UNIVERSITY OF OSLO
Department of Informatics

Identification of user tasks in on-line TV

Master's thesis

Thomas Torp

May 2011

Abstract

The research presented in this master's thesis identifies user behavior and user task in context of on-line TV. In addition, the use of the different tools for data collection, data preparation and analysis of all the collected data. To understand how the user will interact with a system provide the service providers of on-line TV the possibility personalize the content to a specific user. Furthermore, it can be used to develop better on-line TV services, create user profiles, understand what the users do (without any interaction with the user) and in general provide the user with better quality of experience (QoE).

We conducted a controlled experiment with five MSc students from the Informatics department on the University of Oslo. The participants were introduced to two on-line TV providers, where they should perform and provide answers to 11 assignments. All their actions and behavior were logged with the help two logging tools: Noldus uLog to record the low-level user actions and Camtasia Studio to record the computer desktop, user behavior from webcam and sound. In addition to the recorded data, background information about the users was collected before the experiment and a interview was conducted at the end to further understand why the user choose to do what they did.

The experiment led to the identification of 15 different user tasks over two categories. The high-level users tasks consist of 4 different user tasks where they all depend on the low-level user tasks. 11 user tasks are categorized as low-level user tasks, where they are focusing on how the user interacts with the on-line TV service to achieve the high-level user tasks. Furthermore, it has also been identified some user types that are closely related to the high-level user tasks.

The results show that it is possible to identify what the users are doing in some extent. However, it will be hard to conclude that the user is watching some content or not, in relation to if the user is present or not. This must be further investigated to find a solution to this problem. Furthermore, it is identified that the user task in on-line TV can in some extent be found in other TV services. The user tasks that differ, is the one that are service specific.

Acknowledgements

First of all, I would like to thank Amela Karahasanovic for her guidance and support as my supervisor throughout this thesis. All her time has been very valuable to me. I would also want to thank SINTEF for providing me with all the needed equipment to carry out the experiment. Furthermore, the support-team and developers at Noldus has been at some help to comprehend and use their products for user logging and data analysis.

I would also want to thank my mother and father for letting me stay home, at no expense, under all my study time on the University in Oslo. Furthermore, I thank my girlfriend for taking the time to read over this thesis.

A special thanks goes to my fellow master students that did take their time to participate in my experiment. Without them the experiment could not be completed.

Oslo, May 2011
Thomas Torp

Contents

List of Figures

List of Tables

Acronyms

QoE	Quality of Experience
QoS	Quality of Service
VoD	Video on Demand
iTV	Interactive television
IPTV	Internet Protocol television
DVR	Digital video recorder
HCI	Human-computer interaction
HTA	Hierarchical task analysis
CCT	Cognitive complexity theory
GOMS	Goals, operators, methods, and selection rules

Definitions

The on-line TV term

There is at this point a range of terms and definitions of what on-line TV is (otherwise known as Internet TV, or Internet television). There is no agreement on a definition. It comes with different names: web TV, Internet Protocol television (IPTV), enhanced TV, personal TV, and interactive TV (iTV), which signify slightly different things. All the different terms will not be defined or discussed, however the term on-line TV will be defined for use in this thesis.

(Noam et al., 2004, p.4) define Internet-TV (that is the same as on-line TV) as:

1. *Internet TV is the adoption of an Internet-like interface in accessing and watching television-a new form of video navigation over the Internet. Internet TV is a more interactive approach to controlling the television experience with the ability to obtain all stores of ancillary information while watching television.*

2. *Internet TV is the use of the home TV set to view Internet sites, as offered by WebTV Networks, perhaps in conjunction with conventional television viewing. These kind of applications of Internet TV create an interactive television experience called Internet-enhanced TV.*

3. *Internet TV is the use of the Internet protocol to store and transmit video, both at the TV studio and also to various locations. Rather than storing and transmitting digital video as a continuous steam of bits, the digital video is packetized into the packets specified by the Internet protocol.*

The on-line TV term that will be used in this thesis is defined as about with the following as a supplement: On-line TV is providing the user with access to video servers that store existing programs and making them available for viewing at any time, which is distributed via the Internet.

Quality of Service (QoS)

The term Quality of Service (QoS) is extensively used today, especially for IP-based multimedia services. Networks and systems are gradually being designed in consideration of end-to-end performance required by user applications. (Zapater and Bressan, 2007) However, do (Evans and Filsfils, 2007) state that the term QoS mean different things to different peoples, hence it should be defined early what it means in this context. Firstly, is service defined as a description of the overall treatment of a customers traffic across a particular domain. Secondly, is quality defined as the underlaying requirements for an application which can be defined in term of jitter, packet loss, throughput, service availability, and per flow sequence preservation. (Evans and Filsfils, 2007, 87) Furthermore, (Lervold et al., 2010) has defined QoS as a service that typically measures certain set of network parameters. They also provide an graphical overview over QoS and QoE that is presented in figure 1.

Quality of Experience (QoE)

Quality of Experience is how the end user is experiencing the program or service one are delivering to him/her. Furthermore, the R2D2 project defined QoE as the overall acceptability of an application or service, as perceived subjectively by the end-user. It includes the complete end-to-end system effects (client, terminal, network, services infrastructure, etc.) and may be influenced by user expectations and context. Rest of this MSc will take use of the definition of QoE in the R2D2 project.
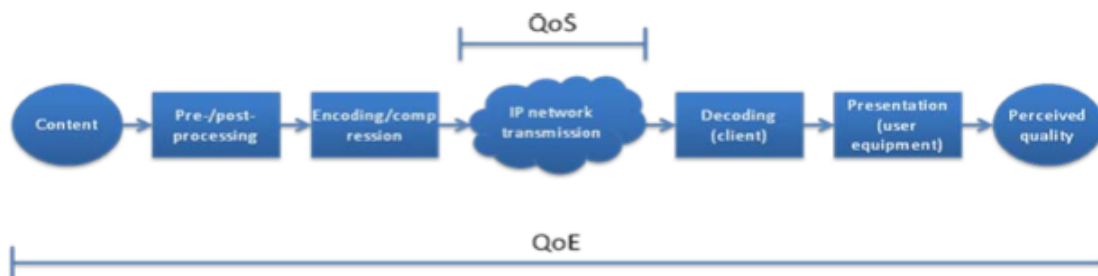
Figure 1: QoS vs. QoE for video streaming (Lervold et al., 2010)

Chapter 1

Introduction

1.1   Motivation

The ways we are watching TV is changing. Users are starting to discover the possibilities with use of time shifting on digital video recorders (DVRs). In the last few years the use has increased significantly. Time shifting is the recording of a program to a storage medium to be viewed or listened to at a time more convenient to the consumer. In the U.S. only did the use of time shifting increase by 17.9% in the third quarter of 2010 in relation to the same period in 2009. A regular American watch nearly 10 and half hours of time shifted TV per month at the end of 2010.(The_Nielsen_Company, 2010) On a DVR device does the user have the possibility to record and watch live shows later on. However, to combine both when and where will give more flexibility. Internet provides the flexibility where a consumer can choose both when and where. Furthermore, to combine this with the TV concept will provide new possibilities for the consumer.

Interactivity is a key element in the field of human-computer interaction, which started to pay attention to the design of user interfaces for interactive TV (iTV) in the mid-1990s.(Bernhaupt et al., 2007) iTV have been on the market for a while, however on-line TV is a relative new service. On-line TV provides the user with an opportunity to choose when and where to see some content. Furthermore, the use of Internet is shortening the connection between users.

Online video content is in a rapid growth and it is getting harder every day for users to find what they are really looking for. Only on youtube.com it is uploaded 24 hours of content every minute.(Youtube.com, 2010) With this amount of content to browse, the users will jump over a lot of content that may be in some interest. To abstract the relevant content and give personalized content back to the user will be an important task, when the content grow.

To understand what users really are going to do when they are browsing content is

crucial. One can provide users better quality of experience (QoE) depending on many factors. When one can predict the user task one can both give a more accurate prediction about what the user wants to watch, however one can also regulate the quality of service (QoS). Furthermore, (Shaikh et al., 2010) have identified that when QoE and QoS is combined in the right way, a service will get more activity.

Numerous studies have been done to predict user behavior and user task on the web, however most of these studies are on web search and news recommendation. These studies are not considering what context the user is in. One cannot know if the user are looking for something specific or just browsing through the content. To understand this, one need to consider the relationship between task and behavior (Gutschmidt and Cap, 2008).

On-line TV and video on demand (VoD) have been investigated in some way to better understand when, how and why people are watching TV on mobile devices (Miyauchi et al., 2009). In the study from (Miyauchi et al., 2009) they identified that users want to make as little effort as possible to find interesting content. Furthermore, they also want to be comfortable when watching the content. From the usage of mobile-TV in (Cui et al., 2007) it is suggested that personal TV is a more accurate description for this kind of service.

To predict the user behavior and user task one need to consider all of these factors. Furthermore, when all of these factors are combined one are able to generate user profiles to individual users. This will give a possibility to give more accurate personal content, better viewing experience and in general increase the users QoE.

One should also consider that almost all studies that are on mobile on-line TV are from Japan, China and Korea. (Miyauchi et al., 2009) (Cui et al., 2007) However, to be able to give the community some study from Scandinavia, more precisely Norway, will provide much better insight in how very modern users interact with on-line TV services. Furthermore, the need for further investigation of on-line TV is high, and will lead to increased understanding of users and users intentions in the on-line TV context.

1.2 Objective

The objective of this research was to identify user tasks in the context of on-line TV with the help of an experiment where users shall complete a range of assignments and all their action will be recorded in a range of ways.

We wanted to investigate (all in context of on-line TV):

- How do users perform tasks when watching, browsing and searching video content?

- Which user task and user behaviors that are taken use of when watching, browsing and searching video content?

- Can user tasks be identified from low-level data?

- How different logging tools can help us to answer the above posed questions?

1.3 Research Method

A controlled experiment was conducted, with five participants. The participants were students in their fifth year, at the Informatics study on the University of Oslo. The participants were introduced to two on-line TV services, one was a TV provider and the second was a newspaper provider that used videos to supplement the articles. On both on-line TV services the participants should carry out a range of assignments. All their actions were logged, video and sound recorded, and verbal protocols were collected. At the end of each session an interview was conducted.

1.4 Research Context

This master thesis is a part of the Road to media-aware user-Dependent self-aDaptive Networks (R2D2). R2D2 is an ongoing 3-year research project lunched in 2009 by a group of eleven small, medium and large companies from telecommunications industry, universities and research institutes: Lund Uversity (LTH), Ericsson AB (EAB), Fundacin ROBOTIKER (ROBOTIKER-TECNALIA), GDT, Stiftelsen SINTEF (SIN), Telefonica Investigacion y Desarrollo (TID), Centre Tecnologic de Telecomunicacions de Catalunya - CTTC, IKUSI (IKU), PRIMETEL (PTL), TELNET-RI (TEL), Uninett (UNI).

The main outcome of R2D2Networks will be a user-dependant media-aware self-aDaptive network management system that will help the network to learn what, when, where and how different services and resources are to be used. It will evolve the current home and access/aggregation networks towards architecture with a better use of the connectivity resources and an optimization of the provided service quality. The proposed solution is based on a resource manager that dynamically optimizes the network resources, and also the Quality of Experience (QoE) of the customers.(Celtic-initiative, 2009)

1.5 Contribution

There are several contributions of this research to the R2D2 project:

1. Findings with respect to earlier research.

2. Creation of a study.

3. Identification of user tasks.

4. General understanding of how users interact with on-line TV.

1.5.1 Findings with respect to earlier research

There is at this time a small amount of related work in the context of on-line TV. There-fore, it is crucial to identify the work that has been done in the context of on-line TV or closely related to the context. Firstly, it has been identified ten articles closely related to user behavior. Secondly, six articles related to user behavior in context of video and lastly, it has been discovered four studies where three have identified user task and user behavior closely related on-line TV. Additionally, the last study has presented another method for identification of user task in the context of video streaming (that is as closely to on-line TV that one can get). This method focus on taking the sessions/requests for a video to identify what the user are doing. E.g. if the user is watching the video for the first time, is fast forwarding, etc. However, this method has not been used in this thesis, duo to creation of a different study design and furthermore wanted to understand the users more accurate with help of observing the users and interviewing them.

1.5.2 Creation of a study

A contribution of this research is the development of a study that includes the design of the experiment, how to prepare the data for analysis and how the data can be analyzed. The experiment (questionnaire, experiment assignments and interview) has been created in collaboration with my supervisor Amela Karahasanovic.

Furthermore, it has been identified how to prepare the data collected in the experiment to include in different statistical programs like The Observert XT and Microsoft Excel. The low-level data is a rich source of information on the participant's actions during the study. We have cleaned and extracted this data to get detailed information of each participant's actions during the assignments. The recorded video have been prepared and coded with a coding scheme. The coding scheme has been developed in an iterative process.

Lastly, it is presented a way to analyze the low-level data with The Observer XT and how one can convert the low-level data to numerical IDs based on the different actions/events that is logged by the Noldus uLog. It is further presented how the combined uLog data and coded video can identify what the users are doing.

1.5.3 Identified user task

Another contribution of this research is the identified user task. It has been identified 4 high-level user task and 11 low-level user tasks, where the all high-level user tasks depends on the low-level user tasks. The low-level user tasks is the interactions with the

on-line TV service and the high-level user tasks is what the user want to achieve with the low-level user tasks. To understand how the high-level user tasks depends on the low-level user tasks a diagram has been created to illustrate the relationships.

1.5.4 General understanding of how users interact with on-line TV

In addition to the identified user task it has been identified some user types (watching and searching user type) and different affects on the user behavior. There is identified three affects that make the users operate in another way than they normally would do. How the Internet page design that embeds the video player, how the user normally find content (e.g. searching or clicking hyper links) and the length of the video. The experiences in how user is interacting with the on-line TV service should be included in later studies to be able to collect better data and create better results.

1.6 Chapter Overview

Chapter 2 Related work	This chapter presents an overview of the related work. The identification of related work includes information on user task, user behavior and on-line TV. Some of the identified related work is not directly related to on-line TV. However, the research identified on user task and user behavior is mostly in an Internet context, which is close to what we are looking for. Four articles are presenting some user task / user behaviors in the context of on-line TV.
Chapter 3 Method	This chapter gives a detailed description of the controlled experiment. The experiment is designed by me and my supervisor, Amela Karahasanovic.
Chapter 4 Data preparation and tools	This chapter describes the data preparation and the used tools in the study. Furthermore, the coding scheme are presented.
Chapter 5 Analysis	This chapter describes the analysis of the data collected by uLog, screen capture and the interview.

Chapter 6 Results	The results of the study are presented and discussed in this chapter. Firstly, the identified low- and high-level user tasks are presented. Furthermore, problems identifying the high-level user task from the low-level user tasks are presented. Thereafter, two types of users are presented. Additional home versus experiment is presented with some statements on how participant operate differently in the experimental environment. Finally, some time specific findings are presented and the results is discussed.
Chapter 7 Validity	The most important threats to the validity of the experiment are discussed in this chapter.
Chapter 8 Conclusions and Future Work	This chapter summarizes the contribution of the research and suggests further research.
Appendix A Background questionnaire	Questionnaire used to state the participants' background: their education, online TV experience and experience with social medias.
Appendix B Interview guide	Guide used as basis for interviews with the participants.
Appendix C uLog settings	The uLog settings used to log the participants actions on the computer.

Chapter 2

Related work

This chapter gives an overview of empirical studies conducted in the field of understanding user task and user behavior in an on-line TV context. The method to identify related work to understanding user task and user behavior in on-line TV context is described in section 2.1. Section 2.2 gives an overview of empirical studies on user behavior in general and user behavior in video context. Section 2.3 gives an overview of user task identification methods and some identified user tasks. Section 2.4 gives an overview of how user profiles are generated and how user task and user behavior is used to create user profiles. Section 2.5 gives an overview of data collection methods and analysis. Section 2.6 gives an overview of the methods for identification of low-level data. Section 2.7 will summarize this chapter.

2.1 Identification of Related Work

To identify the related work I have searched in digital libraries and databases.

In order to find related work I searched the following digital libraries and reference databases: ACM Digital Library, IEEE Xplore, Springer Link and DUO. I have also used the search engines Google and Google Scholar. The keywords used during the search have been combined in different expressions and are the following:

- User
- Task
- Behavior
- Predicting
- Video / on-line TV
- Analysis

- Low-level data

The search was initially performed in October 2010 and repeated in January 2011. The initial search on user behavior and user task gave about 31,000 results. By adding quotes around the search words, the number of results was considerably reduced (31,000 to 3200). I filtered out the irrelevant articles, conference proceedings and theses by first reading the title and then the abstract. Additionally have references in relevant articles been used to find other relevant articles. This work resulted in 10 articles related to User Behavior, 6 articles related to user behavior in the context of video and 4 articles closely related to User Task in TV. Hence no relevant articles were found on on-line TV, except content that has been identified from the R2D2 project.

2.2 Empirical studies on user behavior

Behavior of users have been investigated for a long time. As early as in 1983 (Gould and Lewis, 1983) did state that all computer systems need early focus on users and tasks.

> *First, designers must understand who the users will be. This understanding is arrived at in part by directly studying their cognitive, behavioral, anthropometric, and attitudinal characteristics, and in part by studying the nature of the work expected to be accomplished.* (Gould and Lewis, 1983)

Several studies have been conducted to identify user behavior on the Internet. Table 2.1 and table 2.2 summarize the findings. Table 2.1 gives an general overview and table 2.2 an overview with more focus on on-line TV. Table 2.1 list studies together with description of participants. Furthermore, the third column describes type of the studies, some with detail data. Finally, we give the purpose of the study.

Study	Participants	Type of study	Purpose
(Montabert et al., 2009)	14 HCI intro students	Questionnaire, User Study	Requirements analysis, task modeling
(Szabo and Huberman, 2010)	Digg.com:560000 Youtube.com: (not described)	Logs. Digg.com: 29mill diggs and 2.7mill submissions. Youtube.com: View count series of 7146 Videos daily	Predicting the popularity of online content
(Gutschmidt and Cap, 2008)	20 University students and employees in different departments 26.6 old	Logs (Firefox plugin)	Investigate the relationship between the task and the behavior of Internet users. Find out which attributes of the behavior are essential for a differentiation of user tasks.
(Attenberg et al., 2009)	4000000 anonymous users	Logs(browser plugin) Over one billion individual page requests	Identifying patterns in user behavior and predicting user behavior.
(Liu et al., 2010)	Part1: 16848 anonymous users from 10 different countries and regions. Part2: 10000 anonymous users.	Logs	Develop an news recommendation system in google news.
(Adar et al., 2007)	(not described)	Logs (MSN: 15mill query's and clickthroughs. AOL: 12.5mill query's. BLOG: 15mill posts on 3mill blogs. CNN.com: 950 articles. BBC.co.uk: 12000 articles. TV.com: 2457 shows with 2 pages per show.)	Understanding past behaviors and predicting future behaviors.

(Hassan et al., 2010)	Random 1000 user sessions. (1 session = 3 days)	Logs. Yahoo search engine.	Address problems of predicting, modeling user behavior.
(Cheng et al., 2010)	Only US users	Logs(browser plugin) 3 billion anonymous page views	Actively predicting search intent. Create a series of technologies to fulfill this task
(Maia et al., 2008)	Youtube subscription network users	Logs (web crawler) Top 100 from web crawler. Uploads, videos watched, date joined.	Create an methodology for characterizing and identifying user behavior in online social networks.
(Agichtein and Zheng, 2006)	(not described)	Logs. 7000 labeled queries. 1.2mill search requests and 10mill user interactions from the requests	Propose an effective approach of leveraging millions of past user interactions. Predicting.

Table 2.1: User behavior in general

Almost all of the studies depend on raw user data in the form of different logs. Only (Montabert et al., 2009) did include users with the help of questionnaires and user study. (Gutschmidt and Cap, 2008) have some interacting with users, however the study focused on data logs.

It was only two studies (Montabert et al., 2009) and (Gutschmidt and Cap, 2008) that was done in a controlled environment, rest of the studies have got their content from some service and content providers e.g. yahoo (Hassan et al., 2010), youtube (Szabo and Huberman, 2010) (Maia et al., 2008) and some other providers. The data was in some cases logged with the help of plugins[1] to the browser. Plugins provide enhanced logging options, in relation to the logs that have been pulled from the server side. However, a problem with using plugins is that the user must accept to use the plugin and accept to be logged.

Furthermore, almost all studies in an Internet context depend on some kind of data logs, that is analyzed to identify user behavior. This is may then implying that in an Internet context you can depend just on the analysis you get from data logs. However, it should be said that it depends on the service, because people behave different if they are on a news page or a social network page. Since the users are not included in the analysis or study questions like what the user actually wanted with an request will be unanswered.

---

[1]A plugin is a set of software components that adds specific abilities to a larger software application.

Table 2.2 list studies together with number of participants. Furthermore, column 3 describes type of the studies, some with detail data. Finally, we give the purpose of the study.

Study	Participants	Type of study	Purpose
(Vilas et al., 2005)	(not described)	Case study, Logs. 15000 requests	To improve the service configuration and content selection. Help to develop service models for VoD services.
(Miyauchi et al., 2009)	11 Persons 15-46+ old	User study, Interview	To find out when, how and why. Try to identify issues from different types of users and devices.
(Baluja et al., 2008)	5.2 mill anonymous users from one geographical place	Logs. 29mill video views	To create a method to provide personalized video suggestions for users
(Acharya and Smith, 2000)	Students on university network	Logs. 6 months of data	Do accesses to videos exhibit temporal locality? How frequently are videos accessed compared to HTML documents? Do users exhibit any specific browsing patterns when viewing videos?
(Benevenuto et al., 2008)	Youtube users	Logs(web crawler) Collected data	Understand user behavior in social network created essentially by video interactions.
(Cui et al., 2007)	4 male and 4 female users	Contextual- and in-dept interviews and observation	Given the difference between the mostly stationary television viewing habits and the inherent portability of the mobile phone use what kind of user experience is possible?

Table 2.2: User behavior in video content

The difference between the studies that have focus on Internet versus focus on video, is that the video studies depends more on the users. Furthermore, studies on video content take use of user studies (Miyauchi et al., 2009), observation (Cui et al., 2007), interviews and logs (Benevenuto et al., 2008) (Acharya and Smith, 2000) (Baluja et al., 2008) (Vilas et al., 2005). However, a wider range of methods is used to understand users when it is video content that is used. This may be because a video can contain a lot of information and it's harder to understand what a user wants with the video content.

When comparing usage of regular text content (like browsing news pages) to the use of video content is there a lot of differences. News pages have a lot of text and the user will be more active on the page. The searching, reading and link clicking will make the user more active. Furthermore, this gives the possibility to understand in more detail if the user found the content they were looking for. However, the user may stay watching specific content longer then text content, because video content contains a lot more information. The interaction with the users will give the possibility to understand what they really want to achieve when they browsed some video content.

2.3 User Task

The international organization for standardization gives the following definition of a task: *"Tasks are the activities undertaken to achieve a goal."* Furthermore, a goal is defined as: *"an intended outcome".*(ISO, 1998) When a user is browsing the Internet to watch some TV, it is important to understand the intention. Did the user fail to watch the original program, just browsing through some online content or trying to achieve some information about a specific subject? This is just a small range of possible user tasks, and how to identify these user tasks can be difficult.

Identification of user tasks has been one of the biggest challenges for the human-computer interaction (HCI) community.(Crystal and Ellington, 2004) (Rauterberg and Fjeld, 1998) The understanding of how users are going to achieve a certain goal, depending on what actions path the user chose to use, can be very helpful to identify what they will do. Furthermore, it can help to create a new system or even just updating an existing system to users needs. To identify user tasks one need to collect a range of data to analyze. This can be done in several possible ways and some of these are cognitive approaches. Cognitive approach is to understand the user in a mental way or to focus on the actions that are used to complete a task.

Furthermore, one need to analyze the data one have collected. (Crystal and Ellington, 2004) have walked through the field of task analysis in an HCI context. They have critically reviewed task analysis models and techniques. They chosen to split it up in three parts: Technical Methods that include Hierarchical Task Analysis (HTA), Conceptual Methods that include Cognitive Task Analysis (CTA), Cognitive Complexity Theory (CCT), GOMS and Work Process Methods that include Activity theory. One problem

they have identified is that the methods have become increasingly complex over the time, therefore they are just ignored in practice. The problem lies in the trade off between efficiency and effectiveness. Many new approaches have been proposed, however most of them have lack in empirical evaluation. Another problem with task analysis methods is mention in (Daabaj, 2000). They have tried out a variety of task analysis methods and the problem is that most task analysis methods have a number of weaknesses. The biggest problem identified was to know when to stop the analysis process.

2.3.1 User task identification

User task identification has been performed in (Kunert and Kunert, 2009), (Darnell, 2007) and (Bernhaupt et al., 2007). These studies have been done in context of either Interactive television(iTV) or Digital TV. Table 2.2 list studies together with number of participants. Furthermore, the third column describes average age. Finally how the general information about the participants is gathered and the method of the study.

Study	Participants	Age	General information	Method
(Kunert and Kunert, 2009)	54 (27 female, 27 male, 8 focus groups)	$\sim 25,6$	questionnaires	Focus group
(Darnell, 2007)	10 (7 female, 3 male)	$\sim 40,9$	questionnaires	Video-, sound recording and interview
(Bernhaupt et al., 2007)	11 (5 female, 6 male)	$\sim 30,6$	questionnaires	Usability lab and field study

Table 2.3: Studies on how to identify user tasks in a TV context

(Kunert and Kunert, 2009) was using focus groups to identify tasks in iTV applications. Focus groups are a kind of group interview and is described as: "A focus groups is a structured group discussion about a specific topic, moderated by a trained group leader". They also mention that focus group is a method that often is used for new technologies when little is known about users needs and wishes.

(Darnell, 2007) was recording users in their own home. They had three 2-hour visits to the participant's home. During the first visit the equipment was set up in the home and the participant were interviewed. On the second visit the recording was reviewed and analyzed by the experimenter and participant. The third, and last visit, the equipment was removed and the recording was reviewed and analyzed again.

The usability lab in (Bernhaupt et al., 2007) was created to look like a normal living room. The room was filled with comfortable furnitures, sweets, drinks and magazines

on a coffee table. This will give a more non-artificial environment and will overcome reported shortcomings when testing iTV. They also did have a field study in October 2004 where 80 households were selected. All interaction was stored in a database as sessions. The study in general has more focus on usability then actual identification of tasks. However, it is still relevant because it has much information about what kind of tasks that is hard for the users to complete and how they interact with the system.

2.3.2   User task identification from low-level data

(Qiu and Cui, 2010) has identified that it is challenging to gain insightful knowledge of the characteristics of user behavior from low-level data in context of online video streaming. It is hard to model user behaviors because of:

- Extraction of behavior information usually requires tedious effort on mining the video trace datasets, which often contains huge amount of data.

- The high volatility if user behavior makes it difficult to identify the behavior patterns and study the relationship between them.

(Qiu and Cui, 2010) used a database from MSN video, consisting of trace records from about 5.000.000 video streaming sessions of year 2007 and 2008. The raw database was parsed to remove all redundant information, and the items that was left for later use was: video ID, source data rate, video length, player ID, sessions start time, butter time, buffer count and session duration. Buffer time is the total time spent on buffering the video file during a session. Buffer count is the number of interrupts duo to data buffering during a session. It has been identified seven behavior patterns:

1. Start (B0): the beginning of an access to a video file. The user has never accessed this file before. Here by "access" we mean a sequence of behavior patterns corresponding to a group. In other words, a group of sessions belong to the same "access".

2. Stop (B1): the end of an access to a video file. After this behavior, the user will not re-access the file during the monitored period;

3. Jump forward (B2): compared to the end point of last session in the timeline of the video, the user's video player jumps a certain distance towards the end of the video.

4. Jump backward (B3): compared to the end point of last session in the timeline of the video, the user's video player jumps a certain distance towards the beginning of the video file.

5. Replay (B4): the user finished playing the video before this session, and he/she restarts to play the video from the beginning point.

6. Pause (B5): the user's video player starts from the time point where it stopped in last session.

7. Return (B6): the user returns and keep playing the video.

Furthermore, the most interesting part is where they presents how these user behavior patterns can be used to identify what the user is doing with only investigate the access time to a video. In Figure 2.1 there is an illustration of some user behavior patterns.
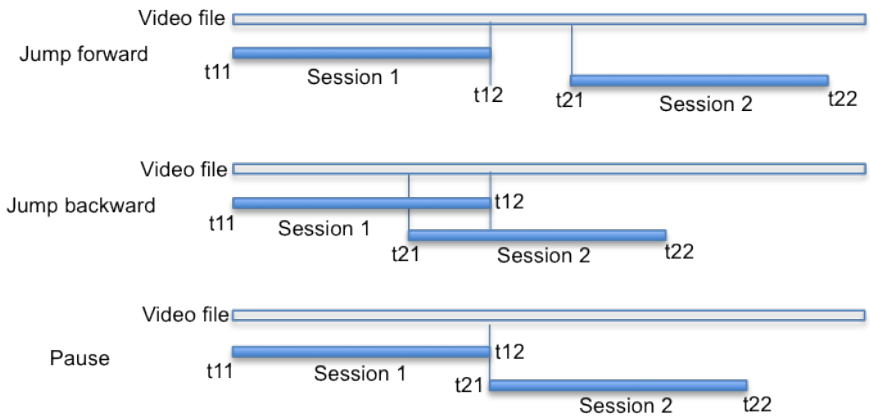
Figure 2.1: User behavior patterns

When the user behavior patterns is defined it is possible to translate between session groups and user behavior patterns. Figure 2.2 illustrates an example of translating a group of sessions info user behavior patterns. The "start time" is the starting point of a session in the time line of the video and the "end time" is the end point in the timeline.
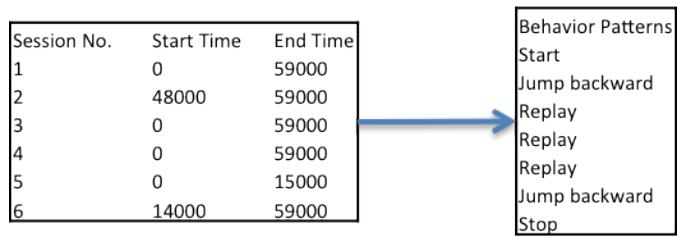
Figure 2.2: Mapping session records to behavior patterns

2.3.3 Findings

(Kunert and Kunert, 2009) identified 40 generic user task. All can be seen in Table 2.4.

Generic user task category	Generic user task
Using application basic	Starting the application
	Exiting the application
	Hiding/showing the application
	Going back to the start page
Viewing content	Accessing a content item
	Viewing a content item
	Switching between content items
Searching	Searching for specific content
	Searching for specific function
Participating	Voting/answering a multiple-choice question
	Allocating items
	Completing text
	Sending the answer
	Influencing a program characters actions by sending ideas for the dramaturgy of a program
	Participating in a "virtual participant" in a program
Communicating	Communicating with participant of a program
	Communicating with known users
	Communicating with the broadcaster
	Communicating with the moderator
	Communicating with unknown users
Entering text	Entering text
Betting	Betting
Shopping	Browsing for products
	Searching for specific product
	Selecting a product
	Ordering a product
Getting help	Using help chapter
Personalizing the application	Personalizing transparency
	Personalizing font size
	Activating captions
	Activating sign-language interpreter
	Changing text or audio language
	Changing ticker speed

Table 2.4: Generic user tasks

(Darnell, 2007) combined observation and interactions with the system. All identified user task are listed in Table 2.5. All user tasks were not included in the article, however just shown as "most frequent high-level behaviors" and "most frequently-used functions".

System	Task/Behavior
DVR - high-level	Skip ads
	Pause
	End of show - find another
	Boring scene - change channel
	Begin watching TV - find show
	Double take - replay
	Skip boring content
DVR - functions	Fast forward
	Pause
	Go to recording
	Play recording
	Go to guide
	30 sec skip
	Delete recording
	Tune from guide
Non-DVR - high-level	Find something to watch during ads
	Chat with co-viewer during ads
	Leave room during ads
	Boring scene - change channels
	End show - find another
	Find ad is preview - watch it
Non-DVR - functions	Punch channel in fullscreen TV
	Go to guide
	Recall (previous channel) fullscreen TV
	Arrow in guide
	Exit guide
	Punch channel in guide
	Tune from guide
	Adjust volume
	Surf in fullscreen TV

Table 2.5: Darnell User Task/Behaviour

(Bernhaupt et al., 2007) found that the system design in iTV has a lot to say if the user tasks would have a successful outcome. To improve the user task success rate the design approach should be user-centered. Furthermore, it should be taken into account the characteristics, experience, expectations, and preferences of the various user

groups.

As mentions earlier has (Qiu and Cui, 2010) identified seven user behaviors / user task that is closely related to how users interact with an video player. All the identified behaviors are possible to identify from the low-level database sessions and they do know what the users is doing when entering some video content. Furthermore, they also did identify that one user do just access one video file once.

In the MSc. thesis done by (Rogne, 2010), within the R2D2 project, some user tasks related to on-line TV have been identified from a questionnaire with 2401 responders (Table 2.6).

Task (watching)	Procentage (100% together)
Specific show	57%
Other	13%
Missed show	11%
Check the program	6%
Smaller categories	6%
Serie preview	4%
Watch previous shows again	3%

Table 2.6: User tasks from earlie R2D2 experiment

More then half of the respondents are using on-line TV to watch a specific program. Furthermore, 11% was on the service because they missed a show, 6% check the program, 6% smaller categories (No access to the TV channel, Subscription, No access to TV, No TV or TV is occupied), 4% to watch sneak preview, 3% to watch a program again and 13% others that don't fit the other categories. All categories is in some way linked to a specific program and it looks like users are not browsing content when they know what they want to watch on the online TV.

2.4 User Profile

User profile[2] in general depends on either explicit or implicit techniques to generate comprehensive user profiles. Most generation of user profiles have been depending on the user adding related information about preferences. However, they are now starting to combine both explicit and implicit techniques. (Takama and Muto, 2007).(Shin et al., 2009) Figure 2.3 shows how they combine explicit information with implicit information. (The diagram was created from diagram from (Shin et al., 2009) and (Chen and Teng,

---

[2]A user profile is a profile about a user with different kind of data. This can be both personal information like email and computer generated information like what kind of content you normally search for.

2009)). The user provides the necessary personal information and the system logs the user behavior. A profile is generated and the content provider has now the opportunity to provide the user with more specific content.
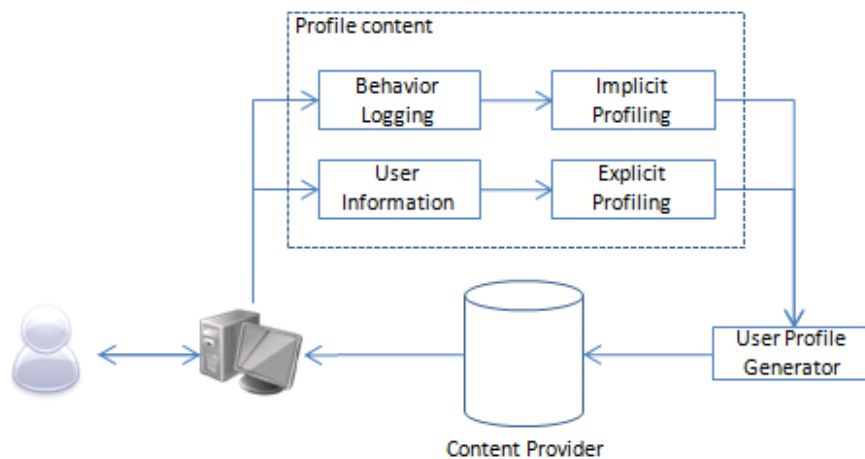
Figure 2.3: User profile generation scheme.

In (Chen and Teng, 2009), (Shin et al., 2009) and (Takama and Muto, 2007) they use explicit information provided by the users and implicit user behavior patterns to find out what the user actually wants. The user behavior patterns are logged with the program title, program type, the time user spent watching the program, the total program time and the watching date. When combining these variables with the information provided from the user, it will be possible to generate user profiles.

Identification of user profiles is done by measuring specific variables. The typical way is to use watch time and attributes of TV programs. The watch time can be pulled from the respective servers and the attributes about a TV program can be found in the meta data. If a user switch from one channel to another the program was uninteresting. If the user continues to watch the channel/program it is assumed that it was interesting.(Takama and Muto, 2007) (Chen and Teng, 2009)

2.5 Methods for data collection and analysis

Different kind of data, e.g. information about a participant or what the participant does, can be collected either manually by the participants filling out forms, manually supported by tools, in interview, or automatically by tools. (Wohlin et al., 2000, 79)

An advantage of using questionnaires is that it does not require much effort for the experimenter, since the experimenter does not have to actively take part in the collection. A drawback is that there is no possibility for the experimenter to directly reveal inconsistencies, uncertainties and flaws in the forms etc. This type of faults cannot be

revealed until after the data collection or if it is a interview, where the experimenter has the possibility to communicate better with the participants during the data collection. A drawback is of course that it requires more effort from the experimenter. (Wohlin et al., 2000, 79) and (Yin, 1994) gives an overview over strengths and weaknesses in different kind of sources. Table 2.7 gives an overview of the six major data sources.

Source of evidence	Strengths	Weaknesses
Documentation	**Stable** - can be reviewed repeatedly. **Unobtrusive** - not created as a result of the case study. **Exact** - contains exact names, references, and details of an event. **Broad coverage** - long span of time, many events, and many settings	**Retrievability** - can be low. biased selectivity, if collection is incomplete. **Reported bias** - reflects (unknown) bias of author. **Access** - may be deliberately blocked.
Archival Records	[Same as above for documentation] **Precise** and **quantitative**	[Same as above for documentation] **Accesibility** due to privacy reasons.
Interviews	**Targeted** - focuses directly on case study topic. **Insightful** - provides perceived causal inference.	Bias due to poorly **Constructed questions** . **Response bias**. **Inaccuracies** due to poor recall. **Reflexivity** - interviewee gives what interviewer want to hear.
Direct Observation	**Reality** - covers events in real time. **Contextual** - covers context of event	**Time** - consuming. **Selectivity** - unless broad coverage. **Reflexivity** - event may proceed differently because it is being observed. **Cost** - hours needed by human observers
Participant Observation	[Same as above for direct observations]. Insightful into **interpersonal** behavior and motives	[Same as above for direct observation]. Bias due to **investigators manipulation** of events.

| Physical Artifacts | Insightful into **culture** features. Insightful into **technical operations**. | **Selectivity** **Availability** |

Table 2.7: Six Sources of Evidence: Strengths and Weaknesses (Yin, 1994, 80)

Furthermore, (Yin, 1994) has identified three approaches to data collection:

- Use of multiple sources of evidence. When this approach is used, it allows an investigator to address a broader range of historical, attitudinal, and behavioral issues. Thus any finding or conclusion in a case study is likely to be much more convincing and accurate if it is based on several different sources of information. (Yin, 1994, 92)

- Use of case study database. When this approach is used every case study project should strive to develop a formal, presentable database, so that, in principle, other investigator can review the evidence directly and not be limited to the written reports. In this manner, a case study database markedly increases the reliability of the entire case study.(Yin, 1994, 95)

- Use of a Chain of Evidence. When this approach is used it will allow an external observer - the reader of the case study, for example - to follow the derivation of any evidence from the initial research questions to the case study conclusions. More over, this external observer should be able to trace the steps in either direction (from conclusions back to initial research questions or from questions to conclusions). (Yin, 1994, 98)

All should strive to achieve these approaches, however I will state that the most important part is approach number two. This will as it states, give investigators directly access to the evidence and data collected. This may give other studies better insight in the study and will make it easier to reuse the material. Furthermore, it can give other researchers, which are investigation, other aspects of the same topic, data that they may want to reuse in their research. This will cause the divergence to be shorter to perform the same type of experiments. Of course it is some problems when companies don't want to share the data which has been collected from their services to the public domain. This should be respected, however the data should be shared for use by other researches when it's possible.

2.5.1 Automatic collection of data

Automatic collections of data do not need any interaction to be collected. However, an issue to take into consideration is what type of low level data that will contribute best to

understand the user task. In the context of Internet (Gutschmidt et al., 2008) conducted a study on what kind of user behavior data best could identify the user task. The first result of the exploratory study gives a hint that small attributes could give a better understanding of the user task. Small attributes stands for average page view duration, number of page views per minute and the number of different URL requests.

(Yuan et al., 2008) has also investigated how to automatic identify user tasks and concluded with the following: Firstly, anchor-based features are more effective to identify user tasks than click-through-based features. Where anchor-based is e.g. the attribute name in a link and click-through is how the user got to a page depending on link clicks; secondly, the number of sites is more reliable than the number of links. Where number of sites that links to an destination URL and number of links is all links on one page linking to the destination URL; Thirdly, click-distribution-based features are more effective than session-based ones. (Lee et al., 2005) found out that a combination of past user-click behavior and anchor-link distribution can identify the goal of a query. The experiment showed that they could identify goals from 90% of the queries.

However, automatic data collection may delay the interaction with the user, and does influence the result. If the data collecting programs uses to many resources on the compute, the system will be less responsive. (Kantner, 1998) in 1998 identified five difficulties when recording behavior on the web:

- Users of Web sites can take numerous paths

- Links and buttons that have similar names but different destinations

- Users can traverse many individual web pages

- Web-page users often cycle through pages repeatedly

- Recording detailed behavior on dynamically generated pages

Even though it is related to delay it should be considered. The Internet is a huge portal for content and therefore it is possible for the user to take a range of paths to some content. Furthermore, since the Internet contains a range of Internet pages different type of elements will be named the same. This can make it difficult to understand what was, e.g. clicked. The Internet is also in an active alteration and it will be hard to have a good overview of the content on one Internet page.

2.6 Low-level data

Low-level data is typically data on user actions with the focus on keystrokes, mouse movement and other devices.(Renaud and Gray, 2004) (Thomas and Karahasanović, 2009) (Coman, 2007) There is a number of tools to captures these kind of actions and (Hilbert, 1998) have examined computer-aided techniques used by HCI practitioners and researchers to extract usability- related information from human-computer interaction

event. Furthermore, (Renaud and Gray, 2004) have claimed that these techniques are mostly aimed at the usability of evaluation of one specific application.

One of the main issues of using low-level data is that it can come from a lot of different sources and grow in complexity. In (Torii et al., 1999) the authors did propose the Ginger-2 Computer-Aided Empirical Software Engineering framework for logging purposes. It consisted of the following type of logs: emacs commands, windows commands, mouse movements, keystrokes, motion measurements, skin resistance level measurement, eye tracking and video recording. This generates a large quantity of data that can be difficult to process.

(Renaud and Gray, 2004) have experienced that interpreting low-level data was a non-trivial task. They used a system called GRUMPS, that is a system that collects low-level data and can deliver it to the investigator. Furthermore, (Thomas and Karahasanović, 2009) have used the GRUMPS system to collect data in four studies. In total was 6.5M user interactions from 238 participants was collected. The quantity of data can do the analysis process hard and demanding. Furthermore, it is also difficult to relate low-level command logs with users intentions (Karahasanović et al., 2005).

It will be very efficient to have tools that can clean and understand the low-level data. (Thomas and Karahasanović, 2009) have done further analysis of how to clean and use the data. In the figure 2.4 it is showed how to clean and interpret the data.
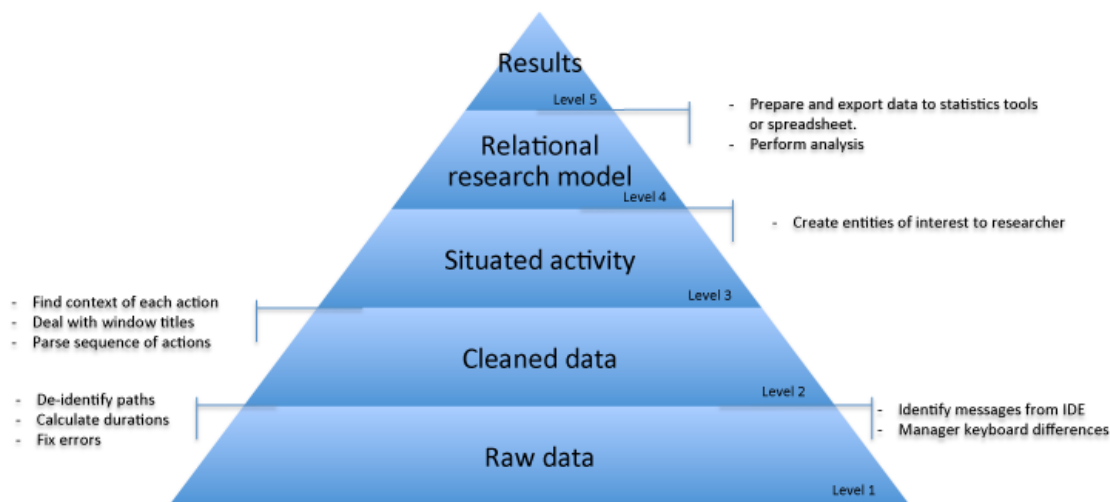
Figure 2.4: Data transformation pyramid (Thomas and Karahasanović, 2009)

Short described level 1 is the collected data with different information, e.g. mouse click with a respective timestamp. Level 2 is cleaning the data of corrupt data, remove data that can reveal the users identity (username and password) and other bogus data. Level 3 deals with finding the context of some actions, like where the keystroke occurred. On Level 4 the researcher must create special purpose entities to facilitate the analysis of a

group of studies/experiments. On level 5 the data are ready to be imported to different statistics tools.

2.6.1 Coding

Coding is a technique used by researchers for organizing, managing and retrieving data. The technique is a systematic way in which to condense extensive data sets into smaller analyzable units through the creation of categories and concepts derived from the data.

In most cases are values extracted from qualitative data to quantitative variables (often collected from observation or interview) in order to perform some type of quantitative or statistical analysis. (Seaman, 1999, p. 563) An example of coding qualitative data to quantitative data can be:

"Tom, Shirley, and Fred were the only participants in the meeting"

will be coded to

num-participants = 3

One issue about the coding is the loss of some information (the name of the participants). Furthermore, the coding scheme should be well formed and reviewed until the information loss is as minimal as possible. (Seaman, 1999, p. 565)

Coding video data of human behavior are difficult to analyze, as this type of information does not lend itself to automation. The coding scheme is different from coding text. For example, a researcher analyzing videos of pair programmers might want to note occurrences of disagreements between the programmers. One of her code categories would be "disagreements". (Burr, 2006) If a disagreement occur, it will be noted in the video time line that a disagreement occurred.

(Jacko and Stephanidis, 2003, p. 479) identified that it is surprisingly few theoretical constructs and approaches for behavior coding in usability context. It was indicated by the survey that the ISO standards 9241-10 and 9241-11 could serve as a constructive foundation. Furthermore, since it is a standard it should, in some range, be used to create the coding scheme.

2.7 Summary

User behavior and user task have been investigated a lot in an Internet context, however one need more research on these subjects in an on-line TV context. If one can get a good overview over how to identify user behavior and user task, can one predict and provide better content to the users. It could benefit the service providers if they understand users better and then can provide a better Internet service when needed.

Many ways are used to identify user behavior in the context of Internet. However, it is not done a lot of research in on-line TV context. (Kunert and Kunert, 2009) and (Darnell, 2007) identified a good range of tasks that a user may do when using iTV or digital TV. The experience and identified user tasks may apply over to the perspective on on-line TV. However, almost all previous studies have depended on the users stating what they wish to do. Only (Qiu and Cui, 2010) has done some user behavior identification without the involvement of users. They take use of session that included the user id, date, time stamp where the video started and ended. With these elements they identified that it is possible to identify if the user was starting to watch a new video, replaying, jumping forward, jumping backward, pausing and stopping the video.

Generation of user profiles can be done in two ways, with explicit and implicit information. Explicit information will come from the user and the implicit information will come from user behavior and user tasks. User behavior and user task must be logged and in some way be analyzed to create the right user profile for one user.

With this research I intend to extend the previous work by using low-level data to identify user behavior and user tasks in context of on-line TV. Earlier studies have been using much interaction with the user, however, there are no studies using low-level data for identification of user tasks. Furthermore, some information will be collected with pre-questionnaire and interview after the experiment. The identified user behavior and user task will serve as a foundation in further research and development of on-line TV services. The user behavior and user tasks will be used later in the R2D2 project to create user profiles. User profiles can be used to provide the users with e.g. independent content, different QoS depending of the context of service use and much more.

Chapter 3

Method

This chapter gives a detailed description of the experiment done in a controlled lab environment in the pilot-study. Amela Karahasanovic and my self designed the experiment. The description of the experiment, data collection methods, data collection tools and assignments are described in chapter 3.1-3.3, 3.4 and 3.5.

3.1 Experiment Overview

The experiment was hosted in a room at the department of informatics at the university of Oslo. The surrounding environment was design to give the participants a feeling that they are in an home environment. Furthermore, it was assigned one observer to one participant, that was going to observe how the participants behaved. The observer was sitting in the same room as the participant.

All the test computers are connected to the Internet through wifi . All data was, after an observation, transferred over to the observer computer for coding and analysis. The experiment setup can be seen graphical in figure 3.1. It illustrates the participant with a test computer and one person that should be observing the participant. At the end of the experiment the data collected will be transfered to the observer computer.

The main outcome of the experiment is to collect a range of data of users using on-line TV. The data will help to identify user tasks and to find search strategies in the context of on-line TV.
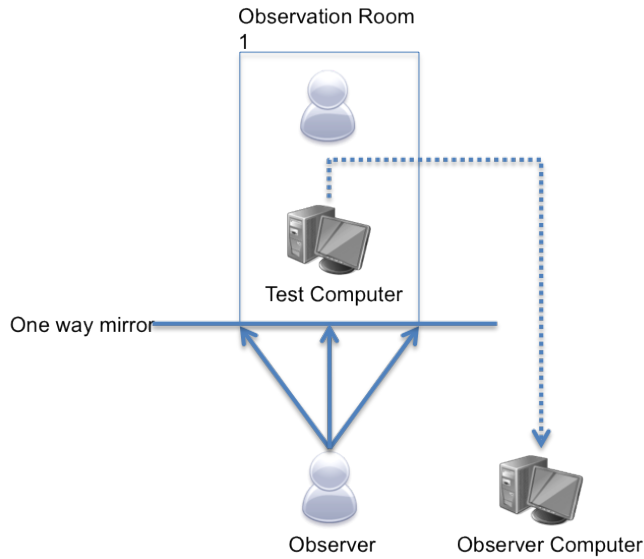
Figure 3.1: Experiment Design.

3.2 Participants

The participants in this experiment were five MSc students studying Informatics at the fifth year on the Department of Informatics on the University of Oslo. The experiment was conducted in a office at the Department of Informatics at the University of Oslo over one-hour sessions, where one participants was present in the office at one time. All participants were paid for their time in the experiment.

Table 3.1 presents the participants background and their experience with on-line TV. All of the participants were male, and the age was in the range 23-26. They had some experience with using on-line TV, where over half of them used it more than once a month. Furthermore, all participants, except one, used social media on a daily basis. They were all used to share content on social medias.

Participant	Sex	Age	In education	Education
1	Male	25	Yes	University
2	Male	23	Yes	University
3	Male	23	Yes	University
4	Male	23	Yes	University
5	Male	26	Yes	University

Participant	Aftenposten TV	NRK on-line TV	Social networks	Sharing
1	Never	Less than once a month	Daily	Less than once a month
2	More often than once a week	More often than once a week	Daily	More often than once a month
3	More often than once a month	More often than once a month	Daily	More often than once a week
4	Less than once a month	Less than once a month	Daily	More often than once a month
5	More often than once a month	More often than once a week	Daily	Daily

Table 3.1: Participant data collected from pre-questionnaire

3.3 Experimental Procedure

A pre-test with one MSc. student and one employee at SINTEF were conducted to ensure that the equipment, assignments and tools were appropriate and working as they should.

The study consisted of three parts:

- The first part was a pre-questionnaire to gather background information about the participants. It included both personal information and some information about the experience with on-line TV.

- The second part was the experiment. All the participants came one by one and were welcomed. They were informed about the experience, the main outcome and what they were going to do for the next hour. They were informed about how the experiment was going to be conducted, this includes video/sound recording and action recording.

- The third, and last part, was a interview with the participants. It were asked a few questions about the experiment and on-line TV in general. All questions can be read in Appendix B - Interview guide.

3.4 Data collection and tools

A range of data has been collected, some with the help of automated logging and other manually by e.g. taking notes. Table 3.2 shows all the data that has been collected in the experiment.

Type of data	Reason for use
Low-level data	All actions done by the participant on the computer during the experiment, logged with the help of Noldus uLog and further analyzed in Noldus Observer XT.
Video	Record user behavior (desktop recording and facial expression), with the help of Camtasia Studio.
Observation	Able to understand the participant better than just from video recording.
Notes	If any unforeseen events occurred during the experiment.
Interview	Understand the intension's behind the participants actions.

Table 3.2: Collected data in the experiment

3.4.1 Noldus uLog

uLog is a tool for automatically recording user-system interaction. It records keystrokes, scrolls, and mouse clicks. One can see which button was clicked and whether or not it was double-clicked. On top of this, uLog can log individual characters as well as complete text strings so one can read exactly what the test participant has entered. uLog tells one which applications are active, which menus are used in the user interface, which pages are opened in the web browser, and much more. (Noldus_uLog, 2010)

3.4.2 Camtasia Studio

Camtasia Studio is a screen video capture software created by TechSmith. The user can define an area of the screen or the full window that should be recorded. This is set before the recording begins. Furthermore, is possible to record audio from a microphone or from the computer. Camtasia Studio allows to take use of your webcam to place it on the desktop recording. This gives the possibility to both watch the desktop of the user and observer the behavior.

3.4.3 Noldus Observer XT

The Observer XT is the professional and user-friendly event logging software for the collection, analysis, and presentation of observational data. The data that has been analyzed can be presented in range of charts, there is also possible to transfer the data to other tools for generation of more complex charts. It is able to combine and manage a large amount of equipment for data gathering. One are able to see the data synchronous with for example video or sound of the objects that one are observing. However, the data needs to be synced manually by the user. (Noldus_ObserverXT, 2010)

3.5 Assignments

The assignments were split up in two parts. The first part is conducted on `http://www.nrk.no/nett-tv` and the second part on `http://www.aftenposten.no/webtv`. In both parts the participants was asked to conduct an easy exercise assignment to get some practice with interface on the on-line TV page. Furthermore, the participant was provided with the real assignments. All assignments should be answered, when it's relevant. Both the exercise assignments and the 5-minute assignments (assignment 1 and assignment 6) did not have any answers. Furthermore, assignment 5 (the iPad assignment) had no answer. After completing the experiment, it has been decided to not analyze the use of iPad (assignment 5) in this thesis. [1]

---

[1]It has been decided to have an independent study with the iPad at later date.

Each assignment was created in such matter to provide us with different kind of data. Some assignments was the participant able to choose what to do (assignment 1 and assignment 6). These was created to provide us with data where the participant was browsing content and at some point will start to watch some content. Furthermore, the rest were created with a more detailed assignment with a direct answer. However, assignment 3 was created as an open-ended question. The answer gave the participant the possibility to decide when the assignment was properly answered. In the assignments that were created to be used on the newspaper page, it was possible to find the answer in both the article text and the video. Assignment 8 and assignment 9 was focusing on the participant to post some content on a social media (Facebook or twitter). It was up to the participant to pick the social media he or she was most comfortable with. The participant was provided with an account to use, for the reason that not all want to use their own account. These assignments will provide us the with both browsing and sharing to a social media from the Internet page with help of sharing buttons on the Internet page.

3.5.1   NRK

NRK provides the user with the opportunity to watch live TV or watch programs that has been shown on live TV for free. The content that have been shown on live TV are made public after a short time. The content is public as long as NRK have the royalties to the program.

Exercise assignment 1

Go on NRK on-line TV and locate your favorite program. Start to look at it. Turn the sound off and on. See the program in full screen. Go back to small screen. Fast forward and back in the program. Go back to NRK on-line TV's website.

Assignment 1

You have a 5-minutes coffee break and you want to see NRK on-line TV on your PC. Go `http://www.nrk.no/nett-tv` and do what you like. We interrupt you after 5 minutes.

Assignment 2

What was the program "Rundt neste sving: Sekken" about and what kind of work has Vidar Svenning?

Assignment 3

What happened to "Arild Rypdal" in the "Glimt av Norge", which ran the "05/13/2010"

Assignment 4

Look at the 10:12 episode of "mat lyst" and answer the following questions 1. What is the name of the woman who is dressed in purple sweater and brown apron? 2. When a potato froze in winter, what did it become then?

Assignment 5

You are sitting in the bus and have access to NRK on-line TV iPad.
Go `http://m.nrk.no/m/nett-tv` and do what you like. We interrupt you after 5 minutes.

3.5.2 Aftenposten

Aftenposten is providing the users with an combination of video and text when browsing news.

Exercise assignment 2

Go Aftenposten TV Video http://www.aftenposten.no/webtv and find out a sports news that interests you. Start to look at it. Turn the sound off and on. See the program in full screen. Go back to small screen. Fast forward and back in the program. Return to overview with all the videos.

Assignment 6

You're at home and have access to Aftenposten.no Video on your PC.
Go to `http://www.aftenposten.no/webtv` and do what you like. We interrupt you after 5 minutes.

Assignment 7

Go Aftenposten.no and find the clip "Dødstallene stiger etter flom". In the clip, find out how many will be affected by the floods in Australia.

Assignment 8

Locate the clip "Kolonnen av biler styrer seg selv," and share the clip on facebook / twitter.

Username: (Removed)
Password: (Removed)
(The same username / password for both facebook and twitter)

Assignment 9

Find a video you like on aftenposten.no and share it on facebook / twitter.

Username: (Removed)
Password: (Removed)
(The same username / password for both facebook and twitter)

3.6   Interviews

The interviews were semi structured (Seaman, 1999), with a mix of open-ended and specific questions, to allow unforeseen types of information to be recorded. Under the first interview it was one interviewer and one observer that conducted the interview. At all later interviews, only the interviewer was present during the interview. The interviewer was responsible for conducting the interview and asking the questions. The observer's job was to take notes and pose additional questions after the interview if the interviewer had overlooked something. The interview guide is given in Appendix B. Each interview was audio recorded. Pieces of the interview relevant to the subject of the study were transcribed.

Chapter 4

Data preparation and tools

This chapter describes the data preparation of the data collected by uLog, Camtasia, questionnaire and final interview. Furthermore, different settings used is also described. Data presented here are from the pre-test and will be the basic for the main experiment. The preparation of the low-level data is presented in section 4.1, screen capture and audio/video data in 4.2, observation notes in section 4.3, interview in section 4.4 and the created coding scheme in section 4.5. Furthermore, have some lessons learned been described in section 4.6.

First observation

In the pre-test, from the first explorative study, we decided to log everything because we did not know what would serve our purpose. This resulted in huge amount of complex data. Further understanding of what to log and how to process it will be a big time saver for the data preparation in the main experiment.

The collected data from the first session differs in size and complexity. E.g. in the low-level data it wasn't hard to identify the meaning of each action, however it was hard identify the purpose of the action. The data collected from the different sources are listed in Table 4.1. The data in Table 4.1 is from one participant and list the type of tool together with what kind of data it provides. Furthermore, the amount of data is listed. The data collected for the other participant will increase or decrease in relation to the first participant presented in Table 4.1

Tool	Type	Amount
uLog	low-level data	4000 lines
Camtasia	Desktop, sound and webcam recording	12GB of data
Observation	Text	4 Small pages of notes
Interview	Sound	10 minutes of recordings

Table 4.1: Data collected from one participant

All data preparation will follow the basic steps proposed in (Thomas and Karahasanović, 2009), that can be seen in Figure 2.4.

4.1 Low-level data

The raw data from uLog was big and complex and contained a lot of information. Early on it was identified that much of the logged events was irrelevant for the project. The first, and most obvious, were the startup events and end events. The startup events consisted of some interaction with uLog and the startup commands from Camtasia. The end events were the stop commands from Camtasia and the stop commands from uLog.

Furthermore, uLog was creating a control image for every event. All images were stored in the logging file and not in a separate image file; therefore the uLog file was even harder to read. The main reason to remove the logging of event with a respective image is because we logged the desktop with Camtasia. Since uLog, Observer and Noldus have noway to resave the file to the wanted format, all images was removed with a regular expression[1]. A regular expression can be used in many text programs or programming languages. In text editor programs, like NotePad++[2], it can be used with the search and replace function. The regular expression used to remove the images, with explanatory text:

< ControlImage >([ˆ <>]] </ControlImage >*
First find < ControlImage >then skip every character that is not <or >and
end with </ControlImage >. The reason for using ˆ <>is so one don't skip
*the end tags for the image, since * just will take any character.*

---

[1]Regular expression, also referred to as regex or regexp, provides a concise and flexible means for matching strings of text, such as particular characters, words, or patterns of characters.

[2]Notepad++ is a free (as in "free speech" and also as in "free beer") source code editor and Notepad replacement that supports several languages. Running in the MS Windows environment, its use is governed by GPL License. `http://notepad-plus-plus.org/`

uLog saves some information about all the application that started and ended. This should not be logged. The main reason for this is because the only program one are using is Internet Explorer. Furthermore, it does give one information about all other programs and processes, e.g. do this include dllhost, SearchProtocolHost, SearchFilterHost and more. They do not just appear when the user starts a new program, however also when the computer start programs self. Dllhost, SearchProtocolHost and SearchFilterHost is a process that is used by Microsoft Windows Operating System and will not affect our result when removed. Since it is used by Microsoft Windows Operating System only it will be possible that our findings not will be transferable to other operating systems. Even thought this is removed do one still have information about what program that is in focus. Furthermore, one can from that identify that a new program has been started. It cannot be identified when the program starts directly, however do new programs get focus when it starts. Therefore, if a program gets focus that hasn't been seen before, a new program has started. One can then only get the programs started by the participant and not by the internal processes by the Operating System.

```
<OBS_EVENT>
        <OBS_EVENT_TIMESTAMP>
        2011−01−31  11:54:31.5738609
        </OBS_EVENT_TIMESTAMP>
        <OBS_EVENT_BEHAVIOR NAME=" Application_started">
        <OBS_EVENT_BEHAVIOR_MODIFIER CLASS=" ApplicationName">
        dllhost
        </OBS_EVENT_BEHAVIOR_MODIFIER>
        </OBS_EVENT_BEHAVIOR>
</OBS_EVENT>
```

Listing 4.1: Example of event application started. Application ended is just the same, only the name is changed.

The event window activated is relevant, however it is also providing data that one don't need. A lot of the window active events contain information about events triggered by the system and not the user. Furthermore, it is useful to know which window the user is interacting with. The only events that should be removed or just ignored are the one that is triggered by the system and is not critical for the project.

```
<OBS_EVENT>
        <OBS_EVENT_TIMESTAMP>
        2011−01−31  11:54:36.4031371
        </OBS_EVENT_TIMESTAMP>
        <OBS_EVENT_BEHAVIOR NAME="Window_Activated">
        </OBS_EVENT_BEHAVIOR>
        <OBS_EVENT_COMMENT>
        window:[VG Nett − Forsiden − VG Nett −
        Windows Internet Explorer]
        </OBS_EVENT_COMMENT>
</OBS_EVENT>
```

Listing 4.2: Example of event window activated

In this experiment the participant should only be needed to use one window under the experiment, e.g. the Internet Explorer window to show the on-line TV services. However, one cannot foresee that a participant only will use one instance of Internet explorer or other programs and should therefore continue to log window active.

Furthermore, uLog has an extensive range of possibilities of logging. All possible logging settings in uLog are located in Appendix C and the settings used in this experiment is also presented.

4.2   Screen capture and audio/video data

Camtasia gave us the possibility to capture desktop behavior and webcam recording to capture users behavior with focus on face expressions. Furthermore, sound was recorded to capture unforeseen events and questions if the participant had any. The advantage of using Camtasia was fewer programs to comprehend and relate to. The only disadvantage was that Camtasia created a huge file with a big overhead. E.g. one hour of recording yields 12GB of data in the Camtasia project format. Different options of saving the video file with sound is possible, however it is no good video codec for compression in Camtasia. Stored with the option "save for youtube", created a video file with the size 9 GB. Furthermore, the transfer process to different computers will take long time and programs could have problems with it. Event though the price of hard drives is rather low now a day, it is less price consuming to convert the data to decrease the file size. More about file conversion and problems with conversion , can be read about in section 4.6 (Lesson learned).

4.3 Observation data

After the pre-test, we realized that the observation notes didn't need to be very detailed. The observation notes didn't provide anymore information than was already provided by the video-, sound- and webcam-recording. Some examples on what information notes was providing:

(1) *Watching and taking notes.*

(2) *Watching and take a look on the question.*

(3) *Correcting answer.*

(4) *Lean back.*

(5) *Pushes pause when writing the answer.*

Most of the notes can be identified from different video and sound sources. However, one cannot distinguish between when a participant is just taking notes, writing an answer or correcting an answer. Nevertheless this will not be a problem, for the reason that one is more interested in the interaction with the computer rather than the interaction with different sources. Furthermore, are the answers the participants have written not that important to the observation. In further experiments and observations with the same amount of data it is possible to remove the observation notes. However, a notebook should be present if an unforeseen event should happen, that not could be identified from the video and sound source.

4.4 Interview data

Each participant had to give some feedback after the experiment was completed. Firstly, it gave us the possibility to identify if the tasks were working as they should. Secondly, get a good insight in how the participants choose to do what they did under the experiment. E.g. why they choose to search for content or just click there way to content. Lastly, it gave the participant the chance to state other thoughts they made up under or after the experiment. All questions can be seen in the interview guide in Appendix B (Interview Guide).

For each recorded interview I identified the start of each question and wrote down the answer. Some statements from the participants were not interesting or related to the questions and were therefore removed. E.g. it could be when a participant asked questions related to what kind of answer we wanted. Furthermore, the answers was used as is, to give the best impression about the participants thoughts.

4.5   Coding Scheme

The coding schema was developed in an iterative process with respect to earlier studies done by (Bernhaupt et al., 2007), (Darnell, 2007) and (Kunert and Kunert, 2009). The problem was to understand the distinction between on-line TV and iTV. The coding scheme created in the first iteration is presented in Table 4.2. Table 4.3 presents the coding scheme further used in the experiment. I ended up with a rather general coding scheme that can be used on a large range of on-line TV Internet sites.

Table 4.2 and Table 4.3 lists the event group with respective events and list the event type. The two possible types are state event and point event. A state event has a start and an end, additionally a point event is an event that happens at a point in time. Some groups are mutual exclusive, where the previous event is active until a new event is activated.

Event Group	Event	Type
Looking	Looking on computer	State event
(Mutual exclusive)	Looking on assignment	State event
	Looking else where	State event
Actions	No Action	Point event
	Typing	Point event
	Link Clicked	Point event
	Play	Point event
	Pause	Point event
	Stop	Point event
	Fast Forward	Point event
	Fast Rewind	Point event
	Fullscreen	Point event
	Sound Up	Point event
	Sound Down	Point event
	Mouse Movement	Point event
	Mouse Click	Point event
Video	isPlaying:YES	State event
(Mutual exclusive)	isPlaying:NO	State event

Table 4.2: Coding schema first developed

39

Event Group	Event	Type
Other	Typing	Point event
	Link Clicked	Point event
Look	Looking on computer	State event
(Mutual exclusive)	Looking else where	State event
Fullscreen	Fullscreen on	State event
(Mutual exclusive)	Fullscreen off	State event
Sound	Initial sound	State event
(Mutual exclusive)	Sound adjustment	State event
	Mute/Unmute	State event
Video	No Video	State event
(Mutual exclusive)	Play	State event
	Pause	State event
	Stop	State event
	Seek	State event

Table 4.3: Coding schema developed for use in experiment

At first I started up with to many events too handle when coding the observation data. E.g. I did have mouse movement as an event and found quickly that it was a bad idea, because one need to stop the video almost every second and the mouse is moved a lot during the experiment. Furthermore, different types of looking events were used. Some types of looking events were removed and two more general looking events where further used (*looking on computer* and *looking elsewhere*).

Some events were not fully removed, just redefined or moved. *Seek* was a more generalized event from fast forward and fast backward. Furthermore, sound up and sound down was redefined to *sound adjustment*. Both these events, that were redefined, didn't loose any critical information. However, when the coding was done it was much cleaner then when it was coded with events that were not redefined. Some events was just moved to other categories because they should operate in different time spaces. E.g. it is beneficial that play and full screen are in different categories. It is then possible to identify the relation between them. This is not possible if they are in the same group, since they will exclude each other.

4.6 Lessons Learned

Some lessons have been learned after using a range of programs for collecting data and analyze data. Firstly, uLog and the observer were the programs that gave me the hardest time. uLog has different save settings like other programs, however uLog has no opportunity to change this after a observation has been started. The problem was

then that it was chosen to save in a normal .XML format, and uLog didn't warn about the observer not handling this format. The observation could then not be used, however the Observer did support a format called .ODX that is in an XML format. Noldus, the developers of both programs, didn't have any converter from XML to ODX. It was then created a converter in JAVA and the XML file was coded to the right format after a short period. However, the Observer XT was updated to the new Observer XT 10.1 and the ODX stopped working again. Therefore it should be tested if the program can use the format that they provide, even thought it may be supported in an earlier version.

When the video was exported from the Camtasia Studio project file, the file size was 9 GB. We used Windows Movie Maker for further compression and gave a compression rate of 95.27%. Even though the compression rate is high, the video and sound quality are more then good enough for our purpose. The settings used in windows movie maker to get this compression rate is shown in the input file column Table 4.4.

Property	Input File	Output File
Extension	wmv	mpeg
File size	395.04M	-
Frame size	1280*720	1920*1080
Video Compression	wmv3	MPEG-2
Video Bit Rate (bps)	8000.00k	5000k
Frame Rate	30.000	Original
Audio format		AC3
Audio Bit Rate (bps)	128.02k	256k
Audio Frequency (Hz)	48000	48000(high)
Audio Channels	Stereo	Stereo

Table 4.4: Video compression settings

On the other side the compression created a problem with the Observer XT 10.1. In earlier version, Observer XT 7, the file format was supported, however from version 8 it was discontinued. Conversion of the video file must be done again, now to a format that the Observer XT 10.1 could handle. People at Noldus advised me to use the OJOsoft Total Video Converter[3] and provided me with some different settings to use. They advised me to use either MPEG1 or MPEG2 video encoding, however it was some problems with only selecting the compression settings with MPEG1 or MPEG2. The solution was to use exactly the same settings as Noldus, even thought the file size become slightly larger. The settings advised by Noldus can be seen in the output file column in Table 4.4.

When an experiment includes different type of Internet pages it may be smart to create an independent coding scheme. Each Internet page and Internet service can differ a lot

---

[3]OJOsoft Total Video Converter is a video converter which can convert video files between almost any video format at an unmatchable speed and high quality. It is easy to use. `http://www.ojosoft.com/`

and a independent scheme would create some advantages. Some of these advantages can be:

1. Used in other experiment with the same objective.

2. Can give the possibility to understand differences between users on different Internet pages.

3. Do not need to understand more than one coding scheme.

4. Overall save some time early in the experiment, because the coding scheme is already provided.

The identification of the right independent coding scheme for a range of Internet sites can be a demanding task. However, when the independent coding scheme is created it can be used in a range of situations. Furthermore, it can just be tweaked when the need for a more detailed scheme is required.

Chapter 5

Analysis

This chapter describes the analysis of the data collected by uLog, Camtasia, questionnaire and final interview. In section 5.1 it is presented how the user tasks have been identified from the low-level data collected from uLog. Section 5.2 presents how the low-level data from uLog has been combined with the screen capture to identify and verify user tasks. Furthermore, the reason for triangulating sources is presented in section 5.3 and lastly how long time the data preparation and analysis consumed in section 5.4.

To identify user tasks I analyzed the data collected from uLog in all assignments, except from assignment 5 that involved iPad. All assignments gave a good range of data that can identify what the users were doing. Assignment 1 and assignment 6 were mainly used to identify when the users was watching a specific video. This was because all participants choose to find a video they wanted to view. All the other assignments was depending more on different questions that needed to be answered and would identify where users are trying to find specific information in a video.

While the data collected from uLog and video recording can tell us what the participants did, it says little about why the participants chose different approaches. This is an advantage of having an interview at the end of the experiment. Not only do they provide information about the participants' actions, they also, to some extent, give us explanations for those actions.

5.1   Identification of user tasks from low-level data

Names on actions logged with uLog, in combination with the desktop recording have been used to identify user tasks. Some actions do not interfere with other actions in some user tasks. E.g. sound adjustment is not interfering with the user viewing some specific content. Furthermore, every user task will handle user actions in different ways

depending on what kind of user task it is.

5.1.1 uLog data

Observer XT 10.1 has the possibility to visualize the observed data and the data collected from uLog. Figure 5.1 is a visualization of the uLog data collected from one session in the experiment. It gives a brief overview of the user actions through the session. Furthermore, the red line (that was added manually by me) describe when an assignment, in the experiment, has started or ended. Figure 5.1 contains all actions done performed by one participant, categorized in a high-level manner.
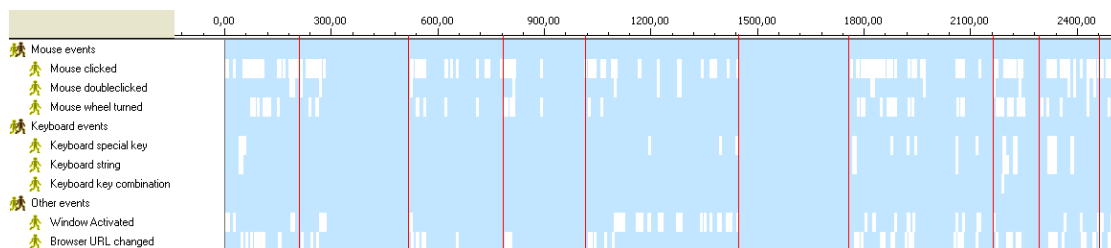

Figure 5.1: Observer XT visualized uLog data collected

Each recorded action, where an action is a small line in the figure, contains more information than presented in figure 5.1. When one positions the mouse over an action, additional information will be revealed. E.g. a mouse click can have the comment that the play button was pushed. Since it is a crucial part to know what was clicked, the visualization must be shown in another format. Each action is logged by uLog as a point event, e.g. that gives the opportunity to see when an action happened in time. However, it is hard to make a good visualization of all the actions in a session.

Microsoft Excel was used for further visualization of the data with respect to a time line. Visualization in Microsoft Excel gave a much better overview of when an action happened and how long it was between two actions. To generate the spreadsheet for Microsoft Excel a small Java program was used to extract the user task depending on the conversion table 5.1.[1] The complexity of the Java program was small and did just comprehend the participants in this thesis. In later experiments a more complex program should be created to fully understand all actions. The Java program extracted all actions from the assignment and combined it with the time stamp. Each action had to be combined with what happened and what was clicked. All extracted data was then filled in each column, where the action got an ID instead of the text. The need for a number ID is to let the Microsoft Excel be able to interpret the number to generate a graph. Table 5.1 shows an example where each ID is connected to a user action. As

---

[1]The java program was created by me and contained different combination of if-statements that compared strings. This thesis will not describe the program in detail, because the complexity is low.

mentioned before the connections in Table 5.1 is not all possible connections, however these actions is used by the participants in the experiment in this thesis.

Action	Coded ID
Browser URL changed	1
Mouse wheel turned	2
keyboard string	3
mouse clicked + hyperlink	4
mouse clicked + back/forward = browser navigation	5
mouse clicked + Pause	6
mouse clicked + mute	7
mouse clicked + slider + seek	8
mouse clicked + slider + Volume	9
mouse clicked + edit	10
mouse doubleclicked + pane + EVRVideoHandler = Fullscreen	11
keyboard special key = keyboard interaction	12

Table 5.1: uLog events combined and coded to numerical ID

In figure 5.2 assignment 1 can be seen carried out by one participant. Table 5.1 is used to convert from the text to the right ID. Each point in the graph is a user action and is pinned to a point in time. Each point on the horizontal axis is the time in seconds.
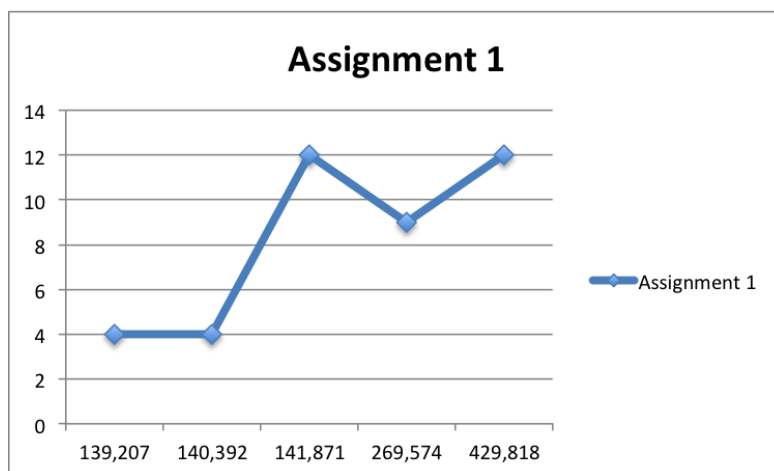
Figure 5.2: Assignment 1: example on browsing and watching

From figure 5.2 it can be identified that the participant is browsing, however the participant is also watching some content between action 3 and 5. The user is starting with clicking hyper links until the participant find something interesting and the user actions stops. The user action, volume changed, is used while watching the video content. However, it does not mean that the participant have stopped watching the video

content, rather that the sound was too high or low. The issue is that the play button was never clicked, and therefore it cannot be identified that the participant is watching some content from only analyzing the low-level data. It must be combined with some other data, e.g. the on-line TV provider receives a video request when the participant access a video. This can identify when a participant has started to watch a video. It can also be identified from the screen capture that the participant started to watch some content.

Figure 5.1 can be used to identify what the participant is doing, however just in a shallow way. Column two, after the second red line, in figure 5.1 shows that the assignment starts with a range of mouse click events were the participant is browsing for content. Since one also have visualized the uLog data in Microsoft Excel one can identify that the participant found some content. Furthermore, the video did start and all events stopped until the participant was interrupted. In some degree this can show that the participant is watching some content. However, events can show up after the video has been started. Therefore, do some events matter to identify if the participant is actually watching the content or not. The same issue arise as mention before, the play button have not been clicked.

5.1.2   Determining the context of actions

It is essential to know the context where an action occurs, to understand what the user or system was doing. (Thomas and Karahasanović, 2009, p1195) In most cases the participant will only be handling one window when carrying out the experiment. Therefore, the most important context switch will be full screen/no full screen. This will be essential for understanding if the participant is watching content or not. Depending on the context actions will interfere in different matters. Furthermore, there is two additionally context's that is important to understand in this experiment: if the participant is watching or not and which window the participant is using. Short described when the on-line TV window is out of focus the participant is most likely not watching the content.

Watching versus not watching context

When the participants is watching some content actions will operate in a different way in proportion to when the participant is not watching some content. In some cases will e.g. the action seek take the participant out of a watching state and into a not watching state. This is due to that the participant is searching for something, e.g. a more interesting part of the video. Table 5.2 shows the kind of actions that triggers the participant to go from a watching state to not watching and the opposite.

	Play	Stop	Pause	Seek	Full screen	Sound change	Link clicked
Watching		X	X	X			X
Not Watching	X		X				

<div align="center">Table 5.2: Actions that trigger to go from on state to another</div>

Differentiate window context

As mentioned earlier the participant is supposed to interact only with one window, i.e. Internet Explorer with the two different Internet pages used. However, it cannot be disregarded that the participants may act unexpected. This may include opening Internet Explorer windows and different other programs. In general it is easy to identify when the participant changed the context to another window. In uLog the window / context change is reported as an action in the log file:

> *Behavior: Window Activated*
> *Comment: window:[Google - Windows Internet Explorer]*

It is straightforward to understand that a window was activated and the window that was activated was Internet Explorer with an Internet page with the title Google. The issue is to differentiate Internet Explorer windows with the same titles. However, to differentiate windows, will in most cases not be a problem, because Internet pages in general do not have the same titles and therefore it will not be an issue.

The most interesting context switch is when the participant goes to full screen mode. It will mostly be the same as when a window got activated. In spite of this the element names do often lack in proper naming (read more in chapter 6.1.1) and therefore must be combined with other actions. One way is to identify when the participant has clicked the full screen button and enters full screen mode. However, the participant do also have the possibility to just double click inside the video window and this is difficult to identify. Therefore it should be tested and investigated how to identify what actions happened when going into e.g. full screen. An example of going into full screen:

> *Mouse doubleclicked: pane:[EVRVideoHandler]*
> *Window Activated: window:[NoCaptionsView]*
> *Window Activated: window:[WMPTransition]*
> *Window Activated: window:[NoCaptionsView]*

In the example about entering full screen it is hard to understand that it is actually full screen that is entered. It can only be identified if one has enough technical knowledge or to combine the logged actions with e.g. recorded video. In the example EVRVideo-Handler is the video window and the rest is logged when the video enter full screen mode.

5.1.3   Screen capture data

To identify what the participant is doing when only analyzing the low-level data, the understanding of when actions occur and what the action means, must be understood. With uLog data this is a difficult task since all actions are point events. However, the coding scheme created in Table 4.3 introduces actions to be created with state events. This will make it possible to identify how long a participant was looking on a video before an action triggers the participant to move to a not watching state. This is only possible when the video has been coded.

To visualize the data collected with screen capture the video must be coded. Therefore must the entire recorded video be watched and coded before further analysis. For every action that was observed the video had to be stopped and the action recorded. In figure 5.3 one observation is coded and visualized in Observer XT. It looks much like figure 5.1 from the uLog data, however it is much easier to identify the duration of each action. The red line, which was manually added by me, describes when an assignment in the experiment has started or ended.
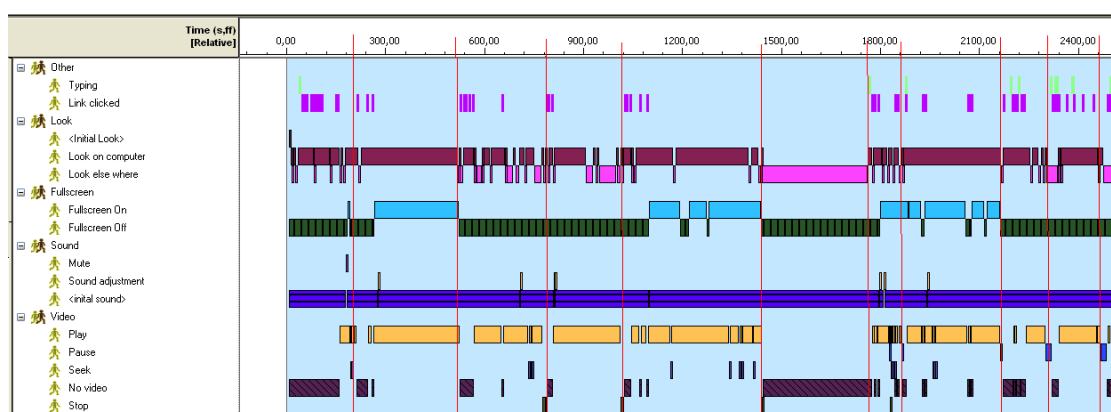
Figure 5.3: Observer XT visualized screen capture data collected and coded.

The Observer XT can be used in a wide range to identify different connection between actions. This can be done with different data profiles to extract data. It is now possible, since the video has been coded, to use the analyze function in the Observer XT and get some statistics about the actions that has been used.

5.1.4   Use of data profiles in Observer XT to filter data

The Observer XT can extract data depending on different data profiles. Each data profile has all data as input and some data as output, depending on how the data profile is created. The equation in the middle of the input and the output can be created from different expressions with AND- and OR-operators. It can extract the actions one

specifies, e.g. when the participant is watching content and not watching content. In figure 5.4 there is an example of a data profile where it is used a filter to include only play, stop, pause and seek behaviors in the analysis. Firstly, the start box contains all information one has gathered in the project/experiment. Secondly, a filter is added were one can extract only the wanted behaviors/actions. Lastly, the result will be presented. The result can then be used to create spreadsheets and charts.
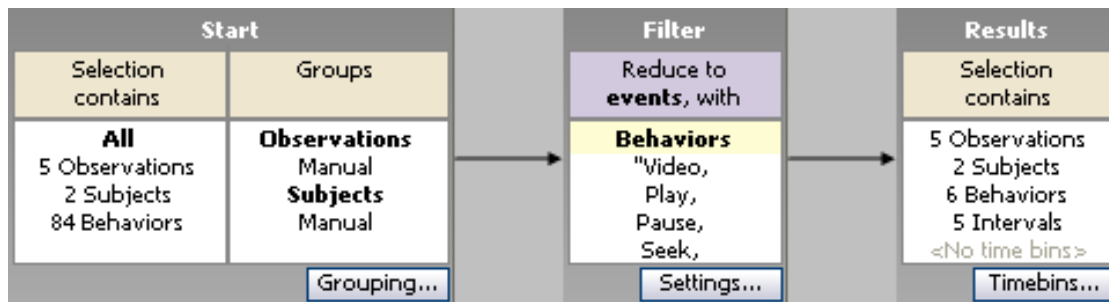
Figure 5.4: Observer XT - Data profiles.

5.1.5 Data collected from use of data profiles

The Observer XT can use the data extracted to create a spreadsheet. A range of different statistic options, e.g. how many times one behavior/action occur, can be chosen and the layout of the spreadsheet can be manipulated. It gives a good overview of the collected data and can be exported for later use in other statistic programs like Microsoft Excel. Furthermore, the spreadsheet can be used to generate different kind of charts in the Observer XT. However, the charts generated in Observer XT is not possible to edit. Therefore chart should be generated in statistic programs like Microsoft Excel, since it is much more flexible and will give a better result.

In Table 5.3 the participants is shown with the total amount of time spent in full screen on, full screen off and play. It is generated in the Observer XT with the help of a data profile where only fullscreen on/off and play is presented.

	Participant 1	Participant 2	Participant 3	Participant 4	Participant 5
Fullscreen On	349,9	893,96	1668,38	1908,4	115,3
Fullscreen Off	2054,18	1621,6	1096,24	833,16	2509,79
Play	1419,09	1638,55	1895,33	1742,62	1789,56

Table 5.3: Fullscreen versus not fullscreen (In seconds)

Figure 5.5 presents the same data for one participant as Table 5.3 did for all. However, it is more explanatory to visualize it than presenting it in a table with numbers. One can also identify when both play and full screen is on, which can identify that a participant is watching a video. Furthermore, it is not possible to identify from the spreadsheet or with the help of data profiles in the Observer XT. This is because one cannot create data profiles where e.g. one state is combined with another state. E.g. extract the data where full screen and play is active. (More about this issue can be read more about in the manual to the Observer XT, in the section of data profiles.)
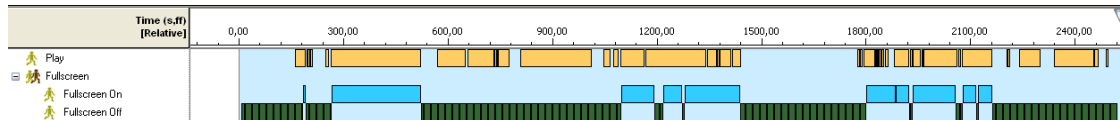
Figure 5.5: Visualization of one participant in full screen versus not full screen

A range of combinations can be generated with the help of data profiles in Observer XT. Each data profile give the opportunity to find and extract just the data one want to use.

5.2 Combining visualized uLog and screen capture data

The combination of visualized uLog and screen capture data gives the opportunity compare and validate them. Figure 5.6 shows both data types in a time line. The red line, which is manually added by me, describes when a assignment has started or ended.

Figure 5.6 is used for the main analysis to investigate the possibility to identify tasks from low-level data. All actions from uLog are synchronized relative to the video recorded in the experiment, thus all actions happens in the right location in the video. It is now possible to identify what actions a participant used to carry out a assignment. To analyze the data it is still a need to use the Observert XT, because all information from uLog cannot be seen in the diagram. However, different amount of information can be extracted from the diagram even though some information is missing. One type of information is when the participant is watching some content do almost all low-level actions stop until the participant is watching no more. (See the second column in Table 5.6 for an example. The participant start by using the mouse to navigate and when the participant has found a video almost all actions do stop. The only action that occurs is the participant adjusting the audio volume.)
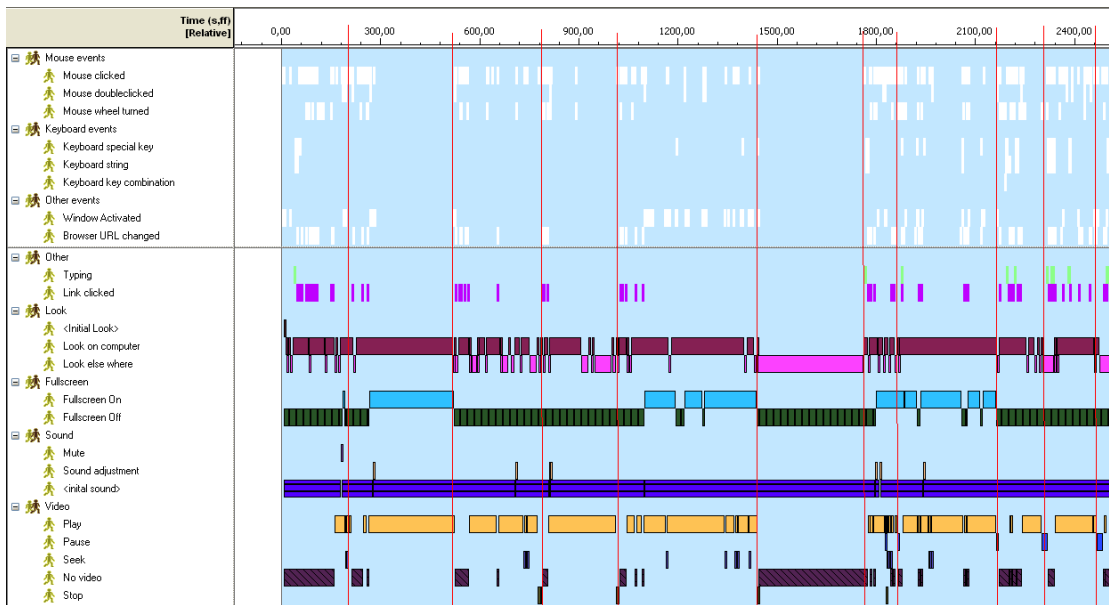
Figure 5.6: Combined visualization of uLog and screen capture

5.3   Triangulation of sources

The purpose of triangulation is to check the accuracy by examining evidence from different data sources. The evidence might come from different sources, be collected using different methods, be analyzed using different methods, have different forms (interviews, observations, documents, etc.), or come from a different study altogether. (Seaman, 1999) Since qualitative research is often met with some skepticism, triangulation used is used to obtain a trustworthy result. Furthermore, it has been identified that case studies using multiple sources of evidence were rated more highly, in terms of their overall quality. (Yin, 1994, p92)
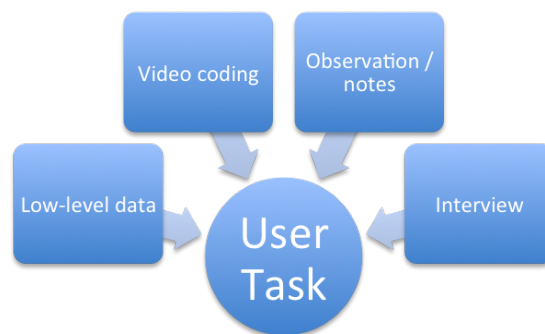
Figure 5.7: Triangulation of sources

51

In this experiment the following sources was used: Low-level data from the participants actions with sound and video recording. Furthermore, notes was taken and the participants interviewed. All of these sources are used to give a more trustworthy and accurate result. The main reason for collection all these data, is to have the possibility to see if the identified user tasks can be found in both the low-level data and the coded video. Furthermore, the recorded interviews can help to understand what the participant actually wanted to do.

All sources have been used in different levels where one source is included to verify each level. Table 5.4 shows each level with respect to the source of evidence. Furthermore, it is stated what the source should verify. Each level of verification embraces all previous levels.

Verification Level	Source	Should verify
0	User Tasks	What users do
1	Low-level data	Level 0
2	Video coding	Level 1
3	Observation and notes	Level 2
4	Interview	Level 3 (More explanatory than verifying)

Table 5.4: Level of verification of sources

5.4 Time consumption in analysis

The different stages of the analysis were both time consuming and a demanding task. Table 5.5 shows the time consumption of the different stages in the analysis. Firstly, is what part of the analysis shown with how long that part took. Furthermore, a comment is written about the stage.

Stage	Time	Comment
Preparation of experiment	15 min	Prepare the room and setup of the system
Experiment	1 hour	Each experiment took around 1 hour to complete.
Video conversion	2 hours	Each recorded video file had to be converted to a smaller format and to a working format that can be used by the Observer XT. This should be done over night to save time.

Video coding	1 hour	Each video needed to be coded. A video was around 45 min long. (Each video coding need to be verified and will take up some more time)
Sum	4.25 hours	

Table 5.5: Time consumption for one participant.

Table 5.5 is for one participant only and therefore must be added up by the number of participant that completed the experiment. The time used for only one participant is rather high and different unforeseen events must be taken into account. Video conversion is the most time consuming part of the analysis. The Observer XT supports the format Camtasia stores the video file in, however the file size is large. Furthermore, since the file is large it will create problems with storing and transferring of the file. For each participant a storage space of around 10GB must be free to use and the transfer time of a 10GB file differs from system to system. With a transfer rate of 15 MB/s, it will take around 1 hour 29 minutes. It takes shorter time to only transfer the raw video (if one count in the conversion time), however is the storage need higher. Furthermore, if the need to copy the video to another computer, it will use 1 hour 29 minutes each time the video needs to be copied. On the other side if the video is converted the size is around 2GB and uses only around 18 minutes to be transferred. It saves both size and time even thought it takes more time to convert the video. However, most video conversion programs can take a list of files for conversion and can be converted over night. Both the transfer and the conversion can be done over night, however people will in most cases transfer files when it is needed and therefore the file should be as small to make the transfer time decrease.

Even though disk space is cheap these days, one should think about the transfer time if more than one person should be using the data. If only one person is going to take use of the data, the conversion can be dropped and just take use of an external hard drive that is connected to the computer. However, if the data should be used by more than one person it should be converted, thus the transfer time and the data consumption will decrease.

Chapter 6

Results

This chapter presents the results of the study. The objective of this research was to identify user behavior and user tasks in an on-line TV context. Section 6.1 presents the results regarding low-level user task and problems with cryptic naming of elements in low-level data. Section 6.2 presents the high-level user task and an illustration of how the high-level user tasks depend on the low-level user tasks. Section 6.3 discuss problems with differentiate high-level user tasks. In section 6.4 two user types in on-line TV presented. How the participants work on the on-line TV service in the experiment versus at home is presented in section 6.5. Furthermore, some time specific findings are presented in section 6.6. Finally, section 6.7 discuss the results in relation to earlier studies, how some national events may interfere with the regular on-line TV usage and lastly a other method for identification of user behavior in online video streaming is discussed.

By observing the participants and analyzing the low-level data has it been identified a range of user tasks. Furthermore, it has been identified that it is possible to identify almost all user task from the low-level data. The identified user tasks are categorized in two main categories, high- and low-level user tasks. The high-level user tasks focus on what the user are doing and the low-level user tasks focus on how the users achieve the high-level user tasks. The high-level user tasks are hard to identify and is because the high-level user tasks is a combination of one or more low-level user tasks. It is needed to combine them in the right way and they must occur in the right sequence. E.g. if the play button is pushed for the first time the video will start, however if it has been pushed on a earlier stage the video will stop[1]. Furthermore, if the user has moved further on the Internet page, a new video will start playing if the play button is pushed again. On some Internet pages the video do start by them self and no low-level actions are logged. Therefore, the identification of what the user is doing a hard and complex task.

---

[1]The videos will stop/pause because the play button has been changed to a pause button when the play button has been pushed to start the video.

6.1 Identified low-level User Tasks

Based on the analysis of the data collected by uLog and the screen capture, it was identified 11 low-level user tasks over two main categories: general and video player (Table 6.1). The video category contains all actions that are generated from the video player and the general category is actions created from the user.

Category	User Task
General	Mouse click
	Typing
Video Player	Play
	Stop
	Pause
	Seek
	Sound adjustment
	Mute
	Unmute
	Fullscreen On
	Fullscreen Off

Table 6.1: Identified Low-Level User Tasks

These low-level user tasks can all be identified from the low-level data collected with uLog. However, some of the user tasks need some decrypting before it could be identified what kind user task it was. User tasks could be logged as cryptic words and not e.g. full screen. Therefore, the user task must be decrypted before it could be identified.

6.1.1 Cryptic element names in low-level data

The on-line TV providers have a technical issue when creating Internet pages to serve the purpose of showing video content. If they want to use different kind of low-level data to identify what users are doing, the different elements on the Internet page and in the actual video player should be named in a way that can be understood by others. After the analysis it has been identified that both Internet pages and the video players in some cases are using non-descriptive names on elements. E.g. when the user double clicks inside a video player to get full screen mode it is logged with EVRVideoHandler in uLog. The user went to full screen mode when the combination of EVRVideoHandler and mouse double click occurs. However, one cannot conclude that the user now is in a full screen mode. This is because different Internet pages use different video players and therefore not always will call it EVRVideoHandler. Furthermore, EVRVideoHandler do not identify anything about what was clicked; it only identify that something

called EVRVideoHandler was clicked. The only way to understand what the meaning of EVRVideoHandler, is to have a basic technical knowledge of the player or do analysis of the data collected with respect to recorded video or other observations.

There are also some events in the uLog data which are named Custom and therefore is almost impossible to identify what kind of event it was. In some cases the event do have a comment that can identify the type of the event. In other cases when there is no comment present it is only possible to know what the action is in an experiment context. This is because it is possible to supplement the uLog data with the recorded video.

This has also been mention by (Thomas and Karahasanović, 2009) to be an issue. A string can apply in many situations in the same application, e.g. merely "Microsoft Internet Explorer", or be common to many applications. Therefore (Thomas and Karahasanović, 2009) have created a set of guidelines, and guideline G6 is the one that treats this issue. Guideline G6 is described as: *"Design window titles, paths and other strings (to be captured from the system under investigation) in such a way as to minimize ambiguity and maximize contextual information when viewed outside the originally expected context."*

6.2   Identified high-level User Tasks

Based on the analysis of the data collected by uLog and the screen capture, it was identified four high-level user tasks. Each user task is detailed described, additionally it is also shown if the user tasks could be identified from low-level data. The identified user tasks are:

  1 - *User watching content where the user is actually watching a video.*

  2 - *User browsing content where the user is browsing for a any video interesting.*

  3 - *User searching for some specific content to watch.*

  4 - *User posting content on social medias.*

6.2.1   User watching content

The participant is watching content when a video starts playing. The high-level user task can be identified in the low-level data by looking on when the play button was clicked or the video was started. The problem here is to know when to set the participant in watching mode. A video can start even though the participant not are going to watch the video. E.g. on the newspaper page the video was started regardless of whether the participant was going to see the video or not. Furthermore, the participant may seek to a point in the video to start where they ended last time they watched the video. It was

identified that when a participant was in full screen modus was a video being watched. However, this cannot by it self say if the participant is watching a video. It should be combined with play and if the video is running.

Full screen mode

At an early stage in the experiment it was thought that the full screen mode was used only for watching content. In most cases the participants was watching content when they was in full screen mode, however one did choose full screen mode when browsing for information in a video. Full screen mode can therefore not by it self identify that a user is watching some content. Even though full screen mode was in most cases used by the participants to watch video (4/5 participants) and it will in most cases give the correct answer. Participants had different comprehension of why to use the full screen mode or not. Some examples of participant's statements:

> "I used the full screen mode because I was just interested in the video. The other content around the video is just disturbing."

> "I used the full screen mode when I watch video because I don't want to watch video on a small screen."

Both participants, from the statements above, watched the video in full screen mode, however the second participant was also seeking in the video. In most cases full screen mode can identify that the participant is watching the video or going to watch it.

Identification from low-level data

There is, at this time, no complete answer how to identify when the user is watching a video. The lack of data and to know when to set the user in watching mode is the main problem. However, one way is to depend on when the user has clicked the play button, clicked a link that start a video or entered full screen in combination where the video has been started.

If the video is starting without the play button has been clicked, it will be hard to know from the low-level data if the video has started playing. The problem is to identify what kind of links that are going to start a video. E.g. on the on-line TV page the participants can choose a range of entry points to a video. On the other hand the provider of the video will know when the video player makes a request for the video. Therefore it should be no problem for the provider to know when a video has started. However, even thought the service provider knows if the video is started or not it is still hard to identify if the participant is watching or not. The participant may choose to seek the video, do other actions on the computer or just wants to listen to the sound from the video.

6.2.2 User browsing content

When a participant is browsing for content the intention is to find any content that can fill some dead time[2]. This user task will only occur when the participant does not know what to watch. This high-level user task can be identified from the low-level by analyzing how the participant found some content. When the participant is browsing, one will not use the search functionality due to not knowing what one want to watch and therefore not knowing what one should search for. However, a reason to search is when the genre is known and the participant will search after e.g. comedy.

Identification from low-level data

The participant is browsing content when the low-level data records hyper link clicks. It is important that it is hyper link clicks and not e.g. a search button with search text that was clicked. This will place the participant in the searching user task. As mention earlier the participant may want to search for a specific genre. The question here is if the participant is still in a browsing user task or if the participant now is in a searching content user task. The participant is determined to watch a specific content in a genre, however the participant does not know what to watch in that specific genre. Therefore the participant is in a user task where he is browsing and searching at the same time. This must be further investigated to determine if it should be a high-level user task that combines both user task or if the participant is in a searching user task, if one knows anything about the content to watch.

6.2.3 User searching content

When the participant is searching for content the intention is to find some specific content. The participant will know the name of the content and be more targeted. Therefore it will be more likely that the participant will take use of the search functionality and this will be the best way to find out if the participant is searching for some specific content. However, some participants choose to click their way to the content even thought they had the name of the content. This is related to how participants are used to find content on the Internet pages. More about these findings can be read about in Section 6.3.

Identification from low-level data

In the low-level data it is in some cases possible to identify when the participants is using the search functionality. However, in most cases it is impossible, mainly because of the missing naming of the elements. The only way to know if the participant is

---

[2]Time in which someone or something is inactive or unable to act productively.

searching some content is when the search button is pushed. When the participants searches with writing some text and pushes the return button on the keyboard it cannot be identified if the participants is searching. It cannot be identified duo to the problem that it is not unambiguous that the participant is searching when the combination of typing and return is used. The combination of typing and return can also be used when the participant wants to change the browser URL.

6.2.4 User posting content on social medias

Most Internet pages have the possibility to post content on various social medias[3]. It is often some type of button or a link that the participant can click. In some cases participants did chose not to use these buttons, instead just copy the link and post it by them self. If the second option is used it is hard for the providers of the content to identify if some content has been shared on social medias.

Identification from low-level data

From the low-level data it must be investigated what link was clicked to identify if the participant is posting some content. In the other case where the participants choose to copy the link and share it by them self, it must be identified by looking on the link the participant is typing in the browser window. If the address text the participant typed in the address box in the browser is equal to e.g. `www.facebook.com` and the contents address has been copied, the participant may want to share some content. However, the providers cannot do this without permission from the participant because of privacy protection[4] issues.

6.2.5 Understanding the user task

It has been created a state diagram[5] (Figure 6.1) to show how the low-level user tasks is interfering with the high-level user-tasks. It shows what different user tasks the user is able to use in the different cases. E.g. when the user use the play user task, it will be disabled the play user task and it will be possible to use pause or stop user task. If the pause or stop user task is used, the play user task will be able to use again. Furthermore, the use of the play user task will move the user over to a watching user task. When the

---

[3]Social media are media for social interaction, using highly accessible and scalable communication techniques. `www.wikipedia.org/wiki/social_media`

[4]Privacy protection concerns the right to keep your private life to yourself, a fundamental right in a society ruled by justice. The ideal is that everyone should be in full possession and control of his or her own personal information. `regjeringen.no/en/dep/fad/Selected-topics/Privacy-protection.html`

[5]State diagrams show system states and events that cause transitions from one state to another. They do not show the flow of data within the system but may include additional information on the computations carried out in each state.(Sommerville, 2011, p135)

pause or stop user task the user will be moved into a browsing or searching user task. The boxes show the different states that the user can be in and all boxes are surrounded with a rectangle. The upper part where the user task play has been used, the user is in a watching user task/state. The lower part is the user in a browsing/searching user task/state because the stop or pause user task are used.

The user task's sound and seek are separated as own boxes because they are identical for all states. If the user choose to adjust the sound or seek the video, the user will be returned back to the same state the user were in before one started to adjust the sound or seek the video. As seen in figure 6.1 the sound box is outside both high-level user tasks/states. This is because it do not interfere with the user watching or browsing content. On the other hand the seek is inside the high-level user task/state of browsing/searching. The seek user tasks takes the user out of the watching user task/state until the user is finished with the seeking. When the seeking is done the user will be returned into a watching user task/state.

The initial state of the diagram is in the bottom left corner. Here, the user do not watch any content and it is very likely that the user is browsing or searching for some content. The user then have some possible paths; *play*, *mute* and *fullscreen on*. The possible user tasks the user can take use of are located close to the active state box. It can also be identified as the parts in the box that is not scribbled. When the user choose to use the user task play, the user will both be placed in another state box and switch the high-level user task to watching. The new state box will now have scribbled the play user task and removed the scribble from stop and pause user task.

Watching

Browsing / Searching

Sound[1]

Seek[2]

Play	Stop
Pause	Seek
Mute	Unmute
Fullscreen on	Fullscreen off
Sound	Link clicked

Fullscreen on / Fullscreen off

Stop/Pause

Mute / Unmute

Play
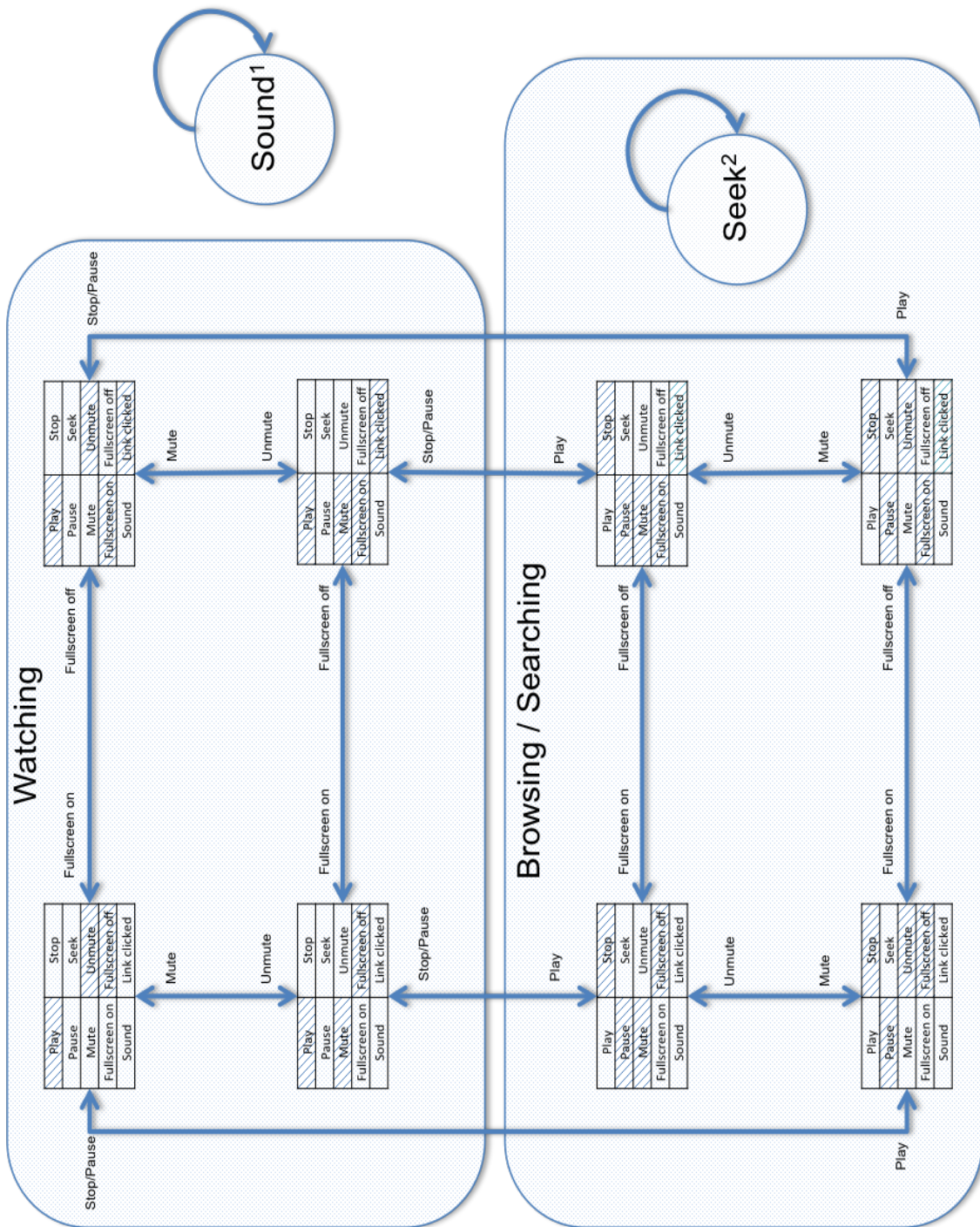
1. Sound adjustments do not interfere with the context. E.g. if the user is watching is he still watching after a sound adjustment.
2. Seek do interfere with the context. E.g. if the user is watching and starts to seek, will it trigger a context change. Watching to not watching while he seeks. (If in not watching context, the context change will not be triggered.)

Figure 6.1: State diagram with on-line TV user task

6.3 Problems with identifying high-level user task from low-level data

High-level user task 2 (User browsing content) and high-level user task 3 (User searching for specific content) is hard to differentiate from each other. The main reason is that participants choose to find content in different ways. Some like to click their way through content, and other like to search their way to content. When the participant is clicking around, e.g. browsing content, it can not be identified if the user actually is just browsing or looking for a specific content.

Participants choose different approaches to find the respective video in each assignment. How they decided to find the right video was influenced by three properties: web page design, what kind of page[6] and how the participant is used to find content. It is well known that people's use of computers change over time and that new and unexpected uses are found for most new systems. (Nielsen, 1994)

The first used service provider in the experiment is an on-line TV provider that got the award for best web place in 2010. The content is restricted in some mater due to copyright reasons and therefore the content will not be that extensive. The on-line TV provider has a good overview of the content and are representing it in a good way, due to this the participants will have more ways to find the content they are looking for.
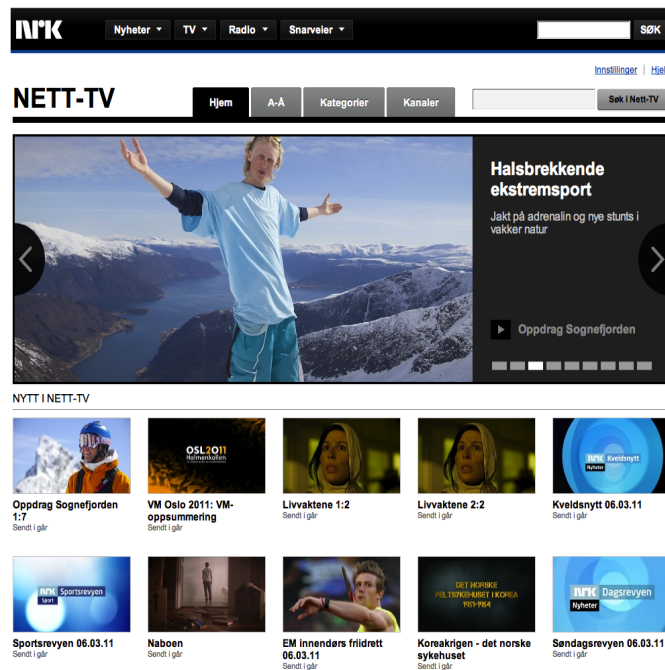

Figure 6.2: On-line TV used in experiment

---

[6]If the page's intention is to show only news, videos, news and videos, etc.

The second service was an online newspaper with the opportunity to combine both news article and video. They do not have the problem with copyright, for the reason that they create the content their self. However, due to the large quantity of news articles, with or without video, makes it difficult for the participant to locate the wanted content.
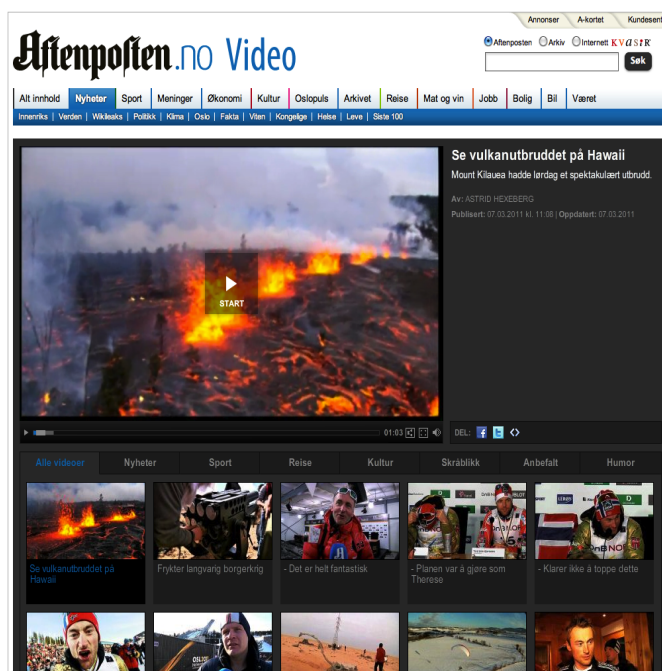
Figure 6.3: Online newspaper used in experiment

On the on-line TV three out of five choose to find the content with the search box in all assignments. One of those three was choosing the search option because he did not have enough knowledge about the interface. Two out of five choose to find the content by using the navigation on the web page because they felt that it was easy to find the right content.

Furthermore, on the newspaper page all participants choose to search the newspaper page. The main reason is because there is no good way to get an overview of the content with video. There is only a small range of the videos that are categorized, and it may also be hard for the users to know what category a specific video is in. The categories are also inconsistent, where e.g. the category *all videos* just contains the 15 newest videos and not all videos.

Some participants did give some statements about how they found the right content:

> "The easiest way to find the right content is to use the letter, A-Å link, overview. Further on just click from there until the right content is found.",
> On-line TV page

> *"On both sites I choose to use the pages search function, further I used the internal browser search option.", on-line TV page and newspaper page*

> *"On the on-line TV page I choose to use the letter overview. On the newspaper page I just searched for the right content with the search function.", On-line TV page and newspaper page*

In addition to the statements, the desktop recording did prove that participants in most cases did do what they stated. However, some of the participant's did use some click navigation even thought they stated that they only used search options.

How the participant found the content was also influenced by how the participant is used to find content. From the statements on page 65, all participants did state that they normally searched after content on other pages or searched by themselves with specific search words.

The participants were more active on the newspaper page. From figure 5.6 it can be observed that the participants is more active. Column 7 to column 11 one can observe that a participant is more active than in column 1 to column 6. Some of the reasons for this are that the newspaper page contains articles with short videos, and therefore must the participants more often find new content to watch. One the on-line TV page is the purpose to provide users with episodes in series and these will in general never be shorter than the videos on the newspaper page. Therefore, will the users be less active when they have found some content on the on-line TV page, than they will on the newspaper page (because the video clips will be shorter and they will need to find some new content).

6.4   User types in on-line TV

Participants choose different ways to identify the answers to each assignment in the experiment. In general there were two ways to find the answer, either to watch the whole video or seek through the video. In table 6.2 it is presented that the participants was using most time playing/watching a video. The participants that were a watching user type had a small seek count and the participant that was a searching user type had a high seek count.

6.4.1   Watching user type

In general all participants did choose to watch the whole video when they were needed to answer any questions. Some participants in the interview stated the reason for this after the experiment was conducted. They were worried that if they seek the video they would miss the position in the video where the answer to the questions was. All did seek

some in the video, however it was mostly in the start of the video or when they were getting bored by the video.

6.4.2 Search user type

As stated in section 6.4.1, participants normally was just watching the video, however it was one participant, see Participant 3 in Table 6.2, that stood out from the rest. Table 6.2 shows the time used in each behavior in the observation.

	Participant 1	Participant 2	Participant 3	Participant 4	Participant 5
Play	1419,09	1638,55	1895,33	1742,62	1789,56
Pause	324,4	50,22	82,87	65,62	77,6
Seek	64,47	19,33	200,6	72,13	88,1
Stop	4,33	25,77	-	7,13	-
No video	591,79	781,69	585,83	854,06	669,83

Table 6.2: Time duration in user tasks (In seconds)

Participant 3 used over the double amount of time seeking the video versus all the other participants. This indicates that the participant may be a person that likes to search for the content or just was being impatient. However, it was observed, during the experiment, that participants started to seek when they did not find the answer after some time watching the video. Participant 3 had both the longest seeking time, and participant 3 did spend the longest time to complete the experiment. The relationship between the complete time and seek time is that participant 3 in many cases did seek past the point where the answer was in the video.

6.5 Home environment versus experiment

Participants operate in a different mater depending on how they like to find information. In a home environment the participants would not search for a specific video like in the experiment. In the interview the participant did state how they worked in home environment versus in the experiment:

> "When I want to find a video at home, I would look at links that have been posted by users on www. reddit. com. And I would choose the videos with the highest ranking. Furthermore, when I need some information I would search for specific words connected to what I need. I always have a goal when I'm searching."

> "At home I will never search for a specific video. I would rather just see the one that shows up on the front page. Further I would just keep on browse the content. However, if I was going to find some information related to school or work would I just search for what I need."

> "At home I would search with the help of Google. For example I would search for a specific content on Google with "program :site the_site". I then expect to get directly to the right content."

From the interviews the participants do state that they in most cases would never search for specific content by them self, however they would rather watch what other users have been watching before them. E.g. from recommendation or from the most rated video clip.

6.6 Time specific findings

Different time specific findings have been identified and will be presented in this section. It is interesting to investigate what the participant did and for how long they did it. Firstly, the different assignment completion time is presented, additionally time consumption of different user task is presented.

6.6.1 Assignment completion time

Assignment completion time has been formed from the time when the participant started an assignment until it was completed. The completion time for assignment 1, assignment 5 and assignment 6 was always around 300 seconds because it was stated in the assignment description that the participant should be interrupted after five minutes. Furthermore, it was up to each participant to decide when the assignment was completed.
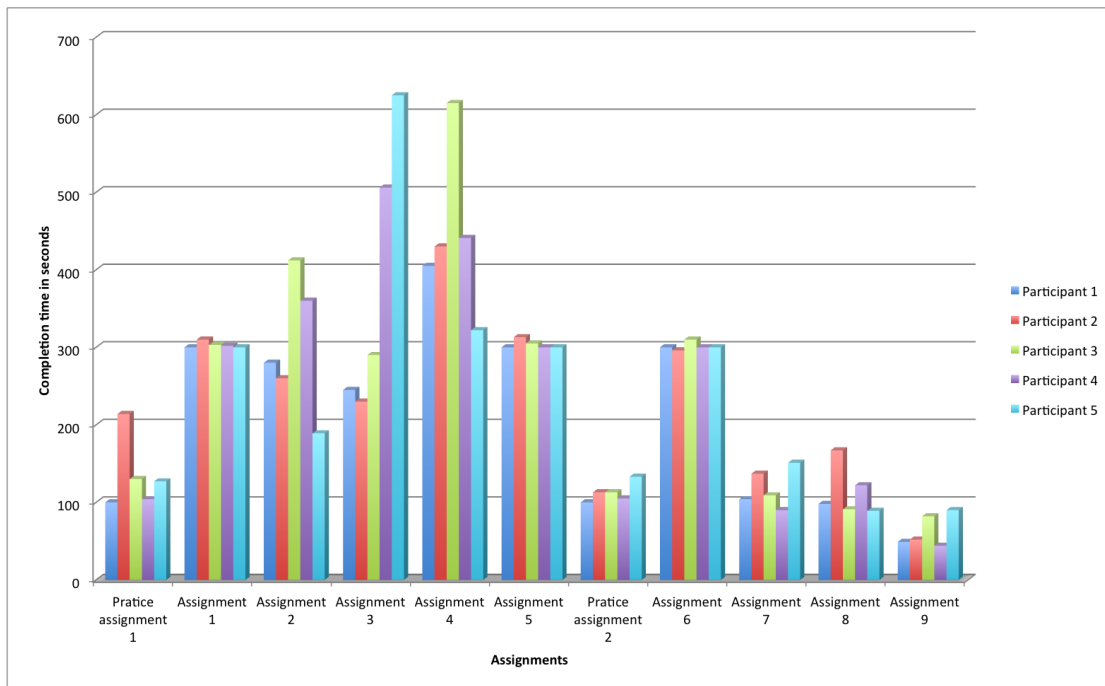
Figure 6.4: Assignment completion time

From figure 6.4 it can be seen that participants differs a lot in completion time on some assignments. Assignment 2 through assignment 4 has a huge differ in completion time. Assignment 3 differs most with one participant completing it in 230 seconds and another participant completing it in 625 seconds. Each participant interprets an assignment in different ways. Thus each participant will comprehend in a different way how detailed an answer should be. E.g. the participant may feel the need to watch the whole video or just the first few minutes. The length in answers do not differ much, however the answers from participants that looked through the hole video did contain more detailed information.

> I didn't know how detailed the answer should be, so I wanted to watch most of the video until I felt that the answer was answered good enough.

> I felt I was finished when I had an cursory answer. I didn't think I needed to go into more detail.

6.6.2 Low-level user task consumption time

Figure 6.5 illustrates how each low-level user task consumes time. Each low-level user task is a state, therefore it is possible to identify how long the participants is in a state. The no video state is not a user task, nevertheless it is used to distinguish between when a video is playing and not.
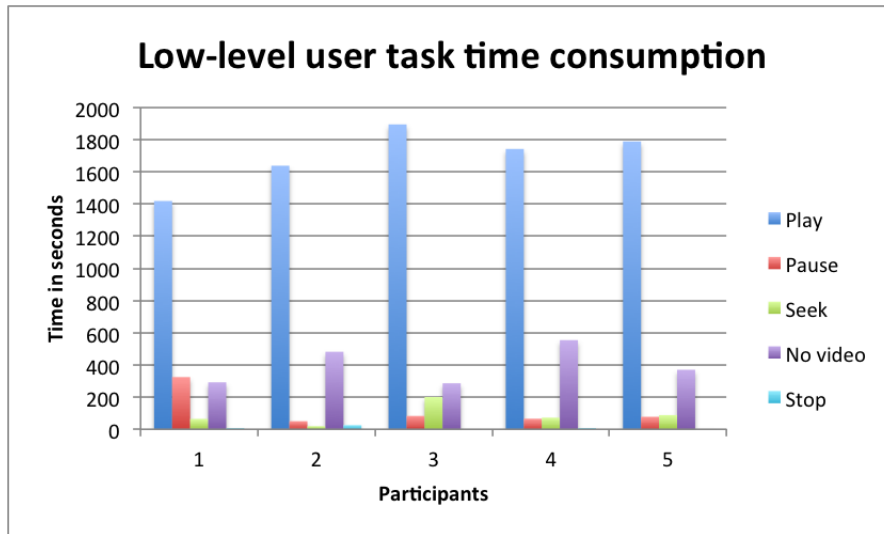
Figure 6.5: User task consumption time

From figure 6.5 it can be identified that the participant is mostly in the state where they are watching some content. Furthermore, it can be observed that participant almost never stops a video, however if it is stopped the pause button is used. The reason that the stop button is not used, is because when the participant is finished with watching a video the participants click a hyper link and the video is stopped. Therefore the no video user task will be activated and not the stop user task. It can also be mention that the play button is almost never used, since the video starts by it self in most cases. The video will start automatic depending on the entry point to a video. By entry point do one mean what linked that is click to find the video. One link may lead to an Internet page where the video doesn't start by it self or that the video does start by it self.

Since the stop button is almost never used it is a possibility that is no more needed in a video player. The pause button provides the same functionality as the stop button and therefore if it is removed one will not lose any functionality. However, it should be taken in consideration that the stop button is a well-known function and by removing it, may lead to confusion.

6.7 Discussion

In this section I will firstly discuss my work and thereafter-in relation to earlier findings. Secondly, discuss one exception to the regular on-line TV watching behavior. Lastly, discuss one relevant method to identify user task in online video streaming services.

6.7.1   Discussion in relation to earlier work

The identified user task is mostly the same that has been identified in (Kunert and Kunert, 2009) and (Darnell, 2007). However, some of the user tasks identified in (Kunert and Kunert, 2009) and (Darnell, 2007) has been identified to be irrelevant. This is related to that some of the user tasks depend on the system and not the general TV context. E.g. to go to recording, delete recording and skip ads. Furthermore, (Kunert and Kunert, 2009) do has a more high-level view of the user tasks, where (Darnell, 2007) have a more low-level view. This distinguish can be a result of difference between how the user tasks has been collected. Where (Kunert and Kunert, 2009) is using focus groups to identify user task, (Darnell, 2007) is using different recordings and interviews. A result of this is that when taking use of focus groups is the user task on what the goal is and when recordings are used, the user task is more on how to achieve the goal.

In on-line TV communication with other participants is not that much used, like it is in iTV. However, this is a phenomenon is getting more used in on-line TV. It is always possible to communicate with other users indirectly with posting content on e.g. www.facebook.com, however it is not that normal to communicate with other users or the broadcaster on a live session in on-line TV. On-line TV is starting to take use of this when larger events is taking place. They let the users interact with other users and a specialist in the field during the live video feed. This was last done under the ski world championship and was only possible to get access to if the users chose to watch it on the on-line TV service. Furthermore, all participants in the experiment did use social networking and are posting content more than once a month. Therefore it does look like that the integration with social media and communication, should be further pursued in development of on-line TV services.

The user task that has been identified in the experiment is more technical and is further used to identify what the users want to do without their interaction. The user task has been categorized in two general categories, low- and high-level user tasks. When comparing the identified low- and high-level user tasks with the user tasks from (Kunert and Kunert, 2009) and (Darnell, 2007) all identified user tasks can be transfered over to both iTV and TV with DVR. Therefore the distance between the different TV services may be smaller than first expected. The difference lies more in how the the service is developed, with respect to the different functionality, rather than in how TV is used in different services. Of course users do use the service different when the functionality differs, however the general TV related user task like play, pause and seek will be the same across all services.

In (Kunert and Kunert, 2009) some of the user task are so general that they are almost useless. The user task *starting the application* and *exiting the application* is general for all possible applications that is not worth mentioning when it should be obvious that a user must start the service in some way, before using it. Therefore these user tasks has not been used or identified in this thesis, when it is obvious that this is something the user must do to use the service.

Furthermore, the identified user task has been used to create a visualization (Figure 6.1) to further understand how they relate to each other. It is a rather custom state diagram, where more than one state is included inside one state. This gives better overview and less states to handle. With use of a normal state diagram should all states be one state, therefore will the diagram be more complex and increase in size. None of the earlier studies has used this method to show how the user tasks relate and hopefully other will use it in later studies with or without relation to on-line TV. Furthermore, should it be mentioned that the diagram is not a fully developed. It contains only what I have identified and additional identification of user task should be added to the diagram to be fully able to understand how all user task affects each other. It will also make it possible to extend and define the high-level user task even more. This will increase the probability to identify if a user is watching, browsing or searching for content.

6.7.2   Affects on user behavior in on-line TV

How the participants accomplished the assignments on the different Internet pages used was influenced by some sources. Firstly, the participants were affected by how the Internet pages were designed. Secondly, how long the video length was and lastly, how they were used to find content.

Internet page design affecting user behavior

During the experiment it has been identified that users are more active on the newspaper page versus the on-line TV page. One early indication from the users is that the newspaper page had two navigation bars, when one was browsing the video content. They indicated that they were insecure of which navigation bar to use and some times ended up with using the global navigation instead of the video specific navigation. When the global navigation was used and e.g. *sport-link* were click. The participant will be taken to the sport section of the Internet page and not the sport section inside the video section. This was identified to be a design issue/usability problem and was creating more interaction with the Internet page/service than was needed by the participants to complete an assignments. Even thought that some participants had been on the newspaper page before, they still stated that the page was poorly designed and they were irritated by it. Therefore, some participants may have been more active than they intended to be. It will both effect the collected data and it will be harder to identify what the participant wanted with an action. This is a topic that may be left out when one is trying to identify how users interact with a service. In the context of on-line TV may the service be embedded in an Internet page (as in this experiment) and this may be forgotten. The service is no longer just they video player, however also the Internet page. The users are then not only interacting with the video player, however is now also interacting with both the video player and the surrounding elements. Where the

elements could be hyper links, pictures and so on. Therefore, should the video player and the surrounding elements in some way be identified as one element.

If one want a more focus on how users just is interacting with the video player and the on-line TV in general, should the surrounding elements be extracted and only the video player be left on the Internet page. This will turn the focus over to the video player and the irritating elements will no longer be a problem. Furthermore, the data that will be collected is now on how users interact with the video player. It will then not be a combination of the video player and the Internet page design.

Habit affecting user behavior

During the interview some of the participant in the experiment did state that they e.g. search the Internet page rather than clicking their way to the content they were looking for. Furthermore, some participants stated that they choose to search both Internet pages because they did not have enough knowledge about the navigation structure. As stated by the participants does the habit affect how to find content one is looking for. Therefore, users that normally navigate by search are less likely to use hyper links as their primary navigation. Even thought it is identified with the collected data, it cannot be concluded before it has been studied even more.

Video length affecting user behavior

On the newspaper page were the videos shorter in length than on the on-line TV page. This is obviously because the on-line TV provides video programs to the users (that have been shown on a normal television). In difference to the newspaper page that provide the users with a short video to supplement the news articles that is presented with text. It has been identified that when the video length is short the users are forced to find new content earlier than if the video length is long. New actions/behaviors must be used before one has found some new content to watch. Because of this will a user be more active on the newspaper page versus the on-line TV page. Furthermore, must one consider that even thought the video length is longer on the on-line TV page, may the user be very active until one has found some interesting content. However, when the user has found some content to watch, the actions will stop until the video is finished or the user is interrupting the watching session in someway.

6.7.3 The exception to regular on-line TV usage

When the user has started a video are one forced to assume that the user is also watching it. Because to know if a user is watching the video or doing other actions on the computer is hard, if not impossible. After checking this up with experienced system developers, I

almost got the same answers from all the developers. The quote under summarizes the feedback to the problem best:

> *In practice it is very difficult, if not impossible. There is no dedicated event in the browser when you replace the window. You may be able to cheat a kind of "inverse check", if you still have the focus or not, but this is by no means trivial, and probably very different implementation from browser to browser.*

Furthermore, when the video is playing and the user only is listening to the sound, this is a problem that cannot be identified. It is the same problem as mention earlier. A good example where the sound is the main purpose to *watch* the video is the latest World Championship[7] in Norway that was hold the last month. Most of the World Championship started during the work hours each day. However, this did not stop people *watching* the World Championship. They solved the problem with work by just letting the live video run in the background and just listen to the sound. Whenever the reporters said something interesting the video was watched for a short while.

After a few days of observation of my own work place was it identified that most peoples used this trend. Furthermore, the feedback from my co-workers did confirm my suspicions. Under is some of the feedbacks from my co-workers to the question: "How do you watch the World Championship at work?".

> *I was watching the World Championship in intervals, this way I also could do some work.*

> *I just let the video go in the background of my browser while I did other work related stuff. Whenever I heard the commentator said anything about the Norwegian participants I switched to the video player and watched it for a while.*

This is of course an exception to the regular usage of the service and in most cases will only occur when big events are taking place.

6.7.4   Different method of user task identification

The way to identify user task / user behaviors, that was done in (Qiu and Cui, 2010), has not been used in this master thesis. However, some of the ways of thinking have been used, e.g. that after a period of time after the video has started, the user will be in a watching state. The reason is the type of low-level data that was used. We had the possibility to identify what the user did just by looking on the name of the action, e.g. play button clicked. Furthermore, we did not have the possibility to get access to the video requests from our logging programs. The results from (Qiu and Cui, 2010) clearly shows that it is a possible way to identify what the users are doing. There should also not be any problems with mapping the sessions to behavior patterns in real time.

---

[7]Nordic world ski championship. `www.oslo2011.no`. Hosted 23. February - 6 March.

The main reason to consider this approach is the time consumption. The data used for analysis is automatic gathered without any interaction with the users of a service (e.g. where one has experiment with participants that need to be observed, like the experiment in this thesis). Therefore, one can use the time one used on the experiment on other parts of the study. The time consuming part in (Qiu and Cui, 2010) is the generation and extraction of the data. Furthermore, the analysis process will be as time consuming as in any other studies. The downside with e.g. a database with trace records is the lack of understanding of the users. However, (Qiu and Cui, 2010) do focus more on the technical aspects of understanding user task / user behavior, than I have done. Where I have been focusing on the low-level data to identify user tasks, with the help from the users to understand the intentions behind each user task. Even though (Qiu and Cui, 2010) do not understand the users in the same extent as studies that include users in the experiment, one could develop a system design with this approach and carry out user tests to understand the users. Based on the discussion, this approach should be further investigated. Identify if this approach can identify user tasks in an on-line TV service and could be applied to the system design of an on-line TV service.

Chapter 7

Validity

The most important threats to the validity of the experiment are discussed in this chapter.

7.1 Experimental Design

In this experiment we wanted to log all user actions in order to identify if it was possible to understand what the participants was doing with the help of low-level data. Most materials was found online, e.g. the Internet pages used that contained all videos etc. All other materials were presented on paper to the participants.

We wanted to make the experimental environment as close to a home environment as possible. This is in general hard since the participants know that they are in an experiment. Even thought the participant stated that they didn't think about the surroundings after a short while into the experiment.

All assignments were given on paper and it was two assignments on each side of the paper. None of the assignments was depending on each other and therefore should not be a problem. The only problem that some participant stated was that they, in a home environment, would not use the on-line TV like they did with the assignments. However, they did think that the assignments were realistic.

7.2 Assignments

All assignments was created by me and further approved by my supervisor, in collaboration with the service providers. The main problem with the assignments is that one can never know if the content used in the assignments will be removed. Therefore was

the video content in the assignments selected specifically, because the content was created by the provider and did have a date it should be removed from the Internet page. Furthermore, the service provider did indicate that the videos without a date would in most cases not be removed

Participants approved all assignments to be easy to understand and not that hard to complete. The only problem that was mentioned was about assignment 2 and assignment 3. The assignments were created in a way that the participant did not know when the answer was well formed. This lead to that some participants ended the assignment fast, where other was undetermined and used longer time to complete the assignment.

7.3 Participants

All the participants were informatics students that will complete their theses this semester (age 23 to 26). They may be more technical than the normal user in Norway. Some of the participant states that they did not use on-line TV services at a normal basic. Even thought they will be more adaptive to new technology since they have grown up with technology. A group of people with different technology background will give better knowledge of what people are doing and give a better variety of low-level data.

7.4 Questionnaire and Interview

My supervisor did created the questionnaire and interview in a previous study. Both the questionnaire and interview was redefined to fit this experiment. In the interview it was supposed to be an interviewer and an observer. This was not always the possibility in my interviews due to difficulties to fit the schedule to both my supervisor and me. In the later studies it should be two person in the interview so important questions is not let out and two persons will be able to interpret what the participants is stating in different ways. The interviewer and the observer then can provide the interview with even more questions to reveal even more relevant information. Even though it can also be only one interviewer since there is only one person that should be interview at once.

7.5 Analysis

All data collected by uLog have been extensively used in this study with the video recording as a supplement. With the uLog data it occur different range of problems. The process of data cleaning and preparation was difficult. Even though I carefully verified all data, there may be actions that have been left out of the analysis. All actions from uLog have been subjective analyzed and may therefore not have been understood

in the right way. Furthermore, some uLog data was lost when a participant was done with the experiment.

The interview was used to understand what the participant's intention were and some basic information about how they worked. All recorded sound from the interview were transcribed for the purpose of being able to quote what the participants stated in the interview. Some words and expressions may have been transcribed wrong, because it was done only by me. Considering the amount of data collected the coding and analysis should be done in collaboration with other people. This will give a more precise and valid analysis. At this point the analysis by me will be used as a foundation in the main experiment, however it will most likely be identified some problems in the main experiment, that must be taken in consideration.

Chapter 8

Conclusions and future work

The objective of this research was (1) to identify how users interact with a on-line TV service, (2) which user task and user behaviors that are taken use of when watching, browsing and searching video content, (3) can user tasks be identified from low-level data and (4) how different logging tools can help use to answer the previous questions. To address these issues, we conducted a controlled experiment with five MSc at the Department of Informatics at the University of Oslo.

The experiment has been conducted to provide future studies with insight in how on-line TV users interact with a system. This has led to the identification of 4 high-level user tasks and 11 low-level user tasks. There has also been identified two kinds of user types that are closely related to the high-level user tasks. All high-level user tasks (Section 6.2) and low-level user tasks (Table 6.1) are more detailed described in the results.

8.1 Contribution

This section presents the contributions of the research to the R2D2 project.

8.1.1 Findings with respect to earlier research

In this thesis one of the main issues have been to identify related work. At this point, the research lacks in the identification of the user task and user behavior in the context of on-line TV. There has been a range of studies on user task and user behavior in the context of Internet in general. However, in the context of TV and video there is only a small amount of related work. It has been identified ten articles related to user task/user behaviors in general. Six articles related to user task/user behaviors in relation to video and only four studies that has identified some user task/user behaviors in the context of TV. One out of the four studies has in the context of video streaming purposed a

possible way to identify what user task/user behavior a user is trying to achieve, with the help of video sessions/requests.(Qiu and Cui, 2010) In relation to the work that has been done in this thesis, the method used in (Qiu and Cui, 2010) will consume less time. However, they do have a less focus on understand the users because they do not include users in the study. They do only focus on the collected data that has been collected from the service providers database.

The identified user task differ some what from what was identified in (Bernhaupt et al., 2007), (Darnell, 2007) and (Kunert and Kunert, 2009). This is related to data collection concepts that have been used in the articles. All was focusing on iTV and interviewing the participants. Only (Bernhaupt et al., 2007) did some recording, however this was only video data and was analyzed in cooperation with the participants. I have taken this further with identifying user tasks from low-level data in an on-line TV context.

8.1.2 Creation of a study

We have created a study that includes the design of a controlled experiment, how to prepare the collected data for analysis and how the data can be analyzed. The experiment design were created to provide us with events/actions that could identify when the user is watching, searching and browsing for content. Furthermore, a questionnaire and a interview where created to provide us with additional information about the users (the questionnaire) and what the intentions behind the collected actions from uLog were (the interview).

Furthermore, I have identified how to prepare and take use of the collected data. Both the Observer XT and Microsoft Excel has been used to analyze the data. With the Observer XT I was just able to conduct only a cursory analysis. Therefore, with the help of Microsoft Excel I was able to take the analysis further. I had to create a short JAVA program where I converted the events/actions to numerical IDs to create graphs that could illustrate what the users is doing. The small JAVA program is not presented in this thesis because it was not that complex and included only different combinations of string comparison. I have with prepared data been able to present how to identify user tasks from the combined low-level data and the coded video. In addition was the illustrations (graphs) created with Microsoft Excel provided a more detailed overview of the low-level data collected with uLog.

With respect to the identified user tasks I have been able to create a state diagram (Figure 6.1) to present how the individual low-level user tasks affects the high-level user tasks. By affecting I mean how the low-level user tasks can move the user from a watching state to a browsing/searching state. The state diagram differs somewhat from the regular state diagrams. This is because I include a more than one state in a state. Combining states in a state box does this and provides a clearer overview of how the low-level states affect the high-level user task.

8.1.3 Identified User Tasks

The high-level user tasks all depend on the lower low-level user tasks. Each high-level user task is both easy and hard to be identified. The easy way is when it can be read right from the collected low-level data what the user is doing. To be able to identify it the easy way, the user must follow a path where clicking does all. E.g. the user is watching some content when the user has clicked a range of links and at some point clicked the play button. The problem starts when the play button is not clicked (because the video can start by it self, when entering the Internet page containing the video). This leaves no action in the log; only the action when the user clicked a link is logged before the video starts. Therefore, if one do not know that a specific link leads to a video, one can not identify that a user is watching some content. This problem needs further investigation to develop an ability to identify when a video has been started.

The best way at this moment to identify if a user is watching some content, is when the full screen mode of the video is activated. When the full screen mode is activated the user is interested in only the video, however the user can still e.g. seek in the video. When the user is seeking in the video, the user will not watch the video until the seeking has been completed.

8.1.4 Affects on user behavior

Different affects on the user behavior have been identified and should be considered. There have been identified three things that affect the user behaviors in the context of on-line TV. Firstly, the participants were affected by how the Internet pages were designed. Secondly, how long the video length was and lastly, how they were used to find content. Because of these affects the user will behave in a different way than they would if they not were affected. The Internet page design can cause annoyance and move the focus from the assignments to the design of the page. Therefore, one should consider this and one could also consider presenting a stand-alone video player to the participants in an experiment. Since the user differ that much because of the length of the video it may be considered to divide it into two independent parts. Where one part is on the newspaper page (containing the short videos) and the second part is on the on-line TV (containing the longer videos). How the users is used to find content, one cannot do anything about. This must just be dealt with when conducting the analysis.

8.2 Difficulties

Several difficulties have been identified during the analysis process. This is mostly connected to how to identify the user task from the low-level data. In some cases uLog is reporting different actions without a good naming on the action. This is related to the

players the service providers are using. Some elements are named in an incomprehensible way, where it in some cases is not possible to understand the action before it is combined with the video recording. This leads to the problem where it in some cases is not possible to identify the user task. Even if it does it will not work in a general matter. Video players used by the service providers differ and elements will not be named in the same way.

It has been hard to identify the connection between the different actions logged by uLog to the right user task. One part of the problem is the identification of the context the actions happened in. It is in some cases hard to identify the context of an action, because of the naming of the window change. Some windows may have the same name, and therefor it will be difficult differentiate them. This must be further investigated and solved in later studies.

The data preparation has been a demanding task where it is hard to get all data in the right form and to fit all together. Even thought programs state that they e.g. handle a video format it should be converted a small file to verify that the conversion settings works. It is not always the case that one setting type will work, because it may be one setting that is wrong or the manual is not updated.

8.3 Future Work

The use of on-line TV service has been in a rapidly growth the last few years and users are starting to explore the flexibility of the on-line TV service. There is still much research that will take place over the next few years within this area and the research presented in this thesis can serve as a foundation for further research that will try to identify what users are doing with only the help of low-level data.

A more automatic way to analyze the data should be examined. The parts where the excel sheets and diagrams is made manually, should be done automatic. This can be done with different programming languages with the help of statistical frameworks and libraries. This makes the time used decrease drastic and make the analysis process less complex.

Future research should take use of the identified user task to develop a on-line TV service. This can then give the service providers the opportunity to personalize the on-line TV service with e.g. user profiles. Each user can then be given a better experience, where the service can adapt it self to the environment.

The next step in the R2D2 Networks project will be to conduct the main study to further explore the use of low-level data in context of on-line TV. Further use of my result in combination with the main study, one will be able to further investigate how user interact with the on-line TV service. Additionally, the different user task could be explored in more detail and it is possible to define them more in-depth.

Bibliography

S. Acharya and B. Smith. Characterizing user access to videos on the world wide web. In *In Proceedings of MMCN*, 2000.

E. Adar, D. S. Weld, B. N. Bershad, and S. S. Gribble. Why we search: visualizing and predicting user behavior. In *WWW '07: Proceedings of the 16th international conference on World Wide Web*, pages 161–170, New York, NY, USA, 2007. ACM. ISBN 978-1-59593-654-7. doi: http://doi.acm.org/10.1145/1242572.1242595.

E. Agichtein and Z. Zheng. Identifying "best bet" web search results by mining past user behavior. In *KDD '06: Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 902–908, New York, NY, USA, 2006. ACM. ISBN 1-59593-339-5. doi: http://doi.acm.org/10.1145/1150402. 1150526.

J. Attenberg, S. Pandey, and T. Suel. Modeling and predicting user behavior in sponsored search. In *KDD '09: Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 1067–1076, New York, NY, USA, 2009. ACM. ISBN 978-1-60558-495-9. doi: http://doi.acm.org/10.1145/1557019.1557135.

S. Baluja, R. Seth, D. Sivakumar, Y. Jing, J. Yagnik, S. Kumar, D. Ravichandran, and M. Aly. Video suggestion and discovery for youtube: taking random walks through the view graph. In *WWW '08: Proceeding of the 17th international conference on World Wide Web*, pages 895–904, New York, NY, USA, 2008. ACM. ISBN 978-1-60558-085-2. doi: http://doi.acm.org/10.1145/1367497.1367618.

F. Benevenuto, F. Duarte, T. Rodrigues, V. A. Almeida, J. M. Almeida, and K. W. Ross. Understanding video interactions in youtube. In *MM '08: Proceeding of the 16th ACM international conference on Multimedia*, pages 761–764, New York, NY, USA, 2008. ACM. ISBN 978-1-60558-303-7. doi: http://doi.acm.org/10.1145/1459359.1459480.

R. Bernhaupt, M. Obrist, and M. Tscheligi. Usability and usage of itv services: lessons learned in an austrian field trial. *Comput. Entertain.*, 5, April 2007. ISSN 1544-3574. doi: http://doi.acm.org/10.1145/1279540.1279546. URL `http://doi.acm.org/10.1145/1279540.1279546`.

B. Burr. Vaca: a tool for qualitative video analysis. In *CHI '06 extended abstracts on Human factors in computing systems*, CHI EA '06, pages 622–627, New York, NY, USA, 2006. ACM. ISBN 1-59593-298-4. doi: http://doi.acm.org/10.1145/1125451.1125580. URL `http://doi.acm.org/10.1145/1125451.1125580`.

Celtic-initiative. Road to media-aware user-dependant self-adaptive networks. `http://www.celtic-initiative.org/~pub/Project-leaflets/Webquality/r2d2-lq.pdf`, 2009. [Online; accessed 4-October-2010].

K.-C. Chen and W.-G. Teng. Adopting user profiles and behavior patterns in a web-tv recommendation system. In *Consumer Electronics, 2009. ISCE '09. IEEE 13th International Symposium on*, pages 320 –324, May 2009. doi: 10.1109/ISCE.2009.5157045.

Z. Cheng, B. Gao, and T.-Y. Liu. Actively predicting diverse search intent from user browsing behaviors. In *WWW '10: Proceedings of the 19th international conference on World wide web*, pages 221–230, New York, NY, USA, 2010. ACM. ISBN 978-1-60558-799-8. doi: http://doi.acm.org/10.1145/1772690.1772714.

I. D. Coman. An analysis of developers' tasks using low-level, automatically collected data. In *Proceedings of the the 6th joint meeting of the European software engineering conference and the ACM SIGSOFT symposium on The foundations of software engineering*, ESEC-FSE '07, pages 579–582, New York, NY, USA, 2007. ACM. ISBN 978-1-59593-811-4. doi: http://doi.acm.org/10.1145/1287624.1287715. URL `http://doi.acm.org/10.1145/1287624.1287715`.

A. Crystal and B. Ellington. Task analysis and human-computer interaction: approaches, techniques, and levels of analysis. In *x*. Citeseer, 2004.

Y. Cui, J. Chipchase, and Y. Jung. Personal tv: a qualitative study of mobile tv users. In *EuroITV'07: Proceedings of the 5th European conference on Interactive TV*, pages 195–204, Berlin, Heidelberg, 2007. Springer-Verlag. ISBN 978-3-540-72558-9.

Y. Daabaj. The use of task analysis methods in support of the development of interactive systems. In *x*, page 781, 2000. doi: 10.1109/ICSE.2000.870502.

M. J. Darnell. How do people really interact with tv?: naturalistic observations of digital tv and digital video recorder users. *Comput. Entertain.*, 5, April 2007. ISSN 1544-3574. doi: http://doi.acm.org/10.1145/1279540.1279550. URL `http://doi.acm.org/10.1145/1279540.1279550`.

J. W. Evans and C. Filsfils. *Deploying IP and MPLS QoS for Multiservice Networks: Theory & Practice*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2007. ISBN 0123705495.

J. D. Gould and C. Lewis. Designing for usability—key principles and what designers think. In *CHI '83: Proceedings of the SIGCHI conference on Human Factors in*

Computing Systems, pages 50–53, New York, NY, USA, 1983. ACM. ISBN 0-89791-121-0. doi: http://doi.acm.org/10.1145/800045.801579.

A. Gutschmidt and C. H. Cap. User behavior under the microscope - can the behavior indicate a web user's task? In *WEBIST (2)*, pages 215–222, 2008.

A. Gutschmidt, C. H. Cap, and F. W. Nerdinger. Paving the path to automatic user task identification, 2008.

A. Hassan, R. Jones, and K. L. Klinkner. Beyond dcg: user behavior as a predictor of a successful search. In *WSDM '10: Proceedings of the third ACM international conference on Web search and data mining*, pages 221–230, New York, NY, USA, 2010. ACM. ISBN 978-1-60558-889-6. doi: http://doi.acm.org/10.1145/1718487.1718515.

D. M. Hilbert. A survey of computer-aided techniques for extracting usability information from user interface events, 1998.

ISO. Iso 9241-11:1998 - ergonomic requirements for office work with visual display terminals (vdts) - parts 11: Guidance on usability. `http://www.iso.org/iso/catalogue_detail.htm?csnumber=16883`, 1998. [Online; accessed 13-January-2010].

J. A. Jacko and C. Stephanidis. *Human-Computer Interaction: Theory and Practice (part 1), Volume 1 (Human Factors and Ergonomics)*. CRC Press, 1 edition, September 2003. ISBN 0805849300. URL `http://www.amazon.com/exec/obidos/redirect?tag=citeulike07-20\&path=ASIN/0805849300`.

L. Kantner. Following a fast-moving target: Recording user behavior in web usability testing, 1998. URL `http://www.teced.com/PDFs/commonlk98.pdf`.

A. Karahasanović, B. Anda, E. Arisholm, S. E. Hove, M. Jørgensen, D. I. Sjøberg, and R. Welland. Collecting feedback during software engineering experiments. *Empirical Softw. Engg.*, 10:113–147, April 2005. ISSN 1382-3256. doi: 10.1007/s10664-004-6189-4. URL `http://portal.acm.org/citation.cfm?id=1046956.1046966`.

T. Kunert and T. Kunert. User tasks and requirements for itv applications. In *User-Centered Interaction Design Patterns for Interactive Digital Television Applications*, Human-Computer Interaction Series, pages 85–98. Springer London, 2009. ISBN 978-1-84882-275-7. URL `http://www.springerlink.com/content/h844642102489608/`.

U. Lee, Z. Liu, and J. Cho. Automatic identification of user goals in web search. In *WWW '05: Proceedings of the 14th international conference on World Wide Web*, pages 391–400, New York, NY, USA, 2005. ACM. ISBN 1-59593-046-9. doi: http://doi.acm.org/10.1145/1060745.1060804.

M. G. Lervold, L. Xing, and A. Perkis. Quality of experience in internet television. 2010.

J. Liu, P. Dolan, and E. R. Pedersen. Personalized news recommendation based on click behavior. In *IUI '10: Proceeding of the 14th international conference on Intelligent user interfaces*, pages 31–40, New York, NY, USA, 2010. ACM. ISBN 978-1-60558-515-4. doi: http://doi.acm.org/10.1145/1719970.1719976.

M. Maia, J. Almeida, and V. Almeida. Identifying user behavior in online social networks. In *SocialNets '08: Proceedings of the 1st Workshop on Social Network Systems*, pages 1–6, New York, NY, USA, 2008. ACM. ISBN 978-1-60558-124-8. doi: http://doi.acm.org/10.1145/1435497.1435498.

K. Miyauchi, T. Sugahara, and H. Oda. Relax or study?: A qualitative user study on the usage of live mobile tv and mobile video. *Comput. Entertain.*, 7(3):1–20, 2009. ISSN 1544-3574. doi: http://doi.acm.org/10.1145/1594943.1594955.

C. Montabert, D. Scott McCrickard, W. W. Winchester, and M. A. Perez-Quinones. An integrative approach to requirements analysis: How task models support requirements reuse in a user-centric design framework. *Interact. Comput.*, 21(4):304–315, 2009. ISSN 0953-5438. doi: http://dx.doi.org/10.1016/j.intcom.2009.06.003.

J. Nielsen. Goal composition: Extending task analysis to predict things people may want to do. *Available via:[February 2010]*, 1994. [Online; accessed 24-February-2010].

E. Noam, J. Groebel, and D. Gerbarg. *Internet television.* Lawrence Erlbaum Assoc Inc, 2004. ISBN 0805843051.

Noldus_ObserverXT. Noldus observer xt information. `http://www.noldus.com/human-behavior-research/products/the-observer-xt`, 2010. [Online; accessed 8-April-2010].

Noldus_uLog. Noldus ulog information. `http://www.noldus.com/ulog/features`, 2010. [Online; accessed 8-April-2010].

F. Qiu and Y. Cui. An analysis of user behavior in online video streaming. In *Proceedings of the international workshop on Very-large-scale multimedia corpus, mining and retrieval*, VLS-MCMR '10, pages 49–54, New York, NY, USA, 2010. ACM. ISBN 978-1-4503-0166-4. doi: http://doi.acm.org/10.1145/1878137.1878149. URL `http://doi.acm.org/10.1145/1878137.1878149`.

M. Rauterberg and M. Fjeld. Task analysis in human-computer interaction - supporting action regulation theory by simulation, 1998.

K. Renaud and P. Gray. Making sense of low-level usage data to understand user activities. In *Proceedings of the 2004 annual research conference of the South African institute of computer scientists and information technologists on IT research in developing countries*, SAICSIT '04, pages 115–124, , Republic of South Africa, 2004. South African Institute for Computer Scientists and Information Technologists. URL `http://portal.acm.org/citation.cfm?id=1035053.1035067`.

S. H. Rogne. Msc. motivation and user satisfaction - a study of web-tv users. May 2010.

C. B. Seaman. Qualitative methods in empirical studies of software engineering. *IEEE Trans. Softw. Eng.*, 25:557–572, July 1999. ISSN 0098-5589. doi: 10.1109/32.799955. URL http://portal.acm.org/citation.cfm?id=322677.322687.

J. Shaikh, M. Fiedler, and D. Collange. Quality of experience from user and network perspectives. *Annals of Telecommunications*, 65:47–57, 2010. ISSN 0003-4347. URL http://dx.doi.org/10.1007/s12243-009-0142-x. 10.1007/s12243-009-0142-x.

H. Shin, M. Lee, and E. Kim. Personalized digital tv content recommendation with integration of user behavior profiling and multimodal content rating. *Consumer Electronics, IEEE Transactions on*, 55(3):1417 –1423, 2009. ISSN 0098-3063. doi: 10.1109/TCE.2009.5278008.

I. Sommerville. *Software Engineering (9th Edition)*. Pearson Addison Wesley, 2011. ISBN 0137035152.

G. Szabo and B. A. Huberman. Predicting the popularity of online content. *Commun. ACM*, 53(8):80–88, 2010. ISSN 0001-0782. doi: http://doi.acm.org/10.1145/1787234. 1787254.

Y. Takama and Y. Muto. Profile generation from tv watching behavior using sentiment analysis. In *Web Intelligence and Intelligent Agent Technology Workshops, 2007 IEEE/WIC/ACM International Conferences on*, pages 191 –194, 2007. doi: 10.1109/WI-IATW.2007.17.

The_Nielsen_Company. State of the media - tv usage trends: Q3 and q4 2010. http://www.nielsen.com/us/en/insights/reports-downloads/2011/state-of-the-media-tv-usage-trends-q3-and-q4-2010.html, 2010. [Online; accessed 15-March-2010].

R. C. Thomas and A. Karahasanović. Experience with an extensible workspace for analysis of low-level usage data. *Softw. Pract. Exper.*, 39:1185–1213, September 2009. ISSN 0038-0644. doi: 10.1002/spe.v39:14. URL http://portal.acm.org/citation.cfm?id=1600267.1600269.

K. Torii, K.-i. Matsumoto, K. Nakakoji, Y. Takada, S. Takada, and K. Shima. Ginger2: An environment for computer-aided empirical software engineering. *IEEE Trans. Softw. Eng.*, 25:474–492, July 1999. ISSN 0098-5589. doi: 10.1109/32.799942. URL http://portal.acm.org/citation.cfm?id=322677.322682.

M. Vilas, X. G. Paneda, R. Garcia, D. Melendi, and V. G. Garcia. User behaviour analysis of a video-on-demand service with a wide variety of subjects and lengths. In *EUROMICRO '05: Proceedings of the 31st EUROMICRO Conference on Software Engineering and Advanced Applications*, pages 330–337, Washington, DC, USA, 2005. IEEE Computer Society. ISBN 0-7695-2431-1. doi: http://dx.doi.org/10.1109/EUROMICRO.2005.63.

C. Wohlin, P. Runeson, M. Höst, M. C. Ohlsson, B. Regnell, and A. Wesslén. *Experimentation in software engineering: an introduction*. Kluwer Academic Publishers, Norwell, MA, USA, 2000. ISBN 0-7923-8682-5.

R. K. Yin. *Case study research. Design and methods, 2nd ed.* SAGE Publications, Thousand Oaks, London, New Delhi, 1994. ISBN 0-8039-5663-0.

Youtube.com. Youtube statistics. `http://youtube-global.blogspot.com/2010/03/oops-poq-surprise24-hours-of-video-all.html`, 2010. [Online; accessed 4-October-2010].

X. Yuan, Z. Dou, L. Zhang, and F. Liu. Automatic user goals identification based on anchor text and click-through data. *Wuhan University Journal of Natural Sciences*, 13:495–500, 2008. ISSN 1007-1202. URL `http://dx.doi.org/10.1007/s11859-008-0423-7`. 10.1007/s11859-008-0423-7.

M. Zapater and G. Bressan. A proposed approach for quality of experience assurance of iptv. In *Digital Society, 2007. ICDS '07. First International Conference on the*, page 25, 2007. doi: 10.1109/ICDS.2007.4.

Appendix A

Background Questionnaire

This questionnaire need to be filled out before the experiment. The information will exclusively be used as part of the experiment, and will be treated confidentially.

1. How old are you?

_____

2. Sex
   ☐ Female
   ☐ Male

3. Are you in education?
   ☐ Yes
   ☐ No

4. Highest education completed
   ☐ Primary
   ☐ Secondary
   ☐ Collage
   ☐ University
   ☐ Other _____

5. How often do you watch Aftenposten TV?
   ☐ Daily
   ☐ More often than once a week
   ☐ More often than once a month
   ☐ Less than once a month
   ☐ Never

6. How often do you watch NRK NETT-TV?
   ☐ Daily
   ☐ More often than once a week

☐ More often than once a month
☐ Less than once a month
☐ Never

7. How often do you use any social media? (Facebook, twitter, etc.)?
☐ Daily
☐ More often than once a week
☐ More often than once a month
☐ Less than once a month
☐ Never

8. How often do you share videos with others through Facebook, twitter or e-mail?

☐ Daily
☐ More often than once a week
☐ More often than once a month
☐ Less than once a month
☐ Never

Appendix B

Interview Guide

The interview guide translated for use in this thesis and were originally written in Norwegian.

BEFORE:

Assistant

- Checks batteries and voice recorders

- Brings:

    - Voice recorder, microphone (battery in it)

    - Pen and paper

    - Interview guide (2 pieces, one for the interviewer, one for the observer)

- At the end of each the day, should the files be downloaded from the voice recorder to the Observer PC (Import the files into the Noldus Observer XT). NOTHING IS DELETED FROM VOICE RECORDER!

DURING AND AFTER:

The interviewer [Amela]:

- Leads the interview based on the guide and records it on the voice recorder. If it interview stops, shall the observer break in.

- Print background information about the participant, so one can read it before the interview starts.

The observer [Thomas]:

- Monitors

- If it interview stops, shall the observer break in.

- make sure everyone gets to speak and monitors the time.

- Take notes (paper, PC, smart pen), do not go overboard with this, one have audio recordings.

- Notes are stored on the Observer PC (imported into the Observer PC)

Interview Guide
One provides the participant with the first level question. Level 2 is only used if they do not mention it.

PART I (10 minutes)

- The method - How you worked with the tasks?
  - Describe the process (what did you do first, second, etc.)
  - What procedures considered to use?
  - Do you work different from normal (similarities / differences)
  - Are you used to watch on-line TV (NRK online-TV; Aftenposten TV)?
  - How was it to find the content you were asked to find?
  - How was it to share content?
  - What was especially easy, difficult? Why? What was difficult? How did you solve the difficulties?

PART I (5 minutes)

- How was it to participate in the experiment?
  - Time pressure, stress
  - The information you got (mail info, information at the start of the experiment, help along the way)
  - Task formulations
  - Difficulty level / realism of the tasks,
  - Logging and the presence of an observer (were you distracted, was hard to write)
  - (All answers for themselves) More on the participants background, if anything is interesting (we have their background form)

Appendix C

uLog settings

Table C.1 shows what uLog can log and Table C.2 shows what uLog will write to the log file when the session is finished. The settings uLog is providing have one column for each top-level setting and the low-level settings in it. The settings used in the experiment are listed in the table besides each setting with a YES or NO statement.

Keyboard		Mouse		Other	
String	YES	Mouse click	YES	Address change in browser	YES
String event/-Timestamp	YES	Wheel	YES	Windows Active	YES
Special key	YES	Drag	NO	Message Box	NO
Key combination	YES	Hover	NO	Application started	NO
Single character	YES	Control Focus Changed	NO	Application ended	NO

Table C.1: What should be logged: uLog settings input

General		Mouse event		Control info	
Date	YES	Mouse button pushed	YES	Type	YES
Time event/- Timestamp	YES	Wheel turn direction	NO	Caption (lable or name)	YES
Milliseconds	YES	Cursor X	NO	Value	YES
Event type	YES	Cursor Y	NO	Password	YES
Event Description	YES			Control framework	NO
		Distance (rel)	NO	Control class	NO
		Distance (tot)	NO	Controll image	NO
				Control type specific info	NO
Event Info		Keyboard		Other Info	
Application	YES	Value	YES	Clipboard content	NO
Window Text	YES	Modifiers	YES	Browser URL	YES
				Custom event info	NO

Table C.2: What should be written to file: uLog settings output