

# Towards a Quantitative Model of Mobile Phone Usage Ireland – a Preliminary Study

**Ronan Farrell, Emmett Carolan,  
Séamus C. McLoone, Seán F. McLoone**

*Centre for Telecommunications Research,*

*Callan Institute, National University of Ireland, Maynooth*

*email : ronan.farrell@nuim.ie, ecarolan@eeng.nuim.ie,*

*seamus.mcloone@nuim.ie, sean.mcloone@nuim.ie*

---

**Abstract—** Mobile phone networks were traditionally voice and SMS messaging networks which made them easy to quantify and model. More recently with the advent of broadband data connections and smart phones, mobile phone networks are being used for a wide range of purposes using ever more bandwidth. This paper utilises data provided by an Irish operator to provide an initial perspective on the current uses of mobile phone networks as a within Ireland. This work serves as a precursor to the development of a quantitative model of existing behaviour which in turn will help predict future requirements.

**Keywords –** mobile phones, classification, optimisation.

---

## I INTRODUCTION

Mobile phones have become pervasive in our society with penetration rates near or exceeding 100% in many developed countries. However the use of mobile phones has changed significantly over their short history. In the early 1990's, mobile phones were used solely for voice. With the advent of GSM phones, we saw the development of SMS messaging with over 13 billion messages sent in 2011 in Ireland [1]. Data services over GSM were limited but showed the potential that was only released with the development of 3G phones and the new generation of smart phones. We are now seeing mobile broadband data services rapidly growing and reducing the need for dedicated voice and messaging services. This trend towards a unified data-only mobile communication system will be enhanced with the new 4G communication technologies where much faster data connections will become available. As the mobile phone networks challenge to become the dominant access technology for the internet, the number of usage modes will continue to increase, with peer-to-peer video services, sensor networks and machine-to-machine communications.

Despite the rapid changes in technology, communications infrastructure changes much slower and most telecommunications operators will soon maintain multiple mobile communication schemes (2G, 3G and 4G technologies) on a number of different frequency bands (potentially 900, 1700, 1800, 1900, 2100, and 2300 MHz). Each communication scheme and frequency has its own

strengths and weaknesses and may be optimally deployed for different usages. To optimally manage the network infrastructure for a rapidly changing user base, it will be necessary to partition traffic across the various networks. To enable this optimisation, models will need to be developed that allow the effect of such partitioning to be evaluated and quantified.

As a precursor to the development of such a model, it is important to have a quantitative appreciation for the use to which existing and future networks are being used. There has been some work done on classification of mobile data usage but this has either been from a coarse overview perspective or from a more sociological perspective. This paper presents an initial investigation into the usage modes for mobile phones within Ireland, with a classification of the data usage into particular use categories. The objective of the paper is to provide a basis for the development of such a model and the subsequent challenge of heterogeneous network optimisation.

The paper will first present a background to the existing heterogeneous network as rationale for the work. This is followed by a description of the data used for the analysis. Some initial insights from the analysis are then presented.

The data for this work has been provided by Meteor as a partner in the SFI Centre for Telecommunications Research (CTVR). No information has been provided that would allow for the identification of individuals or the extraction of

other personal information. Finally for the convenience of the reader, a brief nomenclature is provided in Table 1.

## II HETEROGENOUS MOBILE NETWORKS

Mobile phone networks are often presented as single technology networks, such as GSM or LTE for example. In practice networks consist of a variety of technologies that are integrated to offer users a seamless service (Figure 1). This manifests itself in the technology used by the user and the operator. Most 3G phones offer both a 2G and a 3G service, across multiple bands. This is to ensure service where the 3G service is insufficient. Similarly the operators must integrate the various technologies so that client devices can seamlessly move between technologies and that communications can be connected between devices on different networks. This has resulted in packet and circuit switched networks being integrated. Emulation of circuit switch behaviour (typically voice) by the packet switched networks is required to allow integration with both existing POTS telephone systems and

BSC	Basestation (BS) Controller (2G)
RNC	Radio Network Controller
SGSN	Serving GPRS Support Node
MME	Mobility Management Entity
POTS	Plain old telephone service
GSM	Groupe Spécial Mobile
LTE	Long Term Evolution
CDR	Call Data Records
AMR	Adaptive Multirate
GPRS	General Packet Radio Service
wCDMA	Wideband Code Division Multiple Access
HSPDA	Hi-Speed Download Packet Access

Table 1: Nomenclature of Terms

older mobile communications systems. The complexity of these networks continues to increase as new technologies develop with pressures landing on different aspects of the network as applications develop. For example, the recent moves to smart-

phone apps has increased the signalling and mobile management aspects of the SGSN requiring significant new investment.

Telecommunication operators therefore have the challenge of optimising their network for capital investment and operational costs. Each network technology has different characteristics that may make them more suitable for some forms of communications than others. As a brief overview of these characteristics:

- **2G/GSM @ 900MHz:** is mature, capital costs are low, installed infrastructure, high quality voice network, large coverage range, some data services, but very limited. Inefficient use of spectrum (spectrum efficiency of 0.17 to 0.45 bits/Hz/site). Low energy consumption
- **3G/wCDMA @ 1900 MHz:** capital costs are still significant, can deliver high data rates for users (up to 8.4 with HSDPA), poor range. Higher energy consumption, especially in data-mode.
- **4G/LTEadvanced:** capital costs will be very high for early deployments, could offer up to 1 Gbps of data per site. Will be energy expensive. Short range for high bandwidth services.

The situation will become more complex as technologies become frequency flexible and the regulatory environment evolves. This however is still a few years away.

## III THE DATA

This research is based on call-data-records (CDR) data provided by Meteor for a period of time in 2011. The data provides a record of all the data events requested by users, detailing the basestation sector that serviced the request, the time and duration of the request, the access technology, the data

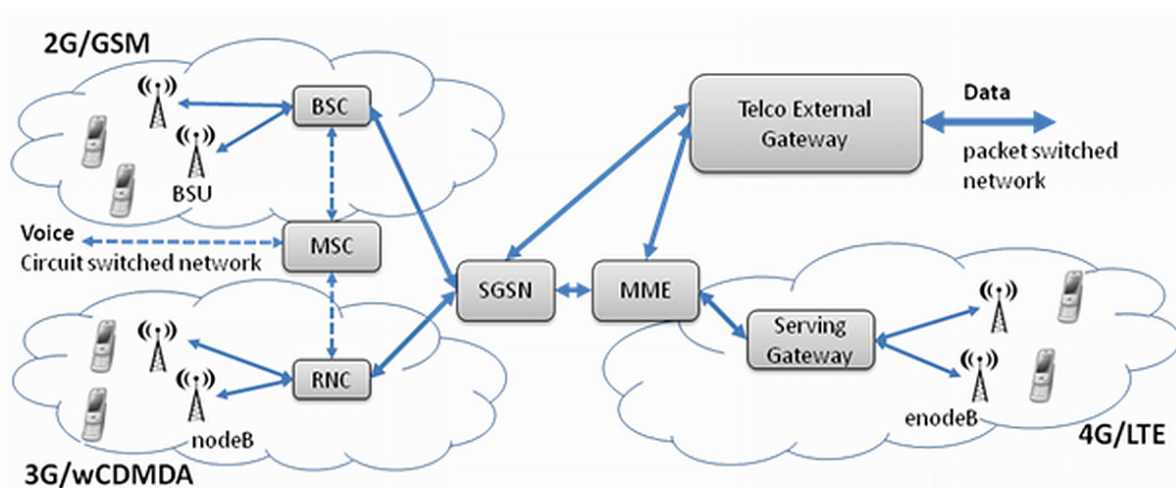


Figure 1: Architecture of a Heterogenous Mobile Telecommunications Network

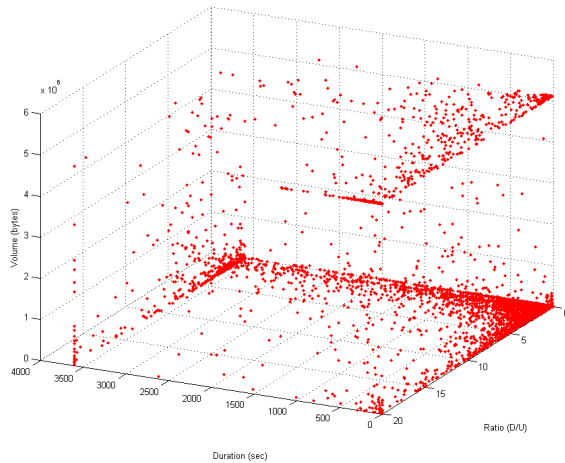


Figure 2: Plot of the data usage characteristics for one day (sampled 1:1000).

uploaded and downloaded in that session. All personal information was removed. During the period of investigation, approximately 1 million users made a phone call, sent a text message or used the data network. In addition, over 400,000 users utilised a data connection, with an average of 6.3 million data sessions per day.

These different use modes and data types are treated differently in different network technologies. To provide insight on the overall usage requirements of the network, they have been converted into their bit-rate equivalent. Voice and SMS traffic have been included as comparable data services – under the assumption that they would be treated as such in a pure packet-switched network, for example LTE. Voice is encoded in mobile phone networks using adaptive multirate (AMR) codecs. In GSM and wCDMA, a narrowband AMR scheme is used and with a typical data rate of 12.2 kbps [6]. A higher quality wideband AMR is used in LTE and offers superior quality at a data range of 12.5 kbps [6, 7].

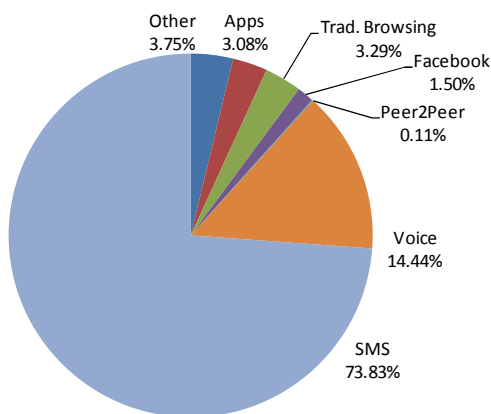


Figure 3: Distribution of communication events by type for the period April 1st - 11th, 2011.

Higher and lower data rates are possible, but for this paper a rate of 12.5 kbps will be used in converting voice channels to an equivalent data session. Text message will be treated as a 1 kB message with 1 second duration. This is excessive, but in practice these messages are so small that individual data rates and durations are meaningless. Multimedia messaging has not been included as it is negligible since the advent of 3G networks.

#### IV THE ANALYSIS AND INITIAL INSIGHTS

The work in this paper is a preliminary analysis of the traffic through the network. The objective is to classify the usage types and to then ascertain the relative impact of these usage types on the mobile phone network. There is other published work in this area [2-4] for other countries, but every society is unique in the rate of technology update and the use of mobile communications media. This work is the first quantitative analysis of the data traffic across the mobile networks for Ireland.

The key parameters for classifying the data are duration, data-throughput, data-rate, the ratio-of upload to download, mobility, time-of-day and location. Mobility is of particular interest due to the challenges of ensuring seamless handover between cells while mobile. Unfortunately, we were unable to complete the mobility analysis in time for this paper and will be part of our future work.

Figure 2 shows the clustering of activities according to duration, download/upload ratio and data volume. The plot suggests that there are a number of clusters representing usage modes. With some additional analysis, the following categories of usage can be identified:

- **Short rapid communications (Apps):** these activities correspond to very short, small quantities of data, generally less than 10 kB and for less than 60 seconds. This is representative, of GPS updates, app interactions, advertising

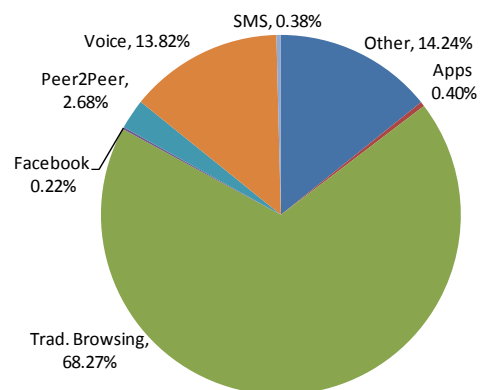


Figure 4: Distribution of communication events by data volume for the period April 1st - 11th, 2011.

updates. Typical phones: Samsung GalaxyS, HTC Wildfire/Desire, Nokia C3.

- **Long duration, large volumes, mostly download (traditional):** Long duration, large quantities of data, mostly download: usually several Mbytes and over a long time. This is the traditional asymmetric usage mode of downloading webpages, videos and other media consumption. Typical Phones: USB Dongles, Samsung GalaxyS, HTC Wildfire/Desire
- **Fast, high-data rates, mostly download, medium duration (YOUTUBE):** These sessions are short bursts of high speed data, which would be representative of video downloads. Typical phones: overwhelmingly USB dongles.
- **Long time connections, low data volumes, similar upload/download ratios (Facebook):** in these sessions, we're seeing as much traffic uploaded as downloaded and consistently over a period of time as the link does not get terminated. This is indicative of peer-to-peer communication

but with insufficient data rates for voice or video, thus text. Typical phones: Samsung Galaxy, HTC Wildfire, LG GT540.

- **Similar download/upload ratios, significant data volumes, less than 10 minutes (P2P Video/Voice/Files):** this suggests 1:1 communications with equal upload and download rates. Average data rates are 120 kbps. Alternatively it could be file-sharing but the ratios should be more in favour of download than upload in that case. Typical phones: USB Dongles, Samsung Galaxy, HTC Desire.

In the datasets, each network event comes with the mobile phone TAC identifier, allowing phone types to be identified. The phones identified were the most popular ones in those categories for Meteor and are reflective of the operator offerings. They will differ for different operators and are only listed to highlight what devices are used for each service.

Traditional SMS messaging falls into the

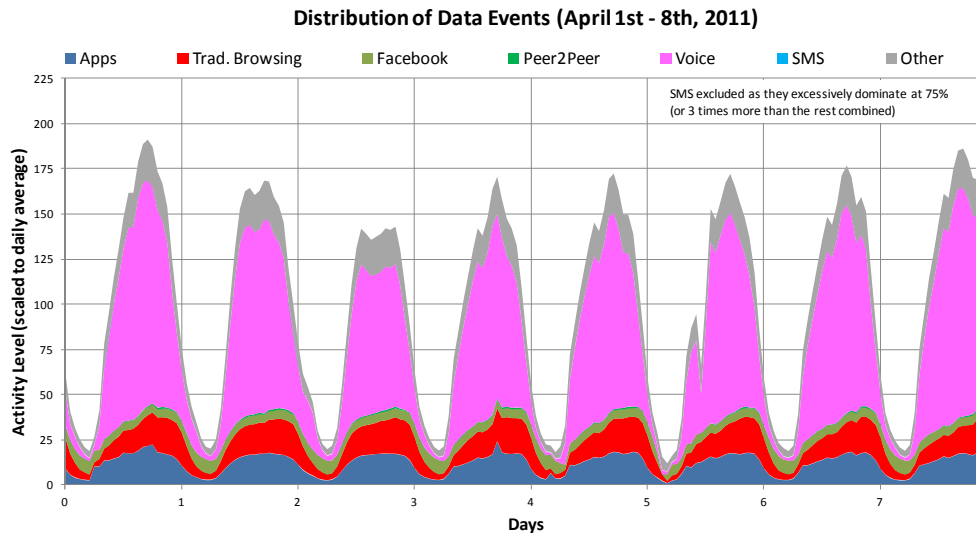


Figure 5: Variation in usage modes over an 8 day period. SMS not included due to large volumes.

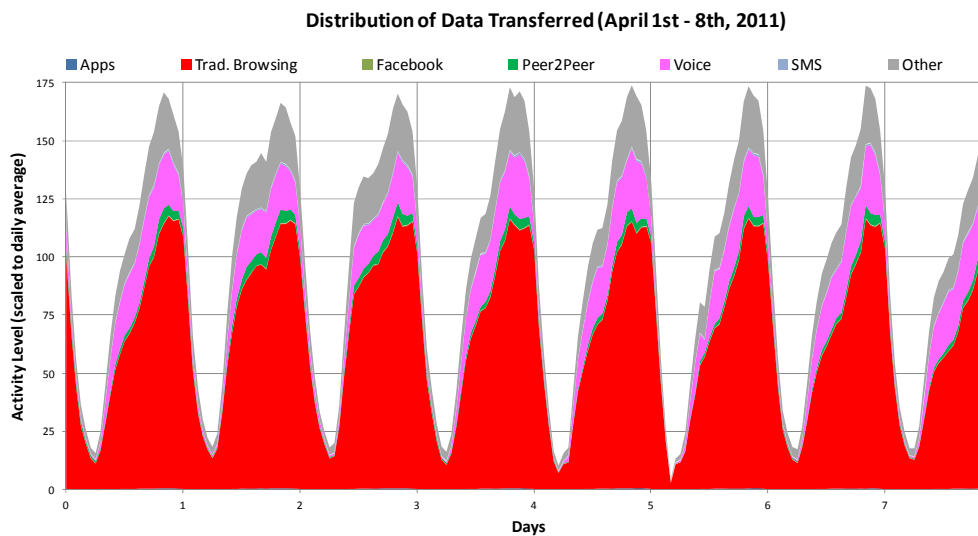


Figure 6: Distribution of the sources of the data generated over an 8 day period.

short-duration category and voice communications as a peer-to-peer communications. We can coarsely classify the bulk of the communications into the these five categories, which are quantified in Table 2. These categories represented over 60% of the activities recorded and 95% of the data transfer (Figures 2 and 3).

	Time	D/U Ratio	Volume
Apps	< 120 s	---	< 256kB
Traditional	> 300 s	>5	> 1 MB
Peer2Peer	<600 s	>0.5, <1.5	> 256 kB
Youtube	< 300 s	>5	> 4 MB
Facebook	> 3590 s	< 3	< 256 kB

Table 2: Classification of Communication Events

Using these classifications, and other information on voice and SMS usage, we can see that these five usage types account for a large majority of the traffic across the network and that the relative weight of each type is highly variable across the day (Figure 3). It is worth noting that the greatest use of

the mobile phone network is still SMS and voice with 80% of the events but only for 14% of the equivalent network data load. The respective weights for these different usages are highly dependent on the replacement rate for the phones and the economic status of the user population. Thus, there may be different variations on the mixture of uses for other networks – depending on their usage base and the financial incentives that may be organised.

Figures 7 and 8 show the variation in these different categories over time. We did not have available the voice and SMS figures for 2010 however they have reportedly stayed constant over the period. It is interesting to note that over 2010 there was little change in data usage, but the six months leading to April 2011 saw a 120% increase in network data volume. Most categories of usage saw significant increases but 75% of the data transferred can still be categorised as traditional browsing.

Data Sessions by Type

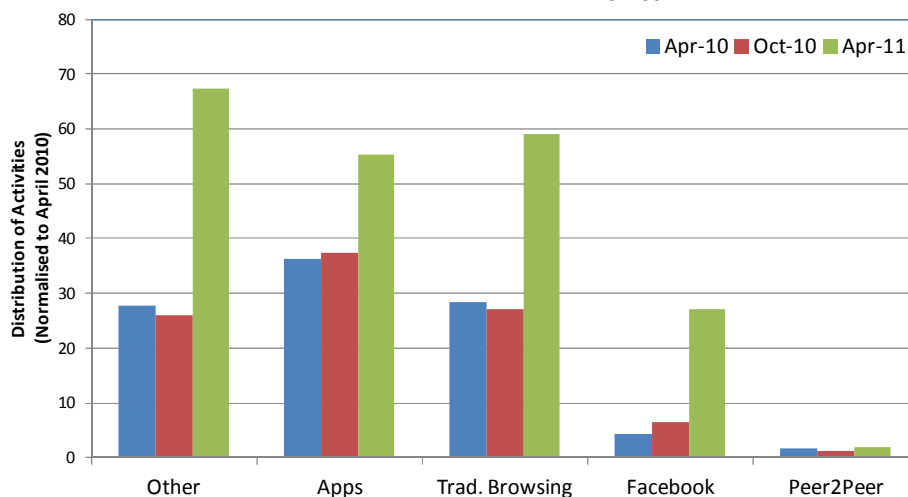


Figure 7: Distribution of communication events by type for the period April 1st - 11th, 2011. SMS events not included as they account for 75% of all traffic.

Data Sessions by Data Transferred

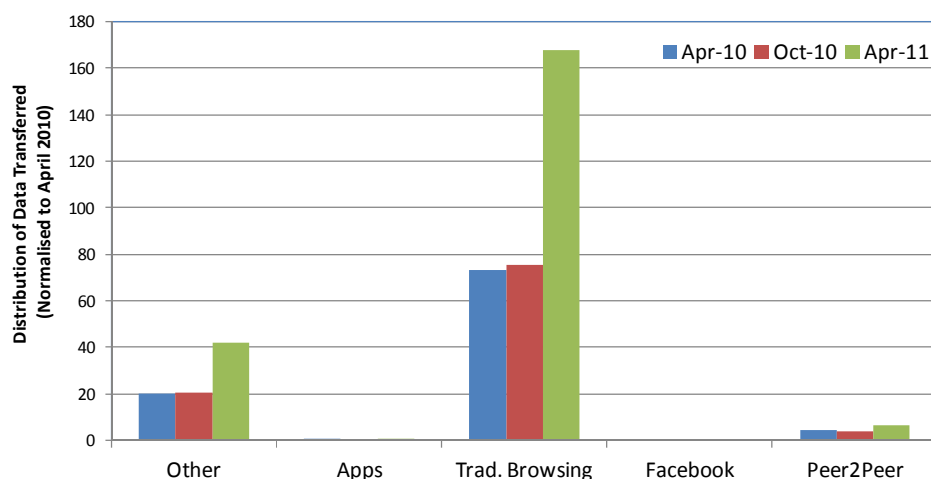


Figure 8: Distribution of communication events by data volume for the period April 1st - 11th, 2011.

## V STEPS TOWARDS A QUANTITATIVE MODEL

This early analysis provides us with some insights in the direction that a quantitative model needs to take. The next step will be to undertake a more rigorous clustering of the usage modes to identify the key modes of usage. These usage modes will need to be correlated with phone types and information about the users - such as age.

One of the key challenges for a mobile operator is the provisioning of redundancy in the network to support mobility. If we can correctly partition usage between nomadic (mostly stationary but moving between locations) and truly mobile, then substantial savings can be made on network provisioning. These initial results suggest the tantalising prospect that the bulk of the data traffic is nomadic in nature and that true high bandwidth mobile usage is rare. This could change with the development of peer-2-peer video calling services.

The other step required in any quantitative model is to identify not only the quantity of data that flows but when and where. Figures 4 and 5 display a time-shifted circadian rhythm to the usage of the mobile phone network - peaking typically between 8-11PM. Other recent work [5] suggests that there is also a strong location dependency on usage modes with different parts of an urban region demonstrating very different usage modes. This is most likely to do with different socioeconomic and demographic characteristics for these regions, but this needs to be explored further. In summary some of the issues that need to be considered are listed below:

- **movement patterns:** work and recreation patterns - for example commuters travelling to work
- **environment:** some early work suggests that high density communities communicate more per person than lower density communities.
- **age:** anecdotally young users are more intensive than older users - this needs to be quantified and verified.
- **prosperity:** smart phones and data services require money yet studies have shown that modern phones are considered a necessity and not a luxury in many societies. The impact of this is unclear.

Once a traffic model has been generated and tested, the next step will be to use it to optimise the traffic flow across the different networks that are available to a mobile phone operator. This could lead to both improvements in service and reductions in cost and energy consumption.

## VI CONCLUSIONS

The purpose of this paper was to provide some insight into the use of mobile phone networks and to identify the structure for a quantitative model. It is clear that most uses of the network can be categorise as either transitory, peer-to-peer or traditional internet browsing. Even though this paper only presents a preliminary analysis, it shows the significant growth in data usage and the now-dominant impact on the network. It shows that traditional browsing behaviours still dominate the data traffic demand though customers are using their phones for a growing variety of other lower-bandwidth applications.

## VII ACKNOWLEDGEMENTS

This work has been supported through the SFI Centre for Telecommunications Research (SFI-CE-I1853). We gratefully acknowledge the support of Meteor for all their assistance with this project.

## REFERENCES

- [1] COMREG's Mobile Telecommunication Market Statistics, [www.comstat.ie](http://www.comstat.ie)
- [2] R.E. Rice, J.E. Katz, "Comparing internet and mobile phone usage: digital divides of usage, adoption, and dropouts", Elsevier Telecommunications Policy, Volume 27, Issues 8-9, pp. 597-623, 2003
- [3] Santi Phithakkitnukoon, Teerayut "Activity-Aware Map: Identifying Human Daily Activity Pattern Using Mobile Phone Data", Lecture Notes in Computer Science, Human Behaviour Understanding, vol. 6219/2010, pp. 14-25, DOI: 10.1007/978-3-642-14715-9\_3, 2010.
- [4] Z. Shafiq, L. Ji, A. Liu, J. Pang, and J. Wang, "Characterizing Geospatial Dynamics of Application Usage in a 3G Cellular Data Network", Proceedings of IEEE INFOCOM, Orlando, FL, March 2012.
- [5] U. Paul, "Understanding traffic dynamics in cellular data networks", Proceedings of IEEE INFOCOM, 2011.
- [6] B. Bessette et al., "The Adaptive Multi-Rate Wideband Speech Codec (AMR-WB)," IEEE Trans. Speech and Audio Processing, vol. 10, no. 8, pp.620-36., 2002
- [7] H. Taddei et al., "Evaluation of AMR-NB and AMR-WB in Packet-Switched Conversational Communications," Proceedings of the IEEE International Conference on Multimedia and Expo, 2004.