Tracking the Evolution and Diversity in Network Usage of Smartphones

Kensuke Fukuda
National Institute of Informatics / Sokendai
kensuke@nii.ac.jp

Hirochika Asai
University of Tokyo
panda@hongo.wide.ad.jp

Kenichi Nagami
Intec
nagami@inetcore.com

ABSTRACT

We analyze the evolution of smartphone usage from a dataset obtained from three, 15-day-long, user-side, measurements with over 1500 recruited smartphone users in the Greater Tokyo area from 2013 to 2015. This dataset shows users across a diverse range of networks; cellular access (3G to LTE), WiFi access (2.4 to 5GHz), deployment of more public WiFi access points (APs), as they use diverse applications such as video, file synchronization, and major software updates.

Our analysis shows that smartphone users select appropriate network interfaces taking into account the deployment of emerging technologies, their bandwidth demand, and their economic constraints. Thus, users show diversity in both how *much* traffic they send, as well as on what networks they send it. We show that users are gradually but steadily adopting WiFi at home, in offices, and public spaces over these three years. The majority of light users have been shifting their traffic to WiFi. Heavy hitters acquire more bandwidth via WiFi, especially at home. The percentage of users explicitly turning off their WiFi interface during the day decreases from 50% to 40%. Our results highlight that the offloading environment has been improved during the three years, with more than 40% of WiFi users connecting to multiple WiFi APs in one day. WiFi offload at offices is still limited in our dataset due to a few accessible APs, but WiFi APs in public spaces have been an alternative to cellular access for users who request not only simple connectivity but also bandwidth-consuming applications such as video streaming and software updates.

Categories and Subject Descriptors

C.2.3 [**Computer-Communication Networks**]: Network Operations—*Network monitoring*

General Terms

Measurement

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.
IMC'15, October 28–30, 2015, Tokyo, Japan.
Copyright is held by the owner/author(s). Publication rights licensed to ACM.
ACM 978-1-4503-3848-6/15/10 ...$15.00.
DOI: http://dx.doi.org/10.1145/2815675.2815697.

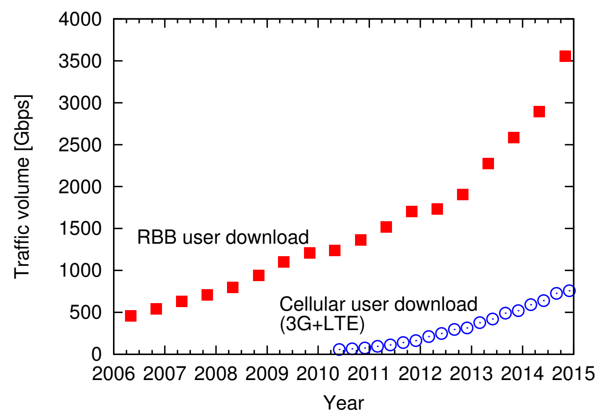


Figure 1: Growth in residential broadband and cellular traffic in Japan [34].

Keywords

Smartphone traffic; LTE; WiFi; WiFi offload

1. INTRODUCTION

In recent years the deployment of high-speed Internet at home and bandwidth-consuming applications have been driving the growth of residential broadband traffic. Following this growth, smartphone traffic has had a large impact on cellular providers. Figure 1 illustrates the growth in the total residential broadband and cellular (3G+LTE) download traffic volumes in Japan [34]. The total residential broadband traffic was estimated from the traffic volume measured at the customer edges of six Japanese commercial Internet service providers (ISPs). The total cellular traffic was measured in the backbones of four Japanese cellular providers covering most cellular users in Japan. The cellular traffic volume from smartphones accounted for 20% of the residential broadband traffic volume at the end of 2014. One reason for this trend is the rapid deployment of smartphones with rich applications. Most Japanese smartphone users are charged a flat rate with a soft bandwidth cap where user's maximum bandwidth is limited (e.g., 128kbps) during peak hours for a few days if the previous three days download volume surpasses a threshold (e.g., 1GB); thus, most light users are not concerned about saving their cellular traffic volume.

A crucial issue for cellular providers is the limited resource of channels in cellular networks. Cellular providers employ several methods to minimize quality of service/quality of experience (QoS/QoE) degradation. One commonly used technique to mitigate this issue is to apply a bandwidth cap to heavy hitters. Second, providers are deploying new capacity such as the new 4G LTE network reaching full coverage in Japanese cities. Finally, they are currently encouraging their customers to offload cellular traffic to WiFi networks (WiFi offloading). To enable WiFi offloading, cellular providers are deploying new free WiFi access points (APs) for their customers in public places (e.g., cafes, metro stations, streets, and airports). Furthermore, they have begun to give customers free home WiFi routers with the expectation that these will increase offloading at home to residential broadband. The Japanese government also plans to promote the deployment of more free APs in public spaces towards the 2020 Olympic Games in Tokyo [35]. This promotion is expected with the goal of supporting Internet access for foreign visitors. Rapid deployment of better WiFi coverage enables cellular providers to offload 3G/LTE traffic to WiFi networks in many places if they appropriately manage to lead their users to these networks.

The success of these WiFi deployment hinges on this key question: *how do smartphone users select a network from the alternatives available to them?* Prior work has not supported a straightforward answer to this question due to the difficulty in tracking all traffic flows upstream.

To answer this question, we characterize smartphone traffic behavior from measurements at the user-side. We previously gave preliminary analysis of a two-days-long Android smartphone measurement in [18]. Here we go further, presenting analysis of three measurements, each 15 days long, covering over 1500 Android and iPhone users in the Greater Tokyo area from 2013 to 2015. We characterize the evolution of smartphone usage in terms of cellular access, WiFi access, geolocation, application, device OS, and bandwidth cap.

The main findings of the paper are as follows. (1) Smartphone traffic is becoming the main player in home networks (§ 3.1, § 3.2, and § 4). We estimate that the daily offloaded traffic volume for light users accounts for 12% of residential broadband light user traffic volume. Also, the estimated total offloaded traffic volume accounts for 28% of the total broadband traffic volume. (2) WiFi traffic volume has increased more than cellular traffic volume due to the large contribution of heavy hitters. Even for light users (i.e., median), the WiFi traffic is higher than that of cellular traffic as of 2015 (§ 3.2, and § 4). (3) Users properly select network interfaces; the number of users explicitly turning off WiFi has decreased by 10% in three years (§ 3.3). Also, the number of associated WiFi networks per user has increased. As of 2015, 40% of WiFi users are associated with at least two APs in one day (§ 3.4). The traffic volume to public and office WiFi APs, however, only accounts for 2% of the total WiFi traffic volume. (4) The signal strength of the associated WiFi network is generally high; however, 12% of the public WiFi networks exhibit subpar quality (§ 3.4, and § 3.5). We further confirm the rapid deployment of 5GHz APs in public WiFi networks though dominant APs at home and office are still 2.4GHz APs. (5) Bandwidth-consuming applications (e.g., video, online storage) have become more popular (§ 3.6) in WiFi networks. We see indications that users use such bandwidth-consuming applications on public WiFi networks (§ 3.7). Also, software updates require WiFi networks by default, thus the timing of software update can be delayed for users without home APs. For security-critical updates, this delay may leave users vulnerable longer.

year	duration	#And	#iOS	#total	%LTE
2013	07 Mar - 22 Mar	948	807	1755	25%
2014	28 Feb - 22 Mar	887	789	1676	70%
2015	25 Feb - 25 Mar	835	781	1616	80%

Table 1: Overview of datasets.

Overall, we characterize the diverse network usage of smartphones under certain demands (e.g., rich applications) and constraints (e.g., bandwidth cap and network availability). Our results show *slow but clear growth of WiFi traffic offloading during these three years.* In particular, the deployment of public WiFi networks provides users both simple network connectivity and also more bandwidth for bandwidth-intensive applications such as video streaming and software updates. However, we find that WiFi offloading at offices is still limited.

2. METHODOLOGY AND DATASET

We developed measurement software for Android[1] and iPhone[2]. The software supports iPhone 4-6 series for iOS, and Android devices with OS version > 2.1. It runs in the background and records several pieces of device information such as byte and packet counts per network interface, application information, battery status, geolocation information, network information (e.g, BSSID, ESSID for WiFi), and a unique random device ID. While most information is similar across both OSes, there are some differences due to what information is available. Android OS reports application categories and their total traffic volume, but iOS has no interface to obtain the traffic volume per application. Similarly, Android OS reports non-associated WiFi AP information as well as associated WiFi APs if the WiFi interface is turned on, whereas iOS only reports the associated AP information. In concern for energy use and privacy, we report only coarse geolocation (5km precision) and do not report throughput. The software collects statistics every 10 minutes and uploads this data to a central server. If the upload fails the software caches the data and sends it later.

We conducted three measurement campaigns in March of 2013, 2014, and 2015 (Table 1)[3]. Each campaign used an independently recruited group of subjects from the greater Tokyo area, with more than 800 Android and 700 iPhone users. Each recruiting and selection process was done by a marketing research company in consideration of the market share of major Japanese cellular providers. During the campaigns, a wide variety of users were selected as listed in Table 2, and they were requested to use their smartphones as usual. The male-female ratio of the users was about 50:50. We also conducted a user survey at the end of the measurements in order to better understand user's behavior that cannot be seen in the network data (see details in § 4.2). The number of analyzed unique device IDs was about 1600-

[1]https://play.google.com/store/apps/details?id=com.inetcore.linkspeed

[2]https://itunes.apple.com/jp/app/li-yong-shi-tai-diao-zha/id597320740?l=en&mt=8

[3]Our work has been institutionally reviewed and approved as human-subjects research (NII260128, NII270129).

Occupation	Percentage		
	2013	2014	2015
government worker	2.1	3.4	2.4
office worker	20.0	20.1	23.6
engineer	16.7	14.7	16.6
worker (other)	12.8	13.7	13.2
professional	2.4	2.0	2.8
self-owned business	6.1	6.7	5.6
part timer	9.0	10.1	10.6
housewife	15.0	14.2	13.3
student	9.6	8.3	2.7
other	6.3	6.8	7.1

Table 2: User survey: user demographics.

1700 over three years. This number includes non-recruited users who installed the measurement software from respective app stores.

We intend to characterize users behaviors depending on their exchanged traffic volume. Throughout the paper, we refer to light users as those whose daily download traffic ranges from the 40th to 60th percentiles, and heavy hitters as users whose daily download traffic is ranked in the top 5%. Note that as daily user traffic volume is highly variable, one user may be a light user one day and heavy hitter on another.

To focus on typical use, we cleaned our data to avoid distortion due to two atypical events. First, we removed tethering traffic data from the datasets since such data have different traffic characteristics.

Second, we removed traffic relating to the iOS update in 2015 from our main analysis. Apple updated iOS(to iOS 8.2) during our measurement in 2015. We identified each devices update (around March 10th, 2015) and ignored all user traffic data in the day and the next day of the update for our main analysis due to its huge data size (over 500MB). We discuss traffic from this update specifically in § 3.7.

Possible measurement biases: An online market research company selected these participants from pooled users, so they are likely more advanced users than the average smartphone user. The availability of public WiFi networks in metro stations, cafes, shops, and on the streets is higher in Tokyo than in different areas. Most commuters in this area use public transportation (e.g., trains, subways, and buses) rather than personal cars. Thus, the probability to encounter public WiFi networks is likely high, and resulting WiFi traffic volume in public spaces is also high.

3. ANALYSIS OF SMARTPHONE TRAFFIC

3.1 Aggregated traffic behavior

We first examine the aggregated traffic behavior of cellular (3G and LTE) and WiFi networks in our datasets. Figure 2 indicates the weekly variations in the aggregated traffic volume in March 2015. The notations TX and RX in the figure are traffic volumes from and to smartphones, respectively. As expected, the WiFi volume exceeds the cellular volume. Thus the traditional measurements of cellular traffic such as Figure 1 *underrepresent* all smartphone activity due to WiFi offload. The ratio of WiFi traffic to the total traffic increases over time from 59% in 2013 to 67% in 2015. Also, the cellular traffic is mainly composed of LTE traffic accounting for 32% in 2013 but 80% in 2015 (see also Table 1).

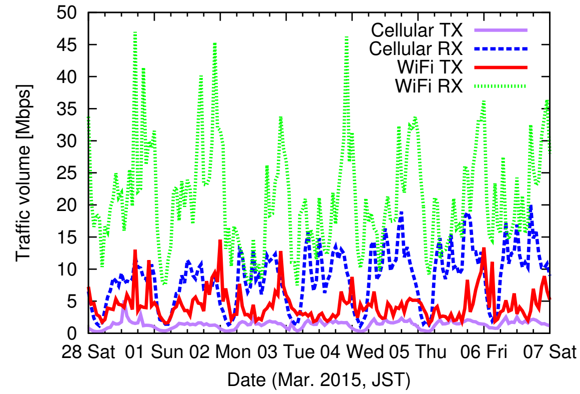


Figure 2: Aggregated traffic volume.

We see that network selection depends strongly on the time of day, with cellular traffic peaks corresponding to commute times and WiFi peaks to evening times at home. Three traffic peaks in cellular RX are affected by human activities in the morning (8am), noon (12am), and evening (7-9pm). Morning and evening peaks corresponded to the peak time of commuting mainly by public transportation in the Greater Tokyo area. In contrast, major peaks of the WiFi RX (11pm-1am) are right after the evening peak of cellular RX (9pm), though we also confirm peaks for morning and noon in WiFi RX. Cellular traffic on weekends is smaller than that on weekdays, while WiFi traffic is the opposite. Thus, these observations clearly present different temporal usage of WiFi and cellular networks.

3.2 Daily user traffic volume

We next focus on the daily traffic usage pattern per user. Our goal is to understand the impact of light users and heavy hitters by characterizing traffic distribution for each user as it changes over three years.

Daily total traffic: We first show that daily traffic increases each year, and users show a large variation in how much they send. Figure 3 shows the CDF of the daily traffic volume starting from midnight per user per day in a semi-log plot. Note that a single user appears one time for each day in the trace. We omitted users that downloaded less than 0.1MB. The shape of the curves is close to a unimodal distribution. The traffic volumes obviously increase over time, and the RX traffic volumes are about five times larger than the TX traffic volumes. While these two observations have been reported in residential broadband traffic [10], the results clearly show the client-server type user behavior.

Daily cellular and WiFi traffic: Next, we present that daily WiFi and cellular traffic volumes per user are largely distributed. Figure 4 presents the CDFs of the daily user traffic volume for cellular and WiFi network interfaces in 2015. As expected, the download and upload traffic volumes are highly skewed. On one hand, 8% of cellular interfaces and 20% of WiFi interfaces do not send and receive any data. On the other hand, the data shows that nearly all users respect traffic caps, with only 1.4% of users exceeding the 1GB per 3-day soft bandwidth cap. We examine the details of the bandwidth cap in § 3.8. When unconstrained by limits, we see a much longer tail of heavy users—the top heavy hitter downloaded 11GB in one day.

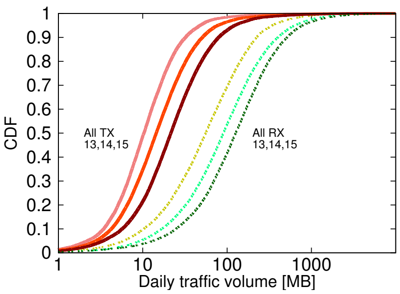


Figure 3: CDFs of daily total traffic volume per user.

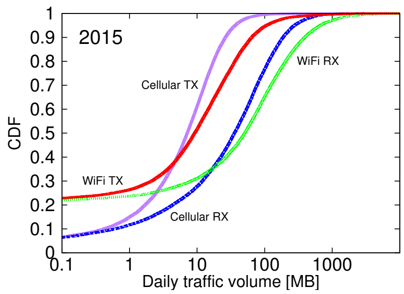


Figure 4: CDFs of daily traffic volume per type.

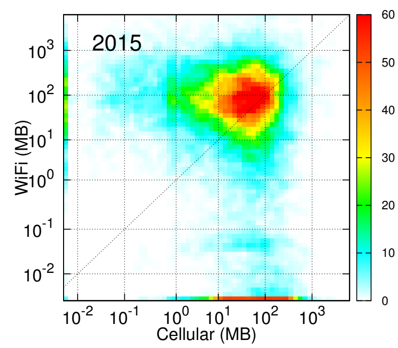


Figure 5: Daily traffic volume per user.

median	2013	2014	2015	AGR
All	57.9	90.3	126.5	48%
Cell	19.5	27.6	35.6	35%
WiFi	9.2	24.3	50.7	134%
mean	**2013**	**2014**	**2015**	**AGR**
All	102.9	179.9	239.5	53%
Cell	42.2	58.5	71.5	30%
WiFi	60.7	121.5	168.1	66%

Table 3: Daily download traffic volume per user (MB/day) and annual growth rate.

Traffic growth: We emphasize the difference of growth in median and mean daily user traffic volume; they reflect traffic behavior of light users and heavy hitters, respectively. To investigate traffic growth, we list the median and mean daily user download traffic volumes in Table 3. The median traffic indicates the typical light user's traffic volume, and the mean traffic reflects traffic from all users, but is biased by heavy hitters. The annual growth rates (AGRs) are obtained by linear fit.

We see that users download more traffic via WiFi in late years. In fact, the traffic growth of median WiFi RX is significant while the median cellular RX is higher than the median WiFi RX in 2013. Thus, light users overlook the importance of WiFi offloading in 2013, but they change their behavior depending on the increase in traffic volume. Similarly, the growth of the mean WiFi traffic volume is high. Heavy hitters are also heavier in WiFi than cellular in mean traffic volume.

3.3 WiFi vs. Cellular

To assess the impact of traffic offloading we next examine of how users select between WiFi and cellular network interfaces.

3.3.1 Aggregated view

WiFi-intensive and cellular-intensive users: We introduce three types of users (cellular-intensive, WiFi-intensive, and mixed user) from the usage of network interfaces. Figure 5 shows a heat map of traffic volumes (log-log scale), for each user and each day by network (cellular and WiFi). The color bar indicates the number of users per day for a specific traffic ratio. The diagonal line indicates reference users who downloaded equal amounts of data from both cellular and WiFi networks.

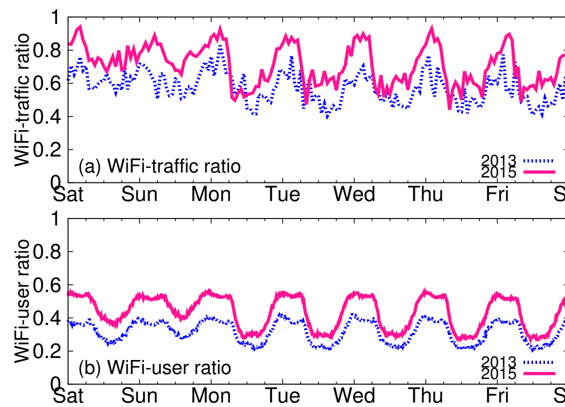


Figure 6: (a) WiFi-traffic ratio and (b) WiFi-user ratio.

We find the existence of three typical user types; *Cellular-intensive* users do not use their WiFi interfaces due to no available WiFi APs or no WiFi configuration. *WiFi-intensive* users only rely on their WiFi interfaces to avoid their cellular traffic fees. *Mixed* users select both cellular and WiFi network interfaces. The figure indicates that 22% of users are cellular-intensive and 8% are WiFi-intensive users. We see a decrease in cellular-intensive users: from 35% in 2013 to 22% in 2015, while the fraction of WiFi-intensive users is stable during the three years. Interestingly, both WiFi-intensive and cellular-intensive users show a wide spread in traffic volume; thus either can be heavy hitters.

User-level WiFi offloading: This data shows that mixed-network users often offload traffic to WiFi networks. Users above the diagonal are evidence of offloading, since their WiFi traffic exceeds their cellular traffic. The concentration of users above the diagonal shows WiFi offloading is common. Although 55% of mixed users are above the diagonal, a large number of users remain below it, suggesting WiFi usage can be further improved, especially for high cellular traffic volume users.

3.3.2 WiFi-traffic ratio and WiFi-user ratio

We introduce two metrics to highlight the benefit of using WiFi networks: The *WiFi-traffic ratio* is defined as the WiFi download traffic volume divided by the total download traffic volume in one-hour time bins. A ratio close to 1.0 in-

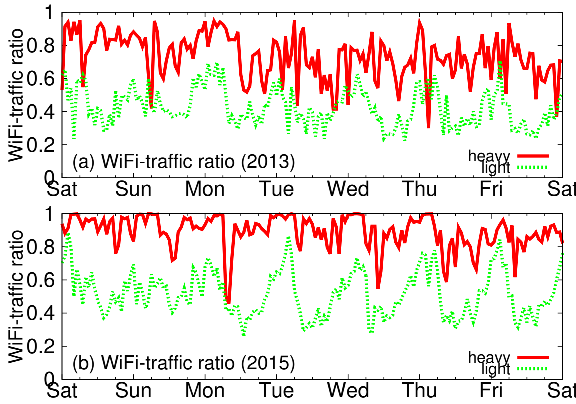


Figure 7: WiFi-traffic ratio (2013 and 2015).

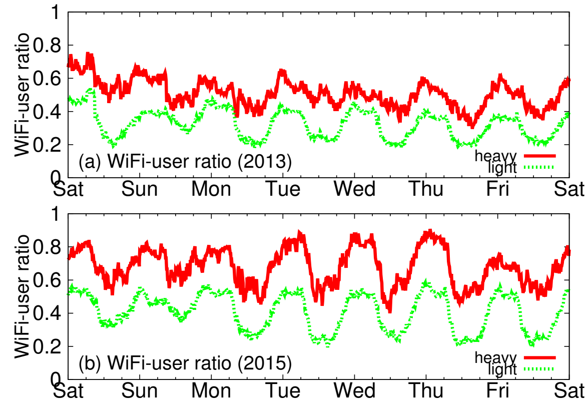


Figure 8: WiFi-user ratio (2013 and 2015).

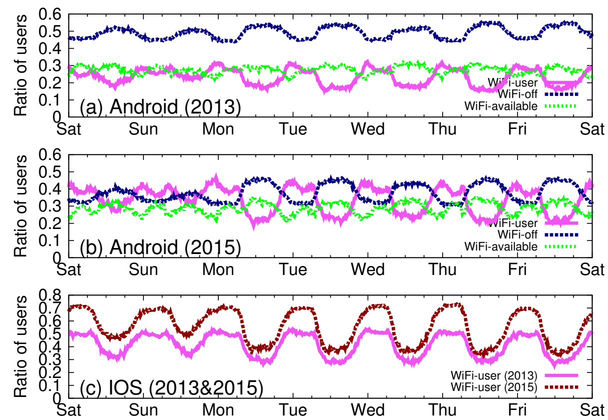


Figure 9: Ratio of users: (a) Android 2013, (b) Android 2015, and (c) iOS 2013 & 2015.

dicates that most traffic volume is downloaded via WiFi. The second metric is the *WiFi-user ratio*, defined as the ratio of the number of users associating to WiFi networks in each time bin. A ratio close to 1.0 means that, at that time, most users are using WiFi.

We see that the WiFi-traffic ratio increases showing increasing use of WiFi offloading. Figure 6(a) compares the WiFi-traffic ratio over each day of the week in 2013 with 2015. We observe that the WiFi-traffic ratio varies between 0.4 and 0.9 with clear diurnal trend. WiFi traffic is the highest from 11pm to 2am and the lowest in weekday afternoons. This temporal pattern differs from the simple WiFi traffic volume (Figure 2). Traffic penetration via WiFi is common at night though cellular traffic is still dominant in some time periods during the daytime. This data also confirms that WiFi offload is increasing, with the mean WiFi-traffic ratio growing to 0.71 in 2015 from 0.58 in 2013.

WiFi-user ratio shows that more users connect to WiFi networks in later years, however *only 50% of users connect to WiFi during peak hours.* We examine the temporal variation in the WiFi-user ratio in Figure 6(b). The number of users peaks from 9pm to 2am, while 10am to 6pm is the off-peak. The mean WiFi-user ratio also increases from 32% in 2013 to 48% in 2015. Thus, not only heavy hitters but also light users recognize the benefit of connecting to WiFi. However, only 50% of devices use WiFi even in the peak time.

3.3.3 Difference in users

We now highlight the difference in the WiFi-traffic ratio and WiFi-user ratio between light users and heavy hitters. We emphasize that in 2015, heavy hitters offload most of their traffic volume to WiFi, and that the ratio of offload traffic for light users also increases over the years. Figure 7 represents changes in WiFi-traffic ratio from 2013 to 2015. Heavy hitters already offload most of their traffic to WiFi and this trend shows no obvious diurnal pattern in 2013 (mean: 73%) and 2015 (mean: 89%). For light users, the ratio increased over the three years (from 42% to 52% in mean), particularly its diurnal pattern is much clearer in 2015, characterized by stable activity on weekends and typical daily trend on weekdays.

We see a large increase of WiFi-user ratio especially for heavy-hitters during the three years. Figure 8 illustrates the increase of the WiFi-user ratio. The difference between heavy hitters and light users is dispersed over the three years. In 2013, about 60% of heavy hitters connect to WiFi during their peak traffic time (mean: 51%). In 2015, more than 80% of heavy hitters use WiFi during their peak time (mean: 68%). These results suggest that some heavy hitters in 2013 are mainly connected to cellular networks, but this trend weakens as traffic volume increases over time.

3.3.4 Difference in device OS

We show that there are differences in offloading patterns correlated with device OS. We first define more detailed cellular and WiFi usage to quantify the potential opportunity of WiFi offloading. For Android devices, we categorize users into three groups, as shown in Figure 9(a). The first category is users connecting their device to WiFi networks (*WiFi-user*; pink curve). The second category consists of *WiFi-off users* whose WiFi interfaces are off (blue curve). The third consists of *WiFi-available users* who use cellular interfaces as the primary link but still turn on their WiFi interface (green curve). Note that the WiFi-off users explicitly turn off the WiFi interface and the WiFi-available users do not connect to any WiFi AP.

We confirm that *40% of Android devices explicitly turn off their WiFi interface, but this ratio decreases during the three years.* The ratio of the WiFi-available users is stable at around 0.25 in both 2013 and 2015, meaning that a quarter of the users could offload their traffic to WiFi if any appropriate APs are discovered. We discuss these users later in § 3.5. In comparison, the variation in the ratio of WiFi-off users is opposite of that of WiFi users; the peak time is business hours (10am-6pm). Our results show that nearly 50% of Android users *explicitly* turn off their WiFi devices during the day in 2013 (blue curve). However, the situation of the low WiFi usage improves in 2015, demonstrating that the ratio of WiFi-off users drops from 50% to 40% (Figure 9(b)).

WiFi connectivity of iOS is higher than that of Android. For iOS devices, the software does not report detailed information about the WiFi interface. However, we conclude that iOS devices connect to WiFi 30% more than do Android devices, as shown in Figure 9(c). We also confirm that there is no difference in the WiFi-user ratios among three cellular carriers providing iPhones. Thus, WiFi-user ratio differs between the two device OSes rather than cellular carriers.

3.4 Usage of WiFi networks

Next we illustrate the results of user adoption of available WiFi networks at home, in offices, and in public places.

3.4.1 Home, public, and other WiFi networks

We first define the locations of WiFi APs users connected to. We identify each WiFi AP users associate with by its (BSSID, ESSID) pair (the MAC address of AP and its network name). We then categorize WiFi networks into these three types:

Home: We identify home locations as the most common (BSSID, ESSID) pair to which each device connects during at least 70% of the time between 10pm and 6am in one day. Note that users do not always have at least one home network. In fact, the percentage of users with estimated home AP increases over time; 66% in 2013, 73% in 2014 and 79% in 2015. These numbers are roughly consistent with the result of our user survey in Table 8. Table 4 lists the numbers of detected APs.

Public: We identify public networks based on well known ESSID names (e.g., 0000docomo, 0001softbank, eduroam). These services are often deployed by Japanese cellular providers for their customers (§ 1), or by free/commercial WiFi providers (e.g., 7Spot and Metro_Free_Wi-Fi). We categorize FON APs that use a public ESSID at home as a home network because we find that some users connected to the public ESSID over 24 hours instead of using the default private ESSID by FON. The number of detected public APs doubles in three years as shown in the table.

Other: Other associated pairs are mainly located at offices or mobile WiFi APs. It also includes some open APs provided by shops and hotels. We further estimate APs at offices when they (1) mainly connect between 11am and 5pm on weekdays and (2) are not otherwise classified as home, public or mobile APs. The number of estimated office APs is stable over the years.

type	2013	2014	2015
home	1139	1223	1289
public	5041	9302	10481
other	545	673	664
(office)	166	168	166
total	6725	11198	12434

Table 4: Number of estimated APs.

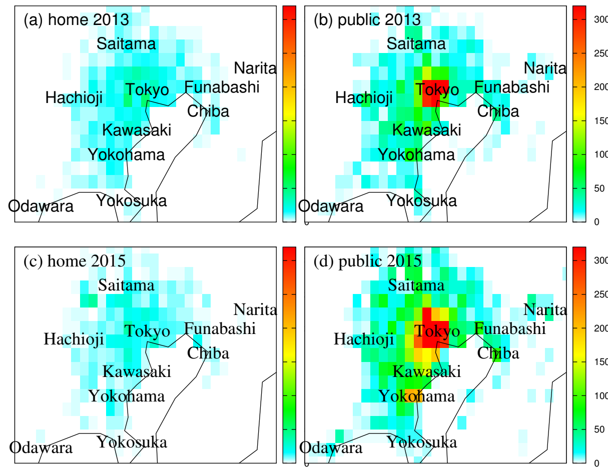


Figure 10: Number of associated unique APs per 5km cell: (a) home in 2013, (b) public in 2013, (c) home in 2015, and (d) public in 2015.

This classification shows a wide deployment of APs especially in public spaces. Figure 10 visualizes the locations and densities of associated APs in the Greater Tokyo area. Each cell is a 5km square, with color showing identified AP density. The four quadrants of the figure show home and public distribution in 2013 and 2015. Home networks are widely dispersed, reflecting many residential areas. There are only minor differences between the two datasets.

This data suggests that *coverage of public WiFi networks is broad in the Greater Tokyo.* The maps of public networks highlight strong densities in the downtown Tokyo area. The number of associated public networks is over 300 in the highest density areas (Shinjuku and Shibuya areas). We also confirm the wide coverage of public APs even far from the downtown area. By comparing the two datasets, furthermore, high-density areas are more dispersed in 2015. For example, the number of cells with at least one AP increases from 229 to 265, and that of cells with larger than 100 APs increases from 10 to 23. This result suggests that users can potentially connect to widely deployed public WiFi APs if devices are appropriately configured.

However, *the contribution of public and office APs is still small when measured by traffic volume.* Figure 11 represents the weekly traffic variations for home, public, and office in 2013 and 2015. The major contribution of WiFi traffic volume is home networks (95% of the total volume). The public and office WiFi traffic volumes are much smaller (4% of the total volume), though those volumes double during this period. The diurnal patterns of public and office WiFi are in opposition to that of home WiFi. These clear patterns in 2015 emphasize the increase of users associated to these WiFi networks.

3.4.2 Access patterns of WiFi networks: location

We first show the number of associated APs increases over time, and that daily traffic volume per user does not correlate to mobility pattern. Figure 12 illustrates the number of associated APs per user per day for all users (A), heavy hitters (H), and light users (L). We find evidence that *users connect to multiple WiFi APs in late years.* In 2013, 70%

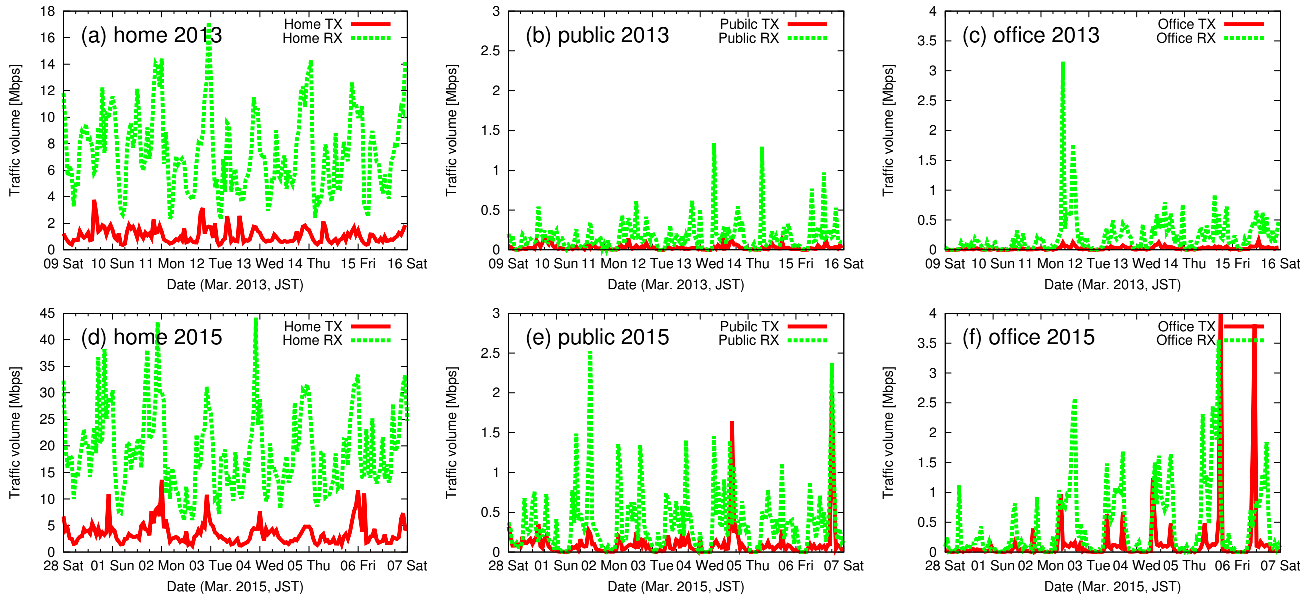


Figure 11: WiFi traffic volume: (a) home in 2013, (b) public in 2013, (c) office in 2013, (d) home in 2015, (e) public in 2015, and (f) office in 2015.

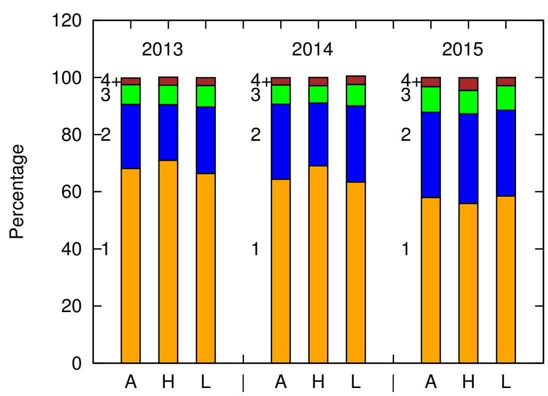


Figure 12: Percentage of the number of associated APs per day per user: All users (A), heavy hitters (H), and light users (L).

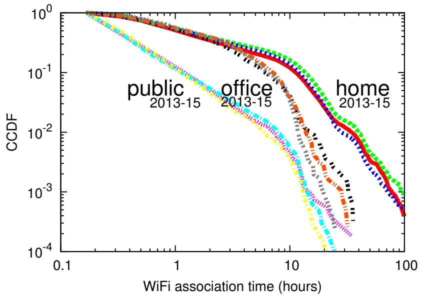
ESSIDs per day	HPO	2013	2014	2015
1	100	54.7%	52.6%	46.4%
	010	3.0%	2.4%	2.4%
	001	10.5%	9.4%	9.2%
2	110	8.2%	10.0%	9.0%
	101	10.7%	12.9%	16.5%
	011	1.4%	1.3%	1.7%
	020	0.6%	0.3%	0.3%
	002	1.5%	1.8%	2.4%
3	102	1.8%	2.0%	2.7%
	120	1.9%	1.4%	1.4%
	111	2.2%	2.3%	3.4%
	012	0.3%	0.4%	0.6%
	021	0.4%	0.2%	0.3%
	003	0.3%	0.4%	0.5%
4+	—	2.3%	2.5%	3.2%

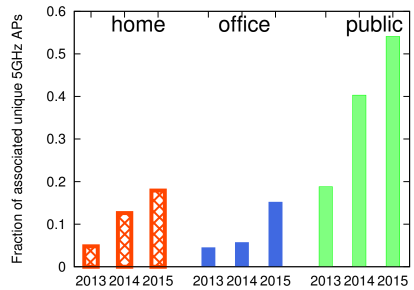
Table 5: Breakdown of number of associated APs (user·day): home (H), public (P), and other (O) networks.

of users only connect to one AP per day, but it decreases 10% in three years. Moreover, we do not see a significant difference between heavy hitters and light users. Thus, user traffic volume does not correlate to the mobility patterns.

Next we detail that the main usage patterns of WiFi networks are still only at home, and that the use of multiple APs has been more common. Table 5 lists the breakdown of the number of associated ESSIDs per device in one day for all users. The HPO column indicates the possible combinations of WiFi network types; home (H), public (P), and other (O) networks. For example, HPO = 120 represents the number of devices that connect to one home network, two public networks, and no other network in one day. For users using one AP, the percentage of using home WiFi decreases from 55% to 46%, though it is still dominant. About 10% of the users do not have any home WiFi network but connect to other WiFi networks. They have no broadband wired links at home for WiFi offloading; however, they are aware of WiFi traffic offloading. For users using two APs, the top connection pattern is one home and one other network. However, our manual inspection reveals that iOS devices connect more to public networks than to other networks whereas Android devices connect oppositely. We also confirm low percentages (2.2-3.4%) of users who connect to more than three networks in one day. The maximum number of associated ESSIDs per user per day is eight in our datasets: one home, four public, and three other networks.

We further show that the distribution of user connection durations to one WiFi network do not change over time. Figure 13 shows the complementary cumulative distribution of the consecutive time with the same AP in a log-log plot. All plots are characterized by long tails with a cutoff. Ninety percent of the users connect for less than 12 hours for home networks, 8 hours for office networks, and 1 hour for public





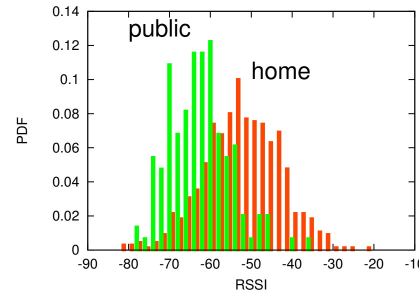


Figure 13: CCDFs of WiFi connection duration.

Figure 14: Fractions of associated unique 5GHz APs.

Figure 15: PDFs of WiFi RSSI for associated APs (2015).

networks. The results of home and office are consistent with typical user activity by definition. No clear difference is observed between the three datasets, meaning that *the usage patterns in connection duration do not change for any WiFi network type from 2013 to 2015.*

3.4.3 Access patterns of WiFi networks: 2.4 and 5GHz

Recent WiFi APs operate in two frequency bands: 2.4 and 5GHz. We find that *deployment of 5GHz APs has been rapid for public networks, but most home and office networks are at 2.4GHz.* Deployment of 2.4GHz APs is wider than 5GHz because of their earlier availability on the market. However, the 2.4GHz spectrum is less robust against noise. We investigate the deployment status of 5GHz APs at home, in offices, and in public spaces. Figure 14 shows the fraction of unique associated 5GHz APs to the total unique associated APs for the three years. The fraction of 5GHz APs varies across homes, offices, and public spaces. Still less than 20% at home and office APs are 5GHz, i.e., most WiFi APs mainly operate in 2.4GHz. For public WiFi, we see that more than half are 5GHz APs in 2015. We believe that wide use of 5GHz in public location is due to recent aggressive rollout of WiFi, while long device lifecycle means that home and office have not yet needed to upgrade.

3.4.4 Quality of WiFi networks: signal strength

There are several factors that affect the QoS/QoE of smartphones, in which the received signal strength indication (RSSI) is a common metric for quantifying the strength of the receiver-side signal in wireless communication. A larger RSSI represents a stronger signal, and an RSSI larger than -70dBm is generally better for WiFi connectivity. For example, it has been reported that the retransmission probability of a TCP session is $\approx 10\%$ for RSSI $= -70$dBm and rapidly increases for RSSI < -70dBm in a dataset measured at a conference WiFi network [41].

The signal strength at home is generally sufficient for network connectivity. We calculate the maximum RSSI of each associated AP in 2.4GHz, and plot the probability density of RSSIs of associated home and public networks in Figure 15. Home networks are characterized by a bell shape with a mean of -54dBm, consistent with those in French home networks [13, 14]. Three percent of the home networks report a weaker RSSI than -70dBm. We also confirm that the shape of the plots of office networks is similar to that of home networks.

On the other hand, *a small percentage of associated public APs indicate subpar quality.* The plot of public WiFi networks clearly shifts to smaller RSSIs around -60dBm. For narrowing the cell size in public WiFi, the mean signal strength can be weaker than that of home WiFi networks. As the result, we confirm that 12% of associated public WiFi networks indicate weaker signal strength (< -70dBm), likely causing poor QoS/QoE to the users. This low-quality connectivity is a reason to deploy more 5GHz APs in public WiFi networks.

3.4.5 Quality of WiFi networks: channel usage

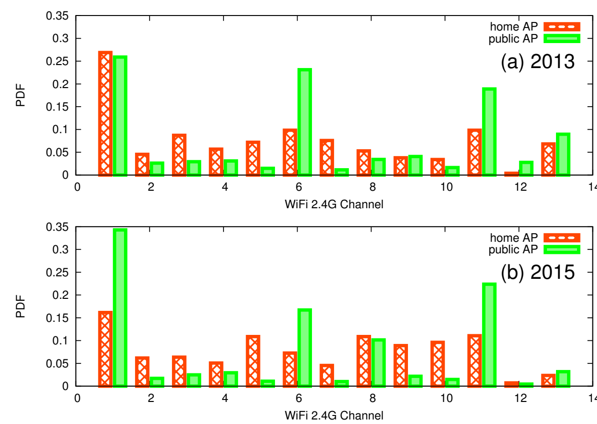


Figure 16: Associated 2.4GHz channels (2013 and 2015).

Finally, we investigate the efficacy of channel selection of APs. In addition to the RSSI, channel selection of WiFi networks is an important factor that affects the performance of the WiFi networks. In an IEEE 802.11b/g/n network, 13 channels are available in the 2.4GHz band, and two BSSIDs using neighboring channels may interfere with each other due to the overlap of channel bandwidth (i.e., cross channel interference). It is known that at least a five-channel interval is necessary to avoid cross channel interference between two channels. Some recent home-use APs have a function for detecting and avoiding channel interference. In addition, public WiFi providers design and deploy their APs to avoid cross channel interference between their APs; however, they can still interfere with other public APs.

We confirm that *deployment of public APs is generally well designed, and the concentration on the default channel has been relaxed in home APs.* Figure 16 shows the probability density of the associated channels in the 2.4GHz band for Android devices. Devices connected mainly to Ch1, Ch6, and Ch11 in public WiFi APs, indicating that devices are set up to minimize interference. Such observation is also recently reported in [38, 6]. In 2013, the numerous connections to Ch1 at home are, however, potentially causing more channel interference. This poor setting is likely due to the lack of intelligent channel selection mechanisms at the APs. However, we observe improvement in avoiding possible interference in later years; WiFi channels at home are more dispersed and have less concentration on Ch1 in 2015.

3.5 Availability of public WiFi APs

To understand how much cellular traffic *WiFi-available* users can offload to *public* WiFi networks, we next look at WiFi coverage, signal strength, and how much traffic they can offload.

We first suggest a wide coverage of potentially available public WiFi APs. We sum up all detected public networks (all) and detected public networks with a signal strength strong enough to be associated with (i.e., RSSI $\geq$ -70dBm) (strong), in a 5km cell. We confirm that 2.4GHz public WiFi networks are widely deployed in the Greater Tokyo area in our data, though the coverage of our data is imperfect due to the limited number of recruited users. In the downtown area (Shinjuku, Shibuya area), for example, the number of public WiFi networks with strong signal is over 10000 in one cell, while the number of detected networks is about 45000 in one cell. Comparing the two datasets (2013 and 2015), the number of cells with more than 100 strong 2.4GHz APs increases from 102 to 139, and that with more than 100 strong 5GHz APs is from 10 to 84. This observation indicates that deployment of 5GHz APs in public networks has largely improved.

Next, we identify that WiFi-available users (§ 3.3.4) encounter a few strong and available public WiFi APs, and that the number of such APs increases. Figure 17 represents the complementary cumulative distribution of the number of detected public APs per WiFi-available device per 10 minutes in 2015. The plot labeled as "strong" represents the number of detected networks with strong maximum signal strength, and "all" represents the number of all detected networks. Both plots represent that a few users see many APs but most see only a few. We observe that most users (90%) see fewer than 10 2.4GHz APs. As expected, the number of strong WiFi networks is much smaller than all detected WiFi networks. The data shows that 5GHz deployment is still early, since only a few users (10%) could find strong 5GHz APs, and only slightly more (30%) can find any 5GHz. However, this deployment is better than that in 2013; 10% for all APs and 3% for strong APs. Tail behavior of strong 2.4GHz and strong 5GHz APs resembles each other, suggesting WiFi APs operating in both bands.

Finally, we infer that *WiFi offloading to available public WiFi networks can reduce 15-20% of daily cellular traffic for WiFi-available users.* As described in § 1, cellular providers deploy free public WiFi service to their customers. Devices whose WiFi interfaces are enabled but not associated with any WiFi networks (i.e., WiFi-available user) can offload their cellular traffic to the public WiFi networks of their providers. Examining of such available WiFi networks for WiFi-available users, we confirm that 60% of WiFi-available users have opportunities to connect to stable public WiFi networks. We sum up cellular download traffic volume during these periods as possible offloading traffic, then conclude that 15-20% of daily cellular traffic volume for such WiFi-available users can be transferred to public WiFi networks.

	Cell home		**Cell other**		**WiFi home**		**WiFi public**	
	type	%	type	%	type	%	type	%
2013	brows.	38.0	brows.	38.5	brows.	28.0	brows.	44.1
	social	7.3	comm.	7.7	social	6.8	social	4.0
	comm.	6.2	social	7.6	comm.	4.3	life	3.3
	video	5.7	news	2.6	video	4.0	comm.	3.0
	news	2.0	video	2.1	news	3.5	news	2.9
2014	brows.	36.4	brows.	31.4	video	30.4	dload	22.5
	video	7.4	comm.	9.9	brows.	20.7	brows.	21.9
	comm.	7.4	video	8.0	comm.	6.5	video	13.8
	social	6.3	news	6.6	news	6.0	life	4.9
	news	6.2	game	6.3	dload	4.7	health	3.2
2015	brows.	28.3	brows.	28.3	video	25.4	brows.	24.0
	video	11.0	comm.	12.7	brows.	20.0	video	19.6
	comm.	9.5	video	12.0	dload	11.1	dload	9.9
	social	7.9	news	7.6	comm	7.4	life	4.1
	news	5.8	social	6.9	social	4.7	comm	3.6

Table 6: Top application categories, as ranked by RX traffic volume.

3.6 Application breakdown

WiFi offloading often requires application support, so we now examine applications, their traffic volumes, and their support for offload.

We first group popular applications into 26 categories in Google Play. Major application categories are: browser, social networking (e.g., Facebook, Twitter), video and media (e.g., Youtube, Nicovideo), communication (e.g., Line, email clients), news, gaming, music, travel, shopping, downloading, entertainment (e.g., lottery, survey), tools (e.g., printer, speed test), productivity (e.g., online file storage service), and lifestyle (e.g., restaurant info, cooking). Note that the browser category includes web use; some web use may overlap with other categories (such as videos and social networks that are accessed from the browser instead of an application).

We distinguish application usage at home or other places with WiFi and cellular network interfaces. Table 6 and Table 7 show the percentage of the traffic volume (sum of RX and TX) generated from the five largest Android application categories in each year, broken out by network type and location. We infer home locations of cellular networks with the same classification technique described in § 3.4.1.

First of all, *users are more likely to use high-bitrate applications like video when they are on free networks like WiFi.* The ratio of video traffic increases especially in WiFi networks over the years, while we see that most common application categories in all scenarios are: browser, social networking, video, and communications. Thus, watching video with a smartphone has been a casual usage, though the users understand that a rich and low-cost network environment is preferable for this purpose. Furthermore, we observed an increase in video traffic in public WiFi networks. Thus, *users demand for public WiFi networks is changing from simple connectivity-sensitive to more bandwidth-consuming applications.* Instead, most traffic of video and media category in cellular at home is due to cellular-intensive users (without APs at home).

We also confirmed *some applications force users to use WiFi networks to reduce traffic volume in cellular networks.* The percentage of the productivity category increases in up-

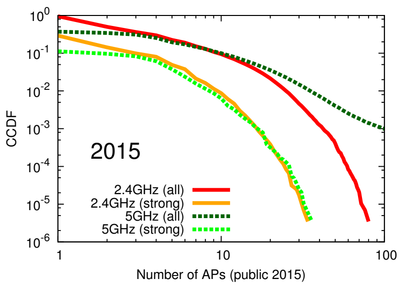


Figure 17: CCDFs of the number of detected public WiFi networks per device per 10 min.

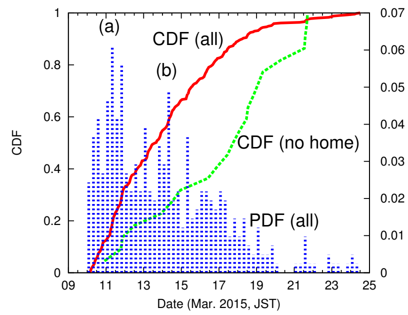


Figure 18: Software update timing.

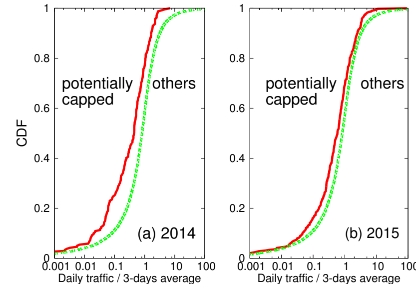


Figure 19: Effect of soft bandwidth cap.

	Cell home		Cell other		WiFi home		WiFi public	
	type	%	type	%	type	%	type	%
2013	brows.	23.0	brows.	24.5	social	24.8	brows.	33.8
	video	16.1	social	15.1	brows.	18.4	social	9.1
	social	14.2	video	10.6	video	14.7	comm.	4.2
	comm.	10.8	comm.	6.6	prod.	13.8	prod.	3.5
	prod.	2.1	prod.	3.0	comm	5.2	video	2.1
2014	brows.	31.0	brows.	28.5	prod.	39.5	brows.	41.8
	social	9.5	social	15.8	video	18.7	prod.	15.2
	prod.	9.0	comm.	11.5	brows.	15.7	life	6.9
	video	9.0	video	7.5	social	6.9	comm.	5.0
	comm.	8.5	game	3.8	comm.	5.8	news	4.6
2015	brows.	33.7	brows.	29.7	brows	19.9	video	35.6
	comm	11.2	comm	14.5	prod	15.7	brows.	25.8
	social	11.0	social	11.3	video	12.7	prod	5.2
	game	9.9	life	5.9	social	10.5	comm.	5.0
	news	3.7	game	5.1	comm.	8.0	busi	4.6

Table 7: Top application categories, as ranked by TX traffic volume.

load volume. In particular, in WiFi networks at home, this category accounts for a high percentage of traffic volume due to online storage software that uploads/downloads large files only if a WiFi interface is available. Thus, applications seem to play a major role in promoting WiFi offloading.

Finally, for light users, the contribution of video applications to download traffic becomes small. We investigate the application usage of light users in 2015 (data omitted due to space). The results are almost consistent with those for all users. A notable difference is a low contribution of video traffic; it is not ranked in the top five categories in both cellular and WiFi networks.

3.7 Software update

While § 3.6 looked at typical application use, in 2015 we captured an unusual event where many iOS devices carried out software upgrades. Apple only allows iOS upgrades on WiFi, not cellular, by default. Thus, this event provides a case study of application-forced WiFi offloading.

The impact of the update on traffic volume is high; the size of the update is 565MB which is more than ten times larger than the daily median download traffic volume (50.7MB), and it is more than half of the soft bandwidth cap limit.

Timing of software updates follows a typical flash crowd, with a large burst and a long tail. Figure 18 shows the time-of-day of updates in CDF and PDF since we first see an iOS update on March 10th, 2015 (JST). In two weeks, 58% of all iPhone devices are updated to iOS 8.2. Half users updated in the first four days, with 10% of on the first day (cf. CDF (all)). We also confirm two broad peaks; one is the first and second days (label (a) in the figure), and the other is the first weekend (label (b)). Considering the percentage of the un-updated devices, the time for migrating to the new version would have a longer tail.

We find that *availability of WiFi is important to timely deploy updates; lack of WiFi can delay OS upgrades.* In fact, only 14% of users without inferred home APs updated their device OS. Moreover, the CDF of updates without home APs demonstrates that the updates of such users are delayed. The median update delay of users with home WiFi and without is 3.5 days. For security-critical updates, this delay may leave users vulnerable longer.

Furthermore, we observed that some users without home WiFi seem to go out-of-their way to access public WiFi to obtain the update. We inspect 19 updated devices that have no inferred home APs. *Eleven users without home AP update their devices via public WiFi*, and two via office APs. This observation is unexpected but reasonable behavior that reduces update costs.

3.8 Effect of soft bandwidth cap

Japanese cellular providers commonly limit heavy hitter's maximum bandwidth especially in peak hours. A typical bandwidth cap begins after 1GB is received over the previous three days. The download speed of users over the cap will be limited (e.g., 128kbps) during peak hours for the next few days. However, two cellular providers relaxed this policy in February 2015, so we expect less impact from bandwidth caps in 2015 compared to prior years.

We detect potentially capped users from their current and past traffic volumes, and find that the ratio of such users is small. The number of potentially capped users grows each year, from 0.5% in 2013, 0.8% in 2014, to 1.4% in 2015.

We can also see the effects of bandwidth limitation after exceeding the cap. To observe this effect, we plot the ratios of the daily traffic volume per user to the mean daily traffic volumes for the previous three days in 2014 and 2015 in Figure 19, where "potentially capped" represents users whose traffic volume in the previous three days exceeds the threshold (= 1GB). The gap between potentially capped users and others indicates the effect of the soft bandwidth cap. The download traffic volume of capped users decreases due to the cap in 2014. Typically, 45% of capped users download less than half the mean of three-day traffic volume. On the other hand, only 30% of other users download the same percent-

WiFi AP	home 13	14	15	office 13	14	15	public 13	14	15
yes	70.4	72.9	78.2	31.6	25.6	28.0	44.9	47.9	53.6
no	25.9	22.8	18.6	63.2	69.0	67.4	48.9	44.9	40.9
NA	3.7	4.3	3.2	5.1	5.4	4.6	6.1	7.2	5.5

Table 8: User survey: associated WiFi APs during the measurements (%: 2013/2014/2015).

Reason	home 13	14	15	office 13	14	15	public 13	14	15
No available APs	33	34	40	46	49	52	25	24	23
Difficult to set up	32	27	21	16	15	11	31	31	25
No configuration	48	35	32	33	25	22	43	31	29
Battery drain	18	14	15	16	9	7	25	18	13
Failed	5	6	8	7	7	7	9	8	11
Security issue	NA	6	14	NA	9	14	NA	15	35
LTE is enough	NA	25	21	NA	12	10	NA	22	23
Other	6	5	5	12	10	10	9	5	4

Table 9: User survey: reasons for unavailability of WiFi APs (%). Multiple answers were allowed.

age of traffic volume. Further manual inspection clarified that 65% of capped users have no inferred home APs.

However, some users that exceed the threshold for bandwidth capping but appear to not be penalized with slower bit rates. A reason of this large traffic volume is likely download traffic in off-peak hours.

By comparing the two figures, we confirm that *the gap between the capped users and others is small in 2015, showing the change of the policy.* The figure shows that the gap is 0.15 in 2015 at the median while 0.29 in 2014. This change also supports the reason of traffic increase of bandwidth-consuming applications in cellular networks from 2014 to 2015 (Table 6).

4. IMPLICATIONS

Next, we provide implications on the contribution of smartphone traffic on residential broadband traffic, connectivity and quality of WiFi networks, and bandwidth-consuming applications.

4.1 Large impact of home WiFi offload

We start by examining the impact of offloaded traffic volume on the total volume of residential broadband traffic. As shown in Figure 1, the nationwide cellular traffic volume (3G+LTE) corresponds to 20% of the estimated nationwide residential wired broadband traffic volume [34], so a large amount of smartphone traffic offloading will suggest an increase in broadband traffic. In our measurements, the median cellular download traffic volume is 36MB/day and the median WiFi download traffic is 51MB/day in 2015 (§ 3.2). Thus, 58% of smartphone traffic is WiFi; a 1.4:1 ratio of WiFi-to-cellular traffic. This ratio is higher than that in Cicso's report (45% in 2014) [11]. As a result, we roughly estimate 28% (= 20%*1.4) of the total residential broadband traffic volume as WiFi traffic by smartphones because 95% of WiFi traffic is at home (§ 3.4.1).

We also compare smartphone traffic volume with broadband traffic volume per user. A different source of residential broadband traffic data reported that the median download traffic volume of a residential customer per day in a Japanese ISP is 436MB/day as of 2015 [9]. We again infer with this volume that traffic share of one smartphone in a home network is 12% (= 51MB / 436MB).

Thus, *the growth in offloaded WiFi traffic even for light users has had a large impact on legacy broadband ISPs, where it is hidden in the traffic volume of heavy hitters.*

4.2 Connectivity of WiFi networks

To understand the behavior of user that cannot be seen in the network data, we had all our users fill out a questionnaire. In the post questionnaire, we asked two questions about WiFi APs to users. The first question is "where did you connect to WiFi APs in three locations (home, office, and public)?" (Table 8), to compare with the results of our location-based analysis (§ 3.4.2). The second one is "why did you not connect to WiFi APs in these three locations?" (Table 9), to understand possible causes of unavailability of WiFi networks.

The answer to the first question shows that the percentage of associated APs at home and in public increased while that for office network is low and stable. The percentages of home WiFi APs are consistent with our estimation (Table 5), but, high percentages of public WiFi differ from our estimation. This gap represents the difference between user's recognition and actual connectivity; users think they have more connectivity than they really do in public WiFi networks. However, the results of the associated WiFi networks demonstrate that *the percentage of users with multiple APs has been increasing, and it is more than 40% in 2015* (§ 3.4).

From the measurement results and the answer to the second question (Table 9), we summarize the plausible reasons for the low availability of WiFi networks during daytime as follows. (1) Users report that connecting private smartphones to WiFi networks at offices (Bring your own device; BYOD) is still not common in Japan. It is supported by high percentages of the answer "there is no deployment of APs" in Table 9. In addition, a major reason of "Other" for office WiFi networks is due to the security policy that does not allow employees to connect their personal smartphone to the office WiFi networks. Our traffic data showed a significant increase in office WiFi traffic volume (Figure 11) though the number of inferred office APs (Table 4) remains stable during the three years.

(2) Configuring WiFi APs is difficult. Configuration problems decrease in public WiFi networks over time in the survey. Users report in our questionnaire that entering WiFi passwords into their mobile devices is often complicated. To reduce complexity, since 2013 WiFi APs from cellular providers use SIM-based authentication [25] without user and password information. As a result, smartphone customers can use WiFi APs from their provider with no manual actions. Coupled with the deployment of more access points (Figure 10), we believe this simplification in their use helps increase the degree of WiFi offload.

(3) Many users have no incentive to use public WiFi networks because the current quality of cellular networks is enough and are charged a flat fee per month. High percentage of the answer "communication speed in LTE is enough" in Table 9 supports this claim. Nonetheless, we estimated that 15-20% of the traffic data can be offloaded if users appropriately configure the settings for public WiFi networks (§ 3.5).

(4) Finally, we found that security with public WiFi was a significant concern, and that battery life was not.

4.3 Quality of public WiFi networks

We discuss how public WiFi networks can improve their quality. The quality and accessibility of WiFi networks have been improved over the years. However, 12% of associated public WiFi networks using 2.4GHz are characterized by a low RSSI (§ 3.4.4). Such low-quality network environments are likely a reason for users hesitating to use the public WiFi network, though over half of the associated public networks have sufficient quality. To cope with this problem, device OSes should implement tests for the quality of a WiFi network before starting to offload to prevent the degradation of user QoE.

The coordinated deployment of public APs is a difficult problem in reality because they use unlicensed bands. To mitigate performance degradation due to channel interference among different public WiFi networks, deployment of APs that support multiple providers by announcing multiple ESSIDs from a single AP should be promoted, especially in downtown areas. In fact, we confirmed this type of APs in our dataset, by checking similar BSSIDs assigned to different providers. It will be also cost-effective to use these APs for providing free WiFi APs to foreign visitors instead of deploying new APs, towards the 2020 Olympic Games. Extended usage of such APs has already been discussed in the context of providing free WiFi in disaster areas after the Tohoku earthquake.

We also observed a significant deployment of 5GHz APs in public spaces (§ 3.4.3). 5GHz APs are more robust against noise and helpful in improving the quality of WiFi network [40, 46]. This migration also contributes to increases in availability, use, and traffic in public WiFi networks. Furthermore, the limited spatial and more reliable nature of 5GHz APs may allow new types of applications to emerge, such as fine-grained targeting advertisement.

4.4 Bandwidth-consuming applications

We observed a large growth of bandwidth-consuming application traffic (§ 3.6). Users with WiFi access download more video (Table 7), suggesting that users understand that rich-bandwidth and low-cost WiFi networks are available, and changing their usage accordingly. Similarly, other bandwidth-consuming applications contribute to WiFi offloading. For example, iOS shows a pop-up message to users to promote the use of WiFi for software updates. We see a similar change in upload use when WiFi is available. For example, WiFi has been largely used for the productivity category, which includes online file storage (Table 6).

This intelligent network connectivity management by software has two different perspectives. On one hand, it will be more common and necessary to cope with bandwidth-consuming applications in the future, particularly to avoid bandwidth caps. On the other hand, *it causes another adaptive behavior of users.* Some users rely on public WiFi networks for their software updates (§ 3.7). Furthermore, we observed an increase in video traffic even in public WiFi networks. User demand is towards casual use of bandwidth-consumption applications in public WiFi networks. This trend may introduce a change of the economic cost and the QoS model (i.e., traffic engineering) of public and free WiFi networks, because low-cost connectivity is the main advantage with current WiFi networks.

5. RELATED WORK

Cellular smartphone usage: Quantifying the diverse usage of smartphones has been one of the hot topics in traffic measurement and analysis. Some studies have pointed out that the flow-level performance of the 3G/LTE cellular network is affected by many aspects such as the differences in the cellular providers, types of devices, transport protocols, and mobility [32, 28, 29, 27, 37, 19, 42, 17]. The application usage and resulting performance in cellular networks also indicate large diversity [47, 16, 28, 51, 44, 27]. In particular, the usage pattern of applications in smartphones depends on the mobility and geographical region of users [47, 51, 52]. Recent studies addressed the impact of user measurement environment on network performance [21, 43]. There are several approaches to characterize this diversity such as application identification [50, 36], app store usage [49], and geolocation mapping [3, 48]. Another direction of current research is to quantify and model energy consumption of smartphones [27, 15, 7]. Regarding bandwidth caps, a recent experiment concerning cellular service pricing shows that time-dependent pricing helps reduce peak traffic volume in cellular networks [30].

WiFi smartphone usage: Some studies focused on WiFi smartphone usage in campus networks [26, 22, 8], city-level networks [1], or networks in public transportations [24]. They argued that typical applications of such WiFi networks were HTTP, more specifically video, social networking, and software updates. In addition, the deployment of WiFi APs and their interference are discussed [2].

Throughput/delay performance of WiFi and 3G/LTE networks: Several performance metrics have been compared using vehicle-based [4, 20, 12] and speedtest-based [45] measurements. They showed that the throughput and delay performance of WiFi networks are basically better than those of cellular networks, and they are also highly variable depending on the environment, different from broadband wired access. Furthermore, in the context of traffic offloading, some studies evaluated the performance gain of WiFi offloading on the basis of WiFi APs discovered using vehicle based measurements [31, 23].

WiFi networks: Home WiFi usage has been analyzed in detail [33, 39, 13, 14]. An earlier report pointed out a low percentage of traffic volume from mobile devices at home in 2009 [33]. Also, the performance [40, 46], the traffic ratio [33], and the availability [13, 14] of home WiFi networks have been intensively studied. A large analysis of ESSIDs focused on the social relationships of users [5]. A recent large-scale survey of WiFi networks revealed application usage and channel interference in the wild [6].

The main difference between our work and others is to provide a comprehensive view of smartphone usage in cellular and WiFi networks from measurements with over 1500 users, for better understanding the adoption of smartphone users to avaialble WiFi networks. We previously explained the preliminary results of WiFi traffic offloading performance based on a two-day measurement with 450 Android users [18]. However, the measurement period was short and the available information was limited mainly to traffic volume without application and geographical information. For the current study, we analyzed three 15-day-long detailed datasets and provided more comprehensive results including the differences regarding device OS, location, and applications.

6. CONCLUDING REMARKS

We measured WiFi and 3G/LTE smartphone usage from more than 1500 Android/iOS smartphone users in the Greater Tokyo area for two weeks each in 2013, 2014, and 2015. Our results showed that smartphone users select WiFi and cellular networks based on several tradeoffs. In particular, we highlighted slow but clear adoption of available WiFi environment at home, in offices, and in public spaces. Light users have been offloading their traffic to WiFi networks, while heavy hitters totally rely on WiFi networks for their traffic offload. Moreover, we confirmed that wide deployment of WiFi APs enables users to obtain more opportunities to offload their traffic to WiFi networks. Users who have access to public WiFi send more traffic, and WiFi-availability results in heavier use of video and is required for large software updates. This adoption indicates that users recognize public WiFi networks as a part of the common network infrastructure.

Acknowledgments

We thank our shepherd Dina Papagiannaki, the IMC reviewers, Kenjiro Cho, Romain Fontugne, John Heidemann, Johan Mazel, and João Taveira Araújo for their valuable comments, and the Ministry of Internal Affairs and Communications, Japan for its support.

7. REFERENCES

[1] M. Afanasyev, T. Chen, G. M. Voelker, and A. C. Snoeren. Analysis of a mixed-use urban WiFi network: When metropolitan becomes neapolitan. In *IMC'08*, pages 85–98, Vouliagmeni, Greece, Oct 2008.

[2] A. Akella, G. Judd, S. Seshan, and P. Steenkiste. Self-management in chaotic wireless deployments. In *MobiCom'05*, pages 185–199, Cologne, Germany, Aug 2005.

[3] M. Balakrishnan, I. Mohomed, and V. Ramasubramanian. Where's that phone?: Geolocating IP addresses on 3G networks. In *IMC'09*, pages 294–300, Chicago, IL, Nov 2009.

[4] A. Balasubramanian, R. Mahajan, and A. Venkataramani. Augmenting mobile 3G using WiFi. In *MobiSys'10*, pages 209–222, San Francisco, CA, Jun 2010.

[5] M. V. Barbera, A. Epasto, A. Mei, V. C. Perta, and J. Stefa. Signals from the crowd: Uncovering social relationships through smartphone probes. In *IMC'13*, pages 265–276, Barcelona, Spain, Oct 2013.

[6] S. Biswas, J. Bicket, E. Wong, R. Musaloiu-E, A. Bhartia, and D. Aguayo. Large-scale measurements of wireless network behavior. In *SIGCOMM'15*, pages 153–165, London, UK, Aug 2015.

[7] X. Chen, N. Ding, A. Jindal, Y. Hu, M. Gupta, and R. Vannithamby. Smartphone energy drain in the wild: Analysis and implications. In *SIGMETRICS'15*, pages 151–164, Portland, OR, Jun 2015.

[8] X. Chen, R. Jin, K. Suh, B. Wang, and W. Wei. Network performance of smart mobile handhelds in a university campus WiFi network. In *IMC'12*, pages 315–328, Boston, MA, 2013.

[9] K. Cho. Broadband traffic report. In *Internet Infrastructure Review*, volume 28, pages 28–33. Internet Initiative Japan, Aug 2015.

[10] K. Cho, K. Fukuda, H. Esaki, and A. Kato. Observing slow crustal movement in residential user traffic. In *CoNEXT'08*, page p.12, Madrid, Spain, Dec 2008.

[11] Cisco. Cisco visual networking index. http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white_paper_c11-520862.html.

[12] P. Deshpande, X. Hou, and S. R. Das. Performance comparison of 3G and metro-scale WiFi for vehicular network access. In *IMC'10*, pages 301–307, Melbourne, Australia, Nov 2010.

[13] L. DiCioccio, R. Teixeira, and C. Rosenberg. Characterizing home networks with homenet profiler. Technical report, Technicolor Technical Report CR-PRL-2011-09-0001, 2011.

[14] L. DiCioccio, R. Teixeira, and C. Rosenberg. Measuring home networks with homenet profiler. In *PAM'13*, pages 176–186, Hong Kong, China, Mar 2013.

[15] N. Ding, D. Wagner, X. Chen, A. Pathak, Y. Hu, and A. Rice. Characterizing and modeling the impact of wireless signal strength on smartphone battery drain. In *SIGMETRICS'13*, pages 29–40, Pittsburgh, PA, Jun 2013.

[16] H. Falaki, R. Mahajan, S. Kandula, D. Lymberopoulos, R. Govindan, and D. Estrin. Diversity in smartphone usage. In *MobiSys'10*, pages 179–194, San Francisco, CA, Jun 2010.

[17] P. Fiadino, M. Schiavone, and P. Casas. Vivisecting whatsapp in cellular netoworks: Servers, flows, and quality of experience. In *TMA'15*, pages 49–63, Barcelona, Spain, Apr 2015.

[18] K. Fukuda and K. Nagami. A measurement of mobile traffic offloading. In *PAM'13*, pages 73–82, Hong Kong, China, Mar 2013.

[19] Z. Gao, A. Venkataramani, J. F. Kurose, and S. Heimlicher. Towards a quantitative comparison of location-independent network architectures. In *SIGCOMM'14*, pages 259–270, Chicago, IL, Aug 2014.

[20] R. Gass and C. Diot. An experimental performance comparison of 3G and Wi-Fi. In *PAM'10*, pages 71–80, Zurich, Switzerland, Apr 2010.

[21] A. Gember, A. Akella, J. Pang, A. Varshavsky, and R. Caceres. Obtaining in-context measurements of cellular network performance. In *IMC'12*, pages 287–317, Boston, MA, Nov 2013.

[22] A. Gember, A. Anand, and A. Akella. A comparative study of handheld and non-handheld traffic in campus Wi-Fi networks. In *PAM'11*, pages 173–183, Mar 2011.

[23] B. Han, P. Hui, V. A. Kumar, M. V. Marathe, G. Pei, and A. Srinivasan. Cellular traffic offloading through opportunistic communications: A case study. In *CHANTS'10*, pages 31–38, Chicago IL, Sep 2010.

[24] J. Hare, L. Hartung, and S. Banerjee. Beyond deployments and testbeds: Experiences with public usage on vehicular WiFi hotspots. In *MobiSys'12*, pages 393–405, Low Wood Bay, UK, Jun 2012.

[25] H. Haverinen and J. Salowey. EAP-SIM Authentication. RFC4186, Jan 2006.

[26] T. Henderson, D. Kotz, and I. Abyzov. The changing usage of a mature campus-wide wireless network. In *MobiCom'04*, pages 187–201, Philadelphia, PA, 2004.

[27] J. Huang, F. Qian, Y. Guo, Y. Zhou, Q. Xu, Z. M. Mao, S. Sen, and O. Spatscheck. An In-depth study of LTE: Effect of network protocol and application behavior on performance. In *SIGCOMM'13*, pages 363–374, Hong Kong, China, Aug 2013.

[28] J. Huang, Q. Xu, B. Tiwana, Z. M. Mao, M. Zhang, and P. Bahl. Anatomizing application performance differences on smartphones. In *MobiSys'10*, pages 165–178, San Francisco, CA, Jun 2010.

[29] 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 *MICNET'09*, pages 19–24, Beijing, China, Oct 2009.

[30] C. Joe-Wong, S. Ha, S. Sen, and M. Chiang. Do mobile data plans affect usage? results from a pricing trial with isp customers. In *PAM'15*, pages 96–108, New York, NY, Mar 2015.

[31] K. Lee, I. Rhee, J. Lee, S. Chong, and Y. Yi. Mobile data offloading: How much can WiFi deliver? In *CoNEXT'10*, page 12, Philadelphia, PA, Dec 2010.

[32] X. Liu, A. Sridharan, S. Machiraju, M. Seshadri, and H. Zang. Experiences in a 3G network: interplay between the wireless channel and applications. In *MobiCom'08*, pages 211–222, San Francisco, CA, Sep 2008.

[33] G. Maier, F. Schneider, and A. Feldmann. A first look at mobile hand-held device traffic. In *PAM'10*, pages 161–170, Zurich, Switzerland, Apr 2010.

[34] Ministry of Internal Affairs and Communications. Information and communications statistics database. http://www.soumu.go.jp/johotsusintokei/field/tsuushin06.html.

[35] Ministry of Internal Affiars and Communications. SAQ^2 Japan Project: Action Plan for Establishing ICT Use Environment for Foreign Tourists. http://www.soumu.go.jp/main_content/000306547.pdf, 2014.

[36] S. Miskovic, G. M. Lee, Y. Liao, and M. Baldi. AppPrint: automatic fingerprinting of mobile applications in network traffic. In *PAM'15*, pages 57–69, New York, NY, Mar 2015.

[37] A. Nikravesh, D. R. Choffnes, E. Katz-Bassett, Z. M. Mao, and M. Welsh. Mobile network performance from user devices: A longitudinal, multidimensional analysis. In *PAM'14*, pages 12–22, Mar 2014.

[38] Ofcom. Study on the use of Wi-Fi for metropolitan area applications. http://stakeholders.ofcom.org.uk/market-data-research/other/technology-research/2013/wifi-met-area/, 2013.

[39] K. Papagiannaki, M. Yarvis, and W. S. Conner. Experimental characterization of home wireless networks and design implications. In *INFOCOM'06*, page 13, Barcelona, Spain, Apr 2006.

[40] A. Patro, S. Govindan, and S. Banerjee. Observing home wireless experience through wifi aps. In *MobiCom'13*, pages 339–350, Miami, FL, Sep 2013.

[41] M. Rodrig, C. Reis, R. Mahajan, D. Wetherall, and J. Zahorjan. Measurement-based characterization of 802.11 in a hotspot setting. In *E-WIND'05*, pages 5–10, Philadelphia, PA, Aug 2005.

[42] P. Romirer-Maierhofer, M. Schiavone, and A. D'Alconzo. Device-specific traffic characterization for root cause analysis in cellular networks. In *TMA'15*, pages 49–63, Barcelona, Spain, Apr 2015.

[43] M. Z. Shafiq, L. Ji, A. X. Liu, J. Pang, S. Venkataraman, and J. Wang. A first look at cellular network performance during crowded events. In *SIGMETRICS'13*, pages 17–28, Pittsuburgh, PA, Jun 2013.

[44] M. Z. Shafiq, L. Ji, A. X. Liu, and J. Wang. Characterizing and modeling internet traffic dynamics of cellular devices. In *SIGMETRICS'11*, pages 305–316, San Jose, CA, Jun 2011.

[45] J. Sommers and P. Barford. Cell vs. WiFi: On the performance of metro area mobile connections. In *IMC'12*, pages 301–314, Boston, MA, Nov 2012.

[46] S. Sundaresan, N. Feamster, and R. Teixeira. Measuring the performance of user traffic in home wireless networks. In *PAM'15*, pages 305–317, New York, NY, Mar 2015.

[47] I. Trestian, S. Ranjan, A. Kuzmanovic, and A. Nucci. Measuring serendipity: Connecting people, locations and interests in a mobile 3G network. In *IMC'09*, pages 267–279, Chicago, IL, Nov 2009.

[48] S. Triukose, S. Ardon, A. Mahanti, and A. Seth. Geolocating IP addresses in cellular data networks. In *PAM'12*, pages 158–167, Vienna, Austria, 2012.

[49] N. Viennot, E. Garcia, and J. Nieh. A measurement study of google play. In *SIGMETRICS'14*, pages 221–233, Austin, TX, Jun 2014.

[50] X. Wei, L. Gomez, L. Neamtiu, and M. Faloutsos. Profiledroid: Multi-layer profiling of android applications. In *MobciCom'12*, page 12, Istanbul, Turkey, Aug 2012.

[51] Q. Xu, J. Erman, A. Gerber, Z. Mao, J. Pang, and S. Venkataraman. Identifying diverse usage behaviors of smartphone apps. In *IMC'11*, pages 329–344, Berlin, Germany, Nov 2011.

[52] Z. Zhu, G. Cao, R. Keralapura, and A. Nucci. Characterizing data services in a 3G network: Usage, mobility and access issues. In *ICC'11*, page 6, Kyoto, 2011.