Characterization of Mobile Web Quality of Experience using a non-intrusive, context-aware, mobile-to-cloud system approach

by

Elias Allayiotis

A thesis submitted in partial fulfilment for the requirements for the degree of Doctor of Philosophy at the University of Central Lancashire

June 2017

University of Central Lancashire

STUDENT DECLARATION FORM

I declare that while registered as a candidate for the research degree, I have not been a registered candidate or enrolled student for another award of the University or other academic or professional institution.

I declare that no material contained in the thesis has been used in any other submission for an academic award and is solely my own work.

Signature of Candidate

Type of Award **Doctor of Philosophy**

School **School of Computing, Engineering and Physical Sciences**

Abstract

This study presents a modelling approach for quantifying the Mobile Web Quality of Experience (MWQoE). It builds on current QoE and Web QoE research, and by fusing together data that is available on modern mobile devices, constructs a novel MWQoE model that is user-centered, context-aware and non-intrusive (does not depend on user feedback). This study identifies the factors which affect Web QoE and measures their effect on it in mobile scenarios. Moreover, this study explores scenarios in which Web QoE can be effectively characterized and enhanced, delivering a novel Mobile-to-Cloud system for the continuous evaluation of MWQoE in real-world environments. The significance of defining and evaluating MWQoE is identified. Specifically, MWQoE can be used by online providers to uncover customer insights and illustrate how the experience in using their products is perceived by their customers. In fact, MWQoE can be considered an important key performance indicator showing the technology acceptance or *satisfiability* of customers for a specific web product or service.

The MWQoE model uses Context Spaces Theory together with Bayesian Networks and works under uncertainty. It is derived organically from collected live user data and is evaluated by using context data from different domains (User, Device, Network) together with Web usage metrics (such as Page Load Timings, HTTP Request Size, etc.). The MWQoE metric uses Utility Theory and delivers a quantified characterization of MWQoE on a single scale for specific scenarios. These scenarios and the effect of their attributes to MWQoE are evaluated by actual user responses from a lab experiment. Results show that the MWQoE model provides either the same or more conservative characterizations of the user's experience in 93% of all tests, a feature that works for the benefit of the user, and in general, across all scenarios and user profiles, provides a close characterization of user satisfiability (Fair vs Good). In addition, a novel Mobile-to-Cloud system (BetterX) is proposed, designed and implemented, providing the practical and continuous evaluation of the MWQoE metric in real-world environments. This thesis shows that the Mobile Web Quality of Experience evaluation and prediction in using a non-intrusive approach is possible and that the MWQoE model can provide the basis of an enhanced, more accurate model applicable in real-world scenarios which can characterize the user's *satisfiability* in mobile web browsing.

Acknowledgements

This work has been a challenging yet inspiring 3-year journey in which I was fortunate enough to meet, work and collaborate with some truly remarkable people.

First, I would like to express my gratitude to the 2013 PhD scholarship committee for giving me the opportunity to join the program and become part of the research community. Thank you, Dr. Lee Chatfield, Dr. Irene Polycarpou and Dr. Josephina Antoniou for your vote of confidence.

I am very grateful to Prof. Janet Read, for her early contribution of ideas and for her guidance and support throughout the program; and to Dr. Martin Bateman and Dr. Daniel Fitton for allowing me time for discussions and for contributing their invaluable feedback.

A big thank you to my examining committee, Professor Panayiotis Zaphiris, Dr Nearchos Paspallis and Dr Martin Varley for a memorable discussion and an intense, albeit vital debate. Your suggestions and feedback widened the perspective of this thesis and enhanced my work.

A very special appreciation goes out to all the lab participants, both in and out of campus, for contributing their time and helping me complete my experiments; to Emma Hesketh and the UCLan research office for offering much needed resources and for their continuous support to all of us throughout the program; and to the administrative team of UCLan Cyprus for making life much easier in handling logistics, travel itineraries and budgets. Moreover, I would like to thank Amazon Web Services for being kind enough to sponsor my work through their Research Grants program.

To all open source contributors, whose work greatly aided in building various software components, including Piotr Wittchen and his work in Network Events for Android, Jan Odvarko and his work in Firefox Developer Tools and Neal Lathia and his work on Sensor Manager for Android. Thank you for making my development work much easier.

I am forever grateful to all my good friends and my family, whose support and encouragement has always been giving me the strength and confidence to pursue new endeavors and overcome the hardest of challenges in both work and life. I've been very fortunate in life to be surrounded by truly remarkable people whose love, understanding and encouragement has been a personal source of never-ending energy to strive though hardships and shine through solaces. To my cat Phoebe for keeping me company on those never ending nights of reading, writing and fighting with the numbers during data analysis, and for allowing me a comfortable ambience at home. To Eleni, my partner for 15 years and now a good friend: words cannot begin to describe the appreciation and gratitude that I have for you for being there with me, by me, through everything. You've touched everything that I've accomplished in life and for this I am, and will forever be grateful.

Last but (definitely) not least, I would like to thank my advisor, Dr. Josephina Antoniou, an amazing academic, a talented researcher and above all a good friend. I couldn't have done this without Josephina's guidance, enthusiasm and immense knowledge. I cannot imagine having a better advisor and mentor for my PhD study and for that I will be forever indebted. For all those early morning meetings, for the countless cup of coffees we had while discussing this work, and especially for her encouragement and being there for me through hectic schedules, working through trains and plane flights trying to put the final touches before the deadline. I look back at those times with gratification: we've made the destination and the journey was a magnificent one.

Thank you all!

Elias.

Contents

List of Tables

List of Figures

Chapter 1 Introduction

Quality of Experience (QoE) is the field of study which examines and measures the factors that characterize the user's perception of quality in interacting with specific technologies. QoE aims to define *what makes users happy*, i.e. what technology features are better received and why. QoE is very important to content and application providers since it outlines the experience of their clients in using their products, and it is considered an important key performance indicator, as it illustrates *user satisfaction* and *technology acceptance* of technology users.

Current QoE definitions are conceptual and therefore impractical for real-world scenarios while Web QoE models are highly segmented as they deal with only a subset of factors relating to experience. Academia's focus thus far has been primarily on QoE for video and voice adaptations and much less work has been done on Web QoE. Moreover, given that the hardware configurations of mobile devices are completely different from desktops and laptops, and that the mobile user's context is radically different from that of the desktop or laptop user, additional considerations for Web QoE should be included in its definition and evaluation. The aim of this study is to provide a more fine-grained characterization of Mobile Web QoE, one that is specific to selected mobile web browsing scenarios, one that can be used in real-world live settings and one that does not depend on user feedback.

This thesis answers the question of ***how to model, measure and enhance the Web QoE of mobile device users in a practical manner applicable to real-world settings***. In doing so, this thesis hypothesizes that it can use the sensing capabilities of modern mobile devices to capture data in a non-intrusive manner and use context-awareness to measure, model and predict the Web QoE of mobile device users in certain scenarios. To contextualize the aims and hypotheses of the study better, the research question is broken into 3 sub-questions:

"*What are the design and technical considerations of an end-to-end system which can efficiently measure and enhance MWQoE in a non-intrusive manner?*" As stated in Cherubini & Oliver (2009) and in Nakhimovsky (2009), QoE models and methods need to be practical and work in an automated manner. Since the aim of this study is to provide a MWQoE model that can be adopted by industry, it should be applicable in real-world scenarios and work in a non-intrusive way. This study hypothesizes that it can infer the *web session user intention* and the *web session user profile* by using a context-aware approach. This is achieved by fusing together anonymous data from 25 attributes (*Table 6*) and using Context Spaces Theory (Padovitz *et al.*, 2005) together with Bayesian Network (Jensen & Nielsen, 2007) and Utility Theory (Ligeza, 1995) to deliver a MWQoE model and metric in a non-intrusive manner. The context of the device, the network, the user and the web session is inferred using relationships and observations from literature as well as findings from this study's data analysis. In doing so, this study considers the technical aspects of the design and implementation of an end-to-end system to deliver MWQoE, and provides a proof-of-concept implementation.

"*What factors affect MWQoE and how do they affect it?*" As shown in the Related Work section below, there is a plethora of attributes which affect QoE. In specific to Web QoE, Page Load time has been identified the main driver in Ameigeiras *et al.* (2010), Hoßfeld *et al.* (2011), Sebastian Egger *et al.* (2012), Strohmeier *et al.* (2014), Bischoff (2016) and Szabo *et al.* (2016) since it has been shown that the faster the page loads, the better the experience. But given the adopted definition of QoE (International Telecommunication Union, 2006), there is a plethora of other attributes which can cause the perception of the user to fluctuate. Moreover, the perceived Web QoE can differ from the perceived Mobile Web QoE given that context in a mobile setting can be very different from context in a fixed location. This study hypothesizes that it can provide a more fine-grained definition and model of Mobile Web QoE; one that is specific to the mobile web, one that is organically derived from live mobile web data and one that considers far more attributes than existing models found in literature. In doing so, this study depends on the data sensing capabilities of modern mobile devices to capture a rich attribute dataset and identify and measure the effect of those attributes on MWQoE.

"*Can we identify real-world scenarios in which we can predict and enhance MWQoE with reasonable confidence?*". The predicate to answer this question is the development of the MWQoE model and metric. The hypothesis here is that the effectiveness of the MWQoE model derived in a non-intrusive manner from anonymous live user data and without user submitted ratings (user opinions), can be evaluated in a lab experiment. In doing so, the real-world scenarios identified and modelled in MWQoE are simulated in a controlled environment where the user provides feedback on specific web browsing scenarios so that the user Mean Opinion Score (MOS) can be evaluated against the MWQoE metric in comparable scenarios. This approach examines the effectiveness of the MWQoE metric as a characterization of the user's experience in specific scenarios. Moreover, specific content adaptation measures in a lab experiment are tested to uncover their effect and potential enhancement to MWQoE. This is attempted on the assumption that MWQoE can be enhanced on the client side by adapting content and content-delivery for the user.

This study expands the definition of Web QoE from '*the interactive services that are based on the HTTP protocol and accessed via a browser*' (Hoßfeld *et al.*, 2011) to mobile devices. Therefore, this study defines the Mobile Web QoE (MWQoE) as *the experience perceived by users when accessing the web from a mobile device browser*. Given the pervasiveness of mobile web and the 4+ billion mobile devices in use today, from which almost half are smartphones (Statista Inc, 2016), MWQoE is considered an important and current research topic and is the primary subject of this thesis. Moreover, the importance of MWQoE is identified in industry by the fact that online providers continuously strive to deliver content faster by reducing web page load times (Strohmeier *et al.*, 2014), while at the same time they aim to improve the user journey through their digital services by carefully crafting features and web session sequences that are better received by their customers (Lanoue, 2015). There is an ever-increasing shift towards improving content delivery speed while simplifying and redesigning user interfaces for ease of use, all aimed towards improving the quality of experience. However, the effectiveness of such endeavors on the actual user experience is hard to be evaluated outside test labs and in real-world settings. While web analytics are becoming an absolute necessity for any site or service that needs further insights into their user base, they fall short in describing *satisfiability* or *technology acceptance*.

The contributions of this study are delivered via the *Better Experience* (BetterX) system. BetterX is the proposed, designed and implemented Mobile-to-Cloud solution which delivers the novel MWQoE model and metric. Thus, this study delivers 3 innovative major contributions: the MWQoE model (Section 3.4.4.6), the MWQoE metric (Section 3.4.4.6) and the BetterX system (Section 3.2.1.1.2). Moreover, this study delivers 3 minor contributions: The BetterX database (Section 3.4.1), the relational schema for storing HTTP metrics (Figure 22), and the findings from testing QoE impact features in a lab experiment (Section 4.3.7, 4.3.8). The BetterX database is the largest Mobile Web QoE database assembled from live user activity with more than 1.3 million readings. The Web Metrics Entity-Relationship diagram is the first published attempt to provide a normalized relational schema to anonymous HTTP archive data attributes. Finally, the findings from a lab experiment are presented and discussed, and the effect of 2 client-side content-delivery features on user *satisfiability* and subsequently to MWQoE is identified.

BetterX differs from other QoE implementations since it captures live user data from the BetterX Android App in a non-intrusive manner without any user intervention and questionnaires. BetterX uses a novel fusion of data sources – such as context data and web metrics – and offloads data processing to BetterX Cloud – a novel four-layer Cloud orchestration which allows processing optimizations at all layers. BetterX Cloud processes web session data and aligns the user's context and user's web session metrics to infer the *web session user intention*, the *web session user profile* and the different context states; all of which constitute the MWQoE Bayesian model (Section 3.4.4).

The MWQoE metric's value is that it can be used as a guide to trigger client-side MWQoE enhancements to the user. The BetterX solution defines and measures MWQoE as well as provides the system to ultimately enhance it in specific scenarios. Moreover, BetterX enables a continuous MWQoE evaluation on infrastructure which is flexible and scalable, making it a tool with high industry adaptation potential.

This novel MWQoE approach is guided by the 3 questions presented in Strohmeier *et al.* (2014) which are answered in order to provide a definition of Web QoE: *What are the key influencing factors*, *What are the user tasks*, and *What is the mapping between Web QoE and network metrics?* In doing so, this thesis delivers a MWQoE characterization in a non-intrusive context-aware manner, delivers a single MWQoE metric on a scale from 0 to 1 which is evaluated via actual user responses and provides an end-to-end MCC system for the continuous evaluation of MWQoE in real-world mobile user scenarios. The effectiveness of the MWQoE metric is identified since it closely reflects MOS in a comparison across all observed scenarios while confirming the importance of the MWQoE model parameters by the users.

1.1 Thesis Structure

Background & Related Work gives a detailed overview of QoE works in both industry and academia. In Background, the industry's motivations and trends to measure and enhance QoE are illustrated. In Related Work, the past and current QoE literature in video, voice and web adaptations is reviewed. Moreover, the most important QoE approaches are outlined as well as the Context-Aware adaptations of Web QoE. The review starts with the early establishment of QoS (Quality of Service) metrics and the gradual paradigm shift to a more user-centric approach that is QoE, and the examination of the user's attitude toward quality. In addition, the similarities and differences between QoE and User Experience (UX) studies are presented and discussed as they are both considered part of the broader field of "User Studies".

MWQoE Model Generation outlines the approach in developing the MWQoE model and metric, the adopted research methodologies and the analysis, design and implementation of all aspects of the BetterX system including the technical considerations of using Cloud as an offloading platform. Furthermore, it discusses attribute selection for the MWQoE model, provides an analysis of the collected BetterX dataset and concludes with a detailed review of the generated user profiles and the generated context-state models which are compiled together to deliver the MWQoE model and metric.

MWQoE Model Evaluation outlines the approach on evaluating the effectiveness of MWQoE against user data collected in a lab experiment via questionnaires. This is considered the secondary data collection phase for this study and presents how the user Mean Opinion Scores (MOS) in selected web session scenarios compares against the MWQoE metric, as well as confirms the importance of the selected factors which have been included in the design of the MWQoE model from findings of the lab data.

MWQoE Discussion presents a thorough analysis of the BetterX system, the MWQoE model and the MWQoE metric as well as the additional 3 contributions of this study. A detailed comparison of each of the 6 contributions to related literature is provided together with a discussion of the effectiveness and applicability of the contributions to industry and academia.

Finally, **Conclusion** re-iterates the contributions of the study and discusses how each contribution relates to the aims of the study and the main research question. An additional discussion is provided which paves the road for future work drawn from the collected datasets, findings and other observations from data analysis.

Chapter 2 Background & Related Work

Background provides an overview of the attempts to measure and enhance the user experience in industry. ***Related Work*** presents a literature review of QoE, Web QoE, as well as works in the field of UX. The need to approach QoE from a user-centered point of view and provide practical real-world approaches is identified in industry and in literature, and it is noted by this study. Moreover, the need for an interdisciplinary approach is identified as well; one that provides a more fine-grained practical model which can be used in real-world scenarios and captures a user-centered QoE. Most definitions that are found in literature, although comprehensive, they are conceptual and their actualization in a practical manner is missing, although the need for it is evident (as it is explained in this section). Moreover, practical models found in literature are limited to a small subset of attributes that QoE definitions are comprised of and do not combine factors of both human needs and expectations together in a unified form.

2.1 Background

Attempts to measure and enhance QoE are found in both industry and academia. In industry, companies like Google and Mozilla have been researching ways to keep their users satisfied and improve their QoE by speeding up content delivery; either by reducing content size or improving the delivery process. In the same spirit, academia has been involved in QoE by identifying factors which affect the user experience, measuring their effect on users and studying ways to improve it.

In an effort to address the limitations of the HTTP protocol (Fielding *et al.*, 1999) and devise the means to deliver content faster to users, Google developed *Spdy* (Google Inc., 2013). *Spdy* was an application-layer protocol, designed primarily to reduce Page Load time by 50%. *Spdy's* differentiator was that it used a single TCP connection per domain, as opposed to many, and used compression on headers. The *Spdy* protocol was the major influencer and driver of the newly adopted HTTP/2 (IETF HTTP Working Group, 2015) protocol by the Web Performance Group (W3C). HTTP/2 achieved page load time speedup by using binary instead of textual content transmission, being multiplexed instead of blocking, using header compression to reduce overhead and allowing for push responses from the server to client proactively.

In addition to *Spdy* which made the content delivery process faster, Google developed *Webp* (Google Inc., 2016), a new image format which reduced the size of the traditional PNG by 26% consequently minimizing the total transmission time. Mozilla on the other hand developed *Electrolysis* (Mozilla, 2016), its most recent and most noticeable attempt to increase page responsiveness. As reported by Mannes (2016), the newly designed multi-process architecture for Mozilla's Firefox browser increased page responsiveness by an astounding 400 to 700 percent increasing Firefox user's satisfaction.

At the same time, efforts to measure web application performance have been developed by industry with the W3C taking the lead to develop the web performance timing specifications (W3C, 2015). The web performance timing specifications broke down timings in each functional component of the client-server architecture and outlined ways in which each component's performance could be measured. For example, the navigation timings specification enabled the measurement of the complete end-to-end user latency, while the server timings specification enabled the client to collect from the server request-response performance metrics allowing for app delivery optimizations.

These attempts to measure and speedup web application performance have been motivated by the effect of Page Load time on user experience. It has been shown that the more time a user waits the less positive is the experience, the higher the drop-off probability and the less revenues for the provider (Souders, 2009). Therefore, the Page Load effect to QoE has been a major motivator to service and content providers.

In *Velocity 2009* (Souders, 2009) – one of world's most important web performance conferences - the extent of the Page Load time effect has been the main talking point. Bing showed that a 2 second slowdown of search results could reduce revenues by 4% where Google discovered that doubling of Page Load time could result in a 25% user drop-off. Shopzilla on the other hand observed a 25% increase in page views and a 7-12% increase in revenues by a speed-up of Page Load times from an average of 7 seconds to an average of 2 seconds. In addition, and as it was stated in Eaton (2012), 1 in 4 users abandon a site if the page takes more than 4 seconds to load. As the same time, Amazon showed that a 1 second delay in Page Load could cost their operations an astonishing 1.6 billion dollars in sales annually.

The significance of Page Load time in industry, its effect on user experience and consequently to the revenues of content and service providers, has undoubtedly been one of the major motivators in the attempts by industry outlined above. It has been a major driver in this study as well, since Page Load time has been considered in this study's MWQoE model as a major factor which its effect on Web QoE has been observed during data exploration (Section 3.4.2) and the weight of its effect evaluated via actual user responses in a lab study (Chapter 4) .

However, latest trends in industry as well as in academia shifted focus for measuring the web experience from a Page Load perspective to measuring the web experience from a design and user interface viewpoint. And although Page Load timings are as important today as in the past, web performance is currently taking a more refined stance on experience, with effort being put on the user in designing seamless user experiences and the delivery of enhanced human interactions with web content. As posted on the latest Velocity conference website – *Velocity 2016*: '*Performance for the people: It's not just about page-load time anymore—the focus is shifting increasingly to seamless user experiences and human interactions. Design and design thinking is changing the way we approach our work, products, and teams*.' (O'Reilly Media, 2016).

This noticeable shift in industry from Page Load metrics to User Experience (UX) (Lanoue, 2015), from objective metrics calculation to subjective metrics inference, has been reflected in academia with the shift from QoS metrics to QoE. As research suggests, even good metrics are not important if a user is unhappy about a service being offered (Ickin, Wac & Fiedler, 2012). Therefore, even if QoS metrics are excellent, interfaces may be difficult for users to navigate, the accessing device may be performing poorly or a user may just be having a bad day, making his/her experience less than *excellent*. And, even though it is paramount for application and service providers to be mindful of Page Load timings and measure the level of impact of the network on *user satisfaction*, there is a need to examine the user experience in a more interdisciplinary and user-centered approach to identify and measure subjective user related factors. QoE reflects the user's perception as it illustrates the expectations and reference of the user to the network, the device and the specific application being used.

Considering this shift in trend and the apparent need in industry for experience evaluations, this study takes a more refined "technical" approach which uses a novel QoE method to construct the Mobile Web QoE (MWQoE) model (Chapter 3) together with a UX-type lab approach to evaluate the effectiveness of the generated model and metric with user opinion (Chapter 4). While recognizing the advantages and disadvantages of both QoE and UX studies (as outlined in this section), this work takes a "technical" standpoint in terms of an experience evaluation. The effect of not only web metrics such as the Page Load time are examined, but focus is given on the user by examining the user's context and the effect of the user's context on Mobile Web QoE (MWQoE). In addition, *usability* aspects of the user's experience, although outlined and noted in the lab experiment, are not included in the MWQoE model where a more fine-grained characterization of current Web QoE models is aimed and delivered, by the inclusion of device and network metrics as well as web metrics.

2.2 Related Work

2.2.1 Defining QoE

QoE can be understood as the quality perceived by the user in using an application or service. Given that there is current literature dealing with the criteria for measuring QoE for specific scenarios, there is still no universally accepted definition actualized to produce a quantifiable metric in real-world settings. Nevertheless, it is accepted that the term QoE refers to the perception of the user about the quality of a service or network. More specifically, QoE is defined in Schatz *et al.* (2013) as the overall acceptability of an application or service, as perceived subjectively by the end user.

As in Industry, currently QoE academic research is the result of a paradigm shift to a more user-centric environment from a network-centered one which did not fully capture the user's experience into a distinctive metric. Throughout the years the research community has focused on the Quality of Service (QoS) as the standard quality concept. QoS defines the capacity of a specific network to give a service at a guaranteed service level. QoS contains all components, capacities, and systems in the network that guarantee the continuation of the arranged service quality between the client and the network. QoS is well defined with metrics such as throughput, delay, jitter etc.

The primary question to consider within the literature is a formal definition of QoE. Various definitions of QoE are published, which tend to be similar but still vary to an extent, for instance:

- QoE concentrates on the client and is considered as the accumulation of all the observation components of the network and execution with respect to desires of the clients (Strohmeier et al. 2014).
- QoE applies to any sort of network collaboration, for example, it may be seen as impacted by three variables: the client's substance inclination regarding needs and objectives, the network over which the substance is gotten to, and the gadget with which the client interfaces with the network (De Moor *et al.*, 2010)
- QoE is a multi-dimensional development of discernment and conduct of a client, which speaks to his/her passionate, cognitive, and behavioral reactions (Schatz *et al.*, 2013)
- QoE can be characterized as the qualitative measure of the everyday experience the client gets when he utilizes the administrations he is subscribed to including experiences, for example, blackouts, nature of picture, velocity of the fast Internet administration, inactivity and postponement, client administration, and so forth (Reichl *et al.*, 2010).

The adopted QoE definition for this study is derived from ITU-T (International Telecommunication Union, 2006) which defines it as follows: '*The overall acceptability of an application service, as perceived subjectively by end-users. QoE includes the complete end-to-end system effects (client, terminal, network, services infrastructure, etc.). The overall acceptability may be influenced by a series of factors including user expectations, the usage context, the device usability and the user's personality.*'

In regards to Web QoE and in realizing the shift from QoS to QoE, the Advancing Customer Experience (ACE) framework (Schatz & Egger, 2011) was formed by the Telecommunications Research Center in Vienna and in collaboration with mobile network operators as an attempt to provide the methodology and

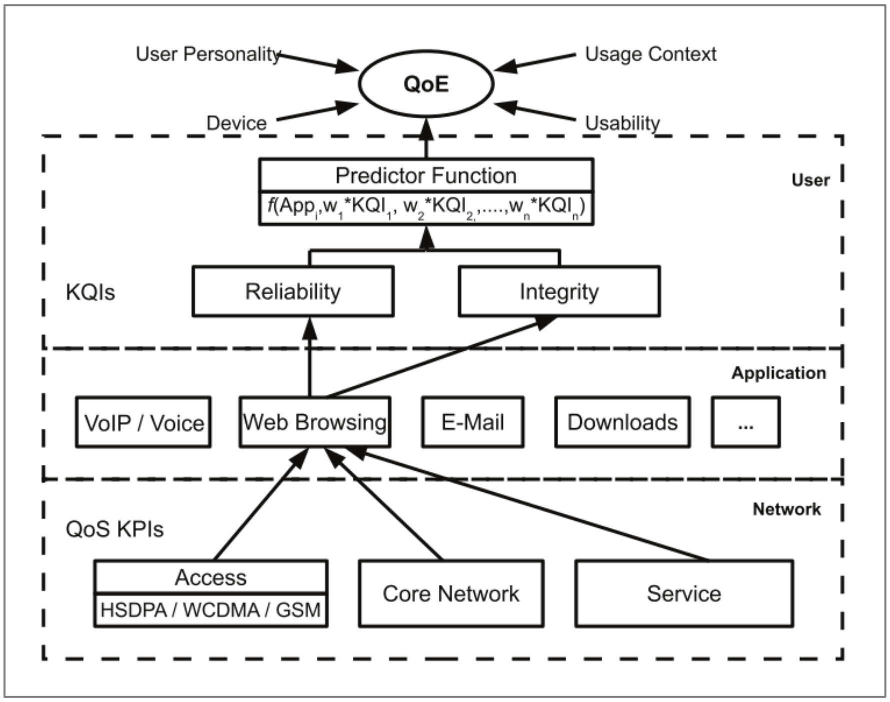


Figure 1 The ACE QoE Modeling Framework (Schatz & Egger, 2011)

context for future QoE experiments. ACE's goal was to investigate the links between QoS and the perceived QoE in a layered approach as shown in *Figure 1*. This modeling framework regarded QoE as a multidisciplinary, user-centric study which was affected not only by QoS Key Performance Indicators (KPI) but with the user's personality, the device context, the perceived usability and the usage context of the task at hand. Moreover, the study concluded two important QoE realizations. The first was that results from lab and field studies can vary considerably in regards to QoE even when the same tests and QoS parameters are used. The second was that context can have a substantial effect on QoE ratings and it is preferred to be tested in the field rather than in the lab.

In another study and in addition to the above findings, the QoE interaction model in the communication ecosystem was presented by Rehman Laghari & Connelly (2012) which defined QoE as '*a blueprint of all human subjective and objective quality needs and experiences arising from the interaction of a person with technology and with business entities in a particular context*'. Rehman Laghari, Crespi and Connely gave an interdisciplinary view of a QoE integrated framework which combined cognitive science, psychology, business and technology. In doing so, they regarded user context to influence experience and have a significant impact of QoE; thereby classifying contextual entities and mapping their causal ("cause-effect") relationships with each other. Moreover, they provided a taxonomy of QoE relevant attributes and their interactions.

Adding to the connection of QoE and cognitive science, the *WQL* hypothesis (Sebastian Egger *et al.*, 2012) used the Weber-Fechner law of psychophysics, which considered time as a stimulus, to describe the smallest acceptable difference between sequential levels. In doing so, *WQL* revealed a logarithmic relationship between QoS and QoE which stated that bandwidth and waiting time are very closely correlated – the more a user waits the less satisfying the experience. Moreover, (Sebastian Egger *et al.*, 2012) stated that there is a difference between the technical Page Load time and the user perceived Page Load time.

In another study, the *IQX* hypothesis (Fiedler, Hossfeld, & Tran-Gia, 2010) revealed a negative exponential relationship between QoE and one degrading QoS attribute. The *IQX* hypothesis was validated for streaming services in which the QoE was expressed as the Mean Opinion Score (MOS).

The major QoE influence factors were identified in (Liotou, Tsolkas & Passas, 2016) and were broken down into Service Independent and Service Dependent. Service Independent factors included Network/ Transport/ Physical layer factors and Common (usability) type factors. Service Dependent factors included factors that were specific to the service being tested such as video, VoIP, download type, etc. Moreover, QoE model types were classified into 3 categories: subjective, objective and hybrid (a fusion between subjective and objective).

2.2.2 QoE on Voice and Video

In examining QoE for voice and video applications, most studies were focused on examining the correlations of QoS with QoE for attributes such as bandwidth, delay, loss, etc. The results of these studies were validated mostly by using the Mean Opinion Score (MOS) in Lingfen Sun and Ifeachor (2002, 2004, 2006) and Hoßfeld and Binzenhöfer (2008). The MOS of an audiovisual sequence was simply the average of the valuations made by users as actual quality of sequence, as measured on a scale from 1 (lowest quality) to 5 (best quality). Following this paradigm, the goal of a QoE metric was to correlate best with the MOS obtained from a subjective quality test with real users. However, it was observed that MOS evaluations tended to be computationally intensive, cumbersome, not repeatable, and hard to adapt to real time quality assessment (Hosfeld, Fiedler & Zinner, 2011).

In the case of video QoE, the well-known video quality metrics such as the Video Quality Metric (VQM) and Structural Similarity Index (SSIM) illustrated that they are a significant improvement over the traditional video quality metrics such as MSE and PSNR (Liotou, Tsolkas & Passas, 2016). This was achieved by the consideration of some inherent characteristics such as the form of distortion and structural information which improved the confidence level of quality evaluation. However, it was found that both cannot reliably replace subjective evaluation.

In regards to QoE-aware wireless video streaming, Scalable Video Coding (SVC) (Schwarz, Marpe & Wiegand, 2007) stands out with rate adaptation capabilities coping with bandwidth scarcity and network variation. SVC offers spatial, temporal and quality scalabilities for enhancement in picture size, frame rate and picture fidelity, respectively. QoE-aware streaming is achieved by excluding some enhancement layer packets during transmission to conform to network variation, hardware heterogeneity or users' requirements.

For voice and VoIP applications, Ding and Goubran (2003) proposed a parametric, non-intrusive discourse quality appraisal calculation. The voice payload investigation was performed by the Perceptual Evaluation of Speech Quality (PESQ) (Rix *et al.*, 2001) calculation and the commotion recognition model was fused in the ITU-T E-model (International Telecommunication Union, 2015). In another study by Hoßfeld and Binzenhöfer (2008), the QoE in Skype for UMTS networks was measured with regards to users MOS. In it, the activity profile of the Skype application was inferred from QoS parameters like throughput and jitter.

In Szabo *et al.* (2016), a new QoE model, the *MLQoE*, was introduced which was built on QoS predictors (such as packet loss and delay) via supervised regression and the subjective opinion scores reported by users. The differentiator of the *MLQoE* model was that it fused more network metrics in its feature selection algorithm than previously published models such as the *IQX* by Fiedler, Hossfeld, & Tran-Gia (2010), the *WQL* by Sebastian Egger *et al.* (2012), the *E-model* by International Telecommunication Union (2015) and *PESQ* by Rix *et al.* (2001). Moreover, it was shown the *MLQoE* outperformed all the above models and predicted the user reported QoE (MOS) '*fairly accurately*'.

Finally, in Casas *et al.* (2015), the effect of network conditions to QoE was examined for popular applications such as Youtube, Facebook and Web browsing. A lab study was used to gather MOS on different network configurations and scenarios which showed that the downlink bandwidth fluctuations are key in the effect of QoE in regards to the applications tested. One of the differentiators of the study was that it used passive traffic monitoring, as opposed to active traffic monitoring but lacked the use of user context and regarded QoE only from a QoS standpoint.

It is noted, that QoE in voice and video studies measured QoE from the network perspective, as outlined above, and completely disregarded the user's perspective and the user's context in the examined scenarios.

2.2.3 Web QoE

Web QoE measures the experience of the end user while using the web from a browser (Cecchet *et al.*, 2013). As stated above, most QoE literature deals with voice and video, much less deals with Web QoE and even less than that deals with Web QoE from a mobile perspective. However, given the pervasiveness of mobile device usage (Statista Inc, 2016) and the fact that more than half of web traffic generated comes from mobile users (Naqvi, 2015), there is a need to study Web QoE from the mobile user perspective. The Mobile Web QoE (MWQoE) is defined as the experience of a mobile user accessing web resources via a browser and is the focus of this study.

Regardless of network generation, usage patterns, pricing plan or mobile device, web access speed remains one of the defining characteristics in literature in measuring Web QoE and the overriding determinant of user satisfaction for mobile web services. As mobile network operators battle with network overload, mobile users complain about network speed (Schatz and Egger, 2011). There is a need for improvement in several areas such as data transfer protocols (IETF HTTP Working Group, 2015), website improvements (Souders, 2009) and faster web browser implementations (Mannes, 2016) in order to achieve better browsing speeds and therefore better user satisfaction. At the same time, there is a need for an enhanced Web QoE definition and approach, one that is practical and attempts to capture and measure not only web speed and Page Load, but user-related factors as well, and enable a more *holistic* approach capturing more of the factors which are included in the adopted ITU-T definition.

In regards to defining objective Web QoE models, two recent studies are outlined below. The work of Ameigeiras *et al.* (2010) provides an estimation of Web QoE in regards to web response time. It shows how the user perceived Web QoE (as expressed by MOS) changes in different service response times for the scenario of web page downloading. The work of Sebastian Egger *et al.* (2012) provides a logarithmic mapping of QoS to Web QoE from an interactive web browsing lab experiment. In Sebastian Egger *et al.* (2012), the applicability of the Weber-Fechner law on QoE is demonstrated. Further, it shows that the interaction of the TCP and HTTP protocol with the network in scenarios such as *HTTP pipelining*, *impact of TCP's slow start*, etc. cause a complex, non-linear relationship between network metrics such as bandwidth and wait time. Moreover, it is stated that the application-level Page Load time differs from the network Page Load time since local machine rendering requires additional overhead and can vary dramatically.

The importance in using both objective and quantitative measurements in QoE studies is stated in Brooks and Hestnes (2010). The authors argue that objective measures do not necessarily need to be derived from technology such as QoS, but can also be derived from users, with the involvement of psychological research in measuring user behaviour. It is noted that future QoE models must enable industry to measure and compare services or products efficiently by considering: 1) the technical parameters (QoS), 2) the usage context attributes and 3) the outcome of usage (such as effectiveness and satisfaction). Furthermore, it is illustrated that objective psychological measures will not be dependent on user opinion (such as MOS), since it is shown that '*user behaviour should not rely on user opinion*'. For example, the completion time of a specific task which can be considered a psychological measure can be measured without user opinion.

Strohmeier et al. (2012) shows that the user task, the web content, and the evaluation of the user's fulfillment goal need to be considered when measuring Web QoE. In addition, Schatz and Egger (2012) presents three important findings in regards to Web QoE. The first is that bandwidth has a moderate effect on Web QoE. The second is that device features such as application performance and screen size impact Web QoE more than bandwidth. Finally, the third is that screen size impacts Web QoE.

In addition, the work of Hoßfeld *et al.* (2011) shows that the memory effect (historical experiences of users) is an influencing factor for Web QoE. In this study, the Page Load time is measured via a simulation of a web browser with HTTP traffic. It is shown that the Page Load time '*well aggregates the influences of network transmission on Web QoE*' and that the memory effect of Web QoE is revealed by a difference in MOS levels of web pages with the same QoS metrics. For this reason, it is noted in this study that for calculating Web QoE of a series of web sessions, the previous Web QoE needs to be accounted together with the current QoS readings.

2.2.4 Web QoE Approaches

Several approaches in measuring and enhancing Web QoE have been reviewed for this study. Below is a brief review of the most notable approaches found in current literature.

In Szabo *et al.* (2016), the author utilizes a collaborative QoE-aware resource for redistributing network traffic to web pages. In doing so, the QoE from a network perspective is measured, illustrating the potential to improve download time by 104% in specific scenarios, thus improving the Web QoE.

In Hora *et al.* (2016), a study to examine the effect of WiFi on Web QoE is presented. In it, a web browser emulation is used to measure the Page Load time which is regarded as Web QoE. A total of 4 regression based prediction algorithms are used to predict when the quality of the WiFi signal degrades, therefore predicting QoE reduction. The author argues that potential adaptation of such a technique by ISPs can help them proactively initiate WiFi troubleshooting measures to their clients thus helping them avoid poor Web QoE.

In Cherubini and Oliver (2009), the automatic logging technique is used to record events sensed from mobile device without user intervention. The automatic logging technique succeeds in capturing real-world data outside controlled lab environments. In addition, the Quasi Experimentation is proposed to extend the automatic logging technique in capturing the user's context with micro cameras in real-world environments. Moreover, the method of Mobile User Experience Research in Nakhimovsky (2009) is outlined, aiming at capturing user interaction and system state while recognizing and capturing user context and state, and gathering and managing user self-reports.

In Cecchet *et al.* (2013), an infrastructure is presented called *mBenchLab* which measures the QoE of accessing web services from mobile devices in real-world scenarios. The novelty of *mBenchLab* is that it does not rely on simulation or emulation for data. This is achieved by recording HAR (Odvarko, Jain & Davies, 2012) traces via an Android proxy on web sessions to specific websites which are re-built from their originals internally into a controlled public cloud infrastructure. Analysis of the data is done via replaying the HAR traces and using the latency observed from different usage scenarios (such as laptops, tablets, smartphones) to measure QoE.

A more suitable approach for real-world capturing has been used in Ickin, Wac & Fiedler (2012) where data has been collected in an unobtrusive manner using context sensing from Android users. This was a hybrid model attempt using both quantitative and qualitative procedures but did not produce any practical metrics rather than only generalized results. The user ratings were collected using the Experience Sampling Method (ESM) (Cherubini and Oliver, 2009) and further information from users was gathered via weekly interviews.

Finally, in Mitra, Zaslavsky & Ahlund (2015), the Context-Aware QoE Model (*CaQoEM*) model is presented. *CaQoEM* uses Bayesian networks (Jensen and Nielsen, 2007) and Utility Theory (Ligeza, 1995; Padovitz *et al.*, 2005) to build a context-aware model for measuring QoE on mobile devices. The model uses 3 QoS metrics (codec, packet loss, delay) and the user location as the user's context. The QoS metrics are simulated for a VoIP application in different network scenarios such as *vertical hand-offs*, *network congestion* and *wireless signal fading*. The derived QoE measurement is validated by the opinions of 29 participants in a lab study.

In regards to enhancing the Web QoE in mobile user scenarios, several studies have dealt with the adaptation of web content to the mobile device. Given the deliberation that mobile devices are bound by hardware limitations, such as limited keyboards, small screens and small memory, they are given special consideration when accessing and rendering information from the web. The notable work of Schatz and Egger (2012) revealed the issues in rendering web content on mobile devices and noted that screen size has an effect on Web QoE.

The study of Gong *et al.* (2009) classifies existing adaptive approaches of mobile web applications into three categories: content selection, content personalization and adaptation of the presentation. The World

Wide Web Consortium (W3C) defines adaptation as a process of selection, generation or modification that produces one or more units of perception in response to a single requested resource. Hence, mobile web adaptations are defined as any automatic action that adapts the content and presentation to improve user interaction with mobile handheld devices.

Web pages may contain multiple images, forms, and other dynamic content, which cannot be displayed in the same manner as on regular desktops and laptops. In certain cases, the limitations of the device may result in certain content being unrenderable, or simply not accessible (e.g., navigation via image-maps on a low-resolution device). These issues raise questions about different methods of improving mobile Web access using different languages, formats and architectural designs. Zhou (2007) proposes an architecture called Scalable Browser for mobile devices. The Scalable Browser features include fetch-on-demand, progressive rendering and display on demand navigation style. The overall goal of the research is to enhance the user interface and browsing experience for handheld devices. The Scalable Browser architecture is based on a progressive delivery and rendering process whereby partial contents are rendered to the client. This is achieved by separating HTML pages into: structural data (which determines the style/geometric layout of HTML tree) and semantic data (descendants of structural data). In addition, browsing is aided by converting the HTML pages to an intermediate SVG format, which retains all the features of HTML in order to ease the deployment process (Coles, Meglan & John, 2011).

In Yap and Marshall (2010) a navigation model for web access is proposed which allows existing web content and services to be used on wireless devices. The m-Links system is designed to achieve the following goals: web navigation on small devices, digging into embedded information on web pages for useful data, separation of service from links, and providing an open framework for others to develop services for wireless clients. One main advantage of this scheme is that the entire content from the requested site is not sent back to the client all at once. Pages are summarized in a neat, hierarchical format of links to enable clear navigation. A user is not flooded with the entire contents of a page at the initial step, but receives a list of links through which the user can “dig” for more content (“dig and do” model). The m-Links architecture consists of three main components: the link engine (processes web pages into link data structure); the service manager (returns services appropriate for each link e.g. read, print, send, etc), and, the UI generator (supports different user interfaces for different small devices).

In Zadeh, Wang & Kubica (2007), the authors introduce five methods for summarizing, browsing and progressively disclosing parts of web pages for small handheld devices. Using this scheme, the requested web page is divided into “Semantic Textual Units” (STUs), which can be lists, paragraphs or image ALT tags (images are not displayed) that are arranged in a STU hierarchy. The main contributions of this research include: summarizing Web pages through partitioning into STUs and summarizing the parts. The authors developed and experimented with different summarization schemes involving selecting important, descriptive STU keywords. This summarization process is very important, as it is the core of the progressive disclosure mechanism used for mobile clients. Different keyword extraction techniques and summary-sentence extraction were performed to adequately summarize the STUs. The authors demonstrate with experiments that their summarization scheme in some case proves to be three or four times better than no-summarization schemes as in Yap and Marshall (2010).

2.2.5 Context Aware QoE

The realization that context-awareness can be used in characterizing Web QoE has been shown in Nakhimovsky (2009), Mitra, Ahlund & Zaslavsky (2011), Schatz and Egger (2011), Ickin, Wac & Fiedler (2012), Mitra, Zaslavsky & Ahlund (2015) and Mantoro *et al.* (2011). It has been shown that context-awareness can not only provide a more accurate characterization of Web QoE but provide the basis for enhancing Web QoE. Following is an overview of how context is classified and of what constitutes context-awareness in QoE.

Context is composed by using 4 different types of information: location, identity, activity and time (Dey and Abowd, 1999) and is defined as ‘*any information that can be used to characterize the situation of an entity. An entity is a person, place, or object that is considered relevant to the interaction between a user and an application, including the user and applications themselves*’ (Dey, 2001). The use of the term *context-aware* in any system, method or application shows that context is being considered in the measurement or analysis of the task at hand and is used to enhance the end-product. In the case of QoE, *context-aware* refers to the use of context to measure and enhance the user experience. In doing so, a *context-aware*

QoE method automatically collects context related data and '*lets the application designer decide what information can be used and how*' (Dey, 2001). Moreover, based on the work of the same authors in Dey and Abowd (1999), context is used to present information and services to a user; automatically trigger the execution of a service, and classify and store that information for later use.

In a similar manner as in Dey (2001), the authors in Bettini *et al.* (2010) describe context reasoning and interpretation as '*to abstract from low-level context by creating a new model layer that gets the sensor perceptions as input and generates or triggers system actions*'. As shown in Figure 2, the sensor/context attributes provide the raw dataset from which a semantic interpretation of context is inferred. Then, the inferred context is used in conjunction to define situations as relationships between inferred context states. The situations reflect the user's behavior and the state of the associated system(s). Therefore, situations are used to provide meaning from context sensors where context-aware systems use changes in situations (or scenarios) to trigger system adaptations.

In Padovitz *et al.* (2005), the *Context Spaces Theoretical Model* is presented which fuses sensor data with *MAUT* (Multi-Attribute Utility Theory) (Ligeza, 1995) to provide an approach for context-aware applications. The aim of this approach is to infer the occurrence of situations, as shown in Padovitz *et al.* (2005) and Bettini *et al.* (2010). For example: "Is y happening at x time?" As defined, the situation space is comprised of a collection of context states which are then represented by a collection of context attributes illustrating the state of a specific entity at a specific time. The *relevance* function is used define the importance of attributes (weight of attributes) to a specific situation, relative the other attributes that are part of the situation. In addition, the *contribution* function is used to define the individual contribution (the utility on a scale from 0 to 1) of the attribute to a specific situation. As stated, the *MAUT* '*reflects the evaluation of a particular attribute containment; the higher the probability of an attribute value being within the region, the greater contribution is evaluated for the attribute*' (Padovitz *et al.*, 2005). Basically, *MAUT* considers the definitions outlined above to derive a metric which shows the occurrence of a situation in a single numerical scale from 0 to 1. *MAUT* is used in this study to derive the MWQoE metric from the MWQoE model in a single scale from 0 *Poor MWQoE* to 1 *Excellent MWQoE*.

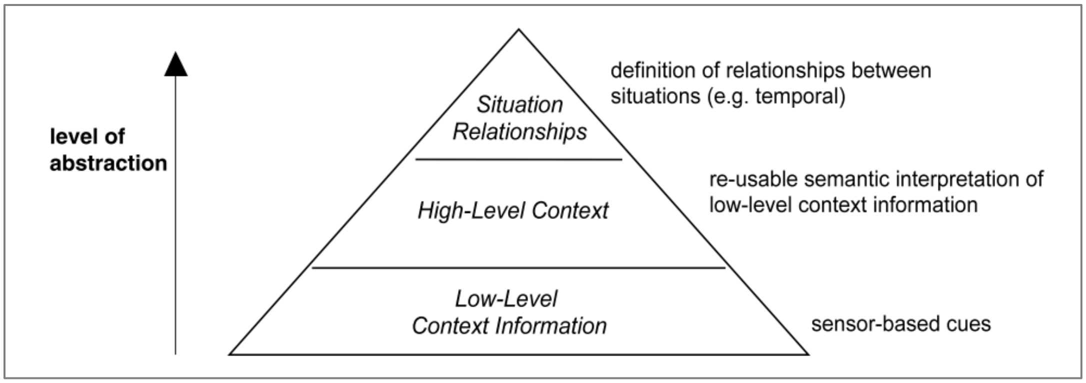


Figure 2 Context Interpretation and Abstraction (Bettini *et al.*, 2010)

2.2.6 QoE in User Studies

Within the broad field of User Studies (E. L.-C. Law *et al.*, 2008) the term *Quality of Experience* constitutes a generic term which is loosely defined, and its definition is based primarily on the approach employed in its evaluation. Human-Computer Interaction (HCI) is considered a part of User Studies which examines the way people interact with software (Preece and Rombach, 1994). HCI focuses on understanding and evaluating the *usability* of software packages while employing design and implementation methods for human-computer interfaces. On one hand, in HCI studies, *User Experience (UX)* (E. Law *et al.*, 2008; Bevan, 2009) is linked with *usability* which is primarily viewed as the quality of the interaction between the user and the system, where the improvement of the interaction between the two is the end goal. On the other hand, and as shown in the previous section of this chapter, QoE approaches take a more technical perspective where the user's satisfaction is linked with the performance of technical parameters (quantifiable measures) such as QoS.

The work of Ballesteros and Segall (2013) distinguishes these two *User Experience* approaches. The first approach named "techno-centric" links the user's satisfaction with technical (mostly network) parameters. In this approach the basic assumption in QoE is that the better the performance of the network, the higher the satisfiability of the user. The second approach named "user-centric" examines *usability*, the way users interact with systems, and how usability and user interaction components affect the user's experience. This approach entails a multi-disciplinary point of view which blends user perception studies with QoE studies and examines the ways human perception can be incorporated into QoE evaluations and management.

Moreover, Ballesteros and Segall (2013) argue that a shift from perception to a more subjective approach where user expectations and experiences are considered, can benefit the HCI community; that the user-centric approach can complement the techno-centric approach since it brings together aspects of the user experience where users are involved.

The work of Preece and Rombach (1994) themes to the interdisciplinary nature of HCI and compares HCI evaluation approaches with Software Engineering (SE) approaches. They argue that there is a need and utility of a common taxonomy and measurement framework for both HCI and SE in '*facilitating communication between the two disciplines*'. They identify HCI as the study which combines computer science with sociology, psychology and anthropology and outline 5 HCI approaches: Laboratory testing, Usability Engineering, Heuristic evaluation, Ethnographic studies and Cognitive Modelling. Laboratory testing is the common lab study in which differences in one or more independent variables are examined. Usability Engineering is the approach where the usability of the product is quantitatively defined and can be evaluated either in a lab experimentation or in a field study (observing users at work using a specific software product). The Heuristic evaluation identifies usability problems when inspecting software systems while the Ethnographic study adds contextual inquiries in HCI evaluation. Finally, Cognitive Modelling uses the user's cognitive and physical behavior to build models using the Goals, Operators, Methods and Selection rules (GOMS).

In addition, two different approaches to HCI are provided by Dillon (2002) and Brooks and Hestnes (2010). Dillon extends the ISO-9241 definition of usability, being a form of experience evaluation, from '*the effectiveness, efficiency, and satisfaction with which specified users can achieve specified context of use*' to QoE being an artifact of calculation based on the user's action, the end-result and the user's emotion. It is argued that usability does not ensure a high quality of experience since each of the defining terms of usability (effectiveness, efficiency and satisfaction) can be problematic in different scenarios. For example, satisfaction, which relates to the user's preference, may not use maximum effectiveness nor efficiency in different scenarios. The proposed user experience definition is defined by Process, Outcome and Affect where Process is the interaction with the system; Outcome refers to all the parameters that capture the gain of the user for using the system and Affect refers to all the psychological/mood elements of the experience. Moreover, Dillon (2002) presents that aesthetics are found to be directly related to usability and that usage time can yield different usability results. For example, a system which is found to be unusable in a lab experiment can be found usable given that the user is given more time with it.

In Brooks and Hestnes (2010), the authors argue that the validity of QoE evaluations is higher when it includes objective measurements and uses quantitate data which is consistent with numerical objective data. By using quantitative numerical data, the validity of QoE evaluations can be explored via statistical descriptions whereas qualitative ratings cannot provide validity since the difference in the numbers does not reflect the distance between each scale position used in the evaluation. The authors argue that '*the study of user behavior should not rely on user opinion*', that QoE approaches should be objective and not subjective, and should be a '*direct measure of either the process or outcome of user behavior*'. In doing so, the authors propose that for higher validity QoE evaluations, ratings by users on specific context and in specific scenarios can be utilized. The ratings can be captured via quantitative ratings (interval or ratio scale) rather than qualitative ratings so that the data can be statistically summarized; therefore, opposing the widely used MOS quality scales (1 (Bad), 2 (Poor), 3 (Fair), 4 (Good), 5 (Excellent)) where the difference in the numbers does not reflect the distance between each scale position.

Moreover, the authors provide a taxonomy and critically analysis of 3 QoE approaches: The first QoE approach is *Testing user-perceived QoS* where subjective data is collected via standardized methods, such as MOS where users are surveyed to express their experience evaluation in specific testing scenarios. The recorded QoS levels are compared and analyzed against user evaluations. It is argued that a major disadvantage to this method is that important changes in QoS may be small enough for users to unconsciously notice although these changes may still affect their behavior. Moreover, it is stated that this method collects only user's perceptions and not user behavior nor user interaction. The second QoE approach is the *Surveying Subjective QoE* method which extends the *Testing user-perceived QoS* method by incorporating measures such as usability and user satisfaction. However, it is still found inappropriate since it is a subjective approach. Finally, the third QoE approach *Modeling Media Quality* success in being objective since it measures technical parameters and being user-centered since it tests parameters which directly influence the user's experience. However, this evaluation is also found to be inappropriate since it performed against user opinion which is flawed as stated above.

In regards to QoE from a web perspective, the work of Hasan and Abuelrub (2011) pulls together website quality criteria from different works in order to provide a unified evaluation framework. It is stated that content quality, design quality, organization quality, and user-friendly quality are the 4 dimensions which can be used to evaluate quality in Web QoE studies. For each dimension the authors provide a series of criteria. For example, for Content Quality, the content needs to be timely, i.e. 'up-to-date' and accurate i.e. from valid sources and without errors. For design quality, the site needs to be attractive and make use of multimedia components like images and videos. For Organization quality, the site needs to have an appropriate structure and be well organized i.e. be usable, easy to use and navigate and reliable, i.e. load fast and be constantly available. For user-friendly quality, the site needs to be customizable based on user preferences and interactive via various communication channels such as chat rooms and suggestion forums.

Chapter 3 MWQoE Model Generation

3.1 Introduction

The Model Generation section outlines and discusses the design of the BetterX system and the collected data list, the methodology procedures and the conceptual model as it was devised in the experiment design, as well as the implementation and technical aspects of the BetterX system including the data attributes selected for the MWQoE model. The section concludes with a presentation of the data analysis results and a detailed examination of the final MWQoE model.

The Model Generation phase, is an applied quantitative data-driven study (Kothari, 2004) which involves measurements of specific attributes of selected context entities and delivers a practical solution which models and measures MWQoE organically from live user data. The *Model methodology* (Elio *et al.*, 2006) is used to construct a model for characterizing MWQoE, and the *Experimental methodology* (Elio *et al.*, 2006) is used to develop the BetterX system used for data collection.

This study uses a novel fusion of context states (Padovitz *et al.*, 2005) to build the MWQoE model and metric using Bayesian Network (Jensen and Nielsen, 2007) and linking it into a MWQoE metric based on Utility Theory (Ligeza, 1995).

The study is context-aware since it uses the user's context to interpret user mobile web browsing scenarios and infer user situations based on context spaces theory (Padovitz *et al.*, 2005). Context-awareness in system research refers to environment, device and network characterizations of particular entities (Dey, 2001). This study uses context in the same manner as in Dey and Abowd (1999) and builds *web session user intentions* in order to infer the intention of the user in a distinct web session, and *web session user profiles* in order to infer the profile of the user for a specific web session. Based on the current state of capabilities of the Android operating system and modern mobile devices, a total of 65 attributes are identified as interesting/useful and are collected. These attributes are mapped into 6 distinct data domains 1) Device, 2) Network, 3) Connection, 4) Web, 5) Sensor, 6) User; and aim in interpreting the situations in which a web session has taken place in a specific location at a specific time. The method of context interpretation, as illustrated in *Figure 2 Context Interpretation and Abstraction (Bettini et al., 2010)*, is used to infer these situations using a three-stage process. The first is the low-level context which is captured by sensors (such as GPS for location context), the second is a semantic interpretation of a high-level context, i.e. the situation, and the third is the relationship(s) between the situations. The idea behind context reasoning and interpretation (Bettini *et al.*, 2010) is to infer 'sensor perceptions' or interpret metrics given an understanding of the user's situation. For MWQoE, this context-aware method is used to generate *Web Session User Intentions* and *Web Session User Profiles* based on observations of the attributes in the 6 data domains (Device, Network, Connection, Web, Sensor, User). As in Bettini *et al.* (2010), the context reasoning method is being used to derive the *Web Session User Intention* from web metrics and context data (high-level context) and use that to derive the *Web Session User Profile* and the scenarios (situations) of mobile web usage.

The underlying theoretical framework for constructing MWQoE is based on Decision Theory, which combines Probability Theory and Utility Theory (Ligeza, 1995). This framework was chosen since the MWQoE approach needed to account for uncertainty given that certain data attributes could not be available in certain situations, for example, the user's location could be unknown if the location service was disabled on the user's mobile device. This unknown factor is referred to as "risk" in decision theory.

The MWQoE metric is derived using Utility theory. The MWQoE metric by definition is highly aligned with the definition of Utility Theory that '*Every state has a degree of usefulness, or utility, to an agent, and that the agent will prefer states with higher utility*' (Ligeza, 1995), such that the higher the utility the higher MWQoE metric for the user. Utilities are easy to use in decision-making. The decision rule in these cases follows simply the following paradigm: *"choose the option with the highest utility"*. This basic rule is the core of deciding on the best experience for the mobile web user. MWQoE is generated by a utility, where the numerically represented values are based on the collected and analyzed data. The element of risk, as stated above, i.e. where there exist unknown parameters/factors, can be incorporated in the decision-making process by using the idea of "expected utility", which represents the idea of "risk" in terms of probabilities that are taken into consideration for the final utility.

Thus, the decision-theory sub-field of the "probability-weighted utility theory" has emerged to represent the probability of each alternative option under different states (of nature or the system under investigation, etc.). According to utility theory, the use of probabilities in a non-subjective evaluation (such as the noon-intrusive data collection approach), are easier to generalize once they have been correctly determined for one person. The accuracy is expected to increase as the time of observation (or data collection in this case) increases. On the contrary, if utilities are subjective then the objective validity of the decision-making process is lost.

In Probability Theory, Bayesian Networks are subjective networks often regarded as *Belief Networks* (Barber, 2010); a factor that correlates with the subjectivity of MWQoE given the nature of user perception. Bayesian Networks are suitable with dealing with uncertainty and form causal probabilistic networks which fit perfectly with the causal nature of the MWQoE attributes as they affect each other, since context affects web metrics and vice versa. They are defined by a set of random variables which are the nodes in the network and a set of directed links are used to pair each node with another – thus forming causal links or relationships within the nodes. A node X that is linked with node Y denotes a direct influence of X to Y. Moreover, the strength of influence for each node to another is defined by conditional probabilities.

Although, many models for decision-making under risk have been developed, this work has selected to work with Bayesian models. The reason is that this set of models tackles the cases where the risk, and hence the probability in the decision model, are subjective and not the result of nature or the system. In fact, the current study makes a characterization of a subjective experience by using a non-intrusive approach and hence must consider that there exist subjective probabilities in the characterization of this experience. However, the model principles consider that on the one hand these subjective beliefs are coherent, and, on the other hand, these subjective beliefs can be based on (and changed according to) a series of observations made by the subject.

In addition, Bayesian Networks have been gaining momentum in academia and in industry, especially in the field of Artificial Intelligence as shown in Friedman and Goldszmidt (1996), Jensen and Nielsen (2007) and Barber (2010), since they can support both Discrete and Continuous datasets, are data-driven and are relatively easy to implement. Moreover, they can be considered *flexible* by the fact that '*the degree of belief in a given scenario will converge with the limiting frequency regardless of the initial degree of belief (hypothesis)*' (Ligeza, 1995).

The MWQoE model treats each data attribute as a Bayesian Network node. The conditional dependence relationships in the Bayesian network are coded based on findings from Fiedler, Hossfeld, & Tran-Gia (2010), Schatz and Egger (2011) and Sebastian Egger *et al.* (2012) as well as findings from this study's data analysis using Naïve Bayes modeling. Naïve Bayes modeling (Lewis, 1998) uncovers hidden relationships of attributes given a set of training data of observations of a specific entity. These relationships uncovered in this study are used as *expert knowledge* which defines the relationship between each context state. The Bayesian Network's predictive nature makes it an ideal framework for this study as the collected dataset can be used to evaluate MWQoE on each observed web session while it is continuously adjusted and fine-tuned with the introduction of new evidence.

In considering the appropriate technology for designing the MWQoE system, Cloud Computing (Cloud) was chosen, since it is used to compliment the processing and storage limitations of mobile devices. Cloud fits well with the spirit of this study since it considers the cost efficiency and scalability of the service provider, making the delivered solution more attractive for industry adaptations. Moreover, it provides large amounts of storage space and delivers demanding computational tasks in a fast and affordable way by distributing computational loads within a network of interconnected devices.

The option of using the device as the *processor* was rejected since it would, on one hand, eliminate the need for data relay from-and-to the Cloud, but on the other hand, would consume valuable device resources. Potential side effects of implementing a computationally intensive model (as it is outlined in Section 3.2.3) could be either a device crash or a degradation of the user's experience due to a poorly performing device; both of which are not acceptable for the purposes of this study. Moreover, the option of using a client-server architecture instead of the Cloud, was found insufficient since it would be harder to scale resources, in both processing and storage for real-world implementations.

The MWQoE model and metric are delivered via an original Mobile Cloud Computing (MCC) (Fernando, Loke & Rahayu, 2013) platform which is designed and implemented in order to enable the continuous evaluation of MWQoE, as well as, provide the framework for delivering potential experience enhancements to the end user in specific web scenarios. This novel end-to-end approach which is user-centered and geared towards evaluating and ultimately improving the user experience uses Cloud as an *experience* *processing hub*. It fully utilizes the capabilities of modern mobile devices such as tablets and smartphones to capture and interpret user context data on a non-intrusive automated manner. Moreover, this approach combines topics in MCC such as computation offloading by Kumar *et al.* (2013) and energy efficiency in sensor applications by Zhuang, Kim & Singh (2010) and Oshin, Poslad & Ma (2012), while it enables further enhancements of the MCC technology by using context-awareness not only for MWQoE but for service provisioning (Ting-Yi Lin *et al.*, 2013) from mobile to cloud and vice versa.

The non-intrusive data collection for the Model Generation Phase is achieved via an offloading procedure of transferring anonymous user data from the mobile device to the Cloud using a custom-built Android application, as shown on *Figure 3*. The anonymous data is processed through a series of Cloud component orchestration and the results are relayed back into the mobile device. This novel approach aims in achieving context-awareness by carefully considering the context of the user, the context of the web session, the context of the network connection and the context of the user's surrounding environment. This is achieved by inferring context from the attributes in each of the 6 data domains collected (Device, Network, Connection, Web, Session, Sensor, User).

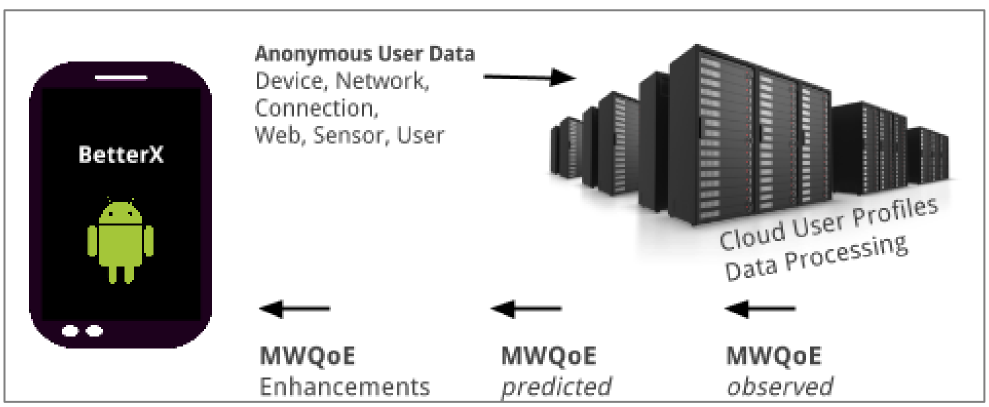


Figure 3 BetterX MCC Architecture

MWQoE is user-centric and measures the experience from the user's mobile device. The complete end-to-end Mobile-to-Cloud solution quantifies MWQoE for selected scenarios and on a per web-session basis. This study adopts the notion that QoE is highly subjective and personal to each distinct user and to a distinct web session and it cannot be handled in a *one size fits all* approach. Therefore, distinct web intentions and user profiles for each web session observed are developed which categorize anonymous data into clusters of QoE characteristics and by using specific and validated relationships characterize the experience of each web session on a unified scale. The importance of using anonymous data is derived from the non-intrusive nature of this study where the privacy of the user (Section 3.3.3.1) and the security of the user's data (Section 3.3.3.2) are considered.

3.2 Design

3.2.1 Participants & Apparatus

The intention of the BetterX data collection was to capture as much information as possible from the BetterX Android app (Figure 3) and filter the attributes which were found useful and were going to be used in the model during data analysis. The aim of the BetterX data collection, i.e. the 1st phase of the project, was to collect anonymous web session information from users all over the world together with context information using an Android application. The data collection phase aimed at getting more than 2000 web sessions together with readings from the user's device, the network, the connection and the device's available sensors. No geographical or other restrictions were set to participants except that of the minimum age limit of 18. Hence, for the 1st data collection phase, the BetterX project was created.

The BetterX project[1] refers to all the research components, including the two data collection phases, the implementation of the BetterX Cloud Setup (Section 3.2.1.1), the BetterX Android app (Section 3.3), and all the pieces of code which were created to actualize the MWQoE model and metric.

The BetterX end-to-end MCC solution enables the continuous data collection, evaluation, and potential enhancement of MWQoE in a live real word setting. It was designed as a turnkey solution for online service and content providers who need more insights into their user's perceived QoE. The design and technical considerations of BetterX Cloud (Allayiotis and Antoniou, 2014) are outlined in the following section.

3.2.1.1 BetterX Cloud – Design & Technical Considerations

BetterX Cloud, as shown in Figure 4 utilizes a unique orchestration of functional components to provide the processing requirements for the MWQoE metric and potential enhancements. The successful content delivery through a functional division of tasks between the mobile device and the Cloud is described by

[1] *http://www.betterx.org*

separating the description into the specific functional components comprising the architecture and handling particular aspects of the delivery. The innovation of this architecture originates from the fact that it allows optimizations to be enabled at all layers of the content delivery process, i.e. in the receiving layer, the processing layer, the storing layer and the sending layer, making this architecture flexible and efficient. For example, the storing layer can be optimized using distributed storage and replication for faster throughput, the processing layer can use parallel processing to minimize processing time, etc. It is a novel approach which collects data from the user side, transmits the data using a context-aware manner to the Cloud, the Cloud acts as a processing hub and interprets the data by building user profiles and compiling the MWQoE metric, and finally, given specific user scenarios, delivers potential enhancements to the user as an attempt to enhance the experience. The cycle of continuous MWQoE evaluation is intended not only to reveal insights about the user's experience, but also to continuously adopt and improve based on new data (evidence) from the user.

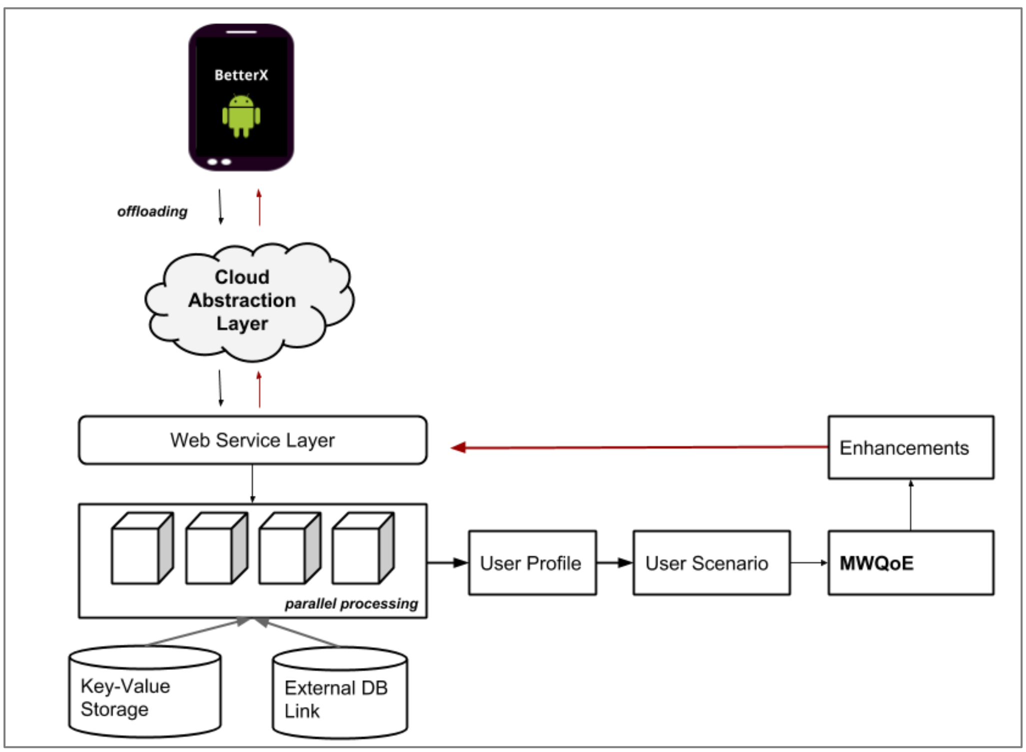


Figure 4 BetterX Cloud – Components Orchestration

3.2.1.1.1 Guiding Cloud Paradigms

The BetterX Cloud has been guided and motivated by 2 important paradigms in the Cloud realm. The first is that Mobile-to-Cloud (MCC) architectures can be an efficient vehicle of migrating computational and storage capabilities from the mobile device as long as costs (in terms of both bandwidth and strain of resources such as power) are carefully considered. The second is that Cloud computing can be used efficiently in many technology principles given that its scalability and parallelism capabilities are fully exploited.

In regards to the second motivation, this study adopts the notion that '*computing, storage, and networking [should] focus on the horizontal scalability of virtualized resources rather than on single node performance*' (Armbrust *et al.*, 2010). This is reflected in the BetterX Design, shown in Figure 4, as considerations for the parallel operations of both data storage, and data processing are realized. In addition, research focusing on replication strategies in Cloud settings that involve optimizing both data processing and network latency (Stuedi and Terry, 2010) is considered. The BetterX approach aims to optimize processing by accommodating parallel execution via replication, and, moving data to nodes that are closer to the location of the user thus minimizing network latency.

Furthermore, BetterX considers the importance of security and privacy of the users. In doing so, it utilizes a user generated security mechanism with user privacy features which secures user data on the mobile device prior to transmitting it (Mowbray and Pearson, 2009). Details of all data encryption and data obfuscation mechanisms employed in this study are found in Section *3.3.3.2 Security Measures* and Section *3.3.3.1User Anonymity* section respectively.

3.2.1.1.2 The 4 Linked Layers of BetterX Cloud

BetterX Cloud is divided into 4 linked layers: Data Receiving layer, Data Processing layer, Data Storing layer, and Data Sending layer (Allayiotis and Antoniou, 2014).

3.2.1.1.2.1 Data Receiving layer

The Data receiving layer (Figure 5) is composed of a RESTful web service (Greaves *et al.*, 2008) which accepts HTTP GET and POST requests from the mobile device. All requests are received from the Data Receiving layer and are sent to the Data Processing layer. Amazon's Elastic Load Balancer (Amazon Web

Services, 2015c), a load balancing mechanism, is deployed to scale requests into multiple processing nodes and ensure high availability of service. Each processing node is an instance of Amazon's Elastic Compute nodes *EC2* (Amazon Web Services, 2010). Scaling is considered in the design of the Data Receiving layer of BetterX since it is intended for real-world adaptations where fault-tolerance measures for application robustness and reliability are considered.

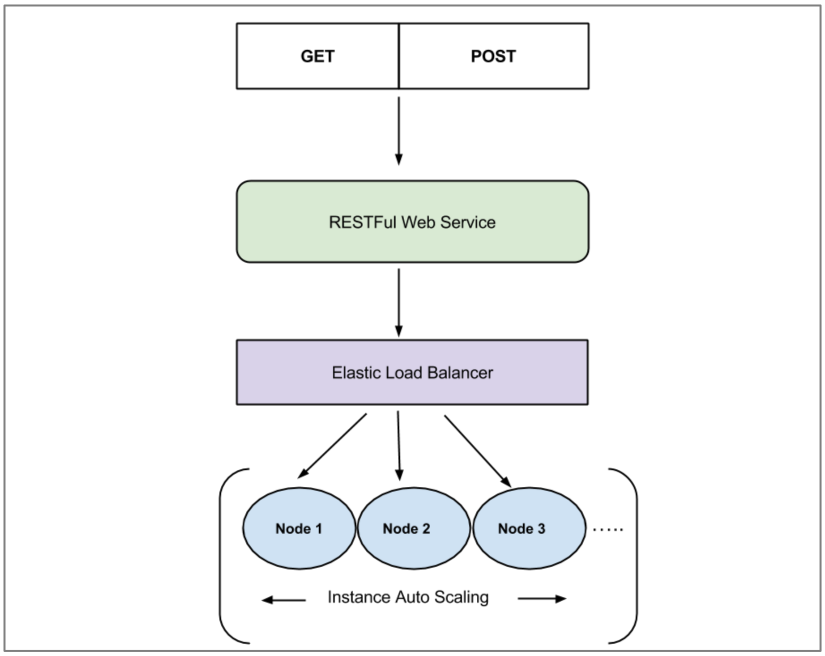


Figure 5 BetterX Cloud - Data Receiving layer

3.2.1.1.2.2 Data Processing layer

The Data Processing layer utilizes the MWQoE predictive model to calculate the MWQoE metric considering the Data Collection Attribute List (Table 1) to derive the Web Session User Intent (Figure 9) and identify the scenarios in which a mobile web session has taken place. The Data Processing layer holds a series of Cloud Processing Nodes (Figure 6 BetterX Cloud - Data Processing layer) each one acting as a district and autonomous processing hub. The Data Processing Node is designed so that it can be executed in parallel and utilize the full capabilities of the Cloud infrastructure.

In addition, the Data Processing layer is designed so that that certain MWQoE enhancement features could potentially become part of the Cloud Processing Node in which instructions for specific MWQoE enhancements could be triggered based on the context of a user in a specific web session, and sent to the mobile device for execution. This study has tested 2 potential enhancement features and examined their effect to Web QoE (Section 4.2.6). However, they are not included in this section since they were not considered in the design of the Model Generation phase but rather were identified in the design of the Model Evaluation phase of the project.

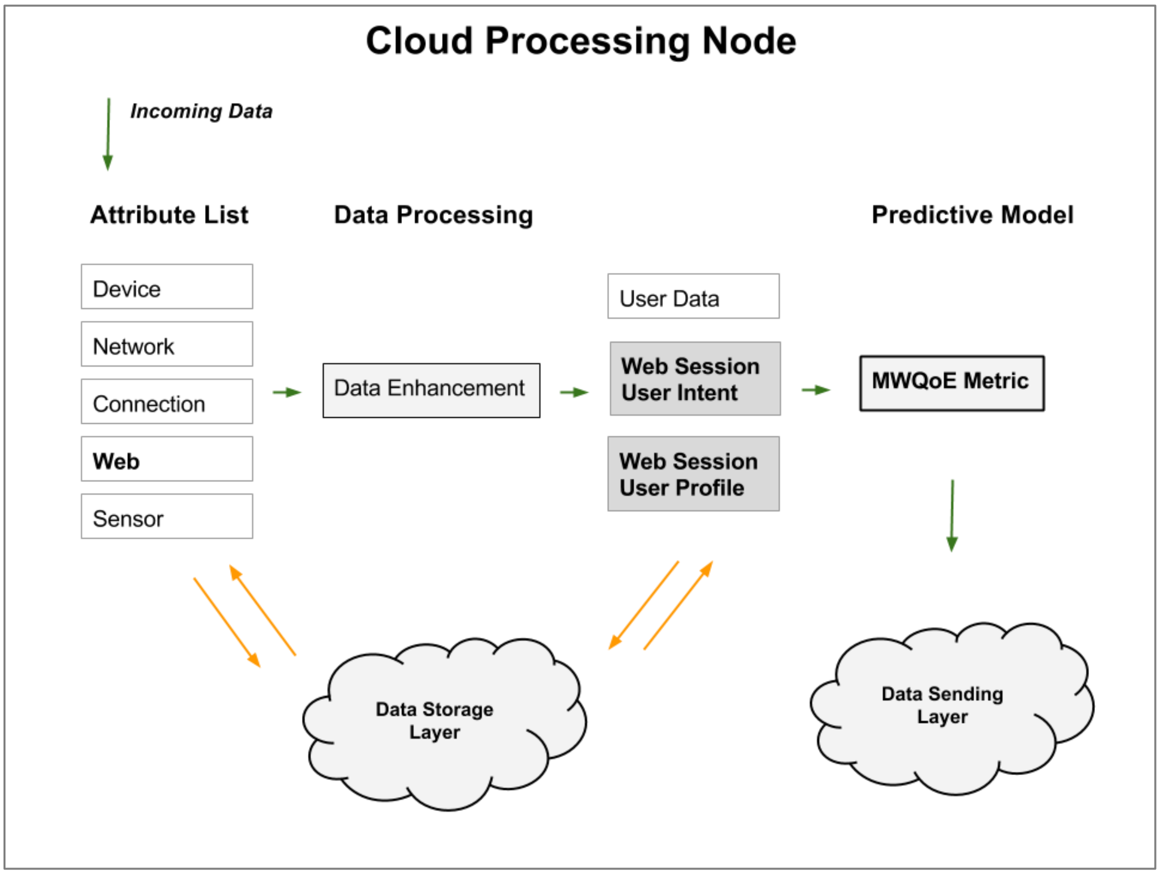


Figure 6 BetterX Cloud - Data Processing layer

3.2.1.1.2.3 Data Storing layer

The results from processing are sent to the Data Storing layer which stores and archives data in the user's database. The Data Storing layer utilizes Apache Cassandra (Lakshman and Malik, 2010) for the Sensor attributes and Oracle MySQL (Oracle Corporation, 2012) for the Device, Network, Connection, Web and User attributes. The structure for the Data Storing node in is in key-value pairs (Figure 7). Apache Cassandra is chosen since it is found to be a fast, highly scalable and fault-tolerance database widely adopted in industry as indicated by Alvaro *et al.* (2010) and Newswire (2012) and used in conjunction with many data analytics cluster computing frameworks, such as Apache Spark (Armbrust *et al.*, 2015) which is one of the data analysis tools used in this study. On the other hand, Oracle MySQL is a widely-used database management system, robust, highly efficient and extensively developed.

3.2.1.1.2.4 Data Sending Layer

The results from Data Processing are sent to the Data Sending layer, as illustrated in Figure 8, which compresses the data and structures them in JSON (Douglas, 2006). It has been shown in literature that compressed JSON format data via a Restful interface is an efficient way of transporting data from the cloud to a mobile device (Gil and Trezentos, 2011).

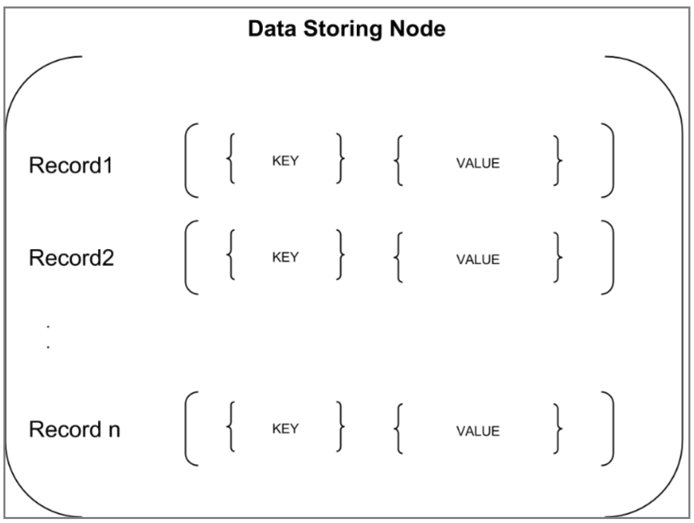


Figure 7 BetterX Cloud - Data Storing layer

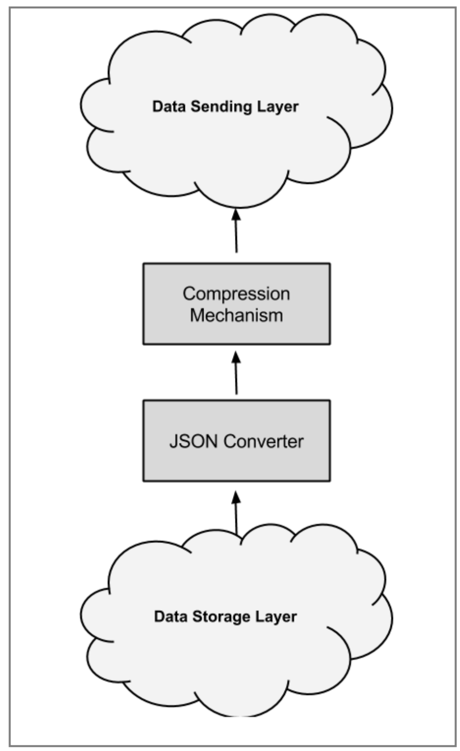


Figure 8 BetterX Cloud - Data Sending layer

3.2.2 Procedures

The list of data attributes designed to be collected by the BetterX Android application covered 6 different domains: 1) Device, 2) Network, 3) Connection, 4) Web, 5) Sensor, and 6) User. The list contained a comprehensive set of attributes which were associated with the adopted ITU-T QoE definition (International Telecommunication Union, 2006) and were made available by the majority of the mobile devices in market today. It was a large dataset composed of 65 attributes aimed in capturing and understanding the user's physical context, the network and connection state of the mobile device, the capabilities of the device and more importantly the web activity. Given the non-intrusive nature of this study, the Data Collection Attribute List, as shown in Table 1*,* was designed to enable an *accurate* interpretation of the user's surrounding conditions when using the web from a mobile device. Context-states were aimed to be derived from the Data Collection Attribute List and to be examined alongside web usage metrics for understanding the user scenarios of each of the observed web sessions and evaluating the experience index (MWQoE metric) for each one of them.

It is important to note that the BetterX MWQoE model was designed to not depend on QoS metrics but directly consider them, as they were examined in previous QoE studies illustrating QoS to QoE correlations (Gao *et al.*, 2009; Gong *et al.*, 2009; Fiedler, M.; Hossfeld, T.; Tran-Gia, 2010; Cui and Biersack, 2012). The BetterX MWQoE approach was user-centered since it viewed connection and network information from the client side rather than the network side. More importantly, the network speeds and device capabilities were linked with content and page loading timings which together with device state/capabilities metrics were interpreted into user perceived page loading time metrics. This was designed to be one of the main differences in the approach of this study as compared to previous QoE studies; that it considered experience as a highly personal matter.

Moreover, the novelty of this approach is found in the user-centered, non-intrusive theme of this study where context-attributes are fused into a single-scale metric. The metric is constructed by observing quantifiable attributes such as network metrics with inferred context-states such as the intention of the user for a specific web session and the importance of the web session to the user (as it is illustrated in Section 3.2.3 Preliminary Conceptual Model).

The Device attributes captured basic information about the device, its capabilities and screen size, as well as a list of applications running and basic phone actions such as phone calls, screen switch offs, etc. The Network attributes captured the state of the network as it was received from the network provider (either Mobile Network or Wi-Fi). The Connection attributes outlined the current and available connections of the device. The Web attributes captured all available web session HTTP metrics from the Application Layer (Day and Zimmermann, 1983) such as the Content Load (the number of milliseconds for the content of the page to be loaded from the initial request) and Page Load (the number of millisecond for the page to be completely loaded; when the *onLoad* event is fired) as well as HTTP POST/GET request details (Fielding *et al.*, 1999). The Web attributes were based on the HTTP Archive format (HAR) (Odvarko, Jain and Davies, 2012). For retaining user anonymity in HTTP request and response data, content was not captured and values from the request and response parameters were stripped away. The Sensor attributes logged data for each of the sensors that were available on the mobile device. The User attributes were captured via a very basic questionnaire upon installation of the BetterX Android App (Figure 12 - Figure 18).

Device
- Timestamp
- Device Id
- Device Manufacturer and Model Number
- Screen Size
- Battery: Level / Status / Temp / Voltage / Health
- Browser Type / Version
- Apps Launches
- Location: Latitude / Longitude / Accuracy
- Screen Status: On/Off
- Phone Events: On Call, etc…

Network
- Link Speed
- Has Internet
- Mobile Status
- Wifi Status
- Wimax Status
- Signal Strength
- Frequency
- SSID
- RSSI
- IP
- Detailed State
- Net Available
- Net Capabilities
- Down Bandwidth
- Up Bandwidth
- Transport Types
- MTU
- TCP Buffer Size

Connection
- Conn. Status: Available / Connected / Connecting
- Network Type
- Roaming
- Strength

Web
- Tab: Id / Status
- Page: Id / Url*
- Page Start Time*
- Page Load Timings
  - Content Load
  - Page Load
  - Blocked
  - DNS
  - Connect
  - Send
  - Wait
  - Receive
  - SSL
- Connection Info
- HTTP Request Details
- HTTP Response Details
- Cache Details (Before and After Request)

Sensor
- Accelerometer
- Gyroscope
- Magnetic Field
- Light
- Temperature
- Step Counter
- GPS

User
- Age
- Gender
- Education
- Timezone
- Location
- Phone Use Freq.
- Web Use Freq.

Table 1 BetterX Android - Data Collection Attribute List

All Timing metrics were captured in milliseconds timestamps in *Coordinated Universal Time* (UTC) and each timestamp was localized to the user's local time using the user's Timezone.

The Collected Attribute list as shown in Table 1 was replaced during data analysis with a *revised* Data Attribute list which kept only the attributes that were found suitable and useful for the MWQoE model (Table 6). The revised data attribute list is discussed in Section *3.3.3.3.*

Certain collected attributes were intended to be used either to extract, classify or estimate further information about the user in order to produce an estimation of the *Web Session User Intent* (Figure 9). The *Web Session User Intent* was designed to be inferred by examining the HTTP Response and Request details (which revealed further insights about the interest of the user for the particular domain), the Web Session Duration, the User Local Time, the Domain and Domain Category, the User Physical Context (Activity, Location) and Frequencies (which revealed potential Time-Location-Domain relationships). The estimated *Web Session User Intent* revealed the *Web Session User Profile* (the classification of the type of the user). For example, the sample data in Table 2 below leaded to the following narrative about the observed user web sessions: *The user is a 20-year-old student (derived from Demographics) sitting (derived from Activity) in a shopping mall (derived from Physical Context) visiting the website of a nearby cinema (Domain, Domain Category) with the intent to purchase movie show time tickets (Frequencies, HTTP Request and Response)*.

Domains were designed to be extracted from URLs (Figure 9) and categorized (Domain Category) using Amazon's Alexa API (Amazon Web Services, 2016). The Web Session Duration was estimated by examining the Tab Status (Open/Closed) from which the web session was generated, the App Launches (if any apps were launched during the web session) and the Web Session Timestamp (the time which the web session was initiated).

In addition, accelerometer data was designed to be used in estimating the User Activity using a feature extraction algorithm by Kwapisz, Weiss & Moore (2011). User Activity was broken down into 6 categories: *walking*, *jogging*, *ascending stairs*, *descending stairs*, *sitting* and *standing*.

Moreover, the captured GPS traces (latitude and longitude) were reverse geocoded using the Google Maps API (Google Inc., 2015) in order to classify the User Location. The User Location was manually examined and classified into distinct User Location Categories which included: Residential Neighborhood, Commercial Location, City Center, Shopping Center, University, Local Store, etc. The User Activity together with the User Location and User Location Category were used to estimate the User's Physical Context.

Domain	Attribute(s)	Processing	Inferred Context
User	Demographics	Get user demographics from database	User Age = 20 User Profile = Student
User	Time (UTC) / User Timezone	Translate time to user local time	User Local Time = 4pm
Sensor	Accelerometer	Process accelerometer data to reveal user activity	User Activity = Sitting
Sensor	GPS (Latitude / Longitude)	Reverse geocode coordinates to reveal location type	User Location Type = Shopping Center
		Web Session 1	
Web	HTTP Request	Domain = www.kcineplex.com Domain Type = Movie Theater Request = GET Resource = /view_showtimes.php	Web Session User Intent = Browsing for movie show times
		➔ Inferred Web Session User Profile = Web User (Browsing)	
		Web Session 2	
Web	HTTP Request	Domain = www.kcineplex.com Domain Type = Movie Theater Request = POST Resource = /buy_tickets.php	Web Session User Intent = Purchase movie ticket(s)
		➔ Inferred Web Session User Profile = Web Buyer (Online Tickets)	

Table 2 Sample Data for User Intent & User Profile

For each of the 6 data domains collected, the study was designed to examine *how*, and *to what extent*, each attribute affected MWQoE. Device data was used to examine how MWQoE was affected by device characteristics, network data was used to examine how MWQoE changes based on network changes and so forth. The Data Collection Attribute List (Table 1) combined with the Web Session User Intent (Figure 9) provided a novel fusion of attributes.

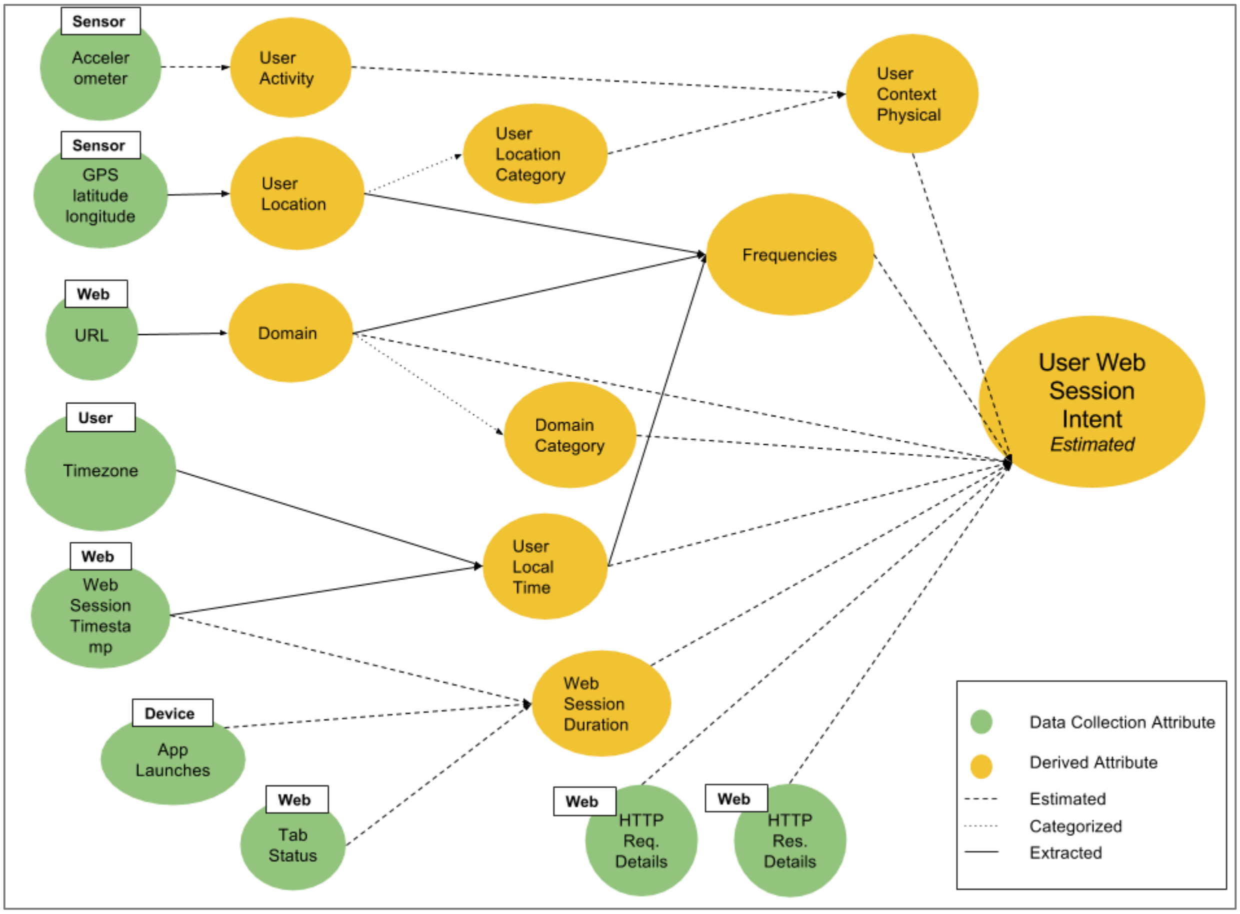


Figure 9 User Web Session Intent

3.2.3 Preliminary Conceptual Model

The MWQoE model at its concept (prior to data analysis) was designed as an artifact of calculation for each of the 6 data domain attributes and the *Web Session User Intent*. As shown in Figure 10, the MWQoE model reflected the QoE of a specific user at a specific time and for a specific mobile web session.

$$MWQoE_{(User,\ Time)} = \boldsymbol{f}\ (Device, Network\ , Connection, Web, Sensor, User, Web\ Session\ Intent)$$

Figure 10 MWQoE Concept

The design of the MWQoE model considered the quantification of the context-states of the Device, Network, Connection, Web, Sensor, User and the Web Session User Intent to deliver the MWQoE metric in a unified scale. The BetterX project, as outlined in the rest of the section, used the conceptual MWQoE model (Figure 10) as the basis for the generation of the final MWQoE model and metric. The BetterX project captured live device data from users all over the world and enabled the identification of relationships of the selected attributes and the quantification of their effect to the user's mobile Web QoE, producing the MWQoE model and metric organically.

3.3 Implementation

For this study, BetterX Cloud was partially implemented, as compared to the initial Cloud infrastructure design elaborated in Section 3.2.1.1. Specifically, although discussed, the implementation of the complete end-to-end BetterX Cloud solution was considered beyond the scope of this thesis since it required a significant amount of time to implement and deploy on cloud infrastructure. All crucial data analysis components for composing and delivering the MWQoE metric were implemented and a proof-of-concept is actualized in this thesis. The BetterX Android was fully implemented per the Data Collection Attribute List (Table 1) and the details of its implementation together with the details of the implementation of the BetterX Cloud are outlined in this section.

BetterX Android is this study's primary data collection vehicle. It is a custom-built Android application which runs as a non-intrusive background service and collects the attributes listed in the Data Collection Attribute List Table 1. *Figure 11 BetterX Components & Services* outlines the major functional components of the application, the attributes that are captured for each component, and the linkage of each component via a network connection to the BetterX Cloud.

BetterX Android consists of 3 data loggers, a Firefox web browser add-on and a file transfer service; all of which are enabled after setup as background services on the Android operating system. The 3 loggers are: 1) Sensor Logger which captures the Sensor data domain attributes (as outlined in Table 1), 2) Network Logger which captures Network and Connection attributes, and 3) Features Logger which captures Device and User attributes. Each logger creates 1 log file and a file transfer service sends all log files generated to an Amazon Simple Storage Service cloud bucket (S3) (Amazon Web Services, 2015b) via a secure connection. Each device transfers files once a day and files are compressed into 1 archive file and encrypted using a public key prior to transmission. Although the files are sent to the Cloud via a secure connection, encryption to each file is added as an additional security layer so that even in a case of unauthorized access to a processing or storage node, the user data stays encrypted. The user is given the option to define when the relay takes place so that it does not interfere with other tasks. The default option is at night when the device is connected to Wi-Fi and while the device is being charged.

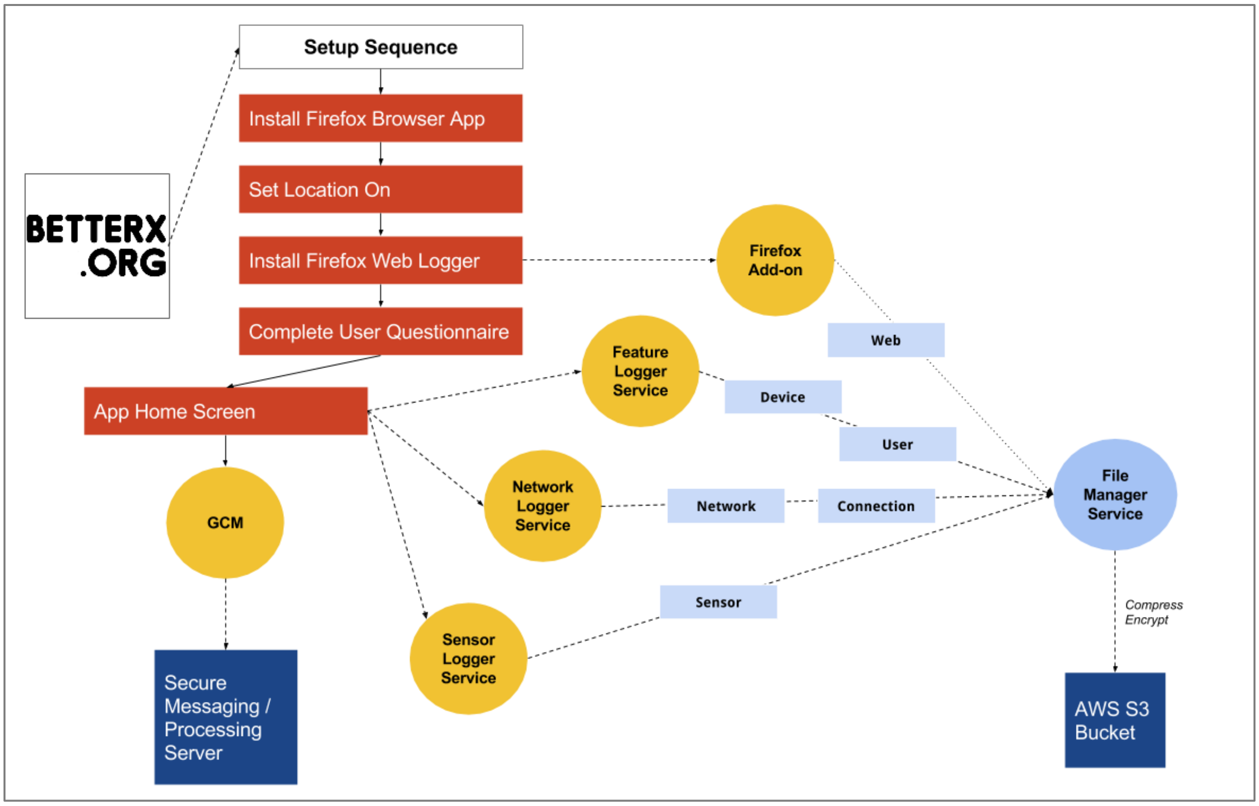


Figure 11 BetterX Components & Services

All Web attributes are captured using a custom Firefox add-on which is installed during setup (Figure 16 - Figure 18). Users are asked to use Firefox as their default mobile web browser so that all user-generated browser web traffic is captured. Although it is noted that Mozilla's Firefox is not the most popular browser for the Android platform (StatCounter, 2016), it is however the only browser which provides an API to capture all the required web metrics from this study.

The loggers save data into JSON formatted flat files during various intervals or when a certain action triggers a change in an attribute being monitored. The File Manager Service collects all generated files in a daily basis and compresses and encrypts them before sending the them to the Cloud. The transmission is done via a secure connection and all other interactions of the app with back-end web services are done via SSL secured connections. Google Cloud Messaging (GCM) is used to push notifications from the Cloud to users and vice versa and it chosen for this study since it allows for user anonymity.

Upon installation, the application generates the User Id (UID) which is unique for every Android device and is used to reference the user in the BetterX system. The application collects anonymous information about the user therefore no user registration screen is provided. For collecting demographic information, a very brief setup questionnaire (Figure 13 - Figure 15) is presented to the user during setup. The questionnaire captures gender, age, education, user location, and asks the user to define how frequently he/she uses the device and the web browser on the device. During setup, the user is asked to accept the BetterX Privacy Policy (*Appendix Privacy Policy)* and End-User License Agreement (*Appendix End-User License Agreement*) both of which guarantee the user's anonymity and outline how the user's data is used throughout the BetterX project.

The BetterX main screen, as shown in Figure 19 displays the UID (or Device Id) of the user and is used as a reference for any communications between the user and the project team. It also shows the status of the background services running on the device. In case of an application error or failure on any of the 3 service loggers, a message is shown to the user asking for the application to be restarted. The main menu (Figure 20) has 3 options: 1) Messages (send and view received messages), 2) Tickets (shows how many raffle tickets have been generated and allows the user to gain additional raffle tickets by sharing the app with contacts on their phone), 3) Support (Frequently Asked Questions section). Further information on Tickets is given in the Campaign section below (S*ection 3.3.3*).

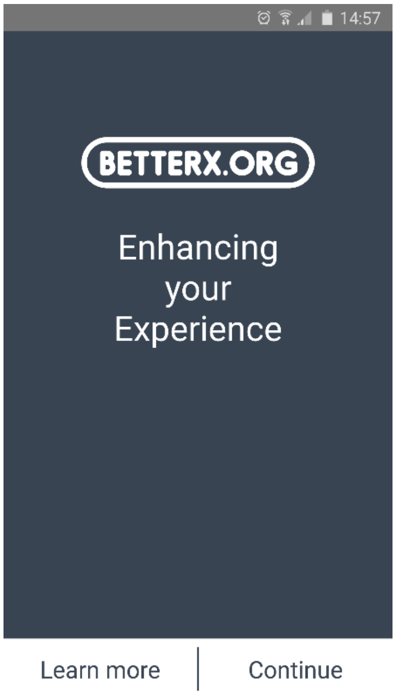


Figure 12 BetterX Android – Setup screen 1

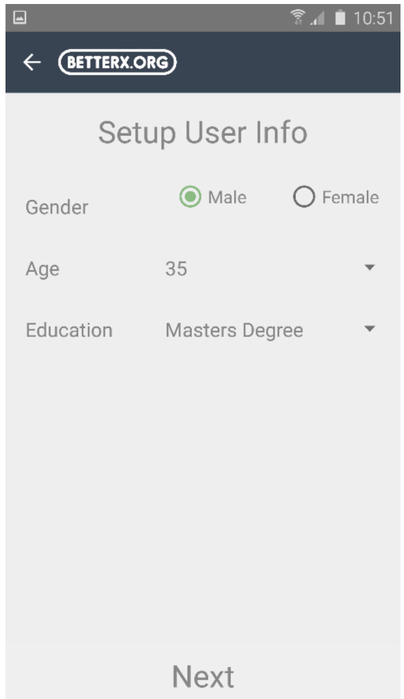


Figure 13 BetterX Android – Setup screen 2

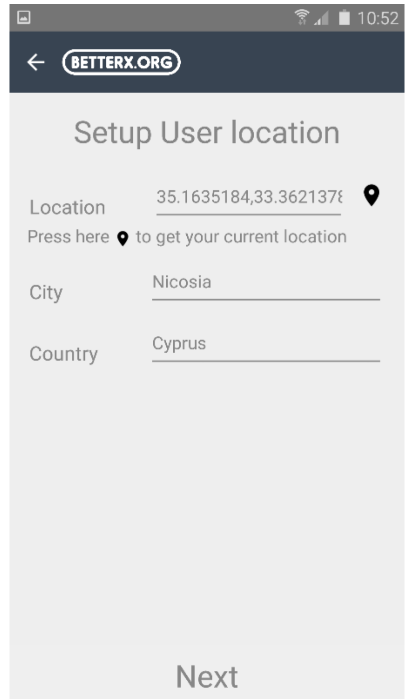


Figure 14 BetterX Android – Setup screen 3

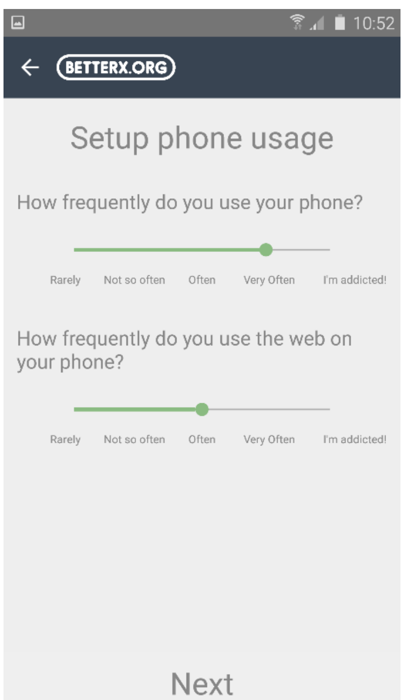


Figure 15 BetterX Android – Setup screen 4

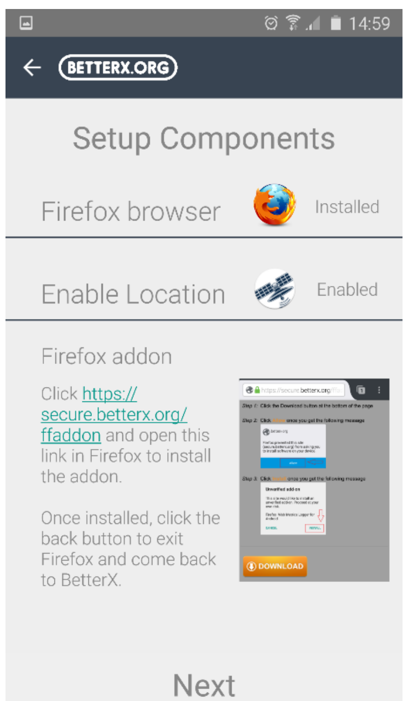


Figure 16 BetterX Android – Setup screen 5

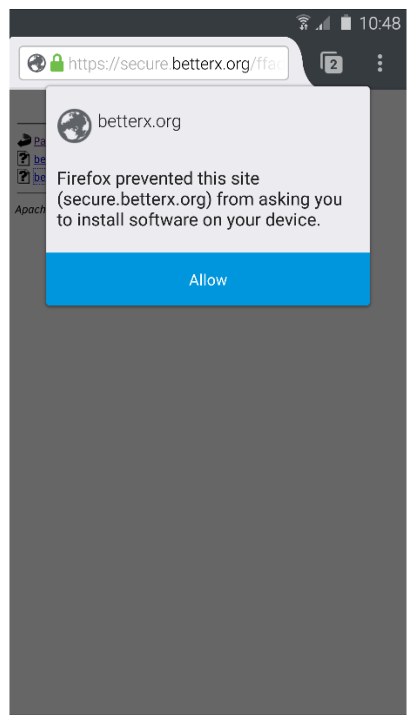


Figure 17 BetterX Android – Setup screen 6

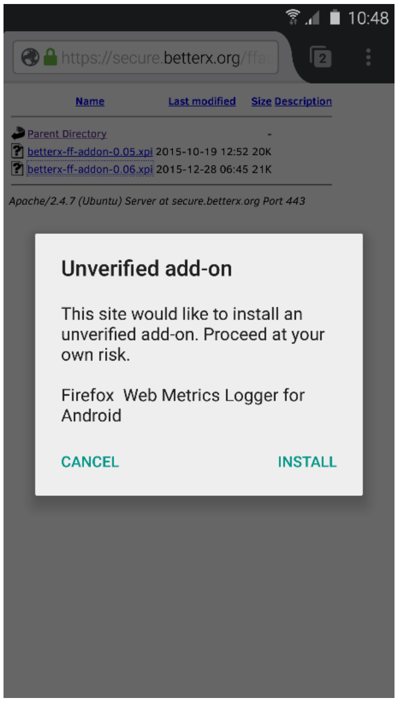


Figure 18 BetterX Android – Setup screen 7

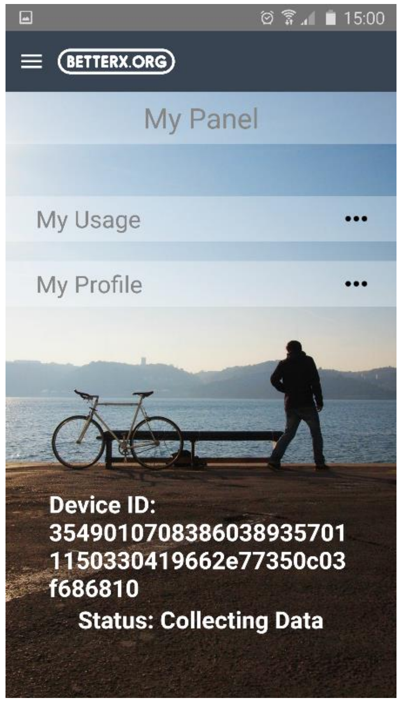


Figure 19 BetterX Android – Home screen 1

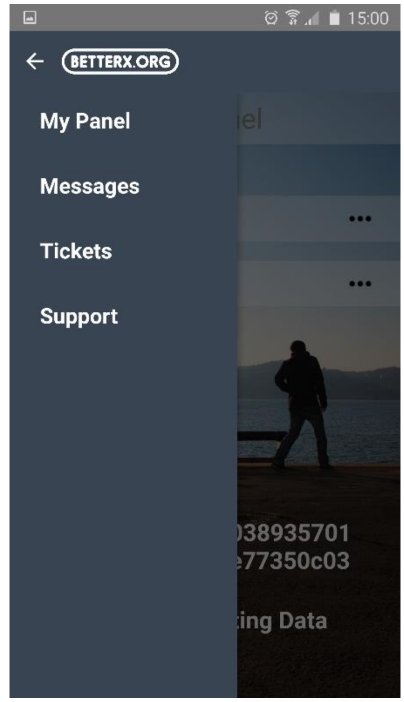


Figure 20 BetterX Android – Home screen 2

3.3.1 Technical Details

The open source *SensorManager* library for Android (Lathia *et al.*, 2013) is used to provide the data from all sensors and the *SensorDataManager*[2] library is used for managing the data in the app. *SensorManager* uses an adaptive sampling mechanism which reduces the volume of the recorded data. The Android platform provides 2 types of settings for capturing sensor data: a fixed-window setting which captures data based on a preset static frequency, and a variable-window setting which captures data based on dynamic frequency. Moreover, each attribute used by *SensorManager* is classified either as a *Pull* or *Push* attribute (Table 3). *Pull* attributes need a call invocation to capture a reading and *Push* attributes broadcast the readings automatically to all subscribed entities. In addition, the Network logger uses the *Network Events Library*[3] to monitor and record the Network and Connection attributes. The Network logger uses 4 distinct state changes to capture new readings: when the Network state changes, when the Signal Strength changes, when the Network Capabilities change and when the Link Properties change.

[2] *https://github.com/emotionsense/SensorDataManager*
[3] *https://github.com/pwittchen/NetworkEvents/*

The File Manager Service uses the UID to name files prior to sending them to AWS S3 storage. This UID is derived from the *IMEI* (International Mobile Station Equipment Identity), the *MEID* (Mobile Equipment Identifier), the *SIMESERIAL* (14-digit SIM card serial number), and the *ANDROID_ID* (Android Device ID); all of which are collected from the Android Platform API. *Figure 21* shows an example of how a sensor file is been named before transmission by the BetterX Android application.

Pull Sensors	**Push Sensors**
Accelerometer	Battery
Gyroscope	Connection State
Location (GPS)	Connection Strength
Magnetic Field	Location
WiFi	Screen
	Phone
	Step Counter

Table 3 BetterX Sensor Classification

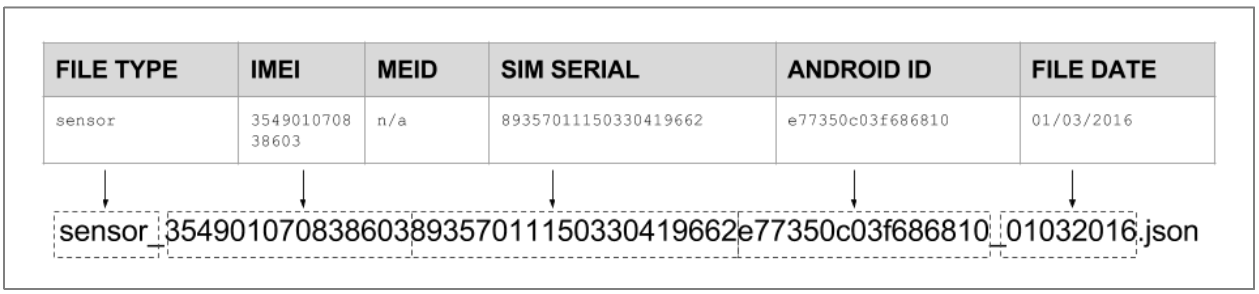


Figure 21 BetterX UID Components

3.3.2 BetterX Data Management

A series of custom python scripts is implemented for importing files from AWS S3 to the processing nodes. The main data loading script copies each zip file to the processing nodes' local disk, unzips and decrypts the contents of the file and parses the name of the file (*Figure 21*) to identify the UID and the file date. A custom python parser validates the format of each file and imports its contents to MySQL (Web, User, Device, Network attributes) and Cassandra (Sensor attributes). There are 6 different types of files: 1) Tickets file which contains the raffle tickets from each user (tickets where used to award participants with prizes at the end of the data collection campaign – see Section 3.3.33.3.3 BetterX Data Collection Campaign for more information), 2) Setup file which contains all the device information of the file plus the setup questionnaire responses, 3) Sensors file which includes all the sensor data, 4) Network file with all the network information, 5) Features file with the log of the device app usage plus the available features of the device and 6) the Web file with all the web metrics captured from the BetterX Firefox Addon. The main loading script moves all successfully loaded files into a secondary storage space to be archived. Data loading is done automatically in batches and every batch is verified by comparing record counts from the loading scripts reports and the database.

A total of 9 python scripts[4] are created for managing BetterX as shown in Table 4. Additionally, a PHP based back-end interface (A*ppendix BetterX Back-End Interface*) is created in one of the Processing Nodes to accommodate easy tracking of files and user messaging.

File Name	**Description**
data-loader.py	BetterX main data loading script which takes all files from an AWS S3 bucket and loads them into a database
db-importer.py	Helper file for the data loader which verifies JSON formats and maps to the appropriate database
db-attributes.py	Helper file for the data loader which maps JSON fields into database fields
reverse-geocode.py	Uses the Google Maps API to reverse geocode location coordinates and derive address and further information about a place
parse-domain.py	Parse a URL into domain, subdomain and extension
alexa-api-search.py	Uses the Alexa API to classify domains
user-stories.py	Creates flat files for each user for each day of use for easier user analysis.
labs-data-import.py	Lab tests data loader
labs-attributes.py	Helper file for Lab tests data loader which maps attributes into database fields.

Table 4 Data Management Scripts

In regards to the MySQL database, various schemas are created for a total of 56 tables. The most important schema is considered to be the Web attribute schema which is a normalized relational representation of the HAR file format (Odvarko, Jain & Davies, 2012) in its anonymous form. Figure 22 shows the Entity-Relationship (ER) Diagram of the Web schema, and it is regarded as one of the minor contributions of this study since it is the first published normalized schema for anonymous HTTP traffic capture.

[4] *The source code of all BetterX files together with the source code of BetterX Android and the BetterX database are available online on the project's website: www.betterx.org*

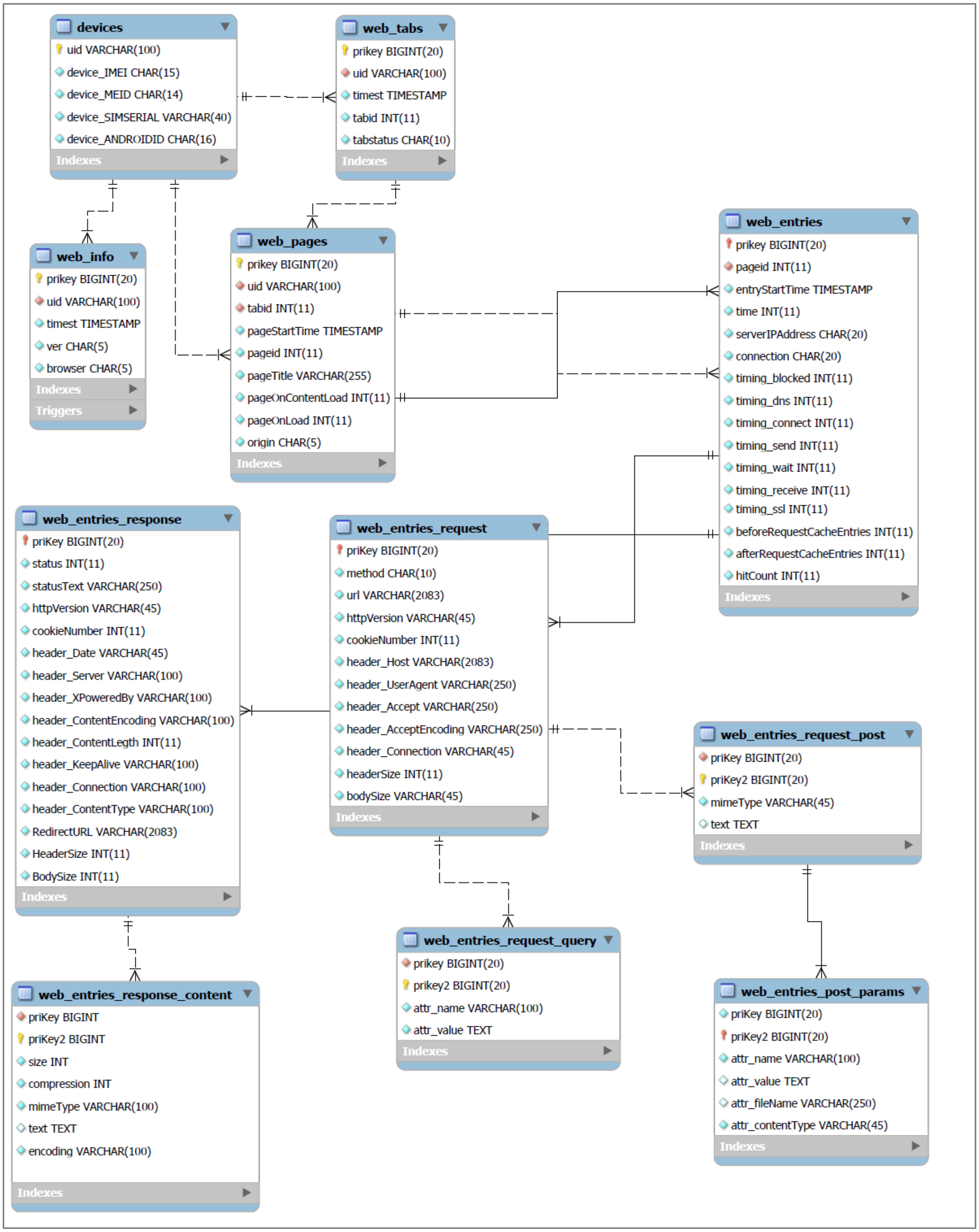


Figure 22 ER Diagram of BetterX Web attributes

In the BetterX Cassandra database a very simple model is created to store the 4 types of Sensor attributes (Table 1): Accelerometer, Gyroscope, Magnetic Field, and Light. For the first 3 tables the primary key is composed of the UID (text datatype) and the millisecond Timestamp (bigint datatype), followed by xAxis, yAxis and zAxis fields (double datatype). For the Light table a *light* integer field is used along with the UID and Millisecond Timestamp key.

3.3.3 BetterX Data Collection Campaign

BetterX was designed to recruit anonymous users online via web advertising and crowd-sourcing techniques and collect data automatically via BetterX Android in a non-intrusive manner. All the BetterX data collection campaigns, the third-party services used, and the data handling endpoints are outlined in Figure 23. Figure 23 shows the different user recruitment methods and how they are linked with each distinct aspect or service employed for data collection and data enhancement.

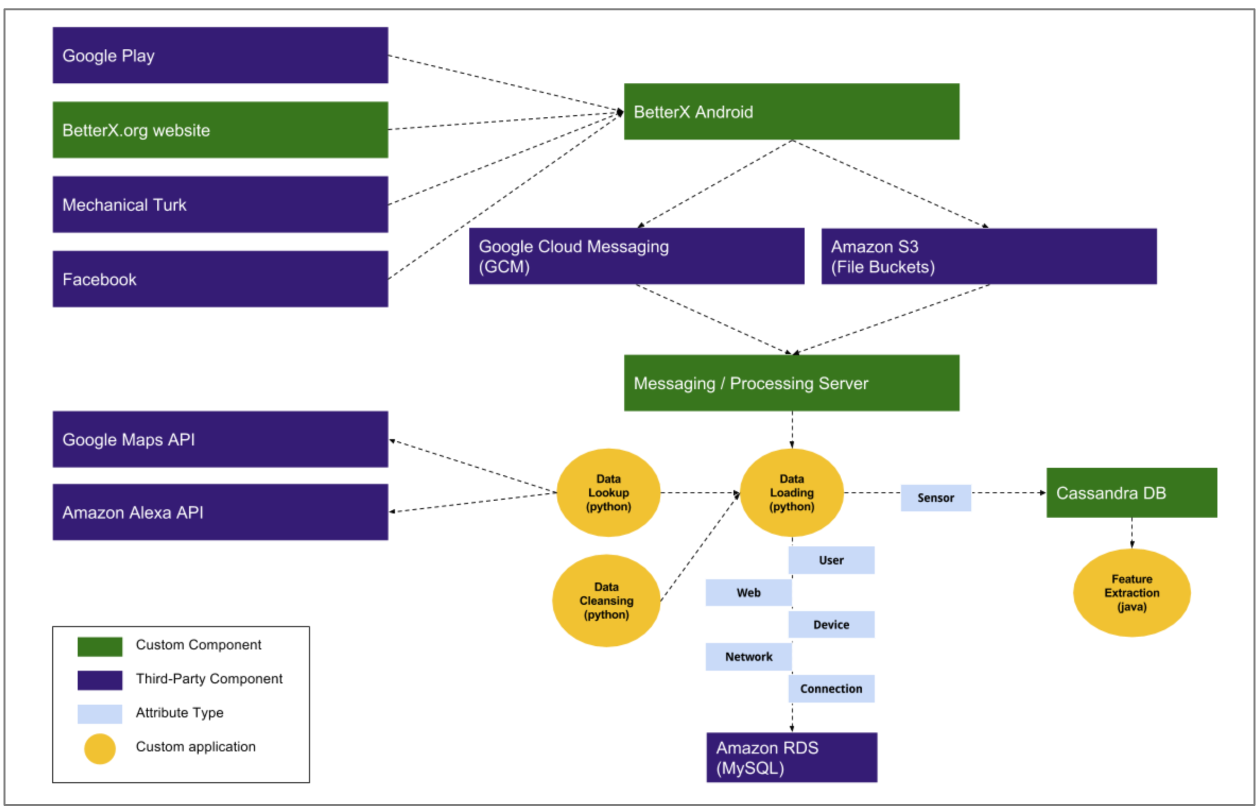


Figure 23 BetterX Campaign, Service & Data Endpoints

In the BetterX Data Collection Campaign, users were recruited to install BetterX Android on their mobile device (either smartphone or tablet) from the project's website (*www.betterx.org*) and from Amazon's

Mechanical Turk platform (Amazon Web Services, 2015a). By using this approach, this study consciously sacrificed control of variables from a lab environment with collecting data from an uncontrolled live environment. The advantage of this approach was that, on one hand, it could yield a large real-world dataset, one that would have been prohibitive to collect in a lab environment given the scope of the study and time/resource limitations. The disadvantage, on the other hand, was that it could not record user opinions on web experience given that data collection was non-intrusive. To supplement this limitation, a second data collection phase was executed in a lab environment where user opinions were recorded.

The primary goal was to yield a rich web usage database from a large and diverse population; something that was found to be prohibitive on a lab setting from both a time perspective and a resource perspective. The quantity of data was essential for this study as the MWQoE metric by definition, as many predictive models of this nature, would get more accurate with more user web session observations. Therefore, this study had to capture as many web session observations as possible.

The project was advertised offline as well via campus posters (*Appendix BetterX Campus Campaign Poster*), leaflet handouts (*Appendix BetterX Campus Leaflet Handout*) and a series of in-class presentations to students. Users were incentivized to participate with a raffle draw which gave each participant the opportunity to win Android tablets at the end of each collection phase. Moreover, participants could invite their friends to install the app (via the Ticket option on the app as shown in *Figure 20* and for each share receive an additional raffle ticket. For the duration of the BetterX project, a total of 5 Android tablets have been awarded to users.

To drive online traffic to the BetterX campaign, the project was advertised on the University's Facebook page and was also promoted by paid ads in both the Google Play Store (*Appendix Google Play Store Page for BetterX Android*) and the Facebook Ads Platform (*Appendix BetterX Facebook Page*).

The majority of recruited online users were drawn from Mechanical Turk. The advantage of Mechanical Turk was that it provided access to a large and demographically diverse pool of online users whose anonymity could be preserved. A monetary reward of $1 USD was awarded for each app installation after 15 days of mobile device usage was collected.

3.3.3.1 User Anonymity

All attributes collected by BetterX Android had no user identifiable information. The web browsing history which was captured by the Web attributes was anonymized by obfuscating the contents of the web page and certain attributes from the web pages' URL. These measures maintained the web page structure and the structure of the URL but removed all content which may have contained user identifiable information. Table 5 outlines all user anonymity measures applied to Web Data.

1	Values were obfuscated using one-way hash functions such as *MD5*
2	The only recognizable part in the URL was the *prefix* (i.e. www), the *domain* (i.e. uclan), the *extension* (i.e. ac.uk) and the *parameter names* (i.e.? user=)
3	All parameter values were obfuscated (i.e.? user=username)
4	All cookie values were obfuscated.
5	All HTTP POST/GET Parameters were obfuscated
6	All HTML content from the Response Body is obfuscated with the exception of HTML tags so the structure of the html page is preserved.

Table 5 User Anonymity Measures

For example, the url *http://www.uclan.ac.uk/students?name=Elias&location=Cyprus* was converted into *http://www.uclan.ac.uk/students?name=&location=* and the HTML code of *<div id=sectionA>text<div>* was converted into *<div id=sectionA></div>*.

3.3.3.2 Security Measures

All user data files were compressed and encrypted on the mobile device before being sent via the network to BetterX Cloud for processing. Encryption was accomplished by using a public key certificate (RSA algorithm, 4,096 bits length). The key was included in the Android app and was used whenever there was data transmission. The private key certificate which could decrypt the data file from the mobile device was stored in the 2 Data Processing Clusters on AWS. For decryption, there were 2 requirements: 1) the private key and 2) a secure passphrase (16-character password). For each data loading batch, the password had been typed in the node's terminal so that the files could be decrypted.

Communication from the mobile device to the cloud was done using a secure (SSL encrypted) connection and access to the server was only allowed by the project's team. Access to BetterX Cloud was done via an *ssh* (secure shell) using a username, a password and private key which was generated upon user registration with Amazon Web Services. Access was allowed only from predefined IP addresses.

3.3.3.3 Data Attributes Selected for MWQoE

During data exploration and analysis, 28 attributes from all the 6 data domains collected (Table 1) were found useful and were included in the MWQoE model, as shown in *Table 6*.

Device	***Web***	***Sensor***
Timestamp	Tab: Id / Status	GPS
Device Id	Page: Id / Url*	
Battery Level	Page Start Time*	***User***
Apps Launches	Page Load Timings	Timezone
Location	Content Load	Location
Screen Status: On/Off	Page Load	
Phone Events: On Call, etc…	Blocked	
	DNS	
Network	Connect	
Link Speed	Send	
Has Internet	Wait	
Signal Strength	Receive	
	HTTP Request Details	
Connection	HTTP Response Details	
Strength		

Table 6 Data Attributes Selected for MWQoE

From the Device domain, the Browser information, the Device Manufacturer, Device Model Number as well as the Screen Size were not used. The Browser Information was collected as a verification attribute to ensure that data used in data analysis came from Firefox versions which have been tested during implementation of the BetterX system. The Device Manufacturer and Device Model Number were collected in order to provide a reference to the user’s device in case further information about the device specs and capabilities were needed. The Screen Size, although it has been identified in literature by Schatz and Egger (2012) as a factor which affects QoE, the actual relationship between screen dimensions and QoE has not been identified and could not be established during data exploration. The inclusion of Screen Size in the MWQoE model is regarded in this study as future work.

From the Network domain, the attributes which were used were the Link Speed, the Signal Strength and the Internet attribute. The rest of the attributes provided background information about the capabilities and the status of the network. These attributes were interesting to observe but were not useful for the intended model since meaningful relationships between them and QoE could not be established. The list of all the collected attributes can be found in *Table 1 BetterX Android - Data Collection Attribute List*.

From the Connection domain, the Connection Strength was used and the Connection status, Network type and Roaming were filtered out since no meaningful relationship could be established for those attributes during data exploration or found in literature.

From the Web domain, all the attributes were used except of the Cache details which were collected to examine the possibility of building a predictive web-prefetch model. The possibility for future work for a web-prefetch model to enhance the Web QoE of mobile users will be considered in the future.

From the Sensor domain, only the GPS sensor readings were used in data analysis and the remaining of the sensors attributes were filtered out. The Accelerometer data together with the Gyroscope data were intended to be used for estimating the user activity using a feature extraction algorithm (Kwapisz, Weiss & Moore, 2011) as shown in Section 3.2.2. However, the feature extraction algorithm was found to provide accurate user activity only with fixed-frequency sensor readings and could not have been applied to the adaptive sensor reading implementation which was used in BetterX Android (Section 3.3.1). Therefore, the user activity characterization could not be established in the Model Generation phase and was not considered in the MWQoE model. In addition, during the Model Generation phase, meaningful relationships between Light, Temperature and Step Counter could not be established nor assumed. Therefore, those attributes were excluded from the MWQoE model as well.

From the User domain, only the user's Timezone and the Location of the user were used in the MWQoE model. The rest of the attributes, for example the user's gender and user's education were not intended to be included in the model rather than used to establish a *user reported sample description* of the collected data.

3.4 Results

3.4.1 Overview

The BetterX sample (collected dataset) was found to be a rich, well-distributed dataset in both demographics and device/web frequencies, as illustrated below in this section, suited to formulate the

MWQoE model having a total of 1,371,500 user readings (all attributes excluding sensors), spanning a total of 2727 distinct mobile web sessions of which a total of 58,515 HTTP Requests have been observed (Table 7). It is a dataset derived from real user activity, capturing device states and user context in a series of scenarios which could not be simulated in a lab test.

The BetterX Data Collection phase lasted for 1 month (from February to March of 2016) in which a total of 2727 web sessions have been observed for 165 distinct domains. A total of 1442 files from 55 devices have been sent from the BetterX Android app to the BetterX Cloud for processing.

Recruited Users	55
Files Processed	1442
Mobile Web Sessions Observed	**2727**
Mobile POST/GET Request Measured	58515
Total User Readings	1371500
Distinct Domains Observed	165

Table 7 BetterX Sample Overview

In regards to demographics, 35% of the users were female and 65% of the users were male, averaging 30 years of age. Figure 24 shows a count of total observations per geographic region. Forty percent of the users came from Cyprus, 31% of the users came from the United States, 20% of the users came from India and the 9% (5 users) from other countries.

Figure 24 BetterX Total Observations per Geographic region

Figure 25 shows the types of web session observed in the BetterX dataset where 61% were found to be work related web sessions (*Working Online*), 10% were found to be *Searching for Products,* 9% *General Web Browsing*, 8% *General Searching* (search engine queries), 7% *News Reading*, 3% *Products Purchasing*, 1% *Logging-in* and 1% (15 web sessions) were *Unknown* since the domain type could not be established.

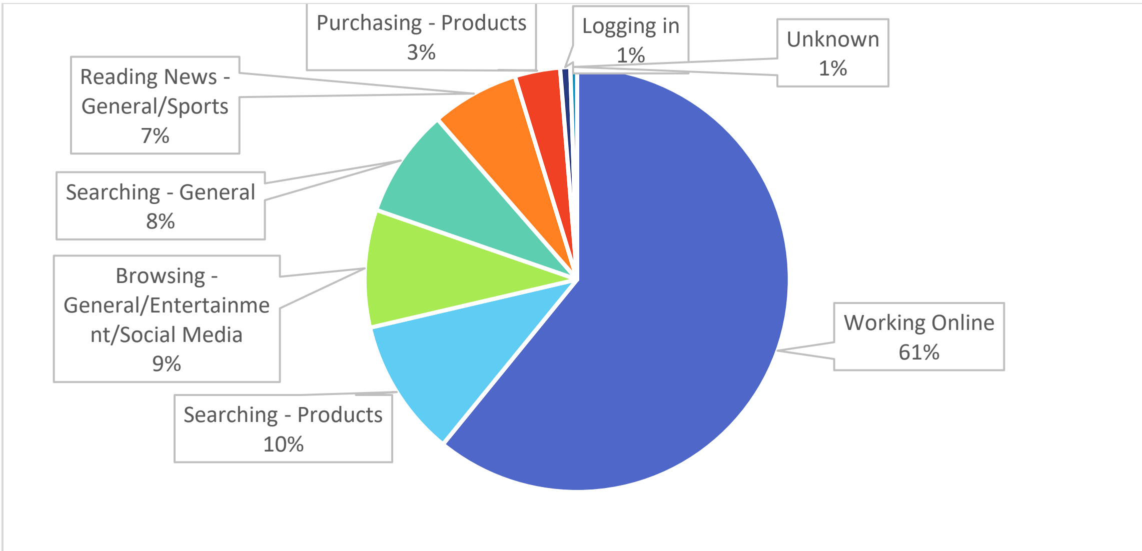


Figure 25 Web Session Types Observed

3.4.2 Data Exploration & Web User Daily Stories

The overall data analysis process used to explore the BetterX dataset was as follows: Table reports were generated via custom SQL queries and then results were saved in MS Excel file format for easy manipulation, graphing and data drill-downs. In addition, *Tableau Desktop*[5] was used to create graphs and maps of selected reports and the *R* statistics language (Zoonekynd, 2007) together with *R-Studio*[6] were used for statistical analysis and data clustering. Finally, the *GeNIe*[7] application was used for Bayesian modeling and building the MWQoE model.

The initial exploration of the dataset was done via a series of custom SQL queries and visualizations. A few of the exploration queries are outlined below.

1. Daily observation counts per device id for all data domains collected
2. Web domain usage report per user demographics
3. Web domain usage report per device id
4. Web domain per Geographic location usage report
5. Web domain per Geographic location broken down per Device Id
6. Web Session duration estimation report per Device Id for each day of usage

After the initial exploration of the data was completed and the overall impression of the dataset was established, then the *Web User Daily Stories* were created for each user. For *Web User Daily Stories* a

[5] *http://www.tableau.com/*
[6] *https://www.rstudio.com/*
[7] *http://www.bayesfusion.com/*

custom PHP application was developed, the BetterX Data Viewer, which allowed selected attributes (*Table 6*) to be visualized for one or more web sessions of each user on specific times and specific locations.

Figure 26 BetterX Data Viewer Menu

As shown in Figure 26, the BetterX Data Viewer enabled user selection via UID, selection by day, time filtering (minimum and maximum time), 2 graph modes (either plot 1 graph per file or plot all files together in 1 one graph for comparison), selection of the specific attributes to view (each attribute was stored in one file) and the option to display the location data on a map. Moreover, it provided a line-by-line graph for selected attributes for a distinct user, as shown in Figure 27. Figure 27 shows the changes of the Content Load, Page Load, and Total HTTP Time as they relate to changes in the Response Body Size attribute for the web session observed by the user (randomly selected) within a specific time-range in a single day.

Figure 27 BetterX Data Viewer Attribute Graphs

The *Web User Daily Stories* provided the detailed analysis of all the context attributes and web metrics available for each web session of each user. The data for the Web User Daily Stories implementation was compiled via a custom python script (Table 4) which used the millisecond timestamp of the attributes to align them around each observed web session. The attributes *before*, *while* and *after* each web session were grouped and reported to identify the conditions *before*, *while* and *after* a web session has taken place, as shown in Figure 28. *This* figure illustrates at a high-level the method for defining the *Web User Daily*

Stories. It shows a sample timeline of device activity about each web session observed, and contextualizes the conditions in which each web session has taken place.

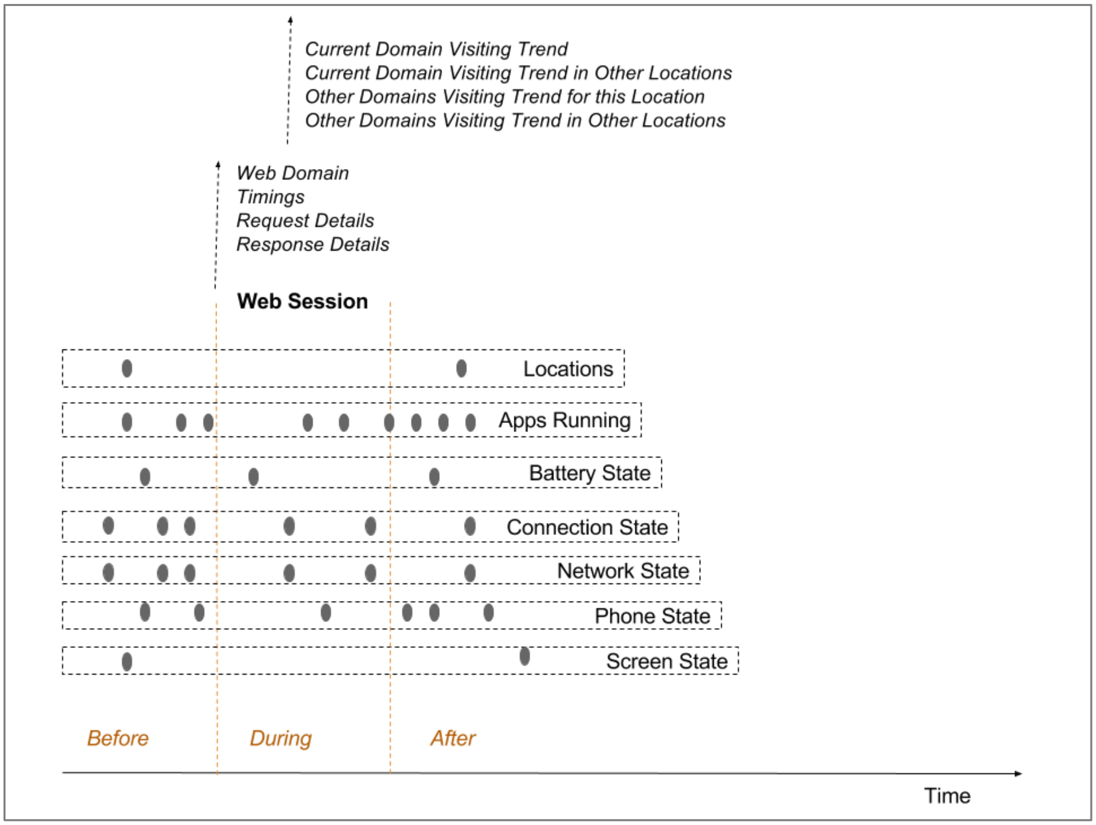


Figure 28 Web User Daily Stories Analysis Sample

For each of the 2727 distinct web sessions observed in the BetterX dataset, context attributes from all 6 domains captured were clustered and chronologically aligned in the manner shown in Figure 28. Each BetterX user was examined separately and insights on not only the web metrics and the context data but potential links and correlations that later become part of the MWQoE model were noted. To better illustrate the use of the *Web User Daily Stories*, a snapshot of it for a user is provided below:

> *'The device id of 35571…600e is from a 31-year-old male from Bharuch, India. The user used Firefox browser v44 and had indicated that is a "very often" mobile phone and mobile web user. The user used the BetterX Android app for 2 days. The readings from the first day, Day 34 (34th day of the year) started at 10am local time. Almost half of the web sessions were from Amazons' Mechanical Turk website (mturk.com). The*

rest of the web use was mainly on online survey websites such as qualtrics.com, soscisurvey.de, appsgeyser.com as well as popular sites such as Google, Facebook and Amazon. The estimated average duration per web session for mturk.com was 261 seconds. The user was located in a residential neighborhood in Gujarat, India. The user stayed at the location throughout the day. In fact, this is the only location registered for this user which was logged during setup. It seems that the user disabled Android's location reporting after the app setup has been completed. The user was connected to the internet on a WiFi connection with fluctuating speeds and the battery level was declining. It is possible that the user disabled location reporting in order to preserve power. Connection Strength with Battery Level need to be analyzed for potential correlations. The user had 71 sessions within one hour from 11am to 12pm on mturk.com and most of them where searches and HIT views with a high number of 'HTTP Blocked' timings. Viewing the Blocked timings with the Page Load timing reveals a potential correlation between the 2 attributes and must be a subject of further analysis. For the second day, Day 35, the web sessions started at 12PM local time and continued throughout the rest of the day. Mostly the traffic was similar to Day 34. It can be inferred from this review of the user that the User Profile is of an Online Crowdsourcing Worker, working from home. The intent of the user for all the sites that are categorized either as crowdsourcing sites or online survey sites can be classified as "working" with intent importance as "High"'.

Analyzing the *Web User Daily Stories* in a day-by-day basis revealed insights about the context of the user, the device context and the network context for each observed web session. This analysis together with the calculation of the Web Session User Intent (as shown in Figure 9), for each of the 2727 web sessions observed enabled the interpretation of the user's motivation and aim for each web session. Sample records with Web User Session Intent are shown in Table 8.

Timestamp	1454831049	1454795277	1454596412	1454856853	1454728055	1454728089	1454535938
Local Time	2/7/2016 7:44	2/6/2016 21:47	2/4/2016 14:33	2/7/2016 14:54	2/6/2016 3:07	2/6/2016 3:08	2/3/2016 21:45
Content Load	1236	3455	12007	5314	5319	3019	1609
Page Load	1263	6327	19899	9471	8668	8210	1732
Blocked	4	3	215	2026	75	6	79
Connect	0	0	0	0	0	0	0
Send	0	0	0	0	0	0	0
Wait	303	486	529	2590	54	160	187
Receive	21	604	848	317	2	0	0
Response Header Size	745	524	343	1046	545	1403	2164
Response Body Size	12024	17382	73863	11487	244	0	20
Domain	mturk.com	imdb.com	philenews.c om	facebook.co m	google.com	facebook.com	netflix.com
Domain type	Croudsourcing	Entertainment	News	Soc. Media	Search	Soc. Media	Entertainment
URL Part	findhits?match=	filmotype/actress?ref_=			url?q=&sa=&ved=&sig2=&usg=		Login?locale= &nextpage=
Status	200	200	200	200	302	302	302
Status Type	OK	OK	OK	OK	Found	Found	Found
Location Type	Commercial	Residential	Residential	Residential	Residential	Residential	Residential
Battery Value	20	89	68	89	32	31	94
Connection	MOBILE	WIFI	WIFI	WIFI	WIFI	WIFI	WIFI
Strength	4	4	4	2	4	4	4
Intent	**Working**	**Browsing – Online Entertainment**	**Reading News**	**Browsing Social Media**	**Searching**	**Browsing Social Media**	**Browsing Online Entertainment**

Table 8 Random Records of Web Session Analysis Table

3.4.3 BetterX User Profiles

Once the data analysis of the Web User Daily Stores was completed and the Web Session User Intent was appended back into the Web Session Analysis Table (Table 8), then the Web Session User Profiles were created and appended back into the same table. The Web Session User Profiles were derived from the Web Session User Intent and the frequency of the Web Session User Intent within the observed time-frame of each distinct web session user. For example, if a user was observed *working* (the web session user intent) on *crowdsourcing* sites (domain type), then user was classified as *worker* for the web sessions that matched that same intent-domain-url part criteria. In a similar manner and after manually reviewing and analyzing all 2727 web sessions, a total of 4 distinct User Profiles was derived: 1) Worker, 2) Buyer, 3) Reader and 4) Web User (Default). The Worker was defined as an individual who performs online work from either a residential or commercial location and uses mainly crowdsourcing sites, data entry sites or online survey sites. The Buyer was defined as an individual whose primary intent on web usage is to search, review and purchase products from ecommerce sites. The Reader was defined as a web user whose primary web activity is to read the news from the web browser via different news providers. And, finally the Web User – Default user type - was defined as an individual whose main web activity was to use search engines to search and get answers for specific topics or questions.

The basis of the MWQoE model was formed once the Web Session User Intent and the Web Session User Profile were appended back into the Web Session Analysis Table.

3.4.4 BetterX MWQoE Model

3.4.4.1 Overview

The MWQoE model is constructed using a Bayesian Network model as the underlying theoretical framework. All attribute relationships that are identified in the data exploration are coded in the Bayesian MWQoE model as e*xpert knowledge*. The MWQoE model is broken down into 4 context sub-models derived from the BetterX Dataset: 1) Timings Context State, 2) Network Context State, 3) Device Context State and 4) Web Intent Importance Context State. The section below outlines each model, its data analysis and generation.

3.4.4.2 Timings Context

The Timings Context State (TCS) as illustrated in the causal links diagram (Figure 29), models the web metrics of each web session. It forms a discrete representation of the amount of time it takes to make available a web page to the user from initial request to completion. TCS is generalized into 5 states (Excellent, Very Good, Good, Fair, Poor) closely resembling the Likert scale (Tullis and Albert, 2013) which is used for MOS evaluation (International Telecommunication Union, 2014). The TCS model feeds from all HTTP timing metrics plus the HTTP Request and Responses Sizes in order to form the causal relationships of each of the selected attributes. The TCS model evaluates the time it takes for a webpage to be made available to the user based on the size of the request, the size of the response and the total time it took to fetch and display the contents of the web page.

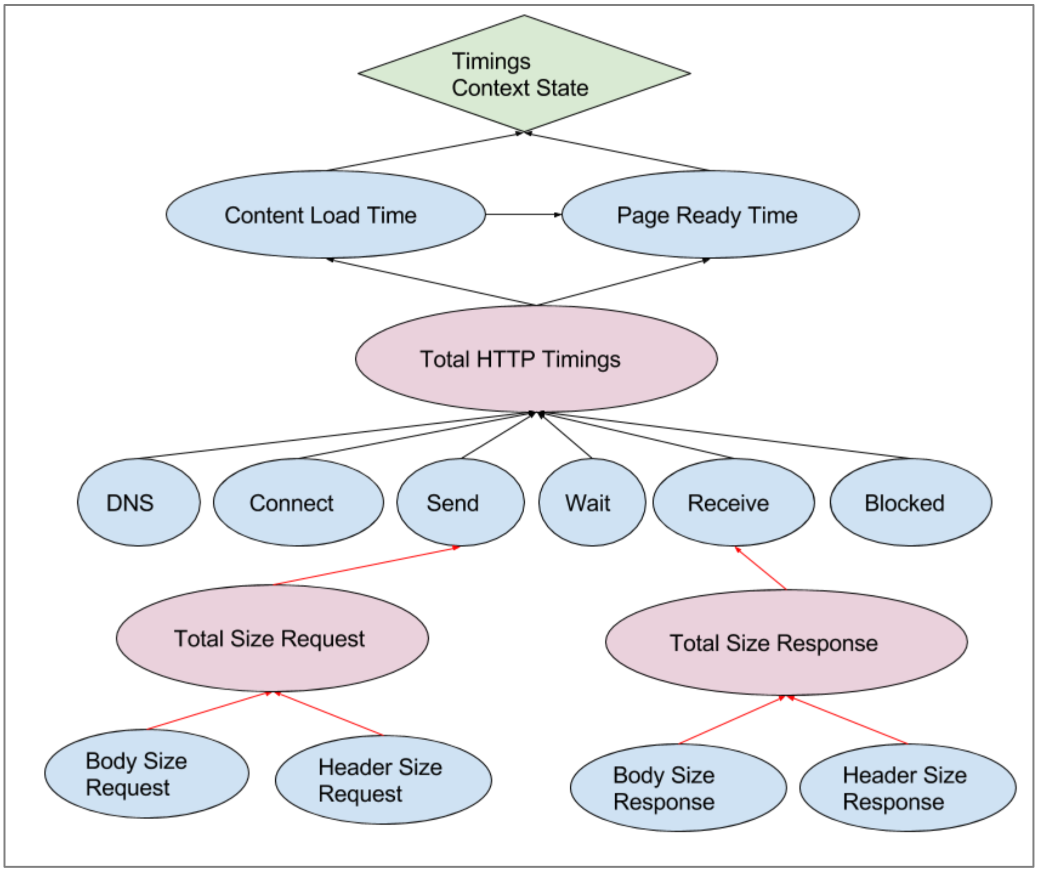


Figure 29 TCS Causal Links

The Total Size of the Request (TRTS) is the sum in bytes of the Body Size Request and the Header Size Request. The Total Size of the Response (TRES) is the sum in bytes of the Body Size Response and the Header Size Response. TRES is linked with the Send attribute since the time that is needed to send a request is proportionally related with the size of it. The same applies with the TRTS and the Receive attribute as the time it takes to receive a response is proportionally related to the size of the response. The Total HTTP Timings (THTTP) is the sum in milliseconds of all the Timing attributes (DNS, Connect, Send, Wait, Receive, Blocked). DNS measures the time it takes to resolve the host name. Connect measures the time required to create a TCP connection. Send is the time to send the HTTP request to the server. Wait is the time waiting for a response from the server. Receive is the time required to read the entire response from the server or cache. Blocked is the time spent waiting in queue for a network.

The Content Load Time (CL) represents the time it takes to load all the content of the page whereas the Page Load Time (PL) represents the total time it takes to make the web page available to the user. Both of these attributes are affected by THTTP. Also CL affects PL since the more time it takes to load the content of the page the more time it takes to have the page ready to the user and vice versa, thus making both CL and PL affect TCS.

After the Casual Links for TCS are defined using the relationships uncovered in data exploration, then the Bayesian Credible Intervals are calculated for each attribute in the model. Credible Intervals (CI) define the Highest Density Region (HDR) of the sample size and constitute the shortest interval which contains 95% of the probability. A Credible interval is to Bayesian Statistics what a Confidence Interval is to Descriptive Statistics. Basically, it is a different approach but with the same goal. For TCS and the rest of the sub-models (NCS, DCS, ICS), 95% CIs are used. *Table 9* lists the attributes used in TCS, their abbreviation codes, their attribute type (*what they measure*) and the lower and upper bounds of their CI's.

Attribute Name	**Attribute Code**	**Attribute Type**	**95% CI Lower**	**95% CI Upper**
Content Load	CL	Timing (ms)	0	9475
Page Load	PL	Timing (ms)	0	20574
Blocked	BLK	Timing (ms)	0	1980
DNS	DNS	Timing (ms)	0	0
Connect	CON	Timing (ms)	0	0
Send	SND	Timing (ms)	0	1
Wait	WAT	Timing (ms)	0	2371
Receive	RCV	Timing (ms)	0	1078
Body Size Request	BSRT	Size (bytes)	0	2078
Body Size Response	BSRE	Size (bytes)	0	45344
Header Size Request	HSRT	Size (bytes)	272	2337
Header Size Response	HSRE	Size (bytes)	0	1565

Table 9 TCS Credible Intervals

The records which did not belong to the Credible Interval range are discarded and the remaining records are used to Cluster and Discretize the data, preparing it for the TCS Bayesian model. The *k-means clustering* algorithm is used for discretizing each attribute into a default of 10 context states (*Bins*) as shown in *Table 10*. *Bins* provide the *definitions* of each context-attribute in the model as they illustrate how metrics are converted into context-states (*discrete characterizations*). For example, the context-state of *CL01* which is the 1st *Bin* of the Content Load attribute, is defined as Content Load of 0-535 milliseconds. Some attributes do not have a large enough range of values in order to form 10 bin sizes, such as the Send attribute which ranged for 0 to 1. Both the DNS and the Connect attributes are all zero so 1 bin is used for each. Once the attributes are discretized in their respective Bins then the TCS is defined as each state of Content Load and Page Load is mapped to a distinct TCS state (Table 11*).* The TCS States are linearly transformed, as shown in Table 12 below and a Utility is assigned for each state. The Utility of TCS illustrates the timing of which the web page is made available to the user with the value 1 being *Excellent* and the value of 0 being *Poor*. The final TCS Bayesian model as it was coded in the GeNIe platform is shown in Figure 30.

Attribute Name	Bin Size	Bins
Content Load	10	[0, 535) [535,1268) [1268,1833) [1833,2430) [2430,3062) [3062,3812) [3812,4813) [4813,6084) [6084,7536) [7536,9469]
Page Load	10	[0, 648) [648, 1691) [1691, 2484) [2484, 3223) [3223, 4071) [4071, 5255) [5255, 7017) [7017, 9947) [9947,14346) [14346,20574]
Blocked	10	[0.0, 58.4) [58.4, 161.0) [161.0, 296.4) [296.4, 447.1) [447.1, 618.3) [618.3, 841.6) [841.6,1142.3) [1142.3,1451.7) [1451.7,1713.9), [1713.9,1948.0]
DNS	1	0.0
Connect	1	0.0
Send	2	0.0 [0.5,1.0]
Wait	10	[0, 118) [118, 296) [296, 483) [483, 644) [644, 819) [819,1015) [1015,1221) [1221,1466) [1466,1821) [1821,2371]
Receive	10	[0.0, 13.6) [13.6, 41.7) [41.7, 78.4) [78.4, 146.8) [146.8, 242.3) [242.3, 333.5) [333.5, 433.5) [433.5, 559.5) [559.5, 743.5) [743.5,1009.0]
Body Size Request	10	[0.0, 21.8) [21.8, 57.1) [57.1, 119.6) [119.6, 256.8) [256.8, 447.5) [447.5, 689.9) [689.9, 979.7) [979.7,1225.8) [1225.8,1610.3) [1610.3,2074.0]
Body Size Response	10	[0, 214) [214, 1688) [1688, 3881) [3881, 5901) [5901, 9134) [9134,12814) [12814,16088) [16088,22750) [22750,34307) [34307,45344]
Header Size Request	10	[0.0, 30.4) [30.4, 80.9) [80.9, 140.6) [140.6, 262.6) [262.6, 447.5) [447.5, 708.9) [708.9,1073.7) [1073.7,1572.9) [1572.9,1955.4) 1955.4,2074.0]
Header Size Response	10	[0, 1399) [1399, 3500) [3500, 5323) [5323, 7311) [7311, 9917) [9917,13013) [13013,16157) [16157,22760) [22760,34307) [34307,45344]
Total Request Size (TRTS)	10	[278, 426) [426, 573) [573, 737) [737, 925) [925,1189) [1189,1486) [1486,1724) [1724,2020) [2020,3101) [3101,4253]
Total Response Size (TRES)	10	[0, 2559) [2559, 6035) [6035, 9486) [9486,12576) [12576,14735) [14735,17040) [17040,21342) [21342,27797) [27797,36658) [36658,46357]
Total HTTP Timings	10	[0, 177) [177, 412) [412, 626) [626, 859) [859,1130) [1130,1411) [1411,1740) [1740,2111) [2111,2660) [2660,4099]

Table 10 TCS Clusters

State Name	State Code	Content Load	Page Load
EXCELLENT	TCS01	CL01 – CL02	PL01 – PL02
VERY GOOD	TCS02	CL03 – CL04	PL03 – PL04
GOOD	TCS03	CL05 – CL06	PL05
FAIR	TCS04	CL07 - CL08	PL06 – PL07
POOR	TCS05	CL09 – CL10	PL08 – PL10

Table 11 TCS Mappings

TCS State	Ordinal Scale	Utility Value Calculation	Utility Value
EXCELLENT	5	=5-1/5-1 = 1	**1**
VERY GOOD	4	=4-1 / 5-1 = 0.75	**0.75**
GOOD	3	=3-1/5-1 = 0.5	**0.5**
FAIR	2	= 2-1 / 5-1 = 0.25	**0.25**
POOR	1	= 1-1 / 5-1 = 0	**0**

Table 12 TCS Utility Calculation

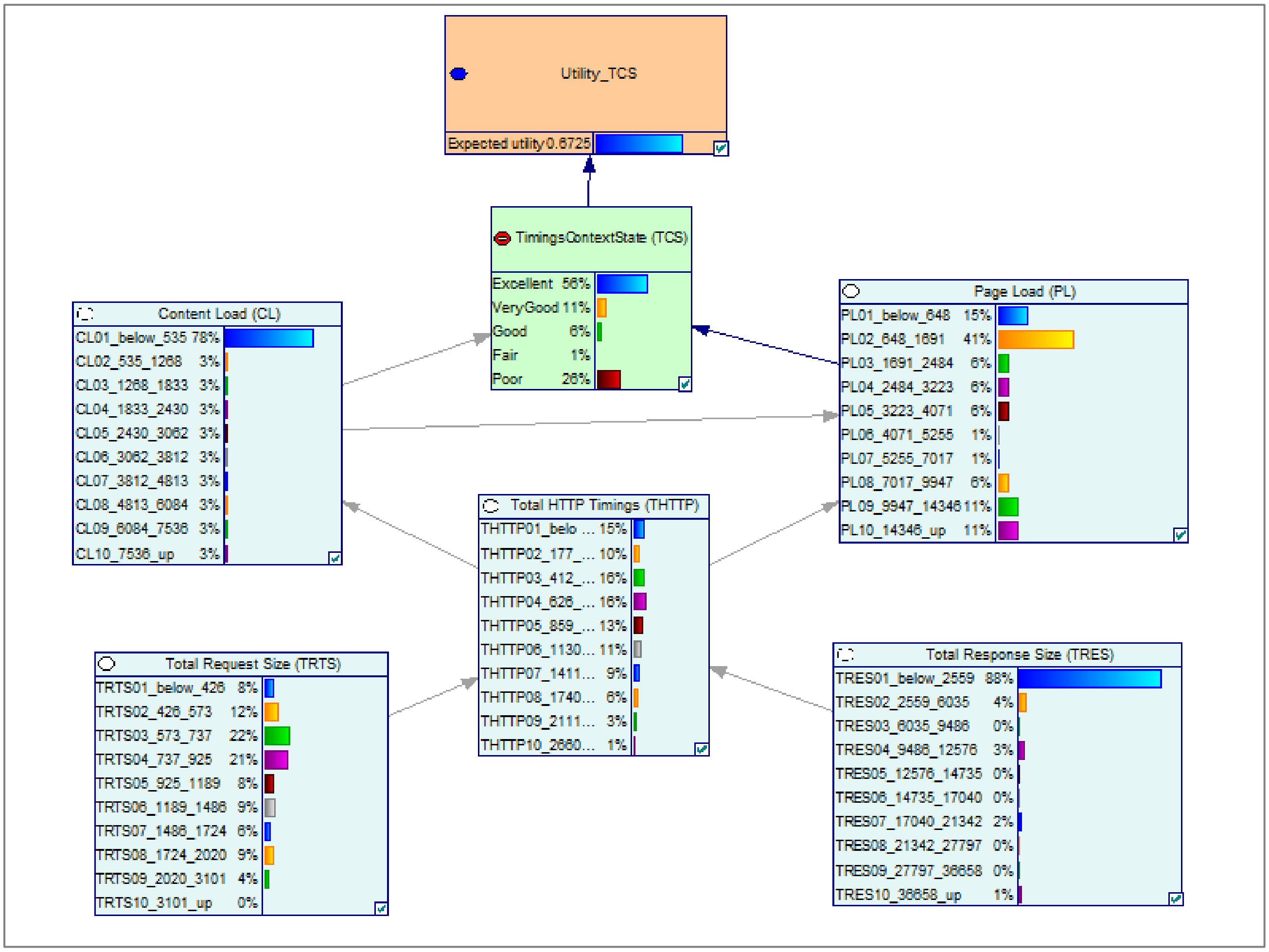


Figure 30 TCS Model

As shown in Table 11, the PL state is being given a higher strength of influence to TCS than the CL state, since content may be loaded completely but the page may not be immediately available to the user due to a high device processing load which can cause a lag in rendering. For example: *TCS(CL04, PL04) = "Very Good"* shows that the web page availability to the user is Very Good and *TCS(CL05,PL05) = "Good"*, however *TCS(CL04,PL05) = "Good"*. Moreover, it is noted that given the TCS model (*Figure 30*), 56% of all web sessions observed were *Excellent*, 11% were *Very Good*, 6% were *Good*, 1% was *Fair* and 26 were *Poor*.

The remaining 3 sub-models (Network, Device, Web Intent Importance) are generated using the same methodology as the TCS model. Credible intervals of 95% are calculated for all attributes which are selected from data exploration and discretized in bins of default size 10 using the k-means clustering algorithm.

3.4.4.3 Network Context

The Network Context State (NCS) as it is illustrated in the causal links diagram in Figure 31 models the Network state for each web session observed. It forms a discrete representation of the quality of the Network as it was received by the mobile device. Using the same methodology as the TCS model, The NCS model is generalized into 5 states (*Excellent*, *Very Good*, *Good*, *Fair*, *Poor*).

The three attributes that compose NCS are: 1) Internet, 2) Network Link Speed and 3) Signal Strength. The Internet (INT) attribute is of a binary format and shows if access to the Internet is available from the current connection. The Internet attribute is of a major strength of influence to NCS (Figure 31) since a web page cannot be accessed without internet. Therefore, when there is no internet connection, NCS is *Poor*, *NCS(INT=0) = "Poor"*, and when there is internet connection NCS can be from *Fair* to *Excellent*, *NCS(INT>0) = "[Fair-Excellent]"*. The Network Link Speed (LSPD) attribute measures the speed of network link in Mbps (Megabits per second) and the Signal Strength (SIGS) is a combination (union) of the Connection Strength (CSTR) and the Network Signal Strength (SIGL). CSTR measures the WiFi connection strength and SIGL measures the Mobile Network connection strength. For simplicity, SIGS is used to represent the strength of both these connections.

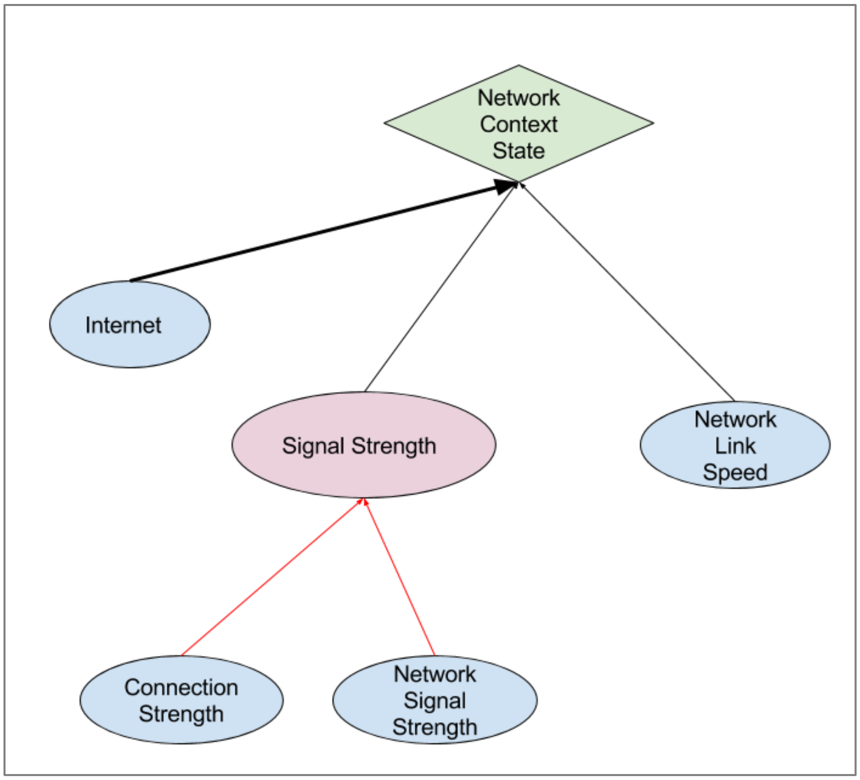


Figure 31 NCS Causal Links

Table 13 shows each attribute used in NCS, the Credible Intervals and the Discretization of each attribute into distinct states. Mapping attribute states together is illustrated in Table 15 - Table 17. NCS is coded into 5 states, NCS01 = *Excellent* and NCS05 = *Poor* as shown in Table 13. SIGS and LSPD are then mapped individually into NCS states. The higher the signal strength and speed of the connection, the higher the NCS state. For SIGS, each one state is mapped into one NCS state in a linear one-to-one manner *Table 15* since both states have 5 attributes each. Therefore, *NCS(SIG0) = NCS05* is *Excellent* and *NCS(SIG5) = NCS01* is *Poor*.

On the other hand, Link Speed (LSPD) is mapped into NCS considering two observations about the collected dataset. The first is that LSPD is a 10 state attribute whereas NCS is a 5 state attribute therefore a one to one mapping is not possible. Second, that the average web page data size is around 1-2Mb which given the Link Speed rates observed, enables very fast downloading for web sessions. Therefore, the LSPD to NCS mapping needs to reflect both observations in which the NCS state should trend higher and not be as evenly distributed as the SIGS to NCS mapping, since the Network speeds in general are found to be very capable of fast downloading even of the most demanding sites.

Attribute Name	Attribute Code	95% CI Lower	95% CI Upper	Bins
Internet	INT	--	--	0,1
Network Link Speed	LSPD	0	72	0.00 [2.54, 8.92) [8.92,18.68) [18.68,30.30) [30.30,37.50) [37.50,43.50) [43.50,50.72) [50.72,55.72) [55.72,63.17) [63.17,72.00]
Connection Strength	CSTR	--	--	0,1,2,3,4,5
Network Signal Strength	SIGL	--	--	0,1,2,3,4
Signal Strength	SIGS	--	--	0,1,2,3,4,5

Table 13 NCS Credible Intervals & Clusters

State Name	State Code
Excellent	NCS01
Very Good	NCS02
Good	NCS03
Fair	NCS04
Poor	NCS05

Table 14 NCS States

Signal Strength Level	NCS State Code
SIG0	NCS05
SIG1	NCS04
SIG2	NCS03
SIG3	NCS03
SIG4	NCS02
SIG5	NCS01

Table 15 SIG - NCS Mappings

Therefore, as shown in Figure 32*,* it is assumed that any Link Speed higher than 43 Mbps delivers an *Excellent* Network State and a Link Speed between 2.5 and 8.9 delivers a *Fair* Network State. The mapping of the LSPD to NCS is outlined in Table 16. For clarity, Figure 32 shows how each distinct Link Speed in Mbps which has been observed in the BetterX dataset maps to each NCS state. For example, Link Speed of 13 Mbps gives *Good* NCS whereas Link Speed of 65 gives *Excellent* NCS.

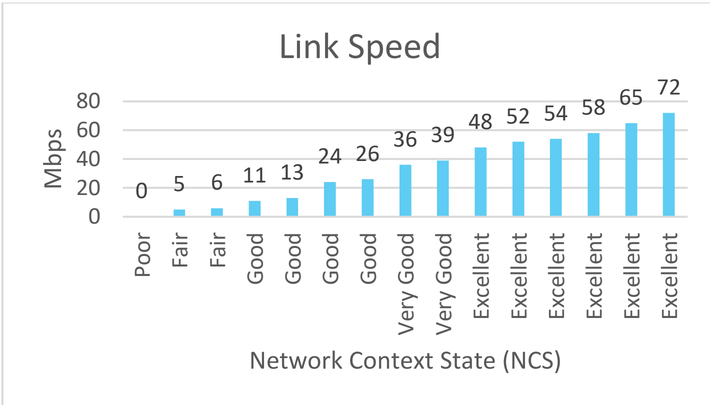


Figure 32 NCS States per Link Speed

Link Speed Level	NCS State Code
LSPD01	NCS05
LSPD02	NCS04
LSPD03	NCS03
LSPD04	NCS03
LSPD05	NCS02
LSP06	NCS02
LSPD07	NCS01
LSPD08	NCS01
LSPD09	NCS01
LSPD10	NCS01

Table 16 LSPD - NCS Mappings

		NCS01	NCS02	NCS03	NCS04	NCS05			NCS01	NCS02	NCS03	NCS04	NCS05
SIG0	**LSPD01**					x	**SIG4**	**LSPD01**			x		
	LSPD02					x		**LSPD02**			x		
	LSPD03					x		**LSPD03**		x			
	LSPD04					x		**LSPD04**		x			
	LSPD05					x		**LSPD05**		x			
	LSPD06					x		**LSPD06**	x				
	LSPD07					x		**LSPD07**	x				
	LSPD08					x		**LSPD08**	x				
	LSPD09					x		**LSPD09**	x				
	LSPD10					x		**LSPD10**	x				
SIG1	**LSPD01**				x		**SIG5**	**LSPD01**			x		
	LSPD02				x			**LSPD02**		x			
	LSPD03				x			**LSPD03**		x			
	LSPD04				x			**LSPD04**	x	x			
	LSPD05			x				**LSPD05**	x				
	LSPD06			x				**LSPD06**	x				
	LSPD07			x				**LSPD07**	x				
	LSPD08		x					**LSPD08**	x				
	LSPD09		x					**LSPD09**	x				
	LSPD10		x					**LSPD10**	x				
SIG2	**LSPD01**				x								
	LSPD02				x								
	LSPD03			x									
	LSPD04			x									
	LSPD05		x										
	LSPD06		x										
	LSPD07		x										
	LSPD08	x											
	LSPD09	x											
	LSPD10	x											
SIG3	**LSPD01**			x									
	LSPD02			x									
	LSPD03			x									
	LSPD04			x									
	LSPD05		x										
	LSPD06		x										
	LSPD07	x											
	LSPD08	x											
	LSPD09	x											
	LSPD10	x											

Table 17 SIG-LSPD-NCS Mappings

The last step for the NCS model is to define the relationship between LSPD, SIGS and NCS. The mapping of these three attributes is shown in Table 17 and a graph of the actual values of both of these attributes is shown against NCS in Figure 33. The mapping reflects the predicate that SIGS and LSPD share the same strength of influence in NCS since slow web page downloading has been observed with both a very low signal strength with a high link speed, and, with high signal strength and low link speed. Therefore, SIGS and LSPD are evenly distributed in NCS. Figure 33 shows that when LSPD = 36 Mbps and SIGS = 3 then NCS = *Very Good*, whereas LSPD = 55 and SIGS = 4 gives NCS = Excellent.

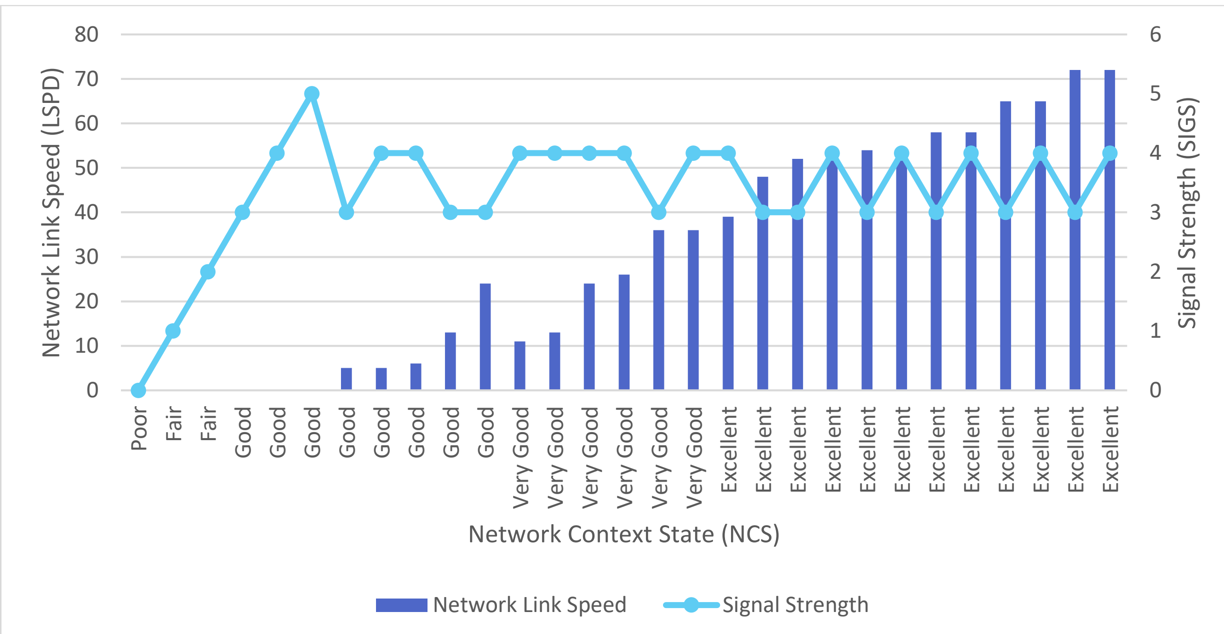


Figure 33 NCS per Signal Strength and Link Speed

Finally, the NCS States are linearly transformed, as in TCS, and a Utility Value is assigned for each state [0 – 1]. The final NCS Bayesian model as it is coded in *GeNIe* is shown in Figure 34. It shows that 56% of all web sessions observed where NCS *Excellent*, 28% were *Very Good*, 3% were *Good*, 3% were *Fair* and 10% were *Poor*.

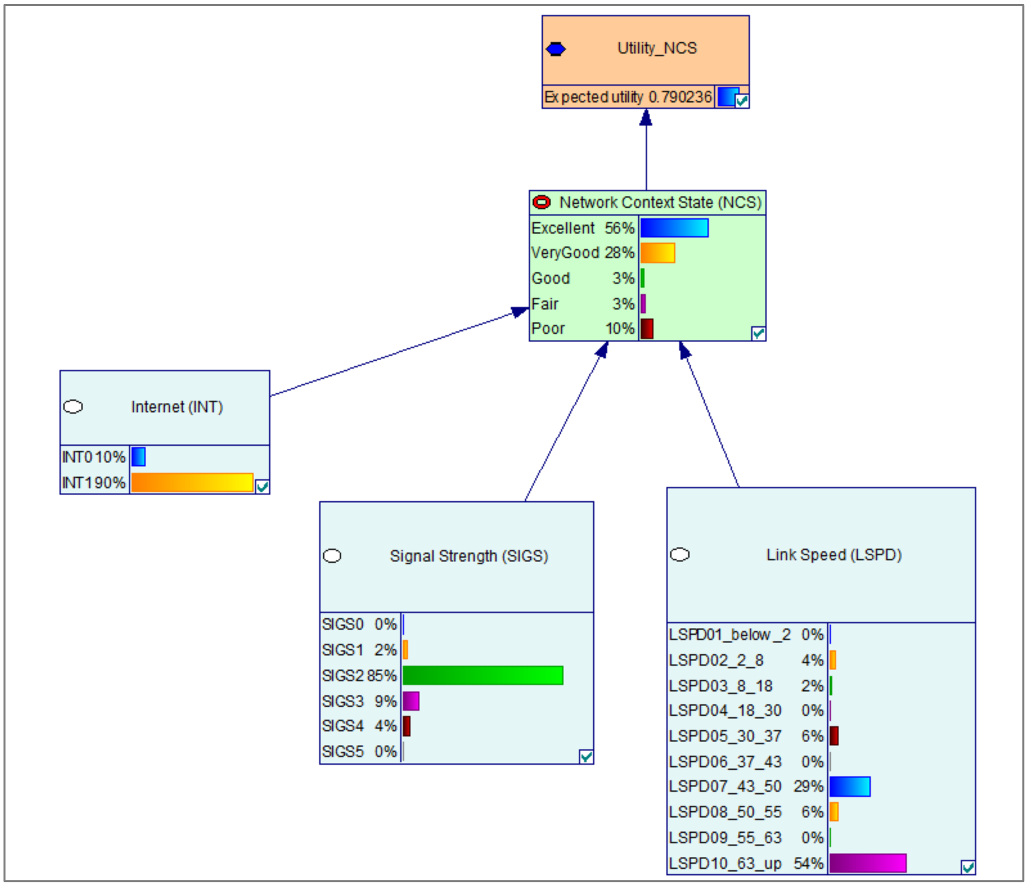


Figure 34 NCS Model

3.4.4.4 Device Context

The Device Context State (DCS) as it is illustrated in the causal links diagram in Figure 35, characterizes the availability of the device in regards to the amount of applications running on the device, the battery state of the device and the amount on telephony traffic on the device. DCS models the device usage frequency in regards to apps, phone and battery. Device usage has an inverse relationship with device availability, i.e. the available resources to use for web session downloading and content rendering depend on the amount of resources used for other tasks on the device such as applications running, the power available of the device and the actual telephony usage. The predicate for DCS is that device resource availability affects web browsing since it has been observed during data exploration that Network requests on slower/busier devices are delayed being sent to the Network and content is rendered in an *indolent* manner.

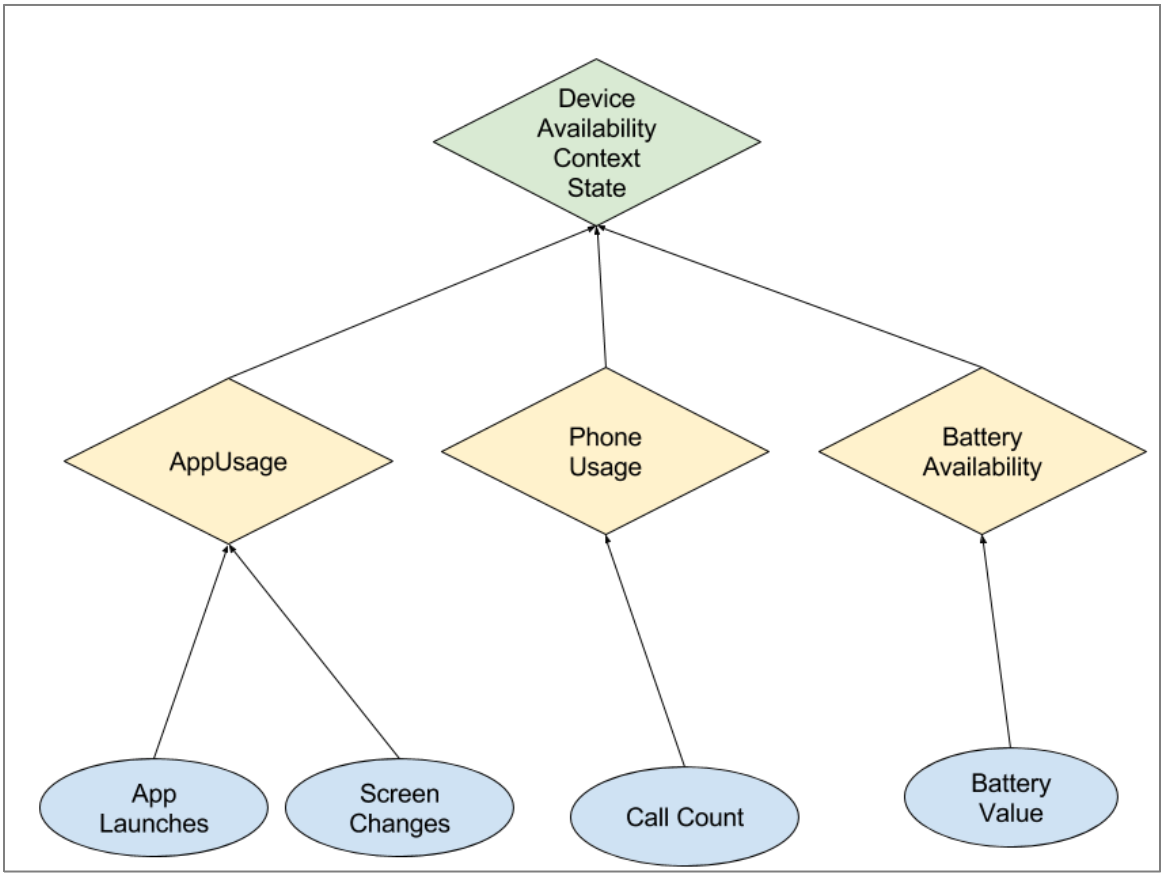


Figure 35 DCS Causal Links

The Table 18 below shows the CI levels and associated clusters for each attribute used in DCS. The Total Calls attribute (CALLS) is the sum of all incoming and outgoing calls within the day. The Screen State Changes (SCREEN) is the sum all times the screen transitioned within the day (screen off/screen on). The App Launches (APPS) is the sum of total app or background service launches within the day. Finally, the Battery Value (BATT) is the available battery reading at each web session.

Apps Usage (APPUS) is defined as an intermediary state which fuses Application usage and Screen changes to give a characterization of the device usage in regards to user interaction with the screen and the apps running on the device. Phone Usage (PHUS) is the intermediary state which characterizes the total call count. Finally, Battery Availability (BATV) is the intermediary state which characterizes the available power on the device.

Attribute Name	Attribute Code	95% CI Lower	95% CI Upper	Bins
Total Calls	CALLS	0	8	0.0 [0.5,1.5) [1.5,2.5) [2.5,4.0) [4.0,5.5) [5.5,6.5) [6.5,7.5) [7.5,8.0]
Screen State Changes	SCREEN	1	72	[1.00, 4.02) [4.02, 5.84) [5.84, 8.78) [8.78,12.87) [12.87,18.66) [18.66,27.67) [27.67,39.86) [39.86,53.40) [53.40,64.11) [64.11,69.00]
App Launches	APPS	0	1091	[0.0, 19.2) [19.2, 55.3) [55.3, 99.8) [99.8, 172.2) [172.2, 342.8) [342.8, 526.5) [526.5, 638.8) [638.8, 799.0) [799.0, 969.2) [969.2,1091.0]
Battery Value	BATT	12	100	[12.0, 13.9) [13.9, 16.5) [16.5, 20.4) [20.4, 27.9) [27.9, 38.6) [38.6, 52.8) [52.8, 68.1) [68.1, 80.2) [80.2, 91.8) [91.8,100.0]

Table 18 DCS Credible Intervals & Clusters

The Table 19-Table 23 show each intermediary state as it is defined in terms of its respective attribute(s).

BATT	Range	BATV
BATT01	Below 13	Low
BATT02	13-16	Low
BATT03	16-20	Low
BATT04	20-27	Low
BATT05	27-38	Medium
BATT06	38-52	Medium
BATT07	52-68	Medium
BATT08	68-80	High
BATT09	80-91	High
BATT010	91+	High

Table 19 BATT - BATV Mappings

CALL	Range	PHUS
CALLS1	0	Low
CALLS2	1	Low
CALLS3	2	Medium
CALLS4	3-4	Medium
CALLS5	5	Medium
CALLS6	6	Medium
CALLS7	7	High
CALLS8	7+	High

Table 20 CALL-PHUS Mappings

APP	Range	APPUS
APPS01	Below 19	Low
APPS02	19-55	Low
APPS03	55-99	Medium
APPS04	99-172	Medium
APPS05	172-342	Medium
APPS06	342-526	Medium
APPS07	526-638	High
APPS08	638-799	High
APPS09	799-969	High
APPS10	969+	High

Table 21 APP-APPUS Mappings

SCREEN	Range	APPUS
SCREEN01	Below 4	Low
SCREEN02	4-5	Low
SCREEN03	5-8	Low
SCREEN04	8-12	Medium
SCREEN05	12-18	Medium
SCREEN06	18-27	Medium
SCREEN07	27-39	Medium
SCREEN08	39-53	High
SCREEN09	53-64	High
SCREEN010	64+	High

Table 22 SCREEN-APPUS Mappings

APPUS	PHUS	BATV	DCS
High	High	High	**Low**
		Medium	**Low**
		Low	**Low**
	Medium	High	**Medium**
		Medium	**Low**
		Low	**Low**
	Low	High	**Medium**
		Medium	**Medium**
		Low	**Low**
Medium	High	High	**Medium**
		Medium	**Medium**
		Low	**Low**
	Medium	High	**High**
		Medium	**Medium**
		Low	**Low**
	Low	High	**High**
		Medium	**Medium**
		Low	**Low**
Low	High	High	**High**
		Medium	**High**
		Low	**Low**
	Medium	High	**High**
		Medium	**High**
		Low	**Low**
	Low	High	**High**
		Medium	**High**
		Low	**Low**

Table 23 APPUS-PHUS-BATV-DCS Mappings

For example, BATT is assumed to be *Low* if battery level is lower than 27%, *High* if its higher than 68% and *Medium* if it's between 27% and 68%. PHUS is *Low* if there are less than 2 calls on the device per day, *Medium* if there are less than 7 calls and *High* if there are more than 7 calls on the device. APPUS is *Low* for less than 55 app launches and less than 8 screen changes, *Medium* for 55-526 app launches and less than 39 screen changes, and *High* for more than 526 app launches and more than 39 screen changes per day.

DCS is composed by combining the three intermediary states (BATV, APPUS, PHUS) using the relationships outlined in Table 23 which shows that a higher strength of influence is given on APPUS and BATV and a lesser one on PHUS since it has been observed that app usage drains the power on the device faster than telephony activity. Battery is used as a factor in the DCS model since the Android operating system has been observed to restrain processing and screen brightness on low battery levels, therefore affecting the availability of device resources. Moreover, PHUS is included in the DCS model since telephony operations on Android devices have been observed to be treated with higher priority than other operations or applications. Therefore, it is logical to assume that the higher the number of calls on the device the potentially less resources available on the device for web downloading and web rendering.

Finally, the DCS States are linearly transformed, and a Utility Value was assigned for each state [0 – 1] as shown in *Table 24*. The final DCS Bayesian, as shown in Figure 36, reveals that 6% of all web sessions observed a *High* DCS, 35% observed a *Medium* DCS and 59% observed a *Low* DCS.

DCS	**Ordinal Scale**	**Utility Value Calculation**	**Utility Value**
High	3	=3-1/3-1 = 1	**1**
Medium	2	=2-1 / 3-1 = 0.5	**0.5**
Low	1	=1-1/3-1 = 0	**0**

Table 24 DCS Utility Calculation

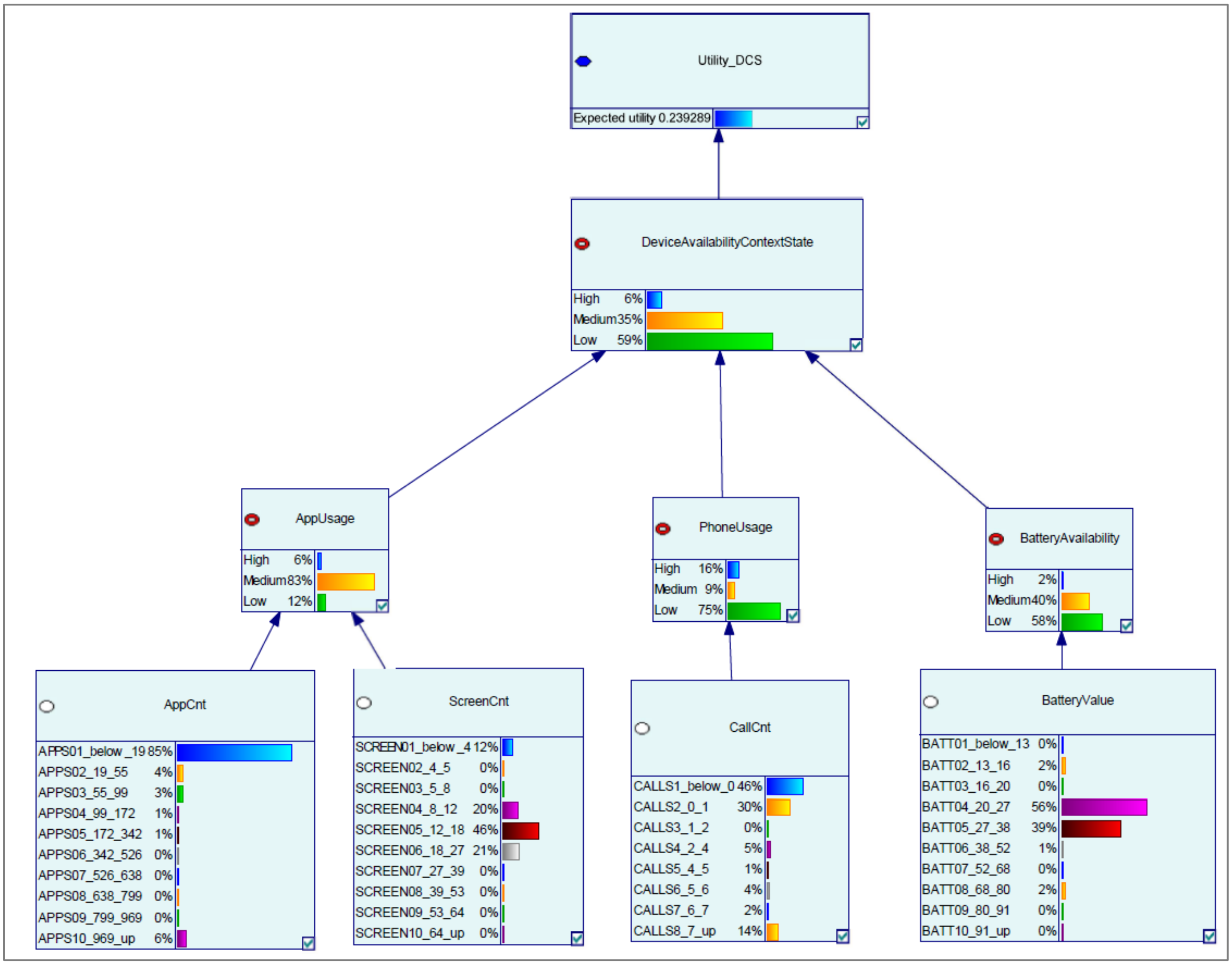


Figure 36 DCS Model

3.4.4.5 Web Intent Importance Context

The Web Intent Importance Context State (ICS) uses 4 attributes: The Location Type (LOC) from which the web session is generated, the Domain Type (DOMT) of the web session's URL, the User Type (UTYPE) or Web Session User Profile, as it has been outlined in *Section 3.4.3 BetterX User Profiles*, and the Web Session User Intent (INTNT) which defines the intention of the user for initiating the web session. The Web Intent Importance Context State provides a characterization of the importance of the web session to the user. ICS is used in the MWQoE model based on the assumption that different types of requests hold different importance levels to users in specific scenarios. For example, a web session request to access work related material can have a higher importance to an online worker than a web request to check the news while sitting in a coffee shop. Therefore, the ICS model regards the user profile, the user context and the type of web site to derive an estimation on the importance of any given web session.

ICS is accomplished via a categorization of each of its associated attributes and the use of Augmented Naïve Bayes (Lewis, 1998) as a data learning mechanism to uncover and quantify the relationships between the attributes. This is done since the relationships for this sub-model between the attributes are difficult to manually define on an attribute by attribute basis. For example, the Web Session User Intent (INTNT) to Location Type (LOC) relationship is not as straight forward to define as the relationships from the previous models (DCS, NCS, TCS) where relationships from literature and current observations are used. Therefore, for the purposes of ICS, a reverse approach is used which is data-driven, allowing for the relationships to be uncovered automatically via the use of the Naïve Bayes algorithm, given the discrete classification and representation of each of the context attributes.

As it has been shown in *Section 3.2.2,* the LOC attribute is created via a reverse geocode of the user's location coordinates to reveal the address and map the location into 7 Location Types: 1) City Center, 2) Commercial Area, 3) Residential Area, 4) Shopping Center, 5) Street/Highway, 6) University, 7) Unknown (location not reported). UTYPE is mapped into the generated 4 Web Session User Profiles: 1) Worker, 2) Buyer, 3) Reader and 4) Web User. DOMT is mapped into 27 different Domain Types: 1) Apps Site,2) Brand Site,3) Classifieds,4) Consumer Electronics,5) Crowdsourcing / Freelancing ite,6) Digital Library,7) Entertainment Site,8) File/Video/Photo Sharing,9) Food Site,10) Internet Gateway,11) News Site,12) Online Ads Distributor,13) Online Portal,14) Online Store,15) Organization Site,16) Personal Site,17) Real-estate Site,18) Rewards Site,19) Search Engine,20) Social Media,21) Surveys,22) Travel Site, 23) Unknown, 24) URL Shortener, 25) Weather Site, 26) Web Hosting, 27) Web Portal. The Web Session User Intent (INTNT) is classified into 12 distinct states as shown in *Table 25* using the method outlined in Figure 9.

User Web Session Intent (INTNT)	BetterX Observations
Browsing - General	6
Browsing - Entertainment	102
Browsing - Products	54
Browsing - Social media	136
Logging in	21
Searching - Products	232
Purchasing - Products	94
Reading News - General	120
Reading News - Sports	62
Searching - General	226
Working Online	1660
Unknown	14

Table 25 BetterX User Web Session Intent Counts

Afterwards, a 11-scenario narrative is created which covers all the 2727 web session observations in the BetterX dataset and defines the assumed importance of each observed web session to the user, and specifically to the Web Session User Profile. For each scenario, an importance level is applied - *Very Important* (ICS05), *Important* (ICS04), *Moderately Important* (ICS03), *Somewhat Important* (ICS02) and *Not Important* (ICS01). Importance levels are applied with priority based on their sequential order as shown in Table 26.

1	All traffic from workers to work sites regardless of location is *important*
2	All traffic from any user type logging in a website from any location is *important*
3	All traffic from an online buyer purchasing products from any location is *important*
4	All traffic from an online reader on news/entertainment site from any location other than residential is *moderately important*
5	All traffic from any user searching for information at any site from any location other than residential is *moderately important*
6	All traffic from online readers on a news/entertainment site from a residential location is *somewhat important*
7	All traffic from any user searching for information at any site from a residential location is *somewhat important*
8	Any type of web browsing from any location by any user is *somewhat important*
9	Any type of web browsing from any location to a social media site is *somewhat important*
10	Any type of user searching from any location other than residential is *somewhat important*
11	All other traffic is *not important*

Table 26 Scenario Importance Narratives

Once all the BetterX web sessions are coded with ICS levels and all the discrete classifications (INTNT, LOC, UTYPE, DOMT) are applied, then the resulting dataset is used to automatically generate the ICS Bayesian model using the Augmented Naïve Bayes Approach. The resulting model is shown in Figure 37 and reveals that based on all web session observations, the User Web Session Intent (INTNT) of the user plays a stronger influence on ICS than the user profile (UTYPE). Moreover, it shows that the Web Session User Intent (INTNT) has a direct relationship with the Web Session User Type (UTYPE) and domain type (DOMT) and to a lesser extent, to location type (LOC). In addition, location type (LOC) has a direct relationship with Web Session User Type (UTYPE) which in turn has a strong relationship with domain type (DOMT).

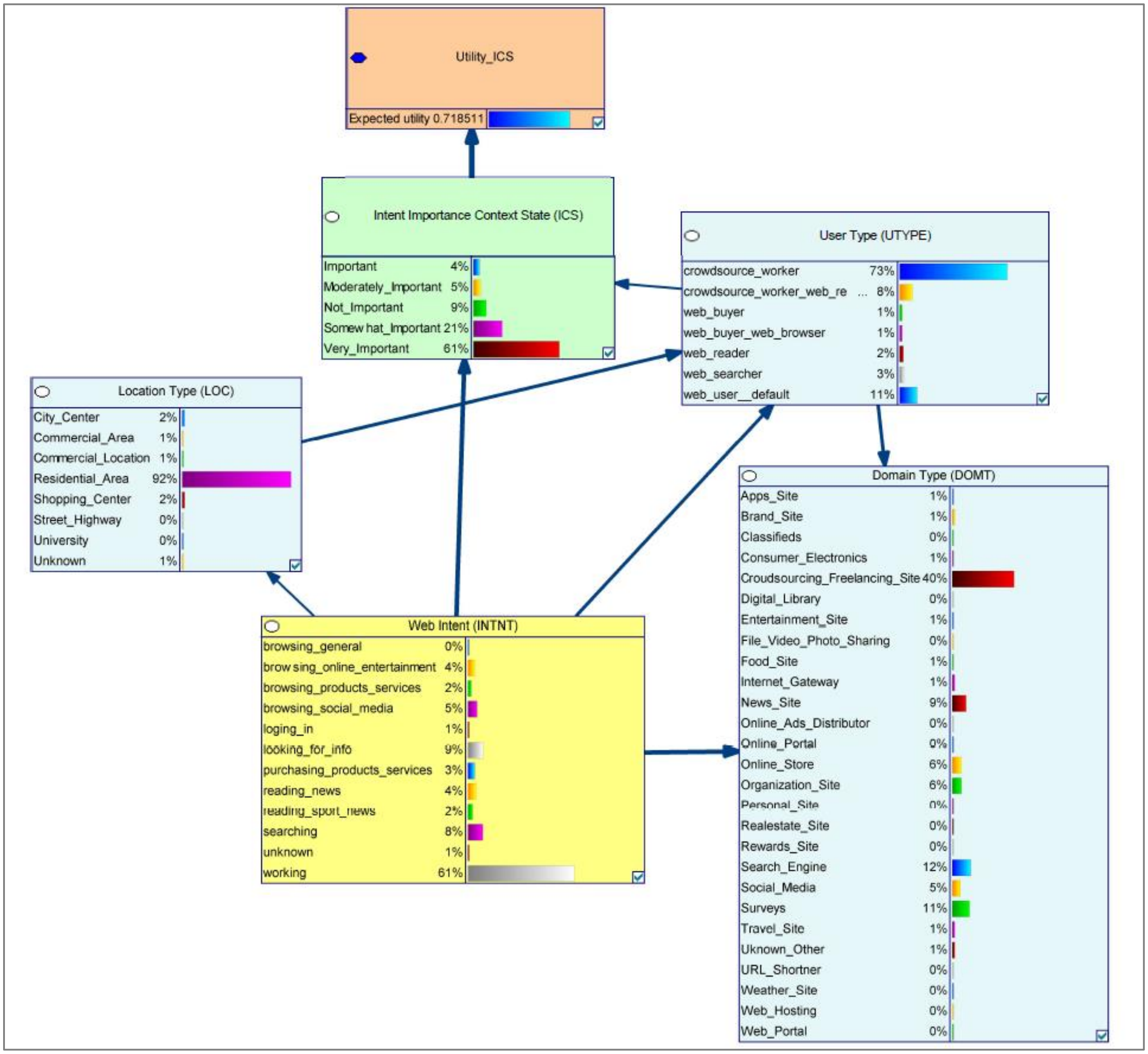


Figure 37 ICS Model

Finally, the ICS States are linearly transformed, and a Utility Value is assigned for each state [0 – 1]. It is noted that based the 11-scenario importance narrative, 61% of all web sessions are identified as *Very Important*, 4% as *Important*, 5% as *Moderately Important*, 21% as *Somewhat Important* and 9% as *Not important*.

3.4.4.6 The Generated Model

The final MWQoE model, as it is illustrated in the causal links diagram (Figure 38), characterizes the experience of the mobile web user in regards to Timings Context (TCS), Device Context (DCS) and Web Intent Importance (ICS). The Network Context State (NCS) is not included in the final MWQoE model since the web metrics (TCS model) have shown that they reflect the state of the network (NCS model) through the fluctuation of values in the HTTP Metrics (THTTP), the Content Load (CL) and Page Load (PL) attributes. The BetterX dataset revealed that the higher the state of NCS, the less the value of THHTP, CL and PL and therefore the faster the web page downloading. Therefore, MWQoE considers TCS and ignores the NCS.

TCS and DCS are fused together to derive the Web Immediacy State (WIS). WIS represents the *immediacy* of the web page which is affected by the web timings (TCS) and the availability of the device (DCS); both attribute states which cause a web page to be either *instantly* available to the user (from initial request to a *ready* state) or have a noticeable *lag*/*delay* in both content fetching and/or content rendering. Once WIS is derived, then it is combined with ICS (Intent Content State) to derive the Intent Weighted Web Immediacy State (IWWIS). IWWIS considers the *immediacy* of the page in regards to the derived Intent Importance of that web page to the user. IWWIS is the MWQoE characterization and it is modeled to reflect the QoE of the user for a particular mobile web session. The experience is reflected as a factor of the *immediacy* of the web page and the *importance* of the web session to the user. The Utility of IWWIS is derived in a linear manner and is the final deliverable of the MWQoE model: the MWQoE metric. The MWQoE metric is linearly transformed on a scale from 0 to 1 as it was initially intended for this study.

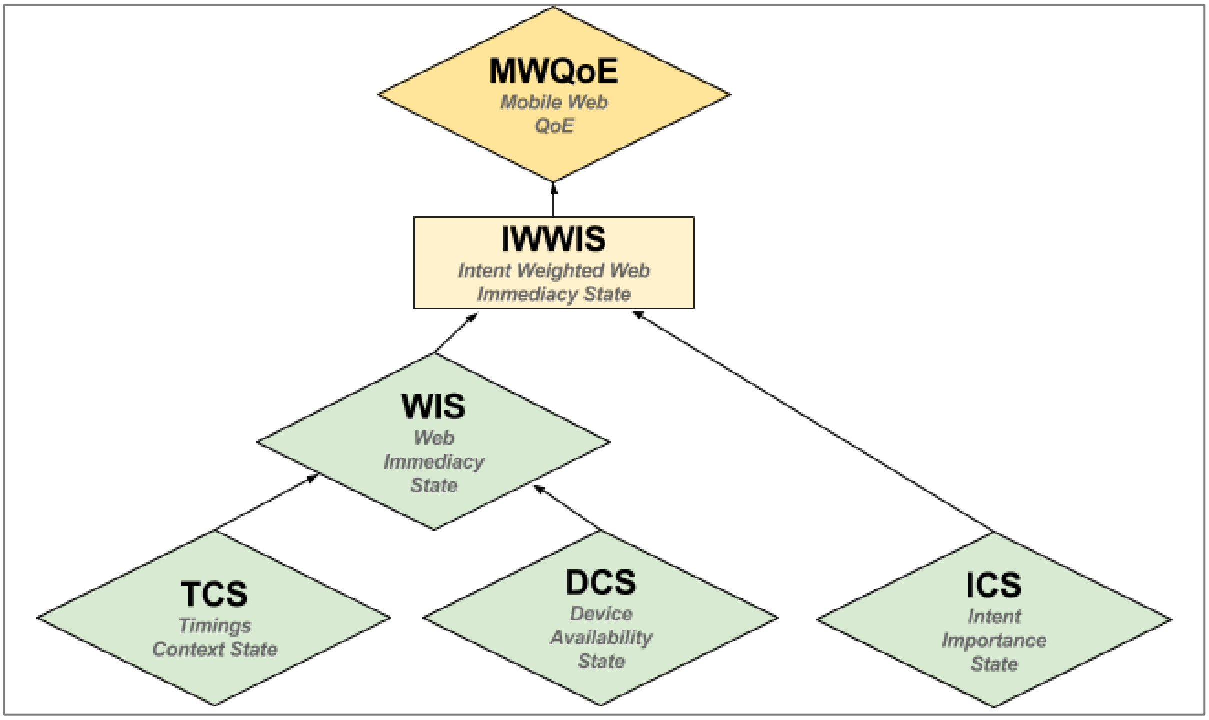


Figure 38 MWQoE Causal Links

Table 27 below shows how each IWWIS State maps via a linear transformation to MWQoE. An *Excellent* IWWIS of a single web session translates into MWQoE=1, whereas a *Poor* IWWIS translates into MWQoE=0.

IWWIS State	State Code	Ordinal Scale	MWQoE (Utility)
EXCELLENT	MWQoE_5	5	1.00
VERY GOOD	MWQoE_4	4	0.75
GOOD	MWQoE_3	3	0.50
FAIR	MWQoE_2	2	0.25
POOR	MWQoE_1	1	0.00

Table 27 IWWIS-MWQoE Mappings

The WIS state is predicated by the assumption that the better the web metrics (lower values) and the device availability, the better the web immediacy for the user and vice versa. WIS is represented by 5 states, *Poor* – *Excellent* (*Table 28*) and gives a slightly higher strength of influence to TCS rather than to DCS given that web attributes have a stronger impact to the *immediacy* of the web page rather than the availability of the web resources, since PL and CL can be affected by DCS, however, DCS cannot be affected by PL and CL. i.e. The Page Load time (PL) can be affected by the processing power and memory available of the device (higher PL with slower machines), but, the processing power of the device cannot be affected by the Page Load time.

WIS	**ICS**	**IWWIS**
Excellent	Very Important	Excellent
	Important	Excellent
	Moderately Important	Excellent
	Somewhat Important	Excellent
	Not Important	Excellent
Very Good	Very Important	Good
	Important	Good
	Moderately Important	Very Good
	Somewhat Important	Very Good
	Not Important	Very Good
Good	Very Important	Fair
	Important	Good
	Moderately Important	Good
	Somewhat Important	Good
	Not Important	Good
Fair	Very Important	Poor
	Important	Poor
	Moderately Important	Fair
	Somewhat Important	Fair
	Not Important	Fair
Poor	Very Important	Poor
	Important	Poor
	Moderately Important	Poor
	Somewhat Important	Poor
	Not Important	Poor

Table 28 WIS-ICS-IWWIS Mappings

TCS	**DCS**	**WIS**
EXCELLENT	High	Excellent
	Medium	Excellent
	Low	Excellent
VERY GOOD	High	Very Good
	Medium	Very Good
	Low	Good
GOOD	High	Good
	Medium	Good
	Low	Fair
FAIR	High	Fair
	Medium	Fair
	Low	Poor
POOR	High	Poor
	Medium	Poor
	Low	Poor

Table 29 TCS-DCS-WIS Mappings

The IWWIS state is composed by fusing together WIS and ICS (Table 29). WIS is given a slightly stronger influence to IWWIS than ICS given the consideration that immediacy is more important to the user than the intention of the web session. Finally, the complete MWQoE Bayesian model is coded in the GeNIe platform, as shown in *Figure 39,* which reveals that the mean MWQoE metric of all observed web sessions is 0.62

(*Good*), that 56% of all web sessions are found to have *Excellent* MWQoE, 2% *Very Good*, 6% *Good*, 6% *Fair* and 30% *Poor*.

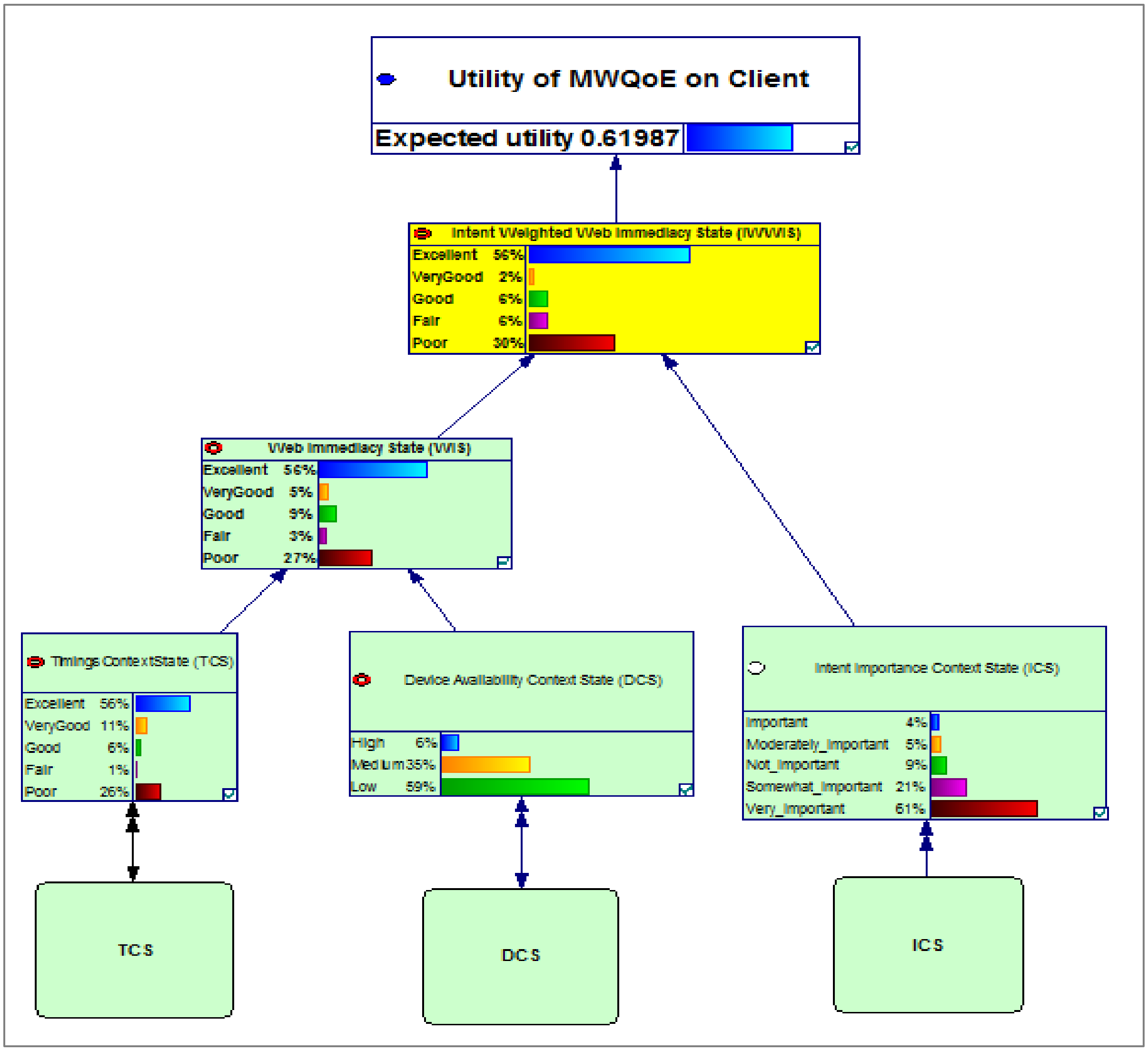


Figure 39 Generated MWQoE Model

Chapter 4 MWQoE Model Evaluation

4.1 Introduction

In the **Model Evaluation Phase**, a Lab Experiment was designed and executed so that the importance and strength of influence of the MWQoE model attributes could be established with mobile web users. It was designed so that users could provide their insights and their own perceived QoE evaluation in specific scenarios identified in the BetterX dataset. This approach enabled MWQoE to be evaluated and compared against user perceptions via an examination of the user's MOS against the generated MWQoE metric; a test which showed that in 93% of scenarios tested, the MWQoE model generated either the same or a more conservative evaluation of the user's MOS, and across all scenarios, provided a close characterization of the user satisfiability (Fair vs Good).

Furthermore, the Lab Experiment tested factors that were not part of the MWQoE model, such as content presentation and content delivery to establish their effect and impact on MWQoE, and examine specific client-side enhancements to improve the MWQoE model for specific scenarios.

The Lab Experiment was designed to provide a user questionnaire using Likert Scales (Kothari, 2004) since this approach has been shown to give '*the most important information about user's perception of the system and their interaction with it*' (Tullis and Albert, 2013). This method was chosen as the most appropriate given the subjectivity and highly personal nature of QoE – as it has been adopted by this study. The goal was to identify lab users that could be matched to the 4 Web Session User Profiles (Worker, Buyer, Reader, Web User) and have them follow a series of web browsing actions (web tasks) drawn from the Web User Daily Stories, i.e. simulate, as closely as possible, the observed web browsing which has been collected and analyzed in the BetterX dataset, Model Generation phase (3-39Chapter 3). In each web task, users had to report on a Likert Scale their *satisfaction level* and answer open-ended questions on the reasons behind their ratings. This approach was designed to enable an evaluation of the MWQoE metric with the user reported MOS. The aim of collecting user satisfaction levels, as mentioned above, was to enable their comparison against the MWQoE metric on comparable situations (relating to the device/network/web domains). This was achieved by simulating the observed scenarios of the BetterX dataset in a controlled lab environment. In addition, the answers to the open-ended questions achieved in collecting further insights about the user's perception in regards to the factors which influence their satisfaction, thus providing solid grounds for future work.

4.2 Design & Implementation

Having stated the above, the aims of the Model Evaluation phase were summarized in 3 categories. The 1st was to **Evaluate** the impact of experience factors that were included in the MWQoE model, the 2nd was to **Test** factors which were not included in MWQoE, and 3rd to observe the **Impact** of 2 additional factors to user satisfaction. The list below provides further information about the aims for each data domain or specific attribute.

1) **Evaluate**: The importance of factors that were directly related to each of the MWQoE context states (TCS, NCS, DCS, ICS) and their effect on MWQoE.
    a. The effect to MWQoE on the immediacy of a web page (WIS state in MWQoE) in regards to the Web Timings (TCS state) and Network Timings (NCS state)
    b. The effect to MWQoE on the responsiveness of the device (DCS state in MWQoE)
    c. The effect to MWQoE on the content relevance per user expectation (ICS state in MWQoE)
2) **Test:** Factors that were not directly related to MWQoE but could potentially become part of it (future work)
    a. The effect to MWQoE of the device screen size
    b. The effect to MWQoE of the ease of use/ease of navigation/interactivity of the web page
3) **Impact**: Factors that could potentially have an impact to MWQoE (enhancement/determent) and needed to be tested.
    a. The effect of Auto-Scrolling via Device Tilting for News Reading while in motion
    b. The effect of simpler web interfaces when purchasing products from e-commerce sites.

In regards to the questionnaires used in Model Evaluation, their structure and content was broken down into 2 sections. The 1st section was the same in both structure and content for all 4 profiles. However, the 2nd section was different in content (web tasks) for each of the 4 profiles but with the same structure. The questionnaire's 1st section included demographic information (Gender, Age, Occupation) and a frequency rating on *mobile device use* and *mobile web use* (same questions as in BetterX Android Setup). Moreover, it included 8 questions with Likert scales for each to capture the user's rating (from 1 to 5, Poor - Excellent) on the importance of each of the following:

1) The immediacy of the web page *(Evaluate the TCS model)*
2) The speed of the mobile network *(Evaluate the NCS model)*
3) The responsiveness of the device *(Evaluate the DCS model)*
4) The quality of the screen *(Test the effect of the quality of the screen)*
5) The size of the screen *(Test the effect of the size of the screen)*

6) The quality of the site being used *(Test the effect of the site's quality)*
7) The relevance of the content in regards to needs/expectations *(Test the effect of context relevance)*
8) The ease of use of the site *(Test the effect of the site's interface)*

The 2nd section included 4 tasks which were different for each profile as each profile tested a different web page which matched the specific profile. Each task involved the use of either a Tablet or a Smartphone and was composed of a series of steps which the user had to follow. The user had to use a mobile device, and for each step of each task give a *satisfiability* rating (from 1-5, Poor – Excellent) after completing it. At the end of each task the user had to provide an overall rating of the experience for the task and answer an open-ended question providing the reasons for his/her rating. Task 1 involved using a tablet and task 2 involved using a smartphone both in an unrestricted Wi-Fi connection which was controlled by the administrator's laptop. Tasks 3 and 4 were the same as tasks 1 and 2 respectively but with the difference that the Wi-Fi connection was capped to 800-900 Bits per second. This was done to measure the difference in MOS of each device with alternating network speeds. The expectation for this was that lower-speeds would yield lower satisfaction levels and higher-speeds would yield higher satisfaction levels (Section 3.4.4.2). The tablet and smartphone were chosen so that a difference in specifications (Display Size, Display Resolution, CPU, GPU, RAM) would uncover the effect of these attributes to the user experience (Section 3.4.4.4). The specs shown in Table 30 show that the tablet had double the display size of the smartphone but the smartphone had almost double the CPU, RAM and triple the screen's pixel density.

Specs[8]	**Samsung Galaxy Tab 2 10.1 P5100**	**Samsung I9505 Galaxy S4**
Released	2012, May	2013, April
Dimensions	256.6 x 175.3 x 9.7 mm	136.6 x 69.8 x 7.9 mm
Weight	588 g (1.30 lb)	130 g (4.59 oz)
Display Type	PLS TFT capacitive touchscreen, 16M colors	Super AMOLED capacitive touchscreen, 16M colors
Display Size	10.1 inches (~65.8% screen-to-body ratio)	5.0 inches (~72.3% screen-to-body ratio)
Display Resolution	800 x 1280 pixels (~149 ppi pixel density)	1080 x 1920 pixels (~441 ppi pixel density)
Multitouch	Yes	Yes
Chipset	TI OMAP 4430	Qualcomm APQ8064T Snapdragon 600
CPU	Dual-core 1.0 GHz Cortex-A9	Quad-core 1.9 GHz Krait 300
GPU	PowerVR SGX540	Adreno 320
Memory Internal	1 GB RAM	2 GB RAM

Table 30 Lab Experiment Device Specs Comparison

[8] *http://www.gsmarena.com/compare.php3?idPhone1=4567&idPhone2=5371*

The BetterX Firefox Addon was installed in both devices logging the Web attributes (Table 1) for every web session performed by the user. The 4 types of questionnaires used and the process followed by the administrator in the lab experiment are outlined below.

4.2.1 Worker Questionnaire

The web site chosen for the Worker Questionnaire (A*ppendix Questionnaire for Worker*) was *mturk.com* as it was the most visited site for the sessions that were classified as *Workers* in the BetterX dataset. The steps for each task were as follows and were chosen since they simulated the process of searching and finding available work on the Mechanical Turk platform:

1. *Go to mturk.com*
2. *Click on the link named View them now to view all the available work on the site.*
3. *Read all of the work titles on the first page and click on one View a HIT in this group*
4. *Review the contents of the page.*

4.2.2 Buyer Questionnaire

The web site chosen for the Buyer Questionnaire (A*ppendix Questionnaire for Buyer*) was *amazon.co.uk*[9] as it was a highly visited ecommerce site in the BetterX dataset. The steps for each task were as follows and were chosen since they simulated the process of searching and finding a specific product on Amazon's platform:

1. *Go to amazon.co.uk*
2. *Search for a white polo short and review the results.*
3. *Locate an item in the first page of the results you wish to review and click on it.*
4. *Review the product page of the item*

The search term *white polo short* was deliberately chosen since it is very common for users to search for *polo shirts* rather than *polo shorts*. The intention of this little tweak in the search term was intentional and based on the observation that the site (amazon.co.uk) would auto-correct and display results to *polo shirts* rather than *polo shorts* which was the intended and pre-defined web session intent for the user, thus diverting the user from the pre-defined web session intent. This was designed to examine how the user's

[9]*Disclosure: Amazon Web Services was a sponsor of this study*

satisfiability changes when their intention and aim for a product search is not satisfied or when the site's auto-correction feature does not correctly recognize the intent of the user.

4.2.3 Reader Questionnaire

The web site chosen for the Reader Questionnaire (*Appendix Questionnaire for Reader*) was *news.google.com* as it was a highly visited news site in the BetterX dataset. The steps for each task were as follows and were chosen since they simulated the process of locating and reading the news from the web browser on the specified site:

1. *Go to news.google.com*
2. *Locate the World section, click on it and review the headlines listed on the first page.*
3. *Pick the one that interests you the most and click on it.*
4. *Review the contents of the news article*

4.2.4 Web User Questionnaire

The web site chosen for the Web User Questionnaire (*Appendix Questionnaire for Web User (Default)*) was *www.google.com* as it was by far the most widely used search engine in the BetterX dataset. The steps for each task were as follows and were chosen since they simulated the process of searching and finding specific information from the web browser on the specified site:

1. *Go to www.google.com*
2. *Search for Toji Japan and review the results of the first page Pick the one that interests you the most and click on it.*
3. *Click on a page you want to review*
4. *Review the contents of the page*

The search term *Toji Japan* was chosen so that users discover a new piece of information from a search engine, eliminating their bias in the search results and mimicking the Web User profile. Given the demographics of the intended sample of the Lab Experiments, it was highly improbable that users would be familiar with this term.

4.2.5 Lab Procedure

Participants were escorted one at a time into a room with comfortable temperature and good lighting, after they've signed the consent form (*Appendix Lab Study Consent Form*). The administrator asked a series of questions to discover how the participants used their mobile device and the sites that they've visited frequently, so that they could be assigned to a matching profile. This was done so that participants could be matched with the most appropriate profile based on their own preference and usage. Afterwards, the administrator handed to the participant the assigned profile questionnaire and the 2 devices (tablet and smartphone) while providing instructions on how to proceed. The administrator logged into a spreadsheet the participant number (as the study was anonymous) and the time the experiment started and finished so that the web metrics could be traced back into each questionnaire using the timestamp in web metrics. Before handing-in the devices, the administrator cleared the Firefox cache in each device so that sites were not available to load from cache. Once the participant completed the second step then the administrator limited the network speed and reset the cache so that tasks could be repeated (tasks 3 and 4).

4.2.6 Lab Procedure for Impact Factors

For the Buyer profile and the Reader profile, an additional step was added after task 4 which entailed for the user to test a potential *QoE impact factor* on the smartphone and report back a satisfiability rating. These impact factors were added in the lab experiment since the aim of the project was not only to identify how to measure and model Web QoE on mobile devices but also to identify ways to enhance it. In doing so 2 potential Web QoE impact factors were designed and implemented after the analysis of the BetterX dataset. These 2 factors, as they are outlined below, aimed to provide client-side changes in either the web content or in the delivery of the web content and test whether they enhance the user's satisfiability.

For the Buyer profile, a custom Firefox CSS Injection add-on was used to hide parts of the web site including advertisements, secondary menus, etc. making *amazon.co.uk's* interface *cleaner* and *less busy*. The aim of the add-on was to test the effect of a cleaner web interface to the user's *satisfiability*. For the Reader profile, a custom Firefox Auto-Scrolling add-on was used which enabled the automatic vertical scroll of the web page when the device was tilted by the user. The page scrolled upwards when the device was tilted away from the user, and downwards when the device was tilted towards the user. The participant was instructed to walk while trying to read the contents of the news page (*news.google.com*) and then attempt to do the same with the add-on enabled so that user satisfiability rating could be compared and the effect of the add-on to the experience of the user could be quantified.

4.2.7 Sample Overview

A total of 31 mobile device users participated in the Lab Experiment; 15 males and 16 females having minimum age of 21, maximum of 56 and average age of 37. Thirty-two percent of the users (10 users) had IT related backgrounds, 39% (12 users) held managerial level positions at various fields, 13% (4 users) were academics of various disciplines and the remaining 16% (5 users) were professionals/consultants in various fields and 1 university student.

Each of the 31 users completed 1 of the 4 available questionnaires. As shown in Figure 40, a total of 14 participants completed the Web User questionnaire, 8 completed the Buyer questionnaire, 7 the Reader questionnaire and 2 the Worker questionnaire. The profile of the Worker was hard to match with participants since there are not too many professionals in Cyprus who work online and have experience in crowdsourcing platforms. The profile of the Web user was the most popular since it was used as the default profile given that all the participants of the experiment had experience in using a search engine for specific enquiries. Since the purpose of the Lab Experiment was to evaluate the effect of attributes which were included in MWQoE and verify expected experiences from the users in all profiles, the results of this experiment are looked at collectively, i.e. as a set of 31 users and not per individual profile, similarly to the Model Generation phase which approached all users with the same data collection methodology and subsequently categorized their experiences. Future work plans to specifically analyse individual profiles, a task that would need a higher user recruitment per profile, especially for the worker profile, which only recruited 2 users in the lab experiment.

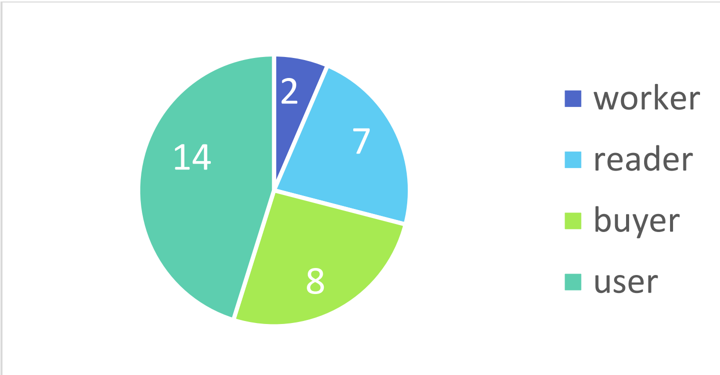


Figure 40 Participants Per User Profile

In regards to the level of experience of each participant relating to mobile device and web usage from mobile device, more than 80% of the sample reported that they use their device *very often* and more than *very often*, whereas, 90% of the sample reported that they use the web from their mobile device *often* and more than *often* (*Figure 41* and *Figure 42* respectively).

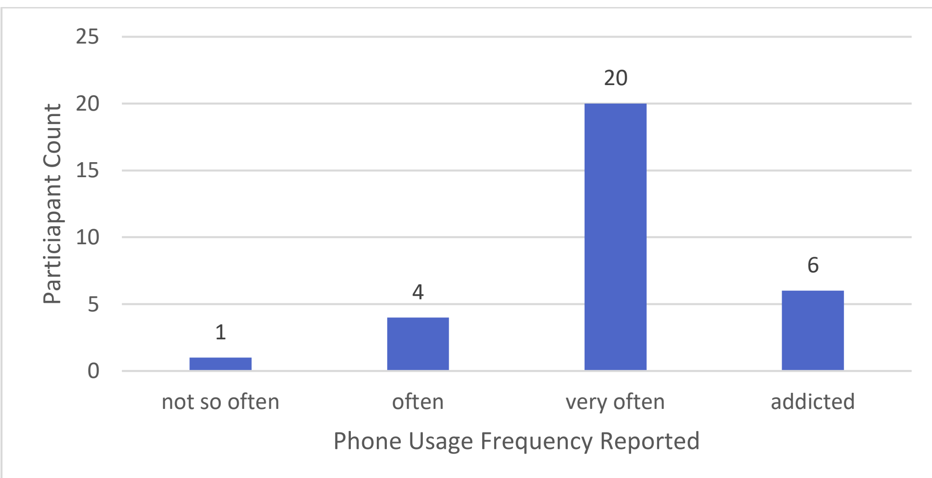


Figure 41 User Reported Phone Usage Frequency

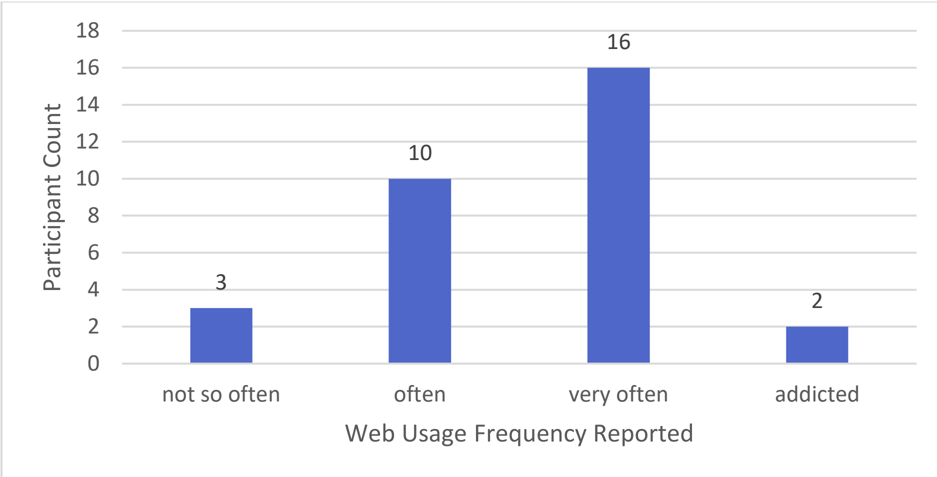


Figure 42 User Reported Web Usage Frequency

It is noted that both the BetterX sample and the Lab sample hold well distributed demographics and have similar and therefore comparable mobile usage experience levels. In regards to device usage, 78% of the BetterX users reported *Very Often* usage and higher, in comparison with 83% of the Lab users which reported the same frequencies. In regards to mobile web use, both BetterX and Lab users reported exactly 58% of *Very Often* usage and higher.

The MOS for each lab participant in each of the 4 tasks is shown in Table 31. One of the 62 phone tasks was not completed, thus no MOS was reported because the participant of that task was running out of time and had to skip 1 of the tests. In addition, 7 out of the 62 tablet tasks were not reported due to battery failures of the tablet during the last tests which forced the tablet to shut down.

Type	Task 1 MOS (Tablet)	Task 2 MOS (Phone)	Task 3 MOS (Tablet)	Task 4 MOS (Phone)
Buyer	1	1	1	1
Buyer	4	5	3	4
Buyer	2	4		4
Buyer	4	5	4	4
Buyer	5	5	3	4
Buyer	4	4	4	4
Buyer	2	3		
Buyer		4		4
Reader	3	3	3	3
Reader	4	5	3	4
Reader	3	4	4	4
Reader	4	5	3	2
Reader	3	4	4	4
Reader	3	5	2	5
Reader	5	4	4	2
Web User (Default)	3	2	4	3
Web User (Default)	5	5	4	5
Web User (Default)	4	5	4	5
Web User (Default)	4	4	5	5
Web User (Default)	3	4	3	4
Web User (Default)	4	4	3	5
Web User (Default)	4	5	2	5
Web User (Default)	4	5	4	5
Web User (Default)	4	5	5	4
Web User (Default)	2	4	3	4
Web User (Default)	4	4	4	4
Web User (Default)	4	3	4	4
Web User (Default)	4	4		4
Web User (Default)		5		5
Worker	4	3	4	3
Worker	2	1	1	2

Table 31 Lab Experiment Participant's MOS in Tasks 1-4

4.3 Results

The results from the lab data analysis confirm the importance of the factors used in the MWQoE model via direct user responses. Factors such as the *Device Responsiveness, Page Immediacy* and *Internet Speed* are confirmed to be the most important factors for users when accessing the web from a smartphone or tablet. Moreover, the effectiveness of the MWQoE model is evaluated in a MOS vs MWQoE comparison across all user profiles, which reveals that MWQoE closely reflects MOS (Section 4.3.6). In addition, the effect of the Network Speed and the effect of the Device characteristics on MOS is measured and relationships coded in the MWQoE model in regards to those attributes are verified. Following is a detailed analysis of all the findings.

4.3.1 Reported Importance of Experience Factors

The importance of all the experience factors which are used in MWQoE and tested in the lab experiment is confirmed, as shown in Figure 43. *Device Responsiveness, Page Immediacy* and *Internet Speed*, factors which map to each of the modelled context states of MWQoE are reported as the most important experience factors by users. The *Device Responsiveness* factor is modeled in MWQoE by the Device Availability State (DCS) and the *Page Immediacy* factor is modelled by the Timings Context State (TCS) which are both combined to provide the Web Immediacy State (WIS). The inclusion of these context states in the MWQoE and their importance is verified by actual user ratings.

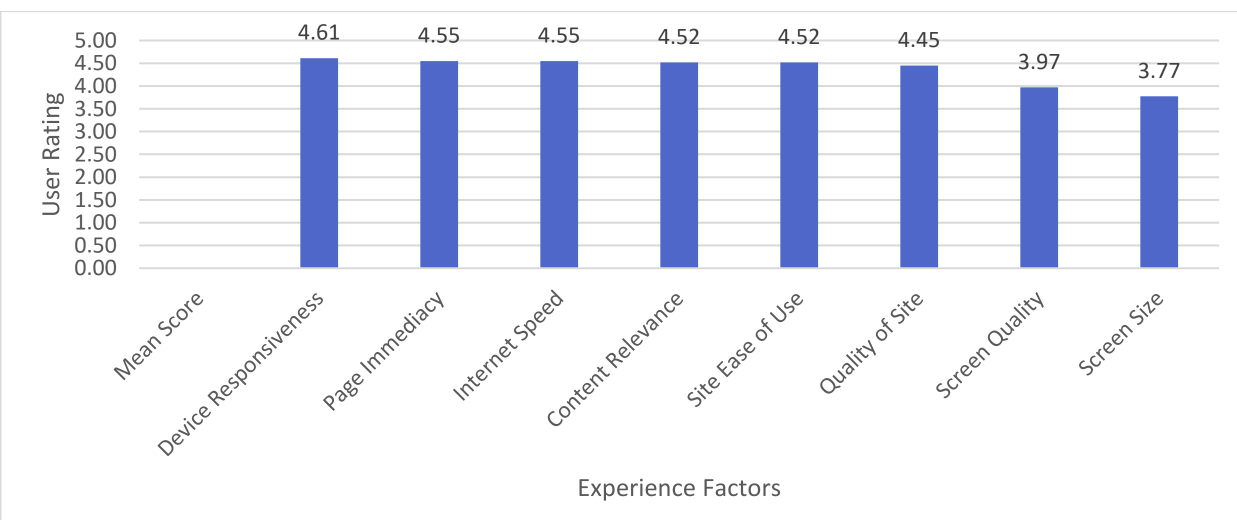


Figure 43 Importance of Experience Factors ranked by Users

Moreover, with negligible difference follow the qualitative attributes of web pages such as *Content Relevance*, *Site Ease of Use* and *Quality of the Site*. These are attributes which are not directly mapped with MWQoE. However, this finding paves the way for future MWQoE model enhancements with the inclusions of these attributes since their importance to QoE is verified by user ratings.

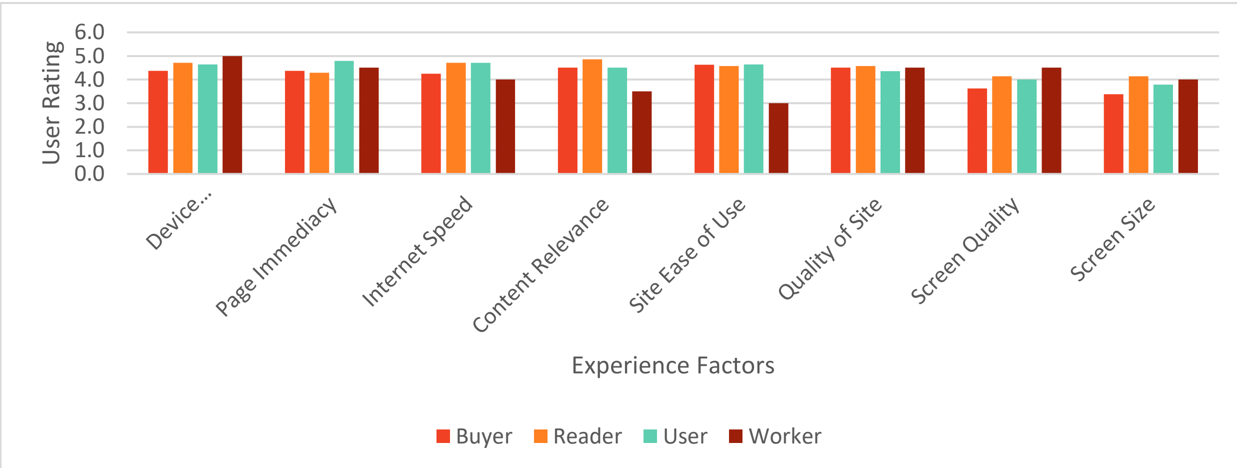


Figure 44 Importance of Experience Factors per profile ranked by Users

As it is shown in Figure 44, the importance of all the selected factors remains the same throughout all 4 user profiles, a fact that is considered in the MWQoE model. Although, a slightly reduced importance of *Site Ease of Use* was reported by Online Workers, it does not directly affect MWQoE since this attribute is not included in the model. Moreover, given the fact that the sample size of the Workers is only 2 users this finding can be considered as immaterial.

4.3.2 Reported MOS per Device

The overall MOS for smartphones in all profiles is found to be 11% higher than the MOS of tablets, as shown in Figure 45. The noticeable difference in performance of the 2 devices is reported by the majority of participants citing performance and screen issues with the tablet. The factors which are modeled in

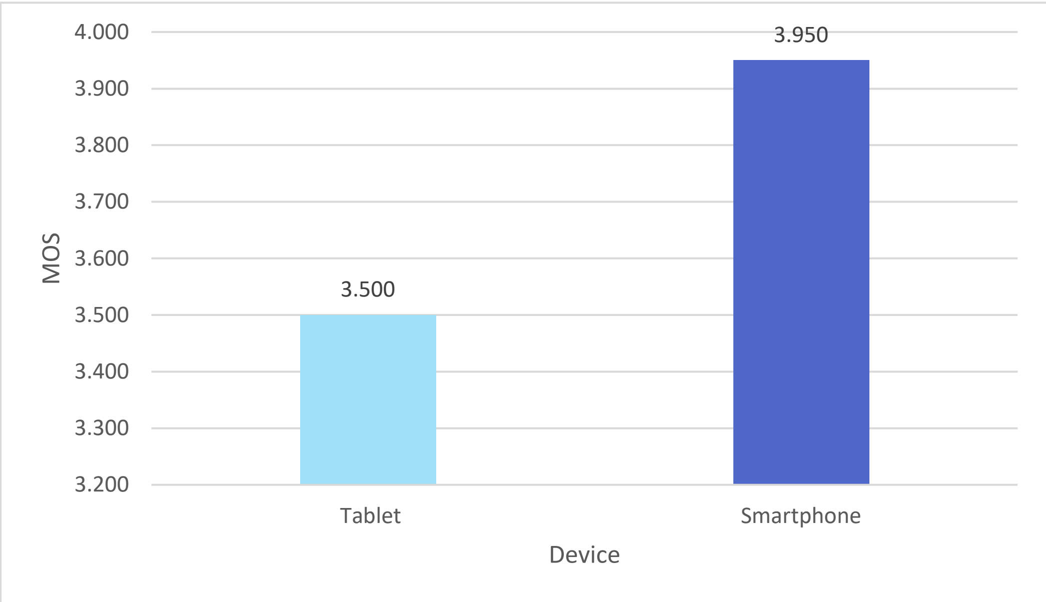


Figure 45 Overall Reported MOS per Device

MWQoE to affect the Device Context State (DCS) and subsequently the *immediacy* of the web page (WIS) are shown to be aligned with the lab test results since a higher performance device (such as the smartphone) produced higher MOS than a device of lower performance (such as the tablet). The tablet having half of the processing and rendering power of the smartphone (Table 30) is reported by users as *irresponsive* and *"delivering choppy sites"* although the speed of the network is the same on both devices. Device features such as CPU, RAM, GPU and screen pixel density are shown to affect the quality of experience of web users.

4.3.3 Reported MOS per Network Speed

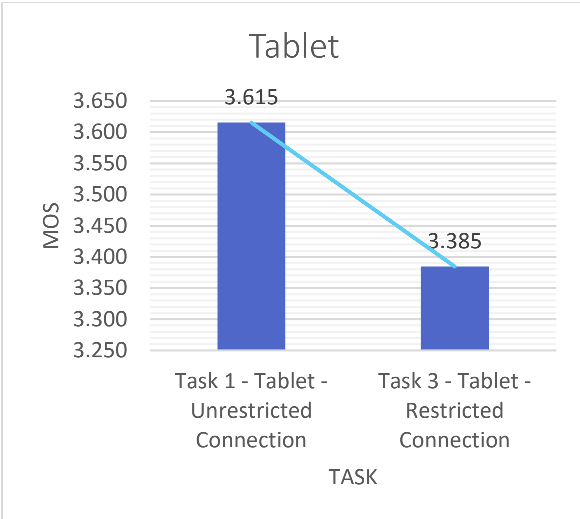


Figure 46 Tablet MOS

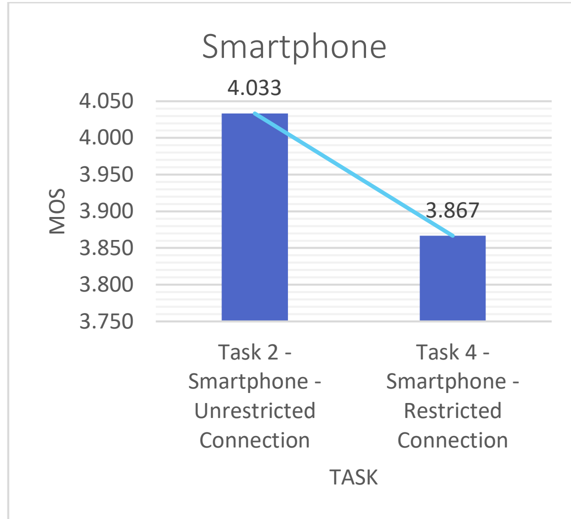


Figure 47 Smartphone MOS

It is verified that the modeled relationship in MWQoE between the speed of the network (NCS) and subsequently to the Web timings context state (TCS) has a direct effect on experience. MOS is decreased by 6% when the speed of the network is restricted on the tablet, and 4% on the smartphone. The slight difference in MOS fluctuation between the devices can be attributed to the fact that the *lag* in overall *web immediacy* was observed to be more evident on the slower machine (tablet) than a faster machine (smartphone). Figure 46 and Figure 47 show the MOS for each device for each of the associated tasks and the drop in MOS for each device with different connection speeds. It is noted that almost half of the users reported the speed decrease on the tablet (Task 3) whereas only a 1 in 5 users reported the speed decrease on the smartphone (Task 4).

4.3.4 Findings per User Profile and Task

For the Buyer profile, users reported that they were not that satisfied with the loading time of the site on the tablet in default network speed, however, they were satisfied with the site on the smartphone in default speed. As shown in Figure 48 and Figure 49, the amazon.co.uk site delivered a different layout and design on the smartphone’s smaller screen which was preferred by the users, as opposed to the design delivered on the tablet. Moreover, it was noted by users that the site’s auto correction feature from *white polo shorts* to *white polo shirts* caused frustration since the user intention was not met and users had to take extra steps to reach to the desired search results.

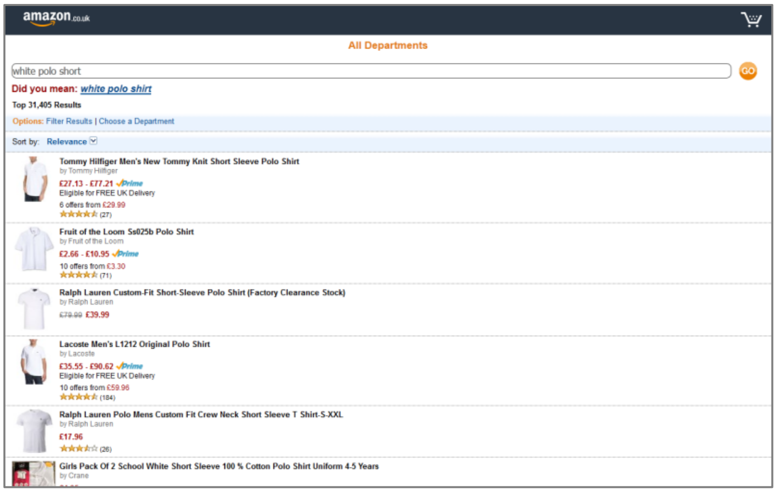


Figure 48 Amazon.co.uk on tablet

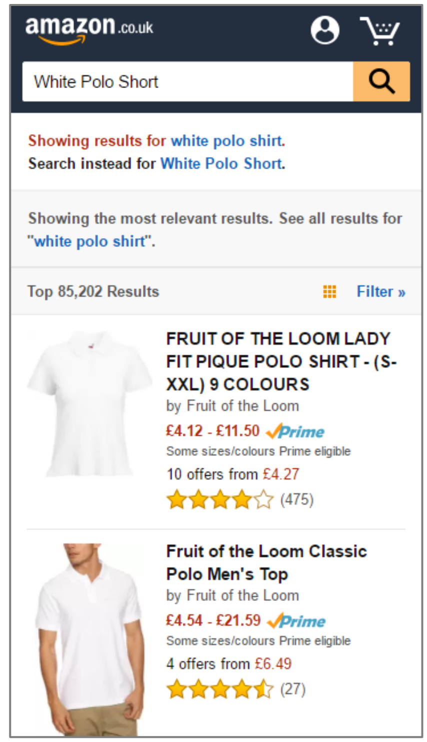


Figure 49 Amazon.co.uk on smartphone

For the Reader profile, half of the users reported that the loading time on the tablet was not satisfactory with the default connection. The connection speed drop was noted primarily on the tablet and only a few noticed it on the smartphone. For the Worker profile, the speed decrease was noted in both devices from both participants and the user's satisfaction was decreased given the design of the site was not optimal for smaller screens. Users reported that they had to zoom in to different parts of the site in order to read the content. For the Web User profile, the speed drop was noted on the tablet and not on the smartphone as users regarded the loading time on the smartphone to be satisfactory in both cases. This can be attributed to the fact that google.com is relatively *light*. Users reported that they liked the design of the results especially on the smartphone for google.com. Moreover, most users liked the content of the search results while others reported that they were not satisfied with it without giving any further explanation.

The page with the highest MOS is the *search results page* of google.com on the smartphone with a score of 4.8, followed by the home page of google.com and the news.google.com home page both on the smartphone with a MOS of 4.4, and the amazon.co.uk home page on the smartphone with a 4.3.

4.3.5 User Bias and Favourable Response Trend

It has been shown in literature that one of the limitations of questionnaires is that users typically tend to give more positive feedback, as they believe '*it makes them look better in the eyes of others*' (Tullis and Albert, 2013). This is named Social Desirability Bias (Nancarrow, Brace & Wright, 2001) and is identified in the responses of the lab participants as a *favorable response trend*.

The questionnaire was designed in such a way that users had to rate each individual web session and then rate the web sessions collectively. Any difference in the calculated average of all web sessions per task with the reported overall average would constitute the favorability rate of the user towards that specific task. During data analysis, the difference of the two scores (calculated vs reported) was considered for all participants. Each difference was marked as favorable (+1 score) if the round-off of the average was towards a higher MOS rating or if the reported score was higher than the calculated; and a non-favorable (-1 score) was applied if the round-off was towards a lower MOS rating or if the reported score was lower than the calculated score. A neutral score of zero was applied when the difference was zero or when the difference was less or equal to 0.25 (calculated minus reported). Then for each task the sum of all ratings was drawn, outlining at which tasks users answered more favorably. Results shown in Figure 50 reveal that Task 3 is the most favorably rated task, and that is the task with the slowest machine (the tablet) on the restricted network speed. It is also noted that in general, users answer much more positively than negatively thus *exaggerating* their satisfaction ratings.

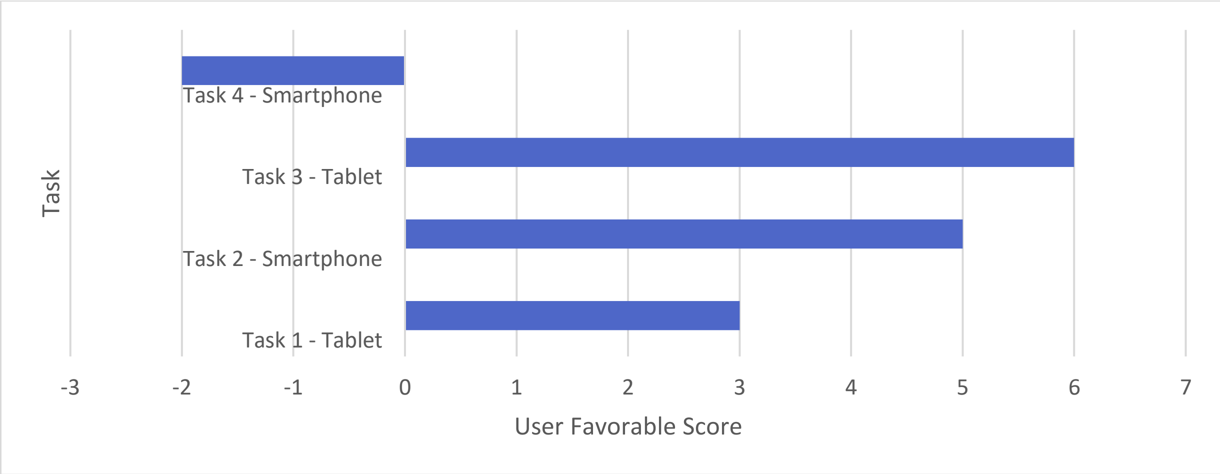


Figure 50 User Favorable Scores

4.3.6 MOS vs MWQoE

The final and most important data analysis task of the lab data is the evaluation of the user reported MOS with the MWQoE metric. The comparison between the user MOS across all profiles closely compares with the MWQoE metric (as shown in Table 32, Figure 51 and Figure 52). The comparison reveals a conservative MWQoE metric, one which provides either the same or a lower evaluation of the user's satisfiability in 93% of all comparisons. This feature works for the benefit of the user, as it has been intended for this project, since a lower metric can provide the basis for potential enhancements to the user. In addition, this comparison reveals that MWQoE is evaluated closely with the actual user averages since 66% of total observations are found to be either the same or close to the model (1 or 2 degrees' difference). Furthermore, and given that this project is the first attempt to provide a unifying metric for MWQoE, the 28% of all observations which have been found to digress further from actual user opinion (3 and 4 degrees' difference) together with the 6% where user opinion is lower than the metric are the basis for improvement for future work with a greater sample size.

The MOS vs MWQoE evaluations for all tasks performed in the lab experiment across all profiles are shown in Table 32. The Context-States are generated for MWQoE based on the data collected from the smartphone and tablet in the lab experiment. The MOS for each task is compared with the MWQoE metric of the same task. The rows in Table 32 which are highlighted in shades of green show that 66% of all comparisons closely reflect user opinion, whereas the records highlighted with shades of orange (28%) reveal a higher digress of the MWQoE metric from user opinion. The records highlighted with grey (6%) indicate the observations where the metric gave a higher evaluation than the user.

In this analysis, the actual Web attributes recorded from each Task are fed into the MWQoE model forming each task evaluation test. The TCS state is evaluated comparing actual TCS metrics from both scenarios (PL, CL, etc.). For the DCS state, tablet tasks are given *Medium* availability and smartphone tasks were given *High* availability. This is done considering the specifications of each device (Table 30) and the fact that the smartphone is faster and has more processing and rendering power than the tablet, therefore has higher availability. In addition, the ICS for each profile is coded in a manner that is aligned with the 11 Scenario Importance Narratives (Table 26):

- Buyer ➔ ICS = *Important*
- Reader ➔ ICS = *Moderately Important*
- Web User ➔ ICS = Somewhat *Important*
- Worker ➔ ICS = *Very Important*

User Profile	**MOS**	**MWQoE**	**Context States**
Reader	Fair (2)	Very Good (4)	*TCS (CL03, PL04, THTTP02, TRTS02, TRES04) = Very Good, DCS (Phone) = High, ICS = Moderately Important*
Reader	Good (3)	Very Good (4)	*TCS (CL04, PL04, THTTP03, TRTS01, TRES03) = Very Good, DCS (Phone) = High, ICS = Moderately Important*
User (Default)	Fair (2)	Good (3)	*TCS (CL03, PL05, THTTP02, TRTS03, TRES04) = Good, DCS (Phone) = High, ICS = Somewhat Important*
User (Default)	Good (3)	Very Good (4)	*TCS (CL02, PL03, THTTP03, TRTS02, TRES03) = Very Good, DCS (Phone) = High, ICS = Somewhat Important*
Buyer	Poor (1)	Poor (1)	*TCS (CL06, PL08, THTTP09, TRTS01, TRES09) = Poor, DCS (Phone) = High, ICS = Important*
Buyer	Good (3)	Good (3)	*TCS (CL05, PL05, THTTP04, TRTS04, TRES09) = Good, DCS (Phone) = High, ICS = Important*
Reader	Very Good (4)	Very Good (4)	*TCS (CL04, PL04, THTTP03, TRTS02, TRES04) = Very Good, DCS (Phone) = High, ICS = Moderately Important*
User (Default)	Good (3)	Good (3)	*TCS (CL02, PL05, THHTP01, , TRES01) = Good, DCS (Tablet) = Medium, ICS = Somewhat Important*
User (Default)	Very Good (4)	Very Good (4)	*TCS (CL02, PL03, THTTP02, TRTS03, TRES03) = Very Good, DCS (Phone) = High, ICS = Somewhat Important*
User (Default)	Very Good (4)	Very Good (4)	*TCS (CL03, PL04, THTTP03, TRTS03, TRES04) = Very Good, DCS (Phone) = High, ICS = Somewhat Important*
Worker	Good (3)	Good (3)	*TCS (CL04, PL04, THTTP02, TRTS01, TRES03) = Very Good, DCS (Phone) = High, ICS = Very Important*
Worker	Poor (1)	Poor (1)	*TCS (CL04, PL07, THTTP02, TRTS01, TRES03) = Fair, DCS (Phone) = High, ICS = Very Important*
Worker	Poor (1)	Poor (1)	*TCS (CL05, PL07, THTTP02, TRTS03, TRES03) = Fair, DCS (Tablet) = Medium, ICS = Very Important*
Buyer	Fair (2)	Poor (1)	*TCS (CL09, PL10, THTTP04, TRTS04,) = Poor, DCS (Tablet) = Medium, ICS = Important*
Buyer	Very Good (4)	Good (3)	*TCS (CL04, PL05, THTTP04, TRTS03, TRES08) = Good, DCS (Phone) = High, ICS = Important*
Buyer	Very Good (4)	Good (3)	*TCS (CL04, PL04, THTTP05, TRTS04, TRES09) = Very Good, DCS (Phone) = High, ICS = Important*
Buyer	Very Good (4)	Good (3)	*TCS (CL05, PL05, THTTP05, TRTS04, TRES09) = Good, DCS (Phone) = High, ICS = Important*
Buyer	Very Good (4)	Good (3)	*TCS (CL05, PL05, THTTP05, TRTS03, TRES09) = Good, DCS (Phone) = High, ICS = Important*
Reader	Fair (2)	Poor (1)	*TCS (CL04, PL10, THTTP03, TRTS04, TRES05) = Poor, DCS (Phone) = High, ICS = Moderately Important*
User (Default)	Excellent (5)	Very Good (4)	*TCS (CL02, PL04, THTTP02, TRTS02, TRES03) = Very Good, DCS (Phone) = High, ICS = Somewhat Important*
User (Default)	Excellent (5)	Very Good (4)	*TCS (CL02, PL04, THTTP03, TRTS02, TRES03) = Very Good, DCS (Phone) = High, ICS = Somewhat Important*
User (Default)	Excellent (5)	Very Good (4)	*TCS (CL02, PL04, THTTP02, TRTS03, TRES03) = Very Good, DCS (Phone) = High, ICS = Somewhat Important*
User (Default)	Excellent (5)	Very Good (4)	*TCS (CL02, PL03, THTTP02, TRTS02, TRES03) = Very Good, DCS (Phone) = High, ICS = Somewhat Important*
User (Default)	Fair (2)	Poor (1)	*TCS (CL08, PL07, THTTP04, TRTS02, TRES04) = Fair, DCS (Tablet) = Medium, ICS = Somewhat Important*
User (Default)	Fair (2)	Poor (1)	*TCS (CL06, PL07, THTTP04, TRTS01, TRES09) = Fair, DCS (Tablet) = Medium, ICS = Somewhat Important*
User (Default)	Excellent (5)	Very Good (4)	*TCS (CL03, PL04, THTTP02, TRTS02, TRES03) = Very Good, DCS (Phone) = High, ICS = Somewhat Important*
User (Default)	Excellent (5)	Very Good (4)	*TCS (CL02, PL04, THTTP02, TRTS02, TRES03) = Very Good, DCS (Phone) = High, ICS = Somewhat Important*
Worker	Very Good (4)	Good (3)	*TCS (CL04, PL04, THTTP02, TRTS03, TRES03) = Very Good, DCS (Tablet) = Medium, ICS = Very Important*
Worker	Very Good (4)	Good (3)	*TCS (CL03, PL05, THTTP02, TRTS03, TRES03) = Good, DCS (Tablet) = Medium, ICS = Very Important*
Worker	Fair (2)	Poor (1)	*TCS (CL04, PL08, THTTP02, TRTS02, TRES03) = Poor, DCS (Tablet) = Medium, ICS = Very Important*
Worker	Fair (2)	Poor (1)	*TCS (CL04, PL08, THTTP02, TRTS01, TRES03) = Poor, DCS (Phone) = High, ICS = Very Important*

Buyer	Very Good (4)	Fair (2)	*TCS (CL05, PL07, THTTP04, TRTS01,) = Fair, DCS (Phone) = High, ICS = Important*
Buyer	Excellent (5)	Good (3)	*TCS (CL04, PL05, THTTP05, TRTS03, TRES09) = Good, DCS (Phone) = High, ICS = Important*
Buyer	Excellent (5)	Good (3)	*TCS (CL05, PL05, THTTP04, TRTS04, TRES09) = Good, DCS (Phone) = High, ICS = Important*
Buyer	Very Good (4)	Fair (2)	*TCS (CL05, PL06, THTTP04, TRTS02, TRES08) = Fair, DCS (Phone) = High, ICS = Important*
Buyer	Very Good (4)	Fair (2)	*TCS (CL05, PL07, THTTP05, TRTS04, TRES09) = Fair, DCS (Phone) = High, ICS = Important*
Reader	Good (3)	Poor (1)	*TCS (CL08, PL07, THTTP08, TRTS02, TRES10) = Fair, DCS (Tablet) = Medium, ICS = Moderately Important*
Reader	Good (3)	Poor (1)	*TCS (CL07, PL06, THTTP03, TRTS02, TRES10) = Fair, DCS (Tablet) = Medium, ICS = Moderately Important*
Reader	Good (3)	Poor (1)	*TCS (CL08, PL08, THTTP02, , TRES01) = Poor, DCS (Phone) = High, ICS = Moderately Important*
Reader	Good (3)	Poor (1)	*TCS (CL09, PL08, THTTP09, TRTS02, TRES10) = Poor, DCS (Tablet) = Medium, ICS = Moderately Important*
Reader	Very Good (4)	Fair (2)	*TCS (CL07, PL06, THTTP02, TRTS01, TRES01) = Fair, DCS (Phone) = High, ICS = Moderately Important*
Reader	Excellent (5)	Good (3)	*TCS (CL05, PL05, THTTP03, TRTS02, TRES04) = Good, DCS (Phone) = High, ICS = Moderately Important*
Reader	Very Good (4)	Fair (2)	*TCS (CL06, PL07, THTTP03, TRTS01, TRES07) = Fair, DCS (Phone) = High, ICS = Moderately Important*
User (Default)	Excellent (5)	Good (3)	*TCS (CL05, PL05, THTTP03, TRTS02, TRES03) = Good, DCS (Tablet) = Medium, ICS = Somewhat Important*
User (Default)	Good (3)	Poor (1)	*TCS (CL06, PL08, THTTP07, TRTS02, TRES07) = Poor, DCS (Tablet) = Medium, ICS = Somewhat Important*
User (Default)	Good (3)	Poor (1)	*TCS (CL10, PL09, THTTP02, TRTS02, TRES10) = Poor, DCS (Tablet) = Medium, ICS = Somewhat Important*
User (Default)	Good (3)	Poor (1)	*TCS (CL05, PL08, THTTP02, TRTS02, TRES05) = Poor, DCS (Tablet) = Medium, ICS = Somewhat Important*
User (Default)	Very Good (4)	Fair (2)	*TCS (CL06, PL07, THTTP03, TRTS02, TRES06) = Fair, DCS (Phone) = High, ICS = Somewhat Important*
User (Default)	Good (3)	Poor (1)	*TCS (CL08, PL08, THTTP07, , TRES10) = Poor, DCS (Tablet) = Medium, ICS = Somewhat Important*
User (Default)	Very Good (4)	Fair (2)	*TCS (CL06, PL07, THTTP02, TRTS02, TRES05) = Fair, DCS (Phone) = High, ICS = Somewhat Important*
User (Default)	Very Good (4)	Fair (2)	*TCS (CL10, PL06, THTTP03, TRTS03, TRES04) = Fair, DCS (Phone) = High, ICS = Somewhat Important*
Buyer	Excellent (5)	Fair (2)	*TCS (CL04, PL06, THTTP04, TRTS01,) = Fair, DCS (Phone) = High, ICS = Important*
Buyer	Very Good (4)	Poor (1)	*TCS (CL09, PL09, THTTP04, TRTS03, TRES10) = Poor, DCS (Tablet) = Medium, ICS = Important*
Buyer	Very Good (4)	Poor (1)	*TCS (CL09, PL10, THTTP03, TRTS03, TRES10) = Poor, DCS (Tablet) = Medium, ICS = Important*
Reader	Excellent (5)	Fair (2)	*TCS (CL07, PL07, THTTP02, TRTS01, TRES01) = Fair, DCS (Phone) = High, ICS = Moderately Important*
Reader	Very Good (4)	Poor (1)	*TCS (CL07, PL07, THTTP02, TRTS01, TRES10) = Fair, DCS (Tablet) = Medium, ICS = Moderately Important*
Reader	Very Good (4)	Poor (1)	*TCS (CL10, PL09, THTTP10, TRTS01, TRES03) = Poor, DCS (Phone) = High, ICS = Moderately Important*
Reader	Very Good (4)	Poor (1)	*TCS (CL08, PL09, THTTP03, TRTS01, TRES01) = Poor, DCS (Phone) = High, ICS = Moderately Important*
User (Default)	Very Good (4)	Poor (1)	*TCS (CL10, PL09, THTTP03, TRTS02, TRES03) = Poor, DCS (Tablet) = Medium, ICS = Somewhat Important*
User (Default)	Very Good (4)	Poor (1)	*TCS (CL10, PL09, THTTP06, TRTS02, TRES09) = Poor, DCS (Tablet) = Medium, ICS = Somewhat Important*
User (Default)	Very Good (4)	Poor (1)	*TCS (CL03, PL06, THTTP02, TRTS02, TRES09) = Fair, DCS (Tablet) = Medium, ICS = Somewhat Important*
User (Default)	Very Good (4)	Poor (1)	*TCS (CL08, PL08, THTTP05, TRTS02,) = Poor, DCS (Tablet) = Medium, ICS = Somewhat Important*
User (Default)	Very Good (4)	Poor (1)	*TCS (CL05, PL10, THTTP02, TRTS03, TRES03) = Poor, DCS (Phone) = High, ICS = Somewhat Important*
User (Default)	Very Good (4)	Poor (1)	*TCS (CL10, PL09, THTTP03, TRTS02, TRES04) = Poor, DCS (Phone) = High, ICS = Somewhat Important*
User (Default)	Very Good (4)	Poor (1)	*TCS (CL09, PL08, THTTP02, TRTS03, TRES02) = Poor, DCS (Phone) = High, ICS = Somewhat Important*
Reader	Excellent (5)	Poor (1)	*TCS (CL04, PL09, THTTP03, TRTS02, TRES06) = Poor, DCS (Phone) = High, ICS = Moderately Important*
Reader	Excellent (5)	Poor (1)	*TCS (CL08, PL10, THTTP09, TRTS01, TRES05) = Poor, DCS (Phone) = High, ICS = Moderately Important*
User (Default)	Excellent (5)	Poor (1)	*TCS (CL08, PL09, THTTP02, TRTS02, TRES03) = Poor, DCS (Phone) = High, ICS = Somewhat Important*
User (Default)	Excellent (5)	Poor (1)	*TCS (CL10, PL09, THTTP02, TRTS03, TRES03) = Poor, DCS (Phone) = High, ICS = Somewhat Important*
User (Default)	Excellent (5)	Poor (1)	*TCS (CL09, PL08, THTTP02, TRTS02, TRES03) = Poor, DCS (Phone) = High, ICS = Somewhat Important*

User (Default)	Excellent (5)	Poor (1)	*TCS (CL05, PL10, THTTP04, TRTS03, TRES04) = Poor, DCS (Phone) = High, ICS = Somewhat Important*

Table 32 MOS vs MWQoE Evaluations (Green 66%, 0-2-degree difference, Orange 28% 3-4-degree difference, Grey 6%, MOS is higher than MWQoE)

The scatter plot (Figure 51) shows the relationship between MWQoE and MOS. The generated trend line of y = 0.2771x + 3.1524 reveals a positive gradient where the y values are the discrete MOS ratings of the user and the x-axis are the MWQoE evaluations generated by the lab experiment device data together with the TCS, DCS and ICS context-states. The positive gradient, evident in Figure 51, shows that the higher the MWQoE prediction, the higher the user reported MOS. The offset from the origin of the line y = 0.2771x+3.1524, crossing the y axis 3.1524 units above zero, shows that the MWQoE model is quite more conservative than the MOS ratings of the users. This finding is reinforced in Figure 52 as well, which shows the distribution of the difference between MWQoE and MOS. Figure 52 shows that in only 6% of the total web session observations MOS was higher than MWQoE and in 94% the MWQoE model provided the same or a more conservative evaluation.

The conservatism of the MWQoE model is a significant result as it works for the benefit of the user since a lower system MWQoE prediction, in comparison with the actual user rating, provides the basis for potential enhancements to the device and subsequently an improvement to the user's MWQoE. Moreover, this finding was expected given the user favourable response trend (Figure 50) which reinforces literature on the Social Desirability Bias (Nancarrow, Brace & Wright, 2001) in which users tend to answer more favourably and exaggerate ratings.

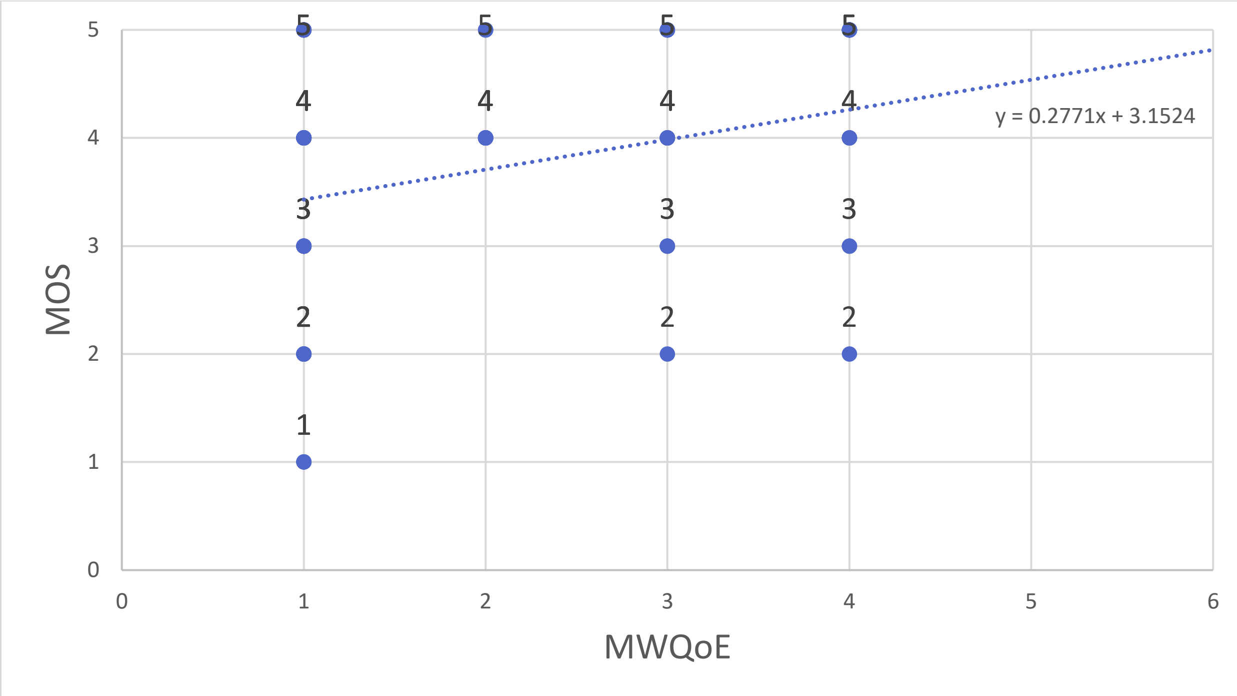


Figure 51 Mapping MWQoE & MOS Relationship

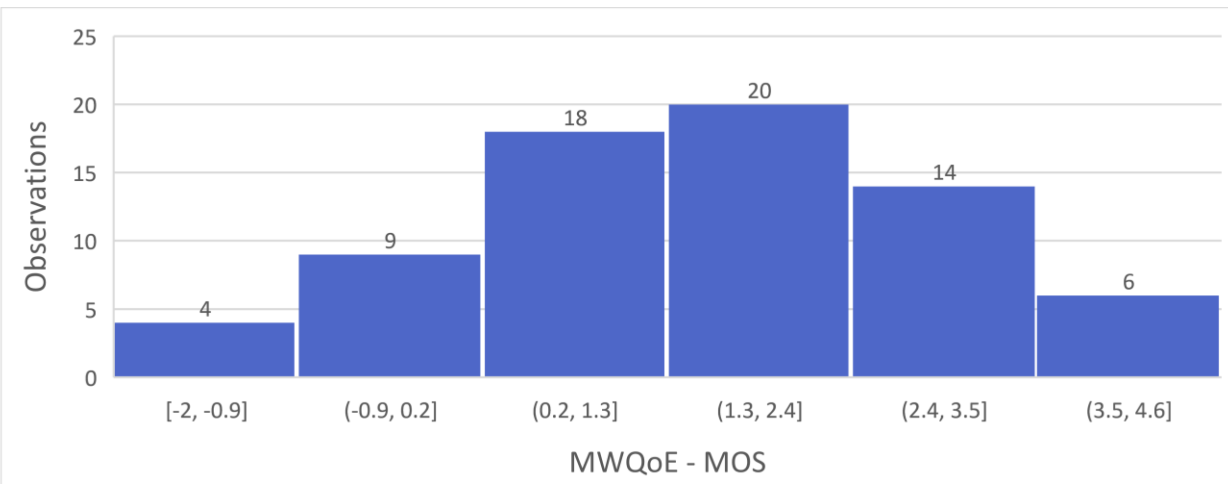


Figure 52 Distribution of the difference between MWQoE & MOS

4.3.7 Additional Findings

The Firefox auto-scrolling add-on via device tilt (Section 4.2.6) is shown to decrease the user's *satisfiability* while the user is in motion (Figure 53). Users cited that they were highly satisfied with the site *news.google.com* when seated, less satisfied while they were walking and much less satisfied when they were walking while the add-on was activated. This was caused by the reported difficulty of the user to read the content while in motion; a difficulty that was found to be increased by the add-ons sensitivity to motion and *non-smooth* scrolling. These factors make it harder for the user to navigate the site and read the content thus decrease the user's QoE. Only 1 out of 5 users reported that they found the add-on useful.

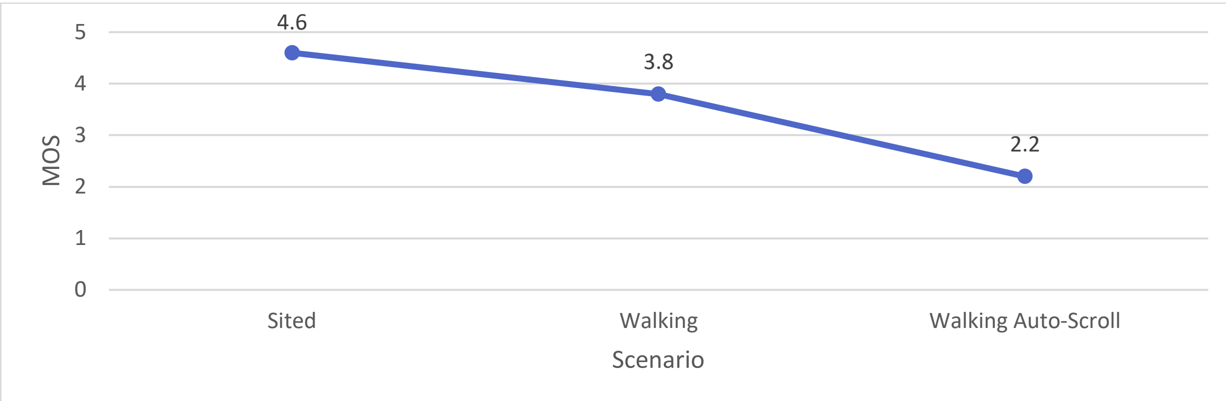


Figure 53 Usage scenarios of news.google.com

In regards to the Firefox CSS Injection add-on which hid certain parts of *amazon.co.uk* and gave the website a cleaner interface, results are inconclusive given the small sample (Figure 54). Users reported that they missed the product suggestion sections of the website and the lengthy product descriptions while others reported slightly higher satisfaction on the edited search result pages.

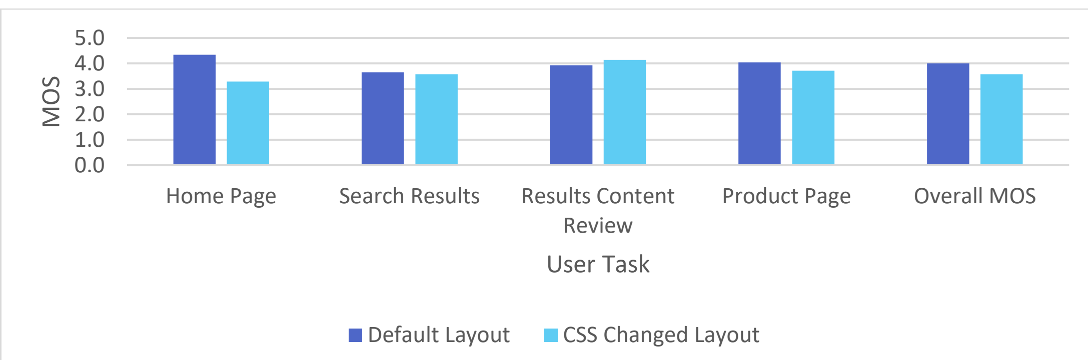


Figure 54 Layout Comparison of amazon.co.uk

4.3.8 Interesting Observations

During data analysis, some interesting findings/observations of secondary importance to this study have been discovered and are briefly outlined below.

The first observation is based on the BetterX dataset and shows that Web requests to the same domain which return an HTTP code of 302 (*URL redirection*) (Fielding *et al.*, 1999) have higher PL and CL timings than web sessions from the same domain which return an HTTP code of 200 (*OK*). Given the HTTP protocol and response status codes, it comes to no surprise that a request which is returned directly without any problems takes less time to be executed than a request which redirects to a new URL. However, it is noted in this study that the effects of HTTP 3xx redirect codes to MWQoE have not been examined in literature and should be included in future work. Figure 55 shows the average CL and PL timings per HTTP code (200 vs 302) of the top 10 domains in the BetterX dataset.

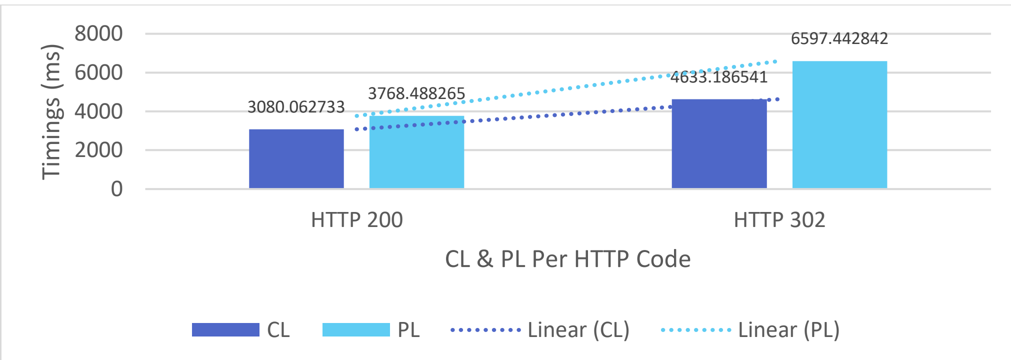


Figure 55 Top 10 BetterX Domain Timings per HTTP Code

The second observation is based on the Lab test data and shows that female participants are more sensitive to the speed decrease on the slower device (tablet) than male participants. In particular, the female participant's MOS dropped by 7% when the connection speed was capped, whereas the male participant's MOS dropped by 2%. Although this study does not intent to examine QoE per gender, however, it considers this observation as interesting and noted it for further investigation.

The third and final interesting observation is that the *user's attitude* affects MOS ratings. With each of the 31 lab participants, the administrator noted in the experiment log the *attitude* of the user based on a visual observation of the *user's composure*. The administrator rated the user as *Negative* if the user seemed frustrated at times or nervous or anxious to complete the test. On the other hand, the administrator rated the user as *Positive* if the user seemed relaxed and not frustrated with the performance of the devices. In that manner, each participant was classified in the dataset and a simple comparative analysis on the average MOS per task of each of the 2 participant types revealed an interesting result: *Positive* users reported slightly better ratings than *Negative* users in each task. This observation is reinforced by the adopted QoE definition that user context and specifically user attitude can have an impact of experience.

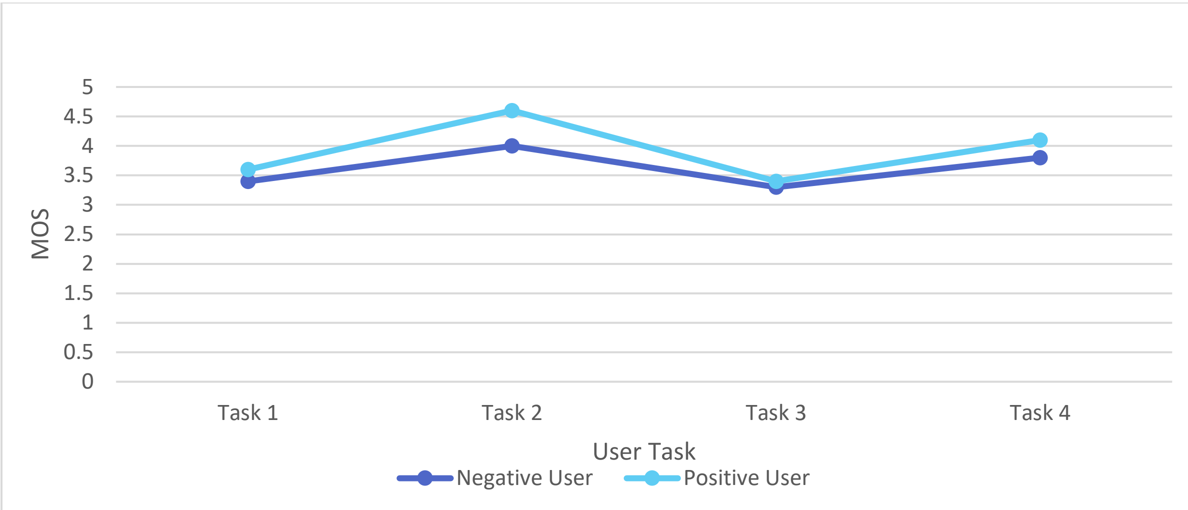


Figure 56 Per Task Reported MOS by Observed User Attitude Classification

Chapter 5 MWQoE Discussion

5.1 Overview

This study demonstrates that the non-intrusive evaluation of the mobile users Web QoE is possible by using a context-aware approach. The results of the **MWQoE Model Evaluation** show that the MWQoE model provides either the same or a more conservative characterization of the user's experience in 93% of all tests, a feature which works for the benefit of the user. Moreover, and in general across all scenarios and user profiles tested, the MWQoE metric provides a close characterization of user satisfiability (Fair vs Good). The MWQoE model is not considered generalizable and is found to be effective and useful in scenarios where the *Web Session User Profile* is either a *Worker*, *Buyer*, *Reader* or *Web Search User* (Section 3.4.3) and where the *Web Session User Intent* can be inferred using the context-aware methods outlined in *Figure 9 User Web Session Intent*. The findings of this study illustrate that the MWQoE model and metric are effective in the above-mentioned scenarios and both can be considered as a *starter framework* for future work; one which has the potential to provide even more insights and enhance the accuracy of the user's satisfiability evaluation. This is reinforced by this employed Bayesian methodology, a data-driven approach which increases the accuracy of the metric evaluation with new evidence (bigger sample size). Therefore, future work on this *starter framework* can be enhanced by new user profiles, more web session observations and a higher number of lab tests, gaining even more user insights and coming even closer to a more accurate characterization of Mobile Web QoE.

This effectiveness of MWQoE is reinforced by the validation of the importance of the MWQoE factors (Section 4.3.1) by lab users, as well as the evaluation of MWQoE against MOS (Section 4.3.6). The MWQoE metric produces a conservative characterization of the user's Web QoE in comparison with user's MOS which was expected, as warranted by the *User Bias and Favourable Response Trend (Section 4.3.5)*. The fluctuations in MOS are found to be higher on the tablet device rather than the smartphone, a finding that reveals that the device performance has a stronger effect on Web QoE than network speed for the scenarios tested; a finding which is aligned with the findings of Sebastian Egger *et al.* (2012) and Schatz and Egger (2012).

In addition, the novel approach employed by this study to infer the scenarios in which web sessions have taken place, such as the *Web Session User Intent* (Figure 9) and the *Web Session User Profile* (Section 3.4.3), is found to be a feasible approach of characterizing Mobile Web QoE via context-awareness since it considers *page latency* from the user's perspective as well as the *user task* as suggested by Strohmeier et al. (2012).

The real-world, non-intrusive approach of MWQoE is considered an improvement from previous approaches found in literature such as Cherubini and Oliver (2009), Nakhimovsky (2009) and Ickin, Wac & Fiedler (2012) since all of them are found to be impractical and not applicable for real-world settings. In Cherubini and Oliver (2009), the approach succeeded in gaining real-world data outside a controlled environment, however, it did not capture nor infer the user's context. The *Quasi Experimentation*, as proposed in Cherubini and Oliver (2009) extended this approach by capturing the user's context with micro cameras and automatic logging, however, it is considered impractical as it imposes constraints on users. The work of Ickin, Wac & Fiedler (2012) succeeded in fusing together quantitative and qualitative user readings, but is found impractical and user-intrusive as well since it uses the Experience Sampling Method (ESM). In addition, the approach of Nakhimovsky (2009) faced practical challenges with the use of diary studies, such as self-reports, and left the burden of data collection to users. Nonetheless, this study's approach is reinforced by Nakhimovsky (2009) and the finding that lab studies were found to produce different results from live user studies.

5.2 MWQoE model

The first major contribution of this study is the novel MWQoE Bayesian model, as stated above, which has been shown to be effective in evaluating user satisfiability in average and across all tested scenarios. MWQoE was found in average to produce conservative evaluations of satisfiability (*Fair*) in comparison with the average MOS (*Good*). MWQoE is based on the widely used ITU-T definition (International Telecommunication Union, 2006) and shows that the Web QoE can be measured in a non-intrusive manner by using context-awareness to infer the context in which a specific web session has taken place in regards to the user's physical context, the user's device and the web metrics. Moreover, MWQoE identifies the factors which affect Web QoE and their strength of influence. The factors which are used to derive the

MWQoE model are shown in the *Table 6 Data Attributes Selected for MWQoE* and are the building blocks of the context states outlined in Figure 30, Figure 34, Figure 36 and Figure 37. Moreover, the importance (or weight) of each of the derived context states to MWQoE is shown in *Table 27 IWWIS-MWQoE Mappings*, *Table 28 WIS-ICS-IWWIS Mappings* and *Table 29 TCS-DCS-WIS Mappings.*

MWQoE is an improvement from current Web QoE models such as Ameigeiras *et al.* (2010) and Sebastian Egger *et al.* (2012). Although both models are practical in their implementation, they express QoE as a function of latency and completely disregard the user's context. MWQoE enhances the definition of the latency factor by considering the user perceived latency. This is achieved by using web metrics together with the inferred device context and a characterization of the availability of resources to render web content on the device; as it is illustrated by the Device Availability Context State (DCS). Moreover Ameigeiras *et al.* (2010) and Sebastian Egger *et al.* (2012) present general Web QoE models and do not consider factors, such as the context of the user and the context of the mobile device, which have been shown to affect Web QoE on mobile user scenarios. In the same manner, the work of Cecchet *et al.* (2013) also provides a very narrow definition of Web QoE since it expresses it only through latency terms. Even though Cecchet *et al.* 2013) is identified in this thesis as a notable practical attempt to measure QoE outside test labs, it fails from the perspective that it uses location context not from the user's perspective, such as the location of the user at a specific time of a web session, rather than the location of the content delivery network which is used to illustrate the effect of the web server's location on website latency. MWQoE succeeds to characterize the Web QoE from a user-centered perspective considering not only web metrics and the user's location as in Ameigeiras *et al.* (2010), S Egger *et al.* (2012) and Cecchet *et al.* (2013) but including the *User Web Session Intention*, a categorization of the user using 4 user profiles, and a characterization of the device's availability to render web content (DCS).

MWQoE is an improvement of Brooks and Hestnes (2010) since MWQoE works in a non-intrusive manner and does not rely on MOS, rather than uses MOS to validate its effectiveness. Moreover, MWQoE considers both QoS metrics and user context, as in Brooks and Hestnes (2010), but does so with a far greater number of attributes and by measuring the user perceived QoS from the mobile device.

Finally, given that MWQoE infers the user's intent by analyzing web session data together with user context data, it is found to be aligned with Schatz and Egger (2012) and Strohmeier *et al.* (2014a). MWQoE implements the suggestions by Strohmeier, Pyykk & Raake (2012) that not only Page Load time needs to be considered in Web QoE but the time to render the web content as well. This is done by the Web Immediacy State (WIS) which fuses web metrics such the Page Load time with device availability metrics considering the above suggestion.

To the best of our knowledge, MWQoE is the first practical user-centered attempt to measure Web QoE on specific mobile scenarios from live user data by using context-awareness in a non-intrusive approach. The approach of using Bayesian Network with Utility Theory is adopted from Mitra, Zaslavsky & Ahlund (2015). The difference is that the work of Mitra, Zaslavsky & Ahlund (2015) does not deal with Web QoE but rather with QoE for VoIP and uses only location context and 3 QoS attributes in a simulated environment. MWQoE does not use simulations rather than real-world data to measure Web QoE and is more fine-grained since it considers 3 distinct context states (TCS, DCS, ICS) which are derived from 25 distinct data attributes. Moreover, the works of Szabo *et al.* (2016) and Hora *et al.* (2016), although notable QoE approaches are been found to completely disregard the user-centered theme of QoE. The first is partial to only network metrics, whereas the second does not address any factors in regards to the accessing device.

The MWQoE model can be further enhanced with the inclusion of parameters which have been validated by users as important for their experience, such as the *site ease of use*, the *site quality* and the *site content relevance*. Future versions of the MWQoE can potentially include characterizations of such nature, providing a more fine-grained MWQoE definition and thus a potentially more accurate MWQoE evaluation. Moreover, a context-aware characterization of these factors can be used as the guide for triggering MWQoE enhancements, such as the simple interface add-on (Section 4.2.6). The *screen size* and the *quality of the screen*, factors which have been reported as important from users, can be considered in a future MWQoE model together with the inclusion of the accelerometer data to infer the user activity.

5.3 MWQoE metric

The MWQoE metric is the Utility of the MWQoE model and is used to provide a unified MWQoE measurement in a bi-polar scale from 0 to 1 as shown in Figure 39. The MWQoE metric delivers the quantified characterization of MWQoE in 4 User Profiles (Worker, Buyer, Reader, Web User). The MWQoE metric is used to benchmark the effectiveness of the MWQoE model with user's MOS from this study's lab experiment. The novelty of the MWQoE metric is derived from the novelty of the MWQoE model. The interpretation of the MWQoE metric allows of a new in-depth analysis of the satisfiability of the mobile device user which illustrates the user's expectations and his/her reference to the network, the device and the specific online service being used.

5.4 BetterX system

BetterX actualizes the measurement of MWQoE in a non-intrusive automated manner which can be utilized by online service and content providers in industry. The BetterX system outlines the analysis, design and implementation of a novel end-to-end Mobile-to-Cloud system for the continuous MWQoE evaluation in real-world scenarios. BetterX is the first published Mobile-to-Cloud system which measures Web QoE in mobile user scenarios using a non-intrusive context-aware approach. It is unique in that it is built in a way that it uses Bayesian Networks together with a novel fusion of data sources to build User Profiles and identify User Web Session Intent in selected scenarios. BetterX constitutes a novel orchestration of Cloud components and considers optimizations at all of its fours system layers (Allayiotis and Antoniou, 2014). Moreover, in presenting the BetterX system, this study does not only regard the satisfiability of the end-user via MWQoE enhancements, but considers the efficiency of the cloud provider as well by enabling data processing optimizations in the design and implementation of the system.

5.5 Additional Contributions

In addition to the MWQoE model, the MWQoE metric and the BetterX system, 3 additional contributions are presented: The findings from testing QoE impact features in a lab experiment, the BetterX database, and the relational schema for storing HTTP metrics.

The findings from the *Impact factors* testing (Section 4.2.6*)* evaluated the effect to *user satisfiability* of 2 client-side content delivery adaptations: an auto-scrolling browser add-on which was activated via device tilting on a specific web site (*news.google.com*) while the user was in motion, and a css-injection add-on that removed certain web components of a specific e-commerce site (*amazon.co.uk*) and delivered a simpler web interface to the user. The auto-scroller add-on was tested on 5 users and the simpler interface add-on with 8 users. The auto-scrolling add-on was found to have a negative impact on user's MOS and consequently to MWQoE. It was reported that only 1 out of the 5 users tested would use it for a quick glance of the news headlines while in motion. The results from the simpler web interface add-on were inconclusive since the add-on received mixed reviews and the overall MOS was slightly less than the overall MOS of the users on the original site (3.6 vs 4.0).

Both impact factor test results are considered as minor contributions of this study since they illustrate a novel client-side approach to Web QoE enhancement never attempted. It is important to note that even though the actual enhancement of MWQoE has not been shown, the approach in attempting to enhance MWQoE on the web browser by tweaking content delivery is found to have advantages over previous attempts. The advantage of this approach, as opposed to work proposed by Zhou (2007) is that it does not introduce additional latency to the user. Zhou (2007) claims that their architecture retains the layout and rendering styles of the original document, improves legibility and provides a better interaction interface. However, the fetch-on demand scheme refers to additional latency that can be introduced between numerous fetches from user's requiring more content. Moreover, the browser add-on scripts in this study were not computationally expensive, as in Yap and Marshall (2010), Zhou (2007) and Coles, Meglan & John (2011).

This study provides the only database of Mobile Web QoE data from actual real-world live usage that others researchers can use and is made available from the project's website. The database provides 58,500 anonymous HTTP traces and metrics of 2727 observed web sessions with a total of 1,371,500 mobile device readings. Every record in the database is fully classified with user demographical information and enhanced with all the additional data analysis classifications such as the user's location type, the web domain type, the *web session user intention* and the *web session user profile*.

Finally, the Entity-Relationship (ER) diagram of the relational schema for storing anonymous HTTP traces into MySQL which has been presented in Figure 22 is considered the third minor contribution of this study. It is the first published attempt to store and represent HAR (Odvarko, Jain & Davies, 2012) into a normalized relational form. It is important to use normalized schemas for storing web data into robust relational databases such as MySQL so that the solution can be easily scaled and be applicable in real-world industry settings.

Chapter 6 Conclusion

This study presents 3 major contributions to the field of QoE studies: the **MWQoE model**, the **MWQoE metric** and the **BetterX system**. These are delivered to answer the main research question of "*How to model, measure and enhance the Web QoE of mobile device users in a practical manner applicable to real-world settings*". The MWQoE model succeeds in modeling Web QoE, the MWQoE metric succeeds in measuring (evaluating) Web QoE and the BetterX system is the Web QoE processing hub which can provide Web QoE enhancements in selected scenarios. It is shown in this study that the Web QoE of mobile users can be evaluated in a non-intrusive manner by using the sensing capabilities of modern mobile devices and the processing capabilities of Cloud to collect data about the device, the user, the web session, the network and by using context-awareness to infer the user profiles and the intention of the user for a web session. The practicality of the novel MWQoE approach, on one hand, is found in the BetterX system, which is designed as an end-to-end MCC system which evaluates Web QoE, and on the other hand, on the fact that the evaluation is non-intrusive; i.e. it does not depend on user feedback. The MWQoE model and metric are generated via data collected and processed by the BetterX system, and both of their effectiveness is established in a lab experiment where users provide feedback on simulated web browsing in a controlled environment.

The novelty of the MWQoE approach which fused together the *Web Session User Intent* with the Timings Context State (TCS) and the Device Context State (DCS) provided a practical novel Web QoE characterization. The proof-of-concept implementation of BetterX delivered by this thesis together with the complete end-to-end design and technical analysis provide a solution to both industry and academia for measuring and enhancing Web QoE for the benefit of both the online provider and the mobile user. Thus, answering the 1st research sub question: "*What are the design and technical considerations of an end-to-end system which can efficiently measure and enhance MWQoE in a non-intrusive manner?*".

The factors which were observed and derived from anonymous live user data in the Model Generation phase of the study and afterward modeled in the MWQoE Bayesian model, such as *Device Responsiveness*, *Page Immediacy* and *Internet Speed,* were found in the lab experiment to be the most important factors affecting Web QoE in the scenarios tested. Moreover, additional findings from the lab experiment suggested further potential enhancements to the MWQoE model; for example, the relationship between HTTP redirect codes to Web QoE, the impact of the quality of the screen to Web QoE, as well as the impact of the user's *attitude,* the content quality and usability of the webpage on the user experience are noted by this study and are considered for future work. Therefore, the 2nd research sub question was answered: "*What factors affect MWQoE and how do they affect it?*".

The 3rd and final research sub question "*Can we identify real-world scenarios in which we can predict and enhance MWQoE with reasonable confidence?*" was answered by the lab experiment findings in which the effectiveness of the MWQoE model and metric have been established by comparing the evaluations of the model with actual user responses. The scenarios tested were a combination of the 4 user types (*Section 3.4.3*), the inferred *Web Session User Intent* (*Section 3.4.4.5*) and the inferred *Web Session User Intent Importance* (*Section 3.4.4.5*). The average evaluation of MWQoE across these scenarios against MOS was Fair vs Good, and in 93% of all cases, the MWQoE model provided either the same or a more conservative evaluation. Therefore, it can be concluded that the MWQoE can provide user satisfiability evaluations with reasonable confidence while at the same time state that it can be further improved by the lab test findings with additional parameters and a larger data collection sample size.

The importance of this thesis is that it extends the attempts of both industry and academia in this field in a practical manner. MWQoE is user-centered as it captures metrics from the user's mobile device and regards the needs of the user as well as the context state of the network and the device. This is completely aligned with the efforts observed in industry to shift from a Page Load Web QoE perspective to a more user-centered approach. This thesis can be considered an important vehicle for evaluating customer *satisfiability*; an important consideration for online providers given that it greatly impacts their revenue. In addition, this thesis delivers contributions to academic research by extending the practical Web QoE models in mobile scenarios in providing a MWQoE construct which is fine-grained and reaches further into the conceptual QoE definitions found in literature.

References

Allayiotis, E. and Antoniou, J. (2014) 'Technical considerations towards mobile user QoE enhancement via Cloud interaction', in *Wireless Communications and Mobile Computing Conference (IWCMC), 2014 International*, pp. 936–941. Available at: http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=6906481.

Alvaro, P., Condie, T., Conway, N., Elmeleegy, K., Hellerstein, J. M. and Sears, R. (2010) 'BOOM Analytics: Exploring Data-Centric, Declarative Programming for the Cloud', in *Proceedings of the 5th European conference on Computer systems - EuroSys '10*. New York, New York, USA: ACM Press, p. 223. doi: 10.1145/1755913.1755937.

Amazon Web Services, I. (2010) *Amazon elastic compute cloud (Amazon EC2)*, *Amazon Elastic Compute Cloud (Amazon EC2)*. Available at: https://aws.amazon.com/ec2/ (Accessed: 20 May 2015).

Amazon Web Services, I. (2015a) *Amazon Mechanical Turk*, *Amazon Mechanical Turk*. Available at: https://www.mturk.com/mturk/welcome (Accessed: 20 May 2015).

Amazon Web Services, I. (2015b) *Amazon Simple Storage Service*, *Amazon Simple Storage Service*. Available at: https://aws.amazon.com/s3/ (Accessed: 20 May 2015).

Amazon Web Services, I. (2015c) *AWS Elastic Load Balancing – Cloud Network Load Balancing*, *AWS Elastic Load Balancing*. Available at: http://aws.amazon.com/elasticloadbalancing/ (Accessed: 23 May 2015).

Amazon Web Services, I. (2016) *Alexa Web Services*, *Alexa Web Services*. Available at: https://aws.amazon.com/alexa/ (Accessed: 1 March 2016).

Ameigeiras, P., Ramos-Munoz, J. J., Navarro-Ortiz, J., Mogensen, P. and Lopez-Soler, J. M. (2010) 'QoE oriented cross-layer design of a resource allocation algorithm in beyond 3G systems', *Computer Communications*. Elsevier B.V., 33(5), pp. 571–582. doi: 10.1016/j.comcom.2009.10.016.

Armbrust, M., Fox, A., Griffith, R., Joseph, A. D., Katz, R., Konwinski, A., Lee, G., Patterson, D., Rabkin, A. and Stoica, I. (2010) 'A View of Cloud Computing', *Communications of the ACM*, 53(4).

Armbrust, M., Ghodsi, A., Zaharia, M., Xin, R. S., Lian, C., Huai, Y., Liu, D., Bradley, J. K., Meng, X., Kaftan, T. and Franklin, M. J. (2015) 'Spark SQL: Relational Data Processing in Spark', in *Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data - SIGMOD '15*, pp. 1383–1394. doi: 10.1145/2723372.2742797.

Ballesteros, L. M. and Segall, Z. (2013) 'Quality of Experience and Human-computer Interaction: A Relation Overview', in *MOBILITY 2013 : The Third International Conference on Mobile Services, Resources, and Users Quality*, pp. 34–40. Available at: http://www.thinkmind.org/index.php?view=article&articleid=mobility_2013_2_30_40079.

Barber, D. (2010) *Bayesian Reasoning and Machine Learning*, *Bayesian Reasoning and Machine Learning*. Cambridge University Press. doi: 10.1017/CBO9780511804779.

Bettini, C., Brdiczka, O., Henricksen, K., Indulska, J., Nicklas, D., Ranganathan, A. and Riboni, D. (2010) 'A survey of context modelling and reasoning techniques', *Pervasive and Mobile Computing*. Elsevier B.V., 6(2), pp. 161–180. doi: 10.1016/j.pmcj.2009.06.002.

Bevan, N. (2009) 'What is The Difference Between The Purpose of Usability and User Experience Evaluation Methods?', in *Proceedings of the Workshop UXEM*, pp. 1–4.

Bischoff, A. (2016) *User-Experience-Aware System Optimisation for Mobile Systems*. University of Southampton.

Brooks, P. and Hestnes, B. (2010) 'User Measures of Quality of Experience: Why Being Objective and

Quantitative Is Important', *IEEE Network*, 24(2), pp. 8–13. doi: 10.1109/MNET.2010.5430138.

Casas, P., Seufert, M., Wamser, F., Gardlo, B., Sackl, A. and Schatz, R. (2015) 'Next to You: Monitoring Quality of Experience in Cellular Networks from the End-devices', *IEEE Transactions on Network and Service Management*, 4537(c), pp. 1–1. doi: 10.1109/TNSM.2016.2537645.

Cecchet, E., Sims, R., He, X. and Shenoy, P. (2013) 'mBenchLab: Measuring QoE of Web applications using mobile devices', in *2013 IEEE/ACM 21st International Symposium on Quality of Service (IWQoS)*. IEEE, pp. 1–10. doi: 10.1109/IWQoS.2013.6550259.

Cherubini, M. and Oliver, N. (2009) 'A Refined Experience Sampling Method to Capture Mobile User Experience', in *International Workshop of Mobile User Experience Research*, pp. 1–12.

Coles, T. R., Meglan, D. and John, N. W. (2011) 'The role of haptics in medical training simulators: A survey of the state of the art', *IEEE Transactions on Haptics*, 4(1), pp. 51–66. doi: 10.1109/TOH.2010.19.

Cui, H. and Biersack, E. (2012) *On the Relationship Between QoS and QoE for Web Sessions*.

Day, J. D. and Zimmermann, H. (1983) 'OSI REFERENCE MODEL.', in *Proceedings of the IEEE*. IEEE Computer Society, pp. 1334–1340. doi: 10.1109/PROC.1983.12775.

Dey, A. K. (2001) 'Understanding and Using Context', *Personal and Ubiquitous Computing*, 5(1), pp. 4–7. doi: 10.1007/s007790170019.

Dey, A. K. and Abowd, G. D. (1999) 'Towards a Better Understanding of Context and Context-Awareness', in *International Symposium on Handheld and Ubiquitous Computing*. Springer Berlin Heidelberg, pp. 304–307.

Dillon, A. (2002) 'Beyond Usability: Process, Outcome and Affect in human computer interactions', *Canadian Journal of Library and Information Science*, 8, pp. 57–69.

Ding, L. and Goubran, R. A. (2003) 'Assessment of effects of packet loss on speech quality in VoIP', in *Proceedings - 2nd IEEE International Workshop on Haptic, Audio and Visual Environments and their Applications, HAVE 2003*, pp. 49–54. doi: 10.1109/HAVE.2003.1244724.

Douglas, C. (2006) *JSON: The Fat-Free Alternative to XML*, *JSON: The Fat-Free Alternative to XML*. Available at: http://www.json.org/fatfree.html (Accessed: 20 June 2015).

Eaton, K. (2012) *How One Second Could Cost Amazon 1.6 Billion In Sales*, *Fast Company*. Available at: http://www.fastcompany.com/1825005/how-one-second-could-cost-amazon-16-billion-sales (Accessed: 5 September 2016).

Egger, S., Hossfeld, T., Schatz, R. and Fiedler, M. (2012) 'Waiting Times in Quality of Experience For Web Based Services', in *QoMEX 2012*, pp. 86–96.

Egger, S., Reichl, P., Hosfeld, T. and Schatz, R. (2012) '"Time is bandwidth"? Narrowing the gap between subjective time perception and Quality of Experience', in *IEEE International Conference on Communications*. IEEE Computer Society, pp. 1325–1330. doi: 10.1109/ICC.2012.6363769.

Elio, R., Hoover, J., Nikolaidis, I., Salavatipour, M., Stewart, L. and Wong, K. (2006) *About Computing Science Research Methodology*. Available at: http://webdocs.cs.ualberta.ca/~c603/readings/research-methods.pdf.

Fernando, N., Loke, S. W. and Rahayu, W. (2013) 'Mobile cloud computing: A survey', *Future Generation Computer Systems*. Elsevier B.V., 29(1), pp. 84–106. doi: 10.1016/j.future.2012.05.023.

Fiedler, M.; Hossfeld, T.; Tran-Gia, P. (2010) 'A generic quantitative relationship between quality of experience and quality of service', *Network, IEEE*, 24(2), pp. 36–41. doi: 10.1109/MNET.2010.5430142.

Fielding, R., Gettys, J., Mogul, J., Frystyk, H., Masinter, L., Leach, P. and Berners-Lee, T. (1999) *Hypertext transfer protocol--HTTP/1.1*.

Friedman, N. and Goldszmidt, M. (1996) ‘Discretizing continuous attributes while learning Bayesian networks’, *Proceedings of the 13th international conference on Machine learning (ICML)*, pp. 157–165. doi: 10.1001/archinte.159.12.1359.

Gao, Z., Chen, J., Qiu, X. and Meng, L. (2009) ‘QoE/QoS driven simulated annealing-based genetic algorithm for Web services selection’, *The Journal of China Universities of Posts and Telecommunications*. Elsevier B.V., 16(September), pp. 102–107. doi: 10.1016/S1005-8885(08)60347-7.

Gil, B. and Trezentos, P. (2011) ‘Impacts of data interchange formats on energy consumption and performance in smartphones’, in *Proceedings of the 2011 Workshop on Open Source and Design of Communication - OSDOC ’11*. Lisbon, Portugal: ACM Press, p. 1. doi: 10.1145/2016716.2016718.

Gong, Y., Yang, F., Huang, L. and Su, S. (2009) ‘Model-Based Approach to Measuring Quality of Experience’, in *2009 First International Conference on Emerging Network Intelligence*. IEEE, pp. 29–32. doi: 10.1109/EMERGING.2009.17.

Google Inc. (2013) *SPDY: An experimental protocol for a faster web*, *Spdy*. Available at: http://www.chromium.org/spdy/spdy-whitepaper (Accessed: 24 January 2014).

Google Inc. (2015) *Google Maps API*, *Google Maps API*. Available at: https://developers.google.com/maps/ (Accessed: 20 May 2010).

Google Inc. (2016) *WebP — A new image format for the Web*, *WebP*. Available at: https://developers.google.com/speed/webp/ (Accessed: 5 September 2016).

Greaves, M., Mika, P., Battle, R. and Benson, E. (2008) ‘Bridging the semantic Web and Web 2.0 with Representational State Transfer (REST)’, *Web Semantics: Science, Services and Agents on the World Wide Web*, 6(1), pp. 61–69. Available at: http://www.sciencedirect.com/science/article/pii/S1570826807000510 (Accessed: 27 December 2013).

Hasan, L. and Abuelrub, E. (2011) ‘Assessing the quality of web sites’, *Applied Computing and Informatics*. King Saud University, 9(1), pp. 11–29. doi: 10.1016/j.aci.2009.03.001.

Hora, D. N. da, Teixeira, R., van Doorselaer, K. and van Oost, K. (2016) ‘Predicting the effect of home Wi-Fi quality on Web QoE’, in *Proceedings of the 2016 workshop on - Internet-QoE ’16*. New York, New York, USA: ACM Press, pp. 13–18. doi: 10.1145/2940136.2940146.

Hosfeld, T., Fiedler, M. and Zinner, T. (2011) ‘The QoE provisioning-delivery-hysteresis and its importance for service provisioning in the Future Internet’, in *2011 7th EURO-NGI Conference on Next Generation Internet Networks*. IEEE, pp. 1–8. doi: 10.1109/NGI.2011.5985864.

Hoßfeld, T., Biedermann, S., Schatz, R., Egger, S., Fiedler, M., Hosfeld, T., Biedermann, S., Schatz, R., Platzer, A., Egger, S. and Fiedler, M. (2011) ‘The memory effect and its implications on Web QoE modeling’, *2011 23rd International Teletraffic Congress (ITC)*, pp. 103–110.

Hoßfeld, T. and Binzenhöfer, A. (2008) ‘Analysis of Skype VoIP traffic in UMTS: End-to-end QoS and QoE measurements’, *Computer Networks*, 52(3), pp. 650–666. doi: 10.1016/j.comnet.2007.10.008.

Ickin, S., Wac, K. and Fiedler, M. (2012) ‘Factors Influencing Quality of Experience of Commonly Used Mobile Applications’, *IEEE Communications Magazine*, pp. 48–56.

IETF HTTP Working Group (2015) *Hypertext Transfer Protocol Version 2 (HTTP/2)*, *Hypertext Transfer Protocol Version 2*. Available at: https://http2.github.io/ (Accessed: 5 September 2016).

International Telecommunication Union (2006) *Vocabulary and effects of transmission parameters on customer opinion of transmission quality, amendment 2 P.10/G.100*.

International Telecommunication Union (2014) *Subjective testing methodology for web browsing*.

International Telecommunication Union (2015) *The E-model: a computational model for use in transmission planning*. Available at: http://handle.itu.int/11.1002/1000/12505.

Jensen, F. V. and Nielsen, T. D. (2007) *Bayesian Networks and Decision Graphs*, *The Knowledge Engineering Review*. New York, NY: Springer New York (Information Science and Statistics). doi: 10.1007/978-0-387-68282-2.

Kothari, C. (2004) *Research methodology: methods and techniques*, *New Age International*. New Age International Publishers. Available at: http://196.29.172.66:8080/jspui/bitstream/123456789/2574/1/Research Methodology.pdf.

Kumar, K., Liu, J., Lu, Y.-H. and Bhargava, B. (2013) 'A Survey of Computation Offloading for Mobile Systems', *Mobile Networks and Applications*, 18(1), pp. 129–140. doi: 10.1007/s11036-012-0368-0.

Kwapisz, J. R., Weiss, G. M. and Moore, S. a. (2011) 'Activity recognition using cell phone accelerometers', *ACM SIGKDD Explorations Newsletter*, 12(2), p. 74. doi: 10.1145/1964897.1964918.

Lakshman, A. and Malik, P. (2010) 'Cassandra', *ACM SIGOPS Operating Systems Review*, 44(2), p. 35. doi: 10.1145/1773912.1773922.

Lanoue, S. (2015) *What is UX Design? 15 User Experience Experts Weigh In*, *User Testing Blog*. Available at: https://www.usertesting.com/blog/2015/09/16/what-is-ux-design-15-user-experience-experts-weigh-in/ (Accessed: 1 January 2016).

Lathia, N., Rachuri, K., Mascolo, C. and Roussos, G. (2013) 'Open source smartphone libraries for computational social science', in *Proceedings of the 2013 ACM conference on Pervasive and ubiquitous computing adjunct publication - UbiComp '13 Adjunct*. New York, New York, USA: ACM Press, pp. 911–920. doi: 10.1145/2494091.2497345.

Law, E. L.-C., Bevan, N., Christou, G., Springett, M. and Lárusdóttir, M. (2008) 'Meaningful Measures: Valid Useful User Experience Measurement (VUUM)', in *International Workshop on Meaningful measure: Valid Useful User Experience Measurement (VUUM)*. Toulouse, France: Institute of Research in Informatics of Toulouse (IRIT), p. 27. doi: 10.1145/2070481.2070498.

Law, E., Roto, V., Vermeeren, A. P. O. S., Kort, J. and Hassenzahl, M. (2008) 'Towards a shared definition of user experience', in *Proceeding of the twenty-sixth annual CHI conference extended abstracts on Human factors in computing systems - CHI '08*. New York, New York, USA: ACM Press, p. 2395. doi: 10.1145/1358628.1358693.

Lewis, D. D. (1998) 'Naive (Bayes) at forty: The independence assumption in information retrieval', in Nédellec, C. and Rouveirol, C. (eds) *Machine Learning: ECML-98: 10th European Conference on Machine Learning Chemnitz, Germany, April 21--23, 1998 Proceedings*. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 4–15. doi: 10.1007/BFb0026666.

Ligeza, A. (1995) *Artificial Intelligence: A Modern Approach*, *Artificial Intelligence*. New Jersey: Prentice Hall, Englewood Cliffs. doi: 10.1016/0925-2312(95)90020-9.

Lingfen Sun and Ifeachor, E. (2004) 'New models for perceived voice quality prediction and their applications in playout buffer optimization for VoIP networks', in *2004 IEEE International Conference on Communications (IEEE Cat. No.04CH37577)*. IEEE, p. 1478–1483 Vol.3. doi: 10.1109/ICC.2004.1312757.

Lingfen Sun and Ifeachor, E. C. (2002) 'Perceived speech quality prediction for voice over IP-based networks', in *2002 IEEE International Conference on Communications. Conference Proceedings. ICC 2002 (Cat. No.02CH37333)*. IEEE, pp. 2573–2577. doi: 10.1109/ICC.2002.997307.

Lingfen Sun and Ifeachor, E. C. (2006) 'Voice quality prediction models and their application in VoIP networks', *IEEE Transactions on Multimedia*, 8(4), pp. 809–820. doi: 10.1109/TMM.2006.876279.

Liotou, E., Tsolkas, D. and Passas, N. (2016) 'A roadmap on QoE metrics and models', in *2016 23rd International Conference on Telecommunications (ICT)*. IEEE, pp. 1–5. doi: 10.1109/ICT.2016.7500363.

Mannes, J. (2016) 'Multi-process Firefox brings 400-700% improvement in responsiveness', *Techcrunch*.

Available at: https://techcrunch.com/2016/09/02/multi-process-firefox-brings-400-700-improvement-in-responsiveness/?ncid=rss (Accessed: 5 September 2016).

Mantoro, T., Olowolayemo, A., Olatunji, S. O., Ayu, M. a. and Osman, A. (2011) ‘Extreme learning machine for user location prediction in mobile environment’, *International Journal of Pervasive Computing and Communications*, 7(2), pp. 162–180. doi: 10.1108/17427371111146446.

Mitra, K., Ahlund, C. and Zaslavsky, A. (2011) ‘A decision-theoretic approach for quality-of-experience measurement and prediction’, in *2011 IEEE International Conference on Multimedia and Expo*. IEEE, pp. 1–4. doi: 10.1109/ICME.2011.6012098.

Mitra, K., Zaslavsky, A. and Ahlund, C. (2015) ‘Context-Aware QoE Modelling, Measurement, and Prediction in Mobile Computing Systems’, *IEEE Transactions on Mobile Computing*, 14(5), pp. 920–936. doi: 10.1109/TMC.2013.155.

De Moor, K., Joseph, W., Ketykó, I., Tanghe, E., Deryckere, T., Martens, L. and De Marez, L. (2010) ‘Linking users’ subjective QoE evaluation to signal strength in an IEEE 802.11 b/g wireless LAN environment’, *EURASIP Journal on Wireless Communications and Networking*. Hindawi Publishing Corp., 2010, p. 6.

Mowbray, M. and Pearson, S. (2009) ‘A client-based privacy manager for cloud computing’, *Proceedings of the Fourth International ICST Conference on COMmunication System softWAre and middlewaRE - COMSWARE ’09*. New York, New York, USA: ACM Press, p. 1. doi: 10.1145/1621890.1621897.

Mozilla (2016) *Electrolysis*, *Mozilla Wiki*. Available at: https://wiki.mozilla.org/Electrolysis (Accessed: 20 May 2016).

Nakhimovsky, Y. (2009) ‘Mobile User Experience Research : Challenges , Methods & Tools’, in *Chi 2009*. Boston, MA, USA: ACM, pp. 4795–4798. doi: 10.1145/1520340.1520743.

Nancarrow, C., Brace, I. and Wright, L. T. (2001) ‘“Tell me Lies, Tell me Sweet Little Lies”: Dealing with Socially Desirable Responses in Market Research’, *The Marketing Review*, 2(1), pp. 55–69. doi: 10.1362/1469347012569427.

Naqvi, S. (2015) *Every millisecond counts in mobile latency*, *Twine Prime Blog*. Available at: http://twinprime.com/every-millisecond-counts-in-mobile-commerce/ (Accessed: 5 September 2016).

Newswire, P. (2012) *Hadoop & Big Data Market - Trends, Geographical Analysis & Worldwide Market Forecasts (2012 - 2017)*, *PR Newswire US. 10/24/2012*. Available at: http://www.reportlinker.com/p01016921/Hadoop-Big-Data-Market-Hardware-Software-Services-Hadoop-as-a-Service-Trends-Geographical-Analysis-Worldwide-Market-Forecasts-.html (Accessed: 27 December 2013).

O’Reilly Media (2016) ‘DevOps & Web Performance’, in O’Reilly Media (ed.) *O’Reilly Velocity Conference*. New York, New York, USA. Available at: http://conferences.oreilly.com/velocity/devops-web-performance-ny (Accessed: 5 September 2016).

Odvarko, J., Jain, A. and Davies, A. (2012) ‘HTTP Archive (HAR) format’, *W3C draft*.

Oracle Corporation (2012) *MySQL Documentation*, *MySQL Reference Manual*. Available at: https://dev.mysql.com/doc/.

Oshin, T. O., Poslad, S. and Ma, A. (2012) ‘Improving the Energy-Efficiency of GPS Based Location Sensing Smartphone Applications’, in *2012 IEEE 11th International Conference on Trust, Security and Privacy in Computing and Communications*. IEEE, pp. 1698–1705. doi: 10.1109/TrustCom.2012.184.

Padovitz, A., Loke, S. W., Zaslavsky, A., Burg, B. and Bartolini, C. (2005) ‘An approach to data fusion for context awareness’, in *International and Interdisciplinary Conference on Modeling and Using Context*. Springer Berlin Heidelberg, pp. 353–367. doi: 10.1.1.60.6380.

Preece, J. and Rombach, H. D. (1994) ‘A taxonomy for combining software engineering and human-computer interaction measurement approaches: towards a common framework', *International Journal of Human-Computer Studies*, 41(4), pp. 553–583. doi: 10.1006/ijhc.1994.1073.

Rehman Laghari, K. and Connelly, K. (2012) 'Toward total quality of experience: A QoE model in a communication ecosystem', *IEEE Communications Magazine*, 50(4), pp. 58–65. doi: 10.1109/MCOM.2012.6178834.

Reichl, P., Egger, S., Member, S., Schatz, R. and D'Alconzo, A. (2010) 'The Logarithmic Nature of QoE and the Role of the Weber-Fechner Law in QoE Assessment', in *Communications (ICC), 2010 IEEE International Conference*. IEEE Computer Society, pp. 1–5.

Rix, A. W., Beerends, J. G., Hollier, M. P. and Hekstra, A. P. (2001) 'Perceptual evaluation of speech quality (PESQ)-a new method for speech quality assessment of telephone networks and codecs', in *2001 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings (Cat. No.01CH37221)*. IEEE, pp. 749–752. doi: 10.1109/ICASSP.2001.941023.

Schatz, R. and Egger, S. (2011) 'Vienna surfing', in *Proceedings of the first ACM SIGCOMM workshop on Measurements up the stack - W-MUST '11*. New York, New York, USA: ACM Press, p. 19. doi: 10.1145/2018602.2018608.

Schatz, R. and Egger, S. (2012) 'On the Impact of Terminal Performance and Screen Size on QoE', in *ETSI Workshop on Selected Items on Telecommunication Quality Matters*. Vienna, Austria: ETSI.

Schatz, R., Hoßfeld, T., Janowski, L. and Egger, S. (2013) 'From packets to people: quality of experience as a new measurement challenge', in *Data traffic monitoring and analysis*. Springer Berlin Heidelberg, pp. 219–263.

Schwarz, H., Marpe, D. and Wiegand, T. (2007) 'Overview of the Scalable Video Coding Extension of the H.264/AVC Standard', *IEEE Transactions on Circuits and Systems for Video Technology*, 17(9), pp. 1103–1120. doi: 10.1109/TCSVT.2007.905532.

Souders, S. (2009) *Velocity and the Bottom Line*, *O'Reilly Radar*. Available at: http://radar.oreilly.com/2009/07/velocity-making-your-site-fast.html (Accessed: 5 September 2016).

StatCounter (2016) *StatCounter Global Stats*, *Global Stats*. Available at: http://gs.statcounter.com/#desktop+mobile+tablet-browser-ww-monthly-201503-201603 (Accessed: 18 April 2016).

Statista Inc (2016) *Number of mobile phone users worldwide from 2013 to 2019 (in billions)*. Available at: http://www.statista.com/statistics/274774/forecast-of-mobile-phone-users-worldwide/.

Strohmeier, D., Egger, S., Raake, A., Hoßfeld, T. and Schatz, R. (2014) *Quality of Experience*, *T-Labs Series in Telecommunication Services*. Edited by S. Möller and A. Raake. Cham: Springer International Publishing (T-Labs Series in Telecommunication Services). doi: 10.1007/978-3-319-02681-7.

Strohmeier, D., Pyykk, S. J. and Raake, A. (2012) 'Toward task-dependent evaluation of Web-QoE : Free exploration vs . " Who Ate What ?"', in *GC'12 Workshop: Quality of Experience for Multimedia Communications Toward*, pp. 1309–1313.

Stuedi, P. and Terry, D. (2010) 'WhereStore : Location-based Data Storage for Mobile Devices Interacting with the Cloud'.

Szabo, G., Racz, S., Malomsoky, S. and Bolle, A. (2016) 'When To Take What: QoE-Aware Resource Redistribution among Web Browsing Users and the Potential of Prioritizing QoE Sensitive Content', in *Communications (ICC), 2016 IEEE International Conference*. IEEE Computer Society, pp. 1–6.

Ting-Yi Lin, Ting-An Lin, Cheng-Hsin Hsu and Chung-Ta King (2013) 'Context-aware decision engine for mobile cloud offloading', in *2013 IEEE Wireless Communications and Networking Conference Workshops (WCNCW)*. National Tsing Hua University, Department of Computer Science, Hsinchu, Taiwan, 34881: IEEE, pp. 111–116. doi: 10.1109/WCNCW.2013.6533324.

Tullis, T. and Albert, B. (2013) *Measuring the User Experience*. 2nd edn. Elsevier Inc.

W3C (2015) *A Primer for Web Performance Timing APIs*. Available at: https://w3c.github.io/perf-timing-primer/ (Accessed: 5 September 2016).

Yap, K. M. and Marshall, A. (2010) 'Investigating Quality of Service Issues for Distributed Haptic Virtual Environments in IP Networks', in *2010 Sixth International Conference on Intelligent Environments*. IEEE, pp. 237–242. doi: 10.1109/IE.2010.50.

Zadeh, M. H., Wang, D. and Kubica, E. (2007) 'Human Factors for Designing a Haptic Interface for Interaction with a Virtual Environment', in *2007 IEEE International Workshop on Haptic, Audio and Visual Environments and Games*. IEEE, pp. 21–26. doi: 10.1109/HAVE.2007.4371580.

Zhou, R. (2007) 'How to quantify user experience: Fuzzy comprehensive evaluation model based on summative usability testing', *2nd International Conference on Usability and Internationalization, UI-HCII 2007*, pp. 564–573. Available at: http://www.scopus.com/inward/record.url?eid=2-s2.0-38149023226&partnerID=40&md5=f42a50e981f632c1746ffbc339def27a.

Zhuang, Z., Kim, K.-H. and Singh, J. P. (2010) 'Improving energy efficiency of location sensing on smartphones', in *Proceedings of the 8th international conference on Mobile systems, applications, and services - MobiSys '10*. New York, New York, USA: ACM Press, p. 315. doi: 10.1145/1814433.1814464.

Zoonekynd, V. (2007) 'Statistics with R', *Bernoulli*, pp. 1–13. doi: 10.1021/jp207563h.

Appendices

1. End User License Agreement (EULA)
2. BetterX Campus Campaign Poster
3. BetterX Campus Leaflet Handout
4. Google Play Store Page for BetterX Android
5. BetterX Facebook Page
6. BetterX Back-End Web Interface
7. Questionnaire for Worker
8. Questionnaire for Buyer
9. Questionnaire for Reader
10. Questionnaire for Web User (Default)
11. Lab Study Consent Form
12. List of Acronyms

Privacy Policy

Published on http://www.betterx.org/privacy.html

The current document outlines the Privacy Policy for the BetterX website (www.betterx.org) and the BetterX Android app. It describes how we collect, use and protect your information. By downloading and installing BetterX Android you accept the terms and conditions in this Privacy Policy and the End-User License Agreement

By using BetterX you understand and agree that we are providing a system which automatically collects different types of anonymous information from your mobile device. All information is encrypted and sent using a secure channel to our data center for processing. The information is accessible only to the team members that are listed on our website. At any given time, you can request a copy of all of your data via the Android app.

The information is processed using various data processing methodologies such as statistical analysis and machine learning. A profile will be created for every user which will include his/her usage patterns and the scenarios in which the experience increased, decreased or stayed the same. These profiles will be available for your review from the app after a brief period of time. Once the data we've collected for your profile is sufficient and your own user model is validated, then the application will automatically attempt to adjust different aspects of your mobile device in an attempt to improve the quality of your experience. Changes may include device resource allocations, features that affect content delivery or the content itself. These changes will be attempted automatically - an indicator will be shown on your status panel - and their efficiency will be measured.

User / Human Subject Definition

"Human subjects" means a living individual about whom an investigator (professional or student) conducting research obtains data through intervention or interaction with the individual.

Information we collect

The term "information", refers to all the data which is collected via the BetterX Android App. This data is the result of the interaction of the user (participant) with the device after giving consent. Data does not refer to any other types of information which may be collected or held by a third party.

Information we collect from your mobile device:

- Unique Device Identifier. This ID is generated during the installation of the app on your mobile device. This ID is anonymous and no Name or other user private information is attached to it.
- Device Details. Different types of information which relates to your mobile device such as model number, device features, sensors available, current Android version, number of installed apps, etc.
- Network Details. Different network metrics such as connectivity details (wifi, 3G, 4G), carrier details, bandwidth, network congestion levels, etc.
- Context Details. Temperature at different times of the day, Light levels at different time of the day and location information (GPS data, data from Network location services).
- Website Metrics. Different metrics for websites visited such as domain name, time on website, interaction with the website, type of context accessed, etc.

Third-Party Links

- Mobile Application: No Third party links such as Facebook or Twitter are linked on the mobile application.
- Third Party Data Storage & Processing Center: We use third-party cloud-based services for offering backend functionalities and because of that, any data we collect might be stored and processed in servers that are situated outside your Country of residence.

Cookies and similar technologies

When you visit the website, we may use cookies to collect information about how you use the site and provide features to you.

How we use your information

We use the information we receive to: Understand what constitutes Quality of Experience for Mobile Device Users and provide the mechanisms which will improve it.

- Understand the factors and the scenarios in which the quality of experience is affected for different types of mobile device users.
- Build models for each user which will accurately reflect the perceived user experience at any given point in time.

Sharing of your information

- We will not rent, sell or share your data to any third parties.
- Our findings and aggregated information about your data will be published in academic publications and be made available to other researchers.

Anonymity

We define user anonymity as follows: The user is not requested to provide any identity information (e.g. legal name, date of birth, birthplace, phone number, social media profiles, etc.) nowhere within the website and application. We provide appropriate functionality in both the website and the application for the users to protect their identity.

User Requests

At any time, the user has every right to request a copy of his/her data via a message from the Android app (anonymous communication). We'll make sure that you receive a copy of your data within a maximum of 2 weeks.

At any time, the user can uninstall the application and request deletion of all of his/her data from our data centers. The user can send us a withdraw / deletion request via a message from the Android app. We'll make sure that all of the user's data has been deleted and send back a confirmation notice.

Changes to our privacy

We may modify or update this Privacy Policy from time to time. In the case we do, we'll make sure that you are notified via a push notification on your mobile device Your continued use of BetterX after any modification to this Privacy Policy will constitute your acceptance of such modification.

How to contact us

If you have any questions about this privacy policy or general questions / issues about BetterX Android, please contact us at hello@betterx.org

Unexpected Installation Issues

In the unlikely event that users/participants experience unexpected outcomes from the usage of the BetterX Android app they can either uninstall the app and/or contact us to provide support either via email (hello@betterx.org) or by using the app's message functionality which is anonymous.

However, given the non-intrusive nature of this project (the app runs on the background), there are no possible adverse effects to participants other than the possibility of their mobile device performing poorly due to a bug in our app or a faulty installation. All issues can immediately be resolved by simply uninstalling the app and withdrawing from the project.

Complaints

If you want to file a complaint about data privacy / data anonymity / data handling protocols, please email UCLan's Ethics officer at OfficerforEthics@uclan.ac.uk

End User License Agreement (EULA)

Published on http://www.betterx.org/eula.pdf

This End-User License Agreement (EULA) is a legal agreement between you and the mentioned author (University of Central Lancashire Cyprus) of this Software for the software product identified above, which includes computer software and may include associated media and “online” or electronic documentation (“SOFTWARE PRODUCT”).

By installing, copying, or otherwise using the SOFTWARE PRODUCT, you agree to be bounded by the terms of this EULA. If you do not agree to the terms of this EULA, do not install or use the SOFTWARE PRODUCT.

SOFTWARE PRODUCT LICENSE

1) BetterX Android is being distributed as Freeware for personal use and educational purpose. You are NOT allowed to make a charge for distributing this Software, whether as a stand-alone product, or as part of a compilation or anthology, nor to use it for supporting your business or customers. It may be distributed freely on any website or through any other distribution mechanism, as long as no part of it is changed in any way.

GRANT OF LICENSE

This EULA grants you the following rights: Installation and Use. You may install and use an unlimited number of copies of the SOFTWARE PRODUCT.

Reproduction and Distribution.

You may reproduce and distribute an unlimited number of copies of the SOFTWARE PRODUCT; provided that each copy shall be a true and complete copy, including all copyright and trademark notices, and shall be accompanied by a copy of this EULA.

Copies of the SOFTWARE PRODUCT may be distributed as a standalone product or included with your own product as long as The SOFTWARE PRODUCT is not sold or included in a product or package that intends to receive benefits through the inclusion of the SOFTWARE PRODUCT.

The SOFTWARE PRODUCT may be included in any free or non-profit packages or products.

DESCRIPTION OF OTHER RIGHTS AND LIMITATIONS

Limitations on Reverse Engineering, Decompilation, Disassembly and change (add, delete or modify) the resources in the compiled the assembly. You may not reverse engineer, decompile, or disassemble the SOFTWARE PRODUCT, except and only to the extent that such activity is expressly permitted by applicable law notwithstanding this limitation.

Update and Maintenance

BetterX Android upgrades are FREE of charge.

Separation of Components

The SOFTWARE PRODUCT is licensed as a single product. Its component parts may not be separated for use on more than one computer.

Software Transfer

You may permanently transfer all of your rights under this EULA, provided the recipient agrees to the terms of this EULA.

Termination

Without prejudice to any other rights, the Author of this Software may terminate this EULA if you fail to comply with the terms and conditions of this EULA. In such event, you must destroy all copies of the SOFTWARE PRODUCT and all of its component parts.

COPYRIGHT

All title and copyrights in and to the SOFTWARE PRODUCT (including but not limited to any images, photographs, clipart, libraries, and examples incorporated into the SOFTWARE PRODUCT), the accompanying printed materials, and any copies of the SOFTWARE PRODUCT are owned by the Author of this Software. The SOFTWARE PRODUCT is protected by copyright laws and international treaty provisions. Therefore, you must treat the SOFTWARE PRODUCT like any other copyrighted material. The licensed users or licensed company can use all functions, example, templates, clipart, libraries and symbols in the SOFTWARE PRODUCT to create new diagrams and distribute the diagrams.

LIMITED WARRANTY

NO WARRANTIES.

The Author of this Software expressly disclaims any warranty for the SOFTWARE PRODUCT. The SOFTWARE PRODUCT and any related documentation is provided “as is” without warranty of any kind, either express or implied, including, without limitation, the implied warranties or merchantability, fitness for a particular purpose, or non-infringement. The entire risk arising out of use or performance of the SOFTWARE PRODUCT remains with you.

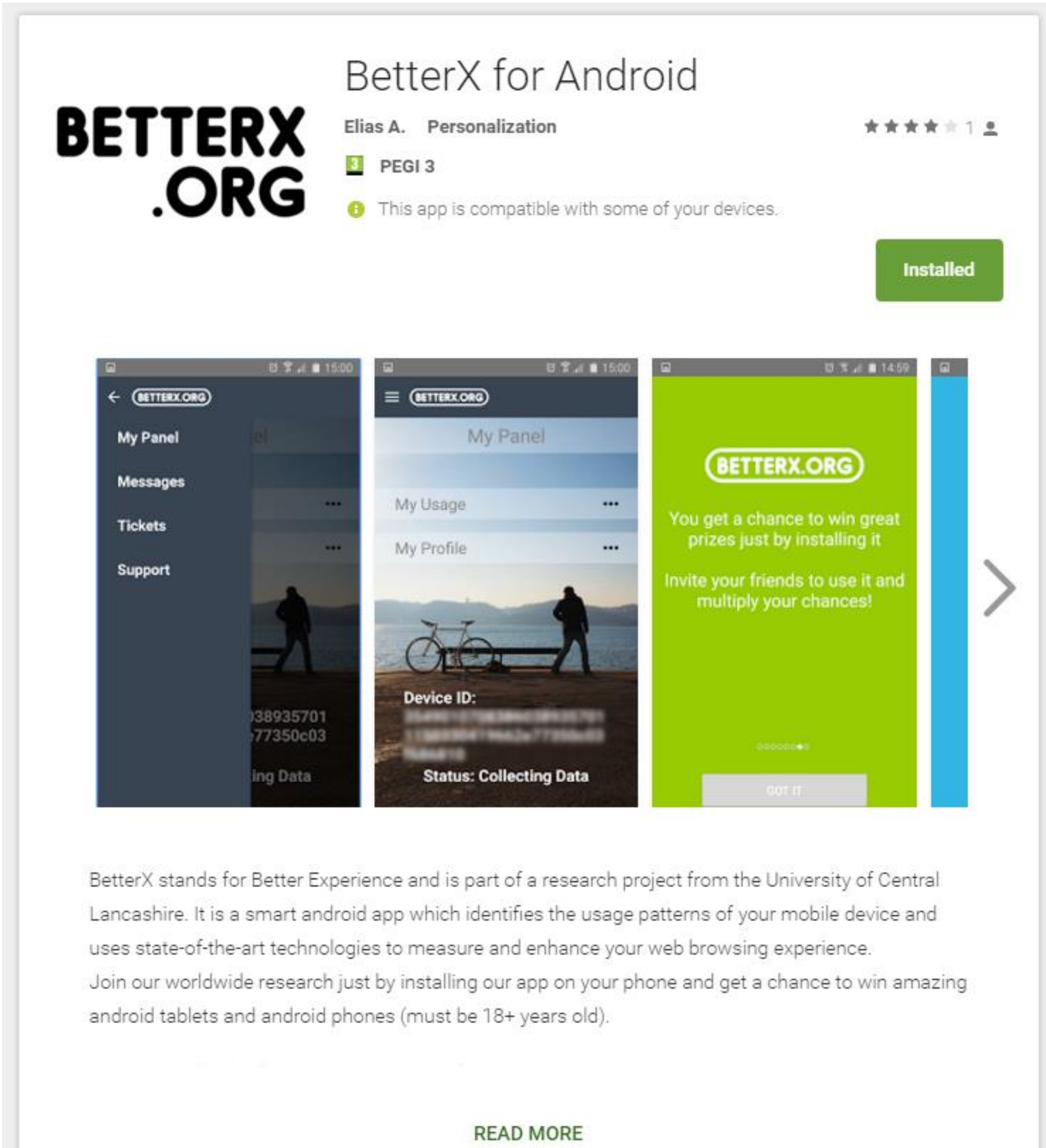
NO LIABILITY FOR DAMAGES

In no event shall the author of this Software be liable for any special, consequential, incidental or indirect damages whatsoever (including, without limitation, damages for loss of business profits, business interruption, loss of business information, or any other pecuniary loss) arising out of the use of or inability to use this product, even if the Author of this Software is aware of the possibility of such damages and known defects.

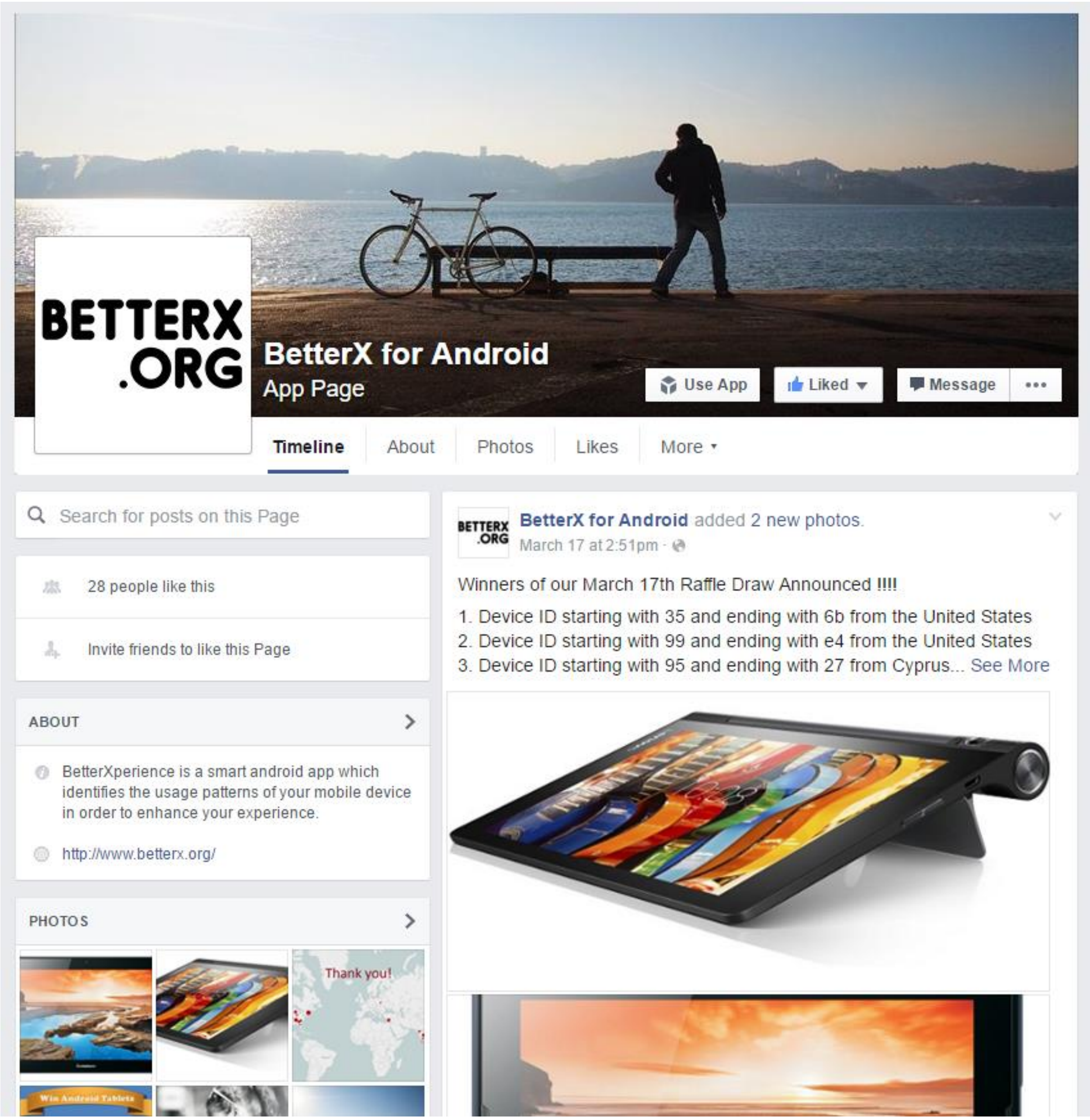
BetterX Campus Campaign Poster

BetterX Campus Leaflet Handout

Google Play Store Page for BetterX Android



BetterX Facebook Page



BetterX Back-End Web Interface

BETTERX.ORG

Files: 1442

Show 10 entries

Search:

No	Zip	Name	Status	ZipFile	ZipFileNo	ZipFileTotal
703	3521660508846	/home/ubuntu/te	OK	1	1	159
704	3521660508846	/home/ubuntu/te	OK	2	1	159
705	3521660508846	/home/ubuntu/te	OK	3	1	159
706	3521660508846	/home/ubuntu/te	256	4	1	159
707	3521660508846	/home/ubuntu/te	OK	1	2	159
708	3521660508846	/home/ubuntu/te	256	2	2	159
709	3521660508846	/home/ubuntu/te	256	1	3	159
710	3521660508846	/home/ubuntu/te	OK	2	3	159
711	3521750767225	/home/ubuntu/te	OK	1	4	159
712	3521750767225	/home/ubuntu/te	OK	2	4	159

Showing 1 to 10 of 1,442 entries

Previous 1 2 3 4 5 … 145 Next

Questionnaire for Worker

Gender: ☐ Male ☐ Female

Age:

Occupation:

How frequently do you use your smartphone?

Rarely	Not so often	Often	Very Often	I'm addicted
☐	☐	☐	☐	☐

How frequently do you use the web on your phone?

Rarely	Not so often	Often	Very Often	I'm addicted
☐	☐	☐	☐	☐

Test Scenario

You are an online worker / freelance professional working from your home office using mainly a smartphone and a tablet. Your job is to source work online and perform all of your work from the web browser.

Question 1 In your opinion as an online worker, how important are the following factors in the **Quality of Experience (or Satisfiability)** of the websites you are using:

	Not Important (1)	**Somewhat Important (2)**	**Moderately Important (3)**	**Important (4)**	**Very Important (5)**
The immediacy of the web page	☐	☐	☐	☐	☐
The speed of the mobile network	☐	☐	☐	☐	☐
The responsiveness of the device	☐	☐	☐	☐	☐
The quality of the screen	☐	☐	☐	☐	☐
The size of the screen	☐	☐	☐	☐	☐
The quality of the site you use for work	☐	☐	☐	☐	☐
The relevance of the work available on the site in regards to your needs/expectations	☐	☐	☐	☐	☐
The ease of use of the site	☐	☐	☐	☐	☐

Task 1 Use the Firefox browser on the Tablet that is provided to you and after the completion of each task outlined below **rank your experience (from 1 to 5)**.

		Poor (1)	**Fair (2)**	**Good (3)**	**Very Good (4)**	**Excellent (5)**
1	Go to **mturk.com**	☐	☐	☐	☐	☐
2	Click on the link named **View them now** to view all the available work on the site.	☐	☐	☐	☐	☐
3	Read all of the work titles on the first page and click on one **View a HIT in this group**	☐	☐	☐	☐	☐
4	Review the contents of the page.	☐	☐	☐	☐	☐
Question How would you rate your overall experience in performing this task?		☐	☐	☐	☐	☐

Question Justify the above rating (provide reasons for it)

Task 2 Use the Firefox browser on the Smartphone that is provided to you and after the completion of each task outlined below **rank your experience (from 1 to 5).**

		Poor (1)	**Fair (2)**	**Good (3)**	**Very Good (4)**	**Excellent (5)**
1	Go to **mturk.com**	☐	☐	☐	☐	☐
2	Click on the link named **View them now** to view all the available work on the site.	☐	☐	☐	☐	☐
3	Read all of the work titles on the first page and click on one **View a HIT in this group**	☐	☐	☐	☐	☐
4	Review the contents of the page.	☐	☐	☐	☐	☐
Question How would you rate your overall experience in performing this task?		☐	☐	☐	☐	☐

Question Justify the above rating (provide reasons for it)

Wait for Admin before proceeding ☺

Task 3 Use the Firefox browser on the Tablet that is provided to you and after the completion of each task outlined below **rank your experience (from 1 to 5).**

		Poor (1)	**Fair (2)**	**Good (3)**	**Very Good (4)**	**Excellent (5)**
1	Go to **mturk.com**	☐	☐	☐	☐	☐
2	Click on the link named **View them now** to view all the available work on the site.	☐	☐	☐	☐	☐
3	Read all of the work titles on the first page and click on one **View a HIT in this group**	☐	☐	☐	☐	☐
4	Review the contents of the page.	☐	☐	☐	☐	☐
Question How would you rate your overall experience in performing this task?		☐	☐	☐	☐	☐

Question Justify the above rating (provide reasons for it)

Task 4 Use the Firefox browser on the Smartphone that is provided to you and after the completion of each task outlined below **rank your experience (from 1 to 5).**

		Poor (1)	**Fair (2)**	**Good (3)**	**Very Good (4)**	**Excellent (5)**
1	Go to **mturk.com**	☐	☐	☐	☐	☐
2	Click on the link named **View them now** to view all the available work on the site.	☐	☐	☐	☐	☐
3	Read all of the work titles on the first page and click on one **View a HIT in this group**	☐	☐	☐	☐	☐

4 Review the contents of the page.	☐	☐	☐	☐	☐
Question How would you rate your overall experience in performing this task?	☐	☐	☐	☐	☐

Question Justify the above rating (provide reasons for it)	

Questionnaire for Buyer

Gender: ☐ Male ☐ Female

Age: [ ]

Occupation: [ ]

How frequently do you use your smartphone?

Rarely	Not so often	Often	Very Often	I'm addicted
☐	☐	☐	☐	☐

How frequently do you use the web on your phone?

Rarely	Not so often	Often	Very Often	I'm addicted
☐	☐	☐	☐	☐

Test Scenario

You are a web buyer. A web buyer is a person who uses the web mostly to purchase products or services online.

Question 1 In your opinion as a web buyer, how important are the following factors in the **Quality of Experience (or Satisfiability)** of the websites you are using:

	Not Important (1)	**Somewhat Important (2)**	**Moderately Important (3)**	**Important (4)**	**Very Important (5)**
The immediacy of the web page	☐	☐	☐	☐	☐
The speed of the mobile network	☐	☐	☐	☐	☐
The responsiveness of the device	☐	☐	☐	☐	☐
The quality of the screen	☐	☐	☐	☐	☐
The size of the screen	☐	☐	☐	☐	☐
The quality of the site that I use for purchases	☐	☐	☐	☐	☐
The relevance of the products available on the site in regards to my needs/expectations	☐	☐	☐	☐	☐
The ease of use of the site	☐	☐	☐	☐	☐

Task 1 Use the Firefox browser on the Tablet that is provided to you and after the completion of each task outlined below **rank your experience (from 1 to 5).**

		Poor (1)	Fair (2)	Good (3)	Very Good (4)	Excellent (5)
1	Go to **amazon.co.uk**	☐	☐	☐	☐	☐
2	Search for a **white polo short** and review the results.	☐	☐	☐	☐	☐
3	Locate an item in the first page of the results you wish to review and click on it.	☐	☐	☐	☐	☐
4	Review the product page of the item	☐	☐	☐	☐	☐
Question How would you rate your overall experience in performing this task?		☐	☐	☐	☐	☐

Question Justify the above rating (provide reasons for it)

Task 2 Use the Firefox browser on the Smartphone that is provided to you and after the completion of each task outlined below **rank your experience (from 1 to 5).**

		Poor (1)	Fair (2)	Good (3)	Very Good (4)	Excellent (5)
1	Go to **amazon.co.uk**	☐	☐	☐	☐	☐
2	Search for a **white polo short** and review the results.	☐	☐	☐	☐	☐
3	Locate an item in the first page of the results you wish to review and click on it.	☐	☐	☐	☐	☐
4	Review the product page of the item	☐	☐	☐	☐	☐
Question How would you rate your overall experience in performing this task?		☐	☐	☐	☐	☐

Question Justify the above rating (provide reasons for it)	

Wait for Admin before proceeding ☺

Task 3 Use the Firefox browser on the Tablet that is provided to you and after the completion of each task outlined below **rank your experience (from 1 to 5).**

		Poor (1)	**Fair (2)**	**Good (3)**	**Very Good (4)**	**Excellent (5)**
1	Go to **amazon.co.uk**	☐	☐	☐	☐	☐
2	Search for a **white polo short** and review the results.	☐	☐	☐	☐	☐
3	Locate an item in the first page of the results you wish to review and click on it.	☐	☐	☐	☐	☐
4	Review the product page of the item	☐	☐	☐	☐	☐
Question How would you rate your overall experience in performing this task?		☐	☐	☐	☐	☐

Question Justify the above rating (provide reasons for it)	

Task 4 Use the Firefox browser on the Smartphone that is provided to you and after the completion of each task outlined below **rank your experience (from 1 to 5).**

		Poor (1)	**Fair (2)**	**Good (3)**	**Very Good (4)**	**Excellent (5)**
1	Go to **amazon.co.uk**	☐	☐	☐	☐	☐
2	Search for a **white polo short** and review the results.	☐	☐	☐	☐	☐
3	Locate an item in the first page of the results you wish to review and click on it.	☐	☐	☐	☐	☐

4	Review the product page of the item	☐	☐	☐	☐	☐
Question How would you rate your overall experience in performing this task?		☐	☐	☐	☐	☐

Question Justify the above rating (provide reasons for it)

Wait for Admin before proceeding ☺

Task 5 Firefox browser on the Tablet that is provided to you and after the completion of each task outlined below **rank your experience (from 1 to 5).**

		Poor (1)	**Fair** (2)	**Good** (3)	**Very Good** (4)	**Excellent** (5)
1	Go to **amazon.co.uk**	☐	☐	☐	☐	☐
2	Search for a **white polo short** and review the results.	☐	☐	☐	☐	☐
3	Locate an item in the first page of the results you wish to review and click on it.	☐	☐	☐	☐	☐
4	Review the product page of the item	☐	☐	☐	☐	☐
Question How would you rate your overall experience in performing this task?		☐	☐	☐	☐	☐

Question Justify the above rating (provide reasons for it)

Task 6 Use the Firefox browser on the Smartphone that is provided to you and after the completion of each task **outlined below rank your experience (from 1 to 5).**

		Poor (1)	**Fair** (2)	**Good** (3)	**Very Good** (4)	**Excellent** (5)
1	Go to **amazon.co.uk**	☐	☐	☐	☐	☐
2	Search for a **white polo short** and review the results.	☐	☐	☐	☐	☐

3	Locate an item in the first page of the results you wish to review and click on it.	☐	☐	☐	☐	☐
4	Review the product page of the item	☐	☐	☐	☐	☐
Question How would you rate your overall experience in performing this task?		☐	☐	☐	☐	☐

Question Justify the above rating (provide reasons for it)

Questionnaire for Reader

Gender: ☐ Male ☐ Female

Age:

Occupation:

How frequently do you use your smartphone?

Rarely	Not so often	Often	Very Often	I'm addicted
☐	☐	☐	☐	☐

How frequently do you use the web on your phone?

Rarely	Not so often	Often	Very Often	I'm addicted
☐	☐	☐	☐	☐

Test Scenario

You are a web reader. A web reader is a user who uses the web mostly to read the news online.

Question 1 In your opinion as a web reader, how important are the following factors in the **Quality of Experience (or Satisfiability)** of the websites you are using:

	Not Important (1)	**Somewhat Important (2)**	**Moderately Important (3)**	**Important (4)**	**Very Important (5)**
The immediacy of the web page	☐	☐	☐	☐	☐
The speed of the mobile network	☐	☐	☐	☐	☐
The responsiveness of the device	☐	☐	☐	☐	☐
The quality of the screen	☐	☐	☐	☐	☐
The size of the screen	☐	☐	☐	☐	☐
The quality of the news site	☐	☐	☐	☐	☐
The relevance of the news available on the site in regards to my needs/expectations	☐	☐	☐	☐	☐
The ease of use of the site	☐	☐	☐	☐	☐

Task 1 Use the Firefox browser on the Tablet that is provided to you and after the completion of each task outlined below **rank your experience (from 1 to 5).**

		Poor (1)	Fair (2)	Good (3)	Very Good (4)	Excellent (5)
1	Go to **news.google.com**	☐	☐	☐	☐	☐
2	Locate the **World** section, click on it and review the headlines listed on the first page.	☐	☐	☐	☐	☐
3	Pick the one that interests you the most and click on it.	☐	☐	☐	☐	☐
4	Review the contents of the news article.	☐	☐	☐	☐	☐
Question How would you rate your overall experience in performing this task?		☐	☐	☐	☐	☐

Question Justify the above rating (provide reasons for it)

Task 2 Use the Firefox browser on the Smartphone that is provided to you and after the completion of each task outlined below **rank your experience (from 1 to 5).**

		Poor (1)	Fair (2)	Good (3)	Very Good (4)	Excellent (5)
1	Go to **news.google.com**	☐	☐	☐	☐	☐
2	Locate the **World** section, click on it and review the headlines listed on the first page.	☐	☐	☐	☐	☐
3	Pick the one that interests you the most and click on it.	☐	☐	☐	☐	☐
4	Review the contents of the news article.	☐	☐	☐	☐	☐
Question How would you rate your overall experience in performing this task?		☐	☐	☐	☐	☐

Question Justify the above rating (provide reasons for it)

Wait for Admin before proceeding ☺

Task 3 Use the Firefox browser on the Tablet that is provided to you and after the completion of each task outlined below **rank your experience (from 1 to 5).**

		Poor (1)	Fair (2)	Good (3)	Very Good (4)	Excellent (5)
1	Go to **news.google.com**	☐	☐	☐	☐	☐
2	Locate the **World** section, click on it and review the headlines listed on the first page.	☐	☐	☐	☐	☐
3	Pick the one that interests you the most and click on it.	☐	☐	☐	☐	☐
4	Review the contents of the news article.	☐	☐	☐	☐	☐
Question How would you rate your overall experience in performing this task?		☐	☐	☐	☐	☐

Question Justify the above rating (provide reasons for it)

Task 4 Use the Firefox browser on the Smartphone that is provided to you and after the completion of each task outlined below **rank your experience (from 1 to 5).**

		Poor (1)	Fair (2)	Good (3)	Very Good (4)	Excellent (5)
1	Go to **news.google.com**	☐	☐	☐	☐	☐
2	Locate the **World** section, click on it and review the headlines listed on the first page.	☐	☐	☐	☐	☐
3	Pick the one that interests you the most and click on it.	☐	☐	☐	☐	☐

4 Review the contents of the news article.	☐	☐	☐	☐	☐
Question How would you rate your overall experience in performing this task?	☐	☐	☐	☐	☐

Question Justify the above rating (provide reasons for it)

Wait for Admin before proceeding ☺

Task 5 Use the Firefox browser on the Smartphone that is provided to you, and **while walking**, attempt to read the news on the webpage **news.google.com.**

	Poor **(1)**	**Fair** **(2)**	**Good** **(3)**	**Very Good** **(4)**	**Excellent** **(5)**
Question How would you rate your overall experience in performing this task?	☐	☐	☐	☐	☐

Question Justify the above rating (provide reasons for it)

Wait for Admin before proceeding ☺

A new feature on the phone has been enabled and by tilting the phone forward the page scrolls down automatically; tilting the phone backwards scrolls up the page.

Task 6 Use the Firefox browser on the Smartphone that is provided to you, and **while walking**, attempt to read the news on the webpage **news.google.com.**

	Poor **(1)**	**Fair** **(2)**	**Good** **(3)**	**Very Good** **(4)**	**Excellent** **(5)**
Question How would you rate your overall experience in performing this task?	☐	☐	☐	☐	☐

Question Justify the above rating (provide reasons for it)

Questionnaire for Web User (Default)

Gender: ☐ Male ☐ Female

Age:

Occupation:

How frequently do you use your smartphone?

Rarely	Not so often	Often	Very Often	I'm addicted
☐	☐	☐	☐	☐

How frequently do you use the web on your phone?

Rarely	Not so often	Often	Very Often	I'm addicted
☐	☐	☐	☐	☐

Test Scenario

You are a web user. A web user is a user of the web who uses it mostly to research topics and answer specific questions.

Question 1 In your opinion as a web user, how important are the following factors in the **Quality of Experience (or Satisfiability)** of the websites you are using:

	Not Important (1)	**Somewhat Important (2)**	**Moderately Important (3)**	**Important (4)**	**Very Important (5)**
The immediacy of the web page	☐	☐	☐	☐	☐
The speed of the mobile network	☐	☐	☐	☐	☐
The responsiveness of the device	☐	☐	☐	☐	☐
The quality of the screen	☐	☐	☐	☐	☐
The size of the screen	☐	☐	☐	☐	☐
The quality of the site that I use to search	☐	☐	☐	☐	☐
The relevance of the information available on the site in regards to my inquiry.	☐	☐	☐	☐	☐
The ease of use of the site	☐	☐	☐	☐	☐

Task 1 Use the Firefox browser on the Tablet that is provided to you and after the completion of each task outlined below **rank your experience (from 1 to 5).**

		Poor (1)	Fair (2)	Good (3)	Very Good (4)	Excellent (5)
1	Go to **www.google.com**	☐	☐	☐	☐	☐
2	Search for **Toji Japan** and review the results of the first page	☐	☐	☐	☐	☐
3	Click on a page you want to review	☐	☐	☐	☐	☐
4	Review the contents of the page	☐	☐	☐	☐	☐
Question How would you rate your overall experience in performing this task?		☐	☐	☐	☐	☐

Question Justify the above rating (provide reasons for it)

Task 2 Use the Firefox browser on the Smartphone that is provided to you and after the completion of each task outlined below **rank your experience (from 1 to 5).**

		Poor (1)	Fair (2)	Good (3)	Very Good (4)	Excellent (5)
1	Go to **www.google.com**	☐	☐	☐	☐	☐
2	Search for **Toji Japan** and review the results of the first page	☐	☐	☐	☐	☐
3	Click on a page you want to review	☐	☐	☐	☐	☐
4	Review the contents of the page	☐	☐	☐	☐	☐
Question How would you rate your overall experience in performing this task?		☐	☐	☐	☐	☐

Question Justify the above rating (provide reasons for it)

Wait for Admin before proceeding ☺

Task 3 Use the Firefox browser on the Tablet that is provided to you and after the completion of each task outlined below **rank your experience (from 1 to 5).**

		Poor (1)	**Fair (2)**	**Good (3)**	**Very Good (4)**	**Excellent (5)**
1	Go to **www.google.com**	☐	☐	☐	☐	☐
2	Search for **Toji Japan** and review the results of the first page	☐	☐	☐	☐	☐
3	Click on a page you want to review	☐	☐	☐	☐	☐
4	Review the contents of the page	☐	☐	☐	☐	☐
Question	How would you rate your overall experience in performing this task?	☐	☐	☐	☐	☐

Question Justify the above rating (provide reasons for it)

Task 4 Use the Firefox browser on the Smartphone that is provided to you and after the completion of each task outlined below **rank your experience (from 1 to 5).**

		Poor (1)	**Fair (2)**	**Good (3)**	**Very Good (4)**	**Excellent (5)**
1	Go to **www.google.com**	☐	☐	☐	☐	☐
2	Search for **Toji Japan** and review the results of the first page	☐	☐	☐	☐	☐
3	Click on a page you want to review	☐	☐	☐	☐	☐

4 Review the contents of the page	☐	☐	☐	☐	☐
Question How would you rate your overall experience in performing this task?	☐	☐	☐	☐	☐

Question Justify the above rating (provide reasons for it)

Wait for Admin before proceeding ☺

Task 5 Use the Firefox browser on the Smartphone that is provided to you, and **while seated**, attempt to search and find information on **Toji Japan** on the webpage **www.google.com.**

	Poor (1)	Fair (2)	Good (3)	Very Good (4)	Excellent (5)
Question How would you rate your overall experience in performing this task?	☐	☐	☐	☐	☐

Question Justify the above rating (provide reasons for it)

Wait for Admin before proceeding ☺

A new feature on the phone has been enabled and by tilting the phone forward the page scrolls down automatically; tilting the phone backwards scrolls up the page.

Task 6 Use the Firefox browser on the Smartphone that is provided to you, and **while walking**, attempt to search and find information on **Toji Japan** on the webpage **www.google.com.**

	Poor (1)	Fair (2)	Good (3)	Very Good (4)	Excellent (5)
Question How would you rate your overall experience in performing this task?	☐	☐	☐	☐	☐

Question Justify the above rating (provide reasons for it)

Lab Study Consent Form

You are being asked to take part in a research study examining how quality of experience is perceived by mobile device users in certain scenarios. Please read this form carefully and ask any questions you may have before agreeing to take part in the study.

What the study is about: The purpose of this study is to identify how users perceive experience in accessing online content from mobile devices. This study examines factors that can potentially affect user experience and evaluates the effect of certain experience enhancement/detraction measures.

What we will ask you to do: If you agree to be in this study, you will be asked to follow a series of specific web related tasks on mobile devices (tablet and smart phone) and answer a brief questionnaire in which you have to state your own opinion of how satisfied/dissatisfied you were on the results of each task. There are 2 types of questions which you will be asked to answer. The first is to give a satisfiability rating (from 1 to 5) for a web visit on a particular site. The second is to rank the importance of factors which you believe affect your experience in different web usage scenarios.

You will not be asked to provide or input any personal identifiable information and therefore no personal information will be recorded in any of the web sessions.

Some of the types of websites you'll be asked to visit are professional freelancing sites, news sites, online stores, search engines, etc. and do not include sites with adult content and/or sensitive material that people may find offensive.

Risks and benefits: There are no anticipated risks in participating in this study**.** There are no benefits to you other than participating and being part of leading Worldwide Quality of Experience research ☺

Compensation: A raffle draw at the end of the lab tests will award 1 lucky participant with a brand new Android tablet.

Data Protection: Your answers will be **confidential**. The records of this study will be kept private. In any sort of report, we make public, we will not include any information that will make it possible to identify you. Research records will be kept in a highly secure storage space; only the researchers will have access to the records.

Taking part is voluntary: Taking part in this study is completely voluntary. You may skip any questions that you do not want to answer. If you decide not to take part or to skip some of the questions, it will not affect your current or future relationship with UCLan. If you decide to take part, you are free to withdraw at any time.

If you have questions: The researchers conducting this study are Elias Allayiotis and Dr. Josephina Antoniou. Please ask any questions you have now. If you have questions later, you may contact Elias at EAllayiotis@uclan.ac.uk. You can reach Dr. Antoniou at JAntoniou@uclan.ac.uk. If you have any questions or concerns regarding your rights as a subject in this study, you may contact UCLan's Ethics Officer at OfficerforEthics@uclan.ac.uk

You will be given a copy of this form to keep for your records.

For more information on this project visit www.betterX.org

Statement of Consent: I have read the above information, and have received answers to any questions I asked. I consent to take part in the study.

Your Signature ______________________________ Date ______________________

Your Name ______________________________ Phone or Email ______________________

List of Acronyms

Acronym	Definition	Acronym	Definition
API	Application Programming Interface	MLQoE	Machine Learning QoE
APPS	App Launches	MOS	Mean Opinion Score
APPUS	App Usage	MSE	Mean Square Error
AWS	Amazon Web Services	MWQoE	Mobile Web Quality of Experience
BATT	Battery Value	NCS	Network Context State
BATV	Battery Availability	PESQ	Perceptual Evaluation of Speech Quality
CALLS	Total Calls attribute	PHUS	Phone Usage
CaQoEM	Context-Aware QoE Model	PL	Page Load Time
CI	Credible Intervals	PSNR	Peak Signal-to-Noise Ratio
CL	Content Load Time	QoE	Quality of Experience
CSTR	Connection Strength	QoS	Quality of Service
DCS	Device Context State	S3	Amazon Simple Storage Service
DOMT	Domain Type	SCREEN	Screen State Changes
ER	Entity Relationship Diagram	SE	Software Engineering
ESM	Experience Sampling Method	SIGL	Network Signal Strength
GCM	Google Cloud Messaging	SIGS	Signal Strength
GOMS	Goals, Operators, Methods and Selection rules	SQL	Structured Query Language
GPS	Global Positioning System	SSIM	Structural Similarity Index
HAR	HTTP Archive Format	SSL	Secure Sockets Layer
HCI	Human-Computer Interaction	STU	Semantic Textual Unit
HDR	Highest Density Region	SVC	Scalable Video Coding
HTTP	Hypertext Transfer Protocol	SVG	Scalable Vector Graphics
ICS	Web Intent Importance Context State	TCP	Transmission Control Protocol
IMEI	International Mobile Equipment Identity	TCS	Timings Context State
INT	Internet	THTTP	Total HTTP Timings
INTNT	Web Session User Intent	TRES	Total Size of the Response
ISP	Internet Service Provider	TRTS	Total Size of the Request
ITU	International Telecommunication Union	UI	User Interface
IWWIS	Intent Weighted Web Immediacy State	UID	User Identification Number
KPI	Key Performance Indicator	UMTS	Universal Mobile Telecommunications System
LOC	Location Type	URL	Uniform Resource Locator
LSPD	Network Link Speed	UTYPE	User Type
MAUT	Multi-Attribute Utility Theory	UX	User Experience
Mbps	Megabits per second	VoIP	Voice over Internet Protocol
MCC	Mobile Cloud Computing	VQM	Video Quality Metric
MEID	Mobile Equipment Identifier	W3C	World Wide Web Consortium
		WIS	Web Immediacy State