Comparing and Contrasting Smartphone and Non-Smartphone Usage

Emmett Carolan, Séamus C. McLoone and Ronan Farrell

The Callan Institute,

National University of Ireland Maynooth, Maynooth, Co. Kildare

E-mail: {ecarolan; seamus.mcloone; ronan.farrell}@eeng.nuim.ie

_____

Abstract— **This paper presents a large-scale network wide analysis of both user behaviour and network dynamics in a live nationwide 2G/3G network. The employed data set tracks over one million subscribers across upwards of ten thousand base stations covering the entirety of the Republic of Ireland. The data set was captured in 2011 and includes a large group of bill-pay and pre-pay smartphone subscribers. The proliferation of smartphones has been a major and recent change to networks worldwide. In light of the recent changes in network access technologies much of the earlier work in this field is now out of date. This paper compares and contrasts the traffic usage of smartphone subscribers and non-smartphone subscribers. A key aim of this work is to quantify and qualify the change brought about by smartphones. Unlike previous studies, we look to the future by treating all traffic including voice calls and SMS as an equivalent data service, as will be the case in 4G. This paper also explores the temporal and spatial properties of both bill-pay and pre-pay smartphone usage and concludes by summarising our key findings & their implications.**

Keywords – Call Detail Records, Smartphone Usage, Mobility, Network Analysis

_____

I   INTRODUCTION

In the past two decades mobile phones and devices utilising the mobile phone network have become ubiquitous in modern society. Mobile phone penetration has approached and in some nations exceeded 100% [1]. Concomitant with this surge in popularity there has been an ever increasing diversification in phone use. In the early 1990's, mobile phones were used solely for voice. The advent of GSM phones paved the way for SMS messaging which saw over 12 billion messages sent during 2011 in Ireland alone [2]. Data services over GSM were limited but showed the potential that is only beginning to be realised with the development of smartphones operating on 3G & 4G.

It is expected that the large changes in how the network is used will result in an enormous growth in the demands placed on the network [3]. The ever increasing adoption of smartphones places new and unique demands on the network infrastructure. In order to meet these demands, it is important to understand the network's traffic dynamics and how smartphone subscribers interact with the network. These key components of resource planning and allocation must be understood if the network is to respond to the challenges it faces.

Several studies have examined spectrum usage and application characteristics in cellular data networks [4, 5, 6]. Much of this prior work uses small scale measurements taken from a few mobile clients. These measurements are then used to understand wireless spectrum usage and to characterise network capacity and performance. However, to understand the entire network usage pattern and subscriber behaviour, it is necessary to perform an analysis of a large scale network-wide data set.

Some studies have been based on a network-wide analysis such as [7, 8] but these studies only consider voice traffic or a user's browsing profile. Studies such as [9] explore network traffic while [10] and [11] explore predictive models of users attributes and communication. More recently, [12] explored data traffic dynamics across an entire network. Although authoritative, the dataset used in [12] was collected in 2007. This is a significant drawback as it predates the widespread adoption of smartphones and the changes this has brought to the network itself and also to user behaviour.

The focus of this paper is to provide an up-to-date measurement-driven examination of smartphone user behaviour on a nation-wide 3G network. A key component of this work will be comparing and contrasting the usage of smartphone subscribers with more traditional subscribers. Particular attention is paid to smartphone subscribers due to the disproportionate impact they have on the network. Unlike other authors we do not limit ourselves by only considering one portion of the network but instead utilise all available data including calls, SMS & cellular data. The goal of this paper is to provide answers to important questions regarding smartphone subscriber mobility, smartphone traffic patterns and smartphone subscriber's spateo-temporal behaviour. Our data set is one full week of nationwide Call Detail Records (CDRs) collected in 2011 from one of the Republic of Ireland's cellular phone networks. The data set includes information on all calls, SMS and cellular data usage of over one million people communicating on a network comprised of over ten thousand base stations.

In the next section of this paper, smartphone subscriber traffic dynamics are examined. Mobility and spatial characteristics of smartphone subscriber usage are then compared and contrasted with that of non-smartphone phone subscribers in section III. Section IV inspects the temporal characteristics of both the non-smartphone phone subscribers and smartphone subscribers. Finally, in section V the implications of the preceding results and observations are reviewed and discussed.

II SUBSCRIBER TRAFFIC

This section concentrates on traffic production from the perspective of the subscribers (in this case, smartphone pre-pay, smartphone bill-pay and non-smartphone).

a) Expressing the Data and Identifying Smartphone Subscribers

For the purposes of this paper voice and SMS are expressed in terms of equivalent 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 multi-rate (AMR) codecs. In GSM and wCDMA, a narrowband AMR scheme is used with a typical data rate of 12.2 kbps [13]. A higher quality wideband AMR is used in LTE and offers superior quality at a data range of 12.5 kbps [13, 14]. 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 200 byte message with 1 second duration. This is an approximation, 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.

Subscribers were matched with their connection device using a TAC (Type Allocation Code) supplied for each subscriber. In this way the identity of smartphone and non-smartphone subscribers could be accurately determined. Smartphone subscribers were further categorised as either bill-pay or pre-pay using subscription information provided by the network provider.

b) Traffic by Subscriber Group

Fig.1 shows the CDF (Cumulative Distribution Function) of the total equivalent traffic generated per subscriber by subscriber group. It highlights the large disparity in traffic generated by members of each group (bear in mind the x-axis is in log scale). Bill-pay smartphone subscribers are clearly using the most traffic with the top 50% consuming almost 10MB or more of total equivalent traffic (cellular data + SMS + calls) per day. Interestingly however, 98% consumed less than 100MB per day. The top 2% of heavy bill-pay smartphone users consumed between 100MB and 3GB per day or between 10 and 300 times the median value for their group.
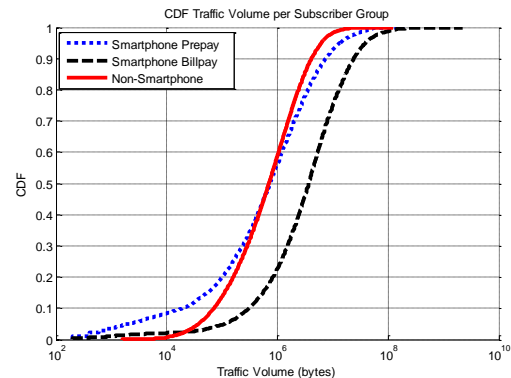

Fig. 1: CDF of total equivalent traffic volume generated by subscriber type

There is an interesting difference between the traffic generated by smartphone subscribers as a function of their subscription type. For example, *the mean smartphone bill-pay subscriber consumes approximately 12MB of total equivalent traffic per day compared to 3.2MB for pre-pay smartphone users.* On this network bill-pay subscribers are more likely to have more expensive smartphones with a higher feature set than pre-pay subscribers. However, this difference in hardware does not account for the disparity in traffic generated. This was found by comparing the traffic generated by bill-pay and pre-pay subscribers who access the network on the same handset. In this instance, the bill-pay subscribers still generated approximately four times more traffic than pre-pay subscribers using the same handset.

Fig. 1 also illustrates *that the bottom 40% of pre-pay smartphone subscribers generate less traffic than the bottom 40% of non-smartphone phone users.* This bottom 40% of pre-pay smartphone subscribers generally use little or no cellular data. Most of the traffic associated with this cohort is the making and receiving of calls and SMS. As shown in Table 1 *the bottom 40% of pre-pay smartphone users are making and receiving significantly less voice calls than the bottom 40% of non-smartphone phone subscribers.* It is possible that these pre-pay smartphone users are communicating in different ways across the network or Wi-Fi (VoIP, instant messaging services, etc.). This could have significant implications for future revenue streams and pricing structures for service providers.

Fig. 2 shows the distribution of total equivalent traffic generation within each of our three groups. A heavy skew in traffic generation is again evident even when examined on the group level. *For example, the top 1% of bill-pay smartphone users generate over 30% of that groups total traffic.* For non-smartphone users, this figure drops to approximately 10%.

Subscriber Type	Mean voice call time in seconds (sent + received)
Pre-pay smartphone (bottom 40% of traffic generators in this cohort)	59
Pre-pay non-smartphone (bottom 40% of traffic generators in this cohort)	102

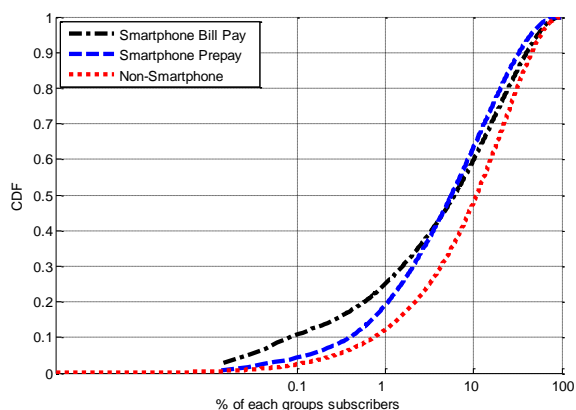Table 1: Mean SMS and voice call use by subscriber type

Fig. 2: CDF of normalised traffic over the percentage of subscribers who generated it broken down by group

III MOBILITY & SPATIAL CHARACTERISTICS

This section will examine mobility and spatial characteristics from the perspective of individual subscribers. Two key metrics of subscriber mobility are examined with the results broken down by subscriber type. Subscriber mobility is an important issue to understand from a network providers perspective as it has implications for hand over management, resource provision etc.

a) Base Stations Visited

Fig. 3 shows the CDF of the number of distinct base stations visited by each subscriber in a day broken down by subscriber group. The figure shows that *smartphone users as a whole connect to a larger number of distinct base stations per day than non-smartphone users.* Bill-pay smartphone users connect to the largest number of distinct base stations in a day (mean distinct base station visits 7.2) compared to their pre-pay counterparts (mean distinct base station visits 6.9). *Non-smartphone phone users are the least mobile group* with about 52% of subscribers connecting to 4 or less base stations. *About 95% of all users visit less than 20 cells, with the remaining 5% of highly mobile subscribers visiting between 20 and 140 cells.*
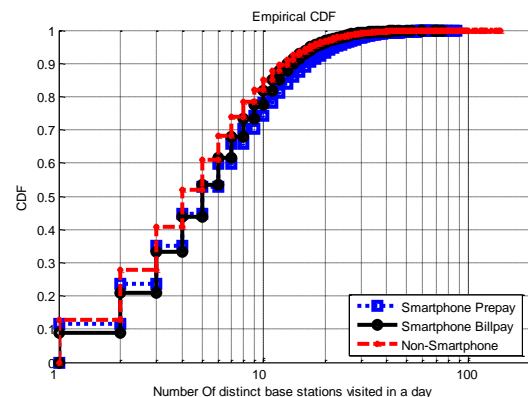

Fig. 3: CDF of the number of distinct base stations visited by a subscriber per day broken down by subscriber group

b) Radius of Gyration

To quantify the physical distance travelled by subscribers we use a measure called the radius of gyration [15]. The radius of gyration is a measure of the linear size occupied by a subscriber's trajectory. It is calculated by averaging the displacement of the recorded locations of the subscriber from a central point. This central point is the centre of mass of the entire trajectory. It is important to note that this captures how widely the subscribers move and not the actual distance travelled. Thus, if a subscriber were to travel in a circle revisiting the same set of base stations the radius of gyration would not increase. However, if the subscriber travelled in a straight line their radius of gyration would increase. Radius of gyration has traditionally been used to study human mobility [15] and has recently been

used to classify mobile subscriber's mobility [12]. The radius of gyration [15] is defined as:

$$r_g = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\vec{r_i} - \overrightarrow{r_{cm}})^2} \qquad (1)$$

where $\vec{r_i}$ represents the i = 1, 2, … n locations recorded for a given user describing his/her trajectory. Recall that the locations are simply the locations of the base stations to which the mobile is connected. The centre of mass point of the user's trajectory is defined as:

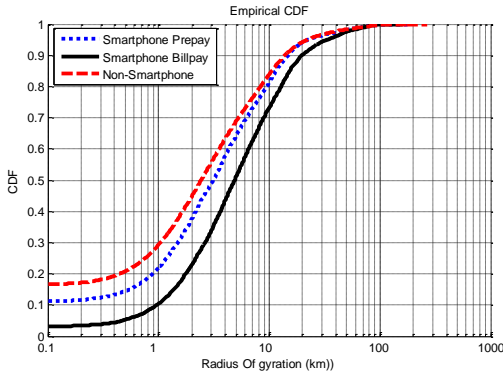$$\overrightarrow{r_{cm}} = \frac{1}{n}\sum_{i=1}^{n}\vec{r_i} \qquad (2)$$

Fig. 4: CDF of the radius of gyration r_g by subscriber category

Fig. 4 shows the radius of gyration r_g broken down by subscriber group. *Bill-pay smartphone users constitute the most mobile group with a median r_g of 5km and a mean r_g of close to 10km. The median r_g for a non-smartphone phone user is 2.5km with the mean value being approximately 6km. Thus, the median bill-pay smartphone user has an r_g two times greater than the non-smartphone phone using counterpart.* Interestingly, 95% of all subscribers have a radius of gyration of less than 50km. 30% of non-smartphone phone users have a radius of gyration of less than 1km. This is significantly higher than the more mobile smartphone group where only 20% of pre-pay and 10% of bill-pay smartphone subscribers have a radius of gyration below 1km.

IV TEMPORAL CHARACTERISTICS

This section examines the temporal characteristics displayed by the individual subscriber groups. In this section the amount of active days in a week, active hours in a day and the total time subscribers (or their phones at least) are actively connected are all quantified. Here, a subscriber is considered as being "connected" or making a connection to the network in three cases 1) during a voice call, 2) when sending/receiving a SMS or 3) when the subscriber has an active cellular data connection open with the network.

Fig. 5 shows the CDF of the number of days that users generate any type of traffic (voice, SMS or data) broken down by subscriber group. *90% of smartphone bill-pay subscribers connect to the network every day of the week compared with 79% of pre-pay smartphone subscribers and 68% of non-smartphone phone users.*
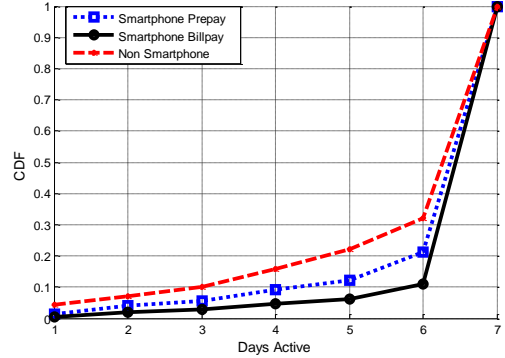

Fig. 5: CDF of number of days users generate any type of traffic on (voice, SMS or data) broken down by subscriber group

Fig. 6 illustrates the large disparity between the number of hours different groups are generating traffic. *Smartphone users are using the network much more often than non-smartphone phone users.* Half of bill-pay smartphone subscribers generate network traffic during 13 or more hours per day, while half of pre-pay smartphone users generate traffic during 12 or more hours per day. 50% of non-smartphone users only generate traffic during 6 or less hours a day.
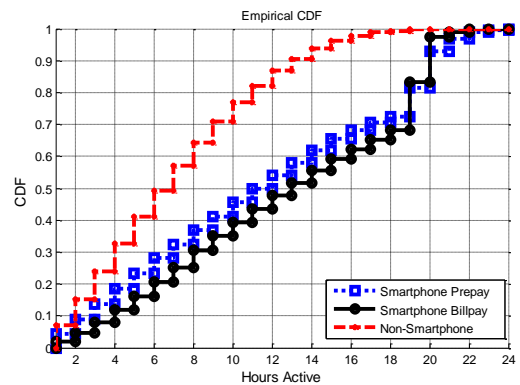

Fig. 6: CDF of number of hours during a typical weekday users generate any type of traffic (voice, SMS or data)

To get a better understanding of the temporal activity of subscribers connected to the network we examine the distribution of the time they spend connected to the network. Time connected is defined as the

duration of all calls made/received and all data sessions routed over the network. Since an SMS does not have duration in the same sense as calls/data sessions we assign each SMS sent/received a duration of one second.
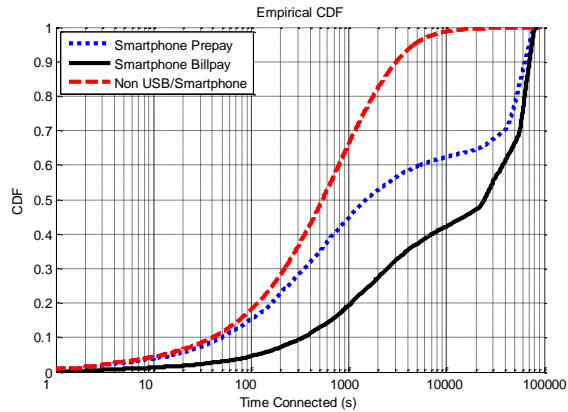

Figure 7: CDF of the time connected to the network per user group

Fig. 7 shows a breakdown of the time connected to the network per day by subscriber group. *Interestingly, the top 50% of bill-pay smartphone subscribers are connected to the network for over 20,000 seconds (5.5 hours) with the groups mean connection time being approximately 30,000 seconds (8.3 hours).* The median bill-pay smartphone subscriber connects to the network for 5.5 hours a day, the median non-smartphone subscriber connects for 10 minutes per day. Thus *the median bill-pay smartphone provider is connected to the network for 33 times as long per day as the average non-smartphone subscriber.*

It was found in section II that the mean bill-pay smartphone subscriber was responsible for about 12MB of total equivalent traffic per day. *Thus, the mean bill-pay smartphone subscriber has a data rate of just 3.2 kbps.* This low data rate is in keeping with the bursty nature of cellular data traffic as identified and discussed in [9]. Thus, users have short sessions of high data rate bursts while e.g. downloading a webpage followed by long periods of inactivity or very low bandwidth usage e.g. reading the web page etc. Thus, there are many subscribers occupying network resources but not using their connection very efficiently.

Interestingly, *the mean pre-pay smartphone subscriber is connected to the network for 19,000 seconds per day (5 hours), 11,000 seconds (3 hours) less than the mean bill-pay subscriber.* Only *10% of non-smartphone phone subscribers use the network for more than 3000 seconds (50 minutes) per day.*

V SUMMARY, OBSERVATIONS & IMPLICATIONS

In this section the key observations of the paper are summarised with a short discussion of the resultant practical implications.

- In our network the mean smartphone bill-pay subscriber generates approximately 12MB of total equivalent traffic per day compared to 3.2MB for pre-pay smartphone users. Interestingly, 40% of pre-pay smartphone subscribers are not using cellular data connections. If service providers wish to entice these subscribers to use their data connections then they may need to make more attractive data packages available to pre-pay customers.

- The bottom 40% (in terms of total equivalent traffic) of pre-pay smartphone users are making and receiving significantly less voice calls than their counterparts using non-smartphone phones (bottom 40% of pre-pay non-smartphone phones). It is possible that these pre-pay smartphone users are communicating in different ways (VoIP, instant messaging services etc.). Using these free alternative communication methods on your handset is becoming feasible for more and more people as Wi-Fi access spreads. This could have significant implications for future revenue streams for service providers.

- The top 1% of bill-pay smartphone users generate over 30% of that groups total traffic. For traditional non-smartphone phone users, this figure is 10%. As smartphone penetration rises will this disparity worsen or improve? From a service providers point of view they may wish to alter pricing structures to reduce this disparity and improve overall network performance for the majority of subscribers.

- Bill-pay smartphone users are the most mobile group followed by pre pay smartphone subscribers. Non-smartphone subscribers are the least mobile group. As smartphones become more prevalent and non-smartphones less so this will have implications for mobility management.

- 90% of bill-pay smartphone users are connecting to the network 7 days a week with a median connection time of 5.5 hours per day over the course of a week. This compares with the non-smartphone phone users where only 68% connect 7days a week with a median connection time of only 10 minutes per day. Thus, the median bill-

pay smartphone subscriber is connecting to the network for 33 times as long as the median non-smartphone phone subscriber per day. However, much of the time smartphone subscribers are connected to the network they are inefficiently using their connection at low data rates due to their bursty data usage. Thus, network providers may wish to come up with new software schemes to reduce this inefficient use of resources.

VI ACKNOWLEDGEMENTS

VII REFERENCES

[1] Y. F. Chuang, "Pull-and-suck effects in Taiwan mobile phone subscribers switching intentions," *Telecommunications Policy,* vol. 35, pp. 128-140, 2011.

[2] COMREG. (2012, 30 Oct). *COMREG's Mobile Telecommunication Market Statistics*. Available: www.comstat.ie/

[3] Cisco, "Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2011–2016," http://www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ns827/white_paper_c11-520862.html2012.

[4] K. Mattar, A. Sridharan, H. Zang, I. Matta, and A. Bestavros, "TCP over CDMA2000 networks: A cross-layer measurement study," *Passive and Active Network Measurement,* pp. 94-104, 2007.

[5] K. Pentikousis, M. Palola, M. Jurvansuu, and P. Perala, "Active goodput measurements from a public 3G/UMTS network," *Communications Letters, IEEE,* vol. 9, pp. 802-804, 2005.

[6] P. Reichl and M. Umlauft, "Project WISQY: A measurement-based end-to-end application-level performance comparison of 2.5 G and 3G networks," in *Wireless Telecommunications Symposium, 2005*, 2005, pp. 9-14.

[7] R. Keralapura, A. Nucci, Z. L. Zhang, and L. Gao, "Profiling users in a 3g network using hourglass co-clustering," 2010, pp. 341-352.

[8] D. Willkomm, S. Machiraju, J. Bolot, and A. Wolisz, "Primary users in cellular networks: A large-scale measurement study," 2008, pp. 1-11.

[9] R. Farrell, E. Carolan, S. McLoone, C., and S. McLoone, F., "Towards a Quantitative Model of Mobile Phone Usage Ireland – a Preliminary Study," presented at the ISSC, NUI Maynooth, Ireland, 2012.

[10] E. Carolan, S. McLoone, S. McLoone, and R. Farrell, "Analysing Ireland's Interurban Communication Network using Call Data Records," presented at the ISSC, NUI Maynooth, 2012.

[11] R. Farrell, "Ascertaining Age from Mobile Phone Usage," presented at the ISSC, NUI Maynooth Ireland, 2012.

[12] U. Paul, A. P. Subramanian, M. M. Buddhikot, and S. R. Das, "Understanding traffic dynamics in cellular data networks," in *INFOCOM*, 2011, pp. 882-890.

[13] B. Bessette, R. Salami, R. Lefebvre, M. Jelinek, J. Rotola-Pukkila, J. Vainio, H. Mikkola, and K. Jarvinen, "The adaptive multirate wideband speech codec (AMR-WB)," *Speech and Audio Processing, IEEE Transactions on,* vol. 10, pp. 620-636, 2002.

[14] H. Taddei, I. Varga, L. Gros, C. Quinquis, J. Y. Monfort, F. Mertz, and T. Clevorn, "Evaluation of AMR-NB and AMR-WB in packet switched conversational communications," in *Multimedia and Expo, 2004. ICME'04. 2004 IEEE International Conference on*, 2004, pp. 2003-2006.

[15] M. C. Gonzalez, C. A. Hidalgo, and A. L. Barabási, "Understanding individual human mobility patterns," *Nature,* vol. 453, pp. 779-782, 2008.