Identifying Diverse Usage Behaviors of Smartphone Apps

Qiang Xu
University of Michigan
Ann Arbor, MI
qiangxu@umich.edu

Jeffrey Erman
AT&T Labs Research
Florham Park, NJ
erman@research.att.com

Alexandre Gerber
AT&T Labs Research
Florham Park, NJ
gerber@research.att.com

Z. Morley Mao
University of Michigan
Ann Arbor, MI
zmao@umich.edu

Jeffrey Pang
AT&T Labs Research
Florham Park, NJ
jeffpang@research.att.com

Shobha Venkataraman
AT&T Labs Research
Florham Park, NJ
shvenk@research.att.com

ABSTRACT

Smartphone users are increasingly shifting to using apps as "gateways" to Internet services rather than traditional web browsers. App marketplaces for iOS, Android, and Windows Phone platforms have made it attractive for developers to deploy apps and easy for users to discover and start using many network-enabled apps quickly. For example, it was recently reported that the iOS AppStore has more than 350K apps and more than 10 billion downloads. Furthermore, the appearance of tablets and mobile devices with other form factors, which also use these marketplaces, has increased the diversity in apps and their user population. Despite the increasing importance of apps as gateways to network services, we have a much sparser understanding of how, where, and when they are used compared to traditional web services, particularly at scale. This paper takes a first step in addressing this knowledge gap by presenting results on app usage at a national level using anonymized network measurements from a tier-1 cellular carrier in the U.S. We identify traffic from distinct marketplace apps based on HTTP signatures and present aggregate results on their spatial and temporal prevalence, locality, and correlation.

Categories and Subject Descriptors

C.2.3 [**Computer-Communication Networks**]: Network Operations - Network monitoring; C.4 [**Performance of Systems**]: Measurement techniques

General Terms

Measurement

Keywords

Smartphone Apps, App Usage Behaviors

1. INTRODUCTION

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

IMC'11, November 2–4, 2011, Berlin, Germany.

Copyright 2011 ACM 978-1-4503-1013-0/11/11 ...$10.00.

The number and popularity of mobile apps is rising dramatically due to the accelerating rate of adoption of smartphones. For example, Android has 150K apps and 350K daily activations [12]. Preinstalled with marketplace portals such as the AppStore on iOS, Market on Android, and MarketPlace on Windows Mobile, popular smartphone platforms have made it easy for users to discover and start using many network-enabled apps quickly. By Jan 22, 2011, more than 350K apps are available on the AppStore with downloads of more than 10 billion [1]. Furthermore, the appearance of tablets and mobile devices with other form factors, which also use these marketplaces, has increased the diversity in apps and their user population. The existence of marketplaces and platform APIs have also made it more attractive for some developers to implement apps rather than complete web-based services. Despite the increasing importance of apps as gateways to network services, we have a much sparser understanding of how, where, and when they are used compared to traditional web services, particularly at scale. This paper takes a first step in addressing this knowledge gap.

A previous study found evidence that there is substantial diversity in the way that different people use smartphone apps [7]. However, because the study relied on volunteers using instrumented phones, it was limited to two platforms and less than three hundred users in a few geographic areas. Other studies of mobile application/app usage [11, 18, 22] have been similarly limited in scope. Thus, it is difficult to extrapolate these results to make representative conclusions about spatial locality, temporal variation, and correlation of apps at scale. For example, "where are apps more popular?", "How is their usage distributed across a country?", "How does their usage vary throughout the day?". While there have been studies of smartphone performance at larger scales [14, 2, 10], which use volunteer measurements or network data to obtain measurements at scale, measuring the usage of different apps from these data sources is more challenging. Volunteer measurements are typically obtained by deploying a measurement tool via an app marketplace, but many popular platform APIs do not permit the measurement of other apps in the background, so it is difficult to write an app that captures this information. Network data may contain information that can identify app behaviors, but this information is not typically part of standard traces. To make representative conclusions about apps, we require a network data set that identifies apps in network traffic and contains a significant number of measurements covering a representative number of devices, users, locations, and times.

In this study, we address the limitations by collecting anonymized IP-level networking traces in a large tier-1 cellular network in the U.S. for one week in August 2010. In contrast to previous work, we use signatures based on HTTP headers (included in the IP-level trace) to distinguish the traffic from different apps. Due to the format of `User-Agent` in HTTP headers when mobile apps use standard platform APIs, this technique gives us the ability to gather statistics about each individual app in a marketplace, not just categories of network traffic characterized by port number. Moreover, our work examines the spatial and temporal prevalence, locality, and correlation of apps at a national scale, not just in one area or over a small population of users.

To our best knowledge, our study is the first to investigate the diverse usage behaviors of individual mobile apps at scale. In this study, we make the following five contributions:

- The data set that we use to study mobile apps is significantly more diverse geographically and in user base than previous studies. It covers hundreds of thousands of smartphones throughout the U.S. in a tier-1 cellular network. This allows us to make more generalizable conclusions about smartphone usage patterns.

- We find that a considerable number of popular apps (20%) are *local*, in particular, radio and news apps. In terms of traffic volume, these apps are accountable for 2% of the traffic in the **smartphone apps** category (*i.e.*, all the marketplace apps that can be identified by `User-Agent`) – that is, their user base is limited to a few U.S. states. This suggests significant potential for content optimization in such access networks as LTE and WiFi where content can be placed on servers closer to clients. Furthermore, it suggests that network operators need to understand the impact of different app mixes in different geographical areas to best optimize their network for user experience.

- Despite this diversity in locality, we also find that there are similarities across apps in terms of geographic coverage, diurnal usage patterns, *etc.* For example, we find that some apps have a high likelihood of *co-occurrence* on smartphones – that is, when a user uses one app, he or she is also likely to use another one. Users also use several alternatives for the same type of app (*e.g.*, multiple news apps). These findings suggest that some apps can be treated as a "bundle" when trying to optimize for their user experience and that there may be opportunities for integration.

- We also find that the diurnal patterns of different genres of apps can be remarkably different. For example, news apps are much more frequently used in the early morning, sports apps are more frequently used in the evening, while other apps have diurnal patterns less visible and their usage is more flat during a day. These findings suggest that cloud platforms that host mobile application servers can leverage distinct usage patterns in classes of apps to maximize the utilization of their resources. Furthermore, network operators may be able to leverage these results by optimizing their network for different apps during different times of the day.

- Mobility patterns can be inferred from network access patterns. Some apps are more frequently used when users are moving around; some of them are used more often when users are stationary. Mobility affects connectivity and performance, so bandwidth sensitive apps that are mobile may need to consider techniques to compensate for bandwidth variability. We find that there is a significant degree of diversity in the mobility of apps.

The rest of this paper is organized as follows: Related work is discussed in §2, §3 describes our data set, §4 presents our measurement results, §5 outlines some implications, and we conclude our study in §6.

2. RELATED WORK

A plethora of studies focus on understanding smartphone apps from different perspectives. Among them, studies of smartphone usage have yielded insights into different entities in the mobile computing community, *e.g.*, content providers, network providers, OS vendors, mobile app designers, *etc.* Accordingly, understanding the usage of mobile apps is critical for content providers to generate, optimize, and deliver content, for network providers to allocate radio resources, for OS vendors to support on-device apps, for app designers to implement efficient programs, *etc.* Overall, our study is the first that attempts to address the lack of sufficient knowledge about how, where, and when mobile apps are used at a national scale.

A group of studies attempted to improve the performance of mobile apps via OS infrastructure support [5, 13, 10, 6], *e.g.*, offloading resource intensive computation to cloud [5], providing clean intermediate interface for apps by the OS [13], and signaling mobile devices by network providers via notification channel to save resource [10]. Our study is complementary to these, as it focuses on profiling the usage patterns of mobile apps; we note that the design of supportive infrastructure would also benefit from the knowledge of mobile app usage patterns.

Also related are studies that proposed measurement tools for smartphone devices characterizing either the device performance or the performance of certain apps [14, 30, 21], *e.g.*, 3GTest [14] measures the network performance of popular smartphone platforms, PowerTutor [30] profiles energy consumption of running apps on Android, ARO [21] characterizes the radio resource usage of mobile apps, *etc.* Compared to these studies, we focus on usage patterns of mobile apps rather than their performance, but our work also has implications on resource consumption.

Studies have also proposed creative mobile apps to enhance user experience under mobility [17, 4, 3, 16], *e.g.*, xShare [17] enabling friendly, efficient, and secure phone sharing on existing mobile phones, Escort [4] leading a user to the vicinity of a desired person in a public place, *etc.* Although mobile apps are fixed in our study, our work provides app designers with measurements and directions that can help them improve design decisions.

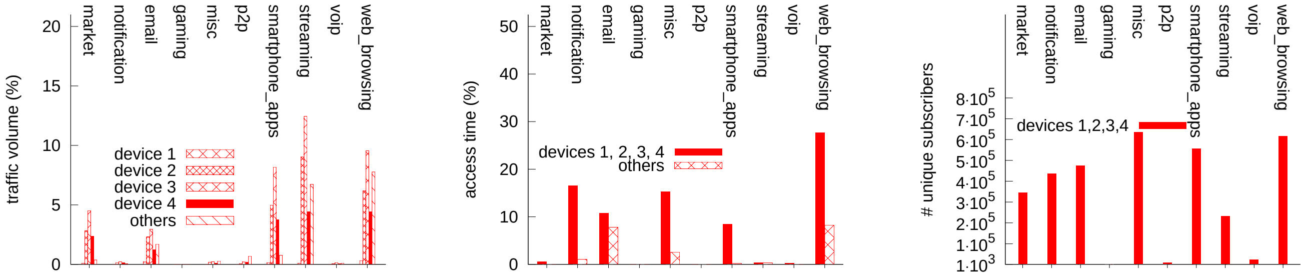
Besides app usage, app selection has been explored as well in context-aware mobile apps recommendation systems [26, 29]. A key requirement for an app recommendation system is to identify the users who share certain similar app interests so that it can predict apps of interest. Understanding patterns in user interests is also part of our study.

On a large scale, there have been studies characterizing the mobile traffic [11, 18, 8, 15] and user interactive behaviors with smartphones [7, 22, 23]. Compared to these studies, our study (1) relies on a data set that can represent the majority of smartphone users across the U.S.; (2) covers the impact of more than one factor, *e.g.*, location, time, device, user, *etc.*; (3) places more emphasis on the usage of smartphones rather than traffic flows, content type, and WiFi usage.

We believe that our study makes an important step in addressing the lack of knowledge of usage behaviors of mobile apps.

3. OVERVIEW OF DATA SET

3.1 Data Set



a. distribution of traffic volume

b. distribution of access time

c. distribution of subscribers

Figure 1: distribution of traffic volume, access time, and subscribers across categories of apps.

In this paper, we use an anonymized data set from a tier-1 cellular network provider in the U.S. It is collected during the week of August 24th, 2010 – August 30th, 2010. The data set contains flow-level information about IP flows carried in PDP Context tunnels (*i.e.*, all data traffic sent to and from cellular devices). This data set is collected from all links between SGSNs and GGSNs in tier-1 network's UMTS core network. Hence, we have a nationwide view of cellular data traffic. Due to volume constraints, only traffic from a uniform random sample of devices is collected. For a random sample of devices, the data contains the following information for each IP flow per minute: the start and the end timestamps, per-flow traffic volume in terms of both the bytes and the number of packets, the device identifier, and the app identifier. All device and subscriber identifiers (*e.g.*, IMSI, IMEI) are anonymized to protect privacy without affecting the usefulness of our analysis. Furthermore, the data sets do not permit reversing the anonymization or re-identification of subscribers.

App identifiers include information about application protocol (*e.g.*, HTTP, DNS, and SIP) and class (*e.g.*, streaming audio, streaming video, web, email). Moreover, given that these popular smartphone platforms include the app's name in the `User-Agent` field when the app uses the standard API to access URL network resources, the marketplace apps can be identified by the `User-Agent` field in HTTP headers. We focus on these apps in this paper and we classify them into the **smartphone apps** category. Note that the browser and YouTube are not included in smartphone apps since they come with the smartphone OS and are not present in the marketplace.

We further categorize smartphone apps by the genre that it is listed under in its platform's marketplace. To find the category, we use the API provided by each smartphone platform's market to search for the app name presented in the HTTP `User-Agent`. While the API typically returns multiple results that match the query (by treating it as a wild card), we manually validated the top apps whether the top first result is correct. Table 1 shows the corresponding validation results of the querying. Upon the response of each query, we have three attitudes: *right*, *wrong*, and *unknown*. We consider a response as right or wrong only if we are confident on the correctness. For example, we believe that a `User-Agent` containing `Pandora` is a music app and another `User-Agent` including `Facebook` is not a sport app. There are some apps that are very hard to tell by us according to their `User-Agent` fields such as `WSDU 4`, which we label as unknown. Additionally, those queries whose responses are empty are labeled as unknown. According to Table 1, we are confident on the correctness of the top first result for the majority of the smartphone apps.

In this paper, we are concerned about four main features per app:

top X apps	right (%)	wrong (%)	unknown (%)
10	8 (80%)	0 (0%)	2 (20%)
20	17 (85%)	1 (5%)	2 (10%)
50	46 (92%)	2 (4%)	2 (4%)
100	91 (91%)	4 (4%)	5 (5%)
200	176 (88%)	5 (3%)	19 (10%)
500	427 (85%)	14 (3%)	69 (14%)

Table 1: accuracy of using `User-Agent` to categorize apps (via manual comparison to app names in the app marketplace).

traffic volume, **access time**, **unique subscribers**, and **locations**. We estimate traffic volume as the sum of the flow byte counts, access time as the sum of the flow durations (with a precision of seconds), and the number of unique subscribers as the number of distinct anonymous device identifiers. There is only one anonymized identifier per distinct device. To determine the location of each device at the time a flow is in progress, we use the cell sector identified in the PDP context used to tunnel the flow. This cell sector is typically recorded when the PDP context begins, when a device moves far enough that the SGSN its traffic routes through changes, switches from 2G to 3G (or vice versa), or switches from 3G to WiFi. While this sector may be slightly stale, previous work [27] showed that they are still almost always accurate to within 40 kilometers. Thus, they suffice for most of our results that only look at U.S. states as distinct regions. For other results we present on sector changes, we may underestimate the number of changes due to this limitation.

In total, the sample data set includes approximately 600K distinct subscribers and approximately 22K distinct smartphone apps.

3.2 Limitation

Our approach to identifying apps using the HTTP `User-Agent` field may miss traffic that does not use the standard platform URL API. However, in Section 3.3 we show that this approach captures a large fraction of traffic that is not email, web browsing, streaming, or a marketplace download (which we identify separately based on other well known heuristics). Obviously, our data set will not capture app usage except when there are network flows. This is acceptable for our study, since we are primarily interested in apps that are gateways to Internet services, not apps that do not use the network.

Another limitation is the time difference when we use these August `User-Agent` fields generated from the trace during the week of August 24th, 2010 – August 30th, 2010, and query them on the current marketplace. Because developers may change the `User-Agent` field in updating their apps, this may result inaccuracy of smartphone app identification. However, according to Table 1, this effect should be small.

3.3 Traffic Summary

Figure 1 shows a summary of all traffic in our data set. Devices 1, 2, 3, and 4 are four major device types in this tier-1 network. Figure 1(a) shows the distribution of traffic volume. We observe that the volume of known smartphone apps traffic is comparable with the traffic volume of web browsing and other HTTP traffic, which is a major motivation for our study. Moreover, the market category also contributes to considerable traffic, which indicates a high demand for smartphone apps from subscribers.

Figure 1(b) shows the distribution of access time of app categories. It is interesting to note that the streaming category is only accountable for a small fraction of the total network access time of all smartphone apps. The gaming, p2p, and voip categories include mostly port and header-identified traffic for common desktop apps. We see that they have a small fraction of both traffic volume and access time, which means that these apps are not common on devices on this cellular network. Figure 1(c) shows the distribution of number of unique subscribers. In this figure, the misc category includes DNS requests, so the misc category roughly has the same number of subscribers as total number of subscribers that we observed in the data set. The smartphone apps and web browsing categories cover almost all the subscribers.

For the remainder of this paper, we only examine traffic in the smartphone apps category.

4. USAGE PATTERNS OF SMARTPHONE APPS

In this section, we investigate how, where, and when smartphone apps are used from spatial, temporal, and user perspectives. We first choose appropriate metrics to evaluate smartphone apps, and then attempt to understand the impact of location, time, user, and app interest accordingly.

4.1 Characterizing Usage with Different Metrics

We begin our analysis by presenting some broad characteristics of smartphone app usage. For our analysis, we choose a number of different natural metrics that profile network activity. We use the following three metrics for each app through most of our analysis: (a) **traffic volume**, defined as the number of bytes consumed by all subscribers using the app; (b) **number of subscribers**, defined as the number of unique subscribers using this app throughout our week-long data set; (c) **network access time**, defined as the total duration summed across all the IP flows generated by the app over our week-long data set.

Figure 2 shows CDFs of these metrics for the apps. For each metric, we aggregate together all the users of a particular app. The long tail of these CDFs directly shows the huge diversity in smartphone apps and their network characteristics. The top app in Figure 2(a) is a "personalized Internet radio app", and is responsible for more than 3TB data in one week, while the majority of smartphone apps generate only 1 – 10 MB over the same time period. Note that this top app is by itself responsible for generating over 50% of the total traffic volume in the smartphone apps category. This dramatic variation in the traffic volume is due to many factors, *e.g.,* app genres, popularity of apps, device types, preferences of the user base, content of apps, *etc.* For example, both news apps and radio apps may provide users with the latest news, but news apps typically deliver most of their content via text while radio apps deliver content via streaming audio; thus, users of these two apps would receive news on their smartphones with a substantial difference in the volume of traffic generated.

We observe a similar variation in Figure 2(b). The top app here is a "social utility connecting people", with a total network access time exceeding 100 years (aggregated across all its users). This app alone contributes to 86% of the total network access time of the smartphone apps category, but the majority of the smartphone apps are seen accessing the network for only about 1 minute - 1 hour. This "social utility" app also has the largest number of unique subscribers, 540,230 according to Figure 2(c). The total number of unique subscribers in our data set is 633,892 by examining the number of unique subscribers with DNS requests in Figure 1. Thus, we may estimate that 6 in every 7 subscribers use this "social utility" app on their smartphones. Recall that this data set contains only a random sample of subscribers, so the numbers here do not reflect the total number of subscribers in the cellular network. Around 60% smartphone apps have no more than 10 unique users in our data set, thus illustrating the long tail of smartphone apps on the market. Because of this long tail, we filter out the smaller apps for some of our analysis as they do not have enough measurements. We discuss this further in §4.2.

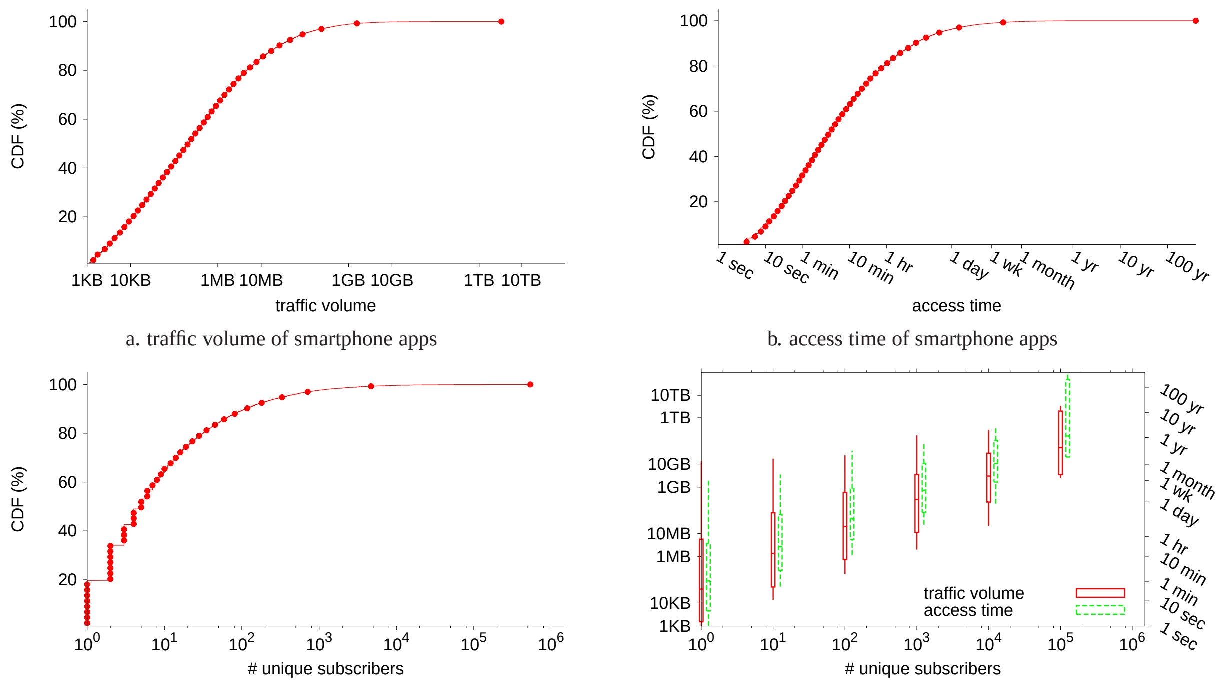
Figure 2(d) shows the correlation between the traffic volume and the number of unique subscribers, and between the access time and the number of unique subscribers. We aggregate the apps with the same $10^{\lfloor \lg N \rfloor}$ (N is the number of unique subscribers) and present the minimum, 25^{th} percentile, median, 75^{th} percentile, and maximum in each aggregation point accordingly. Both the traffic volume and the network access time roughly increase linearly with the number of unique subscribers, but the high variation still exists in the correlation. Due to the high variation, it is difficult to estimate the an app's traffic volume and network access time based only on its number of users. However, given a certain large number of apps together, their traffic volume and access time may be predictable. Accordingly, cellular providers may be able to estimate the radio resource consumption and allocate radio resources.

Figure 3 shows the CDFs of the apps' traffic volume and access time, now normalized by the number of subscribers that use the app. We see a similar variation across apps in these CDFs as well. For example, in Figure 3(a), the app with the largest traffic volume per subscriber consumes 5GB in one week, but the majority of apps consume less than 1MB data per subscriber in the week. Likewise, in Figure 3(b), the app with the longest access time per subscriber lasts for 2 days in one week, while the majority of apps access the network for only 10 seconds - 1 hour per subscriber in the week. From Figure 2 and Figure 3, we also observe apps with very marginal usage in the long tail, *e.g.,* the app consuming only less than 1KB, the app accessing the network less than 10 sec, and the app with only one user. These numbers indicate why we need to filter out these tiny apps for our analysis.

4.2 Popular Smartphone Apps

Figure 2 shows that there are a substantial number of smartphone apps with only 1 subscriber and that 60% of the smartphone apps have no more than 10 unique subscribers. Thus, these apps do not provide enough data for analysis, and, in this section, we explore how to decide systematically which apps can be considered popular and how we can eliminate the effect of apps with marginal usage on our analysis.

In effect, we want to identify the popular smartphone apps based on the numbers of their unique subscribers, but at the same time, we do not want to discriminate against apps with few subscribers that have a significant impact on the network, *i.e.,* generate a lot



a. traffic volume of smartphone apps

b. access time of smartphone apps

c. # unique subscribers of smartphone apps

d. correlation between metrics of smartphone apps

Figure 2: (a)-(c) CDF of volume, access time, and users, with one data point per smartphone app, aggregating users together in one week.

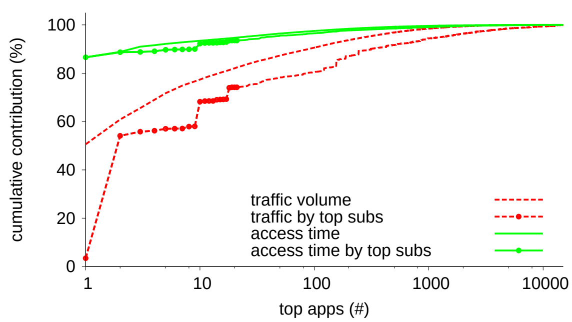


Figure 4: is the number of unique subscribers a good metric for filtering? Contributions of the top X apps to total volume and access time respectively

of traffic or access the network for long time periods. So, we have two questions to answer: (1) is the number of unique subscribers a good metric for filtering? (2) if so, what is a reasonable threshold on the number of unique subscribers?

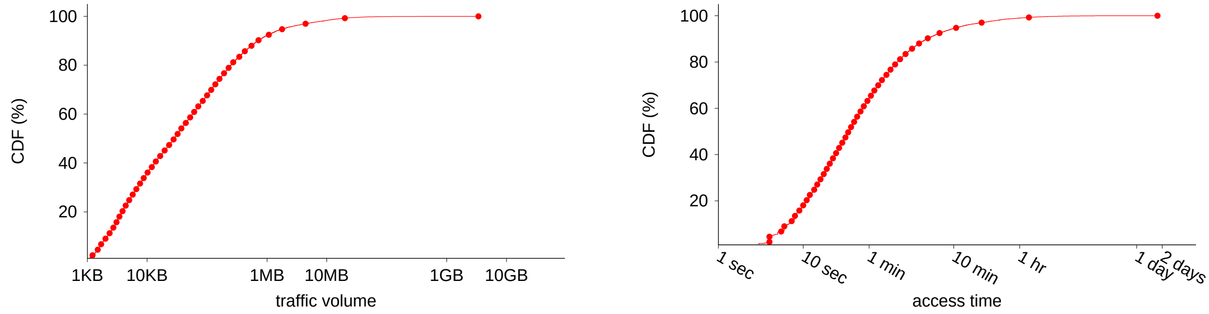
Intuitively, if the number of unique subscribers is a good metric for filtering, the top X apps based on the number of unique subscribers should contribute similar amounts of traffic volume and access times as the top X apps based on the traffic volume or access time. We first compare the contribution of the top X apps based on different metrics. Figure 4 compares the contribution of the top X apps based on the number of unique subscribers against the top X apps based on the traffic volume and based on the network access time. We can observe that the cumulative contributions of the top X apps based on access time and the top X apps based on number of unique subscribers are quite close, by comparing the "access time" and the "access time by top subs". Likewise, the contributions of the top X apps based on traffic volume and number of unique subscribers are also close, although a little difference does exist. We note that over 90% of the total volume and access time is accounted for by the top 1000 apps based on the number of unique subscribers.

Thus, Figure 4 indicates that somewhere above 1000 would be a reasonable boundary to distinguish popular apps from apps in the tail given the 90% coverage. We further explore the marginal nature of the apps ranking above 1000 in Figure 5. Figure 5(a) shows that each app in top 1000 have more than 471 unique subscribers. Figure 5(b) shows the network access time and the traffic volume per user for apps ranking in 1000 – 4000. For both traffic volume and access time, we aggregate every 100 apps into one errorbar that shows the minimum, median, and maximum of every 100 apps. The app accessing network the most consumes only 250 seconds per user in a week, and the app transferring the most data generates only 500 KB per user in a week. Because of these small traffic volumes and short access times, we do not consider these apps to be sufficiently active for our analysis.

Our discussion suggests that a natural threshold would be the top 1000 apps ranked by the number of unique subscribers. Table 2 shows the number of apps in each genre, both for the top 1000 apps as well as all the 22K apps. In the remainder of the paper, we will refer to these top 1000 apps as **popular apps**.

4.3 Spatial Patterns: Distribution of the Geographic Usage of Smartphone Apps

Next, we investigate the diversity of smartphone apps being used by subscribers in different geographic locations. Understanding the spatial usage patterns of smartphone apps suggests ways to improve user experience and performance from many aspects, such as content placement, context-aware applications, and mobile advertisement system. Taking content placement as an example, if content



a. traffic volume of smartphone apps per subscriber   b. access time of smartphone apps per subscriber

Figure 3: metrics of smartphone apps (ii) – averaged based on users.

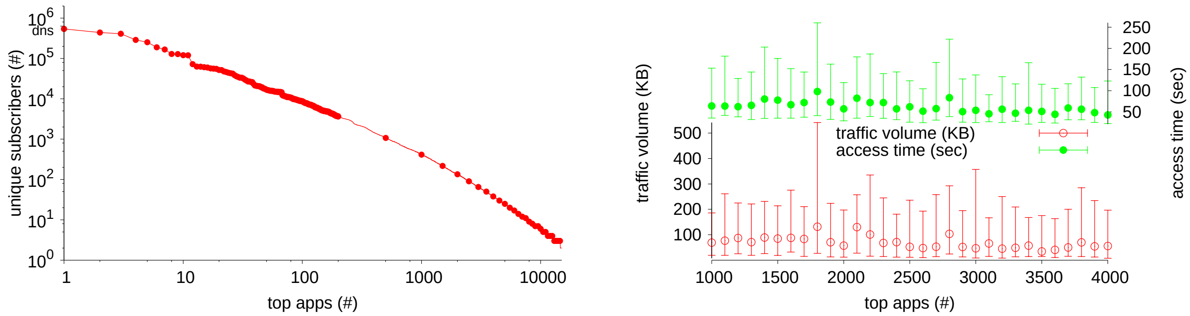


Figure 5: where is the reasonable threshold on the number of unique subscribers?

providers know that some of their apps are most used at certain locations, they may choose to place content close to those locations so that users experience better performance.

4.3.1 Local Smartphone Apps

We first examine whether any apps are **local apps**, *i.e.*, whether the majority of an app's traffic comes from a region. We perform the following analysis: for each app, we divide its traffic by (U.S.) state of the user, and compute the top 1, 3 and 5 state(s) that contribute the most traffic volume or the longest network access times. We expect that if an app's usage is truly localized, most of its traffic or access time (*e.g.*, 90%) will originate from a small number of states.

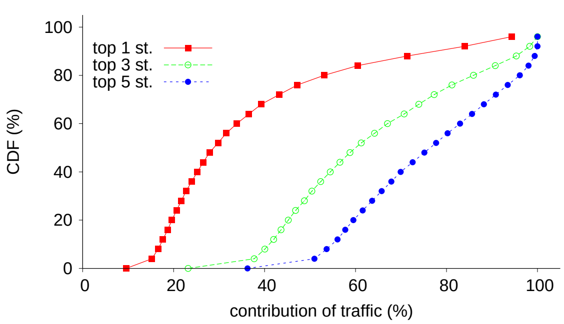


Figure 6: contribution of volume from top X states.

Figure 6 shows the CDF of the fraction of the traffic volume from top 1, 3, and 5 states for top 1000 apps that we have chosen in §4.2. According to Figure 6, 20% of the popular apps have more than 90% of their traffic volume originating from 3 states, 5.8% of the popular apps have 90% of the traffic originating from only 1 state, and 1.7% of the popular apps have all their traffic from 1 state. These 20% apps, which have more than 90% of their traffic volume originating from 3 states, account for 2% traffic in the smartphone apps category. The distribution of the contribution of access time of popular apps are very close to the one of traffic volume. Thus, we see that a significant number of the popular apps are local.

To explore what these local apps are and where they are localized, we examine in more detail the 100 most local apps based on the contribution of the top 3 states; for each of these apps, the top 3 states contribute at least 97% of their total traffic volume. Figure 7 shows the distribution of the top 3 states of the 100 most local apps; we differentiate the rank of the top 3 states for these 100 local apps as well so that we know, for example, California is the state that originated the most traffic for 19 apps, the state originated the second most traffic for 15 apps, and the state originated the third most traffic for 12 apps. As expected, California, Texas, and New York are the states with most local apps – these are the states with large populations of smartphone users. However, there are many states with much smaller populations such as Louisiana, Wyoming, and Kentucky that also have some local apps; upon further analysis, this turn out to be because content from some apps is tailored specifically for users from some regions, *e.g.*, local TV programs, news, radio, weather apps, *etc.* As an example for validation, we show the local apps for Louisiana in Table 3; we see that the six apps that have most of their traffic originating from Louisiana provide TV, news, radio and weather specifically for Louisiana residents.

We also explore the genre-wise breakdown of the local apps, as the genre of an app reflects the content and service type of a smartphone app to a great degree. Table 4 shows the distribution of the genres of the 100 most local apps in Table 7. According to Table 4,

genre	books	business	education	entertainment	finance	games	healthcare	lifestyle	medical	music	navigation	news	photography	productivity	reference	social net.	sports	travel	utilities	weather	unknown
total apps	351	418	643	1827	283	3108	368	1298	205	1296	450	1126	475	527	515	721	787	590	1079	236	5865
popular apps	7	8	13	95	13	199	23	71	4	34	23	89	17	23	26	58	33	29	49	16	170

Table 2: distribution of the genre of apps.

app	description on Google
WWLTV	New Orleans News, Breaking News, Weather ...
KATC	News Coverage at Acadiana-Lafayette, Louisiana ...
KSLANews12	News, Weather and Sports at Shreveport, Louisiana ...
KPLC 7 News	Lake Charles, Louisiana – kplctv.com ...
WBRZ	TV Channel 2 Baton Rouge, LA ...
GoWAFB	Local news, weather ... at Baton Rouge, LA ...

Table 3: local apps from Louisiana.

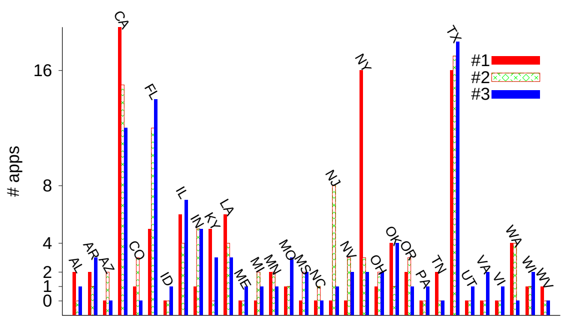


Figure 7: fraction of traffic size in top X states for each smartphone app.

these apps are mostly news, weather, and entertainment apps, likely due to the local nature of their content, *i.e.*, local weather, local news, local TV, *etc.* In music, the local apps are usually online local radio stations. The local education apps are typically created by universities, and mostly used by the local student populations.

4.3.2 National Smartphone Apps

Next, we examine the spatial patterns of smartphone app usage nation-wide. For this analysis, we remove the 100 apps identified as local in the previous analysis (Section 4.3.1), and examine the nation-wide usage of the remaining apps' traffic. We term these remaining apps as **national apps**. Our analysis explores whether certain genres are more popular (or have heavier usage) in some areas than in other areas; in general, we do not expect users to prefer using apps of a specific genre as a function of their geographic location, but our results show that this does happen under certain conditions.

For ease of reference, we term **geographic usage distribution** of a quantity X to be the the empirical probability distribution function

genre	education	entertainment	games	lifestyle	music	navigation	news	photography	productivity	sports	travel	utilities	weather	unknown
# apps	4	5	2	2	7	3	45	1	1	2	3	3	6	16

Table 4: genres of local apps.

(PDF) of X by U.S. state. First, we compute the geographic usage distribution of the unique subscribers and of the aggregate traffic volume generated by all national smartphone apps. We then use them to compute the geographic usage of traffic volume normalized by the number of unique subscribers in the state. Figure 8(a) is the PDF of the geographic usage of the aggregate traffic of all national apps together, while Figure 8(b) is the PDF of the geographic usage of the normalized traffic of all national apps. As expected, California, Texas, New York, Florida, and Illinois are the states that have the highest aggregate traffic from the national apps in Figure 8(a). However, after normalizing the volume to account for the number of subscribers, the distribution looks flatter in Figure 8(b). We perform the rest of our analysis (Figure 8(c-f)) on the normalized traffic, since it makes differences across states be easier identified.

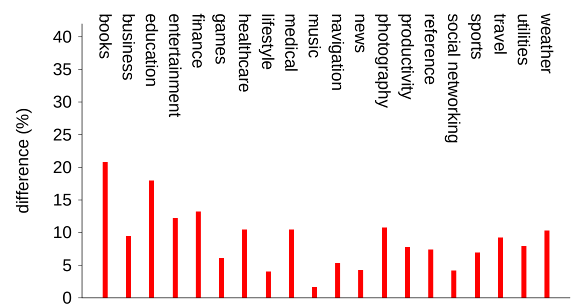
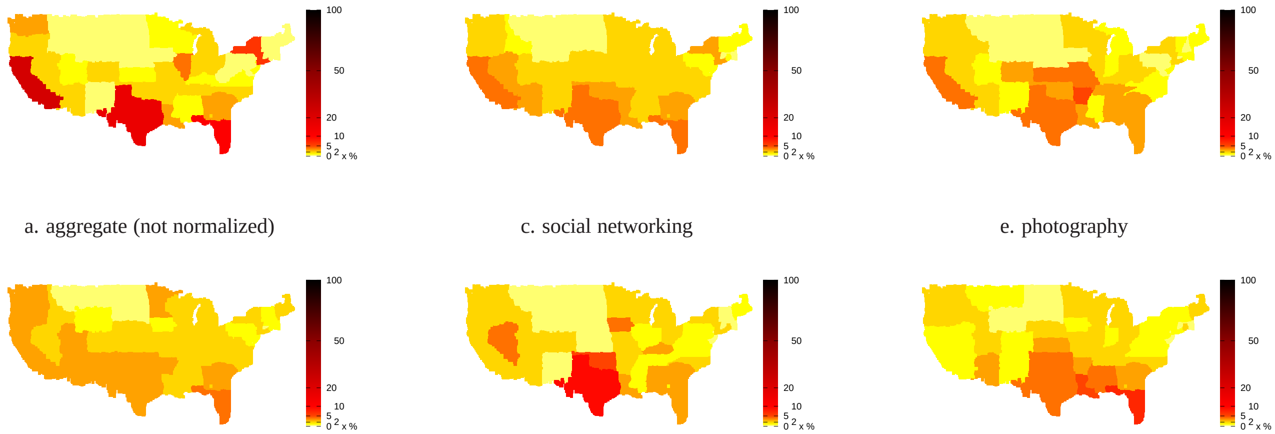


Figure 9: difference in the geographic usage of different app genres

Figure 8(c-f) illustrates the geographic usage of some representative genres (all genres are listed in Table 2). Figure 8(a) demonstrates the PDF of traffic volume from each state of all nation-wide apps. Lifestyle, music, news, and social networking genres have very similar patterns of geographic usage as the aggregate traffic. As an example, we show the social networking apps in Figure 8(c), which is is most similar to the aggregate traffic in Figure 8(b). Education apps in Figure 8(d) appear to be extremely popular in Texas; further analysis revealed that this is because some apps produced by universities (*e.g.*, TAMU) generate a significant fraction of traffic among the education apps. Likewise, Figure 8(f) shows that weather apps seem to be highly used in the south-eastern U.S. This may perhaps have happened because the time periods of our data coincide with the peak hurricane season in those areas [9, 19] – such variable and dangerous weather conditions may cause users to check weather forecasts more frequently.

Next, we measure how far the geographic usage of different genres are from the aggregate geographic usage of the apps. We use the Euclidean distance to measure the distance between a pair of geographic usage distributions. The Euclidean distance between a pair of distributions $[x_1, x_2, \cdots, x_n]$ and $[y_1, y_2, \cdots, y_n]$ is defined



a. aggregate (not normalized) c. social networking e. photography

b. aggregate normalized d. education f. weather

Figure 8: distribution of the geographic usage of apps in different genres.

as $\sqrt{\sum_{i=1}^{n}(x_i - y_i)^2}$. Figure 9 shows the distance between each genre's geographic usage distribution and the distribution of aggregate national apps. We note that some genres, such as books and education, are disproportionately used in some states, while others, such as social networking apps generate traffic more proportionate to the total traffic generated by that state.

The distribution of geographic usage of different apps *within* the same genre may also differ. Figure 10 shows the geographic usage distribution of a number of smartphone apps in the news genre. For this analysis, we select the apps of a few newspapers that are well-known across the entire U.S., and cover news relating to any part of the world. The location of each news app in Figure 10 only reflects where the newspaper headquarter is located. However, some of them have a location indicated in their names (marked with [2] in Figure 10). Although all these apps are used nation-wide, we note that apps whose names have a location indicated seem to be disproportionally preferred at those respective locations. In addition, the Washington D.C. news app seems to be highly preferred in Washington state as well, thus suggesting users look for apps that appear to be local to their region.

4.3.3 *Travel Area of Smartphone Apps*

Our final analysis of the spatial patterns of app usage examines whether individual users use some apps across larger geographic areas than others, *e.g.*, whether Internet radio apps, which users may listen to during their commute, are used across a larger area than books. Such an analysis can help understand what kinds of apps need to be more robust to variations in network quality.

For this analysis, we define the **travel-area** as a smartphone's geographic coverage per individual subscriber over short time, *e.g.*, 6 hours. We use the number of sectors to estimate the geographic coverage, since we do not have access to the device's exact locations. As noted in Section 3, the number of sectors observed in our data set is an underestimate of the actual number of sectors that the device passes through, but our results still give an idea of the relative travel-area of different apps. Specifically, for each app, we compute the average number of unique sectors used by active subscribers in each 6 hour time interval.

Figure 11 shows the distribution of the average travel-area of the top 1000 popular apps. It shows that 10% apps access the network

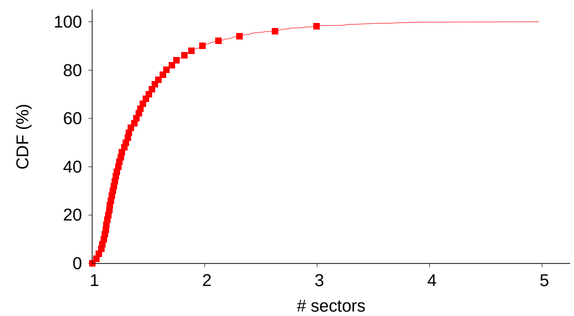


Figure 11: travel-area of apps.

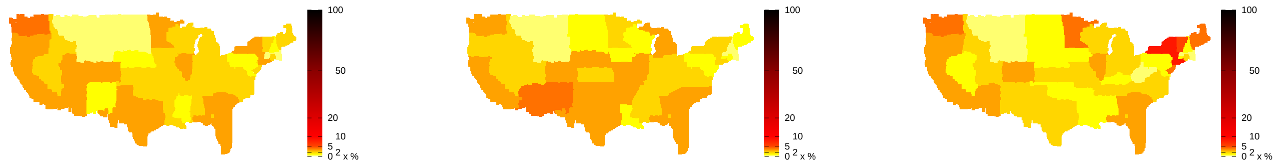
genre	books	business	finance	games	healthcare	lifestyle	music	news	productivity	reference	social net.	sports	travel	unknown
# apps	2	1	4	19	3	5	6	1	3	1	31	2	1	18

Table 5: genres of high travel-area smartphone apps.

from more than two sectors. Thus, our results indicate that a significant fraction of the apps are used when users move around, creating another issue for content caching and delivery techniques. Base stations in future cellular network designs (*e.g.*, LTE) have been considered potential locations for content caching and optimization (since they would be the first IP hop), so significant amount user movement could make it more difficult to cache content appropriately. Table 5 shows that the majority of these apps are games or social networking apps, but there are also a few music and news apps.

4.4 User Patterns: Impact of User Interests on Smartphone Apps Usage

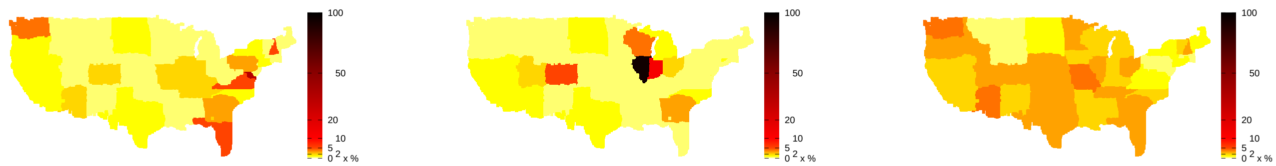
The needs and interests of individual users are the primary factors that inform their usage of apps. Because of user interests, the usage of different apps tends to be correlated. In this section, we analyze the extent to which user apps are correlated. Our analysis has



a. aggregate news apps

b. headquartered in ATLANTA[1]

c. headquartered in NEW YORK[2]



d. headquartered in WASHINGTON, D.C.[2]

e. headquartered in CHICAGO[2]

f. headquartered in VIRGINIA[1]

[1]: names of these apps do not have any location indicated.
[2]: names of these apps include locations.

Figure 10: distribution of the geographic usage of apps in the same genre.

many motivations: knowing what sets of apps are correlated would be helpful for both app developers as well as OS vendors, as they can factor this correlation into their designs and help the apps work better with each other. From a network perspective, such knowledge could help optimize performance or user experience for a set of apps as a bundle, and may also enhance troubleshooting. In addition, app markets can leverage this information for recommending new apps to users.

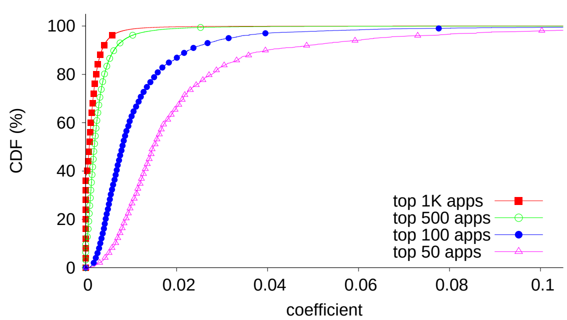


Figure 12: distribution of the Jaccard Similarity Coefficient of the popular apps.

We use the **Jaccard Similarity Coefficient** to quantify the overlap between a pair of apps a and b: we count the number of unique subscribers who have used both a and b, *i.e.*, $joint(a,b)$, and the number of unique subscribers who have used either a or b, *i.e.*, $union(a,b)$. We can obtain $\frac{joint(a,b)}{union(a,b)}$ for all pairs of apps in the popular 1000 apps.

Figure 12 shows the distribution of the Jaccard Similarity Coefficient between the top 1000, 500, 100, and 50 apps. We observe that there is a small fraction of app pairs that have a very high Jaccard Similarity Coefficient. For example, consider a pair of apps a,b whose $\frac{joint(a,b)}{union(a,b)} = 0.05$, and assume that a and b have 2000 unique subscribers together (we know from Figure 5(a) that each of the top 50 most popular apps have over 2000 unique subscribers). Then at this value of the Jaccard Similarity Coefficient, a and b share 100 unique subscribers. Given that there are more than 600,000 unique subscribers in our data set, the overlap of 100 subscribers is unlikely to be due to random chance, indicating that users of app a have a tendency to also use app b. We also note that as we increase the number of popular apps from 50-1000, there is a smaller fraction of app pairs that have a significant overlap in subscribers. This is expected, since an app is more likely to have subscribers overlap with other apps when gets used by more and more subscribers.

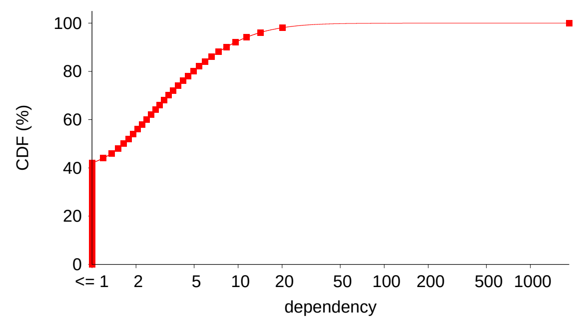


Figure 13: distribution of the dependency between popular apps.

.

Next, we analyze how likely it is for a pair of apps to have a substantial overlap in their users. Our analysis compares the empirical probabilities of a subscriber using each app individually to the empirical probability of a subscriber using both apps together. More precisely, let a,b denote apps, and $Pr[a], Pr[b]$ denote the empirical probabilities of a subscriber using app a,b respectively. Let $Pr[ab]$ denote the empirical probability of a subscriber using both apps a and b. If the subscribers for each app are selected at random from the total population, then we would expect that $Pr[ab]$ to be somewhat close to the product $Pr[a]Pr[b]$. Figure 13 shows the distribution of the ratio $\frac{Pr[ab]}{Pr[a][b]}$ (we term this quantity the **dependency ratio** for ease of reference). It shows that nearly 10% of the app

	books	business	education	entertainment	finance	games	healthcare	lifestyle	medical	music	navigation	news	photography	productivity	reference	social net.	sports	travel	utilities	weather
books	0	0	0	2	0	3	0	1	0	0	0	1	0	0	0	0	0	1	0	0
business	0	0	0	0	0	0	0	4	0	0	0	2	0	0	0	0	0	0	1	0
education	0	0	5	2	0	2	0	0	0	0	2	3	0	0	0	1	0	0	0	0
entertainment	2	0	2	26	0	26	1	16	0	5	1	16	4	3	3	8	4	4	7	0
finance	0	0	0	0	2	0	1	0	0	0	0	4	0	0	0	0	0	1	1	0
games	3	0	2	26	0	110	2	19	1	5	1	3	2	1	13	5	8	12	1	1
healthcare	0	0	0	1	1	2	1	2	1	0	1	8	1	0	0	1	0	0	0	0
lifestyle	1	4	0	16	0	19	2	30	0	5	0	12	0	1	7	6	4	10	6	0
medical	0	0	0	0	0	1	1	0	0	0	0	1	0	0	0	0	0	0	0	0
music	0	0	0	5	0	5	0	5	0	6	0	5	0	1	1	11	3	0	3	0
navigation	0	0	2	1	0	1	1	0	0	0	2	1	0	1	0	1	0	4	0	1
news	1	2	3	16	4	3	8	12	1	5	1	77	1	3	2	12	8	7	4	4
photography	0	0	0	4	0	2	1	0	0	0	0	1	1	1	0	0	0	0	1	0
productivity	0	0	0	3	0	1	0	1	0	1	1	3	1	0	0	9	1	0	0	0
reference	0	0	0	3	0	13	0	7	0	1	0	2	0	0	1	0	1	2	0	1
social networking	0	0	1	8	0	5	1	6	0	11	1	12	0	9	0	32	4	0	3	0
sports	0	0	0	4	0	8	0	4	0	3	0	8	0	1	1	4	13	1	0	1
travel	1	0	0	4	1	12	0	10	0	0	4	7	0	0	2	0	1	9	2	0
utilities	0	1	0	7	1	1	0	6	0	3	0	4	1	0	0	3	0	2	2	0
weather	0	0	0	0	0	1	0	0	0	0	1	4	0	0	1	0	1	0	0	1

Table 6: measuring the dependency between genres: apps with high dependency-ratio.

pairs have a dependency ratio that exceeds 10, and 254 pairs have a ratio exceeding 100.

Table 6 shows the frequency distribution of the genres of these 254 pairs (*i.e.,* pairs with dependency-ratio exceeding 100). We can make two immediate observations from this table. First, apps in the same genre are much more likely have correlated usage. For example, 110 pairs of two games apps that have high dependency-ratio, but games apps are part of a only 230 pairs in total. Second, apps in similar genres are more likely to have high dependence-ratio, *e.g.,* entertainment and games, news and entertainment, entertainment and social networking, travel and navigation, weather and news, social networking and news, *etc.*

There are many reasons that pairs of apps have highly correlated usage. First, many different apps often provide the same type of content in different forms *e.g.,* there may be multiple local news or Internet radio stations targeting the same location, and users often are interested in trying them all out. Or, there may multiple apps that allow users to access the same social networking sites with different user interfaces. A second reason may be that a pair of apps serve similar purpose, but neither may provide complete service on its own *e.g.,* users may have accounts with multiple banks, and need to use each bank's specific app in order to keep track of all their accounts. Yet another reason may be that different apps target similar user interests, and users may try them all out to identify their favorites, *e.g.,* crossword puzzle apps or sudoku apps.

4.5 Temporal Patterns: Distribution of the Traffic over time of Smartphone Apps

Understanding the diurnal patterns of apps is important for several reasons. For example, differences in when certain apps are used can help inform cloud providers on how to best multiplex resources and operators on what to optimize the network for at different times. In this analysis, we compare the traffic volumes and access times consumed by smartphone apps at different hours of the day, both in aggregate as well as for different genres. Our results show that there are diurnal patterns of app usage both in aggregate, as well as by genre, but that the patterns of different genres are noticeably different.

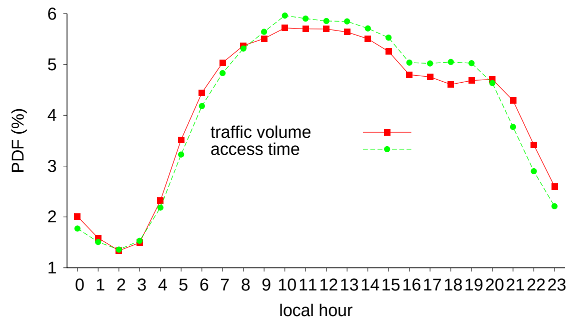
category	# apps	description
entertainment	20	small games, video channels, *etc.*
radio	28	music radio channels, news radio channels, *etc.*
healthcare	12	sleep aid utilities, *etc.*
books	6	bible, references, *etc.*

Table 7: description of late night apps.

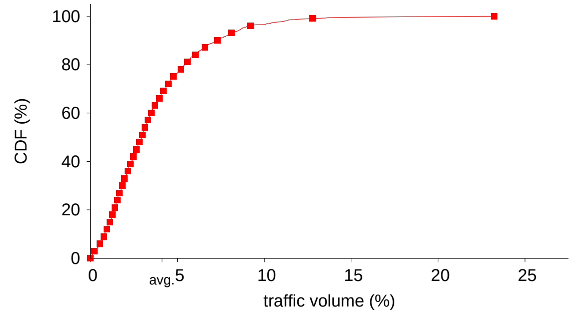
We first investigate the diurnal patterns by aggregating all the popular apps together. For this analysis, we map each flow to the local time of the device's geo-location (based on the sector where the device is connected to the cellular network). Figure 14(a) shows clear diurnal patterns of traffic volume and network access time. Around 1AM – 2AM, the total traffic volume and access time are at their minimum; they start increasing around 4AM, reach the peak usage around noon, start decreasing after 3PM and drop dramatically after 8PM.

In general, apps have more activity during the daytime than at night. However, this may not apply to every popular app. Figure 14(b) shows the distribution of the traffic contribution during late night for popular apps. According to Figure 14(a), in terms of both traffic volume or access time, the time period 1:00 AM – 3:59 AM contributes 4.2% traffic. Even if an app generates uniform traffic every hour of the day, it should generate 12.5% traffic from 1:00 AM to 3:59AM. So, Figure 14(b) indicates that there are some apps that are quite active late at night. We manually investigate these top 66 *late night* apps according to Figure 14(b) that contribute more than 12.5% traffic late at night. Table 7 summarizes the results. It appears that several entertainment and radio apps are used more frequently than expected at night.

Finally, we analyze diurnal patterns across different genres; we expect that different genres of apps to have different usage patterns, since they appeal to different interests. As we did in earlier analysis, we aggregate together the popular apps in the same genre, and compute the distribution of traffic volumes by genre at hourly intervals (again, using the local time of the flow). Figure 15 shows the normalized traffic volume across the day; it clearly shows how different genres do have very different diurnal patterns. In particu-



a. diurnal patterns of traffic volume & access time



b. traffic contribution by late night (1:00 – 3:59 AM)

Figure 14: diurnal patterns.

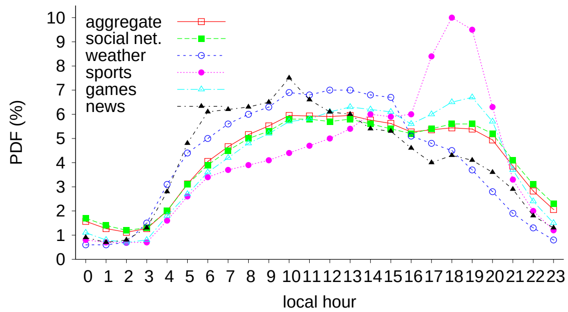


Figure 15: diurnal patterns across different genres.

lar, we see that social network apps have almost exactly the same pattern as the aggregate, but weather and news apps are most frequently used at early morning. Sports apps, on the other hand, peak in the early evening, perhaps because users may watch matches or check scores frequently during those hours. Games apps also peak after standard work hours as we would expect, since that is probably the typical recreation time for most subscribers.

4.6 Device Patterns: Differences Across Platforms

Finally, we compare smartphone app usage across different kinds of devices. We expect that faster devices allow for longer sessions, faster downloads, and more interactivity, thus enhancing the end-user experience. Power users, who use their devices more, may also gravitate to newer and faster devices. We focus on three different devices from the same device family, as we expect device operating system to also affect overall usage patterns. We compare three devices in the same device family but of different generations – we term these device 2, device 3, and device 4 in Figure 1. Device 2 is a HSDPA category 6 device (capable of 3.6Mbps downlink rate), and device 3 and device 4 are in HSDPA category 8 (capable of 7.2Mbps downlink) [24]. Device 2 and device 3 are not HSUPA enabled while device 4 is HSUPA category 6 (capable of 5.76Mbps uplink) [25].

For this analysis, we use slightly different metrics than we have used in the rest of the paper, since our goal is to measure how long a user interacts with the device, and compare these measurements across different devices. For this, we define **individual access time** and the **individual traffic volume** to be the network access time and the traffic volume per flow respectively. We use these metrics for our analysis as we expect the individual access time to provide a measure of how long a user spends with an app, and the individual traffic volume to reflect how much data is transferred each time a user interacts with an app. Obviously not every flow will be larger or longer, but we expect that a device that allows for better interaction will have on average more large flows and longer flows.

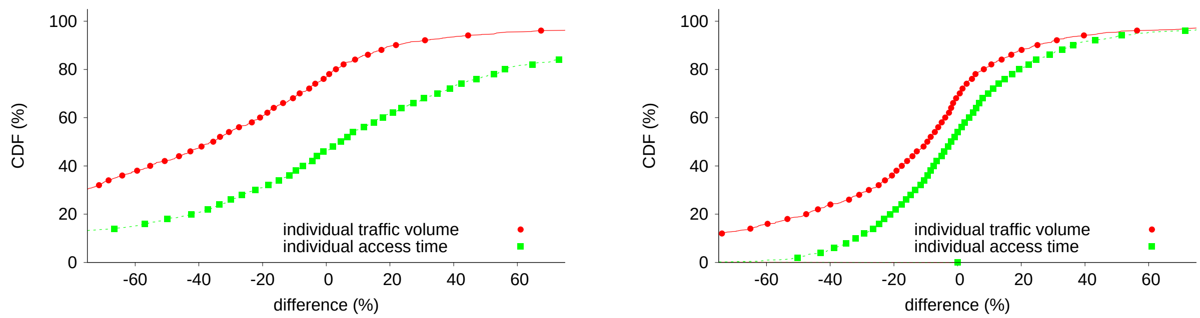
In Figure 16, we compare the individual traffic volume and individual access time between device 2 and device 4, and between device 3 and device 4. For our analysis, we aggregate all the popular apps to first compute directly the individual access time and individual traffic volume for these three platforms, and then compute the relative differences by comparing device 2 against device 4, and device 3 against device 4. According to Figure 16(a), the individual access time for device 2 and device 4 are very close, *i.e.*, the median relative difference is 0. However, individual traffic volume for device 2 is much smaller. The median difference of the individual traffic volume is -30%. Such a big difference indicates that the user experience is substantially different between these two device categories; users of device 4 consume much more data, typically through video. Figure 16(b) shows the comparison between device 3 and device 4; these two device categories are much closer than device 2 and device 4. There may be two explanations. First, a faster device tends to give users better overload experience and encourages them to download more content from the network. Second, power user are more likely upgrade to the latest smartphone, while users not as active may be more likely to keep using their older devices.

5. IMPLICATIONS

In previous sections, we investigate the usage patterns of smartphone apps from spatial, temporal, user, and device perspectives. We believe that our previous observations have important implications for the smartphone community. In this section, we discuss these implications following our previous observations.

5.1 Content Providers

In the analysis of spatial patterns of smartphone usages, we observe a considerable number of local apps (20%) which contribute 2% of the traffic volume in the smartphone apps category. The content provided by these local apps are very deterministic, *e.g.*, news apps, regional radio online services, weather forecast apps, *etc.* Given both the customers and locations for these apps are very closely clustered, content placement and delivery can be further optimized accordingly. It is therefore beneficial to place the content close to GGSNs in the cellular networks [28] for cellular users and place the content close to the geographic location of WiFi users. Besides local apps, for national apps, the distribution of geographic



a. device 2 vs. device 4   b. device 3 vs. device 4

Figure 16: impact of devices used.

coverage is still very dependent on the genre (*e.g.,* weather apps are highly used in the south-eastern U.S.), even the app's name (*e.g.,* the news app headquartered in CHICAGO), *etc.* Therefore, content placement according to the geographic coverage is advisable for both national and local apps.

5.2 Context-Aware Applications

Despite very diverse usage usage patterns across different smartphone apps, they still have some common traits. According to our observations, first, apps in the same genre share similar geographic coverage. Second, some apps share a large set of common users due to the similarity of content and interests, *e.g.,* social networking apps strongly correlate with entertainment apps, music apps, news apps, *etc.* Third, some apps share similar diurnal patterns due to content characteristics, *e.g.,* the peak hours of news apps and weather apps come at early morning.

Context-aware applications can take advantage of the existing similarity/correlation across smartphone apps. Take smartphone apps recommendation systems as an example. Unlike normal PC users, smartphone users depend on apps far more than browsers. Since a smartphone apps recommendation system is the first approach for users to explore various smartphone apps that meet their interests, these systems can be quite important. As the bridge between app marketplace and app customers, if apps recommendation systems can learn user interests and dependency across apps, they can identify more appropriate apps for users, *e.g.,* suggesting gaming fans more entertainment apps and social networking apps.

Another example of context-aware application is advertisement systems, which upon learning user's interests in apps, can deliver more relevant ads to users. Camera or camcorder advertisements may target more smartphone users that use more entertainment and game apps because photography apps are more correlated with entertainment and game apps.

5.3 Network Providers

Besides content providers, cellular network providers also play an important role in content delivery and customization. By understanding the access patterns of smartphone apps, network providers can benefit in allocating radio resource, setting caching policy, compression policy, *etc.*

If a large number of smartphone apps are targeted, their traffic volume and access time roughly have linear correlation with their number of unique subscribers. Accordingly, cellular providers can estimate and allocate radio resources.

We observe that the several few top apps contribute the majority traffic. For example, the app with the largest traffic volume is accountable for 50% of the total traffic volume of the smartphone apps category, and the app with the longest network access time takes 86% of the total network access time of the smartphone apps category. Understanding the usage patterns of these apps, network providers may do certain optimizations case by case.

The temporal patterns of smartphone apps help network providers allocate radio resource. For example, the access time per IP flow helps network providers decide the timers in state promotion [20].

We observe that some smartphone apps have large usage radius, *i.e.,* users of certain social networking apps and games apps are more likely to move around across several base stations. In future, LTE networks will push the first IP hop forward to base stations, which increases the flexibility of content placement and optimization. However, if users frequently move around, the corresponding mobility may increase the complexity to decide where to cache content and what content to cache.

5.4 OS Vendors and Apps Designers

Since smartphones have limited resources, the OS is accountable for resource management, *e.g.,* the push notification on iOS, Android, Windows Phone. Understanding the access patterns of apps on device, OS can add some flexibility to apps and optimize the resource usage. For example, if a user frequently resorts to a certain sleep aid app, then OS may allocate less resource to those apps that may interrupt the user's sleep.

Certain genres of smartphone apps have different characteristics, which may be taken advantage by apps designers. We observe that news and weather apps have distinctive diurnal patterns. Since the content of these apps usually are very time dependent and content fetching time is very predictable, apps designer can implement some prefetching mechanism to reduce the latency perceived by users. Similarly, the content of social networking apps can be prefetched before dinner time.

6. CONCLUDING REMARKS

In this study, we comprehensively investigated the diverse usage patterns of smartphone apps via network measurements from a national level tier-1 cellular network provider in the U.S. Our study is the first attempt in addressing the lack of how, where and when smartphone apps are used at the scale of the entire U.S.

We observed that a considerable fraction of popular apps (20%) are local because their content are expected to serve local users such as news and radio apps. This suggests that there is significant possibility for content optimization in LTE and WiFi access networks where the flexibility of placing content is high.

We also found out that there are similarities across apps in terms of geographic coverage, diurnal usage patterns, *etc.* Certain apps have a high likelihood of co-occurrence – that is, (i) when a user uses one app, he is also likely to use another one; or (ii) users use alternatives for the same type of interests, *e.g.*, multiple news apps, bank apps. These observations suggest that some apps should be treated as a "bundle" when trying to optimize for their user experience. There may be opportunities for integrating these apps together.

Diurnal patterns of smartphone apps can be remarkably different. For instance, news apps are much more frequently used in the early morning while sports apps are more frequently used in the evening. These findings suggest that content providers (*e.g.*, hosted on cloud) can leverage distinct usage patterns in classes of apps to maximize the utilization of their resources.

Many social networking and games apps are more frequently used when users are moving around. Mobility affects connectivity and performance, so bandwidth sensitive content that are mobile may need to consider techniques to compensate for bandwidth variability.

We believe that our findings on the diverse usage patterns of smartphone apps in spatial, temporal, user, device dimensions will motivate future work in the mobile community.

7. REFERENCES

[1] Apple. Apple's App Store Downloads Top 10 Billion. `http://www.apple.com/pr/library/2011/01/22appstore.html`.

[2] A. Balasubramanian, R. Mahajan, and A. Venkataramani. Augmenting Mobile 3G Using WiFi: Measurement, System Design, and Implementation. In *Proc. ACM MOBISYS*, 2010.

[3] X. Bao and R. Roy Choudhury. MoVi: mobile phone based video highlights via collaborative sensing. In *Proc. ACM MOBISYS*, 2010.

[4] I. Constandache, X. Bao, M. Azizyan, and R. R. Choudhury. Did you see Bob?: human localization using mobile phones. In *Proc. ACM MOBICOM*, 2010.

[5] E. Cuervo, A. Balasubramanian, D. ki Cho, A. Wolman, S. Saroiu, R. Ch, and P. Bahl. MAUI: Making smartphones last longer with code offload. In *Proc. ACM MOBISYS*, 2010.

[6] W. Enck, P. Gilbert, B.-G. Chun, L. P. Cox, J. Jung, P. McDaniel, and A. N. Sheth. TaintDroid: an information-flow tracking system for realtime privacy monitoring on smartphones. In *USENIX Symposium on Operating Systems Design and Implementation (OSDI)*, 2010.

[7] H. Falaki, D. Lymberopoulos, R. Mahajan, R. Govindan, S. Kandula, and D. Estrin. Diversity in Smartphone Usage. In *Proc. ACM MOBISYS*, 2010.

[8] H. Falaki, D. Lymberopoulos, R. Mahajan, S. Kandula, and D. Estrin. A first look at traffic on smartphones. In *Proc. ACM SIGCOMM IMC*, 2010.

[9] Federal Emergency Management Agency. Tornado Activity in the United States. `http://www.fema.gov/plan/prevent/saferoom/tsfs02_torn_activity.shtm`.

[10] M. Ficek, T. Pop, P. Vláčil, K. Dufková, L. Kencl, and M. Tomek. Performance study of active tracking in a cellular network using a modular signaling platform. In *Proc. ACM MOBISYS*, 2010.

[11] A. Gember, A. Anand, and A. Akella. A Comparative Study of Handheld and Non-Handheld Traffic in Campus WiFi Networks. In *Proc. International Conference on Passive and Active Network Measurement (PAM)*, 2011.

[12] Google. Eric Schmidt at Mobile World Congress 2011. `http://www.youtube.com/watch?v=ClkQA2Lb_iE&feature=related`.

[13] B. D. Higgins, A. Reda, T. Alperovich, J. Flinn, T. J. Giuli, B. Noble, and D. Watson. Intentional networking: opportunistic exploitation of mobile network diversity. In *Proc. ACM MOBICOM*, 2010.

[14] J. Huang, Q. Xu, B. Tiwana, Z. M. Mao, M. Zhang, and P. Bahl. Anatomizing Application Performance Differences on Smartphones. In *Proc. ACM MOBISYS*, 2010.

[15] R. Keralapura, A. Nucci, Z.-L. Zhang, and L. Gao. Profiling users in a 3g network using hourglass co-clustering. In *Proc. ACM MOBICOM*, 2010.

[16] K. A. Li, T. Y. Sohn, S. Huang, and W. G. Griswold. Peopletones: a system for the detection and notification of buddy proximity on mobile phones. In *Proc. ACM MOBISYS*, 2008.

[17] Y. Liu, A. Rahmati, Y. Huang, H. Jang, L. Zhong, Y. Zhang, and S. Zhang. xShare: supporting impromptu sharing of mobile phones. In *Proc. ACM MOBISYS*, 2009.

[18] G. Maier, F. Schneider, and A. Feldmann. A first look at mobile hand-held device traffic. In *Proc. International Conference on Passive and Active Network Measurement (PAM)*, 2010.

[19] National Hurricane Center. National Hurricane Center. `http://www.nhc.noaa.gov/pdf/TAFB_Trifold.pdf`.

[20] F. Qian, Z. Wang, A. Gerber, Z. M. Mao, S. Sen, and O. Spatscheck. Characterizing radio resource allocation for 3G networks. In *Proc. ACM SIGCOMM IMC*, 2010.

[21] F. Qian, Z. Wang, A. Gerber, Z. M. Mao, S. Sen, and O. Spatscheck. Profiling Resource Usage for Mobile Applications: a Cross-layer Approach. In *Proc. ACM MOBISYS*, 2010.

[22] I. Trestian, S. Ranjan, A. Kuzmanovic, and A. Nucci. Measuring serendipity: connecting people, locations and interests in a mobile 3G network. In *Proc. ACM SIGCOMM IMC*, 2009.

[23] I. Trestian, S. Ranjan, A. Kuzmanovic, and A. Nucci. Taming User-Generated Content in Mobile. Networks via Drop Zones. In *Proc. IEEE INFOCOM*, 2009.

[24] Wikipedia. High-Speed Downlink Packet Access. `http://en.wikipedia.org/wiki/High-Speed_Downlink_Packet_Access`.

[25] Wikipedia. High-Speed Uplink Packet Access. `http://en.wikipedia.org/wiki/High-Speed_Uplink_Packet_Access`.

[26] W. Woerndl, C. Schueller, and R. Wojtech. A Hybrid Recommender System for Context-aware Recommendations of Mobile Applications. In *Proc. IEEE ICDE*, 2007.

[27] Q. Xu, A. Gerber, Z. M. Mao, and J. Pang. Acculoc: Practical localization of peformance measurement in 3g networks. In *Proc. ACM MOBISYS*, 2011.

[28] Q. Xu, J. Huang, Z. Wang, F. Qian, A. Gerber, and Z. M. Mao. Cellular data network infrastructure characterization and implication on mobile content placement. In *Proc. ACM SIGMETRICS*, 2011.

[29] B. Yan and G. Chen. AppJoy: Personalized Mobile Application Discovery. In *Proc. ACM MOBISYS*, 2011.

[30] L. Zhang, B. Tiwana, Z. Qian, Z. Wang, R. P. Dick, Z. M. Mao, and L. Yang. Accurate online power estimation and automatic battery behavior based power model generation for smartphones. In *Proc. IEEE/ACM/IFIP CODES+ISSS*, 2010.