QoMOSN - On the Analysis of Traffic and Quality of Experience in Mobile Online Social Networks

Pedro Casas, Pierdomenico Fiadino, Mirko Schiavonne
FTW - Telecommunications Research Center Vienna
{surname}@ftw.at

***Abstract*—The popularity of Online Social Networks (OSNs) has created an immense amount of research for the analysis of these networks and the traffic produced by popular applications such as Facebook. In this paper we study the networking characteristics of OSNs in mobile networks, analyzing the most popular OSNs in western countries: Facebook and WhatsApp. By analyzing two large-scale traffic traces collected at the cellular network of a major European ISP, we characterize and compare the networking behavior of Facebook and WhatsApp, considering not only the traffic flows but also the network infrastructures hosting them. To enable a better understanding of the Quality of Experience (QoE) requirements of these two applications in the mobile context, we present the results of a subjective QoE study of both Facebook and WhatsApp in smartphones. Our study serves the main purpose of better understanding how major OSNs are provisioned in today's mobile networks, paying special attention to their QoE characteristics. To the best of our knowledge, this is the first paper providing such an analysis using large-scale measurements in cellular networks and subjective QoE tests in smartphones.**

***Keywords*—Traffic Measurements; Quality of Experience; WhatsApp; Facebook; Content Delivery Networks; Mobile Networks.**

I. Introduction

Online Social Networks (OSNs) have rapidly become an integral part of our daily lives, and hundreds of millions of people are nowadays remotely connected through popular OSNs such as Facebook, Google+, Twitter and WhatsApp. The growing popularity of such applications in mobile networks and mobile devices call for new studies which shed light on the network and traffic footprints of OSNs, particularly in cellular networks. In addition, ensuring that the user experience remains at least as good in mobile devices as it results in the more traditional fixed desktop scenario requires a better understanding of the Quality of Experience (QoE) of OSNs in devices such as smartphones. This is especially important for mobile network operators, who need to offer high quality levels to reduce the risks of customers churning for quality dissatisfaction.

Highly popular OSNs such as Facebook have somehow redefined the way the Internet behaves and looks like. Indeed, Facebook is the most popular and widely spread OSN in Europe and the US, with hundreds of millions of users worldwide sharing and accessing content on a daily basis (see http://newsroom.fb.com/key-facts). Facebook content is mainly hosted by the well known Akamai Content Delivery Network (CDN), which represents the most dynamic and widely deployed CDN today, with more than 137,000 servers in more than 85 countries across nearly 1,200 network (see http://www.akamai.com/html/about/facts_figures.html).

There also other types of *modern* OSNs which are becoming extremely popular, especially in the mobile context. Among those, WhatsApp is doubtlessly the leading modern mobile OSN today. Initially developed as an instant messaging application, WhatsApp is today a dominant player in the communication and social sharing of contents worldwide; with half a billion users, it has become the fastest-growing company in history in terms of users. Through WhatsApp, connected users can exchange text messages and share multimedia contents among the different user groups they belong to, including photos, audio and video contents.

In this paper we present QoMOSN, an empirical analysis of the traffic and the Quality of experience of Mobile Online Social Networks. Our study focuses on the characterization of the traffic and the QoE network-side requirements of Facebook and WhatsApp in mobile devices. While many studies have been conducted on the analysis of the interactions among users and the popularity of OSNs, little is known today about the traffic characteristics, the hosting infrastructures, and the QoE-based dimensioning guidelines for such services. Our study is based on an extensive analysis of network traffic flows observed at the core of an operational European cellular network, complemented by subjective QoE tests performed on a lab scenario.

II. Related Work

The analysis of OSNs has been a very fertile domain in the last few years [4]–[9]. Authors in [4] study the power-law and scale-free properties of the interconnection graphs of Flickr, YouTube, LiveJournal, and Orkut, using application-level crawled datasets. The work in [5] present a study on the privacy characteristics of Facebook. Some papers [6], [7] study the new Google+ OSN, particularly in terms of popularity of the OSN, as well as the evolution of connectivity and activity among users. Authors in [8], [9] focus on the temporal dynamics of OSNs in terms of user-interconnections and visited links, using again publicly crawled data of popular OSNs such as Facebook, Twitter, as well as a large Chinese OSN. All these papers rely on crawled web-data and do not take into account the traffic and networking aspects of OSNs. WhatsApp is a very new service, and its study has been so far

The research leading to these results has received funding from the European Union under the FP7 Grant Agreement n. 318627, "mPlane", and was partly performed within the project ACE 3.0 at FTW Vienna.

978-1-4673-7359-3/15/$31.00

quite limited. Some recent papers have partially addressed the characterization of its traffic [1]–[3], but mainly considering an energy-consumption perspective [2], [3]. In terms of QoE network-side requirements for both services, little has been done so far and only some related papers shed some light on the subject [10], [11].

III. OSN Servers and Mobile Traffic

A. Datasets and Analysis Methodology

Our study is conducted on top of two large-scale network traffic traces collected at the core of a European national-wide cellular network in mid 2013 and early 2014. Flows are monitored at the well known Gn interface links between the GGSN and SGSN nodes. Facebook traffic is carried on top of HTTP (we do not consider HTTPS for the study of Facebook, as its usage in 2013 was very limited in mobile devices using the Facebook App or the web-based access), so we rely on a HTTP-based traffic classification tool for cellular traffic called HTTPTag [13] to unveil the corresponding Facebook flows. HTTPTag classification consists in applying pattern matching techniques to the `hostname` field of the HTTP requests. The Facebook dataset consists of one month of HTTP flow traces collected in mid 2013. To preserve user privacy, any user related data (e.g., IMSI, MSISDN) are removed on-the-fly, whereas any payload content beyond HTTP headers is discarded on the fly.

The WhatsApp dataset consists of a complete week of WhatsApp traffic flow traces collected at exactly the same vantage point in early 2014. In the case of WhatsApp all communications are encrypted, so we extended the HTTPTag classification tool to additionally analyze the DNS requests, similar to [12]. In a nutshell, every time a user issues a DNS request for a Fully Qualified Domain Name (FQDN) associated to WhatsApp, HTTPTag creates an entry mapping this user to the server IPs provided in the DNS reply. Each entry is time stamped and contains the TTL replied by the DNS server. Using these mappings, all the subsequent flows between this user and the identified servers are assumed to be WhatsApp flows. To avoid miss-classifications due to out-of-date mappings, every entry expires after a TTL-based time-out. To increase the robustness of the approach, the list of IPs is augmented by adding the list of server IPs signing the TLS/SSL certificates with the string `*.whatsapp.net`. Indeed, our measurements revealed that WhatsApp uses this string to sign all its communications. Finally, we use reverse DNS queries to verify that the list of filtered IPs actually corresponds to a WhatsApp domain.

To identify the FQDNs used by the WhatsApp service, we follow the techniques introduced in our previous work [1], where manual inspection of hybrid measurements was performed. As we found in that study, different third level domain names are used to handle different types of traffic (control, text messages, and multimedia messages). Control and text messages are handled by *chat servers*, whereas multimedia contents (audio, photos and videos) are handled by *multimedia (mm) servers*. As we shall see next, chat and mm servers have very different network footprints. While connections to chat servers are characterized by low data-rate and long duration (especially due to the control messages), media transfers are carried by short and heavy flows.

To study the hosting infrastructures of both OSNs, we complement the traffic datasets with the name of the organization and the Autonomous System (AS) hosting the content, extracted from the MaxMind GeoCity databases (http://www.maxmind.com).

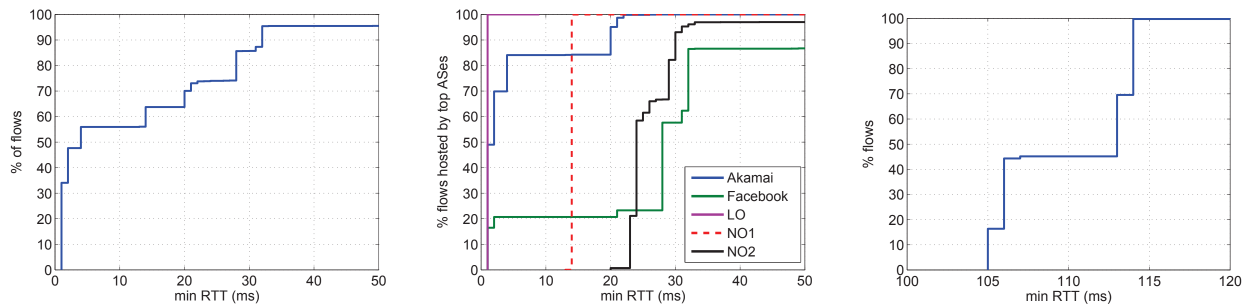
B. Content Delivery Infrastructure

We start by characterizing the Facebook dataset, with a special focus on its underlying hosting/delivery infrastructure. Due to the high number of daily users and the high volumes of served traffic, Facebook follows a sophisticated content delivery strategy. Indeed, we observed more than 6500 server IPs hosting Facebook contents in our traces, distributed across 20 countries and more than 260 different ASes. This confirms the wide-spread presence of several organizations hosting Facebook contents, turning the service provisioning into a very complex scenario. The main organizations/ASes hosting Facebook content are Akamai, Facebook AS (i.e., Facebook itself), the Local Operator (LO) which hosts the vantage point, and two neighbor operators, Neighbor Operator 1 (NO1) and Neighbor Operator 2 (NO2).

Almost 99% of the traffic comes from servers and data centers located in Europe, close to our vantage point, while only 1% of the traffic comes from other continents. This is due to three factors: (i) Akamai, the biggest Facebook content provider, has a very geographically distributed presence, pushing contents as close as possible to end-users; (ii) operators heavily employ local content caching, and large CDNs like Akamai tend to deploy caches inside the ISPs networks, explaining the amount of traffic coming from the local country as well as neighboring countries to the vantage point; (iii) the rest of the traffic is handled directly by Facebook, which has servers split between Ireland (headquarter of Facebook International) and the US.

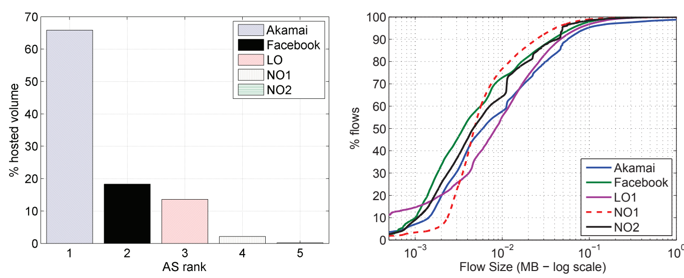
In the case of WhatsApp, the complete one-week server IP mappings resulted in a total of 386 IPs identified as hosting the service, belonging to a single AS called SoftLayer (http://www.softlayer.com, AS number 36351). To avoid biased conclusions about the set of identified IPs from a single vantage point, we performed an active measurements campaign using the RIPE Atlas measurement network (https://atlas.ripe.net/), where we analyzed which IPs were obtained when resolving the same FQDNs from 600 different boxes around the globe during multiple days. These active measurements confirmed that the same set of IPs is always replied, regardless of the geographical location of the requester. SoftLayer is a US-based cloud infrastructure provider consisting of 13 data centers and 17 Points of Presence (PoPs) distributed worldwide.

The WhatsApp hosting infrastructure is completely different. Following the same approach, we observed that despite its geographical distribution, WhatsApp traffic is handled mainly by the data centers in Dallas and Houston, being as such a fully centralized US-based service. While this is likely to change in the near future after Facebook's WhatsApp acquisition, right now, all messages among users outside the US are routed through the core network, unnecessarily consuming additional network resources and potentially impacting the quality of the service.



(a) min RTT to all Facebook server IPs. (b) min RTT to top Facebook hosting ASes. (c) min RTT to all WhatsApp server IPs.

Figure 1. Distribution of overall min RTT to Facebook and WhatsApp server IPs, weighted by the number of flows hosted. In the case of Facebook, results are also discriminated per top hosting ASes.



(a) Shares of hosted volume per org./AS. (b) Distribution of flow sizes.

Figure 2. Hosted volume and distribution of flow sizes per organization.

We additionally investigate the location of the servers from a network topology perspective, considering the distance to the vantage point in terms of Round Trip Time (RTT). The RTT to any specific IP address consists of both the propagation delay and the processing delay, both at destination as well as at every intermediate node. Given a large number of RTT samples to a specific IP address, the minimum RTT values are an approximated measure of the propagation delay, which is directly related to the location of the underlying server. RTT values are obtained from active measurements, performed during the span of each dataset, using a standard ping tool.

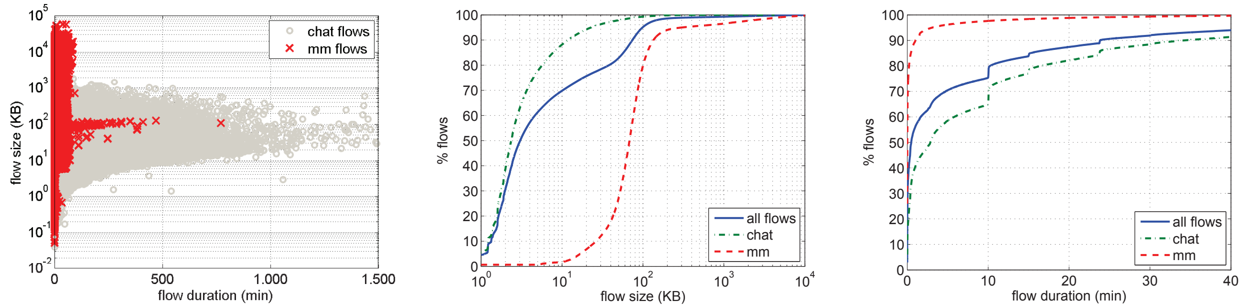
Figure 1 plots the distribution of the minimum RTT to (a) the server IPs hosting Facebook, (b) the top orgs./ASes hosting Facebook, and (c) the server IPs hosting WhatsApp. Values are weighted by the number of flows hosted at each IP, to get a better picture of where the traffic is coming from. As a further confirmation of the geographical diversity in Facebook, the distribution of min RTT presents some steps or "knees", suggesting the existence of different data centers and/or hosting locations. The largest majority of Facebook flows are served by close serves, located at less than 5 ms from the vantage point. As we mentioned, Akamai deploys its servers deep into ISPs, placing content distribution servers inside ISP POPs, which explains the short latency to the vantage point. The LO is the one with shortest delays for all the flows it serves and, along with the NO1, is the one with the least geographical diversity, with only one visible location. Three main steps appear in the CDF of the Facebook servers, which correspond to the headquarters in Ireland (min RTT about 30ms), the servers in the US (min RTT $>$ 100ms), and some servers located at only few milliseconds from the vantage point. Traceroutes to those servers revealed a direct connection to the Internet eXchange Point (IXP) of the local country, explaining the so small delays.

In the case of WhatsApp, the min RTT is always bigger than 100ms, confirming that WhatsApp servers are located outside Europe. Figure 1(c) shows that the service is evenly handled between two different yet potentially very close locations at about 106 ms and 114 ms, which is compatible with our previous findings of WhatsApp servers located in Dallas and Houston. The main finding of this analysis is that even if both OSNs are very popular worldwide, their networking hosting infrastructures follow very different paradigms: based on Akamai's pervasiveness, Facebook is hosted by a highly distributed network architecture, whereas WhatsApp follows a fully centralized and not geo-distributed hosting architecture. Being Brazil, India, Mexico and Russia the fastest growing countries in terms of users (WhatsApp Blog, http://blog.whatsapp.com/), such a centralized hosting infrastructure is likely to become a problematic bottleneck in the near future. Very interesting is the fact that in both cases data locality is probably not maintained, as content might be hosted in other countries outside the local one (in the case of WhatsApp, this is 100% confirmed). In the light of the ever increasing concerns related to privacy and data security, such a geographical distribution might even cause legal jurisdiction issues due to different data privacy protection laws in different countries.

C. Network Traffic Analysis

We focus now on the characteristics of the traffic flows carrying Facebook and WhatsApp contents. Figure 2(a) depicts the volume share of Facebook contents hosted by each org./AS, as well as the flow size distributions. Akamai hosts more than 65% of the total volume observed in our traces, followed by Facebook AS itself with about 19%. In fact, there is a clear distinction on the content sizes handled by both Akamai and Facebook AS: while Akamai hosts the bigger flows, Facebook AS serves only a small share of the service content. Indeed, it is well-known that Akamai serves the static contents of the Facebook service (e.g., photos, songs, videos, etc.), whereas the Facebook AS covers almost exclusively the dynamic contents (e.g., chats, tags, session information, etc.).

To further explore this distinction, Figure 2(b) reports the distribution of the flow sizes served per organization. The CDF suggests that Akamai serves bigger flows than Facebook AS. The remaining ASes tend to host bigger flows than Facebook AS, which is coherent with the fact that ISPs caching is



(a) Flow duration vs. size. (b) Flow size. (c) Flow duration.

Figure 3. Characterization of WhatsApp flows. Whereas mm messages are sent over short-lived flows, text messages result in longer and much smaller flows.

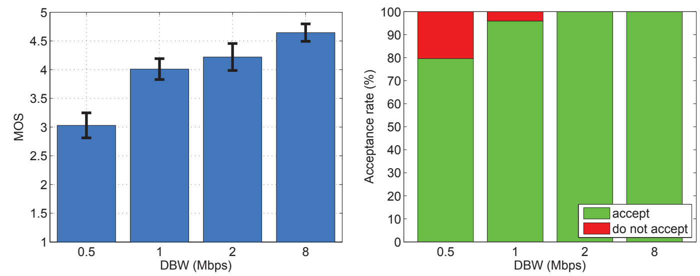
generally done for bigger objects, aiming to reduce the load on the core network.

In terms of WhatsApp traffic, Figure 3 reports the characteristics of the corresponding flows in terms of size and duration. Figure 3(a) shows a scatter plot reporting the flow duration vs. the flow size, discriminating by chat and mm flows. Whereas mm messages are sent over dedicated connections, resulting in short-lived flows, text messages are sent over the same connection used for control data, resulting in much longer flows. For example, some chat flows are active for as much as 62 hours. Figure 3(b) indicates that more than 50% of the mm flows are bigger than 70 KB, with an average flow size of 225 KB. More than 90% of the chat flows are smaller than 10 KB, with an average size of 6.7 KB. In terms of duration, Figure 3(c) shows that more than 90% of the mm flows last less than 1 min (mean duration of 1.8 min), whereas chat flows last on average as much as 17 minutes. The flow duration distribution additionally reveals some clear steps at exactly 10, 15 and 24 minutes, suggesting the usage of an application time-out to terminate long idle connections. This behavior is actually dictated by the operating system of the device, as we show in [1].

IV. QoE in Mobile OSNs

To provide a holistic overview of both Facebook and WhatsApp in mobile devices, we studied the end-user perspective of both applications in terms of their QoE requirements, through a subjective QoE study. The subjective study was performed in a dedicated lab for subjective analysis, compliant with the recommendations provided by the QoE subjective studies standards [14]–[16]: 52 people participated in the study (29 female, 23 male), the average age was 32 years old, with 40 participants being less than 30 years old. Around half of the participants were students and almost 43% were employees, and 70% of the participants completed university or baccalaureate studies. The key benefits of conducting lab studies instead of an analysis in the wild rely on the full control the experimenter has on the overall evaluation process. Indeed, content and context are fully known and controlled, and users are directly briefed and observed on the spot, providing as such tangible and solid results.

Participants performed specific tasks with the Facebook and WhatsApp applications using Android smartphone devices (Samsung Galaxy S4, OS Android 4.4 KitKat) connected to the Internet through independent WiFi access points. The downlink traffic was routed through a modified version of the well-known NetEm network emulator to impose different access network bandwidth and latency profiles. In the case of Facebook, participants were instructed to access the application with a specific user account, browse the timeline of this user, and browse through specific photo albums created for this user. Both the Download BandWidth (DBW) and the access RTT were modified in the case of Facebook. In the WhatsApp tests, participants worked in couples and exchanged specific video files of fixed size (i.e., 5 MB), and the participant downloading the video file was the one providing a QoE feedback.



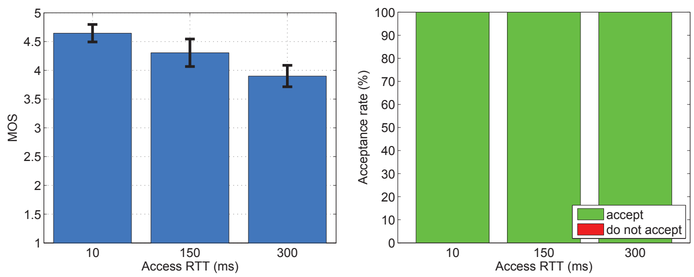
(a) Overall quality (MOS). (b) Acceptance rate.

Figure 4. QoE in Facebook. Overall quality and acceptability for different download bandwidth configurations.

Regarding QoE feedbacks, participants were instructed to rate the *overall experience* according to a continuous ACR-5 MOS scale [14], ranging from "bad" (i.e., MOS = 1) to "excellent" (i.e., MOS = 5). Participants also provided feedback on the acceptability of the application, stating whether they would continue using the application under the corresponding conditions or not. MOS ratings were issued by participants through a custom questionnaire application running on separated laptops, which pops-up immediately after a condition has been tested.

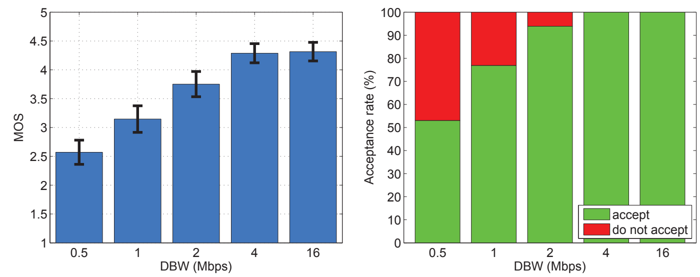
A. *QoE in Facebook Mobile*

Figure 4 reports the results obtained in the Facebook tests for different DBW configurations, considering both (a) the overall quality and (b) the acceptance rate. Before commenting on the results, the reader should note that the maximum MOS ratings declared by the participants are never 5 but somewhere around 4.5. This is a well known phenomenon in QoE studies called *rating scale saturation*, where users hardly employ the



(a) Overall quality (MOS). (b) Acceptance rate.

Figure 5. QoE in Facebook. Overall quality and acceptability for different access delay configurations.



(a) Overall quality (MOS). (b) Acceptance rate.

Figure 6. QoE in WhatsApp. Overall quality and acceptability for different download bandwidth configurations.

limit values of the scale for their ratings [17]. So from now on, we shall consider as optimal quality a MOS ≈ 4.5.

A DBW of 500 kbps is not high enough to reach full user satisfaction in Facebook mobile for Android devices, as participants declared a fair quality with an acceptance rate of about 80%. Still, a DBW of 1 Mbps results in good overall quality, with almost full acceptability of the participants. Excellent QoE results are attained for 8 Mbps, which shows that even if a 2 Mbps DBW allocation is high enough to reach full acceptability (cf. Figure 4), the overall experience of the user can still marginally improve. These DBW thresholds are highly important for network dimensioning, as they allow to understand the boundaries between user satisfaction and over-provisioning of resources. Very interesting is the fact that these QoE requirements in terms of DBW are more restrictive than those we found in [10] for laptops about 2 years ago, evidencing how the Facebook application has been evolving in time, becoming more network resources demanding.

In terms of access latency, Figure 5 shows that QoE degrades when the access RTT increases far beyond 100 ms, but the impacts are not as significant as one might expect a-priori for a browsing-like application, and acceptability seems not to be impacted at all. The main reason for such a result is that the degree of interactivity of the Facebook application is not as high as for other applications such as video-conferencing or gaming, suggesting that all in all, the operator should focus the dimensioning on the DBW rather than the access RTT.

B. QoE in WhatsApp

Figure 6 shows the QoE results for different DBW values. In the case of WhatsApp we only test DBW, as changing the access RTT would result in similar results. Indeed, when considering the QoE of file downloads, download time is in fact the most relevant feature as perceived by the user [11], which is directly linked to anxiety and satisfaction. The Figure shows that users tolerate WhatsApp downloads with a good overall experience and high acceptability as long as the DBW is above 2 Mbps, but user experience heavily degrades for slower connections, resulting in very bad quality for a DBW of 500 kbps. In this case, a DBW threshold of 2 Mbps permits to approximately discriminate between good and bad experience. Given the file size used in the tests, there is a clear saturation effect after 4 Mbps, as QoE does not increase for higher DBW values. Finally, even if the obtained results are partially biased by both the specific file size used in the tests and the participants task briefing, we believe that the study is good in providing guidelines for dimensioning the access network for the case of WhatsApp as well. Even more, the obtained results are similar to those we obtained in [11] for the specific case of Dropbox file sharing, suggesting that the main take aways are potentially more generic than expected when considering file downloads.

V. Conclusions

In this paper we have provided a holistic overview of the networking and QoE characteristics of Facebook and WhatsApp, particularly in the context of mobile networks. By analyzing the traffic collected at the cellular network of a major European ISP, we have characterized the networking behavior of both applications, considering not only the traffic flows but also the hosting infrastructures. We additionally presented the results obtained in subjective QoE tests performed for both applications in smartphones, complementing an end-to-end analysis of these popular OSNs in mobile networks and mobile devices.

References

[1] P. Fiadino et al., "Vivisecting WhatsApp through Large-Scale Measurements in Mobile Networks", extended abstract in *ACM SIGCOMM*, 2014.

[2] E. Vergara et al., "When Mice Consume Like Elephants: Instant Messaging Applications", in *ACM e-Energy*, 2014.

[3] A. Aucinas et al., "Staying Online While Mobile: The Hidden Costs", in *ACM CoNEXT*, 2013.

[4] A. Mislove et al., "Measurement and Analysis of Online Social Networks", in *ACM IMC*, 2007.

[5] Y. Liu et al., "Analyzing Facebook Privacy Settings: User Expectations vs. Reality", in *ACM IMC*, 2011.

[6] G. Magno et al., "New Kid on the Block: Exploring the Google+ Social Graph", in *ACM IMC*, 2012.

[7] R. Gonzalez et al., "Google+ or Google-? Dissecting the Evolution of the New OSN in its First Year", in *ACM WWW*, 2013.

[8] X. Zhao et al., "Multi-scale Dynamics in a Massive Online Social Network", in *ACM IMC*, 2012.

[9] M. Allamanis et al., "Evolution of a Location-based Online Social Network: Analysis and Models", in *ACM IMC*, 2012.

[10] P. Casas et al., "YouTube & Facebook Quality of Experience in Mobile Broadband Networks", in *IEEE Globecom Workshops*, 2012.

[11] P. Casas et al., "A First Look at Quality of Experience in Personal Cloud Storage Services", in *IEEE ICC Workshops*, 2013.

[12] I. Bermudez et al., "DNS to the rescue: Discerning Content and Services in a Tangled Web", in *ACM IMC*, 2012.

[13] P. Fiadino et al., "HTTPTag: A Flexible On-line HTTP Classification System for Operational 3G Networks", in *IEEE INFOCOM*, 2013.

[14] Int. Telecommunication Union, "Methods for Subjective Determination of Transmission Quality", *ITU-T Rec. P.800*, 1996.

[15] Int. Telecommunication Union, "Subjective Video Quality Assessment Methods for Multimedia Applications", *ITU-T Rec. P.910*, 2008.

[16] Int. Telecommunication Union, "Subjective Testing Methodology for Web Browsing", *ITU-T Rec. P.1501*, 2013.

[17] S. Möller, "Assessment and Prediction of Speech Quality in Telecommunications", *Springer - 1st edition*, 2000.