A First Look at Mobile Internet Use in Township Communities in South Africa

Amreesh Phokeer
Melissa Densmore
University of Cape Town
Rondebosch, South Africa
{aphokeer,mdensmore}@cs.uct.ac.za

David Johnson
CSIR
Meraka Institute
Pretoria, South Africa
djohnson@csir.co.za

Nick Feamster
Princeton University
Computer Science Department
Princeton, NJ, USA
feamster@cs.princeton.edu

ABSTRACT
This paper presents a study of mobile data usage in South African townships. In contrast to previous studies, which have studied mobile data usage in developing regions (including South Africa), we focus our study on two townships in South Africa; the extremely resource-constrained nature of townships sheds light, for the first time, on how people in these communities use mobile data. We perform a mixed-methods study, combining quantitative network measurements of mobile app usage with qualitative survey data to gain insights about mobile data usage patterns and the underlying reasons for user behavior concerning mobile data usage. Due to the limited availability of public free Wi-Fi and despite the relatively high cost of mobile data, we find that a typical township user's median mobile data usage is significantly more than Wi-Fi usage. As expected, and consistent with observations of mobile data usage in parts of South Africa with better resources, users tend to favor using Wi-Fi for streaming video applications, such as YouTube. Interestingly, however, unlike users in less resource-constrained settings, township users also consume significant mobile data to update mobile applications, as opposed to relying on Wi-Fi networks for application updates. These behaviors suggest that network and mobile application designers must pay more attention to data usage patterns on cellular networks to provide mobile network architectures that provide more cost-effective mechanisms for tasks such as application update.

CCS Concepts
• Network measurements • Network types • Human Computer Interaction (HCI)

Keywords
Smartphone usage; mobile application popularity; usage patterns

ACM DEV '16, November 17-22, 2016, Nairobi, Kenya
© 2016 ACM. ISBN 978-1-4503-4649-8/16/11…$15.00
DOI: http://dx.doi.org/10.1145/3001913.3001926

1. INTRODUCTION
Despite the tremendous growth in Internet-capable mobile device adoption [1], Internet usage and access to data is limited in South Africa by prohibitive costs and unequal coverage [2]. Yet, the high cost of communication has not deterred the growth of mobile data usage in the less-privileged areas such as in the South African townships[1]. In fact, mobile data usage growth in township areas has outpaced the average usage growth across the whole of South Africa [3].

Yet, mobile data is expensive relative to the incomes of township residents. The World Bank estimates that half of South African's urban population lives in townships and informal settlements, accounting for 38% of working age citizens, and home of nearly 60% of unemployed [4]. In Khayelitsha, one of the poorest areas of Cape Town, the median average monthly income of a family of five is approximately ZAR 1600 (USD 110) [5]. Our study suggested an average monthly expenditure of up to ZAR 100-200 per user. At such a high cost, many users may find mobile data unaffordable; therefore understanding the nature of the usage patterns in lower-income townships is important, both to understand the economic consequences of mobile Internet penetration, as well as to suggest opportunities for network and application architectures to better optimize data use in these settings.

We analyzed the data usage patterns of mobile Internet users living in township communities in South Africa. To this end, seven high school students and seven knowledge workers from two different township communities (Ocean View and Masiphumelele) in Cape Town were recruited to participate in our research study. We performed the study with a mixed-method approach comprising two parts: (1) quantitative measurements of the usage of different mobile applications using the MySpeedTest application [6]; and (2) a survey examining users' behavior concerning mobile Internet usage. With these two methods, we aimed to cross-validate behaviors of mobile usage collected from the measurement application with the responses received from the survey.

In 2015, Mathur *et al.* found that, in contrast to more developed regions, when data is expensive or limited, users have the tendency to be extremely cost-conscious and would employ various strategies to optimize mobile data usage [7]. This situation obviously does not encourage the extensive use of Internet technologies, which could enable resource-constrained communities to share information, communicate, generate content

---

[1] "Townships" refers to urban informal settlements in South Africa, where people were historically displaced during Apartheid period based on their ethnicity. They are the poorest urban communities in South Africa.

and make use of online educational material for their own benefit. It leaves open the complementary but important question of how users in more resource-constrained communities such as townships use mobile applications and consume mobile data. A previous study on broadband measurements in South Africa also revealed interesting data on performance bottlenecks [8]. Yet in contrast, very little is known about Internet connectivity in township communities. By characterizing mobile Internet usage, we attempt to build a solid understanding of the need of cellular networks users from township communities in South Africa.

We also studied the extent to which mobile data traffic is exchanged with users who reside in the same geographic region. Because we do not have access to mobile operators' traffic traces, it is difficult to accurately measure this characteristic. Instead, we studied this question using a survey, which revealed that most of the interactions on social networks are targeted to "friends" who live roughly the in the same locality. This means that users are actually using their expensive and limited data packages to send and receive data to peers living relatively nearby.

The quantitative measurements allow us to investigate how much traffic is being generated for social media, communications, software updates, video streaming, and other applications, as well as how usage is influenced by economic factors such as promotional data packages and zero-rated services. By gathering and analyzing empirical data on how mobile Internet is consumed in township areas, our results can ultimately guide researchers on the needs of mobile phone users, especially in the resource-constrained regions. The outcome of this research can provide important input for the design and deployment of alternative network architectures [9] that could reduce the cost of interconnectivity.

Our study reveals the following findings, several of which contrast with previous studies in South Africa in higher-income communities:

- In contrast to communities with higher incomes, median daily data usage across users is more on cellular data networks than on than Wi-Fi. Qualitative survey results suggest that the relative inaccessibility of public Wi-Fi may induce this behavior.
- In contrast to communities in South Africa with more resources and higher incomes, township users consume significant mobile data on cellular networks to update mobile applications.
- As in other communities, streaming video usage is lower on cellular data networks than on Wi-Fi.

The rest of the paper is organized as follows. We discuss related work pertaining to Internet measurements for mobile data usage in Section 3. In Section 4, we will talk about the research context and give a brief description of the township communities in question. In Section 5 we will discuss our approach, the metrics used and we will give a description of the MySpeedTest application. Finally in Section 6, we present the results of the study including the measurement exercise, the survey, and the semi-structured interviews, before ending with some discussions and closing remarks.

2. BACKGROUND AND RELATED WORK

In this section, we review related work on mobile Internet usage and mobile data measurements. We will present previous works from both the quantitative and qualitative perspective and provide some insights on techniques used to gather usage data. However, besides the specific Internet usage studies and methods, which will be discussed in sections 2.1 to 2.3, it is important to analyze prior research in the field of mobile phone usage in similar demographics in other countries as they provide useful background and contextual information for our study.

2.1 Previous studies of Internet use in urban settings

There has been prior research on mobile application usage in slum communities in different parts of the world. For example, Rangaswamy *et al.* carried out an anthropological study of everyday mobile Internet adoption among teenagers in a low-income urban setting in India [10]. They discovered entertainment to be a major aspect of technology infusion, contributing to the enhancing the ICT related skills and abilities of users. Wyche *et al.* also conducted studies in Viwandani, a slum in Nairobi, articulating some of the ways in which Facebook is used for "hustling", or ad hoc income generation [11]. Both of these studies provide rich perspectives into how slum youth use mobile Internet. Although our study also looks at mobile usage habits of youth, this paper presents a more precise examination of spending and the breadth of application use, rather than trying to provide deep insights into specific ways in which mobile Internet is used. Furthermore, not all low-income urban settings are equal; it is important to understand and characterize use in a variety of contexts.

Sambasivan *et al.* performed an interesting experimental study of how SmartBrowse, a tool to allow users to monitor their Internet usage, helped users to reduce their mobile data expenses but at the same time, increase web pages views [12]. Even though, the MySpeedTest application used in this study could allow users to track data usage, the purpose of our study is to understand usage patterns in normal settings without introducing any form of bias.

2.2 Previous qualitative studies in South African townships

There is a small set of literature found on mobile Internet usage in township communities in South Africa. In 2009, Kreutzer made a study of 66 secondary school grade-11 students in a low-income area in Cape Town [13]. The study revealed that more than 97% of respondents actually owned a mobile phone or used one on a regular basis. The study also suggests that mobile Internet was quite popular with 83% of the respondents accessing the web on a typical day.

In 2011, Donner, Gitau and Marsden studied mobile Internet-only usage in an urban setting in South Africa [14]. They used an ethnographic action research approach to study the challenges and practices of mobile data usage in a resource-constrained setting. Research subjects were observed after being given training and they found out that most of them were still using the Internet on their mobile phones, especially for entertainment and communication - months after receiving the training.

The literature does not mention of any other qualitative study of Internet usage in South African township communities since 2011. The drop in the price of smartphones and faster mobile broadband connectivity (3G/LTE) in those areas completely disrupted the rate at which mobile data is being consumed. Figure 1 shows how data usage drastically evolved for Vodacom with consumption increasing by almost 500% between 2011 and 2015 [15]. This gap therefore further motivates our study on mobile data usage in township communities in South Africa.

2.3 Previous quantitative measurement studies in South Africa

Several other recent studies have performed quantitative examinations of mobile Internet usage. In 2013, Chetty *et al.* used passive and active measurement methods to collect performance and usage data from both home routers and mobile phones [8]. One of the objectives was to compare broadband performance on different connection types and see whether users were getting the performance advertised by their Internet Service Providers (ISPs). We used a combination of measurement tools: BISmark [16] on home routers, MyBroadband [17] and MySpeedTest [6]. They found that (1) users were not getting the advertised speed from their respective ISPs (2) mobile broadband users have a higher throughput than fixed-line users; and (3) high latency to popular websites and services affected performance and quality of service.

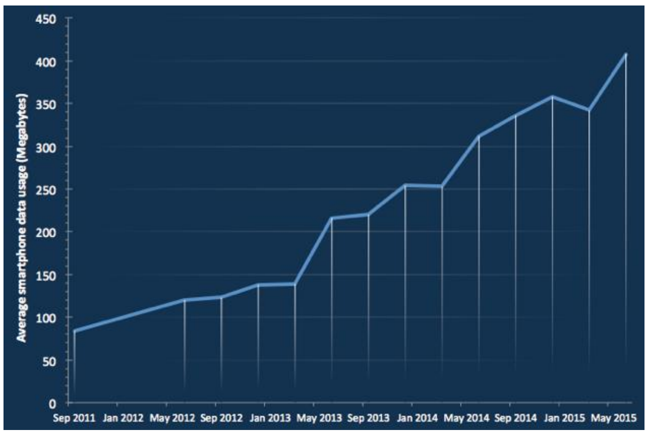

Figure 1: Smartphone data usage (source: Vodacom)

More recently in 2015, Mathur *et al.* used a multi-factor approach triangulating data from three distinct sources: semi-structured interviews, surveys and the MySpeedTest application to study characteristics of mobile broadband usage of high-income versus low-income participants across South Africa. Although the study does not specifically target resource-constrained regions, we expect to find similar patterns especially in terms of application usage. For this study, they interviewed more than 300 participants, made 43 interviews and collected measurement data from 121 mobile devices.

Furthermore, the behavior of mobile Internet users is greatly influenced by mobile pricing practices [18]. In a comparative study of mobile usage between US and South African users, Chen *et al.* found that South African users tend to use more Wi-Fi connections, whenever available, except for zero-rated services provided by their network carrier.

There is currently vigorous debate around the provisioning of "free" services from Over-The-Top (OTT) providers, as it raises questions on anti-competitive practices as well as concerns on net neutrality [19]. A recent study on the quality of service of Facebook's Free Basics service revealed that "paid" and "free" services offer users different experiences, in terms of both content and quality of experience (QoE) [20]. Table 1 gives a list of services that are currently zero-rated by mobile operators in South Africa.

Table 1. List of zero-rated services by mobile operators

Operator	Service	Description
MTN	Wikipedia Momaths	Users can access Wikipedia and MoMaths service for free using Opera Mini.
Cell-C	Whatsapp Freebasics	Whatsapp is unlimited (except voice calling) for R5 per month. Freebasics allows free access to Facebook (no videos and images available) and other free services such as news, classified and Wikipedia. None of the services have images or videos
Vodacom	E-School	Provides zero-rated access to a few educational websites.
Telkom	ShowMax VoD	Free video-on demand service available for premium users only.

2.4 Measurement tools

There are many ways to study the network usage of mobile phones. One way is to capture passive log data at the network operator's level and try to infer statistics on usage. Since it is usually almost impossible to have access to the carrier's data, unless we have some prior agreements, passive log measurements is not an option. The other way to proceed is to collect passive and/or active measurements directly from the mobile device. A few such platforms are available, some of them being proprietary and others open-source. The measurement platform typically consists of a software probe installed on the mobile device and a central database, where measurement data is captured.

The four main measurement Android-based platforms available are: Netalyzr [21], Mobiperf [22], Mobilyzer [23] and MySpeedTest [6]. They are all more or less equivalent, especially that all of them are different implementations of the core library of Mobilyzer. In the design of our experiment, we intentionally decided to only select users with Android smartphones, as they were quite representative of the population where there is a clear dominance of Android phones as opposed to other type of operating systems [24]. We decided to use MySpeedTest as it is open source, widely deployed in South Africa and it proved to be rather efficient based on the experience gathered from the previous studies [7], [8], [18]. Figure 3 shows the control interface and data usage view of the MySpeedTest application.

On the other hand, Koradia *et al.* used a custom-built measurement framework to study the state of cellular data connectivity in rural and urban India [25]. Their analysis mostly targeted cellular network performance for data connection. They performed active measurement during a period of 3 months, amounting to a total amount of 450 hours of data collected using iperf [26]. They tested four different cellular providers from seven different locations. Their measurement architecture consisted of a measurement node (a desktop PC with a 3G dongle), a control server, a measurement server and a data server. As in the study on broadband performance measurement in South Africa [8], they found that throughput on 2G and 3G networks were significantly lower than advertised rates. One important discovery while measuring TCP performance was connection stalls.

3. RESEARCH CONTEXT

Masiphumelele (nicknamed Masi) is a township in Cape Town, South Africa, situated between Kommetjie, Capri Village and Noordhoek occupying roughly one square kilometer. In 2010, the population was estimated at 38000. A number of NGOs such as

Living Hope, MasiCorp and Desmond TuTu Foundation have been working for the past decade to uplift the community through health care, education, youth programs and business development initiatives and there are many opportunities to develop ICT solutions to complement these services. Just five kilometers away, there is Ocean View, another township established in 1968 with approximately 14000 inhabitants (see Figure 2). Both townships currently have no public Wi-Fi and the Cape Town's planned public Wi-Fi project in the townships of Khayelitsha and Mitchell's Plain will not be deployed there in the short term.

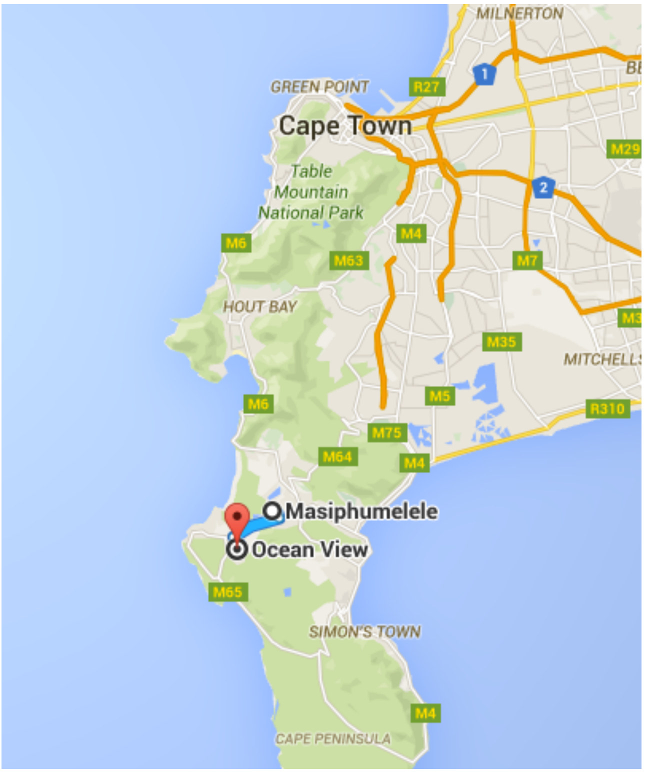

Figure 2: Masiphumelele and Ocean View townships

Current Internet access in Masiphumelele is limited to 3G from the different providers, an Internet Café and limited Internet access at the Library (for example, no YouTube is allowed). In Ocean View, the only publicly accessible Internet service (no Wi-Fi) is at the Library where it is limited to 45 minutes per day per user and users need to get vouchers prior to getting access to the Library computer facilities. As a result most community members access the Internet through cellular connectivity. Ocean View and Masiphumelele are both fairly well covered by GSM and UTMS networks, with some very limited LTE coverage, as this is currently being deployed.

We discovered that most of the users recruited have pre-paid or "pay-as-you-go" mobile plans as opposed to contract plans. They usually buy airtime or data bundles from either the nearby shops or shopping malls. To be able to use contract plans, a user must be able to prove a stable monthly source of income and a bank statement, which automatically disqualifies students and any informal worker. It is therefore very common to see that almost all mobile users in township areas are using prepaid plans. Mobile Internet is available either through time limited data bundles or directly from the airtime available, at a premium cost. Table 2 provides some of the entry-level data plans available, price and validity.

Users living in township areas typically buy data bundles as and when needed, usually multiple times in a week. We have to bear in mind that we are dealing with a population group where more than 50% of the household derives a monthly income of less than R1600 (USD 110) as per a 2011 census from the City of Cape Town [27].

Table 2. Data plans from mobile operators

Operator	Plan 1	Plan 2	Plan 3	Plan 4
MTN	5MB R4 1-day	20MB R12 1-day	50MB R25 3-days	300MB R85 5-days
Voda-com	20MB R5 1-day	100MB R10 1-day	250MB R20 1-day	250MB R60 1-month
Cell-C	20MB R3 1-day	100MB R13 1-day	100MB R25 1-month	300MB R60 1-month
Telkom	25MB R8 1-month	50MB R15 1-month	100MB R29 1-month	250MB R39 1-month

Access to the Internet is therefore a challenge. Not only must users rely on relatively costly mobile Internet connectivity, sometimes with very short life-span, but also they must cope with issues of poor network performance as reported by some interviewees in our study.

4. DATA COLLECTION

In this section, we will discuss how we selected our participants and gathered our datasets and also how representative they are vis-à-vis our population. We then describe some metrics with regards to usage of mobile data connectivity followed by a description of some of the technical and logistical challenges we faced in measuring mobile Internet usage and performance in those two township communities.

4.1 Dataset

We conducted our study on seven high school students from Ocean View and seven knowledge workers from Masiphumelele. We ran our measurement experiment for six weeks (see Table 3), where participants were told to use their mobile phones, just as they would do on any other day. As incentive and at the end of the experiment, for every participating phone, we collected the total amount of data used by the MySpeedTest application on 3G and we topped up the participant's phone with twice the amount that we spent conducting the study. As such we spent between 300 and 400 ZAR to reimburse all our participants.

Table 3. Method and duration

Method	Number of users	Time period
MySpeedTest	14	6 weeks (May-June 2016)
Survey/Interview	14	June 2016

The students were conveniently sampled as they volunteered to participate in this exercise after all grade-10 students were informed about this experiment. Grade-10 students were preferred over lower grades as they were deemed to be at an appropriate maturity level for collaboration with the researchers. Similarly, the knowledge workers were also conveniently sampled as they all work for the NGO Park in Masiphumelele. The 14 users who installed the MySpeedTest applications were surveyed.

4.2 Validity and representativeness

Our sample is rather small to provide good inferential statistics on the whole population of townships in South Africa. However, we argue that this sample gives an indication on potential usage patterns of two important subgroups of a township community, which we believe are the two biggest users of Internet related services, whether it is for communication and social media related activities.

We also acknowledge that conveniently sampling our participants can introduce a bias in our data as argued by Burrell *et al.* [28]. We intentionally selected only participants with Android phones to be able to install the MySpeedTest application. Those with either Blackberries or Windows phone, even though very few could have actually contributed to larger sample diversity. To mitigate this risk, those with non-android phones were interviewed separately and their feedback were recorded on the survey form.

4.3 Usage metrics

To determine usage, we studied the amount of data spent on different classes of applications on a daily basis. By aggregating the data, we then characterized usage as follows:

- Number of applications
- Mean daily usage across all users
- Most used applications on Wi-Fi
- Most used applications on Cellular
- Usage of zero-rated applications

4.4 Data collection tools

We briefly describe the MySpeedTest application and our survey instrument.

4.4.1 *MySpeedTest application*

The MySpeedTest is an Android application developed by the GTNoise Lab at Georgia Tech [6]. The application was used in the study of mobile broadband performance and usage in South Africa [7], [8], [18]. Apart from collecting application usage, the tool also collects data on throughput packet loss, latency and jitter to known online services from the user's smartphone. There is also a feature to collect traceroute data to a specified location as instructed by the user. The application also collects some metadata such as network operator's name, SSID, data cap plan and so on. Table 4 provides a list of relevant tables from which performance and usage data are extracted. In our study, we shall only analyze the usage data and not the performance data.

Figure 3. MySpeedTest app interface

4.4.2 *Survey*

The aim of the survey is two-fold: (1) gather data on parameters that are difficult to measure using quantitative techniques such as price perception or localization of social media contact; and (2) confirm measurements recorded from the MySpeedTest in case results are skewed by outliers. During the survey, we also interviewed the participants to collect some feedback about the general perception of mobile Internet. The survey basically answers the following questions:

- Preferred Internet connection type
- Availability of Wi-Fi access points
- Quality of service
- Amount of time spent on the Internet
- Main activities on mobile Internet
- Price and network reliability perception
- Localization of social media friends

4.5 Challenges

Performing measurements in township areas comes with its set of challenges that we have tried to overcome. First, it is difficult to recruit participants because they are usually uncomfortable about the idea of installing a monitoring application on their smartphone. The school principal actually asked questions about data privacy and how we were going to handle the data collected. Secondly, many people who do have smartphones are absolutely not interested in participating in research studies, even if they were given incentive such as the "double data reimbursement". We therefore concentrated our efforts in recruiting participants from the High School iLearning Centre as well as NGO staff from the Masiphumelele NGO Park. Students were much more interested in participating.

Once the application was installed on the participants' phones, one issue we encountered is that sometimes some of the phones were not collecting any data, whether it was for performance or usage. Those phones were therefore considered as "unresponsive" and were therefore removed from the statistics collected. Out of 23 users recruited, only 14 were active.

Our primary aim was to gather data on mobile Internet usage while avoiding any interventions that might affect usage behavior. Topping up users' mobile phones data as an incentive, would have biased usage patterns. Additionally, some participants did not feel that reimbursing the data used by the MySpeedTest application was a suitable incentive. Some other unresponsive participants said that their mobile phone "was off and they don't really use it as they had no airtime". We decided to exclude those participants from our study.

Finally, the MySpeedTest application separates mobile traffic into cellular and Wi-Fi. However, due to some limitations on the measurement platform, we are currently unable to determine exactly to which access point the client phone has been connected. This limitation does not allow us to separate "public Free Wi-Fi" from "paid Internet Café Wi-Fi". The survey actually helps to fill this gap.

Table 4. List of relevant tables from MySpeedTest database

Table	Description
application	Contains details of the application name and package
application_use	Contains details on the application's network usage (bytes sent and received), whether the application is running in foreground or background
network	Gives information about the network used in a specific measurement. Information such as the network type (Cellular or Wi-Fi), base station ID, GPS coordinates, etc. can be retrieved, if those information are made available.
measurement	Is a metadata table, that contains information about a measurement, the time it was carried out and whether it was manually triggered or scheduled.
device	Contains data on the measurement device. Information such as the network country code, the phone brand and model, software version, data plan type etc. are stored. Confidential data such as phone number are hashed.

5. FINDINGS

We present the results from our empirical measurements and the responses from the survey.

5.1 Measurements

5.1.1 Number of applications

Figure 4 below shows a CDF of the number of applications installed by the 14 participants of this experiment. The number varies from 27 to 55, with a median of 39 applications. For an average user, Google suggests that about 36 applications is the norm [29]. Applications either run in the foreground with user interactions or they are executed as background processes. Some functions such as data backup to the cloud are done in the background. Also, most of the applications installed will also require regular updates, which will ultimately rely on an Internet connection.

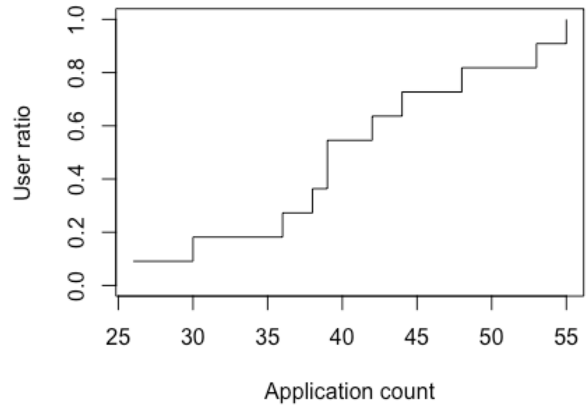

Figure 4. Number of applications installed by users

5.1.2 Daily usage pattern

Figure 5 shows the mean daily usage in MB between Wi-Fi and cellular data across all users for the duration of the study. For 90% of the time, the amount of mobile traffic is slightly higher than Wi-Fi traffic, likely because most of the participants have much easier access to a mobile connection as opposed to a Wi-Fi connection. We can confirm this trend when we look at Figure 6 on the top ten applications and their breakdown in connection types. We see that mobile 3G surpasses Wi-Fi on its own and is far more popular than 2G or 4G, as well.

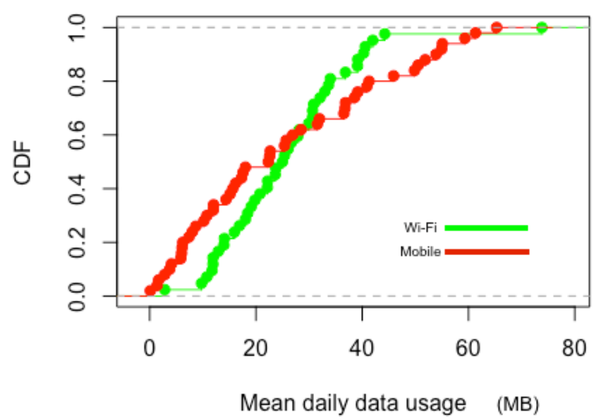

Figure 5. Mean daily usage across all users

5.1.3 Most used applications

We use the amount of data sent and received by application as an approximate proxy for the "popularity" of the application; we do not, however, distinguish between applications running in the foreground versus in the background. Background processes are usually triggered by "administrative" applications such as software updates. However, if the device is infected with viruses and adware, those processes can also run in the background. Figure 6 shows the most used applications on Wi-Fi vs. cellular networks. Mobile 3G is a dominant connection type. The three most used applications are Google Play Store, used to install new applications, followed by Facebook and Chrome.

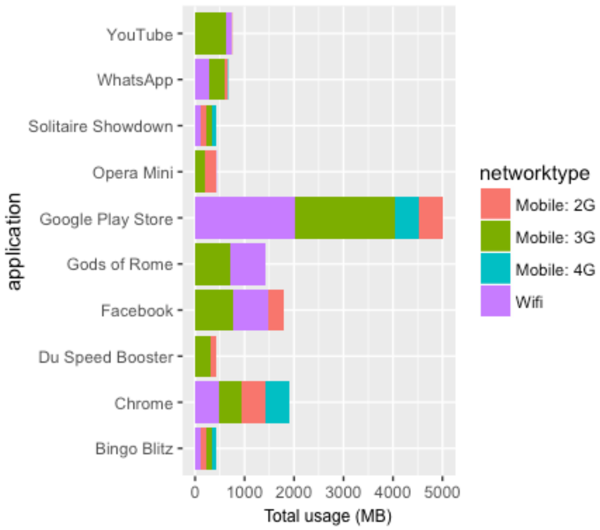

Figure 6. Top 10 applications and breakdown in network connection type and total usage

We were not surprised to see Google Play Store on top of the list. Users tend to install new applications frequently, and updates are downloaded automatically in a background process. Many users do not know that this is the default behavior of their phones and that they should disable updates on mobile data if they want to save on their mobile data usage.

5.1.4 *Popularity by mobile usage*

The bar chart in Figure 7 reveals an interesting trend in the Internet usage for gaming applications. Besides, Google Play Store, Facebook and Chrome, popular applications include Hidden City, Gods of Rome, and Bingo Blitz, which are most popular on the school students' mobile phones. In the future, it might be interesting to make a study of mobile phone usage for games and how gaming behavior affects mobile data usage for people living in those communities. It is interesting to note that, bandwidth greedy applications such as Instagram and YouTube are not very popular both on mobile data and Wi-Fi, as opposed to the current situation with richer demographics, for example, in the USA, YouTube is the second most used application after Facebook [30].

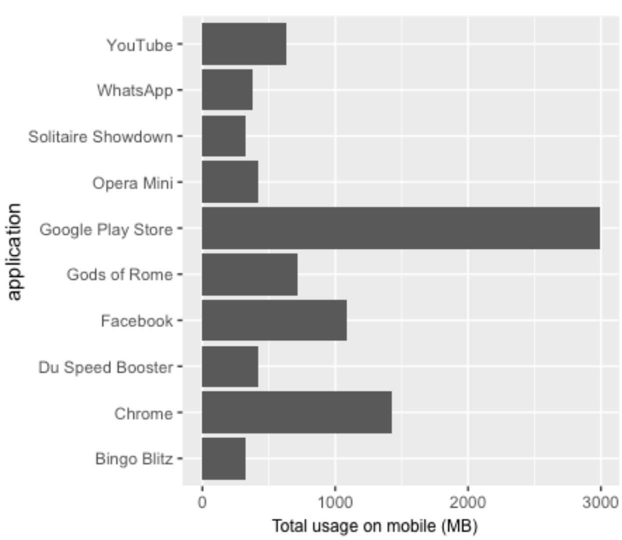

Figure 7. Top ten applications over mobile connections

5.1.5 *Popularity by Wi-Fi usage*

We expected a big difference in the usage patterns for those with access to Wi-Fi. Google Play Store still tops overall usage, but we can see that Chrome is much more utilized on 3G than on Wi-Fi. This result suggests that users tend to spend more time browsing either on news or entertainment, as we found in the survey in Section 5. Interestingly, we found no usage of Opera Mini on Wi-Fi. Opera Mini is usually used to save up bandwidth especially if users are running on limited data bundles. Most probably, Opera Mini is the default browser on some of the participants phone.

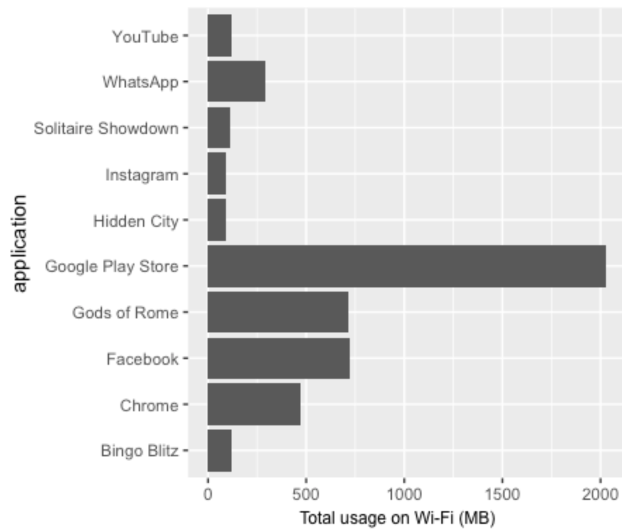

Figure 8. Top ten applications over Wi-Fi

5.1.6 *Usage of zero-rated services*

Currently, mobile operators in South Africa offer three zero-rated services: Whatsapp (ZAR 5 monthly), Freebasics (including Facebook), and Wikipedia Zero (freely accessible using Opera Mini). From the survey we gathered, none of the participants uses Wikipedia Zero or Facebook on Freebasics. Figure 9 shows that Whatsapp is very popular on 3G, irrespective of the mobile operator. Vodacom has the largest share of Whatsapp paid data traffic, compared to Cell-C and MTN. Our data suggests that Cell-C does not appear to offer a big enough incentive to induce subscribers to switch to their network with zero-rated Whatsapp service, but we would need a more detailed study to confirm this observation.

5.2 Feedback from survey and interviews

We individually surveyed the 14 participants and recorded informal feedback on the survey questions, as well. We found that, as expected, all of the devices used were smartphones, with almost 3 out of 14 participants having an additional device such as a tablet that they use to access the Internet. None of the participants has a broadband connection at home, which is not surprising as the deployment of ADSL lines or fiber to the home (FTTH) is not a priority for broadband providers in those areas. Additionally, none of the participants has a mobile data contract plan; all of our users used prepaid data bundles, which also represents the main type of expenditure (78.6%), as opposed to phone calls (11%).

Cellular networks are by far the most used means of Internet connectivity in Ocean View and Masiphumelele. All of the respondents use their capped mobile data to access the Internet on a regular basis. Those who are working in the NGO Park benefit from the free Wi-Fi (not public) installed in there. 4 respondents over 14 claimed they have used it. We believe that even though the participants have access to free Wi-Fi, they do not necessarily use it, as some of them have access to the Internet on their workstations. 7 out of 14 respondents said they do not have access to a Wi-Fi connection, whether paid or free. The students in Ocean View indeed have no other means of accessing the Internet aside from their mobile phones.

Vodacom is the most popular network (6 out of 14), followed by MTN (5 out of 14) and Cell-C (4 out of 14). Almost all the four Cell-C users subscribed to the monthly ZAR 5 unlimited Whatsapp bundle and claimed to have used the Freebasics service at least once. Those two 'zero-rated' services are behind the popularity of this mobile service provider.

In terms of application usage, social media platforms are the most popular. Facebook and Whatsapp top the list followed by YouTube, Gmail, and games. One application that is also very ubiquitous is "ShareIt", a peer-to-peer file sharing application[2]. A few participants (2 out of 14) also mentioned using Opera Mini to browse the Internet. Opera Mini reduces the amount of data transferred by compressing images before they actually reach the mobile phones.

The coverage of mobile data networks and its reliability are both considered "rather fair", although some participants argued that sometimes the "connection is bad" and they had to move to other places to get a better connection.

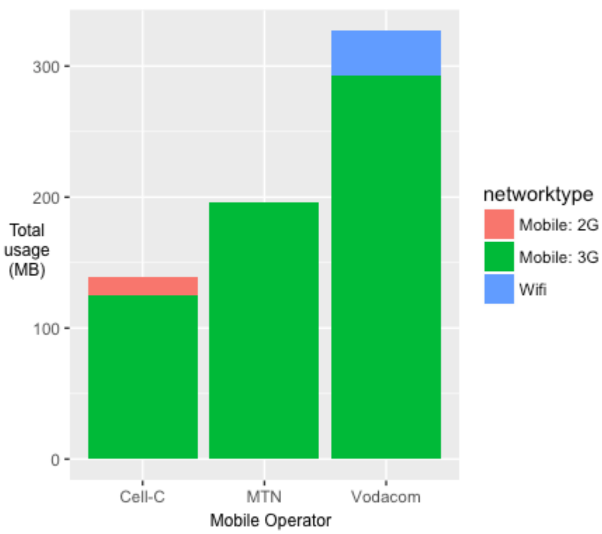

Figure 9. Whatsapp traffic by operators

9 out of 14 participants are opposed to "paying more", even if they would benefit from a better service. 6 out of 14 users argued that they are not quite satisfied with the current level of service provided by their mobile network operator for multiple reasons, price being one of the main reasons. Indeed, almost 9 out of 14 users believe the price of mobile data connectivity is not affordable and is the main reason that discourages them to use their phone to access the Internet. Out of the 14 interviewees, 7 users will typically spend between 60 to 100 ZAR monthly, 5 will spend between 20 to 60 ZAR and the remaining 2 spend between 100 to 200 ZAR. Finally, 9 users agree that keeping track of mobile data expenditure is sometimes problematic, which is consistent with the findings of previous studies.

In terms of social media practices, as we have seen that most of the respondents are very active on a variety of social media platforms. We have tried to understand their interactions with their social media friends; one notable aspect is "locality". 50% of respondents say that they have at least 200 social media friends with 85.7% of them living in the same locality (0 to 10 km), 14.3% living in neighboring communities. Another important aspect to understand was how "nomadic" cellphone users are between the two township communities we studied. We found that there are some movements, such as students from Masiphumelele going to school in Ocean View, or people from Ocean View going to work in Masiphumelele. 4 out the 7 students said that they move at least once or twice in week, the remaining 3 students move on a daily basis. More than 60% of the users would use their mobile phone to access the Internet outside of their home when visiting neighboring places.

6. DISCUSSION

The study reveals interesting trends in the mobile data usage in township communities. It would be interesting to reproduce the same study in another township in South Africa (with a larger sample) and see the outcome. We have found that users in those resource-constrained communities usually do not have the choice of connection type to access the Internet. The more privileged ones will have access to a Wi-Fi connection (if offered freely to the community). For others, they will have to rely on mobile data connectivity where the cost can be relatively prohibitive.

6.1 Most used applications

We found that a large amount of data is used over cellular networks for applications installation and update. Application updates are made available on the Google Play Store and smartphones are automatically synchronized to pull the latest updates. If we multiply the update action by thousands of phones in a community, we end up actually spending a huge amount of data doing the same task i.e. by updating the same set of applications. The idea ultimately, would be to find a solution that allows users to update their phones, without consuming Internet traffic, per se. If we can predict the phone updates and provide them in a "localized" fashion, not only will the updates be faster, but it will also be less costly.

It is also interesting to note that although we do not observe the consolidation into Facebook previously reported in Kenya [11], many of the most used applications are either games or social media. Although our demographic sample consists of 50% students, which could explain the reason behind the high prevalence of games, the survey results also indicated that one of the motivating factors of mobile usage in those two townships is entertainment. As Chirumimalla *et al.* proposed in their paper on "non-productive" activities and desires, the need to have fun and the need for entertainment in one's life is a central developmental aspect [31]. More research is needed to better understand these phenomena.

6.2 File sharing and local traffic

To separate mobile data from local file sharing traffic, we decided not to include the usage of the *ShareIt* application. However, it is interesting to note that 12 out the 14 users were actively exchanging data with their peers using the local file sharing application over Wi-Fi Direct, which enables easy peer-to-peer connectivity. It is much cheaper and faster to share large files locally rather than using an online file sharing service. This

[2] ShareIt does not need data connectivity, as files are transferred to peers using the ad-hoc mode of the phone wireless interface. http://shareit.lenovo.com/

finding confirms what Smyth *et al.* found in their qualitative measurement study on mobile media sharing in urban India [32]. More recently, O'Neil *et al.* explain how file sharing has become a culture part of entertainment in their study of mobile sharing in lower-middle-class Bangalore [33]. We can definitely see similar trends with in our study.

Facebook and Whatsapp are two very popular applications that consume the biggest amount of mobile data, after Google Play Store and Chrome. We have also seen that for a typical user, most of their social media friends are located roughly in the same locality. Therefore, traffic destined for the same place, usually has to travel the world, over expensive links before being sent back to the user living in the same vicinity. A novel approach would be to find a mechanism that can effectively keep local traffic local.

6.3 Zero-rated services

In a resource-constrained setting, one would expect to see a fierce adoption of zero-rated services. However, our study demonstrated that it is not always the case. We found that users have a bad perception of zero-rated services, especially the Freebasics service from Facebook. The fact that Freebasics does not allow users to see pictures and videos does not make the product very attractive. Users actually prefer using "paid" version of Facebook instead of using a "half-cooked" service.

We have also seen that some users would rather use their current subscription of pre-paid mobile data instead of switching onto another network, where zero-rates apply for specific services.

A recent study from the Alliance for Affordable Internet (A4AI) also confirms our findings [34]. They interviewed more than 8000 mobile users across eight countries and found the following:

- Zero-rating did not bring most mobile Internet users online for the first time
- Users typically combine data plans to suit their connectivity needs
- Public Wi-Fi is the primary means of connection for one in five users
- The vast majority of users (82%) prefer access to the full Internet with time or data limitations, if restrictions are imposed.

7. FUTURE WORK

This study did not include more in-depth questions about data usage. For example, we do not know how much data was actively spent (through user interactions) as opposed to data that was passively used for example, by silent application updates. When studying smartphone usage, Falaki *et al.* developed a custom-logging tool that considered the state of the phone while recording usage [33]. This could potentially be a new question to answer that would give more granular information about data usage, especially that we found that the most used application, by far, is Google Play Store (i.e., application installation and updates).

Our study did not directly investigate the needs of users living in those low-income communities. Understand the requirements of the township communities before designing any solution to reduce cost of connectivity in those areas is an important next step.

8. CONCLUSION

Despite significant penetration of cellular data in South African townships, mobile data remains relatively expensive for township residents. In these resource- and income-constrained settings, it is particularly important to understand how people use mobile data so that we can both understand the economic implications of user behavior in these settings and design applications and network architectures that are more cognizant of the high data costs relative to other means of Internet access (e.g., public Wi-Fi). It is also particularly important to understand user behavior in these settings, so that both systems and the broader communities can be designed to help income-limited users reduce the costs they incur for accessing the Internet without sacrificing convenience. Towards this end, we present a first look at mobile data usage in South African townships, using a combination of both quantitative data of application usage from mobile devices and qualitative survey data.

We find that, in contrast to wealthier communities with better resources, township users use cellular data networks relatively more than Wi-Fi networks, both in general and for specific purposes such as application updates. Qualitative survey results suggest that this behavior may be due to the relative inaccessibility of public Wi-Fi in townships. These findings suggest that much work remains, both in the design of mobile applications and network architectures to better optimize cellular data usage and in thinking about how to design communities to increase the accessibility of lower-cost access alternatives, such as public Wi-Fi.

ACKNOWLEDGMENTS

We thank the HPI Research School in ICT4D at the University of Cape Town for financing this research activity. Our gratitude also goes to Torsten Stauch of Appshed and his crew in Masiphumelele and Ocean View for helping recruit students and other mobile Internet users. We are also very much grateful to the Ocean View High School principal, Mr. Klein, for giving us permission to run this experiment on the school premises. Finally, we thank all of the users who participated in this study.

9. REFERENCES

[1] GSMA, "the Mobile Economy Sub-Saharan Africa 2015," *GSM Assoc.*, pp. 1–82, 2015.

[2] A. Gillwald, M. Moyo, and C. Stork, "What is happening in ICT in South Africa: A supply-and demand-side analysis of the ICT sector," *Evid. ICT Policy Action*, no. 7, 2012.

[3] D. McLeod, "Mobile data boom in SA townships," *Techcentral*, 2015. [Online]. Available: http://www.techcentral.co.za/mobile-data-boom-in-sa-townships/56135/. [Accessed: 01-Jun-2016].

[4] World Bank, *Economics of South African townships : special focus on Diepsloot*. 2014.

[5] Strategic Development Information and GIS Department, "City of Cape Town - 2011 Census Suburb Khayelitsha," *2011 Census*, no. July, pp. 1–7, 2013.

[6] S. Muckaden, "Myspeedtest : Active and Passive Measurements of Cellular Data Networks Myspeedtest : Active and Passive," *Proc. ISMA*, no. May, 2013.

[7] A. Mathur, B. Schlotfeldt, and M. Chetty, "A mixed-methods study of mobile users' data usage practices in South Africa," *Proc. 2015 ACM Int. Jt. Conf. Pervasive Ubiquitous Comput. - UbiComp '15*, pp. 1209–1220, 2015.

[8] M. Chetty, S. Sundaresan, S. Muckaden, N. Feamster, and E. Calandro, "Measuring broadband performance in South Africa," *Proc. 4th Annu. Symp. Comput. Dev. - ACM DEV-4 '13*, no. December, pp. 1–10, 2013.

[9] E. J. Saldana, A. Arcia-Moret, B. Braem, A. Sathiaseelan, E. Pietrosemoli, and M. Zennaro, “Alternative Network Deployments: Taxonomy, characterization, technologies and architectures,” *Internet Draft IRTF - Glob. Access to Internet All*, 2016.

[10] N. Rangaswamy and E. Cutrell, “Anthopology, Development, and ICTs: Slums, Youth, and the Mobile Internet in Urban India,” in *ICTD’12*, 2013, vol. 9, no. 2, pp. 51–63.

[11] S. P. Wyche, T. N. Smyth, M. Chetty, P. M. Aoki, and R. E. Grinter, “Deliberate Interactions : Characterizing Technology Use in Nairobi , Kenya,” in *CHI ’10*, 2010.

[12] Nithya Sambasivan, G. Hecht, P. M. Aoki, M. Carrera, and D. Cohn, “SmartBrowse : Design and Evaluation of a Price,” *ICTD2013 Spec. Issue*, vol. 11, no. 1, pp. 21–40, 2015.

[13] T. Kreutzer, “Assessing cell phone usage in a South African township school,” *Int. J. Educ. Dev. Using Inf. Commun. Technol.*, vol. 5, no. 5, pp. 43–57, 2009.

[14] J. Donner, S. Gitau, and G. Marsden, “Exploring Mobile-only Internet Use: Results of a Training Study in Urban South Africa,” *Int. J. Commun.*, vol. 5, p. 24, 2011.

[15] J. Vermeulen, “The massive rise of mobile data in South Africa in 5 charts,” *MyBroadband.co.za*, 2015. [Online]. Available: http://mybroadband.co.za/news/cellular/133090-the-massive-rise-of-mobile-data-in-south-africa-in-5-charts.html. [Accessed: 01-Jun-2016].

[16] S. Sundaresan, S. Burnett, and N. Feamster, “BISmark: A Testbed for Deploying Measurements and Applications in Broadband Access Networks,” *Usenix ATC*, 2014.

[17] SpeedTest, “Ookla SpeedTest,” *MyBroadband.co.za*, 2016. [Online]. Available: http://speedtest.mybroadband.co.za/. [Accessed: 20-Jun-2016].

[18] A. Chen and A. N. Feamster, “Understanding Pricing Effects on Mobile Data Usage,” in *CPR South 2016*, 2016.

[19] A. Futter and A. Gillwald, “Zero-rated internet services: What is to be done ?,” *Broadband 4 Africa*, vol. 1, no. September, pp. 1–10, 2015.

[20] R. Sen, H. A. Pirzada, A. Phokeer, Z. A. Farooq, D. Choffnes, and K. P. Gummadi, “On the Free Bridge Across the Digital Divide : Assessing the Quality of Facebook’s Free Basics Service (In Press),” *To Appear ACM Internet Meas. Conf. 2016 Proc.*

[21] C. Kreibich and N. Weaver, “Netalyzr: illuminating the edge network,” *Proc. 10th ACM SIGCOMM Conf. Internet Meas.*, pp. 246–259, 2010.

[22] J. Huang, C. Chen, Y. Pei, Z. Wang, and Z. Qian, “Mobiperf: Mobile network measurement system,” 2011.

[23] A. Nikravesh, H. Yao, S. Xu, D. Choffnes, and Z. M. Mao, “Mobilyzer: An open platform for controllable mobile network measurements,” *Proc. 13th Annu. Int. Conf. Mob. Syst. Appl. Serv.*, pp. 389–404, 2015.

[24] “Smartphone OS Market Share, 2015 Q2,” *IDC Research*, 2015. [Online]. Available: http://www.idc.com/prodserv/smartphone-os-market-share.jsp. [Accessed: 01-Jun-2016].

[25] Z. Koradia, G. Mannava, A. Raman, G. Aggarwal, V. Ribeiro, A. Seth, S. Ardon, A. Mahanti, and S. Triukose, “First impressions on the state of cellular data connectivity in India,” *Proc. 4th Annu. Symp. Comput. Dev. - ACM DEV-4 ’13*, pp. 1–10, 2013.

[26] A. Tirumala, L. Cottrell, and T. Dunigan, “Measuring end-to-end bandwidth with Iperf using Web100,” *Proc. Passiv. Act. Meas. Work.*, no. April, pp. 1–8, 2003.

[27] Strategic Development Information and GIS Department, “City of Cape Town – 2011 Census Suburb Masiphumelele,” 2013. [Online]. Available: https://www.capetown.gov.za/en/stats/2011CensusSuburbs/2011_Census_CT_Suburb_Masiphumelele_Profile.pdf. [Accessed: 01-Jun-2016].

[28] J. Burrell and S. Hall, “What Constitutes Good ICTD Research?,” *Inf. Technol. Int. Dev.*, vol. 5, no. 3, pp. 82–94, 2009.

[29] J. Tiongson, “Mobile app marketing insights : How consumers really find and use your apps,” 2015.

[30] R. Allan, “The 60 most popular apps of 2016 (so far),” 2016. [Online]. Available: https://www.surveymonkey.com/business/intelligence/most-popular-apps-2016/.

[31] P. Chirumamilla and J. Pal, “Play and power : a ludic design proposal for ICTD,” *Proc. Sixth Int. Conf. Inf. Commun. Technol. Dev.*, vol. 1, pp. 25–33, 2013.

[32] T. Smyth, S. Kumar, I. Medhi, and K. Toyama, “Where there’s a will there’s a way,” *Proc. 28th Int. Conf. Hum. factors Comput. Syst. - CHI ’10*, p. 753, 2010.

[33] J. O. Neill, A. Arbor, J. Chen, and B. Tate, “The Increasing Sophistication of Mobile Media Sharing in Lower-Middle-Class Bangalore,” *ICTD ’16 Proc. Eighth Int. Conf. Inf. Commun. Technol. Dev.*, 2016.

[34] T. H. E. Impacts, O. F. Emerging, M. Data, S. In, and D. Countries, “the Impacts of Emerging Mobile Data Services in Developing Countries November 2015 Research Brief No 1 : the Impacts of Emerging Mobile Data Services in Developing Countries,” no. 1, pp. 1–10, 2015.