

EventSnap: A Smart Networking Based Video Sharing Application

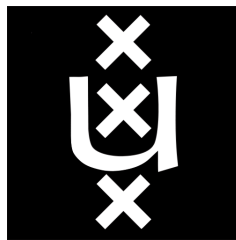
SUBMITTED IN PARTIAL FULLFILLMENT FOR THE DEGREE OF MASTER
OF SCIENCE

FABIJAN BAJO
10184805

MASTER INFORMATION STUDIES
HUMAN-CENTERED MULTIMEDIA

FACULTY OF SCIENCE
UNIVERSITY OF AMSTERDAM

July 7, 2016



1st Supervisor
Dr. P.S. César Garcia
CWI

2nd Supervisor
J.W.M. Kleinrouweler
CWI

EventSnap: A Smart Networking Based Video Sharing Application

Fabijan Bajo

Information Studies, University of Amsterdam
Amsterdam, The Netherlands
fabijan.bajo@student.uva.nl

ABSTRACT

State of the art mobile technology and rapid deployment of WiFi hotspots have empowered the mobile video sharing experience. However, long upload durations and unreliable network behavior, as perceived by the user, are negatively impacting the sharing experience. We have designed, developed and evaluated a mobile video sharing app, equipped with smart networking technology to address this problem. We performed comparative evaluation ($N = 21$), measuring possible enhancements in perceived feelings caused by network aware video sharing and measured global usability. All app features were rated positively, correlating to a high usability and hedonic quality score. Smart networking significantly improved perceived reliability, which especially in the context of crowded network environments is an important discovery.

Keywords

Smart Networking, Bandwidth Distribution, Mobile Video Sharing, Usability

1. INTRODUCTION

Mobile video sharing apps have rapidly grown in popularity and have become an integral part of the online social experience [1]. State of the art mobile technology allows for 4k video recording, where one minute of footage at 30 FPS corresponds to a 375 MB upload [2]. To meet the networking demands, large conventions provide WiFi hotspots, serving hundreds of users at the same time. However, the performance of WiFi hotspots serving locations, such as busy airports has been poor due to unfair bandwidth distribution among clients and traffic asymmetry [3]. We propose an app that incorporates a smart networking architecture as a solution to this problem, focusing on perceived feelings caused by the sharing process. *Smart networking* (SN) corresponds to a computer networking approach where information, such as network load and bandwidth distribution among clients are incorporated in the end-user app to allow for adaptation to the network [4].

When available bandwidth for each user becomes scarce, upload times increase, causing for a negative impact on the video sharing experience [5]. Encompassing moments of temporary interaction disruptions with the system are sources of negative emotional feelings, such as user anxiety. [6, 7]. Furthermore, current networking architectures do not provide access to network state information. Without any user involvement, the majority of system recommendations

and decisions reside on the background, contributing to a lack of control over the sharing process [8]. The "one size fits all" approach leaves no room for second thoughts or personal preference. Moreover, as the network load increases, unexpected upload disruptions may occur, negatively impacting perceived reliability. Trust plays an important role in human-computer interaction, as it helps users to overcome risk and uncertainty [9].

We hypothesize that the incorporation of a SN architecture could enhance the perceived control, reliability and satisfaction of the video sharing process, while attaining high usability. We present the results of 21 user tests with a mobile video sharing app, designed to make use of and communicate with a SN server. Due to complexity of the networking functionality, achieving high usability on the mobile platform can be challenging, which explains the rather twofold character of the hypothesis. A controlled lab-setting provides a suitable testbed for measuring the effects of SN on the sharing process as both "standard" and "smart" sharing can be measured and compared in a single user test, while being able to simulate various network-load scenarios. By executing a consistent procedure, we increase experiment replicability, while allowing for precise control of multiple conditions.

Our study addresses the following research questions:

- To what extent does incorporating a SN architecture enhance the perceived satisfaction, control, and reliability of the video sharing process?
- Is high usability attainable when incorporating a SN architecture and its features into an end-user app?

The remaining of this paper is as follows. Section 2 provides the background and related work. In Section 3, the interface and system description of the proposed application are outlined in detail. Section 4 outlines the methods, followed by the results in Section 5. Section 6 provides the discussion of the results and Section 7 covers the conclusion and future work.

2. BACKGROUND & RELATED WORK

To the best of our knowledge, there are no directly comparable end-user applications that incorporate networking features in a similar way, hence the emphasis on the background component in this Section.

2.1 Smart Networking

There has been a lot of interest recently for a more user-driven networking approach accompanied by more network

awareness with an emphasis on user interaction [10, 11]. From a Human Computer Interaction (HCI) perspective, this entails making network management an easier task for the user and understanding the effects of network performance on perceived user feelings and experience [10]. Chen et al. [12] addressed the impact of network quality and awareness on online gamers, acting as strong indicators of user satisfaction. Yiakoumis et al. [10] advocate that users should guide the management of network traffic, directly reflecting the user's preferences and improving the home networking experience.

The confluence of "Software Defined Networking" (SDN), "Network Function Virtualization" (NFV) and "Mobile Cloud Computing" (MCC) transform the network model and allow users to personalize their experiences in a more responsive and agile manner [11]. SDN enables an efficient and controllable network architecture for managing network traffic [13]. NFV is a way to virtualize network services [13]. MCC, in short, moves data storage from mobile phones into the cloud. The combination of these developments provides mobile apps access to network state information, which together with associated app features is what we define as SN.

With SN, our app development approach combines components from lower network layers with the application layer, therefore relying on close communication between app developer and network operator. Moreover, our research bridges the gap between more technical networking studies that focus on Quality of Service (QoS) and Quality of Experience (QoE), with higher layer end-user components, which are more related to the HCI field. The combination of the user interface description (HCI related) and (back-end) system description (QoE study related) of Section 3 exemplifies this unique approach.

2.2 Sharing Performance & Perceived Feelings

Waiting times perceived by a user during moments where interaction with the system is disrupted are root causes of user-anxiety and irritation [14]. Research in HCI has often stated that feedback on waiting time or process duration has the potential to enhance usability [14, 2]. When analyzing file transfers, visible download and upload times proved most relevant as perceived by the user [2]. A considerable body of literature on the interplay between network performance and user satisfaction can be found in QoE research. For instance, Fiadino et al. [2] measured perceived satisfaction of file sharing through WhatsApp, which they translated into a transfer duration threshold to discriminate between good and bad experience.

Perceived control, representing the user's perception of being in charge of the interaction [15], is a factor of great importance when developing user-driven networking architectures. Many researchers have investigated the ways humans control and interact with computers [16, 17]. In computer networking, control focuses on being involved in the decision-making process of particular networking aspects. For example, MyBoost, a single-button browser extension, allows users to interactively ask for more bandwidth [10].

Perceived reliability, related to human computer trust, focuses on system performance and questions whether the system provides the user with the required advice to make decisions. Can the user rely on the system to function properly? The concept is well discussed in topics as disruption-tolerant networking (DTN) where high reliability and low delivery

latency are important networking factors [18].

2.3 Pragmatics vs Hedonics

When dealing with mobile devices, many contextual factors are involved compared to a traditional website. The hedonic pragmatic model of User Experience (UX), conceptualized by Hassenzahl [19], provides a structured way for measuring usability, while also incorporating subjective feelings. Pragmatic qualities of the model are closely related to the classical concept of usability and focus on task-related aspects ("do-goals") of a particular system [19]. Hedonic qualities, on the other hand, are more related to the systems ability to evoke pleasure and stimulate the psychological well being of the user. These contribute to the product's perceived ability to achieve "be-goals", such as "being competent" for using the product. Due to its simplicity, the hedonic pragmatic model can be seen as a reductionist approach to UX [19], making it a suitable testing tool for lab-tests. Hassenzahl developed a standardized questionnaire called "AttrakDiff 2", assessing pragmatic and hedonic qualities of a product in a succinct and efficient manner [20].

Prior studies incorporating this model include TrainYarn [21], measuring the UX of a public transport app and DinerRouge, which combined the more traditional system usability scale (SUS) with the short version of AttrakDiff [22].

3. EVENTSNAPE: A VIDEO SHARING APP

At the pre-design phase, a preliminary survey of 10 questions regarding video sharing at large events was sent out. 70 respondents answered questions about their smartphone usage habits when sharing video and connecting to WiFi networks (Appendix C), such as preferred network state information during an event. Acquired insights helped with designing certain SN features. Furthermore, the design process was led by literature on network performance vs satisfaction-enhancing UI features [6, 7, 14] and disruption tolerant network concepts [18].

EventSnap is a video sharing app that lets users upload recorded video to a public accessible feed. It's comparable to the YouTube app¹, which transfers videos directly to a public server. What separates EventSnap is the added SN layer. Core SN features of the app include: a live network speed indicator, postpone indicator, upload customization features, such as postponing uploads, resuming disrupted uploads from where they left off and choosing video quality based on predicted upload times. Other important features include upload recommendations and guidance, putting users more in charge of the sharing experience and increasing network awareness.

3.1 User Interface

Only screenshots relevant to this section and the experiment are presented below.

3.1.1 Main Interface

EventSnap uses a tab bar interface, allowing for navigation between the "video library" (Figure 1A), "public feed" and "settings" (Figure 1B) screens. The library screen shows recorded video's from the respective device. A quick launch camera button on the bottom right of each tab (Figure 1) let's users instantly record a video.

¹<https://itunes.apple.com/nl/app/youtube/id544007664?mt=8>

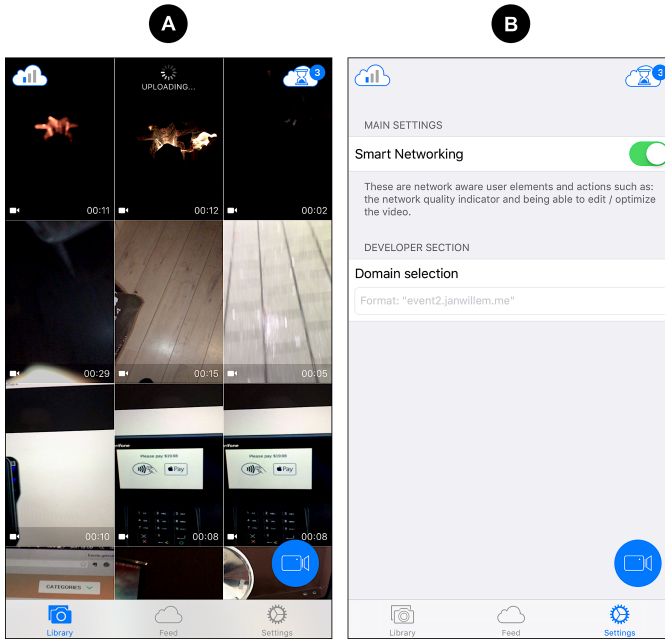


Figure 1:

Screenshots of the video library screen (A) and settings screen (B).

The "network speed indicator" on the top left (Figure 1A, 1B) represents the current network speed. A metaphoric internet cloud with bars, magnified in Figure 2A, shows the current speed rating. From a user perspective, the bars denote a "slow", "medium" and "fast" network. Under the hood, the indicator is mapped to the network load, providing more bars when more bandwidth becomes available. The indicator on the top right (Figure 1A, 1B) is the "upload postpone indicator" (magnified in Figure 2B). Its purpose is presenting the amount of uploads placed in the postpone list (upload queue). When an upload is running on the background (by automatically resuming a prior disrupted upload or triggering queued uploads based on calculated throughput), an activity indicator appears at the top of the screen (Figure 1A, magnified in 2C), indicating an automatic upload launch by the system.

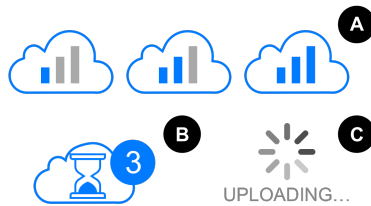


Figure 2:

Enlarged depiction of the speed indicator (A), postpone indicator (B) and activity indicator (C).

Figure 1B shows the preferences screen, where SN can be switched on or off. Switching off SN will hide all related UI elements (i.e., Figure 2) and turn off the system features. It's up to the user to use SN or share without extra networking functionality.

3.1.2 Sharing Video

Either by selecting a video from the library or directly sharing recorded footage, the user is directed to the preview screen (Figure 3A), playing back a preview of the video. Figure 3A presents the preview screen with SN turned on, which compared to no SN, shows a speed indicator above the share button and adds the "settings" button at the bottom right for upload customization. Tapping on the share button with SN on, launches an "action sheet", providing the user with information about the quality in which the video will be encoded and uploaded to the server. Current apps hide these automatic system decisions, losing valuable conversation with the user. When confirmed, the user navigates to the "upload" screen (Figure 3B), showing a speed indicator, upload percentage, progress bar and upload duration. Without SN, only the upload percentage and progress bar are shown.

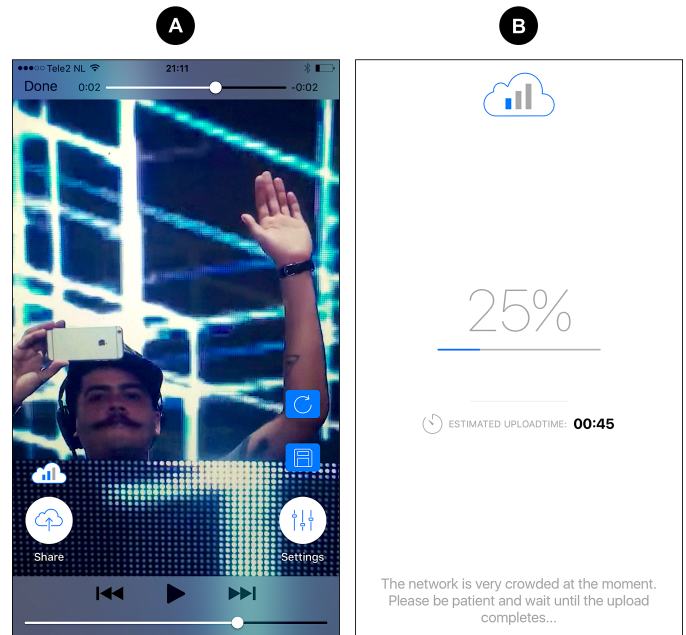


Figure 3:

Screenshots of the preview screen (A) and upload screen (B).

3.1.3 Customizing Uploads

Tapping on the "settings" button at the preview screen (Figure 3A) directs the user to the upload settings screen (Figure 4A). The recommended setup switch (Figure 4B) provides a quick calculated guess on best quality related to the current network load. The upload duration label at the bottom (Figure 4A) shows an approximation of the upload time, refreshing instantly after each user interaction with the customization interface. By horizontally swiping the video quality selection carousel (Figure 4A), a desired video quality can be selected. The thumbnails in the carousel show encoding previews of the video when applying the selected quality. Switching on the postpone switch (Figure 4A) before tapping on the share button (bottom of the screen), will place the upload in a postpone queue, incrementing the postpone indicator afterwards (Figure 2B).

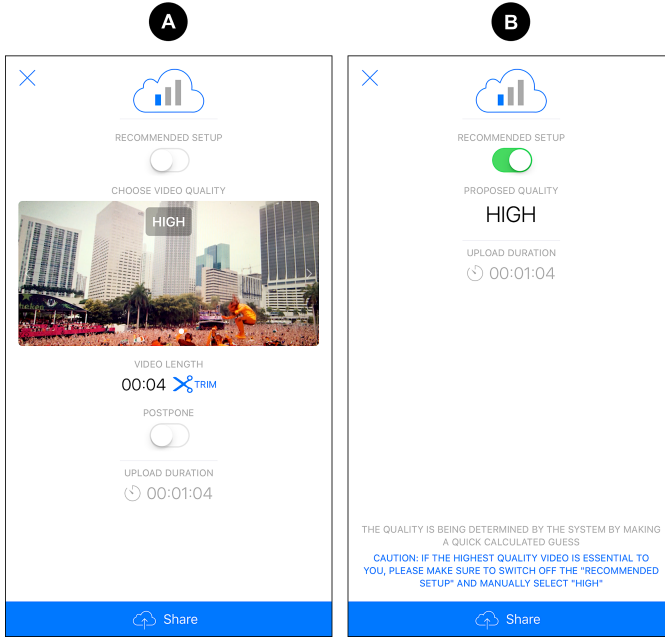


Figure 4:

Screenshots of the upload-settings screen on load (A) and with "system recommendation" switched on (B).

3.2 System Description

3.2.1 Hardware

EventSnap is developed for the iPhone with a minimum requirement of iOS 8 or later. The SN architecture requires programmable hardware, i.e., configurable by the OpenFlow protocol (allows for SDN) and a network controller [4]. On top of the network controller, a service manager must be added to communicate directly with the mobile devices in the network.

The network controller is in charge of configuring and monitoring the network components and has a global view of the clients [4]. The service manager interacts with the controller and mobile devices, allocating required resources for the upload requests. The architecture allows for very predictable uploads in terms of speed and time.

3.2.2 Software

Based on video QoE studies [23, 13], "Low", "Medium" and "High" presets were generated for the app's networking features (Table 1). Choosing the right presets for encoding mobile video is a combination of tradeoffs. At 240p, the "low" preset uses a relatively high video bitrate. This way a minimum acceptable quality level can be guaranteed at the lowest preset, which apart from a lack of detail is acceptable at a low resolution [23]. At 480p, the quality could be described as "very good" with few visible compression artifacts. An average link speed of 2+ Mbps would suffice. The "high" encoding preset produces high definition (HD, 1280 x 720). For users seeking high quality, this preset will suffice.

Upload time predictions from the SN server are put against a maximum waiting threshold, determined by QoE satisfaction scores. Users tolerate transfers up to 20 seconds with a good overall experience [2]. Transfers lasting more than 80

Table 1:
Video encoding presets.

Preset	Resolution	Video	Audio
Low	424 x 240 (240p)	576 kbps	64 kbps
Medium	848 x 480 (480p)	1216 kbps	128 kbps
High	1280 x 720 (720p)	2496 kbps	192 kbps

seconds are considered as "very bad". An 80 second threshold is used to generate video quality recommendations for EventSnap users. The highest quality below the upload duration threshold is recommended to the user.

Without SN, video files are transferred as whole MP4 units, a standard approach among popular video sharing apps. However, with SN, the MP4 file is segmented into smaller data chunks, which are transferred individually. After each successful data chunk transfer, the upload progress is saved to disk, allowing for resumable uploads and disruption tolerance.

The user makes a trade-off between upload time and video quality based on the network speed indicator (Figure 2A) and upload duration prediction at the bottom of the upload settings screen (Figure 4A). When postponing an upload, the upload and respective device are registered by the server. From here on, triggering the upload is in the servers hands. When bandwidth becomes available, the server launches the upload using Apple's push notification system. When dealing with WiFi, network performance decreases rapidly with every new client. Postponing reduces the overall network load, aiming for a fair distribution of available bandwidth resources and overall upload time reduction.

4. METHODS

4.1 Participants

21 subjects (6 female, 15 male), aging between 20 and 35 participated in our 50 minute user test. 15 were students (11 PhD, 4 Masters). The remaining participants included a waiter, car rental attendant, research project coordinator, programmer, app company worker and a Linux system administrator. 10 participants were most familiar with the iOS platform, 10 with Android and 1 with the Windows Phone platform. When questioned about familiarity with the iOS platform, 9 answered being "very familiar", 7 "not very familiar", 3 were "familiar" with the platform, and 1 participant never used it. Participants were recruited using convenience sampling, i.e. selection based on availability and/or accessibility.

4.2 Design

The experiment was designed with the test variables SN (On or Off) and network-load (Quiet or Busy), constituting a total of 8 tasks. All participants tested each setting of the experiment and filled out 2 separate post-test questionnaires after completing the tasks.

The design constitutes three assessment objectives, differing in analysis, measure goals and research question focus. Each component is separately discussed below.

4.2.1 Post-task assessment: Perceived Feelings

Addressing the first research question, perceived satisfaction, control and reliability were measured through comparison of different conditions using 6 questions on a 5-point Likert scale (Table 2). In this comparative experiment, perceived satisfaction relates to the perceived upload time and speed by the user. We developed 2 custom scales, assessing participants on perceived upload time (very short - very long) and perceived upload speed (very slow - very fast). Perceived control relates to the extent in which users feel in control over the sharing process (their influence). Scales for perceived control were adapted from those developed by Agarwal and Karahanna [15], using the control module from the cognitive absorption questionnaire and are presented in table 2. Perceived reliability denotes the user's perceived trust for the system, mainly focusing on whether uploads make it to the server without unexpected disruptions. Perceived reliability questions were adapted from Madsen and Gregor's Human Computer Trust questionnaire, shown in Table 2 [9]. Both perceived control and reliability based questions ranged from strongly disagree to strongly agree.

We used a 2-factor within-subject design, where both SN (On or Off) and network-load (Quiet or Busy) were within-subject factors. The quiet network setting was 10 mbps and the busy network 0.6 mbps, simulating a crowded network environment. To reduce order-effects, we counterbalanced the order of tasks and formed 4 sub-groups, each presented with a different order.

Table 2:

Arrangement of the perceived feelings questionnaire and the merged items.

Category	Question
Satisfaction	- Rate the upload time
	- Rate the upload speed
Control	- When I was using the app, I felt in control
	- I felt I had no control over the interaction with the app
Reliability	- The app provided me with the advise
	I needed to make decisions
	- The systems performs reliable

4.2.2 Post-task assessment: App Features

To gain more user insights on specific SN features and whether they proved useful in completing certain tasks, we extended the base module (Table 2) and added 14 questions (Appendix A2) on perceived satisfaction and perceived usefulness. Satisfaction for this assessment denotes whether the user was satisfied with a particular app feature. These produced single ratings and addressed the second research question by contributing to the global usability assessment. Both satisfaction and usefulness were measured using 5-point Likert scales (strongly disagree - strongly agree) and were adapted and modified from Davis et al. [24] (Appendix A2).

4.2.3 Post-test assessment: SUS & AtrakkDiff

To address the second research question, two global post-test questionnaires assessed the participants on perceived

usability and perceived pragmatic vs hedonic qualities.

To measure global usability, we used the SUS questionnaire, including 10 items on a 5-point Likert scale (strongly disagree - strongly agree) (Appendix A3). The SUS provides a global view of subjective assessments of usability. [25].

The short 10-item version of the AttrakDiff 2 questionnaire was deployed to evaluate pragmatic and hedonic qualities of the app (Appendix A4). The scales contain seven stages between opposing word-pairs, such as, "complicated" vs "simple". Results are described in terms of the four dimensions, PQ (pragmatic quality), HQ-I (identity), HQ-S (stimulation) and ATT (attractiveness) [20].

4.3 Tasks

The experiment was divided into 3 parts, which we counterbalanced and structured as follows (we did not inform users on network behavior prior task execution).

1. Part 1: sharing without SN
 - (a) On a quiet network: upload a 30 second video
 - (b) On a busy network: upload a 30 second video
 - (c) On a busy network: upload a 30 second video and interrupt the upload at 30% by closing the app completely. Restart the app afterwards
2. Part 2: sharing with SN
 - (a) On a quiet network: upload a 30 second video
 - (b) On a busy network: upload a 30 second video
3. Part 3: customizing the upload before sharing
 - (a) On a busy network: upload a 30 second video, but before sharing, handpick the video quality based on provided network information
 - (b) On a busy network: upload a 30 second video using the postpone feature. Lock the app and wait until you receive a push notification
 - (c) On a busy network: upload a 30 second video and interrupt the upload at 30% by closing the app completely. Restart the app afterwards

4.4 Apparatus

The user tests were performed in a lab at CWI in Amsterdam. The room was prepared with two 13,3 inch Apple Mac book pro laptops, an iPhone 6, 6+ and a custom WiFi setup as in [4]. One of the laptops was used by the network operator for controlling and simulating artificial network traffic and the other for answering the questionnaires. The iPhone 6 was used as a remote for controlling the "busyness" of the network from a distance through a custom made app. In combination with the laptop, the network operator was able to control the network very accurately according to specific task requirements. All experimental conditions were run on the iPhone 6+. The network was implemented using an OpenFlow-enabled Raspberry Pi, which acted both as a traffic controller and WiFi access point. The WiFi network was set up using a WiFi dongle (TL-WDN4200 USB adapter).

4.5 Procedure

Participants first signed a consent form about the use of the collected data. A short introduction followed on the topic and goal of the research, after which the test environment was explained. The network operator went briefly over his role and denoted the variable behavior of the network. Next, we briefed the participants on the app with a short

demonstration of the features and described the assessment procedure. When completing a task (explained on a hand-out), a short questionnaire followed related to the task itself. After completing all tasks, the users filled out the SUS and AttrakDiff questionnaire for a global assessment of the app.

5. RESULTS

5.1 Perceived Feelings

All interpreted task combinations in Section 5.1 refer to Table 3, which shows a schematic overview of the performed tasks and related variables. Figures 5 and 6 show plotted medians and interquartile ranges (IQR) for all compared groups. As can be seen in Table 3, customization tasks were all performed on a busy network, which explains the different plot construction shown in Figure 6.

The perceived feelings ratings were interpreted as ordinal values and merged into 3 components using Cronbach's Alpha tests, estimating the average correlations. This resulted in satisfaction (0.84), reliability (0.71) and control (0.74), which scored above the required minimum of 0.7 [26].

For every comparison, such as comparing all "only sharing tasks (Table 3)", an initial Friedman test was conducted to determine whether there was a significant difference somewhere between one of the compared tasks. We chose a significance level of 0.05 to determine whether the null hypothesis (H_0) must be accepted or not. Based on the chi-square distribution (χ^2), which specifies the chance of the hypothesis being accepted, we determine the significance level (p). If $p < H_0$, we know that there are differences somewhere between the tasks, without knowing exactly where those differences lie. To examine between which task comparisons the differences occurred, we run a separate (post hoc) Wilcoxon signed-rank test on each related task combination, which results in a significance level (p -value) for each comparison.

Table 3:

Schematic overview of the tasks from Section 4.3 and related conditions.

Task	SN	Network Load
1 Only sharing	Off	Quiet
2 Only sharing	Off	Busy
3 Upload interruption	Off	Busy
4 Only sharing	On	Quiet
5 Only sharing	On	Busy
6 Customizing (video quality)	On	Busy
7 Customizing (postponing)	On	Busy
8 Upload interruption	On	Busy

Satisfaction (Perceived Upload Time and Speed).

We begin comparing tasks 1, 2, 4 and 5, which are "only sharing" tasks (Table 3). Users performing these tasks, only had to share a video, without extra customization of the upload. SN features for only sharing tasks include, a network speed indicator, upload predictions and extra provided information on system decisions. As shown in Table 3, the tasks differ in conditions they were performed in.

A Friedman test resulted in significance $p = 0.001$ ($\chi^2 = 52.783$). Therefore, a post hoc analysis with Wilcoxon

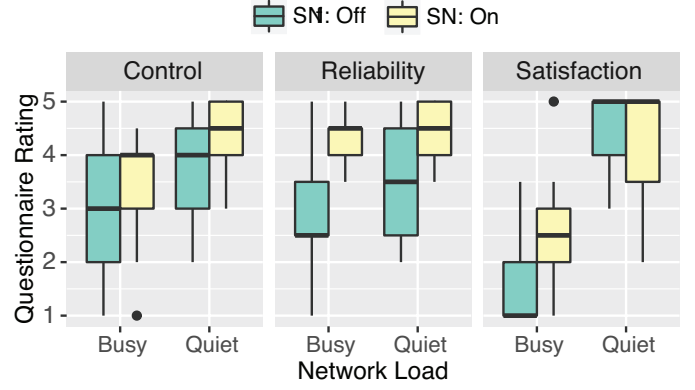


Figure 5:

Boxplots with IQR's for "only sharing" tasks. Variables SN and Network load constitute task conditions as presented in Table 3. The plot compares the following tasks: task 2 (SN: Off with Network: Busy) versus task 5 (SN: On with Network: Busy), and task 1 (SN: Off with Network: Quiet) versus task 4 (SN: On with Network: Quiet). The ratings relate to the questions presented in Table 2 where each perceived "feeling" is the mean of the 2 questions. The dots denote outliers.

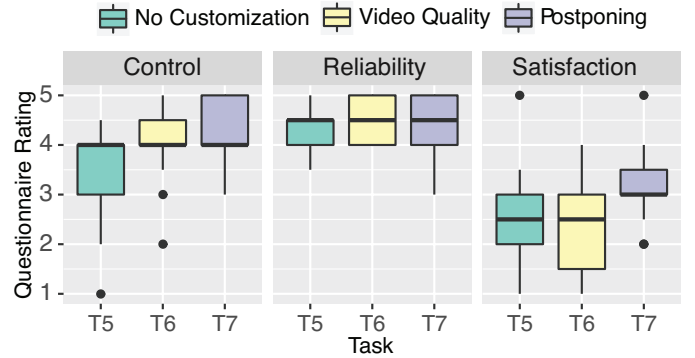


Figure 6:

Boxplots with IQR's for "customization" tasks (SN: On, Network: Busy). The ratings relate to the questions presented in Table 2 where each "feeling" is a combination of 2 questions. The dots denote the outliers.

signed-rank test was conducted to determine the exact cause of the difference.

The satisfaction level did not change significantly when comparing the presence of SN on a quiet network ($p = 0.938$), i.e., turning on SN on a quiet network did not change the user's perceived upload time and speed. The same counts for a busy network (task 2 vs 5, $p = 0.301$).

However, the comparison of tasks 1 vs 2 ($p = 0.003$) and tasks 4 vs 5 did confirm a significant difference. These comparisons confirmed that a quiet network (0.6 mbps) is perceived fast and a busy network (10 mbps) is perceived as slow, basically acting as a control test for further experiments.

Perceived Control.

Only Sharing

We ran a Friedman test on "Only sharing" tasks, 1, 2, 4 and 5 (Table 3), where mean ratings for perceived control ques-

tions (Table 2) were compared. This resulted in a significant difference ($p = 0.002$, $\chi^2 = 24.28$).

The cause of this difference was not due to comparison of tasks 1 and 4 ($p = 0.108$), i.e., adding SN features, such as an upload prediction and speed indicator, did not improve perceived control ratings on a quiet network. A busy network likewise did not cause for changes (task 2 vs 5, $p = 0.301$).

The actual significant differences were measured between tasks 1 vs 2 ($p = 0.046$) which both were performed without SN, but here the network load determined the rating differences. Task 4 vs 5 ($p = 0.014$), showed a significant difference in perceived control, but this time with SN turned on. In short, when looking at perceived control, the network load had more influence on the user ratings than added SN features in this particular experiment.

Customizing

To determine whether additional customization features statistically improved perceived control, we compare task 5 (no customization) with task 6 (picking a custom video quality) and 7 (postponing an upload) (Table 3).

The initial Friedman test resulted in $p = 0.001$ ($\chi^2 = 52.783$). A post hoc analysis was thus needed to determine the source of significance.

There was no significant difference measured between ratings for tasks 5 and 6 ($p = 0.938$), i.e., the perceived control level showed no statistical significance when adding the video quality selection feature. However, the postpone feature did significantly impact perceived control (task 5 vs 7, $p = 0.008$).

Upload Interruption

For tasks 2 (SN: Off) and 8 (SN: On), the user had to intentionally interrupt the upload to experience how the system handles the disruption under different conditions. A two-group comparison based Wilcoxon signed-rank test was conducted for the "upload interruption" combination (Table 3), resulting in significance $p = 0.0002$. With SN turned on (task 8), the upload is resumed when re-launching the app, meaning that the ability to resume a broken or interrupted upload from where it left off increases perceived control significantly.

Perceived Reliability.

Only Sharing

A statistically significant difference in perceived reliability was observed from the "only sharing" combination (Table 3). The initial Friedman test resulted in significance $p = 0.002$ ($\chi^2 = 38.62$).

The cause of this difference was not due to comparison of tasks 1 and 2 ($p = 0.151$), i.e., without SN features, perceived reliability ratings were not impacted by varying network load. Tasks 4 and 5, similarly revealed no significant difference ($p = 0.793$).

The actual significant difference was measured between tasks 1 vs 4 ($p = 0.002$), where smart networking determined the significant improvement in ratings. Task 2 vs 5 ($p = 0.014$) likewise showed a significant difference in perceived reliability. In short when looking at perceived reliability, smart networking features had more influence on the user's ratings than the network load.

Customizing

Additional customization features did not significantly improve perceived reliability, as the Friedman test on tasks 5, 6 and 7 (Table 3) showed no significant difference ($p = 0.133$ ($\chi^2 = 4.01$)). We accepted H_0 , post hoc analysis was not needed.

Upload Interruption

The upload interruption tasks (2 vs 8) were evaluated in terms of reliability. A two-group comparison based Wilcoxon signed-rank test resulted in significance $p = 0.0001$. With SN turned on (task 8), the upload is resumed when re-launching the app, meaning that the ability to resume a broken or interrupted upload from where it left off increases perceived reliability significantly.

5.2 Post-task questions analysis: App Features

Ratings for the app feature related module are plotted in Figures 7 and 8, both showing means grouped by assessment objective (confidence interval = 95%). Appendix B, Tables 4 and 5, present the full question modules with provided Mean (M) values and standard deviations (SD's). The smaller the standard deviation, the more users were at one for a particular question.

5.2.1 Only Sharing

The post-task module for "only sharing" tasks with SN contained 7 additional questions, measuring perceived satisfaction and usefulness of specific UI elements and "smart" system recommendations. The complete Likert-scale items (strongly disagree - strongly agree) are shown in Appendix B (Table 4), together with the Mean values (M) and standard deviations (SD).

As shown in Figure 7, information on the upload duration (Q3, SD = 0.46, M = 4.7) proved most satisfying. This means that when performing "only sharing" tasks (table 3), the SN feature that satisfied most was having information on the upload duration (a prediction of how long the upload will take). When looking at perceived usability, question 7 was rated best (SD = 0.75, M = 4.2), i.e., When performing only sharing tasks (table 3), the most useful SN feature was the network speed indicator to understand the upload time.

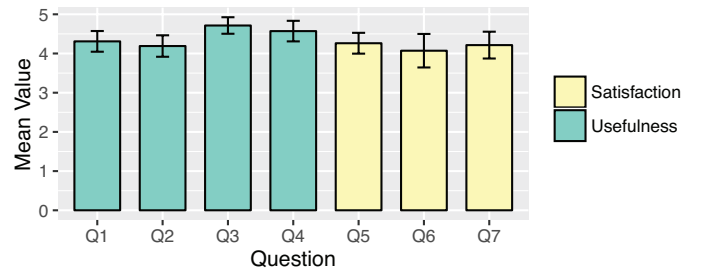


Figure 7:

Summary of the mean ratings for "only sharing" tasks. The full questions are presented in appendix B, Table 4.

5.2.2 Customization

Post-task questionnaires for customization tasks contained 7 additional questions, measuring perceived satisfaction and usefulness of SN features during customization. The complete Likert-scale items (strongly disagree - strongly agree)

are shown in Appendix B (Table 5), together with the Mean values and SD's.

Figure 8 shows the highest mean ratings for Q3 (SD = 0.3, M = 4.9) and Q6 (SD = 0.68, M = 4.4). Question 3 assessed users on whether they liked the resumable upload feature (Appendix B, Table 5), implying that in terms of satisfaction, the SN feature that satisfied most while performing the customization tasks (Table 3), was the resumable upload feature. Question 6 scored best in terms of perceived usefulness when performing customization tasks, meaning that the upload prediction label was perceived as most useful.

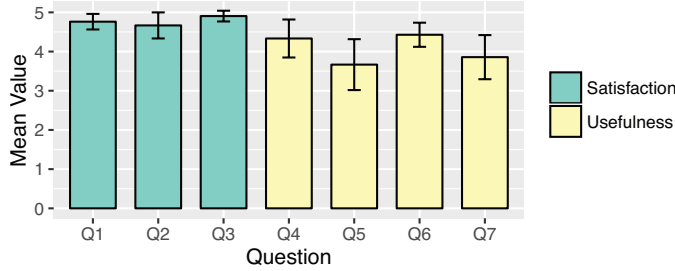


Figure 8:

Summary of the mean ratings for "customization" tasks. The full questions are presented in appendix B, Table 5.

5.3 Post-test questionnaires

5.3.1 Usability (SUS)

The mean SUS score was 87.3 out of 100, corresponding to an adjective rating of "excellent" usability and grade rating of "B" in the qualitative rating scale (Appendix B, Figure 11), specifically developed for interpreting SUS scores [27].

5.3.2 Pragmatics vs Hedonics (AttrakDiff)

Mean values of the word-pairs are shown in Figure 9. Of particular interest are the extreme values, which show which characteristics are critical or particularly well-resolved [20]. All mean scores reside on the right side (positive region) of the vertical scale, describing the app as "clearly structured", "stylish", "creative" and "good" in terms of PQ, HQ-I, HQ-S and ATT, respectively.

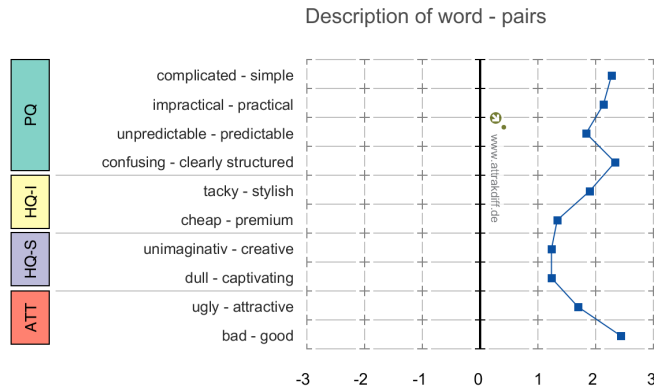


Figure 9:

Attrakdiff results: description of word-pairs.

The average values of the AttrakDiff dimensions for EventSnap are plotted in Figure 10. The best value has been achieved in dimension PQ (SD = 2.16). This value gives an indication that participants rated the app as "usable" in terms of interface usability. The lowest mean value has been achieved in the HQ-S (SD = 1.44) dimension, covering the hedonic quality, "stimulation".

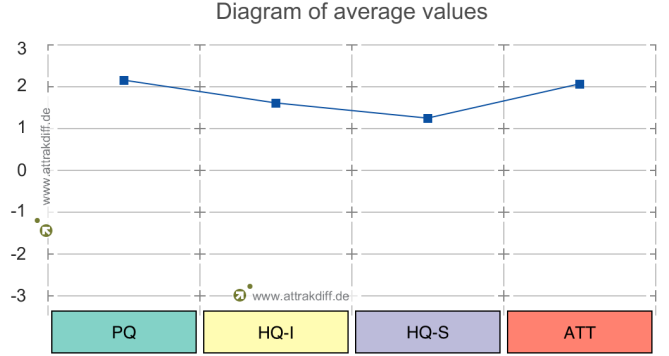


Figure 10:

Attrakdiff results: diagram of average values.

Figure 12 (Appendix B) presents the overall outcome of the questionnaire and classifies the app into a character region. "Desired", the most positive area highlighted in light-gray is where Eventsnap is classified into. The blue rectangle illustrates the confidence interval of the assessment. The small size of the rectangle shows that users were at one when it came to rating PQ (SD = 0.25) and HQ (SD = 0.37).

6. DISCUSSION

Our results suggest that network state information and upload recommendations are highly appreciated during the upload and customization process, suggesting that the users would like to have the SN features in the app and that they cause for an improvement in the sharing process. This correlates with the high usability score from the SUS and the pragmatic quality score from Attrakdiff. Clearness and ease of use proved most valued in the global assessment. Adding technically complicated networking features (e.g. resource allocation) could negatively impact the usability. Even so, by translating them to user friendly interface elements, such as the speed indicator, intuitive recommendation messages and visual customization controls with live feedback, we managed to attain a user-friendly sharing experience and high usability.

Results from the perceived feelings assessment showed large variations in perceived satisfaction, control and reliability. Our findings suggest that SN has no statistical significant effect on perceived satisfaction (perceived upload time and speed). However, it is important to note that the actual upload times stay the same under the hood, regardless of the SN features. Upload predictions and the speed indicator did not manipulate the perceived upload duration, thus in the case of mobile video sharing, predictions and network speed information don't change perceived upload time. On the other hand, users certainly indicated being satisfied with the upload predictions and network speed indicator while sharing and waiting, which is in line with findings from Gronier

et al. [14], stating that feedback (e.g. icons and text messages) on waiting time improves the usability and satisfaction of an interactive system.

SN did significantly improve perceived reliability, which especially in the context of crowded network environments is an important discovery. Users are uncertain about their uploads, making reliability a key factor in the sharing process. People want to know what to expect and whether the things they do will succeed in the end [9]. A good example was the highly rated (both by perceived feelings and app feature assessment) resumable upload feature of EventSnap. By segmenting the video file into smaller data chunks, disrupted uploads were able to continue from where they left of. A much appreciated feature, which proves a valuable addition to the sharing experience. Postponing uploads generally provides reliability to other clients. The user temporarily puts the upload on hold at critical moments, making sure there is enough bandwidth for the network itself. Perhaps if we were able to perform tests with multiple users instead of simulating traffic, the postponing of uploads would impact everyone's sharing experience, causing for positive results across all perceived feelings categories.

We suspect that the predefined upload times on a busy network had a bad impact on the control measure as users generally expected the uploads to go faster after customizations. What surprised us more was that customizing video quality did not significantly improve perceived control. Contrarily, the postpone feature did impact perceived control. Our expectations were the other way around as the video quality feature provides the user with actual choices. A controversial finding as one would think giving the upload in the system's hands, launching the upload whenever it decides to, would have a negative effect on perceived control, in line with Yiakoumis et al. [10], who measured perceived control over the home network.

The contradictory results observed in this study between the app features and perceived feelings based assessment are noteworthy as they underline the difference between having a useful and easy to use application and experiencing positive feelings from the process itself. Hasenzahl's approach to UX underlines this contradiction [19], stating that app features can be linked to two distinct objectives: "optimizing human performance" or "optimizing user satisfaction". However, the statistical comparisons must not be seen as definitive measures of perceived feelings. The exceptional high ratings from the other assessments make a re-design of the experiment worthwhile in future work.

6.1 SN and End-User Products

Though we focused only on video sharing, our approach to SN has broader application. By closely connecting app developers and network operators, users can become more network aware, contributing to a fair networking environment, while still being able to enjoy a highly usable social networking app. SN capabilities will become more important as knowledge about clients, end-users and the network state can be used to customize the network for optimal end-user experiences.

When developing a SN based end-user application, developers should take into account the fact that currently there are no standardized protocols for a direct communication flow between end-user app and SN server. Close communication between developer and network operator allowed for

a custom protocol for EventSnap. Network topology, switch model, and dedicated SN server software all must be taken care of.

6.2 Study Limitations

Though a controlled lab-setting was suitable for our study design, a field experiment would probably result in more natural subjective ratings, resulting in higher ecological validity.

The design of the prototype and complexity of the protocol allowed for one test user at a time, resulting in a rather isolated testing experience. Testing with multiple devices on one SN network would give a better sense of a shared WiFi environment.

With 8 user tasks, a user test of 50 minutes is relatively long. Even though counterbalancing was implemented, we noticed small experiment effects, such as learning effects and fatigue.

7. CONCLUSION

This paper presents the findings of an experiment testing whether the incorporation of a smart networking architecture could enhance perceived feelings towards the sharing experience, while attaining high usability. Smart networking (SN) corresponds to a computer networking approach where information, such as network load and bandwidth distribution among clients are incorporated in the end-user app to allow for adaptation to the network.

To address the first research question, "To what extent does incorporating a SN architecture enhance the perceived satisfaction, control, and reliability of the video sharing process?", we should look at both direct perceived feelings and individual SN features. Though satisfaction and control were not directly impacted by SN features through a specific statistical test, users indicated being highly satisfied with the upload predictions and network speed indicator, which both were perceived highly useful while sharing, waiting and making customizations to the upload. The features can thus be seen as valuable additions to the sharing process and generally adding them makes sense. SN did have a direct significant effect on perceived reliability, which was an important discovery as our research focuses on crowded and unpredictable network environments. Moreover, users are uncertain about their uploads, making reliability a key factor in the sharing process. A good example was the highly rated resumable upload feature, which resumes disrupted uploads from where they left of.

Addressing the second research question, "Is high usability attainable when incorporating a SN architecture and its features into an end-user app?", we confirm that usability is attainable, regardless of the technical complexity of the networking architecture. With a high usability score from the SUS assessment and high pragmatic quality score from the AtrakkDiff assessment, the app received highly positive global post-test ratings. Adding technically complicated networking features (e.g. resource allocation) could negatively impact the usability. Though, by translating them to user friendly interface elements, such as the speed indicator, intuitive recommendation messages and visual customization controls with live feedback, we managed to attain a user-friendly sharing experience and high usability.

Future work could investigate how actual upload times could be improved by improving the system and network itself. Such improvements would cause more satisfaction and

control. Next, deploying the app in a real-world context or on a cellular network is what our research should also aim for in the future. From a video sharing perspective, the ability to send recorded video from user to user might also lead to interesting SN related insights, making users aware of each others networking conditions.

8. ACKNOWLEDGEMENTS

I would like to thank my daily supervisor J.W.M. Kleinrouweler for the useful comments, remarks and as the supplier of the network architecture. Furthermore I would like to thank Dr. P.S. César Garcia as the main supervisor for this project. Also, I like to thank the participants in our experiment, who have willingly shared their time during the user tests.

9. REFERENCES

- [1] Maeve Duggan. Photo and video sharing grow online. *Pew Research Internet Project*, 2013.
- [2] Pierdomenico Fiadino, Mirko Schiavone, and Pedro Casas. Vivisectioning whatsapp in cellular networks: Servers, flows, and quality of experience. In *Traffic Monitoring and Analysis*, pages 49–63. Springer, 2015.
- [3] Arpit Gupta, Jeongki Min, and Injong Rhee. Wifox: Scaling wifi performance for large audience environments. In *Proceedings of the 8th international conference on Emerging networking experiments and technologies*, pages 217–228. ACM, 2012.
- [4] Jan Willem M Kleinrouweler, Sergio Cabrero, and Pablo Cesar. Delivering stable high-quality video: An sdn architecture with dash assisting network elements. *ACM Multimedia Systems Conference*, 2016.
- [5] Jeffrey Erman and Kadangode K Ramakrishnan. Understanding the super-sized traffic of the super bowl. In *Proceedings of the 2013 conference on Internet measurement conference*, pages 353–360. ACM, 2013.
- [6] Carine Lallemand and Guillaume Gronier. Enhancing user experience during waiting time in hci: contributions of cognitive psychology. In *Proceedings of the Designing Interactive Systems Conference*, pages 751–760. ACM, 2012.
- [7] Fiona Fui-Hoon Nah. A study on tolerable waiting time: how long are web users willing to wait? *Behaviour & Information Technology*, 23(3):153–163, 2004.
- [8] Jawaaid A Ghani and Satish P Deshpande. Task characteristics and the experience of optimal flow in human—computer interaction. *The Journal of psychology*, 128(4):381–391, 1994.
- [9] Maria Madsen and Shirley Gregor. Measuring human-computer trust. In *11th australasian conference on information systems*, volume 53, pages 6–8. Citeseer, 2000.
- [10] Yiannis Yiakoumis, Sachin Katti, Te-Yuan Huang, Nick McKeown, Kok-Kiong Yap, and Ramesh Johari. Putting home users in charge of their network. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, pages 1114–1119. ACM, 2012.
- [11] Ericsson. Service innovation through smart networks. pages 1–11, 2014.
- [12] Kuan-Ta Chen, Polly Huang, and Chin-Laung Lei. How sensitive are online gamers to network quality? *Communications of the ACM*, 49(11):34–38, 2006.
- [13] Panagiotis Georgopoulos, Yehia Elkhatib, Matthew Broadbent, Mu Mu, and Nicholas Race. Towards network-wide qoe fairness using openflow-assisted adaptive video streaming. In *Proceedings of the 2013 ACM SIGCOMM workshop on Future human-centric multimedia networking*, pages 15–20. ACM, 2013.
- [14] Guillaume Gronier and Carine Lallemand. How to improve perceived waiting time in hci: A psychological approach. pages 1–6. ACM, 2013.
- [15] Ritu Agarwal and Elena Karahanna. Time flies when you’re having fun: Cognitive absorption and beliefs about information technology usage. *MIS quarterly*, pages 665–694, 2000.
- [16] Sally J McMillan and Jang-Sun Hwang. Measures of perceived interactivity. *Journal of advertising*, 31(3):29–42, 2002.
- [17] Gabe Zichermann and Christopher Cunningham. *Gamification by design: Implementing game mechanics in web and mobile apps*. ” O’Reilly Media, Inc.”, 2011.
- [18] Maurice J Khabbaz, Chadi M Assi, and Wissam F Fawaz. Disruption-tolerant networking: A comprehensive survey on recent developments and persisting challenges. *IEEE Communications Surveys & Tutorials*, 14(2):607–640, 2012.
- [19] Marc Hassenzahl. The thing and i: understanding the relationship between user and product. In *Funology*, pages 31–42. Springer, 2003.
- [20] Usability and design evaluation. <http://attrakdiff.de/>. Accessed: 2016-06-23.
- [21] Tiago Camacho, Marcus Foth, Markus Rittenbruch, and Andry Rakotonirainy. Trainyarn. In *Proceedings of the Annual Meeting of the Australian Special Interest Group for Computer Human Interaction*, pages 455–464. ACM, 2015.
- [22] Adrian Holzer, Bruno Kocher, Denis Gillet, Samuel Bendahan, and Boris Fritscher. Dinerrouge. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*, pages 2187–2192. ACM, 2015.
- [23] JRC Patterson. Video encoding settings for h. 264 excellence. *Technical note, Lighterra, Surfers Paradise*. URL <http://www.lighterra.com/papers/videoencodingh264>, 2012.
- [24] Fred D Davis. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS quarterly*, pages 319–340, 1989.
- [25] John Brooke et al. Sus-a quick and dirty usability scale. *Usability evaluation in industry*, 189(194):4–7, 1996.
- [26] Rob Eisinga, Manfred te Grotenhuis, and Ben Pelzer. The reliability of a two-item scale: Pearson, cronbach, or spearman-brown? *International journal of public health*, pages 1–6, 2013.
- [27] Aaron Bangor, Philip Kortum, and James Miller. Determining what individual sus scores mean: Adding an adjective rating scale. *Journal of usability studies*, 4(3):114–123, 2009.

APPENDIX

A. QUESTIONNAIRES

A.1 Post-task Assessment: Feelings

- Perceived Satisfaction
 - Rate the upload time
 - Rate the upload speed
- Perceived Control
 - When I was using the app, I felt in control
 - I felt I had no control over the interaction with the app
- Perceived Reliability
 - The app provided me with the advise I needed to make decisions
 - The systems performs reliable

A.2 Post-task Assessment: App Features

A.2.1 Only Sharing

- Perceived Satisfaction
 - I liked that the application provided me a recommendation for the video quality
 - I liked the video quality recommendation that the application provided me
 - I liked having information on the upload duration next to the video quality recommendation
 - I liked having information on the network quality while sharing the video
- Perceived Usefulness
 - The video upload time helped me understand the video quality recommendation
 - The network quality indicator helped me to understand the video quality recommendation
 - The network quality indicator helped me to understand the video upload time

A.2.2 Upload Customization

- Perceived Satisfaction
 - I liked that I could set the video quality myself
 - I liked having the option to postpone uploading the video to a later point in time
 - I liked having the option to interrupt the video upload, and to resume it later
- Perceived Usefulness
 - The displayed upload times helped me in picking the video quality
 - The network quality indicator helped me in picking the video quality
 - The displayed upload times could have helped me in my decision to postpone the upload
 - The network quality indicator could have helped me in my decision to postpone the upload

A.3 Post-test Assessment: SUS

- Usability
 - I think that I would like to use this app frequently
 - I found the app unnecessarily complex
 - I thought the app was easy to use
 - I think that I would need the support of a technical person to be able to use this app
 - I found the various functions in this app were well integrated
 - I thought there was too much inconsistency in this app
 - I would imagine that most people would learn to use this app very quickly
 - I found the app very cumbersome to use
 - I felt very confident using the app
 - I needed to learn a lot of things before I could get going with this app

A.4 Post-test Assessment: Attrakdiff

- Pragmatic & Hedonic Qualities
 - Simple - Complicated
 - Ugly - Attractive
 - Practical - impractical
 - Stylish - tacky
 - Predictable - Unpredictable
 - Cheap - Premium
 - Unimaginative - Creative
 - Good - Bad
 - Confusing - Clearly structured
 - Dull - Captivating

B. EXTRA VISUALIZATIONS

Table 4:

Overview of the app feature specific questions for "only sharing" tasks, presented with means and standard deviations (SD's).

Question	Mean	SD
1 I liked that the application provided me a recommendation for the video quality	4.3	0.58
2 I liked the video quality recommendation that the application provided me	4.2	0.6
3 I liked having information on the upload duration next to the video quality recommendation	4.7	0.46
4 I liked having information on the network quality while sharing the video	4.6	0.58
5 The video upload time helped me understand the video quality recommendation	4.3	0.58
6 The network quality indicator helped me to understand the video quality recommendation	4.1	0.94
7 The network quality indicator helped me to understand the video upload time	4.2	0.75

Table 5:

Overview of the app feature specific questions for "customization" tasks, presented with means and standard deviations (SD's).

Question	Mean	SD
1 I liked that I could set the video quality myself	4.8	0.44
2 I liked having the option to postpone uploading the video to a later point in time	4.7	0.73
3 I liked having the option to interrupt the video upload, and to resume it later	4.9	0.3
4 The displayed upload times helped me in picking the video quality	4.3	1.06
5 The network quality indicator helped me in picking the video quality	3.7	1.43
6 The displayed upload times could have helped me in my decision to postpone the upload	4.4	0.68
7 The network quality indicator could have helped me in my decision to postpone the upload	3.9	1.24

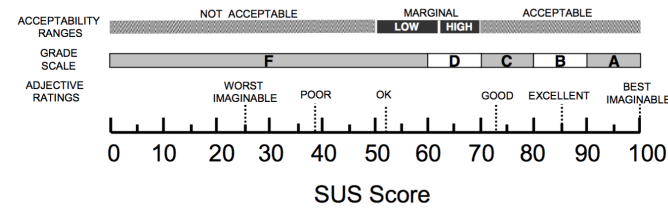


Figure 11:

Meaning of the individual SUS scores: an adjective rating scale [27].

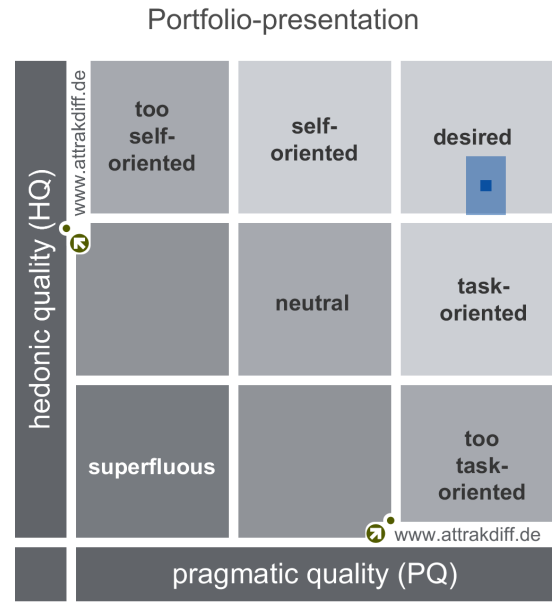


Figure 12:

AttrakDiff results: portfolio presentation with global classification of the app.

C. PRELIMINARY SURVEY

The following questionnaire enquired subjects ($N = 71$) about their smartphone usage habits when sharing video and connecting to WiFi networks. Some of the questions allowed multiple answers.

Video Sharing Habits			
1	Do you share videos with your smartphone?		
	Yes	64.8%	No 18.3%
	Yes, but not my own	16.9%	
2	Who do you share videos with?		
	Family	58.6%	Everyone 34.5%
	Friends	81%	Other 1.7%
3	What apps do you use to share videos?		
	Youtube/Vimeo	36.2%	Skype/Hangouts 15.5%
	Facebook	48.3%	Instagram/Vine 22.4%
	Twitter	19%	Snapchat/Beme 29.3%
	Periscope/Meerkat	0%	Other 3.4%
	WhatsApp/Telegram/iMessage	84.5%	
4	What type of Internet connection do you use to share videos?		
	Telephone provider (3G/4G)	70.7%	Home WiFi 79.3%
	Public WiFi	53.4%	Other 1.7%
	Any WiFi	36.2%	

Video sharing and recording at events			
5	Do you use your smartphone to share videos at events?		
	Yes	50.8%	No 49.2%
6	In which events do you record videos?		
	(Music) festivals	90%	Street festivities 53.3%
	Concerts	60%	Other 0%
	Sport events	60%	
7	If you share these videos, when?		
	I stream it live	23.3%	At home 63.3%
	Right after recorded	76.7%	I don't share 0%
	Later, still at the event	46.7%	Other 0%

Video watching at events			
8	Do you watch videos on your smartphone at events?		
	Yes	50.8%	No 73.2%
	Yes, but only event related	7%	
9	Which videos do you watch at the event?		
	Videos sent to me	84.2%	Live streams 5.3%
	Videos on social media	78.9%	Other 0%
	Videos offered by event	15.8%	

Connectivity			
10	Have you experienced connectivity problems during an event?		
	Often	27.1%	Never 21.4%
	Occasionally	50%	
11	What would you do if a file or a video upload fails?		
	Retry immediately	32.9%	Retry at home 31.4%
	Retry later that day	22.9%	Don't share it 12.9%
12	What would motivate you to use free WiFi at an event?		
	Faster than telephone provider	50%	Event services 17.1%
	Saves my data plan	74.3%	Other 5.7%
13	What network state information would you like to have?		
	If I am connected or not	66.2%	Image quality 15.5%
	My connection speed	60.6%	Location hints 32.4%
	Uploading duration	26.8%	Time hints 23.9%