Title:
Factors influencing quality of experience of commonly used mobile applications

Author:
Selim Ickin, Katarzyna Wac, Markus Fiedler, Lucjan Janowski, Jin-Hyuk Hong, Anind K. Dey

Journal:
IEEE Communications Magazine

Year:
2012
Vol:
50
Issue:
4
Pagination:
48 - 56
URL/DOI to the paper:
10.1109/MCOM.2012.6178833

Factors Influencing Quality of Experience of Commonly-Used Mobile Applications

Selim Ickin, *Student Member, IEEE,* Katarzyna Wac, *Member, IEEE,* Markus Fiedler, *Member, IEEE,* Lucjan Janowski, Jin-Hyuk Hong, *Member, IEEE,* and Anind K. Dey

Abstract—Increasingly, we use mobile applications and services in our daily life activities, to support our needs for information, communication or leisure. However, user acceptance of a mobile application depends on at least two conditions; the application's perceived experience and the appropriateness of the application to the user's context and needs. Yet, we have a weak understanding of a mobile user's Quality of Experience (QoE) and the factors influencing it. This paper presents 4 week long, 29 Android phone users study, where we collected both QoE and underlying network's Quality of Service (QoS) measures through a combination of user, application and network data on the user's phones. We aimed to derive and improve the understanding of users' QoE for a set of widely used mobile applications in users' natural environments and different daily context. We present data acquired in the study and discuss implications for mobile applications design.

1 Introduction

The growing availability of diverse interactive mobile applications, assisting us in different domains of daily life, make their perceived QoE increasingly critical to their acceptance. However, to date, evaluation of QoE has mainly focused on an applications' usability [1], which is evaluated in studies conducted for a limited time in controlled laboratory environments, under the conditions that do not resemble users' natural daily environments. The results of such evaluations help to discover serious and immediate usability issues, but they are unlikely to help in recovering issues that are relevant to real-life situations outside the lab.

These real-life issues involve, amongst others, a non-deterministic QoS [4], and in particular, the performance of the underlying networks' infrastructures supporting the execution of mobile application. The QoS is usually provided at the 'best-effort' level; *i.e.,* without any guarantee by a provider upon its performance. Yet QoS can be critical to the user's QoE, especially for highly interactive mobile applications, which delivery depends on frequent data transfers over the underlying infrastructures.

A common practice for QoE provisioning is that mobile application designers use their own judgment and perception of an application's use as a bellwether to gauge an application's perceived experience [1]. This causes users, whose QoE expectations are not satisfied, to give up using the applications or to switch to another provider. It is estimated that more than half of 200 new Apple's iPhone applications available daily, do not achieve a critical mass of user acceptance and are withdrawn from the store's offer within months from the launch.

The challenge for designers and researchers studying new mobile applications is that no robust scientific methods exist for evaluating applications' perceived QoE in the user's natural environments. Rather, there are qualitative methods for usability evaluation in the Human Computer Interaction (HCI) community [1], and there are quantitative methods for the evaluation of the QoS and performance of the underlying network infrastructures in the data networking community [2]. Due to the dichotomy between these approaches, there are no robust methodologies that combine both types of methods. Our approach is to measure QoE and QoS through a combination of methods with a goal of improving our understanding of factors influencing QoE, and enabling us to derive implications for mobile application design and QoE management.

In [11], which was published in the 3rd week of the ongoing user study, the authors have presented the overview and then the details of the methodology, followed by a very short presentation of the results for a study-in-progress. In this paper, we shortly present the methodology, and then expand the analysis and discussion of the study results in Sections 4-6.

Section 2 of this paper presents related works, 3 - methodology, 4 - results, 5 - factors influencing user's QoE, and Section 6 discusses the role of QoS. Section 7 discusses the results, including the implications for application design, and finally, Section 8 concludes the work and outlines the future work areas.

• *S. Ickin and M. Fiedler are with BTH, Sweden, e-mail: {sic, mfb}@bth.se. K. Wac is with UniGe, Switzerland, e-mail: katarzyna.wac@unige.ch. L. Janowski is with AGH, Poland, e-mail: janowski@kt.agh.edu.pl. J. H. Hong and A. K. Dey are with CMU, USA, e-mail: {hjinh, anind}@cs.cmu.edu.*

This work was partially supported by the US NSF FieldStream (0910754), EU FP7 STREP PERIMETER (FP7-224024), Swiss (SSER C08.0025, NSF PBGEP2-125917, SCIEX 10.076) and EU AAL TraiNutri and WayFiS projects.

2 Related Work

With respect to mobile user experience, according to Jaroucheh *et al.* [10], while modeling usage of pervasive technology and the resulting user's QoE, one should consider the historical as well as current user context and the flexibility of user behavior depending on this context. Similarly, according to Hassenzahl and Tractinsky [9] user experience is influenced by user's internal state, the characteristics of the designed system, the context within which the interaction occurs, and the meaningfulness of the activity. Korhonen *et al.* [12] surveyed mobile users and concluded that the mobile device, task at hand and social context are the most influencing factors of the user's QoE. Similarly, Park *et al.* [13] indicated usability, usefulness and affect as the factors, and Shin *et al.* [14] added enjoyment, as well as network access quality to this list. Reichl *et al.* [3] attempt to evaluate QoE in real user environments by capturing user interaction with the mobile phone, as well as the user context using two different cameras mounted on a large female hat worn by the user. The approach also includes the acquisition of QoS measures on the mobile phone and relating these to QoE. The authors indicate some design implications for video streaming applications.

The most important difference between the existing studies and ours is the fact, that we focus on measuring users' perceived experience in a minimally obtrusive manner on users' (personal) phones, for a set of mobile applications used in their natural daily environments, and we aim to increase the understanding of factors influencing this experience.

3 Methodology

This section presents the design of our methodology, which is detailed further in Wac *et al.* [11].

3.1 Overview

We use mixed methods, incorporating qualitative and quantitative methods in 4 week long user study i) employing a continuous, automatic, unobtrusive context data collection on the user's mobile phone through Context Sensing Software (CSS) application, ii) gathering user feedback on the perceived QoE via an Experience Sampling Method (ESM) executed multiple times per day, and iii) a weekly interview with the user along the Day Reconstruction Method (DRM) method. We focus on already implemented and operational interactive mobile applications, which are available and commonly-used on a typical smartphone.

3.2 Mobile Users and Data Collection

Recruitment: Online Survey

To recruit the participants, we designed an online survey for mobile users to find out how long they have been using mobile, what phone type they have, which network provider they are connected through, what usual usage pattern for voice and data communication they follow, which applications they use, and what users experience in general with the current smartphone. Moreover, we got their socio-economic status. The major criteria for the selection of 30 participants was to own an Android smartphone and to use it frequently in various conditions in daily life. Therefore, 30 users were selected randomly from 430 potential candidates.

Android Sensor Logs: CSS Application

The CSS application unobtrusively collects the information from users' Android phones, like cellular network, Bluetooth, Wireless Fidelity (WiFi) connectivity and its signal strength (RSSI); sent/received data throughput (KB/s), number of calls and SMS's, acceleration, screen orientation and brightness, running applications and user location. Most of the data is collected only when the sensor value changes, *i.e.,* the Android Operating System (OS) updates the CSS with data. As the QoS indicator, we focus on an interactivity of a mobile application and therefore we measure the median Round Trip Time (RTT) for an application-level control message (64B), sent every 3 minutes from the mobile device through the available wireless access network technology to a dedicated server, which is deployed at our university. Simultaneously, Server Response Time (SRT), which is calculated as the time it takes to get an The Hypertext Transfer Protocol (HTTP) request-response with an updated weather information, of the Android smartphone to a dedicated weather application server was also monitored. The CSS sensor logs are immediately written to the phone storage card to minimize the memory allocation on the phone throughout the data collection process, as well as to minimize the risk of data loss. Further details on CSS are provided in [11].

QoE Ratings & Context Logs: ESM

We have employed ESM [5] to gather users' QoE ratings. We have implemented this in the form of a short, mobile-device based survey, which is presented to him/her after using an application. By this way, it's aimed to not influence the experience and behavior of a user. Therefore the survey does not appear after each application usage, but at random times after a random application usage, with a maximum of 8-12 surveys per day. The survey poses questions about i) rated application QoE (poor (1) to excellent (5) based on the ITU recommendation [6] for Mean Opinion Score (MOS), ii) location (home, office/school, street, other indoor, other outdoor), iii) social context (alone, with a person, or with a group), and iv) mobility level (sitting, standing, walking, driving, other). While rating the same application throughout the study or even for a given day, we requested that users do their best at providing independent ratings, while keeping in mind that a rating is a QoE purely subjective, episodic assessment provided on the basis of the given perception of the specific episode of application use. We aimed to capture QoE for a set of widely available mobile applications for entertainment, communication or information purposes such as Internet-based radio, web browsing, online games, video streaming, email, and news. In total, it takes approximately 5 seconds for the user to complete each ESM. According to our tests confirmed by the users, neither CSS nor ESM logs influenced negatively the performance of their smartphone.

Weekly Interview: DRM
We have followed the DRM method [7], inorder to analyze possible relations and causality between QoE ratings, QoS, and user context, we have interviewed the users on a weekly basis regarding their usage patterns and experience on the mobile applications along their previous 24-hour period. During the interview, users explained their responses from the ESM, and these results were compared to the visualized CSS and ESM data logs from the smartphone. This method has been used for fast identification of any inconsistencies between the collected data by CSS and the DRM. This way we could also identify the factors influencing QoE for this particular user.

3.3 Study Participants & Collected Data

The study was conducted for 4 consecutive weeks in February-March 2011. We have recruited 31 Android users with three types of Android OS phones (Motorola, HTC and Samsung), which are subscribed to four providers (Verizon (23 participants), Sprint (4 participants), T-Mobile (3 participants) and AT&T (1 participant)). Two participants dropped out in the first three days of the study due to the battery issues on their older phones (S1, S27). S11 collected only one week of data and then dropped due to an inconvenience. Participants S2, S8 and S9 experienced data logging outages due to malfunctioning software on their phone, or an explicit altering of the logging. Fig. 1 presents participants: (from left): Participant ID(S), gender, profession, phone type, age range, overall MOS as reported in the online survey, overall MOS perception as derived from the study and the percentage of its occurrence within all the collected MOS, number of occurrences of low MOS values: 2 and 1 that are separated by comma, and the total number of MOS ratings collected by the user in the study. None of the participants had accessibility problems related to their phone use and, when asked, none of them admitted that they are adversely affected by the beliefs regarding Electron Magnetic Resonance (EMR) health issues for mobile phone usage. We have collected data logs for 17′699 hours in total, which represent 87.8% of the hours for the overall study duration of 28 days. On average we have missed 84.9 h (*i.e.,* 12.5 %) per a participant (min. 3 h, max. 378 h).

(insert Figure 1 around here)

4 Results for QoE & Context (ESM)

In this section we present results collected for QoE ratings and the user context.

QoE Ratings
In total, we collected 7804 QoE ratings from all the participants with an average of 9.29 ratings per day and per participant. The high ratings (4, 5) are much more frequent than low ones (1, 2) as illustrated in Fig. 2. In general, the participants find their QoE acceptable; they have explained that they learned how to maximize their mobile application usage along their routine activities. The participants exhibited knowledge on circumstances that they can expect particular QoE depending on network coverage.

(insert Figure 2 around here)

Applications
Amongst the applications, for which we have collected QoE logs, there were standard Android-built-in applications, *e.g.,* web or email applications, as well as a variety of specialized ones. After gathering all data from the study, we have identified (from Android market) the following 13 categories of mobile applications (presented in order of descending frequency of usage): 1) communication: talk, skype, fmail, email, gtalk; 2) web: default browser, dolphin; 3) social network applications: okcupid, cooliris, foursquare, facebook, twitter, foursquared, Tumblr, touiteur; 4) productivity tools: astrid, sandbox, calendar, shuffle, callmeter, outofmilk; 5) weather apps: weather, weather service, weatherservice, weathercachingprovider, weathercacheprovider; 6) news: espn, sports, news, penguinsmobile, foxnews, penguinsMob, reddit, newsfox, pittFight; 7) multimedia streaming: listen, youtube, pandora, lastfm; 8) games: Worldwar, WoW, games, poker, zyngawords, words, touchdown; 9) lifestyle apps: horoscope, sparkpeople, diet; 10) finance: stock; 11) shopping: ebay, coupons, starbuckscard, craigslist, starbucks; 12) travel: navigator, maps, locationlistprovider; and 13) other applications.

Context
The applications were used mostly at 'home' then 'office/school' and 'indoor/other' as depicted in Fig. 3. With respect to the social context, the applications were used on average 80% of the time when the user was 'alone'. The mobility levels dominant for an application usage were 'sitting', 'standing', and 'other' specified by the user precisely as 'bed'. Low QoE ratings occur mostly when person is at 'home' or 'school' while being 'alone' and 'sitting'.

(insert Figure 3 around here)

5 Factors Influencing QoE

To derive the factors influencing user's QoE, a word-cloud was created that visualizes the word frequency in user's expressions as a weighted list. Font sizes are set in relation to the frequency of the corresponding word as illustrated in Fig. 4. We have used all the 376 expressions from our 29 user's weekly DRM interviews, as well as data from online survey with 430 entries. Most frequently mentioned key words are applications, mobility, internet, battery, performance (*e.g.,* 'slow', 'freeze') as well as features *e.g.,* , Global Positioning System (GPS), camera, and flash player. Many of these, are difficult to evaluate automatically. We have grouped these words into clusters by using affinity clustering method, which we then have labelled along the identified factor. The coding and grouping of words into clusters have been done by two independent people, and their measure of agreement was 90 %. The most disagreements were related to the interface and the application performance clusters, especially for the cases where the participants were not clear on pinpointing the main issue. Finally, we have distinguished the following factors influencing the user's QoE.

(insert Figure 4 around here)

Application Interface's Design
Application Interface and interaction was mentioned very often; users did not like the position and location of the keys on the smartphone screen, they had difficulty with resizing, web-page scrolling, they did not agree with built-in dictionary items, and they complained about inefficient manual input, *e.g.,* 'fat finger' problem. Some users preferred interacting with a web-based version interface of particular applications, *e.g.,* Facebook, rather that with its widget.

Application Performance
'Freeze', 'sloppy', 'sluggish', 'speed', 'performance', 'usage of memory', and 'sdcard' were the expressions used when referring to a low application performance. Especially, for mobile applications that users previously experienced on a fixed PC, the expectations for performance were high that resulted in low QoE.

For those users, who have had a PC as an alternative device, *e.g.,* to receive/send emails, their rated QoE were limited to reception of emails on the smartphone since they have used the PC for sending emails. The reason for this is that typing on a real keyboard of the PC provides a better experience, especially for long messages. On the other hand, some users preferred a smartphone to run most of the applications. For those users, mobile applications achieved enough usability to enable them not to use a larger and potentially more comfortable PC. In the future, we wish to understand the characteristics of these users and the applications they use on a mobile and PC better.

Two particular participants were adaptive, but at the same time complained about the applications, but had to use it: S4 complained about 'stupid' autocorrect function of messaging, and S16 had to use a specific VPN application with low usability to access a corporate email.

Some mobile users expressed their tolerance, *e.g.,* for worse application's performance when they use it 'on the move'. In parallel, users regularly mixed a network performance with the application performance metrics, *e.g.,* while saying *"Skyping service is incredibly spotty"*, the concern is actually the underlying network connectivity of Skype service, not the application itself.

Battery
Battery efficiency consistently influenced the experience of the mobile users, as it limited their phone usage, especially at the end of the day, when the phone was completely discharged. S12 resolved that issue by carrying an additional battery.

Phone Features
Mobile users noted missing features of the phone, which then hindered their experience, *e.g.,* lack of flash player, lack of personalized alarm clock, lack of special settings for vibrate-only mode, lack of or a faulty GPS, lack of features for privacy settings.

Apps and Data Connectivity Cost
In the online survey, many of the mobile users indicated that the cost of applications and data usage prohibits them from experiencing these applications.

User's Routine
Routine of the user implied that, different sets of applications were used in the morning, in the evening before going to sleep, in the car, and outside the office. The user rating is influenced by user's environment and the importance of the mobile application to the task at hand.

User's Lifestyle
There were highly-ranked applications that supports user's lifestyle choice, *e.g.,* sports, fashion, nutrition, and leisure. They are used on a smartphone due to their convenience of usage while, *e.g.,* being in the gym (for logging of burnt calories), in the cafeteria (for logging caloric intake), or on the street while trying to find a fancy restaurant.

6 Role of QoS

Along the data analysis, we realized that we do not have many evidences of an influence of QoS on the user's QoE. There exist low QoE ratings in our data, however, there are no strong evidence for a low QoS. One of the reasons for that can be, as we have indicated in the previous section, the fact that factors influencing the user's low QoE are different than QoS. Another reason for that can be insufficient granularity of our QoS measurements with limited permissions, *i.e.,* non-rooted phones, to access network level metrics on user phones.

With respect to the QoS, it is influenced by the choice of the wireless access technology, *i.e.*, WLAN, 2.5G, 3G or 4G; therefore influencing QoE. In our study, we observe however that the performance of the access network is not an issue, as users are well-connected and have a choice of networks (as ordered by an increasing nominal capacity): GPRS, $1xRTT$, CDMA, EDGE, UMTS, $EVDO_0$, $EVDO_A$, and HSPA. WiMaX was available for selected users in selected urban locations: S12, S18, S29, and S24. Some number of users, were connected over WiFi at home and office, in order to ensure better QoS. We have observed the diversity in the connectivity through WiFi interface amongst the users, ranging from 0 to 398.5 hours. In total, nine participants have never used WiFi during the study, while six participants never turned the interface OFF that allows the smartphone to connect to any available Wireless Local Area Network (WLAN) access point when detected. We have attempted to investigate the influence of the WiFi signal strength on MOS, however no clear trends were observed.

(insert Figure 5 around here)

The overal mean of RTT values that are collected in the study is 231 ms with a standard deviation of 73 ms. However, it differs per a MOS level (Fig. 5 presents the mean and the standard deviation of SRT and RTT). From the figure, we observe that higher values of SRT and RTT correspond to lower MOS. For both measures, the confidence intervals gets narrower as the MOS increases. The fluctuation in these measures observed for low ratings, especially for MOS values 1 and 2, may be related to the influence of these measures on the application performance. From Fig. 5, we conclude that the recommended server response time for a mobile application assuring the MOS level of 3 is 950 ms, the response time of 850 ms corresponds to the MOS level of 5 and high user experience.

Turning to throughput, *i.e.,* bytes received per second by the smartphone, in Fig. 6 we observe the relation between its average value and the MOS levels; the mean throughput increases with the increase in MOS values, and the confidence interval increase as well. This means that there are many different throughput ranges resulting in the same MOS level.

(insert Figure 6 around here)

The applications, for which there were many low QoE ratings, were streaming multimedia applications like *listen*, *youtube* and *pandora* - audio feeds application, video streaming and real-time radio streaming accordingly. Any participant in our study was using one of these applications in average 1.67 h per day, which involved 50 MB of download traffic and 1.7 MB of uplink traffic per day. A total of 1.38 GB for downlink and 0.5 GB for uplink traffic is observed in 28 days. A participant was running *listen* application in average for 0.8 h, *youtube* for 0.34 h, *pandora* for 0.5 h per day, although most of the usage was observed within a fixed group of 10 participants with a a varying distrubution of population. The applications, *listen*, *youtube* and *pandora* involved around 32.7 MB, 8.15 MB, and 8.36 MB of downlink traffic per day, and 1.03 MB, 0.3 MB, 0.36 MB of uplink traffic, correspondingly. Some of the mobile users using these applications, rated '1' for MOS, being critical of its performance, but some other users were tolerant, knowing that they gain possibility of accessing these applications while being mobile, and paying a performance 'price' for that. Low ratings can be related to the fact, that these applications require high network capacity or, as we learned from participants, some application widgets are buggy and influenced the application performance. Additionally, some participants were launching pre-loaded MPEG-2 Audio Layer III (MP3)s when driving, in the bus, or walking outside.

The 4G (WiMAX) service was rarely used because of its unavailability, as S18 claimed: *"Unfortunately, I don't get 4G in (A). And when I'm in (B), the 4G connection keeps switching on and off, and the notifications are just annoying. So I keep 4G switched off"*. Another (S20) said: *"My phone can operate on a 4G network, but I usually keep it set to 3G because in my experience, the 4G is not considerably faster and just eats up my battery... Generally I keep 4G turned off unless I am doing something network intensive and I know it is available"*. We were surprised to hear that, because according to the results of performance measurements we conducted for the 3G and 4G networks, use of the 4G network results in better QoS parameters than 3G. We presume that the applications used by this particular user, worked sufficiently well on 3G. It seems to contradict with our initial hypothesis that a mobile user always wishes to have the best possible and fastest service. Our future work includes analysis of the data to support or refute this hypothesis.

Additionally, using our DRM, we discovered clusters of users, for which network selection and the resulting QoS depended directly on their phone charging behaviors. Namely, users who were able and willing to charge their phones often preferred the access technologies in the order WiFi-$4G - 3G$, whilst this changes to $3G-$WiFi$-4G$ for the ones who charge their phones less often. A common feeling among our users was that 4G was as good as WLAN but drained too much battery. In addition, 4G coverage is a problem, thus users who are subscribed to providers with 4G support are not necessarily always within the 4G coverage, which leads to the connectivity oscillations between 3G and 4G, resulting in draining extra battery and putting users at risk of instant disconnections.

7 Study Limitations

We recognize study limitations influencing the generalization of results. Firstly, the study has been conducted on self-selective group of mobile users, with self-selective set of applications. We have analyzed the QoE ratings of participants in our study and we have identified some participants who use many applications in a short time. The maximum number of applications that were used within the last 3 minutes of the given QoE rating were 19. We observed that the applications are rated with high MOS values, *e.g.,* 4 and 5, as more number of applications are used within the 3 minutes time period. We hypothesize that either such a user is more advanced and uses many mobile applications, from which he/she is satisfied, or by using so many applications in a short time, this user has no time to pay attention to details of his/her low experience, just grasping the essence of, *e.g.,* information provided by the application for the given situation at hand. It would also be more contributive, if it were possible to clearly pinpoint the type of the access network technologies, *e.g.,* $3G$, $4G$ in relation to the MOS values. In addition, we were not able to capture extreme conditions, *i.e.,* extremely high RTT or SRT values due to lack of granularity of the selected samples for the study. Capturing worst conditions in real-life scenarios is challenging and non-trivial, therefore longer time of the study could be suggested in order to capture all scenarios a user can experience. Moreover, there would have been a wider variety of measurable QoS network metrics, *e.g.,* delay jitter and packet loss, if the smartphones were rooted and privacy concerns were not important.

8 Conclusions and Future Work

In this paper, we have presented our research towards understanding a mobile user's Quality of Experience (QoE) in their natural daily environments. Our approach is a blend of both quantitative and qualitative procedures, where the user becomes an active participant in the research. Firstly, it requires gathering in-situ spontaneous information about the user's mobile experience for a set of widely used mobile applications, by employing the Experience Sampling Method (ESM) for interaction with the user directly after each mobile application usage. Secondly, it requires a retrospective analysis of the user's experience and of the state of factors influencing it, by employing the Day Reconstruction Method (DRM) to assist with the recollection of the past 24 hours. DRM was supportive in validating the collected data through CSS application. In this paper, we have presented an analysis of the collected data by highlighting some factors that impact the user's QoE.

The novelty relates to the factors that influence QoE, including application interface's design, application performance, battery efficiency, phone features, application and connectivity cost, user's routines and the user's lifestyle. These factors go beyond the 'usual' usability, usefulness and user value factors, as already indicated in the literature. Role of Quality of Service (QoS) was also studied that indicates the increase in Server Response Time (SRT) and Round Trip Time (RTT) values and decrease in received throughput reduces the Mean Opinion Score (MOS) values, however, in general, we observed that the users are well-connected and they use their applications mainly while in fixed indoor position and being alone. The recommended SRT for a mobile application assuring the MOS level of 3 is 950 ms.

Our future work includes analysis of other factors and user context influencing this experience and 'grounding them' [8] via additional user studies. The implications for design based factors are numerous; the QoE-management for mobile applications and services must be multi-faceted, and can not only focus on maximizing the QoS of the underlying network infrastructures but also the application's interface design, application performance and mobile device performance factors, given the user's expectations and tasks at hand. In addition, machine learning techniques can be applied to the user data with carefully chosen attributes and appropriate algorithms towards accurate prediction of user's QoE in different context.

References

[1] A. Dix, et al. *Human Computer Interaction*. Prentice Hall, 2004.

[2] K. Salamatian and S. Fdida. Measurement based modelling of quality of service in the internet: A methodological approach. In *Proc. Workshop on Digital Communications, 2001, pp. 158-174.*

[3] P. Reichl, P. Frehlich, L. Baillie, R. Schatz, and A. Dantcheva. The liliput prototype: A wearable test environment for mobile telecommunication applications. In *Proc. CHI 2007, pp.1833-1838.*

[4] In *ITU-T, Definitions of terms related to quality of service. Recommendation E.800, 2008.*

[5] J. M. Hektner and et al. Experience sampling method: Measuring the quality of everyday life. In *Sage Publications Inc., 2006.*

[6] In *ITU-R, Method for the Subjective Assessment of Intermediate Quality Level of Coding Systems, Recommendation BS.1534-1, 2003.*

[7] D. Kahneman and et al. A Survey Method for Characterizing Daily Life Experience: The Day Reconstruction Method. In *Science, 2006: 306(5702), pp. 1776-1780.*

[8] H. van Kranenburg, et al. Grounded contextual reasoning enabling innovative mobile services. In *Proc. ASWN 2005, France.*

[9] M. Hassenzahl and Tractinsky N. User experience - a research agenda. In *Behaviour and Information Technology, 2006.*

[10] Z. Jaroucheh, X. Liu, and S. Smith. Recognize contextual situation in pervasive environments using process mining techniques. In *J of Ambient Intelligence and Humanized Computing, Springer-Verlag, 2010: 2(1), pp. 53-61.*

[11] K. Wac, S. Ickin, J. Hong, L. Janowski, M. Fiedler, A.K. Dey. Studying the Experience of Mobile Applications Used in Different Context of Daily Life. In *Proc. ACM SIGCOMM Workshop W-MUST 2011, pp.7-12.*

[12] Korhonen, H., et al. Analysing user experience of personal mobile products through contextual factors. In *Proc. ACM ICMUM 2010, pp.1-10.*

[13] Park, J., et al. Developing elements of user experience for mobile phones and services. In *Human Factors and Ergonomics in Manufacturing & Service Industries 2011.*

[14] Shin, Y., et al. Examining influencing factors of post-adoption usage of mobile internet. In *Information Systems Frontiers 2011: 12(5), pp. 595-606.*

Bio's

Selim Ickin is a $Ph.D.$ candidate in School of Computing department at Blekinge Institute of Technology

(BTH), Sweden. He has involved in EU STREP FP7 PERIMETER project. He has his B.Sc. degree from the department of Electrical and Electronics Engineering, Bilkent University and M.Sc. degree from the Electrical Engineering department, BTH. He has worked in Nortel Networks and completed his internships in Siemens AB and Schneider Electric.

Katarzyna Wac is since 2010 with the Institute of Services Science of University of Geneva (Switzerland). She holds a BSc and MSc in Computer Science, MSc in Telematics, and PhD in Information Systems. In 2003-2004 she was a researcher at University of Twente, the Netherlands and 2004-2009 at University of Geneva. In 2009-2010 she was on leave at Carnegie Mellon University, Human-Computer Institute (USA). She researches pervasive mobile computing and communication technologies applied in provision of healthcare services.

Lucjan Janowski is an assistant professor at AGH University of Science and Technology. He received his Ph.D. degree in Telecommunications in 2006 from the AGH. During 2007 he worked on a post-doc position in CNRS-LAAS in France. In 2010/2011 he spent half a year on a post-doc position in University of Geneva working on QoE for health applicaitons. His main interests are statistics and probabilistic modelling of subjects and subjective rates used in QoE evaluation.

Prof. Dr.-Ing. Markus Fiedler is holding a PhD degree in Electrical Engineering with focus on ICT and a Docent degree in Telecommunication Systems. Starting from modelling and analysis of traffic in networks, he has been focusing on QoE during the recent years. His interests and experience include performance evaluation, user perception of waiting times, QoE models, security, and instrumented user interfaces. He is Future Internet Cluster co-chair in the ICT domain of the EC.

Jin-Hyuk Hong is a Postdoctoral in the Human-Computer Interaction Institute at Carnegie Mellon University. He performs research at the intersection of machine learning and intelligent systems, and has published over 50 papers on these topics. His current research interests include the use of embedded and mobile sensors to understand human behavior, and the design of novel techniques to build intelligent user interfaces and context-aware systems. He has received the B.S., M.S., and Ph.D. degrees in computer science from Yonsei University, Seoul, Korea, in 2002, 2004, and 2009, respectively.

Anind K. Dey is an Associate Professor in the Human-Computer Interaction Institute at Carnegie Mellon University, and is director of the Ubicomp Lab. He conducts research at the intersection of human-computer interaction, machine learning and ubiquitous computing, and has published over 100 papers in these areas. He serves on the editorial board of IEEE Pervasive and the Personal and Ubiquitous Computing Journal, among others. Anind received his PhD in computer science from Georgia Tech.

S	Gender	Profession	Phone Type	Age	QoE - Survey	QoE - Study	MOS= 2,1	Total No of Rat.
2	M	Customer service	Samsung Captivate	18-24	5	4(47%)	4, 4	218
3	M	Owner, moving company	Motorola Droid	25-35	2	4(61%)	3, 0	181
4	M	Driver	MyTouch 4G	25-35	4	5(77%)	4, 0	233
5	F	Research assistant	HTC Incredible	18-24	5	4(79%)	5, 2	227
6	F	Admin. higher educ.	G2	25-35	4	5(52%)	1, 0	323
7	F	ICT Consultant	Motorola Droid X	25-35	5	5(89%)	5, 1	390
8	M	Web developer	Motorola Droid	25-35	5	4(54%)	4, 0	143
9	F	Medical adm. assis	Motorola Droid	25-35	5	5(66%)	4, 0	197
10	F	Nanny	HTC Incredible	25-35	5	4(60%)	4, 0	543
11	F	Unemployed	Sam. Vib. Galaxy-S	25-35	5	5(68%)	3, 1	62
12	M	Unemployed	HTV Evo (WiMAX)	36-45	4	5(78%)	3, 6	620
13	M	Uni. program mngt	Motorola Droid	25-35	5	4(35%)	25, 3	254
14	M	Contractor	Motorola Droid X	25-35	4	5(63%)	8, 9	369
15	M	Accounts coord.	Motorola Droid	25-35	4	4(84%)	4, 1	196
16	F	Operations analyst	Motorola Droid X	25-35	5	5(57%)	7, 4	327
17	M	System analyst	Motorola Droid	36-45	5	5(48%)	4, 5	240
18	M	ICT consultant	HTC Evo (WiMAX)	25-35	4	5(62%)	5, 0	209
19	M	Teacher	Motorola Droid	25-35	4	5(68%)	4, 18	317
20	F	Admin. assistant	HTC Evo (WiMAX)	25-35	4	4(97%)	1, 1	296
21	M	Univ. student	Motorola Droid	25-35	4	3(57%)	7, 1	195
22	M	Grant admin	HTC Incredible	25-35	5	5(43%)	10, 1	276
23	M	Graduate student	Motorola Droid 2	25-35	2	4(83%)	1, 2	137
24	M	Systems analyst	HTC Evo (WiMAX)	25-35	4	5(94%)	2, 0	198
25	F	Univ. student	Motorola Droid 2	18-24	5	4(63%)	16, 10	386
26	M	Senior adm assis	Motorola Droid	25-35	5	4(55%)	11, 20	253
28	M	Graduate student	Motorola Droid X	25-35	5	5(30%)	34, 15	251
29	M	Paramedic	Motorola Droid	36-45	4	4(48%)	33, 12	213
30	F	Housewife	Motorola Droid X	36-45	5	5(83%)	3, 1	341
31	M	Registered Nurse	Samsung Captivate	25-35	5	4(52%)	1,1	209

Fig. 1: Participants: Demographics and QoE Ratings (S1,S27 dropped out)

*

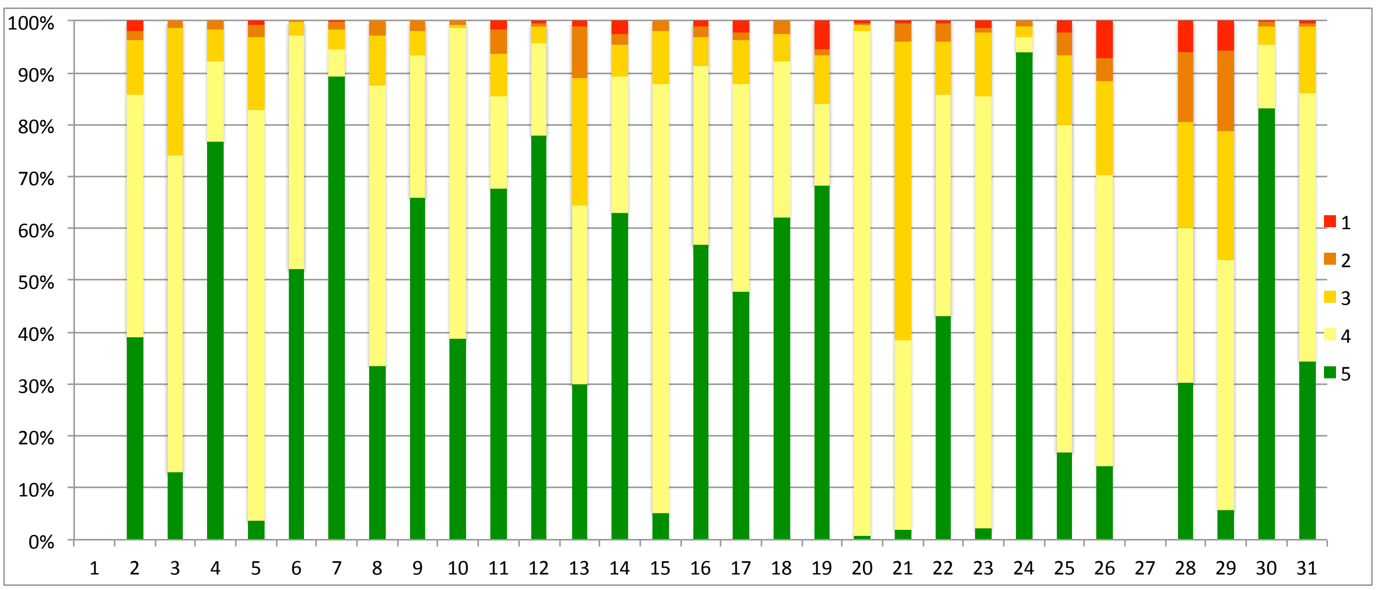

Fig. 2: QoE Ratings Distribution for Mobile Users.

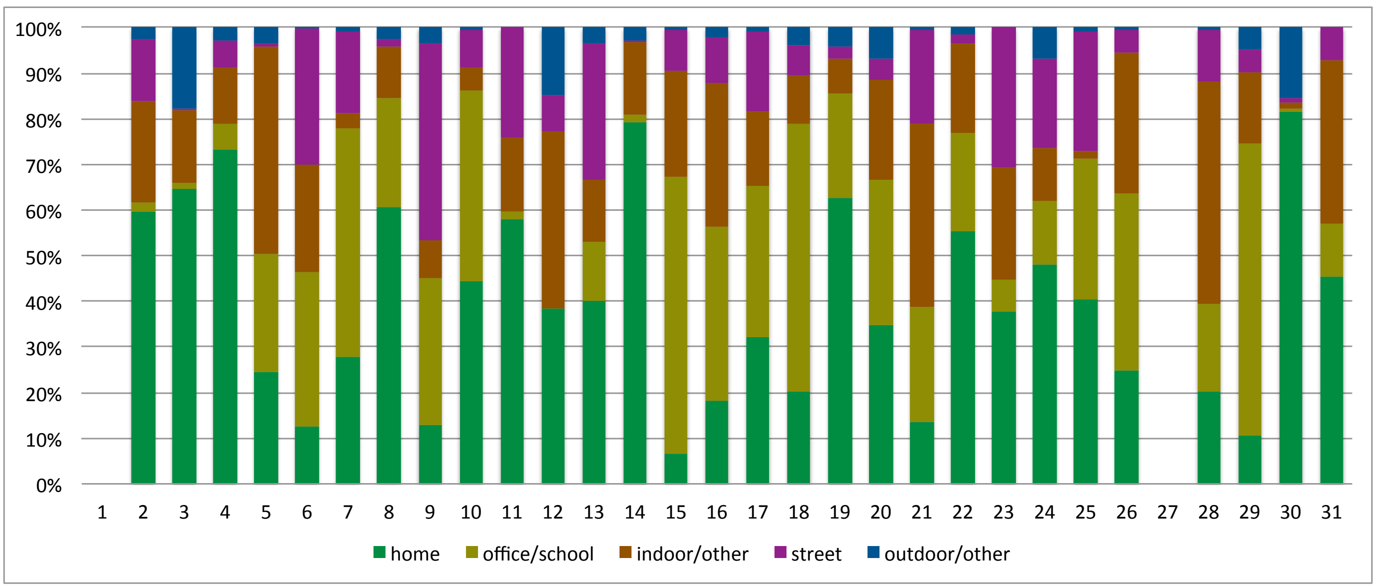

Fig. 3: Mobile Users' Locations Distribution.

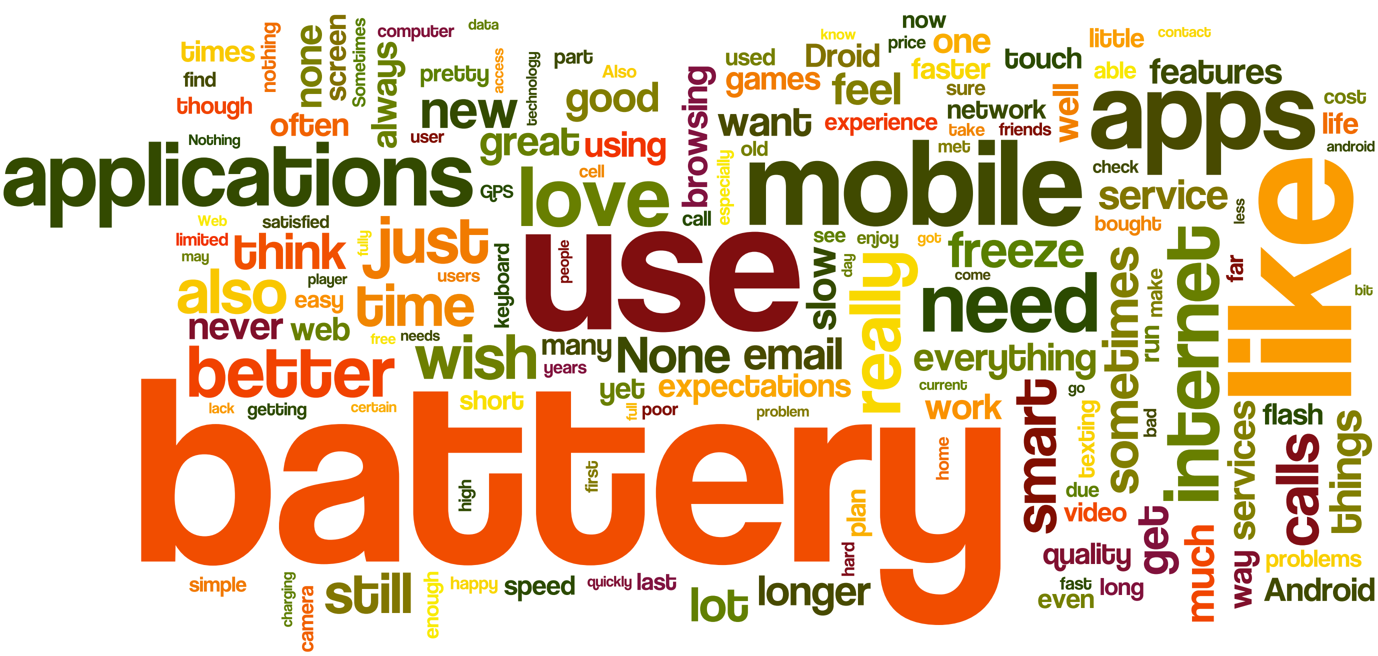

Fig. 4: Expressions in User's Interviews and Surveys

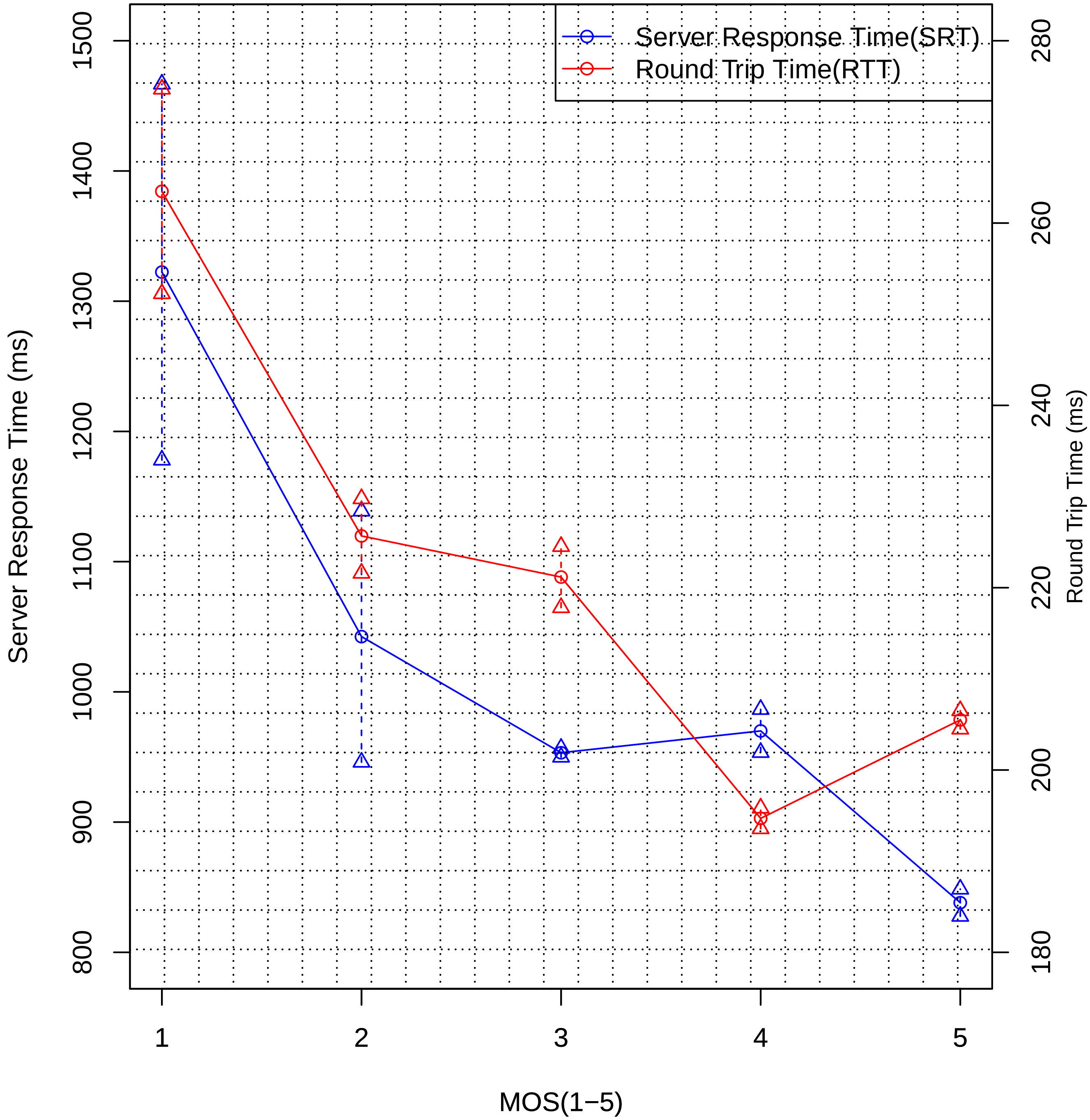

Fig. 5: Mean RTT and SRT (97.5% CI) vs. MOS.

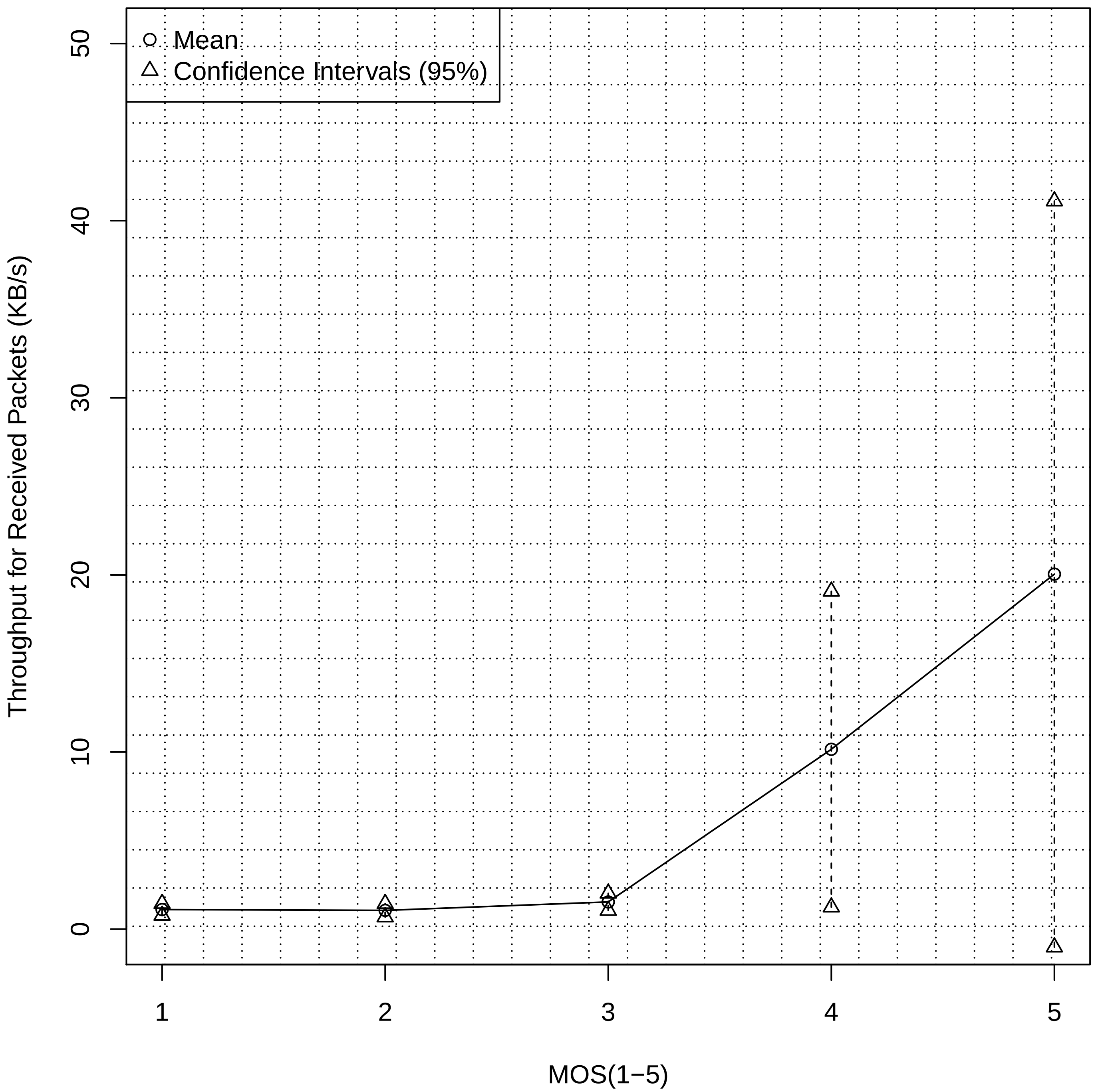

Fig. 6: Mean received throughput (97.5% CI) vs. MOS.