Characterization of Wireless Multi-Device Users

AVEEK K. DAS, PARTH H. PATHAK, CHEN-NEE CHUAH and
PRASANT MOHAPATRA, University of California, Davis

The number of wireless-enabled devices owned by a user has had a huge growth over the last few years. Over one third of adults in the United States currently own three wireless devices - smartphone, laptop and tablet. This paper provides a study of the network usage behavior of today's multi-device users. Using a data collected from a large university campus, we provide a detailed multi-device user (MDU) measurement study of over 30,000 users. The major objective of this work is to study how the presence of multiple wireless devices affect the network usage behavior of users. Specifically, we characterize the usage pattern of the the different device types in terms to total and intermittent usage, how the usage of different devices overlap over time and uncarried device usage statistics. We also study user's preferences of accessing sensitive content and device-specific factors that govern the choice of WiFi encryption type. The study reveals several interesting findings about multi-device users. We see how the usage of tablet and laptops are inter-changeable, how the overall multi-device usage is additive instead of being shared among the devices. We also observe how current DHCP configurations are oblivious to multiple devices which result in inefficient utilization of available IP address space. All the findings about multi-device usage patterns has the potentiality to be utilized by different entities, like app developers, network providers, security researchers, analytics and advertisement systems, to provide more intelligent and informed services to users who have at least two among smartphones, tablets and laptops.

Categories and Subject Descriptors: C.2.3 [**Computer Communication Networks**]: Network Operations—*Network monitoring*

General Terms: Experimentation, Measurement, Security

Additional Key Words and Phrases: Characterization, Campus Network, Multi-Device Users, Network Utilization, Traffic Measurement

ACM Reference Format:
Aveek K. Das, Parth H. Pathak, Chen-Nee Chuah and Prasant Mohapatra, 2015. Characterization of Wireless Multi-Device Users. *ACM Trans. Internet Technol.* V, N, Article A (January YYYY), 19 pages.
DOI: http://dx.doi.org/10.1145/0000000.0000000

1. INTRODUCTION

In the past year, there has been an increase from 26% to 37% - a growth of 42% in one year - in the number of US adults who own a trio of smartphone, laptop and tablet [del 2014]. In addition to this, wireless enabled smart watches and other smart devices are becoming increasingly popular. Thus, we can expect the aforementioned percentages to keep rising in the future, on a consistent basis. Most wireless network measurement studies in the recent past [Falaki et al. 2010a; Huang et al. 2010] are focused on the traffic characterization of a specific user device (in most cases smartphones). Although these studies are very important, they do not include any information about how the device usage behavior changes when a user has other devices, and how all the user's devices are interdependent with respect to their usage patterns. The increase in the number of multi-device users (MDUs) raises some very essential questions such as how such users use their different wireless devices, what content is accessed on each of them, what are their security preferences and expectations, etc. In this paper, we have made an attempt at answering these questions using real network traces from users owning multiple devices in a campus network.

The network usage patterns of different wireless devices, once understood, can be used in many ways to address some crucial issues. A network service provider can use it for efficient resource allocation and planning. For example, to cope with the increasing number of online devices that results in IP address space exhaustion, delaying or revoking IP addresses based on usage pattern can be beneficial to the providers. Proper information about the device that is actively being used by the user can lead to reduction in redundant content delivery by the content providers. Complete usage pattern information of all the devices of a user can be gathered by the advertisers and online analytics providers to get a more complete view of user's online activities beyond the partial view of what is currently available through one device. Lastly, this information can be exploited by the different applications on user's devices to carry out intelligent multi-device coordination that can save energy by turning wireless radio on and off, depending on usage pattern. Although there have been recent efforts [app 2014] in this direction, most applications on today's devices are more or less oblivious to the existence of other devices of the same user.

Author's addresses: A. K. Das, P. H. Pathak and P. Mohapatra, Computer Science Department, University of California, Davis;
C-N. Chuah, Electrical and Computer Engineering Department, University of California, Davis.

DOI: http://dx.doi.org/10.1145/0000000.0000000

Even though such a characterization study for multi-device users has a lot of potential, acquiring real-world network traces for multi-device users itself is a challenge. The main reason for the difficulty is that network traces collected from the access or core networks rarely have any information about user's ownership of devices. In this work, we accomplish a characterization study of multi-device users using wireless network traffic traces collected from a large university campus. We further combine the packet traces with user-device session logs to associate traffic with users. This allows us to monitor fine-grained network usage activity for each user and all her devices. A campus network with a high number of users and wireless devices provides a network which is a good representation of multi-device user environment and make a good choice for MDU study. The characterization study described in this paper is based on data collected for nearly 1,000 access points from the university campus for approximately 30,000 users with the total network packet traces of 23 Terabytes. We classify user's wireless devices into three device types: smartphone, laptop and tablet.

The major findings about MDUs in a campus network as revealed from our work are as follows:

(1) **Device utilization of multi-device users**: When more than one wireless device is possessed by the user, rather than the usage being spread across the multiple devices, the overall network usage increases proportionally to the number of devices. Additionally, the overall time for which a particular device type is used, hardly changes, irrespective of the other devices owned by the user. This indicates that for multi-device users, in a campus network, the overall network usage is additive.
Another interesting observation shows that when users own a tablet, the percentage packets generated by laptops decrease whereas the smartphone usage remains more or less constant. Content accessed by tablets and laptops are also seen to be almost inter-changeable. We also observe that uncarried devices (devices left back at home) have a higher number of very small sessions for tablets (due to background traffic) and a fewer number of long sessions for laptops (due to user assigned tasks). We observe that most of the uncarried traffic is generated from tasks like downloads (high downlink traffic as compared to uplink traffic) and syncing in apps related to mail and social networks.

(2) **ON-OFF usage patterns of devices and efficient DHCP assignment**: Study of the intermittent "ON-OFF" device usage of a specific device type and how it is affected by other devices, show that the usage remains specific to the device type. It is unaffected by the presence of other devices of the same user. In addition, we observe that the average amount of time a handheld device (smartphone and tablets) is continually "ON" is much shorter than DHCP lease times that are assigned by the campus network operators.
We study the period of inactivity of devices after they are assigned IP addresses. This shows very similar behavior for smartphones and tables. In addition, we see that about 7900 handheld device sessions (greater than 100 seconds in length) do not create a network IP packet even after IP is assigned. The corresponding inactivity time in laptops are much smaller.

(3) **Security in multi-device users**: Websites that reveal information personal to users, generally, are accessed more frequently from smartphones and as a result, among the multiple devices of a user, protecting a smartphone against security attacks is most important. Also smartphones create more HTTPs traffic as compared to the other device types of the user. We also observe, that the sensitive content type accessed from each device remains constant and independent of other device's presence.
The selection of WiFi network type (encrypted vs. unencrypted) in the university campus is found to be more correlated to the device-type rather than specific user preferences. We observe that device-specific factors such as convenience of connection to specific network type from certain type of devices significantly affect user's choice. Specifically we see that the handheld devices connect to the encrypted network more as it is more convenient for a highly mobile user, because that network connects automatically. Also, the use of unencrypted network, which requires login every time you connect to it, is proportional to the device screen size.

Our results discussed in this paper are based entirely on the campus WiFi dataset. These results, cannot be generalized to all scenarios. However, since the campus network can be a good representation of a MDU network, the observations discussed in this paper can go a long way in representing multi-device usage patterns. In the rest of the paper, we introduce the dataset and our methodology for device detection in section II. In section III and IV we study in detail multi-device utilization characteristics and the security aspects of multi-device users. Section V includes a discussion on the major findings. After presenting the related work in section VI, we conclude the paper in section VII.

2. DATASET AND METHODOLOGY

2.1. WiFi Network Traces

We collect the network packet traces of the university campus from wireless controllers which connect to WiFi access points (APs). On the controller, the port, through which traffic is forwarded to and from the backbone

Location	Zone A	Zone B
Number of Users	7936	29925
Number of Devices	13729	48284
Number of Packets	19.9 billion	4.8 billion
Number of Access Points	337	696
Total Size	18.821 TB	4.942 TB

Table I: Dataset for characterization study

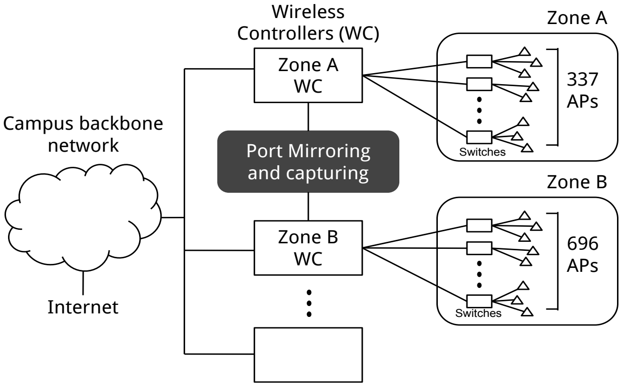


Fig. 1: Campus data capture setup

network, is mirrored to capture the data. The setup for the wireless data capture is shown in Fig. 1. We collect data from the APs of two different areas:

(1) Zone A: includes residential dormitories
(2) Zone B: includes offices, classrooms, cafeterias

The network traces are collected for 8 days for the two zones. A detailed description of the traces (user, packets, size, etc.) can be found in Table I. As seen in the table, in Zone A, the total amount of data is much higher even with significantly lower number of users as compared to Zone B. This signifies that devices at the residential dormitories are connected to the network for a longer durations, which is expected. There is an overlap of 5280 users among the user-sets at two different locations - which indicates our dataset contains network data created by 32581 unique users. All the network data collected at the controllers comprised of both upstream and downstream traffic of the user devices as we are focused on the overall traffic for each device.

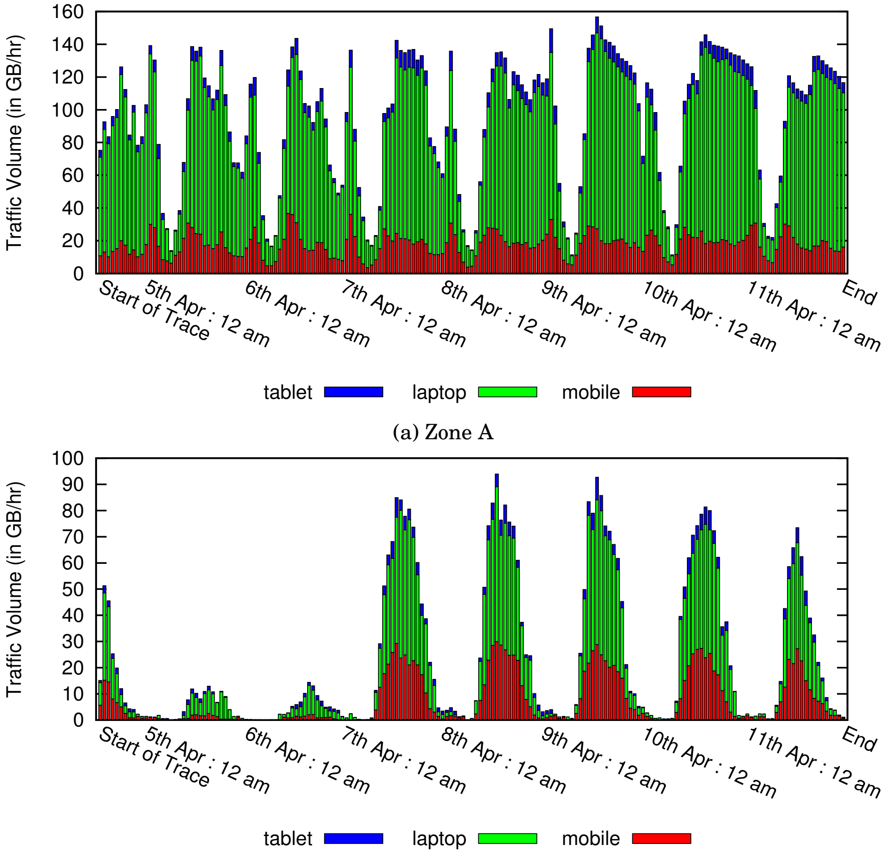
Note that in this study, we only characterize user's *wireless* devices. A multi-device user may also have a wired device such as a desktop computer but the focus of our work are devices that connect to the WiFi network. We also do not include the data created in the cellular network by smartphones as a part of the study. Presence of the wired and cellular network data provides a complete view of each user - but collection from all these data sources is not feasible. All the characterization study presented herewith are based on WiFi data of the campus network.

Fig. 2 represents the total data volume (GigaBytes per hour) variation over the entire duration of our capture from the multiple devices of a user. A comparison of Figs. 2a and 2b show that the data volume at Zone A has two peaks, one after midnight and one during noon, as compared to Zone B, which has one peak around noon. This is indicative of the location category as Zone B includes offices, classrooms, etc. and is expected to have high traffic only during office hours. For the same reason, the traffic in weekends (5th and 6th April) in Zone B is significantly low. Even though its expected that weekends will have higher overall traffic than weekdays - at Zone A - we see similar trend over the entire 8 days. This is because, each of the figures represent the trend of users' devices at that specific location, and not the overall trend of a user. We observe that laptops produce the largest volume of traffic, followed by smartphones and tablets - something that is quite intuitive.

2.2. Network Logs

Since our focus in this work is to understand the characteristics of multi-device users, we also acquire various logs to associate each packet with a user and a device. For this purpose we have two sets of logs:

1) Network Session Logs: The session logs record the association and dissociation times of each device to an AP. The log entries also contain the username, device MAC address, currently assigned IP address and the AP name to which the device is connected. These logs allow us to match each packet with to a user and her device using the IP address.



(a) Zone A

(b) Zone B

Fig. 2: Traffic volume in GB/hr at the two locations over the entire time of trace

Location	Zone A		Zone B	
No. of Devices	User Count	% Users	User Count	% Users
1	3675	46.2	14919	49.9
2	3018	38	12158	40.6
3	992	12.6	2463	8.3
4	195	2.6	313	1
$\geq$ **5**	44	0.6	72	0.2

Table II: Device count distribution

2) Network Address Translation (NAT) Logs: In certain areas, port-based NAT is used for handheld devices on campus. In such cases, we first map packet's public IP address and port to the corresponding private IP address and port using the NAT logs. After the mapping, the network session logs allow us to associate the packet with a user and a device.

Apart from the aforementioned logs, we also use the DHCP association logs. The DHCP association logs provides the *device name* for certain MAC addresses. As we show later, we use this information for detection of device type (smartphone, tablet or laptop).

With the use of the network session logs and the NAT logs, we do a packet-by-packet matching to associate each packet in the network packet traces described in Table I with a unique MAC address and a unique user.

2.3. Data Anonymization

The data collection for the characterization study was performed by the information technology department of a university campus. The department collects network packet traces, logs, etc. of the wireless network for the purpose of monitoring and management. We worked with the department to anonymize collected packet traces and the network logs to remove any information that is specific to an individual before using it for our study. Specifically, we anonymize the IP addresses, the MAC addresses, the usernames, device-names and names of the access points. We employ prefix-preserving anonymization as proposed in [Fan et al. 2004]. The

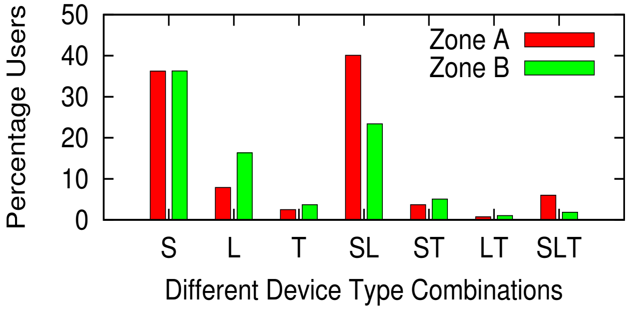


Fig. 3: Device Type Distributions at the two locations

Device Type Combinations	Zone A	Zone B
S	2876	10861
L	626	4890
T	197	1106
SL	3182	6995
ST	293	1516
LT	56	307
SLT	463	556

Table III: Different Device Types

anonymization methods and parameters are kept consistent over all traces and logs in order for us to match packets, users and devices.

2.4. Device Count of Users

After associating each device to a specific user we calculate the number of devices a user owns. The device count variation of users at both locations is represented in Table II. We observe that about 50% of all users have more than one devices, which shows that there is a valid case for multi-device user study in a campus network. However, due the presence of visitors at Zone A and due to transient mobility patterns of users in Zone B, many users show up in our dataset with just one device, increasing the percentage of users with one device type.

In order to remove the effect of visitor devices, we consider a device to be owned by a user if we observe that the device is creating WiFi traffic for multiple days (at least two) using the login credentials of that user. However, if another user, uses the same device, with a different login for a short duration of time, we remove that session's data from our dataset. In cases, when a user borrows a device belonging to someone else, but does not login to the wireless network with his own credentials, we are not able to detect that.

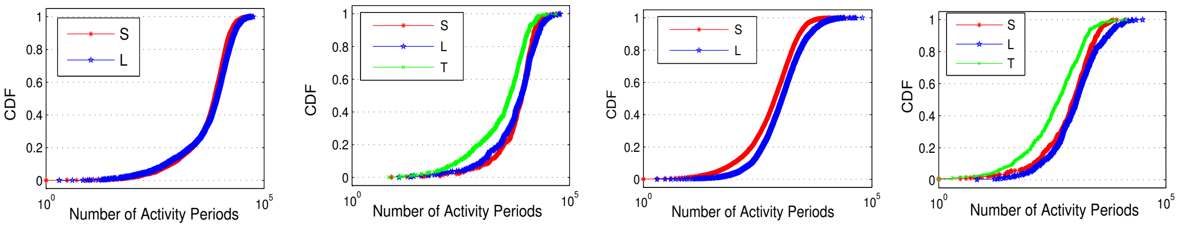
2.5. Device Type Detection

One of the most important steps in our study is the detection of the device type. We limit our observations to three device types - smartphone, laptop and tablet. In addition to these three device types, we observe gaming consoles in our dataset. To accomplish this we combine two different approaches:

2.5.1. DHCP Device Name Mining. The DHCP request message from the device to the server contains the device's hostname. In most of the current platforms such as Windows and Mac OS, the hostname is the device name given by the operating system - e.g.: John-PC. As a result, the DHCP log file mentioned in section 2.2 includes the device-name for some of the MAC addresses. Device names like “John-PC”, “Andy's MacBookPro” or “Trudy-iPhone” have keywords, the presence of which mean that the device is a laptop (in the first two cases) or a smartphone (in the last case). We do a keyword-based search on the DHCP host-names which predicts the device type of the MAC address. Some example keywords are shown in Table IV.

2.5.2. User Agent Parsing and Mining. The user-agent field present in the HTTP GET Request header contains useful information about the device type. We use a combination of the information available (for e.g.: CPU architecture, OS name, browser name, model name, etc.) [Maier et al. 2010] along with the user agent string for device type detection based on keyword-search. A set of keywords are shown in Table IV.

Either of two approaches of device detection, by themselves, is not enough to detect the device type. For certain devices, the user-agent field has no useful device related information, whereas for some users, the DHCP



(a) Zone A - user set SL (b) Zone A - user set SLT (c) Zone B - user set SL (d) Zone B - user set SLT

Fig. 4: Activity period distributions for user sets SL and SLT at zones A and B - the distributions remain the same for specific a device type of a user, irrespective of other user devices

host-name is blank or useless for our purpose. For example, in Android devices, the hostname is hashed for protection of user privacy and has no keywords which can contribute towards device detection. Overall, 57.4% of device type information are detected using the user-agent fields and the rest of 42.6% uses the DHCP host-name information. In certain cases, where there are a low number of user-agent fields and there is no DHCP device name available, the device type remains unclassified. As a result, the device type distribution is slightly different in behavior than the device count distribution we showed above. The percentage of unclassified devices are 3.06% in Zone A and 12.22% in Zone B. Many of the observations in our paper can be further verified based on the difference in the screen size of the devices owned by users. However, since we do not have any screen size information (apart from the few instances when device's brand name is present, like iPhone 6 or Nexus 5) our analysis is based on the three aforementioned device types.

The device type distribution and the number of users in each device type combination are represented in Fig. 3 and Table III. For our analysis, we divide the entire user set into 7 distinct groups: S, L, T, SL, ST, LT and SLT, where a user in set SLT owns smartphone, laptop and tablet. A user set is determined based on the number and type of devices a user owns. The highest number of occurrences of multiple devices is for users with a smartphone and a laptop. In a residential setting (Zone A), the number of users with all three device types (SLT) is higher as compared to Zone B as not all users carry out all their devices. The number of multi-device users with no mobile phones is almost negligible, which is expected as, in present scenario, almost every individual uses and carries around a smartphone. We use the same notation as seen in Fig. 3 to represent the different user sets. In addition, "S(SL)" is a representation of smartphone behavior in the user set having smartphones and laptops, and so on.

3. MULTI-DEVICE UTILIZATION

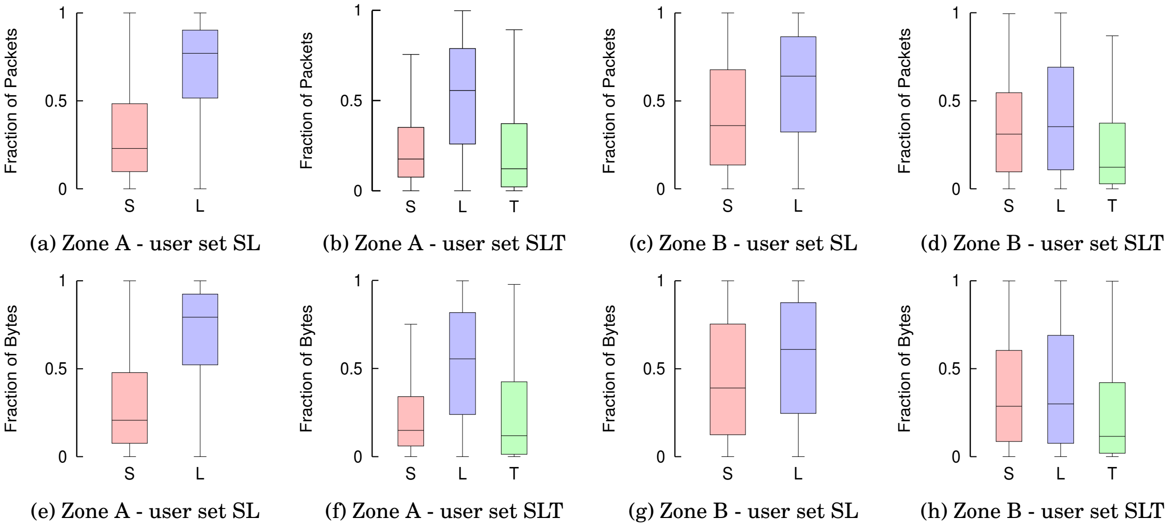
The first question that we address in our multi-device user study is how do the users use their different devices to access the network. We answer the question using two levels of characterization. First, we provide a high-level aggregate characteristics of device usage in terms of time, packets and bytes. We then look at more fine-grained intermittent usage activity (such as ON-OFF usage) in Section 3.2. Note that for all our analysis we consider the network usage as an indication of device usage, as it is known that the maximum network traffic volume is created when a device screen is in [Huang et al. 2012]. We also consider all the packets created by the devices (including TCP control packets, etc.). In this work, utilization is based on the WiFi network usage and we do not consider wired or cellular network usage. The results we have discussed in this section are representative of campus networks.

3.1. Time and Packet Characteristics

1) Activity Period per Device Type: One of the primary indicators of device-usage is the amount of time for which the device generates network packets. A specific network session is not continuous network usage - it is a combination of many activity periods. We define one activity period as a 10-second time interval during which at least one packet was created by the device. As seen in [Das et al. 2014], activity period determined

Detection Method	DHCP Device Name	User Agent Parsing
Mobile Keywords	iPhone, Nokia HTC	Windows Phone, Dalvik Blackberry, Nexus 5
Laptop Keywords	Macintosh, PC Dell, Vaio	amd64, Fedora Ultrabook, Chrome OS
Tablet Keywords	iPad	iPad, Nexus 7, Surface

Table IV: Keywords for device type detection



(a) Zone A - user set SL (b) Zone A - user set SLT (c) Zone B - user set SL (d) Zone B - user set SLT

(e) Zone A - user set SL (f) Zone A - user set SLT (g) Zone B - user set SL (h) Zone B - user set SLT

Fig. 5: Division of traffic volume among different owned devices (a-d: packets, e-h: bytes). We see that inclusion of tablets result in significant decrease in laptop.

using a 10 second window is a significant representative property of wireless network traffic. To understand how the total time usage of various device types varies in presence of other devices, we calculate the number of activity periods created by each device of a user. Fig. 4 shows the Cumulative Density Function (CDF) of the activity periods of devices for users with smartphones and laptop (SL) and users with all three device types (SLT).

• From the CDF representations of activity periods of smartphones at Zone A in different user sets, we see the distributions for smartphones in Fig. 4a (S(SL)) and Fig. 4b (S(SLT)) are identical. A similar trend is seen for all other device types at both Zone A and B.

• As the distributions for different devices remain same for different user sets, we can combine the trends observed at both locations for all device types and claim that in a campus scenario, when a user has more than one device, the overall time the wireless network is accessed, increases rather than the total time getting divided between devices. As a direct result, the amount of time a specific device is used is independent of the presence of other devices. Overlapping activity of different devices, e.g. a user is using her laptop but her phone is also exchanging some traffic, is discussed in details in point 4 later in this section.

2) Percentage Traffic generated per Device Type: Similar to time, the traffic generated by a device is a definitive indicator of the amount of usage of that device. Calculation of the number of packets and bytes created by each device of a user shows how the overall generated traffic by a user is divided between her devices. The distributions of the fraction of packets generated by each device type for different user sets (SL and SLT) are shown in Fig. 5(a-d) in the form of a box-plot. Similarly, Fig. 5(e-h) show the distribution of the fraction of bytes created by each device type. In the plot, each bar represents the distribution that is specific to a particular device type in a unique user set.

• All the representations in Fig. 5 show that laptops create significantly higher traffic compared to other device types. This follows intuitively (also mentioned in [cis 2015]) from the fact that data-extensive websites (like videos, file downloads etc.) are mostly accessed from laptops.

• At the residential dormitories of Zone A (Figs. 5c and 5d) the difference in traffic generated between laptops and handheld devices are much more prominent, as compared to Zone B (which includes classrooms, offices and cafeterias). This is another intuitive location based characteristic that is observed, as handheld devices are used more in a non-residential setting.

• Due to the common set of apps in smartphones and tablets (e.g. Android or iOS apps), it is expected that the presence of a tablet will reduce the percentage of traffic created by the smartphone, as similar content is expected to be accessed in both. However, a look at Figs. 5b and 5d and their comparison with Figs. 5a and 5c will reveal that the inclusion of a tablet device results in a significant drop in the percentage of packets created by laptops but does not, substantially decrease the packets created by smartphones. A similar trend is observed when we look at the bytes created by the different device types in each user set.

3) Content Access from Different Device Types: Time duration, packet count and byte count are good indicators of the amount of usage of the network by a user. But these statistics have no information about the content that is accessed at each location. In this section, we study the different website categories that are accessed by each device type of a user. For this, we employ a keyword based search for different application categories on the information available in packet headers. We specifically look at the full request URI, available in HTTP GET requests, and the DNS queries.

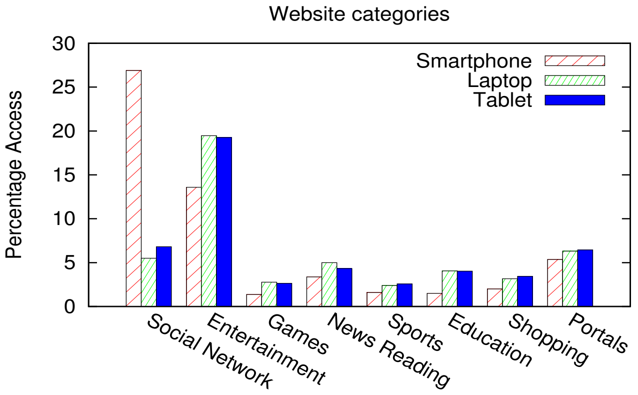
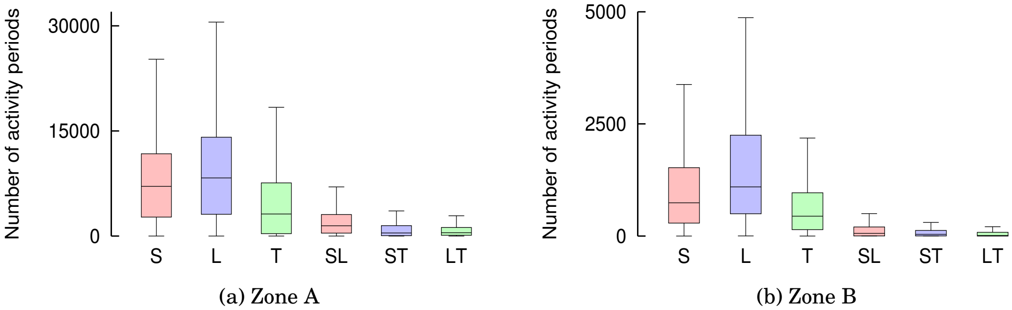


Fig. 6: Access of website categories in different device types: usage in laptops and tablets are almost the same and much different from smartphones.



(a) Zone A

(b) Zone B

Fig. 7: Overlap of activity periods: overall very low overlap. Maximum overlap in smartphones and laptops

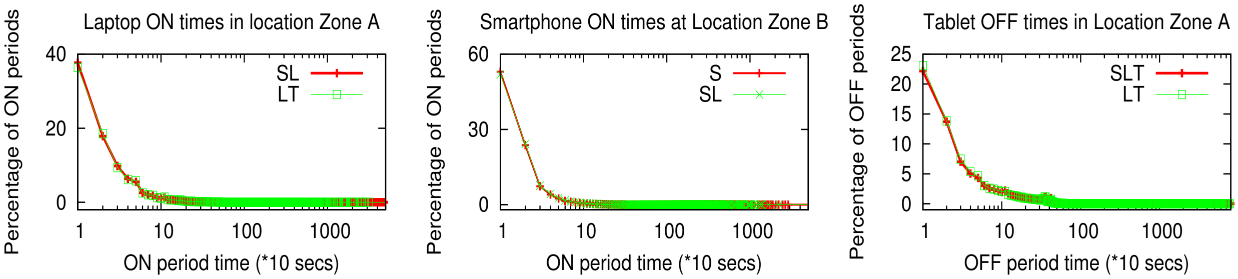
Table V gives an example of the keywords used for the search for the different categories. Fig. 6 shows the different categories of websites as accessed by the three device types among all the users in our dataset. The results are shown as a percentage of access of a particular category among all the websites accessed.

• Fig. 6 shows us that the access of different website categories in a campus environment is very similar between laptops and tablets. The websites accessed via smartphones on the other hand show a different pattern. Because of the presence of similar apps (Android and iOS apps) in smartphones and tablets, we expect that the website access in these devices would have a similar pattern. But our observation proves that in a campus scenario, laptop and tablet usage patterns for a set of overlapping applications, like web browsing, are more or less similar.

• Based on the study of percentage of packets for each device type in Section 3.1, we observe presence of tablets result in the reduction in the usage of laptops. As we see in Fig. 5, the percentage volume of traffic in laptops is significantly affected. This happens because, a set of applications, like Web traffic, are common between the two device types (as seen in Fig. 6) and as a result are offset to tablets from laptops. However this does not signify that the entire traffic between these two device types are interchangeable. There are some applications, like file sharing or live video streaming, that are specific to laptops.

Interest Category	**Keywords**
Social Networks	facebook, twitter, friends, social, plus.google
Entertainment	youtube, netflix, itunes, mp3, video, music
Games	zynga, xbox, games, puzzles, trivia, aws
News and Reading	nytimes, bbc, cnn, blogspot, news, magazine
Sports	espn, mlb, soccer, olympics, fifa, ncaa, nba
Education and Career	.edu, stackoverflow, github, courseera, school
Shopping	craigslist, amazon, ebay, target.com, groupon
Portals	yahoo, google, bing, msn

Table V: Keywords for website detection



(a) (b) (c)

Fig. 8: “ON-OFF” period distributions: Intermittent network usage of the different device types is also independent of other user devices present

4) Device Usage Overlap: Does the presence of more than one device mean that a user accesses the Internet with all her devices at the same time? In this section, we address that question by calculating the total amount of time there is an overlapped usage of two user devices. We calculate the number of activity periods when both user devices were simultaneously active. Fig. 7 shows activity periods of each device types and the overlap times between two pairs of devices. The simultaneous representation helps to compare the overlap times with the actual usage times.

• The overall overlap amount is very low (maximum being 1/4th of the entire time of device usage) as compared to the use of each device type. Comparing between the two locations, we observe more overlap in a residential setting as compared to Zone B. In Zone B, users in many cases, are in motion, and hence instances of two devices being simultaneously used is low.

• The maximum overlap of usage occurs for laptops and mobile phones. This, in a way is intuitive, and shows that a user has a normal tendency to use smartphones even when a laptop device is in use.

• The maximum value of activity period is much higher in Zone A, which follows directly from the fact that usage of devices happen for longer periods in a residential setting.

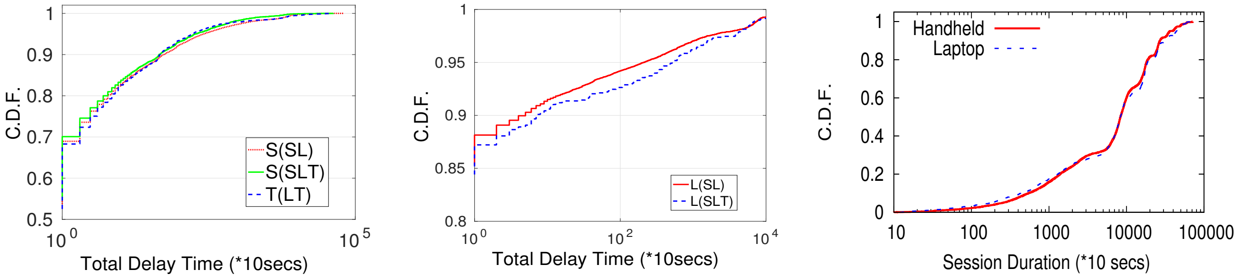
Findings: *(i) From the time and overall traffic characteristics of multi-device users in a campus environment, we observe that the presence of additional user devices does not alter the duration of usage of a specific device. (ii) Another important observation is that the usage of a tablet causes a decrease in the percentage traffic in laptops, whereas the percentage traffic in smartphones remains unaltered. Based on observed traffic volume and content accessed in our campus dataset, we can say that for most web browsing applications, the content usage of laptop and tablet are interchangeable whereas the mobile usage remains unaffected. (iii) The overlapped usage among the multiple devices is very low, with maximum overlap for smartphones and laptops*

3.2. Intermittent Network Usage Characteristics of Devices

1) ON-OFF Network Usage Pattern: We have studied the total amount of time a device was being accessed by users and observed that the behavior is independent of the presence of other devices, in most cases. However, the total usage time does not reveal any information about how a device is used, intermittently. As mentioned before, in our study, we consider the network usage as an indicator of device usage. In most cases a device is not used continuously, but follows an alternating on and off usage behavior. In this paper, we refer to this behavior as the “ON-OFF” device usage pattern. During a WiFi connection, if a packet is created in a 10 second time interval, we call the device “active” in that period. Continuous periods of activity constitute an “ON” period. When the device has a 10-second inactivity period, the “ON” time is over and the “OFF” time starts. Continuous periods of inactivity results in an “OFF” period. The “ON” period starts again when an activity period is observed.

We study how the presence of other devices have an effect on this intermittent user behavior, by calculating the ON-OFF times. Fig. 8 shows the probability mass functions (PMFs) of the “ON” times for laptops in user sets SL and LT, the “ON” times for smartphones in user sets S and SL and the “OFF” times for tablets in user sets SLT and LT.

• The results in the Fig. 8 show that the ON-OFF usage of a device is not affected by the presence of other user devices. The PMF of laptop (and smartphone) ON times is almost identical across both user sets. Similarly, the PMF for the OFF times for tablets is almost identical across different user sets. This substantiates the claim that once a device is connected to the campus WiFi network and is in use by a user - the other devices owned by the same user - does not have an effect on the usage of that device. This is in a way, counter-intuitive, as we would expect the presence of a smartphone affecting the use of laptop (or vice versa), but the observations tell otherwise.



(a) Handheld devices

(b) Laptop devices

(c) C.D.F. of sessions where no IP packets are generated

Fig. 9: a,b: Delay between IP assignment and usage - hand-held devices have a higher delay in creating the first packet after connection as compared to laptops. c: Distribution of sessions when no IP packets are created and the IP assignment is totally wasted.

• However, the ON usage patterns of a smartphone is different from that of the laptop. Such a pattern is seen for all the different device types. This is indicative of the fact that each device type has its own independent way of usage.

Based on the values of "OFF" times, the average inactivity time was calculated to be 100, 170 and 50 seconds for smartphones, laptops and tablets, respectively. Using these "OFF" period values, we recalculate the "ON" period distributions. In this case, we call a device inactive only if the continuous inactivity duration is greater than the average "OFF" duration for that device type. The recalculated "ON" have an average of 6 minutes for smartphones, 15 minutes for laptops and 2 minutes for tablets. The standard DHCP lease time for the campus network is 900 seconds (15 minutes), which is lower than half of the average ON times for handheld devices, as calculated above. A shorter DHCP lease duration during IP assignment, for smartphones and tablets, can help in better utilization of the campus IP address space [Papapanagiotou et al. 2012].

2) Delayed IP Address Assignment: Whenever a device revisits a previously known WiFi network, the device is automatically connected to the wireless network and is assigned an IP address. This IP address is assigned even if a user is not actively using the device. In this section, we study the amount of delay that exists between the time a user is assigned an IP address and the first packet created by the user. We observe that there is a distinct similarity in behavior in this respect between all handheld devices (smartphones and tablets). However, handheld devices behave much differently from the laptop devices. Figs. 9a and 9b shows the CDFs of the total delay times, in the case of handheld devices and laptops for a few representative user sets.

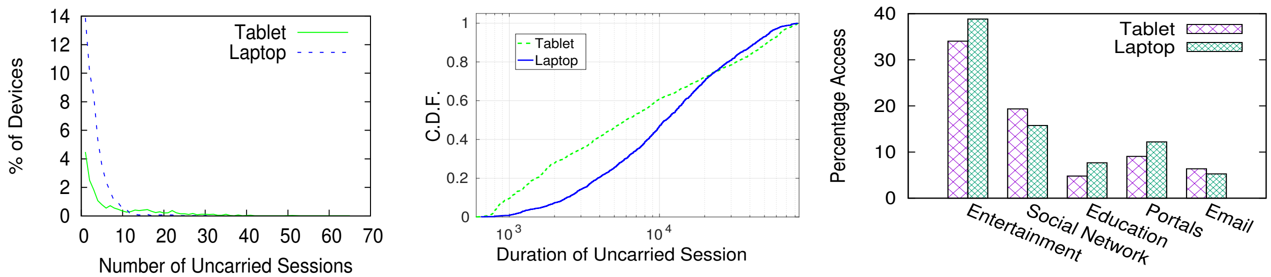
• We observe that nearly 17K (5.6%) sessions in handheld devices have a delay of at least one minute between IP assignment and the creation of the first packet. Overall, in 30% of the handheld device sessions, the device has a delay of at least 10 seconds between the assignment of IP address and the first packet created. However, for laptops this corresponding number is as low as 12%.

• We observe that approximately 9800 (3.2%) sessions assign an IP address to the device, but there are no network packets generated. Among these sessions, majority sessions, specifically 7917, are for handheld devices as compared to 1887 sessions for laptops. All these sessions are greater than 100 seconds in duration. Fig. 9c shows the CDF of the duration of such sessions for both handheld devices and laptops, which reveals that there are long intervals of time when IP addresses are assigned without any network activity.

• The number of sessions with significant delay and, at times, no network activity in handheld devices is approximately 25k. Although this might seem low (8.2%), our dataset is 8 days long and spans different location categories - offices, cafeterias, residences and classrooms. In a scenario, when transient users (or visitors) are very common, e.g.:restaurant, cafeteria, or bus station, this study gives rise to the possible case of not assigning the IP address immediately to handheld devices on entering the vicinity of a WiFi access point of such a location. Ultimately, when the user actually uses the device, a new DHCP request is sent to the server and the IP address is consequently assigned, thus avoiding auto-connection for handheld devices and in turn this can lead to better IP space utilization.

• From the point of view of the end user, efficient DHCP assignment schedules can be implemented to ensure that page load times for user applications are not affected, when delayed DHCP is used. A number of research works [Zhao et al. 2013; Liu et al. 2014] have looked into scheduling of WiFi low power mode or WiFi sleep in an efficient way so that page load times are not compromised and at times, can be even improved. We also believe that, a delayed DHCP assignment system should not increase the load on the DHCP server as the number of IP address requests are the same as when there is no delay.

Findings: *(i) The "ON-OFF" device usage pattern of a specific device type, in a campus network, remains unchanged irrespective of her other devices.(ii) We find that the average duration of continuous activity of*



(a) Number of uncarried sessions by each device type

(b) Duration of uncarried sessions in each device type

(c) Categories of website accessed during uncarried sessions

Fig. 10: Uncarried sessions: laptops create longer sessions, whereas tablets create shorter but more number of sessions. Most of the uncarried sessions are caused by media downloads or by syncing of social networking apps, emails, etc.

Device Type	**Total Uncarried Traffic (GB)**	**Uncarried as a % of overall traffic**	**Uplink Traffic (GB)**	**Downlink Traffic (GB)**	**Downlink:Uplink Traffic**
Tablet	439	9.7	19.4	419.6	21.63
Laptop	1298.5	10.6	77.75	1220.75	15.7

Table VI: Uncarried Session Sizes: Uplink and Downlink Traffic division. High downlink traffic show that downloads form a major part of uncarried session traffic.

handheld devices is much smaller compared to the usual DHCP lease times assigned in the university, which indicates that shorter DHCP lease times can be used for handheld devices.(iii) It is also observed that handheld devices have noticeable difference between the times of IP assignment and creation of first packet (due to auto-connection to WiFi networks). In certain locations, where most users have a transient behavior or are visitors, this phenomenon can be corrected to result in better IP space utilization.

3.3. Uncarried User Devices

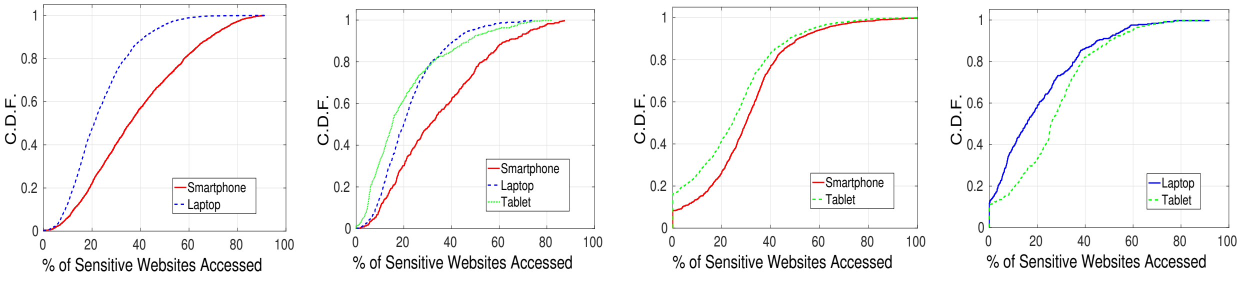
In the previous two sub-sections, we looked at how the different device types are used from the point of view of time, packets, content and intermittent usage. We have also studied the overlapped usage of each device type. In this section, we look at the usage of a specific device type at a residential environment when the user is located elsewhere. Here, we introduce the concept of “uncarried devices”. A user with multiple devices does not carry around all her devices - one or more devices are left back home (or in our case, residential dormitories). These devices, that are left back, are termed “uncarried devices”.

In most cases, a smartphone is carried around by a user and the other device types are uncarried. In our dataset, we check for the creation of a session by a tablet or laptop at the residential location (Zone A) at the same time as a smartphone initiates a network session in the campus at any location apart from Zone A (the residential dormitories or home). When such an instance occurs, we consider the laptop or the tablet to be “uncarried”. Overall, we find 16,963 such uncarried device sessions which in total generate about 1700 GB of uncarried data. Table VI lists the total volume of uncarried traffic that is observed in our dataset.

• The traffic created in the uncarried sessions for tablets are 9.7% of the total traffic (carried+uncarried) generated by those devices, whereas the corresponding number for laptops is 10.6%. This shows that uncarried sessions are non-negligible. In these sessions, the downlink traffic is 21 times the uplink traffic in tablets, and 15 times in laptops. This indicates that bulk of the uncarried traffic is caused without user initiation due to background syncing or ongoing downloads.

• Fig. 10a shows the percentage of devices of each device type that create a specific number of uncarried sessions. It reveals that a large number of the laptops present in our dataset have at least one uncarried sessions. On the other hand, the number of tablets having at least one uncarried sessions is much lower in number. However, the count of uncarried sessions is higher for tablets as compared to laptops. This indicates that a specific tablet, when it is uncarried, creates a large number of sessions. The reason behind such a behavior is that the device keeps connecting to the WiFi network multiple times for syncing of the many tablet apps hence creating background app. traffic at regular intervals.

• A study of the duration of the uncarried sessions by each device type in Fig. 10b shows that the duration is longer in laptop devices. As we have discussed above, the tablet uncarried sessions are usually background app. traffic and hence the sessions are shorter in duration. On the other hand, the laptop sessions having



(a) Zone A: user set SL (b) Zone A: user set SLT (c) Zone B: user set ST (d) Zone B: user set LT

Fig. 11: Access of sensitive websites: percentage sensitive websites are maximum in smartphones - mainly due to social-networking, financial, education and email

longer time duration are indicative of the fact that laptops when left behind are performing some user-assigned tasks like downloads, indicated by the high volume of downlink traffic.

•Fig. 10c shows the categories of different websites to which uncarried sessions create packets. We see that entertainment websites, social networks and portal based apps have highest access during uncarried sessions. Portals and email (also education, as this is a campus network) indicate continuous syncing of apps or open websites to fetch new emails or updates. Social networks indicate background sessions to retrieve notifications. High access of entertainment websites are mainly due to media downloads - which is also confirmed by the high ratio of the downlink to the uplink traffic in these uncarried sessions.

• The creation of shorter but higher number of uncarried sessions by tablets is in a way a redundant delivery of content, as many notifications or updates are sent to both tablets and smartphones. Intelligent co-ordination between the same apps installed in both the devices can be used to eliminate this redundant content delivery - with the knowledge that the smartphone is at a different location from the tablet.

Findings: *(i) In a residential dormitory, when a laptop is the uncarried device, the sessions it creates are longer in length. On the other hand, the uncarried tablets create a larger number but most of them are of shorter duration. (ii) The downlink traffic, in our dataset, is significantly larger than the uplink traffic for uncarried sessions, which indicates that most of the traffic is due to automatic syncing or downloading sort of activity. This is confirmed as most of the uncarried session traffic is to access entertainment websites, social networks and email or portal based apps. (iii) The knowledge that the tablet and mobile are at different locations can lead to reduction in redundant content to the uncarried device.*

4. SECURITY ASPECTS OF MULTI-DEVICE USERS

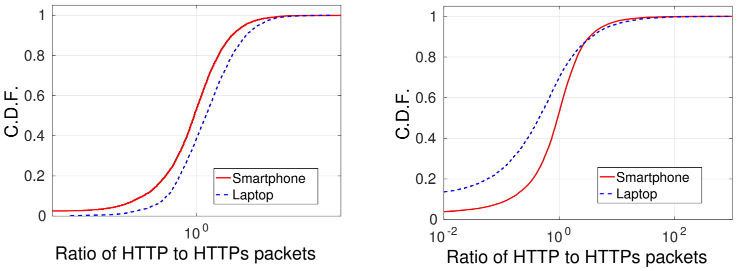
In the wireless networks, there are always security threats, with attacks ranging from D-DOS to spoofing, from malware spread to phishing attacks. For multi-device users in a campus environment, we look at how users of different device types are vulnerable to attacks, based on the websites they access or by connecting to the unencrypted campus wireless network.

4.1. Access of Sensitive Websites

In this section, we study how the device type of a user governs the users' choice of accessing specific websites, specifically websites with content sensitive to users. "Sensitive websites" are defined as websites which reveal information about users preferences or which contains user-sensitive personal information. In addition, websites which require a user to provide log-in information (username and password) are also considered in this category. Major categories of sensitive websites, that we use for our study, are: health, finance, professional, social, productivity and preference. We identify the sensitive websites accessed using keyword-based search on information contained in the packet headers (URIs in GET requests and DNS queries). Table VII shows an example set of the keywords for the different categories that we used to identify the sensitive websites.

Category of Sensitive Websites	Keywords
Health	mydoctor, kaiser, nih.gov, weightwatchers, surgery
Finance	wellsfargo, venmo, paypal, wallet.google, hrblock
Professional	jobsearch, monster, glassdoor
Social	fbcdn, meetup, snapchat, skype, instagram, twitter
Productivity	mail.google, drive.google, slideshare, 4shared
Preference	netflix, groupon, ebay, swagbucks, amazon, wikia

Table VII: Keywords for sensitive website detection



(a) Zone A (b) Zone B

Fig. 12: Ratio of HTTP to HTTPS packets: The number of HTTP packets is higher in Zone A due to use of websites for videos, sports, news,etc.

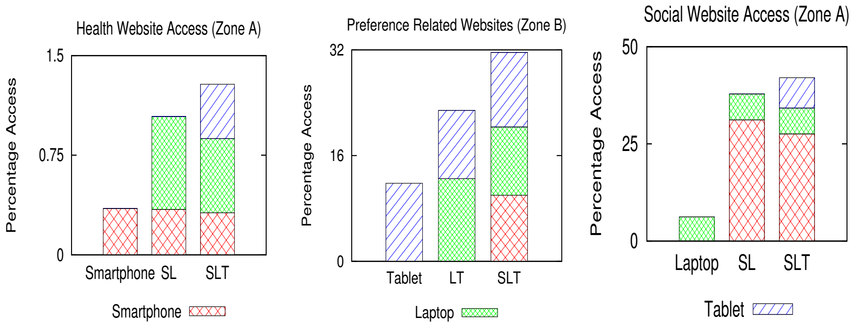


Fig. 13: Access of sensitive website categories: behaviors show overall additive property and consistent use for one specific device type

Fig. 11 represents the percentage of sensitive websites accessed across all the URLs and DNS queries at both Zone A and Zone B. We represent the CDF of sensitive website access for different user sets, based on the device types they carry. In addition, we also look at the ratio of HTTP and HTTPS packets created by smartphones and laptops. Such a representation is shown in Fig. 12. In addition, Fig. 13 shows the access of sensitive website categories from each device type. From this study the major observations include:

• The general pattern of access of sensitive websites in a campus network, as seen in Fig. 11, shows that smartphone devices access sensitive websites more than the other devices. A large part of the smartphone traffic consists of social-networking websites, banking related websites and emails. This contributes to the high access of sensitive websites by smartphones.

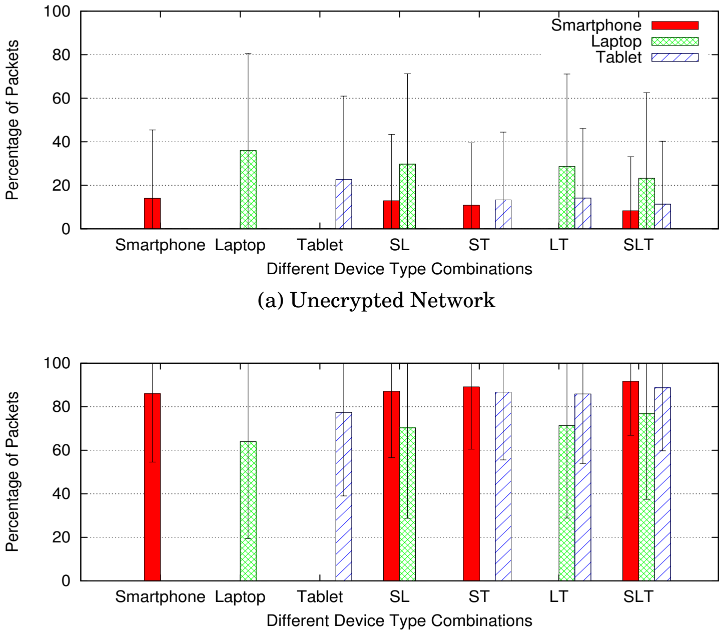
• Another interesting observation is that a specific device type has a consistent amount of access to sensitive websites, irrespective of the presence of other devices. This is in agreement with the observation in section 3.2, where we see that once a device is being used, the presence of other devices does not alter its behavior.

• A comparison of Figs. 11b and 11a with 11d and 11c respectively shows the number of devices with access to no sensitive websites (0%) is significantly lower in the residential scenario of Zone A than Zone B. In Zone B, many devices are accessed for short time periods as compared to dormitories. Also in some cases, the major traffic is background traffic. Thus, many users do not proactively use the devices to access to any sensitive websites.

• Comparison of Figs. 12a and 12b show that Zone A has more HTTP packets than Zone B. This is a contextual location based characteristic, as in the office and work atmosphere of Zone B the websites with HTTPS enabled will be more than in a residential setting. Also in a residential setting there is more usage of the data-extensive websites (sports, news, videos) which mainly operate using HTTP.

• Fig. 12a shows that smartphones have consistently more HTTPS traffic than laptops. HTTPS websites can be considered to be user-sensitive and thus, this result is consistent with our observation in Fig. 11 that smartphones have more access to sensitive websites than other device types.

• The plot depicting the use of specific sensitive website category in Fig. 13 shows that the usage of a particular category does not get shared between multiple devices, but has an overall additive effect. For example, the presence of a banking (or email) app in smartphone gives users more options to check their bank accounts (or emails) - which makes users check these accounts more frequently and causes a general growth in the amount of access. This is similar to the additive nature of activity times, as we have seen in previous sections. Thus, we



(a) Unecrypted Network

(b) Encrypted Network

Fig. 14: Packets in encrypted and unencrypted network: In the unencrypted network, device behaviors are proportional to their screen size

can combine the results to claim, that presence of multiple devices proportionally increases the overall access of the campus WiFi network (both in terms of time and content).

• As observed in the analysis of content from various device types in Fig. 6, we see that smartphones have the maximum access to social networks as compared to other device types (which have very low access to social networks). The use of social networking websites is one of the main reasons behind the high sensitive website and HTTPs access from smartphones. On the other hand, preference related websites and health websites are accessed almost equally from all device types.

Findings: *(i) Sensitive websites constitute a higher percentage of overall content accessed in smartphones as compared to other device types - which indicate that protecting smartphones against security attacks is of utmost importance. (ii) At the same time, we observe that the sensitive website access in the university campus of each device is independent of other device types present, and the overall access patterns is additive. (iii) In addition, we observe that smartphones have higher HTTPS traffic and the HTTP traffic proportion is higher in a residential location as compared to a work/university location.*

4.2. Choice of Encryption in Wireless Network

The wireless network in the campus, from where the data is collected, provides two network options - one is an open wireless network (provides no WiFi encryption), while the other is encrypted. The two different wireless network options are provided from the same access point - so coverage of either network types is never a point of concern on campus. In this section we study how the use of wireless network type in the university campus depends on user's device type and preference.

First, we study the amount of usage which a device is connected to each SSID type. We calculate the percentage usage of a specific wireless network type out of the overall network access. The results shown in Figs. 14a and 14b represent the percentage of packets created via the non-encrypted and encrypted SSID from different device types in all the seven representative user sets in the form of an error plot showing the variation (mean ± standard deviation). Fig. 14 show the access of the unencrypted and encrypted networks is consistent for smartphones, laptops and tablets across different user sets.

On further scrutiny, we can observe the percentage use of the open network is directly proportional to the screen size of the device type. For access to the open wireless network on the campus, students have to provide their login credentials on a portal after connecting to the network and they have to reconnect every time they move to a new access point. For the encrypted network, the password is remembered by the devices and is automatically reconnected every time (without any portal). As expected, the interface of the portal is easier to use and information can be conveniently filled in for devices with bigger screens, which explains the higher usage pattern of the access of unencrypted SSID for devices with bigger screen sizes. The exact opposite trend

X	Y	K-S Stat	p-Value	CV	Hypothesis
smartphone	laptop	0.212	≈0	0.015	Rejected
smartphone	tablet	0.032	0.0037	0.02	Rejected
laptop	tablet	0.182	≈0	0.025	Rejected

Table VIII: K-S Statistic and p-Value

	S	L	T
S	1	0.37	**0.48**
L	0.37	1	0.30
T	**0.48**	0.30	1

(a) Users with 3 devices

	S	L	T
S	1	0.35	**0.57**
L	0.35	1	0.39
T	**0.57**	0.39	1

(b) Users with 2 device

Table IX: Correlation between packet distributions of different user sets

is seen for the encrypted network, as expected, in Fig. 14b. The usage of the encrypted network is higher in smartphones and tablets as compared to laptops.

4.2.1. Selection of Wireless Network Type. Once a device enters a WiFi network, the choice of network type can be made on the basis of a number of factors - like, device type, user's choice, load or location. We discuss each factor in detail with respect to the wireless network on the campus.

Device type dependence: Here, we want to study if the use of the network type in different device types, are inter-related with each other. We consider, as the null hypothesis, the distributions of network type access for all the different device types belong to the same underlying distribution. The alternate hypothesis is that their behaviors are independent. For this purpose, we calculate the two-sample K-S statistic for two empirical distributions (e.g.: smartphone and laptop), say S and L, based on the following equation:

$$\text{K-S statistic} = max(|S(i) - L(i)|), \tag{1}$$

where $S(i)$ denotes the fraction of elements in S with value less than or equal to i and $s(j)$ denotes the fraction of elements in S with values equal to j: $S(i) = \sum_{\forall j \leq i} s(j)$ and $\sum_{\forall j}(j) = 1$. We then compute the p-value, which defines the probability that the null hypothesis is true. A p-value lower than the pre-selected significance level (α = 0.5) indicates the two distributions are different. Another way of interpretation is based on the value of the K-S Statistic - if greater than the critical value - the hypothesis is rejected. The critical value (CV) is calculated as follows:

$$\text{Critical Value} = c(\alpha)\sqrt{\frac{n_1 + n_2}{n_1 \cdot n_2}}, \tag{2}$$

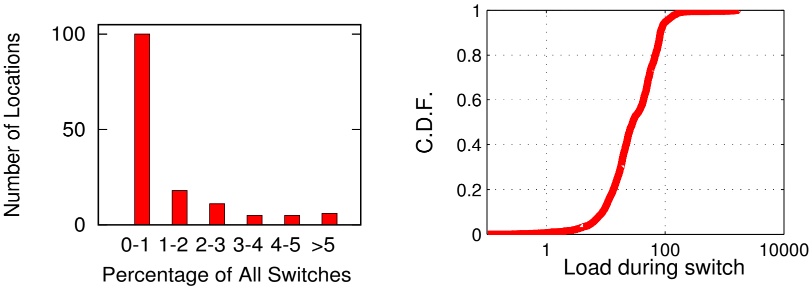
where $c(\alpha)$ is based on the value of α and is equal to 1.36 and n_1 and n_2 are the number of datapoints in each distribution.

The calculated values and conclusions are shown in Table VIII, where X and Y are the two distributions being considered. We observe that the null hypothesis is rejected in all the three cases, hence concluding that the usage of network type by different devices in the campus network are independent of other devices of a user.

User Dependence: The next factor we study is whether the personal choice of users govern the selection of a particular network type for all her users - for example, if a user is security conscious she will ensure to connect all her devices to the encrypted WiFi network at all times. To quantify the dependence we calculate the Pearson's correlation coefficient between the network type usage of two different device types for all the users in a specific user set. We calculate these values for all different MDU user sets (three cases of 2 device types, one case of 3 device type). Table IX shows the correlation values for different user sets.

The results show a higher correlation between the distribution of packets created in each network type for handheld mobile devices as compared to the correlation between other device types. This is observed for all the different multi-device user sets. From Fig. 14, we see the major usage in these cases are of the encrypted network. Handheld devices are in use even when the user is moving around (users with high mobility), thus making the use of open network in this specific university campus inconvenient as users are required to login via the portal whenever they move to a new access point. Thus, users prefer to use a network (the encrypted one) which authenticates automatically in their handheld devices, which explains the comparatively high correlation. Laptop devices do not have such high mobility and hence users' do not have a pre-determined choice of network type in those devices.

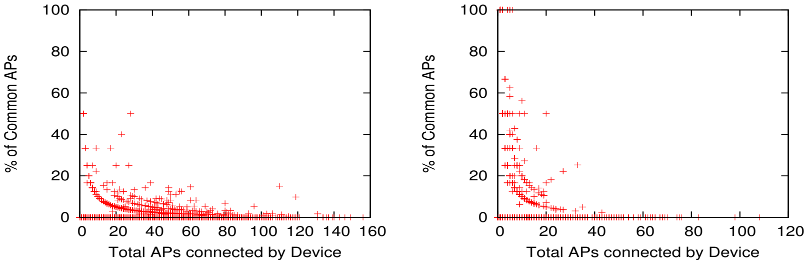
Location and Load Dependence: In addition to user preference or device-type dependence, location and load can also be a factor behind network selection. The location of an access point or the network load at a particular access point can cause a user to select a specific access point encryption type. There might be certain locations where one network type has a better overall performance and as a result, a user is prompted



(a) Change of Network Type across Locations

(b) Load in the Network during SSID switch

Fig. 15: Load and Location Dependence: Contributing factor in some cases, but not the major contribution for selection.



(a) Handheld

(b) Laptops

Fig. 16: Percentage of common access points

to change to that specific network type from the worse performing network. Also, the presence of large number of users can create contention in the network, which might make a certain network not provide enough throughput - prompting the user to switch the network.

Location: To study the location dependence on the wireless network choice, we calculate the number of times a switch of SSID happens at a specific location (building) on campus for all the different device types. Fig. 15a shows the number of locations and the corresponding number of switches, represented as a percentage of all switches, at that location.

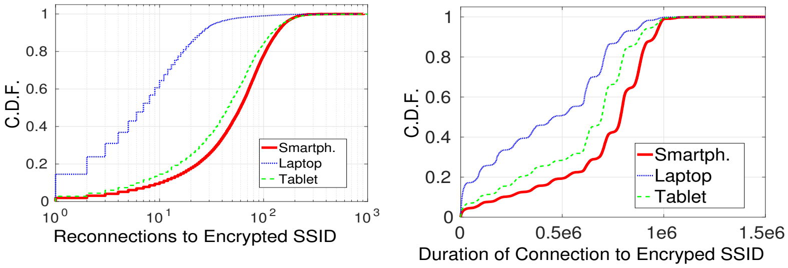
• We observe that most locations have a very low count of SSID switches occurring there. There are only a few number of buildings where more than 2% of the switches occur. From this we can say that there are not a large number of locations where the user is prompted to switch the network type from encrypted to unencrypted network or vice-versa - thus location is unlikely to be the main contributing factor behind network selection.

Load: At an access point a higher device count, than what the AP can handle, can cause the network performance for a new device to be not upto the mark. This can force the user to alter the network type to the network that is serving a smaller number of devices. To see if load is indeed a factor, we calculate the number of devices connected to the other network type in the building, when a user changes the network type she was originally using (encrypted network if the device switches to the unencrypted network). We use this as a coarse-grain measure of the load at the AP instead of the actual traffic volume at that point.

• Fig. 15b is a C.D.F. representation of the number of connected devices when a switch of network type occurs for any of the device types of the user. We see that the most of the switches happen when the number of devices connected to the AP's in the building varies between 10 and 100. Such a count of devices can be considered a low load at most buildings - most of which have 10 to 15 access points. Thus, we can conclude that most of the switches of the network type is hardly caused because of the load in the network.

4.2.2. Characteristics of Switching of Wireless Network Type. The primary question we address in this section is: once a particular device type connects to a wireless network, does it change its network type? Overall statistics show that almost 95% of all devices have no change in the wireless device encryption type over the entire duration of our dataset. The dependence of load and location behind change of wireless network type has been discussed in the previous subsection.

Common Access Points for both network types: In this section, we calculate the number of common access points where a user connects to both the network types as a percentage of the total number of access



(a) Number of re-connections before switching

(b) Time of access before switching

Fig. 17: Connections to the encrypted network and time in the network before switching to the unencrypted network

points to which the user connects. From the scatter plot in Fig. 16 we see the most devices have a low number of common access points. A high percentage is only seen for devices connecting to a low number of access points overall.

We observe similar behavior patterns in smartphones and tablets which further strengthens the claim that handheld usage pattern of a user is correlated. These devices have a fewer number of common access points as they usually keep connected to the encrypted network and does not switch often. Laptops on the other hand have many instances where the number of common APs is a significant amount as they choose to connect to either network type and do not have a pre-defined choice.

Re-connections and Time in same network: Once a user connects to a particular network does the user keep using the same network type in that device all the time? For example, if a user connects to the encrypted network once with her smartphone, does she connect to that same network repeatedly? To study this behavior, we look at the number of times a device is reconnecting to the same network type. We also calculate the time spent in one specific network type before switching to the other network type. Figs. 17a and 17b show, respectively, the cumulative distribution of number of re-connections to the encrypted network and total time before switching to the unencrypted network from the encrypted network type.

Similar behavior for handheld devices (smartphones and tablets), as claimed before, is reconfirmed from Fig. 17. Smartphones and tablets do not switch from the encrypted network often, as seen by the higher number of re-connection instances in Fig. 17a. This is indicative of the earlier observation, where due to the convenience of automatic login to the encrypted network, users keep reconnecting to that network. Handheld devices automatically connect to the previously authenticated encrypted network and the user does not make a choice at every access point she visits. The cumulative distribution of laptops shows fewer re-connections and smaller amount of time spent in the encrypted network. In general, the number of re-connections and time spent in the unencrypted network is lower (as seen in Fig. 14) compared to the encrypted network. Around 40% of users do not reconnect to the unencrypted network more than once, proving that users are in some cases concerned about the security of their devices.

Findings: *(i) The total usage of the unencrypted network in the university campus among different device types is proportional to the device screen size - as larger screens make it more convenient to login to the network using the browser. (ii) The choice of encrypted or unencrypted WiFi network shows loose correlation among different devices of the same user which shows low dependence on user's preferences. Also the choice is unlikely to be dependent on the load or the location of the access point. (iii) On the other hand, the choice is more correlated to the device type which indicates that device-specific factors, which are characteristics of this campus network, such as auto-connect on handheld devices and ease of portal login on laptops play an important role in choosing the network type. (iv) In this specific campus scenario, the use of the encrypted network in handheld devices can be attributed more to the flexibility of usage of the encrypted network, rather than to reasons of security.*

5. DISCUSSION OF RESULTS

In this paper we do a measurement study on multi-device users in a campus wireless network from the point of view of overall network utilization and security aspects in multiple devices of a user. We gained numerous insights through our characterization which can be useful to many entities. We observe that the web browsing patterns in tablets and laptops are very similar for various different interest categories - even though most tablet apps are similar to the ones on smartphones. This information can help app developers as this shows that tablet app development should be more pertinent to laptop-type tasks. This can allow users to offload their laptop access to tablet when mobile. For the same reason, that laptop and tablet network usage is similar, from the perspective of online analytics and advertisers, we observed that mobile combined with laptop or tablet provide a more complete view of user's online footprint as opposed to laptop and mobile or tablet. Presence of multiple devices does not cause the total usage to be shared between device types - but the behavior is additive in nature. As a result, we confirm that any online analytics should span across multiple devices of a user.

Apart from this, since a user with more devices consumes more data overall, schemes that can address content redundancy for content providers as well as device platform developers should be actively investigated. A coordination system can be built which can communicate between the multiple devices of a user and deliver content only to the device which is being accessed at that moment. In addition to reducing the redundant content, this can also make the devices more energy efficient. When a specific device is uncarried or left behind at home, the knowledge that they are at different locations, can help the apps to reduce content delivery to the uncarried device. We also inferred that network operators can improve the IP space utilization by assigning shorter lease times to handheld devices as well as potentially delaying the IP assignment to the handheld devices of multi-device users.

Although expected, we verified that access to sensitive content on mobile platforms is significantly higher, which means protecting against mobile malware is extremely important. Also, we learned that users do not necessarily make an informed decision about the choice of encrypted or unencrypted network, but instead other factors such as convenience of connection to one type on network from a device type affect their choices. This makes it imperative for network providers to make users more informed about the advantages of using a encrypted network.

The characterization study in this paper is based on a campus-wide dataset and the observations are directly applicable to a student population in a university campus. However, for understanding multi-device usage patterns for a larger population and for more generalized inferences, a similar study is required on other representative locations where the daily time-lines and behaviors are different from a university campus. Our dataset includes data from WiFi access points. Thus, our study does not represent user behavior for people who have maximum internet usage using cellular data.

6. RELATED WORK

In recent years, there have been a number of research studies on smartphone characterization. [Falaki et al. 2010a] and [Huang et al. 2010] looked at the usage of smartphones among users, with focus on browsing patterns of users, protocol overhead, radio power usage and management. These studies have been based on data collected via volunteers. Other research efforts have studied the effects of mobility and interaction of users with smartphones at different locations [Trestian et al. 2009], and tried to profile users based on their smartphone usage [Keralapura et al. 2010]. There has also been studies [Falaki et al. 2010b; Xu et al. 2011] that investigated the diversity in users' smartphone interaction patterns and the amount of data consumed. All of these studies are primarily device-centric as they only focus on smartphones and its usage characteristics. Our focus in this work is to explore *user-centric* patterns of network access for multiple devices of the user. Also, as opposed to collecting the data from a single device of volunteers, we have investigated a dataset that can capture network usage pattern of multiple devices of a user.

In a related work, [Gember et al. 2011] have provided a comparative study of overall usage of handheld and non-handheld devices in a campus network. Different from our work, their study mostly characterizes and compares network traffic of handheld and non-handheld devices. However, our objective here is to look at the characteristics of multi-device *users* and how the network access pattern changes for one device in the presence of other user devices.

Apart from the traffic characterization studies, [Papapanagiotou et al. 2012; Chen et al. 2013] have analyzed device connection session lengths for different types of devices. The inferences are shown to be useful in efficient DHCP lease time allocation. In our work, we extend the device-centric view of session lengths to a user-centric view whereby we claimed that delayed IP assignment for devices of multi-device users might be an effective way to improve IP address space utilization.

7. CONCLUSIONS AND FUTURE WORK

In this paper, we presented a detailed characterization of multi-device users in a campus wireless network. Based on the network traces collected for 32,581 users over 8 days, we showed how different network access characteristics of one device of a multi-device user get affected by the presence of other devices. We provide many insights regarding how the characteristics of multi-device users can be useful to various entities such as content providers, advertisers, network operators and application developers. As an extension of this work, we plan to design schemes that can provide improved coordination between the multiple wireless devices of a user. Such a coordination can increase the energy efficiency for user devices as well as decrease the amount of redundant content delivered to all her devices.

REFERENCES

2014. Digital Omnivores Craving More Content Across Devices - Digital Democracy Survey. (2014). www.deloitte.com.

2014. iPhone, iPad, and Mac. Connected like never before. (2014). www.apple.com/ios/ios8/continuity.

2015. Cisco Visual Networking Index. (2015). http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white_paper_c11-520862.html.

Xian Chen, L. Lipsky, Kyoungwon Suh, Bing Wang, and Wei Wei. 2013. Session lengths and IP address usage of smartphones in a university campus WiFi network: Characterization and analytical models. In *Performance Computing and Communications Conference (IPCCC), 2013 IEEE 32nd International*. 1–9. DOI:http://dx.doi.org/10.1109/PCCC.2013.6742781

A.K. Das, P.H. Pathak, Chen-Nee Chuah, and P. Mohapatra. 2014. Contextual localization through network traffic analysis. In *INFOCOM, 2014 Proceedings IEEE*. 925–933. DOI:http://dx.doi.org/10.1109/INFOCOM.2014.6848021

Hossein Falaki, Dimitrios Lymberopoulos, Ratul Mahajan, Srikanth Kandula, and Deborah Estrin. 2010a. A First Look at Traffic on Smartphones. In *Proceedings of the 10th ACM SIGCOMM Conference on Internet Measurement (IMC '10)*. ACM, New York, NY, USA, 281–287. DOI:http://dx.doi.org/10.1145/1879141.1879176

Hossein Falaki, Ratul Mahajan, Srikanth Kandula, Dimitrios Lymberopoulos, Ramesh Govindan, and Deborah Estrin. 2010b. Diversity in Smartphone Usage. In *Proceedings of the 8th International Conference on Mobile Systems, Applications, and Services (MobiSys '10)*. ACM, New York, NY, USA, 179–194. DOI:http://dx.doi.org/10.1145/1814433.1814453

Jinliang Fan, Jun Xu, Mostafa H. Ammar, and Sue B. Moon. 2004. Prefix-preserving IP Address Anonymization: Measurement-based Security Evaluation and a New Cryptography-based Scheme. *Comput. Netw.* 46, 2 (Oct. 2004), 253–272. DOI:http://dx.doi.org/10.1016/j.comnet.2004.03.033

Aaron Gember, Ashok Anand, and Aditya Akella. 2011. A comparative study of handheld and non-handheld traffic in campus Wi-Fi networks. In *Passive and Active Measurement*. Springer, 173–183.

Junxian Huang, Feng Qian, Z. Morley Mao, Subhabrata Sen, and Oliver Spatscheck. 2012. Screen-off Traffic Characterization and Optimization in 3G/4G Networks. In *Proceedings of the 2012 ACM Conference on Internet Measurement Conference (IMC '12)*. ACM, New York, NY, USA, 357–364. DOI:http://dx.doi.org/10.1145/2398776.2398813

Junxian Huang, Qiang Xu, Birjodh Tiwana, Z. Morley Mao, Ming Zhang, and Paramvir Bahl. 2010. Anatomizing Application Performance Differences on Smartphones. In *Proceedings of the 8th International Conference on Mobile Systems, Applications, and Services (MobiSys '10)*. ACM, New York, NY, USA, 165–178. DOI:http://dx.doi.org/10.1145/1814433.1814452

Ram Keralapura, Antonio Nucci, Zhi-Li Zhang, and Lixin Gao. 2010. Profiling Users in a 3G Network Using Hourglass Co-clustering. In *Proceedings of the Sixteenth Annual International Conference on Mobile Computing and Networking (MobiCom '10)*. ACM, New York, NY, USA, 341–352. DOI:http://dx.doi.org/10.1145/1859995.1860034

Lu Liu, Xianghui Cao, Yu Cheng, and Zhisheng Niu. 2014. Energy-Efficient Sleep Scheduling for Delay-Constrained Applications Over WLANs. *IEEE Transactions on Vehicular Technology* 63, 5 (Jun 2014), 2048–2058. DOI:http://dx.doi.org/10.1109/TVT.2014.2313114

Gregor Maier, Fabian Schneider, and Anja Feldmann. 2010. A First Look at Mobile Hand-held Device Traffic. In *Proceedings of the 11th International Conference on Passive and Active Measurement (PAM'10)*. Springer-Verlag, Berlin, Heidelberg, 161–170. http://dl.acm.org/citation.cfm?id=1889324.1889341

Ioannis Papapanagiotou, Erich M. Nahum, and Vasileios Pappas. 2012. Configuring DHCP Leases in the Smartphone Era. In *Proceedings of the 2012 ACM Conference on Internet Measurement Conference (IMC '12)*. ACM, New York, NY, USA, 365–370. DOI:http://dx.doi.org/10.1145/2398776.2398814

Ionut Trestian, Supranamaya Ranjan, Aleksandar Kuzmanovic, and Antonio Nucci. 2009. Measuring Serendipity: Connecting People, Locations and Interests in a Mobile 3G Network. In *Proceedings of the 9th ACM SIGCOMM Conference on Internet Measurement Conference (IMC '09)*. ACM, New York, NY, USA, 267–279. DOI:http://dx.doi.org/10.1145/1644893.1644926

Qiang Xu, Jeffrey Erman, Alexandre Gerber, Zhuoqing Mao, Jeffrey Pang, and Shobha Venkataraman. 2011. Identifying Diverse Usage Behaviors of Smartphone Apps. In *Proceedings of the 2011 ACM SIGCOMM Conference on Internet Measurement Conference (IMC '11)*. ACM, New York, NY, USA, 329–344. DOI:http://dx.doi.org/10.1145/2068816.2068847

Bo Zhao, Qiang Zheng, Guohong Cao, and S. Addepalli. 2013. Energy-Aware Web Browsing in 3G Based Smartphones. In *Distributed Computing Systems (ICDCS), 2013 IEEE 33rd International Conference on*. 165–175. DOI:http://dx.doi.org/10.1109/ICDCS.2013.25