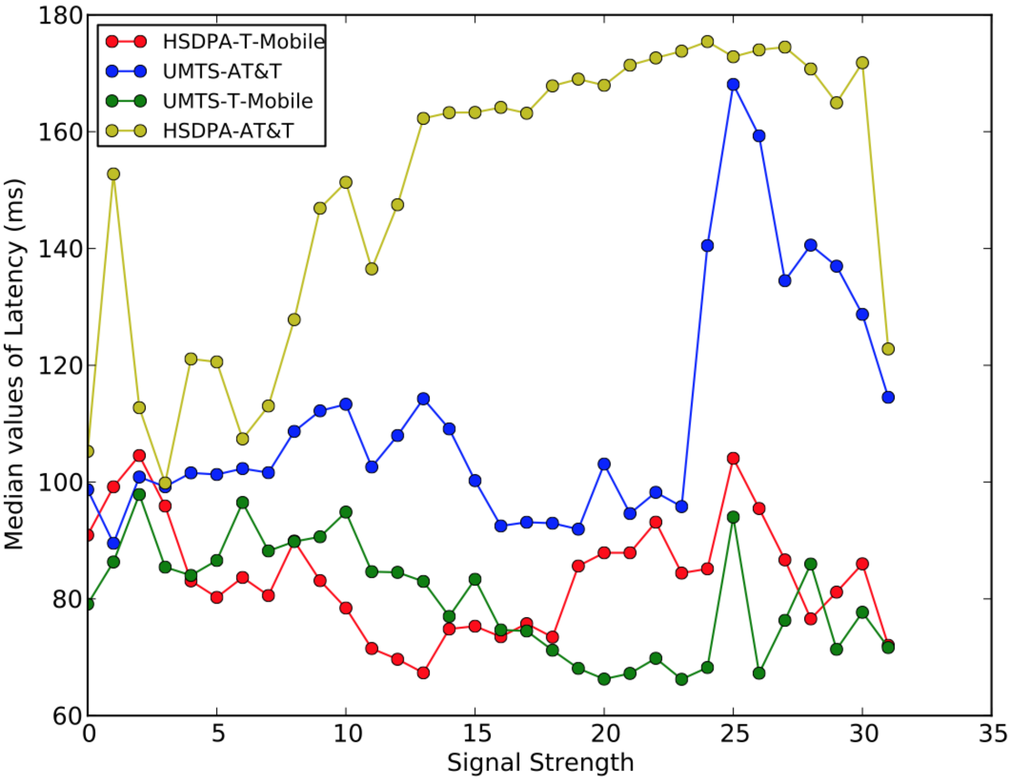


Figure 25: Variation of latency to www.facebook.com with network signal strength for various carriers

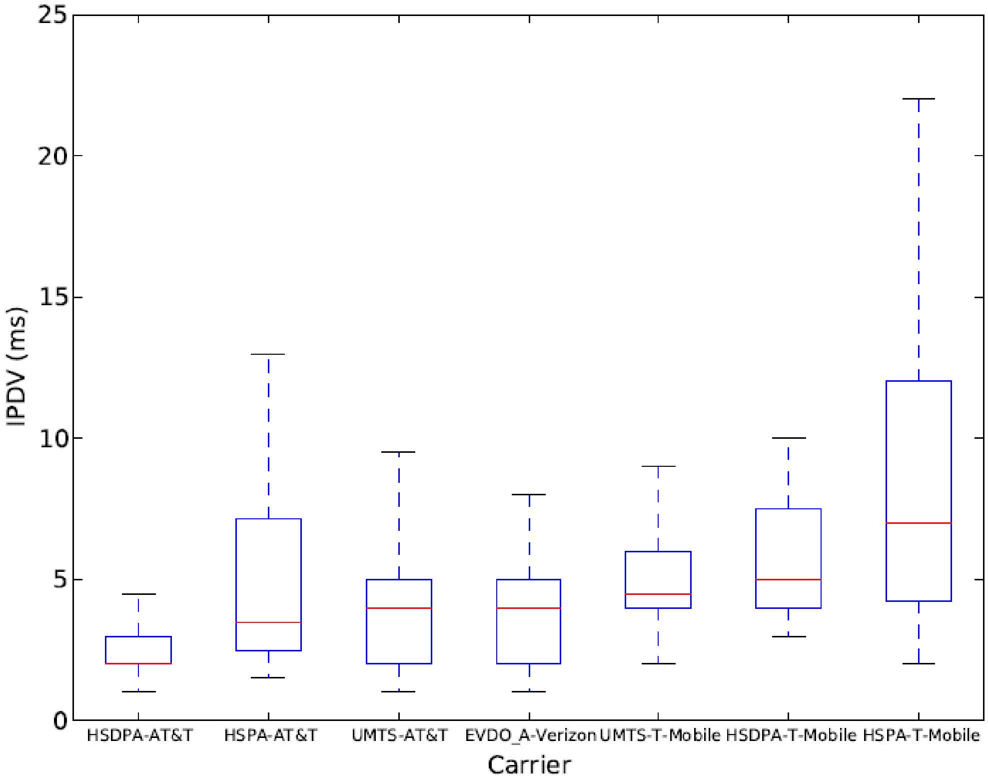


Figure 26: IPDV in 3G networks

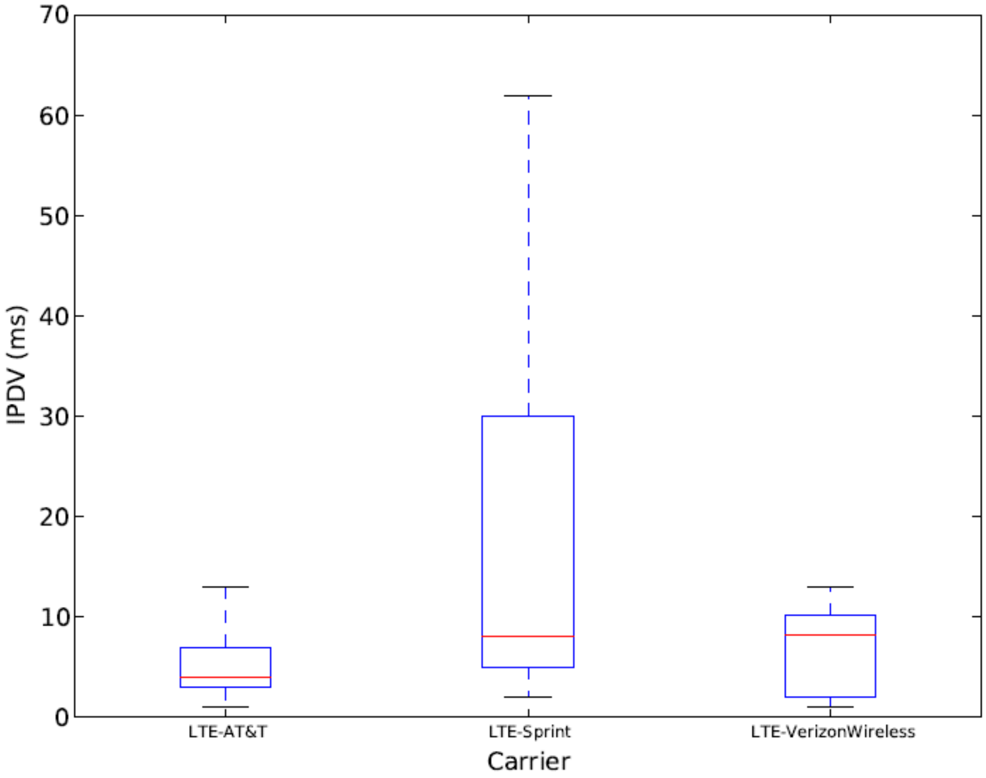


Figure 27: IPDV in 4G networks

4.3 Warm up Measurements

The warm up measurements details the amount of time it takes for a device to acquire a dedicated channel. Previous work has been done in this area by [28, 34, 25]. In order to get large scale measurements of the warm up measurements, rather than from a few rooted devices, MySpeedTest employs this test.

Each warm up measurement contains a train of 40 ICMP ping packets spaced apart by 200 milliseconds. In order to analyze the time it takes to acquire a dedicated channel, the maximum value for round trip time in each train was taken. The remaining packets were normalized by this value to estimate to what percentage of the maximum value, the latencies dropped once a dedicated channel had been acquired.

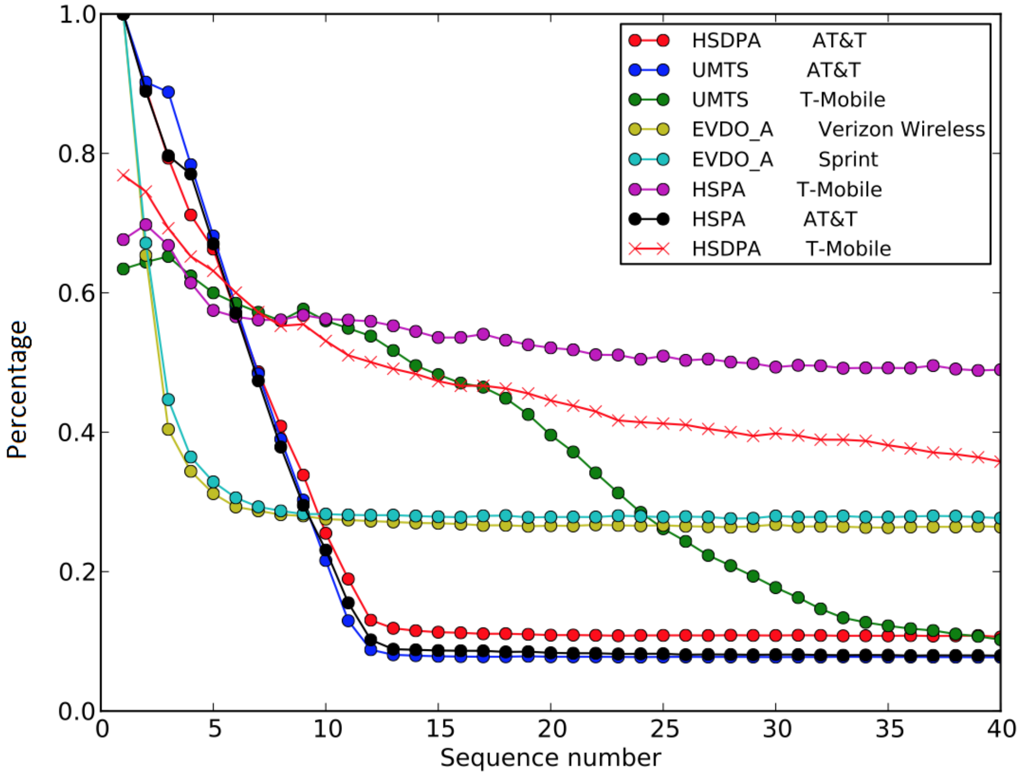


Figure 28: Time to acquire a dedicated channel in 3G networks

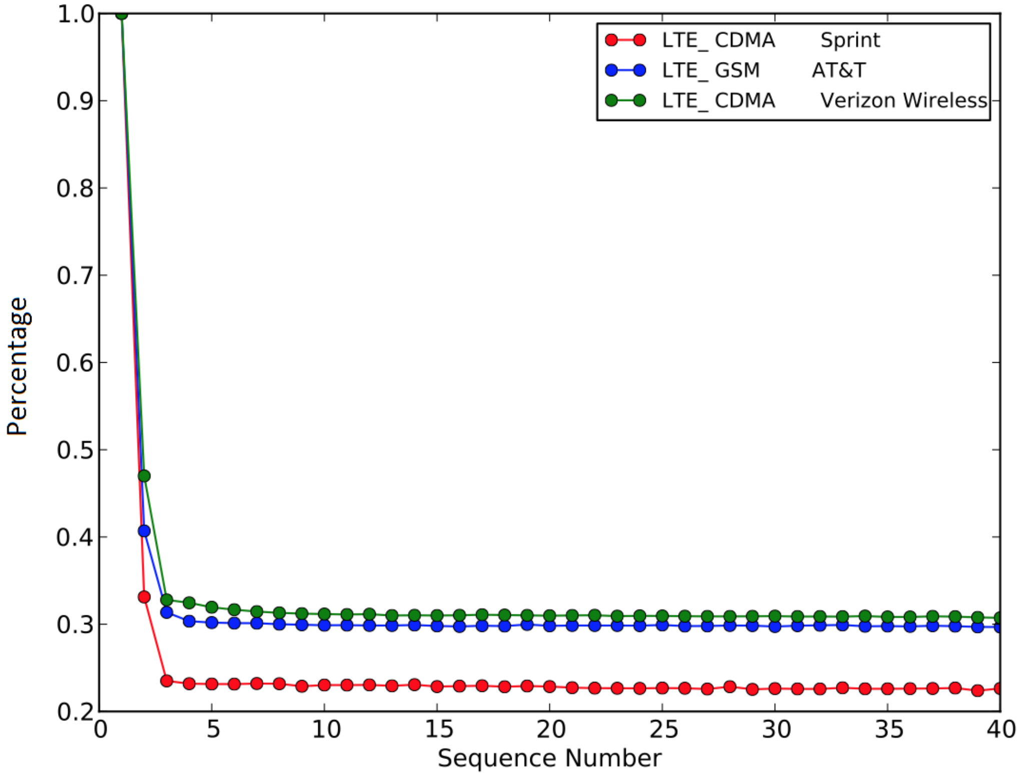


Figure 29: Time to acquire a dedicated channel in 4G networks

Figure 27 shows the result of this analysis for 3G technologies. As can be seen from the figure, the warm up time clearly depends on the service provider and is independent of the network type. AT&T shows a drop to almost 10 percent of the maximum value. It takes about 10 seconds to acquire a dedicated channel in the case of AT&T. VerizonWireless drops to about 25 percent of the maximum value. The time taken though is about 1.2 seconds. T-Mobile and Sprint do not show trends as clearly defined as AT&T and Verizon. In fact the latency in the case T-Mobile seems to gradually decrease to a minimum. Only in the case of T-Mobile, is the maximum value of latency not always for the first packet in the train.

Figure 28 shows the result of the warm up measurements for 4G technologies. In this case, it seems like the time taken to acquire a dedicated channel is the same across service providers at about 600 milliseconds. The work in [28] shows that the state model for LTE is different as compared to the 3G technologies. This new state model is probably used uniformly across service providers and is responsible for the fixed value of warm up time.

CHAPTER V

CONCLUSION AND FUTURE WORK

MySpeedTest has received excellent response in the research community and among users on the Google Play Store. This is for 2 reasons. Firstly, it collects data that has not been gathered before. As mentioned in previous sections, unlike previous studies, the data collected in the case of MySpeedTest takes into consideration the entire context of the measurement. In addition, we collect passive data that will help to better understand user behavior and pinpoint the performance metrics that matter to the user the most.

In Chapter 4 of this thesis the need to separate the data by the different network types was highlighted. The performance variation across the network types is significant and hence these should not be considered together.

Secondly, the user interface of MySpeedTest has ensured a wide global user base in the short span of 4 months. The on device graphs and longitudinal data have resulted in a sharp increase in the number of users. This contributes to the richness of the data set that we are able to gather.

5.1 Future Work

5.1.1 New set of tests

We will soon be adding censorship tests from the OONI measurement suite, [19] so that we can monitor network interference and censorship on cellular networks. This test are in the very nascent stage. The initial set of tests we are considering include:

1. HTTP Header field manipulation - In this test we will vary capitalization in the

header fields of a fixed number of HTTP Requests. If we detect at the server end that the header fields have been modified, then there is some censorship on the route from the client device to the server.

2. HTTP invalid request line - We will send a fixed number of HTTP requests to an echo server. The requests will have invalid request lines (using a randomly generated set of ASCII characters). If we detect that the response from the echo server is different from the request then the requests are being tampered with.

3. Test TCP connections to a specified set of ports - For a dynamic set of IP:PORT pairs we will test TCP connections and reasons for failure if any. The dynamic list will be downloaded intermittently on the client device from a Georgia Tech server.

This is not an exhaustive list of censorship tests that we will add. This is the first set of tests that we are considering. We have not decided the specifics (eg. the list of IP:PORT pairs) of each of the tests or the frequency that we will repeat tests in the background.

5.1.2 Improvement of current tests

The initial data analysis presented in Chapter 4, has resulted in a list of improvements to the current tests. These include:

1. Increase the number of pings to each server from 5 to 15. This is because we have noticed high variability in the ping measurements. In order to gather more statistically accurate data, we intend to increase the number of data points in the ping measurements.

2. Use a dynamic list of ping servers. This list can be populated with the most interesting servers for measurement at the current time. The list will be retrieved

from the server before each measurement.

3. Add latency under load measurements while running the throughput tests. This would give good insight into buffering on the network.

4. Improve the TCP throughput measurements by using a set of globally distributed servers. This will enable far more accurate measurements.

5.1.3 Data Analysis

The analysis of measurement results using MySpeedTest is in its nascent stages. Over 3.5 million measurements have been collected. But many areas remained to be explored. Strong machine learning and statistical models need to be used to correlate the measurement metrics with the various metadata parameters. In addition, the passive measurements have thus far been unexplored. This remains a very interesting area of research. Especially with the global deployment of MySpeedTest. Understanding large scale user behavior with regards to network metrics experienced by applications remains unexplored.

5.2 Partnerships

In order to improve the current set of measurements and get a richer data set we are working on 2 partnerships at the current time -

1. Measurement Lab: [17] has a set of globally distributed measurement servers. They are also experts in making data sets public. Cellular measurements from [30] are already using Measurement Labs infrastructure and have developed an excellent scheduling scheme and collection backend for the tests. Through partnership with M-Lab MySpeedTest will acquire a good scheduling framework, data publishing and data collection backend. In combination with the tests on MySpeedTest and the number of active users, the partnership will be very beneficial.

2. Research ICT Africa: [20]In addition to deployments in developed countries, we are working with Research ICT Africa (RIA) to focus on a deployment in the developing world (specifically, Africa). This will give us unique insight into the performance as seen from developing countries. RIA has already begun recruiting participants for MySpeedTest. This has led to a surge of users in Africa especially South Africa.

REFERENCES

[1] http://venturebeat.com/2013/01/04/google-play-will-hit-a-million-apps-in-2013-probably-sooner-than-the-ios-app-store/.

[2] http://en.wikipedia.org/wiki/History_of_mobile_phones.

[3] http://www.greyfriars.net/gcg/greyweb.nsf/miam/article03.

[4] http://gsmserver.com/articles/gsm_overview.php.

[5] http://en.wikipedia.org/wiki/GSM.

[6] http://www.gsma.com/aboutus/gsm-technology/gprs.

[7] http://www.moxa.com/newsletter/connection/2007/08/.

[8] http://en.wikipedia.org/wiki/High_Speed_Packet_Access.

[9] http://en.wikipedia.org/wiki/Evolved_HSPA.

[10] http://en.wikipedia.org/wiki/CDMA2000.

[11] http://www.gsma.com/aboutus/gsm-technology/lte.

[12] http://developer.android.com/reference/android/telephony/TelephonyManager.html.

[13] http://developer.android.com/sdk/api_diff/9/changes.html.

[14] "Google play developer console." https://play.google.com/apps/publish/.

[15] "Google play store." https://play.google.com/store/apps.

[16] "Iperf." http://sourceforge.net/projects/iperf/.

[17] "Measurement lab." www.measurementlab.net.

[18] "Netperf." http://www.netperf.org/netperf/.

[19] "Open observatory of network interference." https://ooni.torproject.org/docs/tests/.

[20] "Research ict africa." http://researchictafrica.net/.

[21] "Shaperprobe." http://www.cc.gatech.edu/ partha/diffprobe/shaperprobe.html.

[22] "Speedtest.net." http://www.speedtest.net/.

[23] "Strategy analytics." http://www.strategyanalytics.com/.

[24] Almes, G., Kalidindi, S., and Zekauskas, M., "A one-way packet loss metric for ippm," tech. rep., RFC 2680, September, 1999.

[25] Balasubramanian, N., Balasubramanian, A., and Venkataramani, A., "Energy consumption in mobile phones: a measurement study and implications for network applications," in *Proceedings of the 9th ACM SIGCOMM conference on Internet measurement conference*, pp. 280–293, ACM, 2009.

[26] Böhmer, M., Hecht, B., Schöning, J., Krüger, A., and Bauer, G., "Falling asleep with angry birds, facebook and kindle: a large scale study on mobile application usage," in *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services*, pp. 47–56, ACM, 2011.

[27] Demichelis, C. and Chimento, P., "Ip packet delay variation metric for ip performance metrics (ippm)," 2002.

[28] Huang, J., Qian, F., Gerber, A., Mao, Z. M., Sen, S., and Spatscheck, O., "A close examination of performance and power characteristics of 4g lte

networks," in *Proceedings of the 10th international conference on Mobile systems, applications, and services*, pp. 225–238, ACM, 2012.

[29] Huang, J., Qian, F., Mao, Z. M., Sen, S., and Spatscheck, O., "Screen-off traffic characterization and optimization in 3g/4g networks," in *Proceedings of the 2012 ACM conference on Internet measurement conference*, pp. 357–364, ACM, 2012.

[30] Huang, J., Xu, Q., Tiwana, B., Mao, Z. M., Zhang, M., and Bahl, P., "Anatomizing application performance differences on smartphones," in *Proceedings of the 8th international conference on Mobile systems, applications, and services*, pp. 165–178, ACM, 2010.

[31] Paxson, V., Almes, G., Mahdavi, J., and Mathis, M., "Rfc 2330: Framework for ip performance metrics," 1998.

[32] Qian, F., Quah, K. S., Huang, J., Erman, J., Gerber, A., Mao, Z., Sen, S., and Spatscheck, O., "Web caching on smartphones: Ideal vs. reality," in *Proceedings of the 10th international conference on Mobile systems, applications, and services*, pp. 127–140, ACM, 2012.

[33] Qian, F., Wang, Z., Gao, Y., Huang, J., Gerber, A., Mao, Z., Sen, S., and Spatscheck, O., "Periodic transfers in mobile applications: Network-wide origin, impact, and optimization," in *Proceedings of the 21st international conference on World Wide Web*, pp. 51–60, ACM, 2012.

[34] Qian, F., Wang, Z., Gerber, A., Mao, Z. M., Sen, S., and Spatscheck, O., "Characterizing radio resource allocation for 3g networks," in *Proceedings of the 10th ACM SIGCOMM conference on Internet measurement*, pp. 137–150, ACM, 2010.

[35] Sundaresan, S., de Donato, W., Feamster, N., Teixeira, R., Crawford, S., and Pescapè, A., “Broadband internet performance: a view from the gateway,” in *ACM SIGCOMM Computer Communication Review*, vol. 41, pp. 134–145, ACM, 2011.

[36] Xu, Q., Huang, J., Wang, Z., Qian, F., Gerber, A., and Mao, Z. M., “Cellular data network infrastructure characterization and implication on mobile content placement,” in *Proceedings of the ACM SIGMETRICS joint international conference on Measurement and modeling of computer systems*, pp. 317–328, ACM, 2011.