A Measurement of Mobile Traffic Offloading

Kensuke Fukuda[1], Kenichi Nagami[2]

[1] National Institute of Informatics, Japan
[2] INTEC, Inc, Japan

Abstract. A promising way to use limited 3G mobile resources efficiently is 3G mobile traffic offloading through WiFi by the user side. However, we currently do not know enough about how effective the mobile traffic offloading is in the wild. In this paper, we report the results of a two-day-long user-based measurement of mobile traffic offloading by over 400 android smartphone users in Japan. We first explain that the variation of aggregated traffic volume via WiFi is much greater than that via 3G in our dataset. Next, we show that the traffic volume offloading through WiFi is common over whole weekend and weekday night, though weekday rush hours have less chance of traffic offloading. Our results emphasize that a small fraction of users contribute to a large fraction of offload traffic volume. In fact, our per-user level analysis reveals that the top 30% of users downloaded over 90% of their total traffic volume via WiFi. However, bottom 20% of users stuck to 3G only and over 50% of users turned off the WiFi interface in business hours. Also, 17.4% of the total traffic volume was generated by users whose WiFi traffic volume was less than 1MB. We observed that some hybrid users downloaded most of their traffic volume via WiFi in shorter durations. In this sense, there is more room to improve the current traffic offloading by promoting users to use WiFi more effectively. Furthermore, we demonstrate that WiFi offloading is mainly performed by access points (APs) in homes while the use of public WiFi APs is still uncommon in our dataset.

1 Introduction

Smartphones, intelligent mobile phones, are becoming ever more popular around the world. The Ministry of Internal Affairs and Communications of Japan reports that 3G mobile network traffic is now doubling every six months in Japan [13]. This rapid increase in the mobile 3G traffic is a big problem for 3G carriers, because the frequency and bandwidth of the 3G network are limited resources that largely differ from residential FTTH access lines. In addition to the increasing number of users, another reason for this growth is that the monthly fee for a mobile phone is basically a flat-rate. Some 3G carriers have started to force bandwidth capping to heavy-hitters on the basis of their traffic. Furthermore, the 3G carriers promote migration of 3G mobile traffic to high-capacity and less congested fixed networks. For this reason, the offloading of 3G traffic through WiFi (IEEE 802.11{a,b,g,n}) has been attracting more attentions. There are two main usage scenarios of traffic offloading by WiFi. One is to use public

WiFi access points (APs) provided by 3G carriers or other WiFi providers in downtown areas (e.g, cafes, stations, airports) to avoid congestion at 3G base stations. The total number of such public APs provided by 3G carriers is estimated to be over $300,000$ APs in Japan according to their web pages. The other is APs in homes where the high-speed network has been rapidly deployed. In particular, increasing penetration of the fiber access in residential users (over 40%) is reported in Japan [3]. Some 3G carriers started to provide customized WiFi APs to non-professional users so they can easily use WiFi at home. The deployment of WiFi APs in homes accounts for roughly 65% of the total number of residential broadband users.

However, it becomes more and more difficult to understand the behavior of such mobile traffic by traffic offloading, because 3G carriers cannot track such offloaded traffic at their backbone network. Even ISPs providing FTTH services cannot distinguish traffic volume generated by smartphones and others in homes. Thus, in this paper, we intend to characterize the usage of the 3G and WiFi of smartphones in terms of the traffic offloading. We developed special software for android smartphones to measure its usage and collected two day's worth of traffic data from over 400 smartphone monitor users using the measurement software in Japan. The main findings of our measurements are as follows: (1) The traffic offloading in homes is common in our dataset. The total amount of traffic volume via WiFi is much larger than that via 3G. The average traffic offload ratio (i.e., ratio of penetration to WiFi) is 0.64 and the peak traffic offload ratio could reach 0.95, indicating that offloading is effective in terms of traffic volume. (2) However, a small fraction of users contributed a large fraction of traffic offloading. The top 30% of users downloaded over 90% of their traffic volume via WiFi, though 20% of users only used 3G networks. Also, 17.4% of the total traffic volume was generated by users whose WiFi traffic was less than 1MB. In particular, over 50% of users turned off their WiFi interface in business hours, and some hybrid users downloaded most of their traffic volume via WiFi in shorter durations. These results indicate that there is more room for improving the offload by promoting the use of WiFi. (3) WiFi offloading was mainly done by APs at home while public WiFi APs are still not commonly used in our dataset.

2 Dataset and preprocessing

We developed special software to measure the traffic volume via the 3G network and WiFi for android smartphone. It reports the values of the byte and packet of network interfaces of a smartphone to an external server every 10 minutes, as well as the WiFi information (e.g., ESSID, BSSID), 3G network information (e.g., base station information), and device information (e.g., hardware and OS types). For privacy reasons, it does not collect user IDs, GPS information, or application usage. We recruited 435 monitor users who own android smartphones in Japan that were sampled from a thousand potential candidates, considering demography and the market share of the 3G carriers in Japan. Moreover, over 90% of monitor users reported that they have WiFi APs at home. In this sense, the results we will present are likely biased to the behavior of advanced users who have less difficulty using WiFi. The measurement experiment was performed on May 13th (Sun) and 14th (Mon), 2012 (48 hours long).

For preprocessing, we removed the traffic volume by tethering, which means a smartphone simply relays traffic from other devices (i.e., laptop PC) to the Internet, from the dataset. This is because we intended to focus on traffic patterns generated by the smartphone itself, though tethering is a promising application of smartphones. Also, some smartphones have a mobile WiMAX interface (IEEE802.16e) more than 3G and WiFi interfaces, but we removed their traffic.

3 Results

3.1 Global view

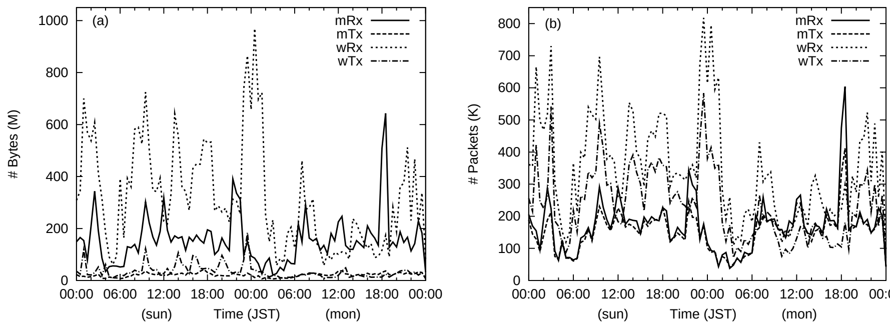

Fig. 1. Traffic variation (bin size = 30 min): (a) bytes and (b) packets.

Figure 1 displays the variation of aggregated traffic volumes and that of the aggregated number of packets in 30-minute bins. Each plot indicates a different type of media: mRx (3G received), mTx (3G sent), wRx (WiFi received), and wTx (WiFi sent). The direction of the traffic is from the view of users (i.e., "received" corresponds to user's download). First, we observed higher WiFi traffic volumes than 3G ones, and the peak of the traffic volume is 1.5 times larger than that of mRx. In particular, the volumes on Sunday are higher than those on Monday, though Monday night is also characterized by high WiFi traffic volume. Thus, the availability of the WiFi network is lower on the weekday in our dataset. Second, we emphasize that peaks in both traffic volumes are not always synchronized, meaning that some users switch the media appropriately depending on the availability. In particular, we confirmed a sharp peak of mRx traffic volume at 6pm on Monday, corresponding to the rush hour in Japan. This peak does not appear in wRx, suggesting that WiFi was hard to use during the rush hours. The same type of non-synchronized peak appears at 9pm on Sunday. The correlation coefficient of time series of mRx bytes and wRx bytes is 0.03, and that of packet based time series is 0.11. Although each user switches between two interfaces exclusively, the variations of aggregated traffic volumes neither positively nor negatively correlate.

The Tx volumes are lower than Rx volumes, likely due to the typical application type of smartphone (i.e., server-client type). In addition, the traffic pattern of Tx resembles that of Rx in 3G packets, while that in WiFi is synchronized but with some gaps. This suggests that typical usage and application of 3G and WiFi are likely different. These results are consistent with the observation that most application traffic is server-client in 3G smartphone traffic [14, 15].

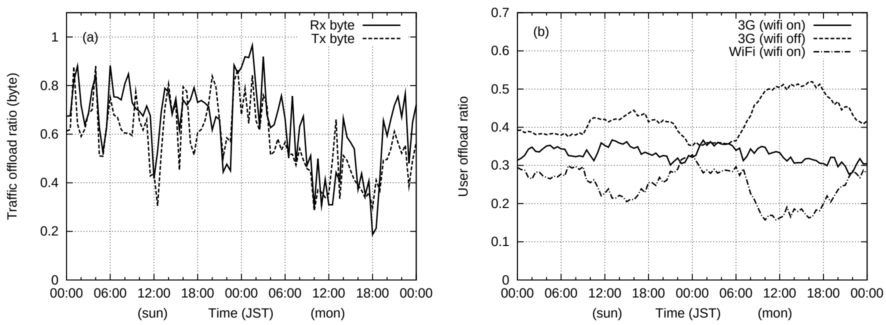

Fig. 2. WiFi offload ratio (bin size = 30min) (a) bytes and (b) users

Next, we investigate the degree of traffic offloading. We define a *traffic offload ratio* as the ratio of WiFi traffic volume to the total volume and a *user offload ratio* as the ratio of the number of WiFi users to the total number of users in 30-minute bins. The ratio closer to 1.0 means the penetration to a WiFi network while that closer to 0.0 means the penetration to a 3G network.

Figure 2 (a) represents the traffic offload ratio over time. The average traffic offload ratio was 0.64 though it varied largely depending on the usage of smartphone; the peak and bottom ratios are 0.97 and 0.19, respectively. The figure highlights the fact that the offloading ratio on Sunday is relatively higher than that on Monday. The average ratio was 0.70 on Sunday and 0.58 on Monday. The lower offloading ratio in the morning and afternoon on Monday suggests fewer opportunities to connect to the Internet via WiFi during work time. As expected, again, the ratio increased on Monday night.

Figure 2 (b) shows the breakdown of users: (1) 3G users whose WiFi interface was also up, (2) 3G users whose WiFi interface was down, and (3) WiFi users. The average user offload ratio corresponding to case (3) was smaller (0.22) than the average traffic offload ratio. We, again, confirm higher ratios during night and lower ones in the afternoon. Only 15% of users connected to WiFi in business hours on Monday, moreover, over 50% of users explicitly turned off their WiFi interface in business hours as shown in case (2). Similarly, the ratio of 3G users whose WiFi was also up is stable (≈ 0.3), indicating that they had few chances to encounter any available APs. In particular, the ratios of 3G users whose WiFi was up and WiFi users in night are closer. This means that WiFi APs were actually effective for almost half of users who turned on WiFi in night. In contrast, only about 35% of users who turned on WiFi interface could download data via WiFi in business hours.

Comparing both figures, we can conclude that the traffic offload was mainly exploited by a relatively smaller number of users. In other words, such heavy users switched their network interfaces explicitly.

3.2 Per-user view

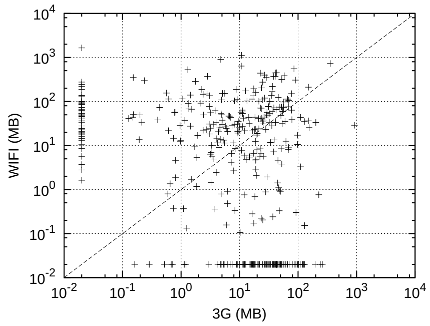

Fig. 3. Scatter plot of 3G and WiFi download traffic volume per user

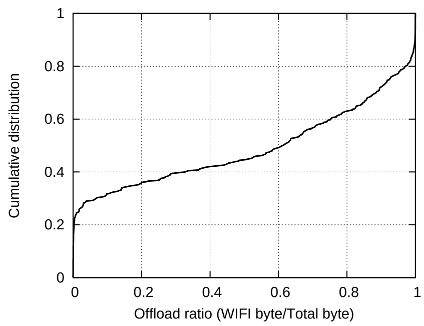

Fig. 4. Cumulative distribution of offload ratio per user

Here, we focus on a microscopic view of traffic offloading. Figure 3 displays the scatter plot of 3G traffic and WiFi traffic volume per user for two days. We confirm horizontal dots in the bottom and vertical dots in the left of the figure, corresponding to the users who only used 3G and WiFi respectively. The former did not use WiFi even at home, and the latter likely saved the fee for 3G network access. A diagonal line in the figure represents users who used 3G and WiFi equally. A non-negligible number of dots below the diagonal, i.e., 3G traffic volume is greater than WiFi traffic volume, show that there is a possibility of increasing traffic offloading. For example, the traffic volume of 3G-only users accounted for 9.6% of the total traffic volume, while that of users whose WiFi traffic is less than 1MB accounted for 17.4% of the total volume.

In addition, Figure 4 displays the cumulative user distribution of the ratio of using WiFi and 3G per user. As explained before, a high (or low) traffic offload ratio corresponds to the penetration of WiFi (or 3G) usage. From the figure, we observe that the 3G-only users accounted for approximately 20% of all users and the WiFi-only users accounted for 10%. The median of users used more WiFi than 3G (0.62). Notably, the top 30% of users switched 90% of traffic volume to WiFi. These results are consistent with the previous results that revealed a relatively small portion of users penetrate to WiFi offloading.

Similarly, Figure 5 shows the relationship between total download traffic volume per user and its traffic offload ratio. We confirm a positive correlation (0.35) between two metrics, indicating that heavy-hitters consume more bandwidth via WiFi and that 3G-only users received less data than offloading users. We conclude that heavy-hitters efficiently use WiFi for their download traffic.

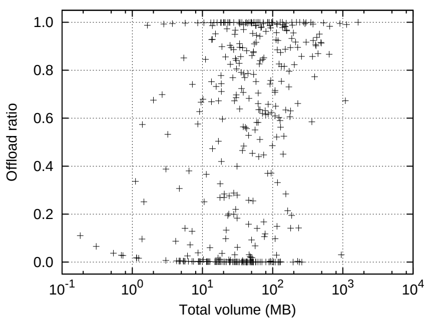

Fig. 5. Total traffic volume and traffic offload ratio

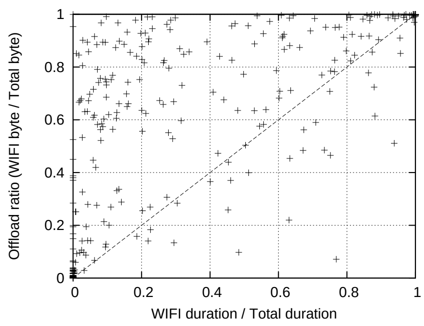

Fig. 6. Ratio of WiFi duration and traffic offload ratio

Finally, we examine the traffic penetration to WiFi and the duration using WiFi interface. Figure 6 displays the scatter plot of the ratio of duration using WiFi to the total duration and the traffic offload ratio per user. The diagonal in the figure indicates the users whose WiFi traffic volume is proportional to its duration. As expected, we see plots concentrated near (0, 0) (i.e., 3G only user) and (1, 1) (i.e., WiFi only user). A notable point, however, is that we still observe plots scattered around lower ratios of the duration and higher offload ratios. This means that these hybrid users downloaded most of their traffic volume via WIFi in shorter periods, consistent with the macroscopic observation in Figure 2.

3.3 WiFi usage

Here, we investigate the location where users associate with WiFi APs. SSID is an identifier of AP in WiFi, and administrators of APs could set their name by themselves, or it could also be left as the default setting. Thus, by categorizing the names of ESSIDs, we could infer the types of location of APs with which users associated. We gathered all SSIDs appearing in the dataset (418 unique ESSIDs) and manually classified them into the following four categories.

- *public* (8 ESSIDs) is SSIDs that 3G carriers freely provide to their customers (e.g., "docomo", "au_WIFI", "0001softbank") and the third-party WiFi carriers provide to their customers (basically at charge) (e.g., "FON", "0033") and administrators freely open to all users (e.g., "freespot").
- *home* (261 ESSIDs) is default ESSIDs when AP manufacturers shipped. Thus, administrators of such APs do not change their ESSID from the default setting. We assumed that such access points are located at home rather

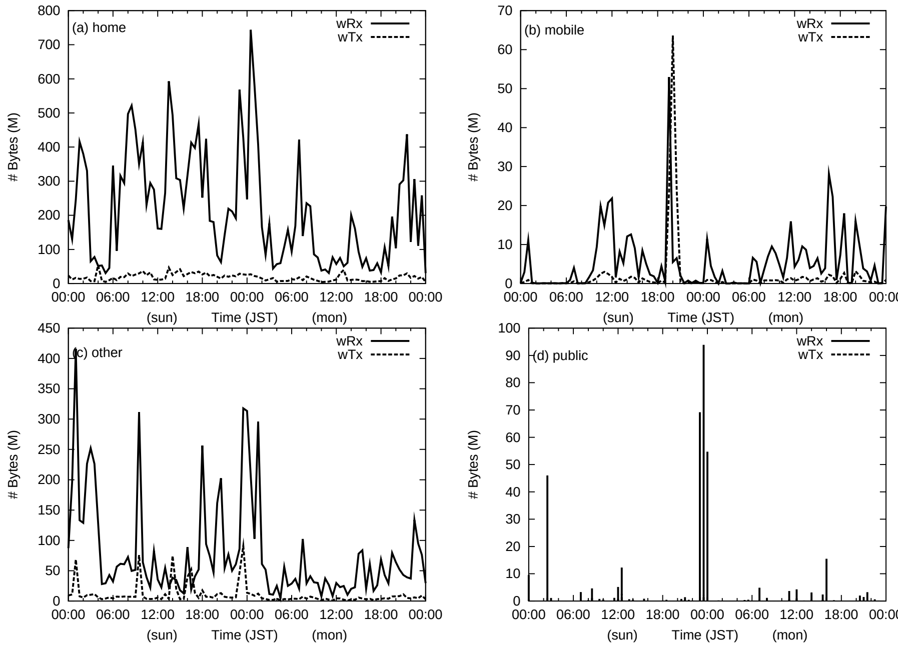

Fig. 7. WiFi traffic usage patterns

than in an office, because the number of devices at home is small and these administrators are likely to be less careful in changing ESSIDs than administrators in office networks.

- *mobile* (19 ESSIDs) is default ESSIDs for a portable WiFi router with a 3G uplink and WiFi down link provided by 3G carriers. The user's smartphone connects to this router via WiFi to obtain an Internet connection.
- *other* (130 ESSIDs) is named ESSIDs, i.e., administrators of APs explicitly changed their ESSIDs. This can be located in homes, offices, shops, etc. Also, it included unclassified ESSIDs.

Figure 7 indicates the variation of traffic volumes for different categories of SSIDs users associated with: (a) home, (b) mobile, (c) other, and (d) public. We confirm that the variation of traffic volume in home dominates the total amount of the WiFi traffic volume shown in Figure 1. Similarly, the traffic variation of the other category is similar to that of home users, indicating that most of these APs are also likely located at home. One interesting point in the mobile category is that its traffic pattern was closer to that of 3G traffic shown in Figure 1 than that of the home category; high traffic in the morning and evening on Sunday and the evening on Monday. The usage pattern of a portable WiFi router is similar to that of the 3G device, indicating that such users save 3G traffic costs by paying the cheaper monthly fee for a portable WiFi router as an alternative. Indeed, the correlation coefficient of wRx bytes of the mobile category and mRx bytes is higher (0.16) than that of wRx bytes of the mobile category and wRx bytes of the home category (−0.01) Also, one unexpected result is a much smaller traffic volume in public WiFi. Sharp and discrete spikes indicate that a small number of users generate traffic volume in a short time; indeed, the biggest peak of the spikes was traffic volume via a FON AP.

In summary, traffic offloading in homes currently works well, though that in public WiFi APs is not very high in our dataset.

4 Related work

There have been many measurement activities to understand wireless network traffic better, including traffic from 3G and WiFi networks.

3G smartphone usage: There have been attempts to characterize 3G smartphone traffic in some countries by measurements at backbone networks or at smartphones. These studies mainly showed the diversity of usage of smartphones in many aspects; differences in device types and carriers [9], user pattern and protocol [4], application [14, 15], geolocation [1], geographical differences [15], and mobility [14, 10, 16]. Related to our work, Ref [14] pointed out the difference in usage of applications depending on the stationarity of users. Our data had no application information, but the penetration of traffic volume to WiFi in homes suggests that the application is used differently inside and outside homes.

WiFi usage: The network usage of campus WiFi networks has also been well studied [8, 6]. They pointed out that the application mixtures in the campus WiFi network differed from those in 3G mobile traffic because a wide variety of devices were connected to the campus WiFi network. Moreover, the WiFi network usage of specialized public transportation has also been analyzed [7]. A recent study of WiFi traffic of hand-held devices focused on home WiFi traffic in residential traffic [12]. It reported that hand-held devices were appeared in up to 3% of residential DSL traffic in 2009.

Availability of 3G and WiFi: 3G and WiFi availability and performance have been compared in [2, 11, 5]. They investigated availability and performance by vehicle and/or walking based measurements. However, some studies only discussed availability of the WiFi network by the appearance of APs rather than actual connectivity.

The originality of our work to others is to characterize the 3G traffic offloading through WiFi on the basis of a large-scale device-based measurement and analysis of a combination of 3G and WiFi traffic.

5 Discussion

Our monitors were recruited by a web-based application and most have APs at home. This means that they are more familiar with using the Internet and smartphones than the average user. Thus, our results are likely biased towards the behavior of such advanced users, and the user and traffic offloading ratios of the current average users will be smaller than in our results. However, these results can be interpreted as corresponding to the situation in the very near future if 3G carriers successfully promote to average users the option of offloading more of their traffic volume to WiFi, considering the fact that the majority of residential users have high-speed Internet connections at home. Even in the current results, the high usage of WiFi was only by a relatively small number of users, and still 17.4% of the total volume was generated by users whose WiFi traffic volume was less than 1MB. In addition, over 50% of users turned off their WiFi interface in business hours, and most of the traffic volume of some hybrid users was downloaded via WiFi in shorter durations. Therefore, the traffic and user offloading ratios could have been higher if the promotion by 3G carriers had been more effective.

Different from the high traffic offload ratio in homes, we observed lower traffic volumes in public WiFi. We cannot currently identify the exact reason for this low availability of public WiFi, but there are several plausible reasons: (1) Most users turned off WiFi connectivity outside the home to save energy. (2) Handover of WiFi APs did not work well due to fast movement of users. (3) Outside of downtown areas, the availability of public WiFi may be not very high. (4) There is wave interference due to a large number of APs at downtown areas. Our results at least demonstrated the possibility of reason (1) being true as shown in Figure 2(b). In particular, the advanced users may proactively save the battery by turning off the WiFi interface. Also, considering the usage of WiFi and 3G networks outside homes and offices, users likely need Internet connection only for e-mail checking or simple web browsing, rather than rich bandwidth applications such as streaming. Such short and simple usage of smartphones generates a smaller amount of traffic volume. In this sense, the availability and connectivity are likely more important than bandwidth for such public WiFi.

Connecting a user's private smartphone to APs at offices is currently not common in Japan because of security policies of companies, and we also confirmed a low traffic volume of named WiFi in office hours. However, some companies have started to allow their employees to connect their private smartphone to APs at offices. In future, WiFi offloading at offices may become more common.

6 Conclusion

We reported the results of our measurement of mobile traffic offloading. We first pointed out that the variation of aggregated traffic volume via WiFi is much greater than that via 3G in our dataset. The average traffic offload ratio was 0.64 and the peak traffic offload ratio could reach 0.95 at midnight. On the other hand, the user offload ratio stayed lower, meaning that a small fraction of users contributed to a large fraction of traffic offloading. In fact, our user level data revealed that the top 30% of users downloaded over 90% of their total traffic volume via WiFi, while 10% of users only used WiFi. However, 20% of users only stuck to 3G, whose traffic volume accounted for 9.4% of the total traffic volume, and over 50% of users turned off their WiFi interface in business hours. Moreover, we observed that some hybrid users downloaded most of their

traffic via WiFi in shorter durations. In this sense, there is more room to improve the current situation of traffic offloading by promoting users to use WiFi more effectively. We also showed that WiFi offloading was mainly performed by APs in homes, and public WiFi APs are still not very commonly used in our dataset.
Acknowledgements: We would like to thank Kenjiro Cho, Romain Fontgune, and the anonymous reviewers for their helpful comments. Also, we thank the Ministry of Internal Affairs and Communications of Japan for its support.

References

1. M. Balakrishnan, I. Mohomed, and V. Ramasubramanian. Where's that phone?: Geolocating IP addresses on 3G networks. In *IMC2009*, pages 294–300, Chicago, IL, Nov 2009.
2. A. Balasubramanian, R. Mahajan, and A. Venkataramani. Augmenting mobile 3G using WiFi. In *MobiSys2010*, pages 209–222, San Francisco, CA, Jun 2010.
3. K. Cho, K. Fukuda, H. Esaki, and A. Kato. Observing slow crustal movement in residential user traffic. In *ACM CoNEXT'08*, page p.12, Madrid, Spain, Dec 2008.
4. H. Falaki, R. Mahajan, S. Kandula, D. Lymberopoulos, R. Govindan, and D. Estrin. Diversity in smartphone usage. In *MobiSys2010*, pages 179–194, San Francisco, CA, Jun 2010.
5. R. Gass and C. Diot. An experimental performance comparison of 3G and Wi-Fi. In *PAM2010*, pages 71–80, Zurich, Switzerland, Apr 2010.
6. A. Gember, A. Anand, and A. Akella. A comparative study of handheld and non-handheld traffic in campus Wi-Fi networks. In *PAM2011*, Mar 2011.
7. J. Hare, L. Hartung, and S. Banerjee. Beyond deployments and testbeds: Experiences with public usage on vehicular WiFi hotspots. In *MobiSys2012*, pages 393–405, Low Wood Bay, UK, Jun 2012.
8. T. Henderson, D. Kotz, and I. Abyzov. The changing usage of a mature campus-wide wireless network. In *MobiCom2004*, pages 187–201, Philadelphia, PA, 2004.
9. J. Huang, Q. Xu, B. Tiwana, Z. M. Mao, M. Zhang, and P. Bahl. Anatomizing application performance differences on smartphones. In *MobiSys2010*, pages 165–178, San Francisco, CA, Jun 2010.
10. K. Jang, M. Han, S. Cho, H.-K. Ryu, J. Lee, Y. Lee, and S. Moon. 3G and 3.5G wireless network performance measured from moving cars and high-speed trains. In *MICNET2009*, pages 19–24, Beijing, China, Oct 2009.
11. K. Lee, I. Rhee, J. Lee, S. Chong, and Y. Yi. Mobile data offloading: How much can WiFi deliver? In *CoNEXT2010*, page 12, Philadelphia, PA, Dec 2010.
12. G. Maier, F. Schneider, and A. Feldmann. A first look at mobile hand-held device traffic. In *PAM2010*, pages 161–170, Zurich, Switzerland, Apr 2010.
13. Ministry of Internal Affairs and Communications. Growth of mobile traffic in Japan. http://www.soumu.go.jp/johotsusintokei/field/tsuushin06.html, 2011.
14. I. Trestian, S. Ranjan, A. Kuzmanovic, and A. Nucci. Measuring serendipity: Connecting people, locations and interests in a mobile 3G network. In *IMC2009*, pages 267–279, Chicago, IL, Nov 2009.
15. Q. Xu, J. Erman, A. Gerber, Z. Mao, J. Pang, and S. Venkataraman. Identifying diverse usage of behaviors of smartphone apps. In *IMC2011*, pages 329–344, Berlin, Germany, Nov 2011.
16. Z. Zhu, G. Cao, R. Keralapura, and A. Nucci. Characterizing data services in a 3G network: Usage, mobility and access issues. In *ICC2011*, page 6, Kyoto, 2011.