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

CCS Concepts: • **Networks → Network measurement**; **Network monitoring;**

Additional Key Words and Phrases: Campus network, multidevice users, network utilization, smartphone, laptop, tablet

ACM Reference Format:
Aveek K. Das, Parth H. Pathak, Chen-Nee Chuah, and Prasant Mohapatra. 2016. Characterization of wireless multidevice users. ACM Trans. Internet Technol. 16, 4, Article 29 (December 2016), 25 pages.
DOI: http://dx.doi.org/10.1145/2955096

1. INTRODUCTION

In the past year, there has been an increase from 26% to 37%—a growth of 42% in 1 year—in the number of U.S. adults who own a trio of a smartphone, laptop, and tablet [Deloitte 2014]. In addition, 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,

Authors' addresses: A. K. Das and P. Mohapatra, Computer Science Department, University of California, Davis, One Shields Avenue, Davis, CA; emails: {akdas, pmohapatra}@ucdavis.edu; P. H. Pathak, Computer Science Department, George Mason Univeriesty, 4400 University Drive, Fairfax, VA; email: phpathak@gmu.edu (work done while at University of California, Davis); C.-N. Chuah, Electrical and Computer Engineering Department, University of California, Davis, One Shields Avenue, Davis, CA; email: chuah@ucdavis.edu.

DOI: http://dx.doi.org/10.1145/2955096

smartphones). Although these studies are important, they do not include any information about how the device usage behavior changes when a user has other devices, and how all of the user's devices are interdependent with respect to their usage patterns. The increase in the number of multidevice users (MDUs) raises some essential questions, such as how such users use their different wireless devices, what content is accessed on each of them, and what their security preferences and expectations are. In this article, 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 resulting in IP address space exhaustion, delaying or revoking IP addresses based on the usage pattern can be beneficial to providers. Proper information about the device that is actively being used by the user can lead to a reduction in redundant content delivery by content providers. Complete usage pattern information of all devices of a user can be gathered by advertisers and online analytics providers to get a more complete view of a user's online activities beyond the partial view of what is currently available through one device. Last, this information can be exploited by the different applications on a user's devices to carry out intelligent multidevice coordination that can save energy by turning wireless radio on and off depending on the usage pattern. Although there have been recent efforts [Apple 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.

Even though such a characterization study for MDUs has a great deal of potential, acquiring real-world network traces for MDUs 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 a user's ownership of devices. In this work, we accomplish a characterization study of MDUs 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-grain network usage activity for each user and all of her devices. A campus network with a high number of users and wireless devices provides a network that is a good representation of an MDU environment and therefore is a good choice for an MDU study. The characterization study described in this article is based on data collected for nearly 1,000 access points (APs) from a university campus for approximately 30,000 users with total network packet traces of 23TB. We classify a user's wireless devices into three device types: the 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 MDUs*: 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 MDUs 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 is also seen to be almost interchangeable. We also observe that uncarried devices (devices left 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 such as 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 shows 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 or tablet) is continually ON is much shorter than DHCP lease times that are assigned by 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 tablets. In addition, we see that about 7,900 handheld device sessions (greater than 100 seconds in length) do not create a network IP packet even after an IP is assigned. The corresponding inactivity time in laptops is much smaller.

(3) *Security in MDUs*: Web sites that reveal information personal to users generally are accessed more frequently from smartphones; as a result, among the multiple devices of a user, protecting a smartphone against security attacks is most important. In addition, 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 presence.

    The selection of WiFi network type (encrypted vs. unencrypted) on 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 a specific network type from certain types of devices significantly affect the user's choice. Specifically, we see that handheld devices connect to the encrypted network more often, as it is more convenient for a highly mobile user because that network connects automatically. In addition, the use of an unencrypted network, which requires a login every time the user connects to it, is proportional to the device screen size.

Our results discussed in this article 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 an MDU network, the observations discussed in this work can go a long way in representing multidevice usage patterns. In the rest of the article, we introduce the dataset and our methodology for device detection in Section 2. In Sections 3 and 4, we study multidevice utilization characteristics and the security aspects of MDUs in detail. Section 5 includes a discussion on the major findings. After presenting the related work in Section 6, we conclude the article in Section 7.

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 APs. On the controller, the port, through which traffic is forwarded to and from the backbone network, is mirrored to capture the data. The setup for the wireless data capture is shown in Figure 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 a significantly lower number of users as compared to Zone B. This signifies that devices at the residential dormitories

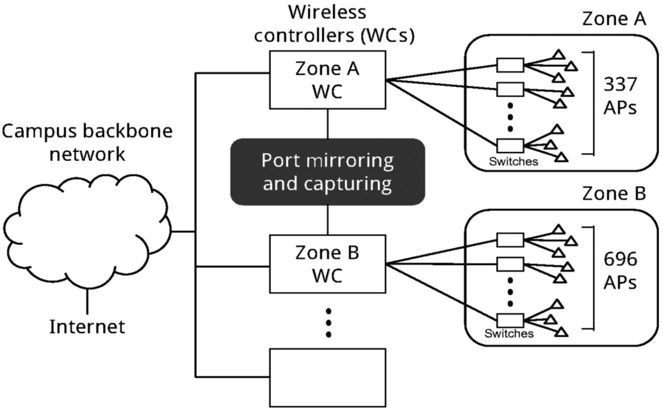


Fig. 1. Campus data capture setup.

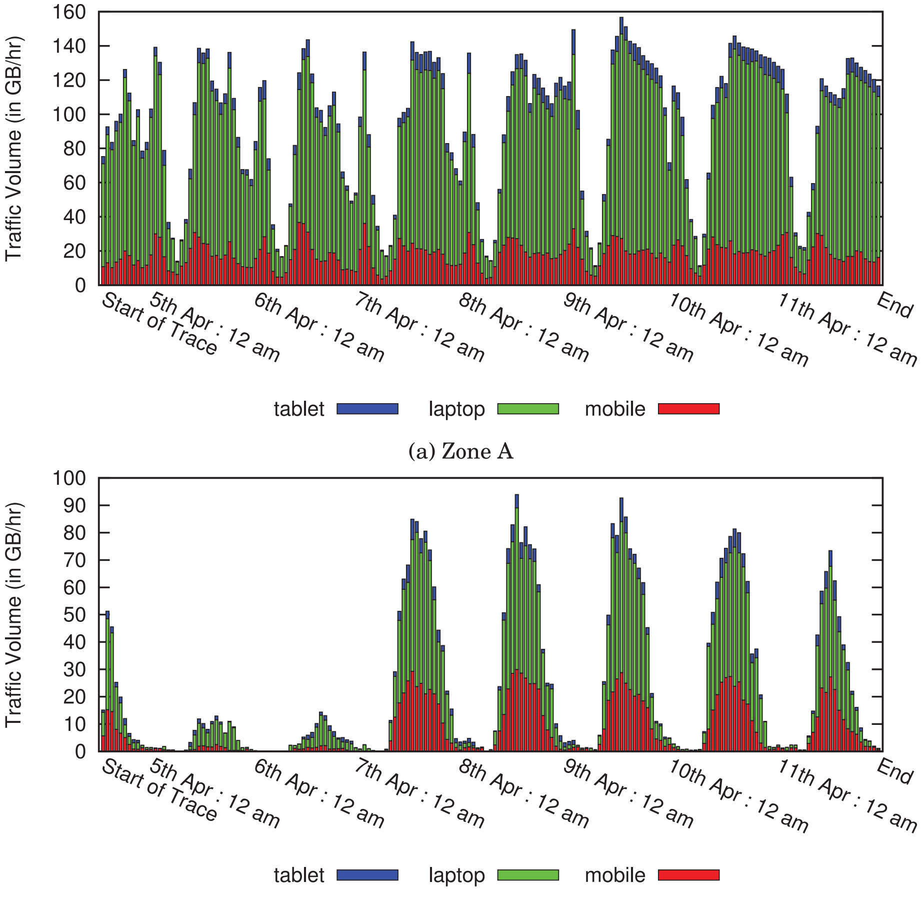
Table I. Dataset for Characterization Study

Location	Zone A	Zone B
Number of Users	7,936	29,925
Number of Devices	13,729	48,284
Number of Packets	19.9 billion	4.8 billion
Number of APs	337	696
Total Size	18.821TB	4.942TB

are connected to the network for a longer duration, which is expected. There is an overlap of 5,280 users among the user sets at two different locations, which indicates that our dataset contains network data created by 32,581 unique users. All network data collected at the controllers was composed 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 a user's wireless devices. An MDU may also have a wired device, such as a desktop computer, but the focus of our work is 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. The presence of wired and cellular network data provides a complete view of each user, but collection from all of these data sources is not feasible. All characterization studies presented herewith are based on WiFi data of the campus network.

Figure 2 represents the total data volume (in gigabytes per hour) variation over the duration of our capture from the multiple devices of a user. A comparison of Figure 2(a) and (b) shows that the data volume at Zone A has two peaks, one after midnight and one at noon, as compared to Zone B, which has one peak at around noon. This is indicative of the location category, as Zone B includes offices, classrooms, and so forth and is expected to have high traffic only during office hours. For the same reason, the traffic on weekends (April 5 and 6) in Zone B is significantly low. Even though it is expected that weekends will have higher overall traffic than weekdays, at Zone A we see a similar trend over the entire 8 days. This is because each of the figures represents 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.



(a) Zone A

(b) Zone B

Fig. 2. Traffic volume in gigabytes per hour at the two locations over the entire time of trace.

2.2. Network Logs

Since our focus in this work is to understand the characteristics of MDUs, 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 AP name to which the device is connected. These logs allow us to match each packet to a user and her device using the IP address.
(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 a 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. DHCP association logs provide the device name for certain MAC addresses. As we show

Table II. Device Count Distribution

Location	Zone A		Zone B	
No. of Devices	User Count	Users (%)	User Count	Users (%)
1	3,675	46.2	14,919	49.9
2	3,018	38	12,158	40.6
3	992	12.6	2,463	8.3
4	195	2.6	313	1
≥ 5	44	0.6	72	0.2

later, we use this information for detection of the device type (smartphone, tablet, or laptop).

With the use of the network session and 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, and so forth 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, MAC addresses, usernames, device names, and names of the APs. We employ prefix-preserving anonymization as proposed in Fan et al. [2004]. The anonymization methods and parameters are kept consistent over all traces and logs 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 device, which shows that there is a valid case for an MDU study in a campus network. However, due to the presence of visitors at Zone A and 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.

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. We are not able to detect when a user borrows another's device but does not login to the wireless network with his own credentials.

2.5. Device Type Detection

One of the most important steps in our study is detection of the device type. We limit our observations to three device types: the smartphone, laptop, and tablet, as seen in Table III. In addition to these three device types, we observe gaming consoles in our dataset. To accomplish this, we combine the following 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 OS (e.g., John-PC). As a result, the DHCP log file mentioned in Section 2.2 includes the device name for some

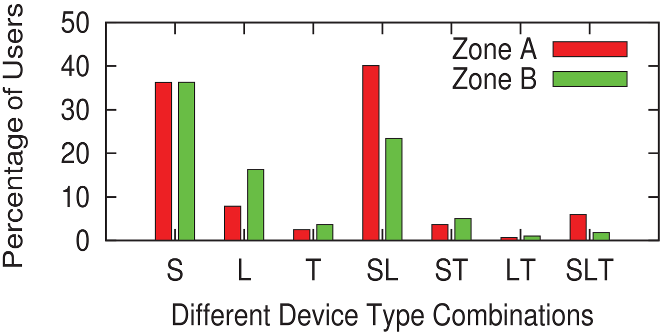


Fig. 3. Device type distributions at the two locations.

of the MAC addresses. Device names like "John-PC," "Andy's MacBookPro," or "Trudy-iPhone" have keywords, the presence of which means 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 hostnames that 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 (e.g., CPU architecture, OS name, browser name, model name), as used in Maier et al. [2010], along with the user agent string for device type detection based on keyword search. A set of keywords is shown in Table IV.

Either of the 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 hostname 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 that can contribute to device detection. Overall, 57.4% of device type information is detected using the user-agent fields and the remaining 42.6% uses the DHCP hostname information. In certain cases where there is 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 from the device count distribution shown previously. The percentage of unclassified devices are 3.06% in Zone A and 12.22% in Zone B. Many of the observations in this article 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 a device's brand name is present, e.g., 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 Figure 3 and Table III. For our analysis, we divide the entire user set into seven distinct groups: S, L, T, SL, ST, LT, and SLT, where a user in set SLT owns a 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 compared to Zone B, as not all users carry out all of their devices. The number of MDUs with no mobile phone is almost negligible, which is expected, as in the present scenario, almost every individual uses and carries around a smartphone. We use the same notation as seen in Figure 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.

Table III. Different Device Types

Device Type Combinations	Zone A	Zone B
S	2,876	10,861
L	626	4,890
T	197	1,106
SL	3,182	6,995
ST	293	1,516
LT	56	307
SLT	463	556

Table IV. Keywords for Device Type Detection

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

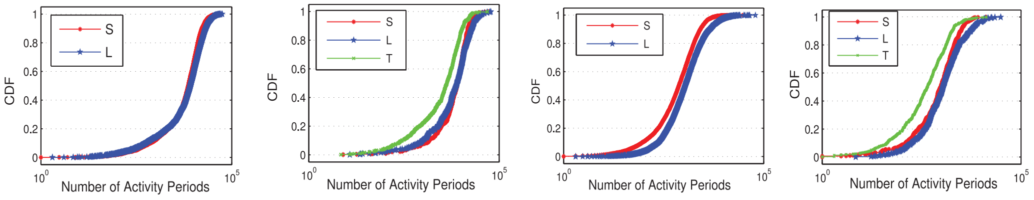
3. MULTIDEVICE UTILIZATION

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

3.1. Time and Packet Characteristics

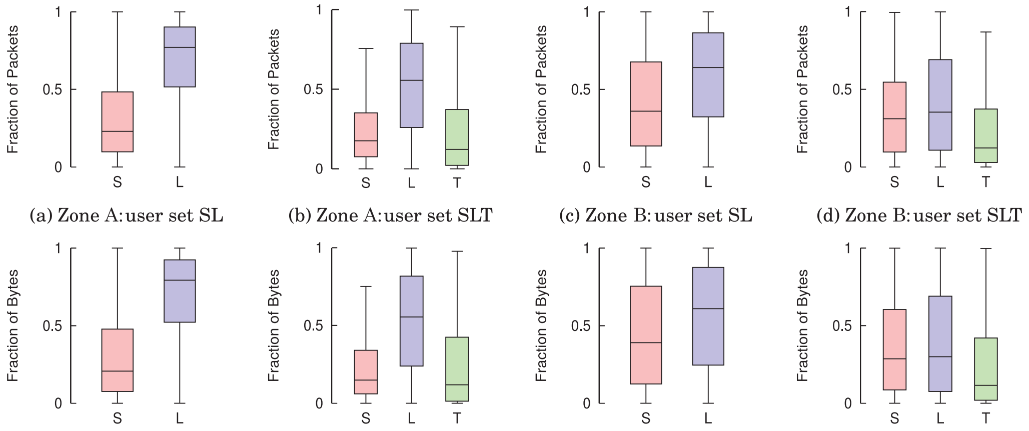
(1) *Activity period per device type*: One of the primary indicators of device usage is the amount of time the device generates network packets. A specific network session is not continuous network usage; it is a combination of many activity periods. We define one activity period as a 10-second time interval during which at least one packet was created by the device. As seen in Das et al. [2014], the activity period determined 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 the presence of other devices, we calculate the number of activity periods created by each device of a user. Figure 4 shows the cumulative density function (CDF) of the activity periods of devices for users with smartphones and laptops (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 that the distributions for smartphones in Figures 4(a) (S(SL)) and 4(b) (S(SLT)) are identical. A similar trend is seen for all other device types at Zones A and B.
- As the distributions for different devices remain the 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 being divided between devices.



(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 the specific device type of a user, irrespective of other user devices.



(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 the inclusion of tablets results in a significant decrease in laptops.

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, such as when 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 of 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 Figure 5(a) through (d) in the form of a box plot. Similarly, Figure 5(e) through (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 representations in Figure 5 show that laptops create significantly higher traffic compared to other device types. This follows intuitively (also mentioned in Cisco [2015]) from the fact that data-extensive Web sites (videos, file downloads, etc.) are mostly accessed from laptops.
- At the residential dormitories of Zone A (Figure 5(c) and (d)), the difference in traffic generated between laptops and handheld devices is much more prominent as compared to Zone B (which includes classrooms, offices, and cafeterias). This is another

Table V. Keywords for Web Site Detection

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

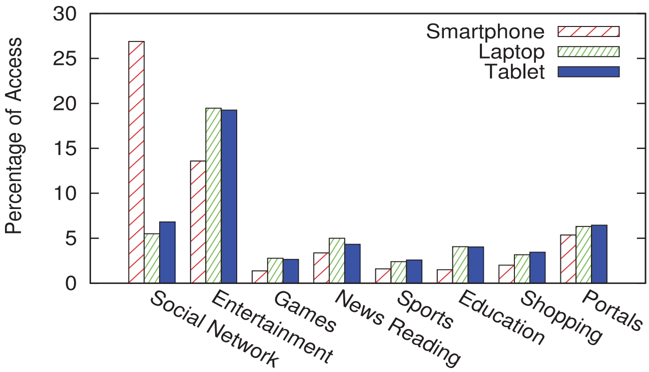


Fig. 6. Access of Web site categories in different device types: usage in laptops and tablets is almost the same but much different from smartphones.

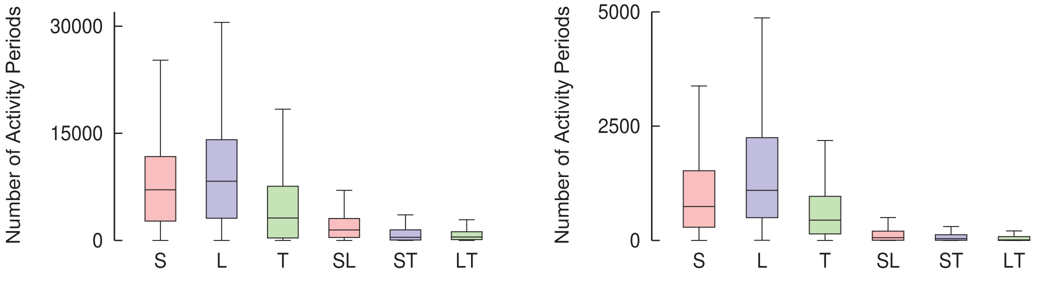
intuitive location-based characteristic that is observed, as handheld devices are used more in a nonresidential setting.

- 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 Figure 5(b) and (d) and their comparison with Figure 5(a) and (c) 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 Web site categories that are accessed by each device type of a user. For this, we employ a keywordbased 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.

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

- Figure 6 shows us that the access of different Web site categories in a campus environment is quite similar between laptops and tablets. However, the Web sites accessed via smartphones show a different pattern. Because of the presence of similar apps (Android and iOS apps) in smartphones and tablets, we expect that Web site access



(a) Zone A (b) Zone B

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

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, such as Web browsing, are more or less similar.

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

(4) *Device usage overlap*: Does the presence of more than one device mean that a user accesses the Internet with all of 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. Figure 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 to the actual usage times.

• The overall overlap amount is very low (maximum being one fourth 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 the instances of two devices being used simultaneously is low.

• The maximum overlap of usage occurs for laptops and mobile phones. In a way, this 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 devices are used for longer periods in a residential setting.

Findings: (i) From the time and overall traffic characteristics of MDUs 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 use of a tablet causes a decrease in the percentage of traffic in laptops, whereas the percentage of 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 laptops and tablets are interchangeable, whereas the mobile usage remains unaffected. (iii) The overlapped usage among multiple devices is very low, with maximum overlap for smartphones and laptops.

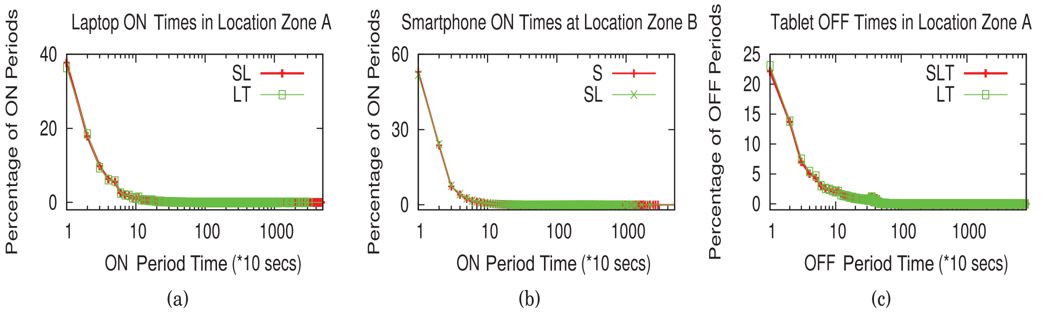


Fig. 8. "ON-OFF" period distributions: intermittent network usage of the different device types is also independent of other user devices present.

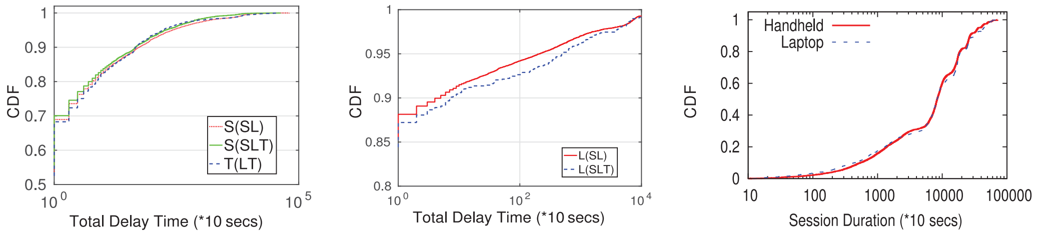
3.2. Intermittent Network Usage Characteristics of Devices

(1) *ON-OFF network usage pattern*: We have studied the total amount of time a device was being accessed by users and observed that in most cases, the behavior is independent of the presence of other devices. However, the total usage time does not reveal any information about how a device is used intermittently. As mentioned earlier, 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 article, 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 result in an OFF period. The ON period starts again when an activity period is observed.

We study how the presence of other devices has an effect on this intermittent user behavior by calculating the ON-OFF times. Figure 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 Figure 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 do not have an effect on the usage of that device. In a way, this is counterintuitive, as we would expect the presence of a smartphone to affect the use of a laptop (or vice versa), but the observations tell otherwise.
- However, the ON usage patterns of a smartphone are different from that of a laptop. Such a pattern is seen for all 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 has 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 earlier.



(a) Handheld devices (b) Laptop devices (c) CDF of sessions where no IP packets are generated

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

A shorter DHCP lease duration during IP assignment for smartphones and tablets can help to better utilize 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 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. Figure 9(a) and (b) show 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 17,000 (5.6%) sessions in handheld devices have a delay of at least 1 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 an IP address and creation of the first packet. However, this corresponding number is as low as 12% for laptops.
- We observe that approximately 9,800 (3.2%) sessions assign an IP address to the device, yet there are no network packets generated. Among these sessions, 7,917 are for handheld devices as compared to 1,887 sessions for laptops. All of these sessions are greater than 100 seconds in duration. Figure 9(c) 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 25,000. 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 in which transient users (or visitors) are very common (e.g., a 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 AP 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 autoconnection for handheld devices; 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. Several research works [Zhao et al. 2013; Liu et al. 2014] have looked into scheduling of a WiFi low power mode or WiFi sleep in an efficient way so that page load times are not compromised and sometimes can even be improved. We also believe that a delayed DHCP assignment system should not increase the load on

Table VI. Uncarried Session Sizes: Uplink and Downlink Traffic Division

Device Type	Total Uncarried Traffic (GB)	Uncarried as Percentage of Overall Traffic	Uplink Traffic (GB)	Downlink Traffic (GB)	Downlink: Uplink Traffic
Tablet	439	9.7	19.4	419.6	21.63
Laptop	1,298.5	10.6	77.75	1,220.75	15.7

Note: High downlink traffic shows that downloads form a major part of uncarried session traffic.

the DHCP server, as the number of IP address requests are the same as when there is no delay.

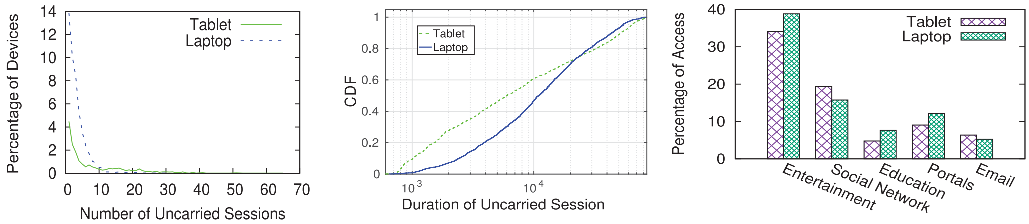
Findings: (i) The ON-OFF device usage pattern of a specific device type in a campus network remains unchanged irrespective of a user's other devices. (ii) We find that the average duration of continuous activity of handheld devices is much smaller compared to the usual DHCP lease times assigned in the university, which indicates that shorter DHCP lease times can be used for handheld devices. (iii) It is also observed that handheld devices have a noticeable difference between the times of IP assignment and creation of the first packet (due to autoconnection to WiFi networks). In certain locations, such as 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 sections, we looked at how 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 use 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 of her devices—one or more devices are left home (or in our case, residential dormitories). These devices 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 a smartphone initiates a network session on 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 1,700GB 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 and uncarried) generated by those devices, whereas the corresponding number for laptops is 10.6%. This shows that uncarried sessions are nonnegligible. In these sessions, the downlink traffic is 21 times the uplink traffic in tablets and 15 times in laptops. This indicates that the bulk of uncarried traffic is caused without user initiation due to background syncing or ongoing downloads.
- Figure 10(a) shows the percentage of devices of each device type creating a specific number of uncarried sessions. It reveals that a large number of laptops present in our dataset have at least one uncarried session. On the other hand, the number of tablets having at least one uncarried session is much lower. However, the count of uncarried sessions is higher for tablets compared to laptops. This indicates that a specific tablet, when uncarried, creates a large number of sessions. The reason behind such 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) 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 greater numbers of sessions. Most uncarried sessions are caused by media downloads or by syncing of social networking apps, emails, and so forth.

• A study of the duration of uncarried sessions by each device type in Figure 10(b) shows that the duration is longer in laptop devices. As discussed earlier, the tablet uncarried sessions are usually background app traffic, and hence the sessions are shorter in duration. On the other hand, the laptop sessions of longer duration are indicative that laptops, when left behind, are performing some user-assigned tasks, such as downloads, which is indicated by the high volume of downlink traffic.

• Figure 10(c) shows the categories of different Web sites to which uncarried sessions create packets. We see that entertainment Web sites, social networks, and portal-based apps have the highest access during uncarried sessions. Portals and email (also education, as this is a campus network) indicate continuous syncing of apps or open Web sites to fetch new emails or updates. Social networks indicate background sessions to retrieve notifications. High access of entertainment Web sites 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 numbers of uncarried sessions by tablets is somewhat of a redundant delivery of content, as many notifications or updates are sent to both tablets and smartphones. Intelligent coordination between the same apps installed on 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 that 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 traffic is due to automatic syncing or downloading sort of activity. This is confirmed because most of the uncarried session traffic is to access entertainment Web sites, social networks, and email or portal-based apps. (iii) The knowledge that the tablet and mobile device are at different locations can lead to a reduction in redundant content to the uncarried device.

4. SECURITY ASPECTS OF MDUS

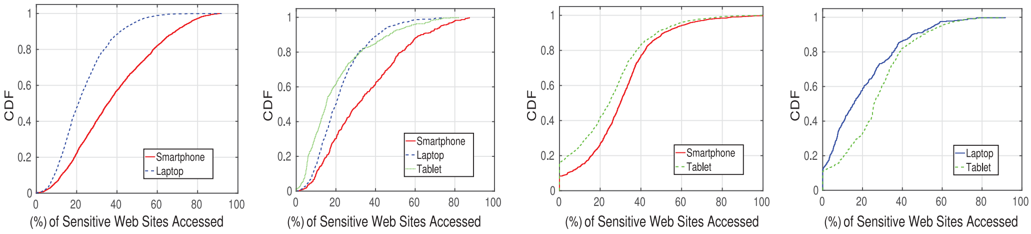
In wireless networks, there are always security threats, with attacks ranging from D-DOS to spoofing and from malware spread to phishing attacks. For MDUs in a campus environment, we look at how users of different device types are vulnerable to attacks based on the Web sites they access or by connecting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

4.1. Access of Sensitive Web Sites

In this section, we study how the device type of a user governs the users' choice of accessing specific Web sites, specifically Web sites with content that is sensitive to

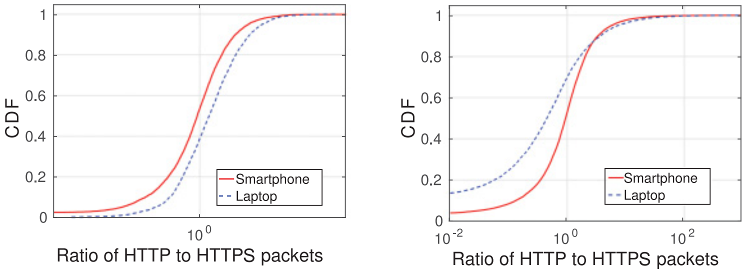
Table VII. Keywords for Sensitive Web Site Detection

Category of Sensitive Web Sites	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



(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 Web sites: the percentage of sensitive Web sites are maximum in smartphones, mainly due to social networking, finance, education, and email.



(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 the use of Web sites for videos, sports, news, and so forth.

users. *Sensitive Web sites* are defined as Web sites that reveal information about user preferences or contain user-sensitive personal information. In addition, Web sites that require a user to provide login information (username and password) are also considered in this category. Major categories of sensitive Web sites that we use for our study are health, finance, professional, social, productivity, and preference. We identify the sensitive Web sites 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 keywords for the different categories we used to identify sensitive Web sites.

Figure 11 represents the percentage of sensitive Web sites accessed across all the URLs and DNS queries at Zones A and B. We represent the CDF of sensitive Web site 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 Figure 12. In addition, Figure 13 shows the access of sensitive Web site categories from each device type. From this study, the major observations include the following:

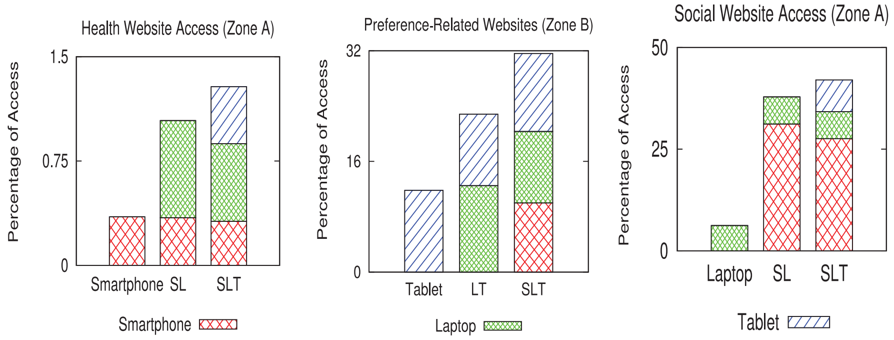


Fig. 13. Access of sensitive Web site categories: behaviors show overall additive property and consistent use for one specific device type.

- The general pattern of access of sensitive Web sites in a campus network, as seen in Figure 11, shows that smartphone devices access sensitive Web sites more than other devices. A large part of smartphone traffic consists of social networking Web sites, banking-related Web sites, and emails. This contributes to the high access of sensitive Web sites by smartphones.
- Another interesting observation is that a specific device type has a consistent amount of access to sensitive Web sites, 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 Figure 11(b) and (a) with 11(d) and (c), respectively, shows that the number of devices with access to no sensitive Web sites (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. Additionally, in some cases, the major traffic is background traffic. Thus, many users do not proactively use the devices to access sensitive Web sites.
- Comparison of Figure 12(a) and (b) shows 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 Web sites with HTTPS enabled will be greater than in a residential setting. In addition, in a residential setting, there is greater use of data-extensive Web sites (sports, news, videos), which mainly operate using HTTP.
- Figure 12(a) shows that smartphones have consistently more HTTPS traffic than laptops. HTTPS Web sites can be considered user sensitive, and thus this result is consistent with our observation in Figure 11 that smartphones have more access to sensitive Web sites than other device types.
- The plot depicting the use of a specific sensitive Web site category in Figure 13 shows that the use 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 a 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 can combine the results to claim that the 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 Figure 6, we see that smartphones have the maximum access to social networks compared to other device types (which have very low access to social networks). High access of social networking Web sites from smartphones is one of the main reasons behind the high

sensitive Web site and HTTPs access from smartphones. On the other hand, preference-related and health Web sites are accessed almost equally from all device types.

Findings: (i) Sensitive Web sites constitute a higher percentage of overall content accessed in smartphones as compared to other device types, which indicates that protecting smartphones against security attacks is of utmost importance. (ii) At the same time, we observe that the sensitive Web site access on 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 that the HTTP traffic proportion is higher in a residential location compared to a work/university location.

4.2. Choice of Encryption in the Wireless Network

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

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

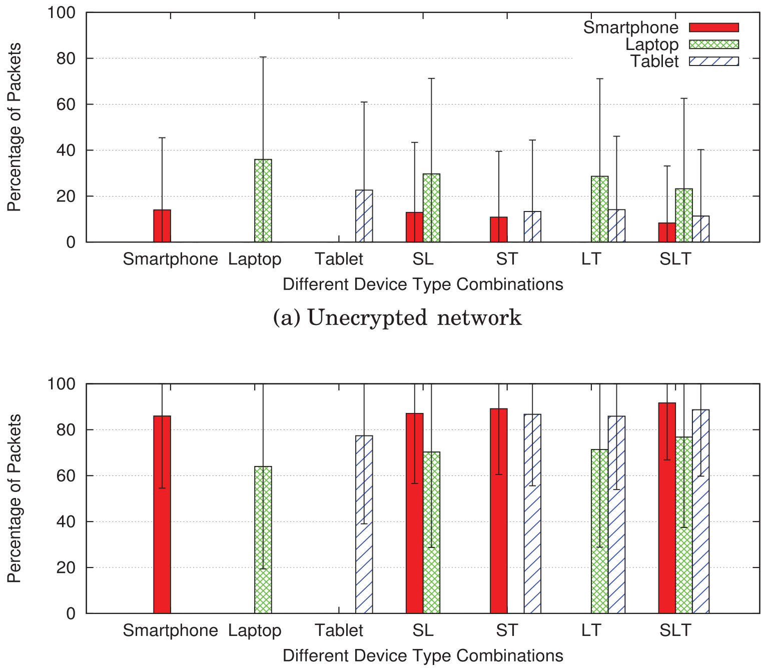
On further scrutiny, we can observe that the percentage of use of the open network is directly proportional to the screen size of the device type. For access to the open wireless network on the campus, students have to provide their login credentials on a portal after connecting to the network, and they have to reconnect every time they move to a new AP. 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 is seen for the encrypted network, as expected, in Figure 14(b). 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 several factors, such as device type, user choice, and location and load. 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 is interrelated. We consider, as the null hypothesis, that the distributions of network type access for all 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



(a) Unecrypted network

(b) Encrypted network

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

Table VIII. K-S Statistic and *p*-Value

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

$\sum_{\forall j}(j) = 1$. We then compute the *p*-value, which defines the probability that the null hypothesis is true. A *p*-value lower than the preselected significance level ($\alpha = 0.5$) indicates that the two distributions are different. Another manner 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 data points 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 three cases, hence concluding that the usage of network type by different devices in the campus network is independent of other devices of a user.

User dependence. The next factor that we study is whether the personal choice of users governs the selection of a particular network type for all of her users. For example, if a user is security conscious, she will be sure to connect all of her devices to the encrypted WiFi network at all times. To quantify the dependence, we calculate the

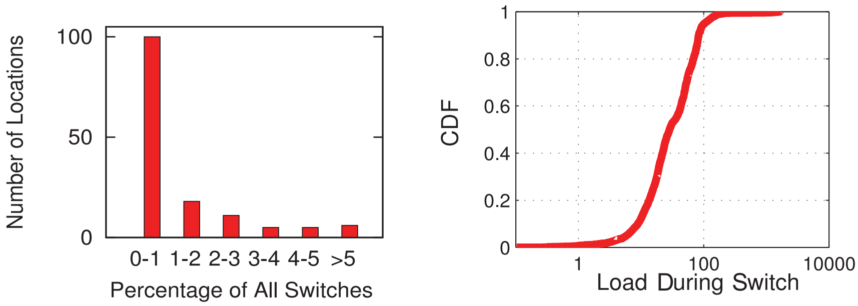
Table IX. Correlation Between Packet Distributions of Different User Sets

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

(a) Users with three 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 two devices



(a) Change of network type across locations

(b) Load in the network during SSID switch

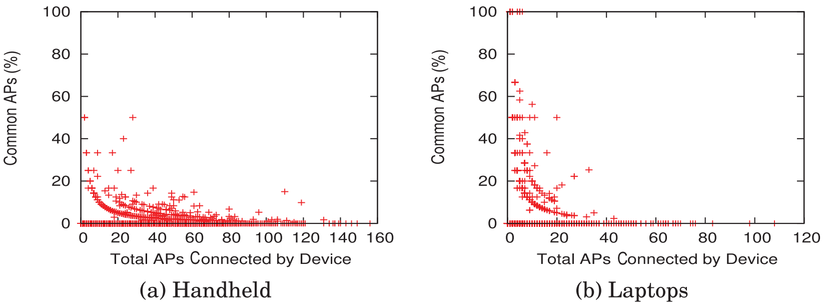
Fig. 15. Load and location dependence: a contributing factor in some cases but not the major contribution for selection.

Pearson correlation coefficient between the network type usage of two different device types for all users in a specific user set. We calculate these values for all different MDU user sets (three cases of two device types, and one case of three device types). 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 compared to the correlation between other device types. This is observed for all different MDU sets. From Figure 14, we see that the major usage in these cases is 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 an open network on this specific university campus inconvenient, as users are required to login via the portal whenever they move to a new AP. Thus, users prefer to use a network (the encrypted one) that 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 predetermined 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 AP or the network load at a particular AP can cause a user to select a specific AP encryption type. There might be certain locations where one network type has better overall performance, and as a result a user is prompted to change to that specific network type from the worse-performing network. In addition, 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 different device types. Figure 15(a) shows the number of locations and the corresponding number of switches, represented as a percentage of all switches, at that location.



(a) Handheld (b) Laptops

Fig. 16. Percentage of common APs.

- 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 an encrypted to unencrypted network or vice versa, and thus location is unlikely to be the main contributing factor behind network selection.

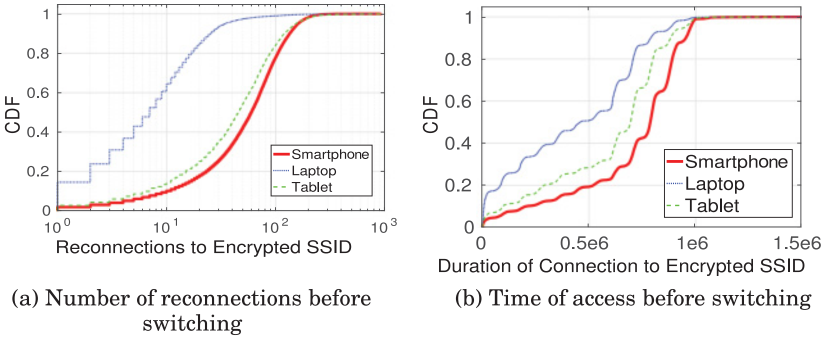
Load: At an AP, a higher device count than what the AP can handle can cause network performance for a new device to be not up to the mark. This can force the user to alter the network type to the network that is serving a smaller number of devices. To see if load is indeed a factor, we calculate the number of devices connected to the other network type in the building when a user changes the network type that she was originally using (an 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.

- Figure 15(b) is a CDF 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 most of the switches happen when the number of devices connected to the APs 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 APs. Thus, we can conclude that most of the switches of the network type are hardly caused because of the load in the network.

4.2.2. Characteristics of Switching of Wireless Network Type. The primary question that we address in this section is the following: 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 was discussed in the previous section.

Common APs for both network types: In this section, we calculate the number of common APs (that is APs where a user connects to both the network types) as a percentage of the total number of APs to which the user connects. From the scatter plot in Figure 16, we see that most devices have a low number of common APs. A high percentage is only seen for devices connecting to a low number of APs overall.

We observe similar behavior patterns in smartphones and tablets, which further strengthens the claim that the handheld usage pattern of a user is correlated. These devices have a fewer number of common APs, as they usually stay connected to the encrypted network and do not switch often. On the other hand, laptops have many instances where the number of common APs is significant, as they choose to connect to either network type and do not have a predefined choice.



(a) Number of reconnections 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.

Reconnections and time in the 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. Figure 17(a) and (b) respectively show the cumulative distribution of number of reconnections 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 Figure 17. Smartphones and tablets do not switch from the encrypted network often, as seen by the higher number of reconnection instances in Figure 17(a). 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 AP she visits. The cumulative distribution of laptops shows fewer reconnections and a smaller amount of time spent in the encrypted network. In general, the number of reconnections and time spent in the unencrypted network is lower (as seen in Figure 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 on 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 a loose correlation among different devices of the same user, which shows low dependence on user preferences. In addition, the choice is unlikely to be dependent on the load or the location of the AP. (iii) On the other hand, the choice is more correlated to the device type, indicating that device-specific factors, which are characteristics of this campus network, such as autoconnect 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 article, we do a measurement study on MDUs on 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 be helpful to 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 a 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 provides a more complete view of a user's online footprint as opposed to laptop and mobile or tablet. The 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 both content providers and device platform developers should be actively investigated. A coordination system can be built that can communicate between the multiple devices of a user and deliver content only to the device that 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 IP space utilization by assigning shorter lease times to handheld devices as well as potentially delaying the IP assignment to the handheld devices of MDUs.

Although expected, we verified that access to sensitive content on mobile platforms is significantly higher, which means that protecting against mobile malware is extremely important. Additionally, we learned that users do not necessarily make an informed decision about the choice of an encrypted or unencrypted network; instead, other factors, such as the convenience of connection to one type on a 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 article is based on a campus-wide dataset, and the observations are directly applicable to a student population on a university campus. However, for understanding multidevice usage patterns for a larger population and for more generalized inferences, a similar study is required in other representative locations where the daily timelines and behaviors are different from a university campus. Our dataset includes data from WiFi APs. Thus, our study does not represent user behavior for people who have maximum Internet usage using cellular data.

6. RELATED WORK

There have been several research studies on smartphone characterization. Falaki et al. [2010a] and Huang et al. [2010] looked at the usage of smartphones among users, with a 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 use [Keralapura et al. 2010]. There have 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 their usage characteristics. Our focus in this work is to explore user-centric patterns of network access for multiple devices of the user. In addition, as opposed to collecting the data from a single device of volunteers, we have investigated a dataset that can capture a 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 nonhandheld devices on a campus network. Different from our work, their study mostly characterizes and compares network traffic of handheld and nonhandheld devices. However, our objective here is to look at the characteristics of MDUs 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] and Chen et al. [2013] 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 MDUs might be an effective way to improve IP address space utilization.

7. CONCLUSIONS AND FUTURE WORK

In this article, we presented a detailed characterization of MDUs on 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 an MDU are affected by the presence of other devices. We provide many insights regarding how the characteristics of MDUs 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 and decrease the amount of redundant content delivered to all of her devices.

REFERENCES

Apple. 2014. iPhone, iPad, and Mac. Connected Like Never Before. Available at http://www.apple.com/ios/ios8/continuity.

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 *Proceedings of the 2013 IEEE 32nd International Performance Computing and Communications Conference (IPCCC'13)*. 1–9. DOI:http://dx.doi.org/10.1109/PCCC.2013.6742781

Cisco. 2015. Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2015–2020 White Paper. Retrieved November 10, 2016, from http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white_paper_c11-520862.html.

A. K. Das, P. H. Pathak, C.-N. Chuah, and P. Mohapatra. 2014. Contextual localization through network traffic analysis. In *Proceedings of the 2014 IEEE INFOCOM Conference (INFOCOM'14)*. 925–933. DOI:http://dx.doi.org/10.1109/INFOCOM.2014.6848021

Deloitte. 2014. Digital Omnivores Craving More Content Across Devices: Digital Democracy Survey. Available at http://www.deloitte.com.

Hossein Falaki, Dimitrios Lymberopoulos, Ratul Mahajan, Srikanth Kandula, and Deborah Estrin. 2010a. A first look at traffic on smartphones. In *Proceedings of the 2010 ACM SIGCOMM Internet Measurement Conference (IMC'10)*. ACM, New York, NY, 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, 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. *Computer Networks* 46, 2, 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 SIGCOMM*

Internet Measurement Conference (IMC'12). ACM, New York, NY, 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, 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 16th Annual International Conference on Mobile Computing and Networking (MobiCom'10)*. ACM, New York, NY, 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, 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)*. 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 SIGCOMM Internet Measurement Conference (IMC'12)*. ACM, New York, NY, 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 2009 ACM SIGCOMM on Internet Measurement Conference (IMC'09)*. ACM, New York, NY, 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 Internet Measurement Conference (IMC'11)*. ACM, New York, NY, 329–344. DOI:http://dx.doi.org/10.1145/2068816.2068847

B. Zhao, Q. Zheng, G. Cao, and S. Addepalli. 2013. Energy-aware Web browsing in 3G based smartphones. In *Proceedings of the 2013 IEEE 33rd International Conference on Distributed Computing Systems (ICDCS'13)*. 165–175. DOI:http://dx.doi.org/10.1109/ICDCS.2013.25

Received July 2015; revised March 2016; accepted June 2016