A Panoramic View of 3G Data/Control-Plane Traffic: Mobile Device Perspective

Xiuqiang He, Patrick Lee, Lujia Pan, Cheng He, John Lui

A Panoramic View of 3G Data/Control-Plane Traffic: Mobile Device Perspective

Xiuqiang He[†], Patrick P. C. Lee[‡], Lujia Pan[†], Cheng He[†], and John C. S. Lui[‡]

[†]Noah's Ark Lab, Huawei Research, China
[‡]Dept of Computer Science & Engineering, The Chinese University of Hong Kong, Hong Kong
{hexiuqiang,panlujia,hecheng}@huawei.com, {pclee,cslui}@cse.cuhk.edu.hk

Abstract. Users can access the Internet via 3G/4G cellular data networks using various types of user devices (e.g., smartphones, tablets, datacards). We conduct a detailed measurement study on the impact of different device types on the data/control-plane performance of a commercial, city-wide 3G cellular data network in China. We present a methodology that correlates different data/control-plane datasets collected at different points in the network core, and identify more than 60K devices of different types per day on average. For the devices we identify, we investigate how their commonly used Internet applications and internal heartbeat mechanisms lead to distinct data/control-plane behaviors. For example, we observe that datacard devices contribute a large volume of IP traffic in the data plane, while smartphones introduce significant resource overhead in the signaling control plane. Our measurement study provides insights for network operators to strategize pricing and resource allocation for the data/control planes of their cellular data networks with regard to the market penetrations of various device types.

Keywords: mobile device traffic; data/control-plane; 3G networks; measurement

1 Introduction

With the wide deployment of 3G/4G wide-area cellular data networks, we have witnessed the tremendous growth of mobile Internet access worldwide. There are many types of devices that enable mobile Internet access, such as smartphones, tablet computers, or datacards attached to laptops/PCs. There have been studies (e.g., [1]) in forecasting the explosive growth of mobile data traffic from these devices. Compared to traditional wireline networks, cellular networks not only have relatively limited data-plane link capacity, but also have higher control-plane overhead that increases the loading of the network core [16] and power consumption of mobile devices [19]. From the perspectives of network operators, it is necessary to understand how the traffic patterns of *different* device types influence the performance of cellular data networks in both data and control planes. Understanding this influence can shed light on how to provision network resources, and if necessary, how to provide differentiated pricing or priority services across different device types or applications.

In this paper, we conduct an in-depth measurement study on the impact of different device types on the data/control-plane performance of a commercial 3G UMTS network

deployed in a metropolitan city in China. We analyze several large-scale datasets, including IP packets with complete payload, UMTS-compliant signaling messages, and radio resource control (RRC) logs. We identify more than 60K devices of different types per day on average, as well as different IP applications. The extensive scales of our datasets allow us to conduct comprehensive analysis. In summary, the contributions of this paper are two-fold:

– We first present a methodology that correlates all these datasets, so that we can study the interactions of data and control traffic of *each* device type. We also characterize the IP applications commonly used by different device types. Note that the formats of our datasets follow the standard 3G specifications. Thus, our methodology is also applicable for general 3G networks, and will be useful for network operators of interest to carry out network planning by collecting the same datasets within their own managed networks.

– We conduct an extensive measurement study mainly based on a 24-hour span of data/control-plane traces. We make the following observations. Even when there are only 7.5% of datacard devices, they contribute disproportionately about 46% of IP data traffic; for smartphone or tablet devices, they contribute less data traffic than datacard devices, but consume significantly high radio resource usage due to the frequent RRC connection setups/releases. Furthermore, we identify that iPhone/iPad devices have some internal heartbeat mechanisms that trigger a substantial number of RRC connection setups/releases, even though they generate minimal data traffic. Our study shows the importance of how mobile devices interact with the control plane and what kind of behavior can be expected.

The remainder of the paper proceeds as follows. Section 2 reviews related work on 3G traffic measurement. Section 3 describes how we correlate different datasets collected at different points of a 3G network. Section 4 presents our findings from the datasets. Finally, Section 5 summarizes the lessons learned from our study.

2   Related Work

There have been measurement studies focusing on in-network data traffic collected inside 3G networks. Kilpi and Lassila [14] analyze the round-trip times of TCP flow data collected at a GPRS/UMTS network. Ridoux *et al.* [22] collect data from a CDMA2000 network, and compare the similarities and differences with wireline data traffic. The DARWIN+ group [3] collects IP packets at a GPRS/UMTS network, and analyzes various issues such as TCP performance [21] and traffic anomalies [2]. The above studies mainly analyze the data-plane performance of 3G networks.

Some recent studies analyze the control-plane performance of 3G networks. Lee *et al.* [16] study the signaling overhead from the security perspective. They show that if an attacker generates data traffic following specific patterns, then a small amount of data traffic can sufficiently trigger heavy signaling overhead in the network core. On the other hand, they only analyze synthetic traces. Qian *et al.* [19] infer and analyze the RRC state transitions of user sessions using data-plane traces, and mainly focus on power consumptions of mobiles. Paul *et al.* [18] collect data packet headers and various signaling messages in a national 3G network, and analyze their temporal and spatial variations. Our work differs from [18] in two aspects. First, our work focuses

on distinguishing different device types and analyzing each of their data/control-plane performance. Also, our datasets contain full IP payload information that enables us to accurately identify more user applications and infer their impact to the network.

Several studies analyze the data traffic behavior of different types of handheld devices. Maier *et al.* [17] analyze the mobile data traffic of handheld devices inside a residential DSL network, and Gember *et al.* [8] compare handheld and non-handheld devices in a campus WiFi network. In the context of 3G cellular data networks, Falaki *et al.* [6, 7] study the smartphone traffic and identify the differences of user behavior based on the traces collected from a number of individual devices. Huang *et al.* [10] propose a tool called 3GTest, which runs on thousands of smartphones over a wide geographic coverage. Each 3GTest-enabled device generates probe traffic to measure the network performance. Instead of collecting data from smartphone devices, Xu *et al.* [23] analyze the IP data traffic collected inside a tier-1 network core, and study the usage behavior of a variety of smartphone applications from the operator's perspective. Note that the above studies mainly focus on the data-plane performance. Our work complements the prior studies by also analyzing the control-plane performance. Specifically, we play the role of a network operator and look into the interactions of both data-plane and control-plane traffic traces collected within a 3G network.

3   Datasets and Preprocessing

We first describe several datasets that are used for our measurement study of a 3G UMTS network. We then explain how we correlate these datasets to obtain the information needed for our analysis. Finally, we address the limitations of our datasets.

3.1   Datasets

Data/control traffic. We capture traffic in both data and control planes from a commercial, city-wide 3G UMTS network in China. Fig. 1 shows the simplified topology of the 3G network we consider. The data path between a *user equipment (UE)* and the Internet traverses a *base station (NodeB)*, a *Radio Network Controller (RNC)*, and the *core network (CN)*, all of which are defined in the 3GPP UMTS standard. Here, we focus on packet-switched domain of the CN, which comprises the *Serving GPRS Support Node (SGSN)* and the *Gateway GPRS Support Node (GGSN)*. In a high level, the CN can be viewed as the central point that relays all UE-Internet data traffic as well as UE-UE data traffic. A 3G data network typically has a hierarchical architecture, i.e., multiple UEs communicate with a NodeB, multiple NodeBs communicate with an RNC, and multiple RNCs communicate with the CN.

We capture traffic in the integrated interface (a.k.a. *Iu-PS*) that connects 16 RNCs and the CN. We collect a 7-day span of traces from November 25, 2010 to December 1, 2010. The traces have a total size of around 13TB with about 27.6 billion packets and 383 million 5-tuple flows, including IP data packets with headers and full payloads and the *Radio Access Network Application Part (RANAP)* signaling messages. RANAP is the UMTS signaling protocol for the radio connection between a UE and the CN.

As shown in Section 4.1, we note that the data traffic patterns are very similar across each day over the 7-day span. Thus, in this paper, we focus on the 24-hour span of traces collected on November 28, 2010. The 24-hour traces have a total size of around 1.9TB with about 4.10 billion packets and 56.5 million 5-tuple flows.
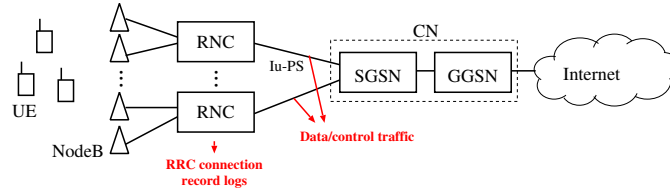
Fig. 1. A simplified architectural view of a 3G UMTS network.

Category	Description
Web browsing	HTTP-based web browsing activities
IM	instant messaging applications such as Yahoo Messenger, MSN, Tencent QQ, etc.
File access	audio/video files, software download/updates, FTP, etc.
Streaming	online video services such as YouTube, PPLive, QQLive, Tudou, etc.
Email	email services such as IMAP, POP3, SMTP, etc.
P2P	file sharing via P2P software such as BitTorrent, Thunder, etc.
Network Admin	network protocols such as DHCP, DNS, MDNS, NetBios, NTP, etc.
Tunneling	tunneling protocols such as Socks4, Socks5, SSL
Others	other successfully identified applications

Table 1. Summary of categories of applications.

Given that our captured data packets contain full payloads, we apply a deep packet inspection (DPI) module to identify the application protocols. The DPI module has been available as part of a commercial product [11]. It maintains a database of signatures of hundreds of applications, and maps each 5-tuple flow to an IP application based on the pre-defined payload signatures. Table 1 summarizes the top application categories that we focus on in our analysis.

RRC connection record logs. A UE sends (receives) data to (from) the Internet via a *Radio Resource Control (RRC)* connection. An RNC keeps the information for every RRC connection that we refer to as an *RRC record*. We collect the log files of the RRC records from all the 16 RNCs in the network (see Fig. 1). The log files span the same period as our data/control traffic trace, and account for a total of 168 million records with total size around 28GB from November 25, 2010 to December 1, 2010. From the RRC records, we then extract the useful fields including: (i) *International Mobile Subscriber Identity (IMSI)*, the unique identification associated with each GSM- or UMTS-based UE, (ii) *type allocation code (TAC)*, the code that uniquely identifies the UE device type and is composed of the initial 8-digit portion of the 15-digit *International Mobile Equipment Identity (IMEI)* code, (iii) RNC-ID, which uniquely identifies an RNC, and (iv) setup/release times of the RRC connection, which enable us to keep track of the connection duration.

Device type mappings. For each TAC, we identify the corresponding device type, including the hardware model and operating system of a UE. We then establish the device type mappings. Such mappings can be obtained from different public sources such as [13, 20]. Table 2 shows the device types considered in our analysis.

3.2 Correlating Datasets

We now describe how we correlate the datasets that cover both data-plane and control-plane information. Our correlation seeks to identify the control-plane behaviors of the

Device type	Description
iPhone	smartphones developed by Apple
Symbian	smartphones using Symbian as the operating system
Windows mobile	smartphones using Windows mobile as the operating system
Blackberry	e-mail and smartphone devices developed by RIM
Android	smartphones using Android as the operating system
Bada	smartphones developed by Samsung
Linux	smartphones using Linux as the operating system
Feature phone	modern low-end mobile phones
iPad	tablet computers developed by Apple
Datacard	laptops or PCs using datacards to access 3G services

Table 2. Summary of device types.
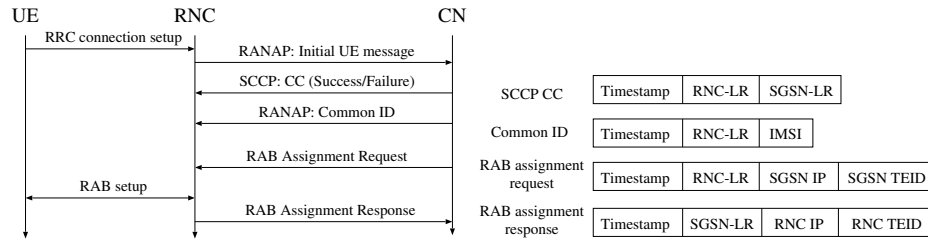
Fig. 2. Signaling message flow of the RRC connection setup.

Fig. 3. Signaling message format (LR denotes local reference).

corresponding data-plane packets generated by different device types and different IP applications.

Extracting signaling messages. We first extract the RANAP signaling messages that are later used for correlation with RRC records. Fig. 2 depicts the signaling message flow of the RRC connection setup. When the UE wants to send or receive data, the RNC first makes a *Signaling Connection Control Part (SCCP)* connection request with the CN, which replies a *SCCP Connection Confirm (SCCP CC)* message. SCCP can be viewed as the transport protocol for RANAP. The CN also replies a *Common ID* message, which contains the IMSI of the UE. The RNC associates the IMSI with the RRC connection for the UE and keeps this association throughout the RRC connection. Note that a single RRC connection may consist of one or multiple *Radio Access Bearers (RABs)* assigned for data communications. Each RAB assignment defines the actual radio resources for transmitting data traffic between the UE and the CN. The CN sends a RAB Assignment Request message to the RNC, which then executes the RAB setup protocol with the UE. Finally, the RNC replies a RAB Assignment Response message to the CN. From this onwards, the UE can send or receive data.

Based on the above signaling message flow, we can identify four signaling messages that are important for our correlation: SCCP CC, Common ID, RAB Assignment Request, and RAB Assignment Response. Fig. 3 summarizes the fields of the signaling messages that we use for our correlation.

Correlating data/control traffic and RRC record logs. The goal of our correlation is to identify the data/control traffic for each RRC connection. One major challenge is that the data/control traffic and the record logs are obtained at different capture points that have *loosely synchronized system clocks* with around 60-150 second difference. Thus, we need to allow tolerance of such capture time differences in our correlation.

The basic idea of our correlation process is similar to the relational join operation by correlating the common fields in different datasets. We elaborate the details as follows:

1. We join the four signaling messages that have identical RNC-LR and SGSN-LR values and are seen within 15 seconds. RNC-LR and SGSN-LR are the local references of the RNC and SGSN associated with the RRC connection, respectively.
2. We join the output record in Step 1 with the RRC record that has the same IMSI field and is logged within 150 second difference. The time difference accommodates the imperfect synchronization among different captured points.
3. We then join the output record in Step 2 with the data packets. The correlation is based on the *Tunnel Endpoint Identifier (TEID)* field, which identifies the data communication tunnel. Each correlated data packet must have the same RNC IP, SGSN IP, and TEID fields as in the RRC record. Also, its timestamp must either be within the duration of the RRC connection, or differ from the setup/release times by no more than 150 seconds. If a data packet matches more than one RRC record, then we correlate it with the RRC record that has the closest setup/release time with the packet timestamp.
4. Finally, we obtain the device information of each packet with the device type mappings based on the TAC field.

The final output is that each data packet is associated with its corresponding RRC record and the device type. We then parse the data packets and perform the analysis.

Implementation. Our correlation is implemented as a MapReduce [4] program and run on a Hadoop [9] platform. This allows us to *parallelize* the correlation analysis. We utilize the *repartitioned sort-merge join* provided by Hadoop to realize the correlation. Our Hadoop platform consists of one namenode and seven datanodes that are connected by Gigabit switches. The platform has 112 CPU cores and 112GB memory in total, and it takes around 10 hours to complete the correlation process for one day of traces.

3.3   Limitations

One limitation of our work is that our datasets were collected in November 2010 (about a year ago from the time of this writing). Given the dramatic expansion of the smartphone market worldwide [12], we expect that there is a dramatic growth of the volume of data/control-plane traffic in 3G networks as of today. Also, there have been continuing version upgrades for smartphone operating systems (e.g., in Android and iPhone/iPad), and such upgrades may change the underlying data transmission behaviors. Nevertheless, since our correlation methodology is based on the standard 3G specifications, it remains applicable for today's 3G networks in general, provided that the same datasets are available. One important future work is to validate the findings of our measurement study with the latest datasets.

Another limitation is that we cannot validate the accuracy of the DPI tool used in our flow-based application identification, given that the tool is part of a commercial

product and is closed-source. Nevertheless, we expect that the identification process is sufficiently accurate for our analysis due to the commercial nature of the tool. Some of our findings (see Section 4) also conform to our intuition. Note that our flow-based application identification can also be substituted by other approaches, such as port-based analysis [6] or dynamic protocol detection [5].

4 Measurement Results

4.1 Overview

We first give a general overview of our captured data traffic. After pre-processing, we identify more than 80% of traffic (out of ∼13TB) that can be successfully associated with the corresponding signaling traffic. Fig. 4 plots the IP traffic volume (per minute) over time in our 7-day traces, and Fig. 5 shows the number of devices identified on each day. We observe that the distributions of each day are fairly stable. In particular, the distributions do not indicate significant differences across weekdays and weekends.

In the following analysis, we only focus on the 1-day traces on November 28, 2010. Given the regularity of the traffic, we expect that our observations are consistent for other days of traffic.
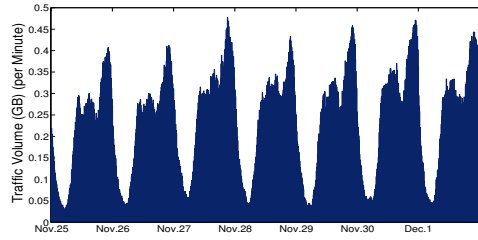

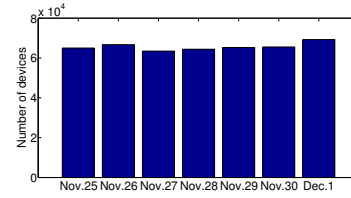**Fig. 4.** Total traffic volume (per minute) over one week.

Fig. 5. Number of devices per day over one week.

Fig. 6 illustrates the number of devices for different types listed in Table 2. There are 64,398 devices in total, in which iPhone leads all devices with a proportion of 32%. Mobile hand-held phones (including smartphones or low-end feature phones) dominate the majority of devices, while iPad and Datacard only account for 7.9% and 7.5% of devices, respectively. We point that over 99% of the devices can be successfully identified with the corresponding device types.

Fig. 7 shows the total traffic volume for each traffic type. Datacard devices contribute 46% of the total traffic, although they only account for 7.5% of devices. We will explore possible reasons for this phenomenon later in Section 4.2.

We now evaluate the control-plane performance of different device types, based on our RRC record logs. Fig. 8 shows the average number of RRC connections per device for each device type. As shown, iPhone triggers the most RRC connections per device (237 times), followed by iPad (174 times). Even datacard devices dominate the largest portion of IP traffic, each datacard device triggers a relatively small number of RRC connections (68 times) compared to smartphones (e.g., Android, Windows Mobile). Note that each RRC connection setup/release triggers a number of signaling messages
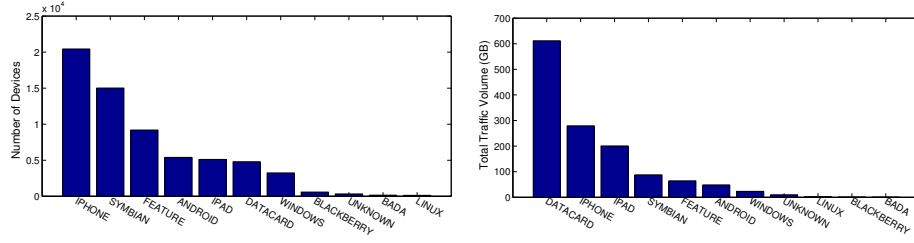
Fig. 6. Number of devices for each device type.   **Fig. 7.** Total traffic volume of each device type.

between a UE and an RNC, and such signaling messages consume radio resources of the cellular network [16]. With the frequent RRC connection setups/releases, we can see that iPhone/iPad devices can bring large signaling overhead of an RNC and consume substantial radio resources of the whole cellular network.

Fig. 9 shows the average RRC connection duration per device for each device type. Note that iPhone has the smallest duration among all (30 seconds). Since iPhone triggers more RRC connections and each RRC connection has a short duration, it implies that iPhone devices inject a high intensity of signaling workload into the network. In contrast, Datacard has the longest RRC connection duration, so we expect that Datacard devices tend to run long-lived data transfer applications that keep an RRC connection active for a long time.

We thus far observe that different device types have distinct data/control-plane performance. In the following subsections, we explore the possible reasons behind.
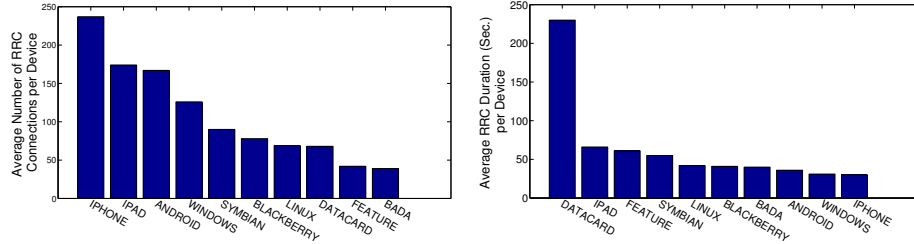

Fig. 8. Average number of RRC connections per device.   **Fig. 9.** Average RRC duration per device.

4.2   Application Characteristics

Let us examine the application usage patterns of different device types, and explore how the traffic compositions of the device types lead to the distinct data/control-plane behaviors that we observe. We first analyze the overall traffic distributions of different IP applications listed in Table 1. Fig. 10 shows the traffic volume distribution of different IP applications, where web browsing (38%), streaming (21%), P2P (10%) and file access (10%) are ranked the top four. IM and Email applications contribute 2% and 1% of the total traffic, respectively. The traffic that cannot be identified, i.e., labeled as "Unknown" in Fig. 10, only contributes 9% of the traffic. In other words, over 90% of total traffic can be successfully identified by our DPI tool.
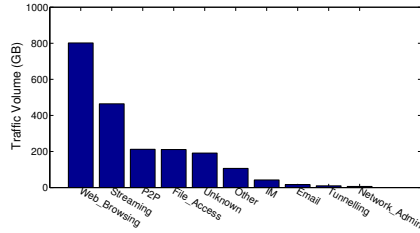
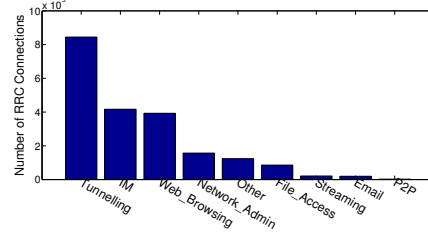**Fig. 10.** Traffic volume of applications.

Fig. 11. Total number of RRC connections of applications.

Fig. 11 shows the total number of RRC connections triggered by each application. In this work, we define that an application triggers an RRC connection if the first data packet transmitted in the RRC connection belongs to the application. We see that Tunneling triggers the most RRC connections among all applications. The reason is that a tunneling session is first established for providing a secure path for any other application protocols, or supporting the roaming service. IM triggers 21% of RRC connections, while generating only 2% of the total traffic (Fig. 10). On the other hand, P2P triggers only 0.1% of the total number of RRC connections, while contributing 11% of the total traffic. This is an important observation because it clearly shows that different applications have *inherently different* data/control-plane behaviors.

Let us investigate the application usage patterns of different device types. Fig. 12 illustrates the total traffic volume of applications for each device type, in which we only look at the top device types that contribute the most traffic. We make two key observations. First, datacard devices contribute 85% and 48% of all P2P and streaming traffic, respectively. Since most datacards are attached to static terminals such as PCs or laptops, these static terminals tend to run long-lived applications. Thus, they contribute a large volume of data-plane traffic even the population of datacard devices is small. Also, P2P and streaming applications trigger very few but long-lived RRC connections, and this validates our previous observations in Figs. 8 and 9. Second, web browsing, streaming, and file access are the top three applications that account for the most traffic volume on smartphones (i.e., iPhone, Android, Symbian, Windows Mobile), and they altogether contribute over 80% of the smartphone traffic.

4.3   Active Devices

We now explore the behaviors of *active* devices and see their impact on the data/control-plane performance. We say that a device is active within time $[t_1, t_2]$ if it starts and releases an RRC connection at times t_1 and t_2, respectively. Figs. 13 and 14 plot the distributions of the traffic volume and the number of active devices (per minute) for different device types over a 24-hour period (on Nov. 28), respectively. As expected, Fig. 13 shows that the traffic volume peak appears in day time, and the traffic trough appears at late night. While the number of active devices in Fig. 14 shows a similar pattern, the peak-trough difference is significantly different. We find that the traffic volume at the trough is 93% less than the peak traffic volume in Fig. 13, while the number of active devices at the trough is only 52% less than the peak value in Fig. 14.

After further investigation, we find that iPhone and iPad have a very similar pattern that is distinct from other device types, i.e., the numbers of active devices of iPhone
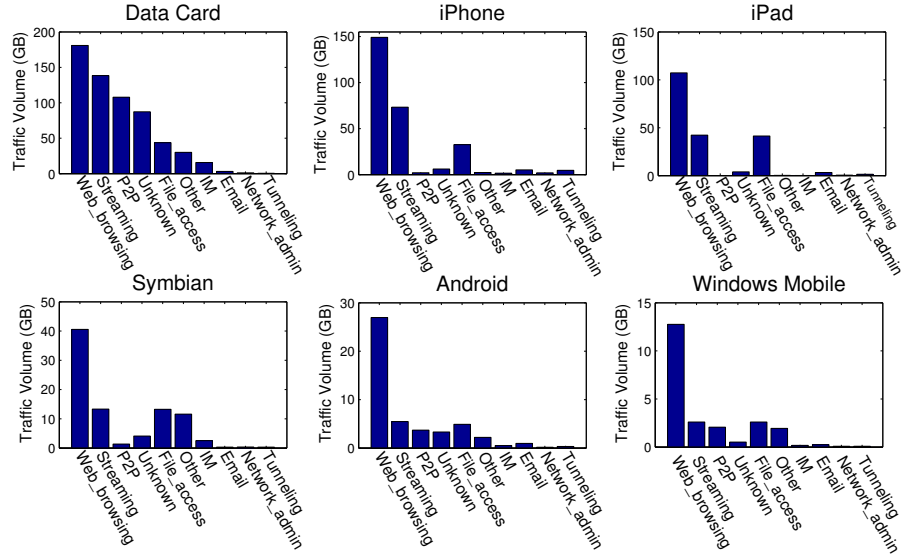
Fig. 12. Traffic volume distribution of applications for each device type (note that the y-axis scale is different for each device type for clear presentation).

and iPad remain quite stable during the 24-hour period. On the other hand, other device types exhibit obvious troughs and peaks in the number of active devices, with the troughs also appearing at late night.

We find that the internal heartbeat mechanism of iPhone/iPad keeps the devices active, even there is no data traffic initiated by user applications. To demonstrate, Fig. 15 shows the probability density of the inter-arrival times of RRC connection setups of iPhone. The inter-arrival times of RRC connections of iPhone occur more often at two values: 64 seconds (18.1%) and 589 seconds (4.9%). Similar observations are also made for iPad. We investigate and find that iPhone/iPad devices are "always-on" and generate heartbeat packets around every 60 seconds, and such packets are sent to Apple servers. Each heartbeat packet triggers an RRC connection, which will be released if there is no data traffic for an idle period that is less than 60 seconds. On the other hand, we do not observe any explicit heartbeat patterns in Android (not shown in the figure), possibly because there are many variants of Android devices and their heartbeat designs are different. In short, the heartbeat packets that we observe in the iPhone/iPad devices introduce many RRC connection setups/releases, which could lead to high signaling overhead in the network core.

Note that there are a few upgrades of iOS (Apple's mobile operating system) in 2011. In iOS 4.2 (which was launched in late November 2010), a technology that enables mobile devices and the network core to collaboratively minimize network congestion and mobile battery consumption [15]. This may change the heartbeat behavior that we observe in our datasets. One future work is to validate such a heartbeat behavior with the latest datasets, based on the same correlation methodology as described in Section 3.2.
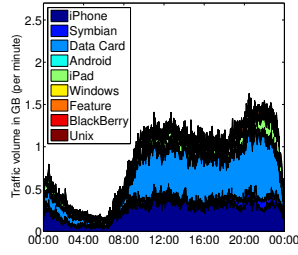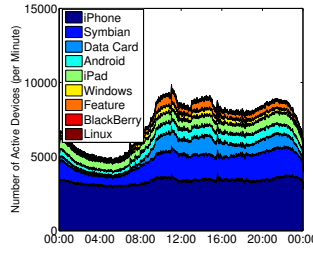
Fig. 13. Traffic volume (per minute) distribution.

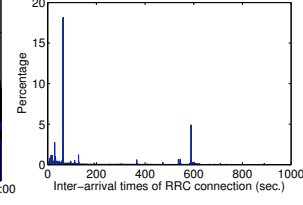**Fig. 14.** Number of active devices (per minute).

Fig. 15. Density of interarrival times of RRC connection setups of iPhone.

5   Summary, Implications, and Conclusions

We now summarize the key observations and potential challenges through the analysis of our massive data traces.

- In the data plane, datacard devices contribute almost 50% of the total traffic, while accounting for only 7% of the device population. Network operators may devise special resource allocation schemes for datacard users. Also, iPhone/iPad devices altogether account for around 40% of devices, and contribute nearly 40% of the total traffic due to their large market shares.
- Our application-based analysis clearly shows that different user groups (distinguished by device types) have distinct user behaviors with specific preferences of choosing applications. For example, in all Apple devices (iPhone and iPad), the dominant applications are web browsing, streaming, and file access, and they altogether contribute more than 90% of iPhone/iPad traffic. From the network operators' perspectives, there are challenging issues of providing QoS guarantees and QoE improvement for different user groups.
- In the signaling control plane, among all device types, iPhone has the highest average number of RRC connections per device. Meanwhile, iPhone/iPad devices have a very similar heartbeat behavior that triggers significantly more RRC connections than any other device types. Recall that each RRC connection setup/release involve a number of signaling messages exchanged between a UE and an RNC [16]. These results reveal that specific device types can increase the signaling overhead to the network.

This paper analyzes the data/control-plane performance of different device types using massive traces collected from the core of a 3G UMTS city-wide network. We describe a correlation methodology that studies the interactions of different data/control-plane datasets, and make several key observations from our analysis that could be of interest to network operators and researchers. Our measurement study motivates the importance of characterizing the data/control-plane workloads of a 3G network and designing adequate strategies for network planning, resource allocation, and pricing. Since our measurement study is based on the datasets that are compliant with the 3G standard, our analysis methodology can be generalized for other 3G operational networks.

Acknowledgment

References

1. Cisco. Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 20102015, Feb 2011.
2. A. D'Alconzo, A. Coluccia, F. Ricciato, and P. Romirer-Maierhofer. A Distribution-Based Approach to Anomaly Detection for 3G Mobile Networks. In *IEEE Globecom*, 2009.
3. DARWIN. `http://userver.ftw.at/~ricciato/darwin/`.
4. J. Dean and S. Ghemawat. Mapreduce: simplified data processing on large clusters. *Commun. ACM*, 51:107–113, January 2008.
5. H. Dreger, A. Feldmann, M. Mai, V. Paxson, and R. Sommer. Dynamic Application-Layer Protocol Analysis for Network Intrusion Detection. In *Proc. of USENIX Security Symp.*, 2006.
6. H. Falaki, D. Lymberopoulos, R. Mahajan, S. Kandula, and D. Estrin. A First Look at Traffic on Smartphones. In *Proc. of ACM IMC*, Nov 2010.
7. H. Falaki, R. Mahajan, S. Kandula, D. Lymberopoulos, R. Govindan, and D. Estrin. Diversity in Smartphone Usage. In *Proc. of ACM MobiSys*, Jun 2010.
8. A. Gember, A. Anand, and A. Akella. A Comparative Study of Handheld and Non-Handheld Traffic in Campus WiFi Networks. In *Proc. of PAM*, 2011.
9. Hadoop. `http://hadoop.apache.org/`.
10. J. Huang, Q. Xu, B. Tiwana, Z. M. Mao, M. Zhang, and P. Bahl. Anatomizing Application Performance Differences on Smartphones. In *Proc. of ACM MobiSys*, 2010.
11. Huawei. MSCG hierarchical DPI solution. `http://www.huawei.com/products/datacomm/catalog.do?id=1219`, 2011.
12. IDC. Worldwide Smartphone Market Expected to Grow 55% in 2011 and Approach Shipments of One Billion in 2015, According to IDC. `http://www.idc.com/getdoc.jsp?containerId=prUS22871611`, Jun 2011.
13. IMEI lookup. `http://imei-number.com/imei-lookup/`.
14. J. Kilpi and P. Lassila. Micro- and macroscopic analysis of RTT variability in GPRS and UMTS networks. In *Proc. of NETWORKING*, 2006.
15. A. Kingsley-Hughes. iOS 4.2 Supports Network Controlled Fast Dormancy. `http://www.zdnet.com/blog/hardware/ios-42-supports-network-controlled-fast-dormancy/10586`, Dec 2010.
16. P. P. C. Lee, T. Bu, and T. Woo. On the detection of signaling DoS attacks on 3G/WiMax wireless networks. *Computer Networks*, 53(15):2601–2616, Oct 2009.
17. G. Maier, F. Schneider, and A. Feldmann. A First Look at Mobile Hand-held Device Traffic. In *Proc. of PAM*, 2010.
18. U. Paul, A. P. Subramanian, M. M. Buddhikot, and S. R. Das. Understanding Traffic Dynamics in Cellular Data Networks. In *Proc. of IEEE INFOCOM*, 2011.
19. F. Qian, Z. Wang, A. Gerber, Z. M. Mao, S. Sen, and O. Spatscheck. Characterizing Radio Resource Allocation for 3G Networks. In *Proc. of ACM IMC*, 2010.
20. radiorraiders. `http://www.radioraiders.com/gsm-IMEI-TAC.php/`.
21. F. Ricciato, F. Vacirca, and M. Karner. Bottleneck detection in UMTS via TCP passive monitoring: a real case. In *Proc. of ACM CoNEXT*, Oct 2005.
22. J. Ridoux, A. Nucci, and D. Veitch. Seeing the difference in IP traffic: wireless versus wireline. In *Proc. of IEEE INFOCOM*, 2006.
23. Q. Xu, J. Erman, A. Gerber, Z. M. Mao, J. Pang, and S. Venkataraman. Identifying Diverse Usage Behaviors of Smartphone Apps. In *Proc. of ACM IMC*, Nov 2011.