Energy-Use Feedback Engineering

Technology and Information Design for Residential Users

Zur Erlangung des akademischen Grades eines
Doktors der Wirtschaftswissenschaften

(Dr. rer. pol.)

von der Fakultät für
Wirtschaftswissenschaften
am Karlsruher Institut für Technologie (KIT)

genehmigte

DISSERTATION

von

Anders Dalén, M.Sc.

Tag der mündlichen Prüfung: 22.10.2015
Referent: Prof. Dr. Christof Weinhardt
Korreferent: Ph.D. Thomas Lawrence (Univ. of Georgia)
Prüfer: Prof. Dr. Thomas Setzer

Karlsruhe, 2015

To Gwen

Acknowledgements

I am very grateful for the opportunity to acknowledge the many people who have helped me to complete this finished thesis manuscript. Firstly, I wish to thank my supervisor Professor Christof Weinhardt for his insightful input over the years that this project was completed. The theoretical expertise and support of my advisors, Professor Jan Krämer and Dr. Christoph Flath, has moreover been vital to this study. I am much obliged to Professor Thomas Setzer and Professor Thomas Lawrence for their thorough consideration of the whole manuscript and for their very constructive feedback during my final seminar.

At the Institute of Information Systems and Management in Karlsruhe, this study has benefited greatly from regular work-in-progress seminars; I am very appreciative of the generous feedback that I have received from all past and present members. Over the years and in every situation, I have been very thankful for the efficient aid of Susanne Heidenreich, Yolanda Hermida and Anette Seiter with all matters administrative, practical and logistical. I also wish to express my gratitude to the bachelor and master students I had the good fortune to supervise: Tobias Ahlers, Johannes Beck, Christoph Großbeier, Henning Quellenberg, Michael Schilling and Carlo Schmidt. Thank you for all that you taught me. To my fellow doctoral students I owe my heartfelt appreciation for years of help, encouragement and great times. I am especially thankful for the friendship and support of Sebastian Chiriac, Dr. Sebastian Gottwalt, Dr. Christian Haas, Dr. Margeret Hall, Anuja Hariharan and Dr. Alexander Schuller. Every visit to the pool, every stroll in the park and every coffee and baked good helped.

Being funded to carry out the inspiring and challenging work of a doctoral dissertation is a privilege. My studies have been supported primarily by the Federal Ministry of Economics and Technology (BMWi) projects MeRegio and iZeus. I have also been incredibly fortunate to receive a innovation stipend from the Karl-Steinbuch Fund together with my dear colleague Dr. Timm Teubner. Furthermore, I am very thankful for the support I received from the Karlsruhe House of Young Scientists (KHYS), which enabled me to carry out research in an international setting. This experience enriched my studies immeasurably. The grant allowed me to spend a term as a Visiting Scholar at UC Berkeley to work in close contact with Professor Dan Wertheimer, develop my research and take part in international intellectual exchange.

I wish to thank all of my wonderful friends and family for their care and welcome visits during the years I have spent completing my doctoral studies. The NyTeknik job advertisements from and the laughs shared with my brother Johan lifted my spirits. I am deeply grateful for the constant backing of my parents, Tore and Ulla. I count myself lucky to have had so much time with you; your unflinching belief in my abilities will always give me the courage to stretch for new opportunities.

I am so thankful for the perspective that my sons, Ruben and Seth, have brought me. Finally, this thesis is dedicated to my wife and best friend Gwen. Her unshakable loyal support steadied my resolve and made this journey possible.

Contents

List of Figures

List of Tables

Part I

Fundamentals

Chapter 1

Introduction

TRIBUS AND MCIRVINE (1971): *"The flow of energy in human societies is regulated by the tiny fraction of the energy that is used for the flow of information."*

BRILLOUIN (1962): *But however tiny, the intricate relationship between energy and information, which make them interchangeable, govern the second law of thermodynamics.*

THE CURRENT ENERGY SYSTEM is largely based on non-renewable sources for generating energy. By definition, these sources will eventually be depleted or not be economically viable for extraction (Sweeney, 1993). The looming climate crisis further amplifies the urgency with which alternatives, preferably renewable ones, must be found. This situation has led to much research and commercial activity to change the current energy-resource composition.

Currently, energy generation from renewable sources is the fastest growing power source in relation to the present capacity worldwide (IEA, 2013). The share of renewable energy is expected to be 25% of the gross energy available globally in 2018, which equals a 40% increase of generation from 2012. Most of this increase is based on on-shore windpower and photovoltaics (IEA, 2013), which are dependent upon weather conditions and the time of day. This means that the ability to manage energy generation is gradually diminished. The energy system, which once was regulated by controlling the supply side, must now activate the demand in this balancing act.

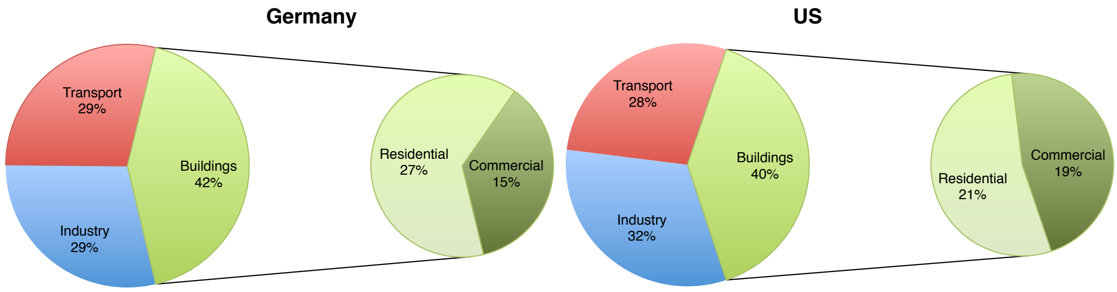

Figure 1.1: German and U.S. energy usage by sector.

The renewed interest in demand-side management (DSM) will help to close the gap between supply and demand. More research is needed to better understand the benefits and challenges of integrating the demand in the system and what information and types of communications technology (ICT) are necessary to activate this potential source for energy regulation (Strbac, 2008; Palensky and Dietrich, 2011).

Energy usage in buildings is one potential focus for DSM activity since households utilizes a substantial portion of the world's generated energy. In Germany the energy usage the building sector amounted to 42% in 2012 (AGEB, 2014), while in the U.S. 40% was used in the same time period (EIA, 2014). Between residential and commercial buildings, the former encompass the main part of this energy use, which is depicted in Figure 1.1. Of the global final energy use, 23% was used in the residential sector, while 27% of the total electricity was used in the same category (IEA, 2014).

Another reason to regulate energy demand in the residential sector is that user behavior has been shown to have a substantial impact on a household's energy usage, even in similar households in terms of size and inhabitant demographics. For example, in a sample of 28 identical townhouses in New Jersey, variation in energy usage was found to be as great as two to one (Socolow, 1978). Since these houses are identical regarding floor plan, interior layout, builder, construction materials, and climate, it is likely that most of the energy-use variance is due to the different behavior of the people in the houses. Furthermore, in houses where there has been a change of residents, it has been found that the energy use of the house with the new residents cannot be predicted from the energy use of the same house with the previous residents (Sonderegger, 1978). Finally, even after houses had been successfully retrofitted (with 20 - 25% savings), the variance in energy usage among the houses remained almost the same as it was before the retrofits took place and the order in which the households' energy-use could be sorted hardly changed (Socolow, 1978). In sum, these examples imply that choices of appliances and individual operation behavior has a significant influence on a household's total energy usage.

Primary energy use in the residential sector has been quite stable since the 1990s (AGEB, 2014). Residential energy users can conserve energy either by efficiency or curtailment measures (Gardner and Stern, 2002). Temporary shifting energy use is also becoming more important to follow intermittent renewable generation (Shaw et al., 2009). However, strategies to conserve energy are not utilized or leveled by other decisions. Technological innovation in energy efficiency is counteracted by the introduction of new appliances and energy services due to rebound effects and a growing economy (Nadel, 2012). This means that even though most major appliances are becoming more efficient, users are operating them more often or buying larger and additional appliances.

Energy-use information is a fundamental requirement to understand and analyze this situation. Unfortunately, attempts to motivate household inhabitants to use energy more actively to help balance the energy supply is inconclusive. Already after the first energy crisis, following the oil embargo in 1973 (Ford, 1997), attempts to influence and control individuals to use energy more conscientiously were tested with mixed success (McDougall et al., 1981). In the 1980s, more research focused on improving energy-use efficiency (Ford, 1997), and as institutions had formed to buffer against future interruptions (International Energy Agency, 2012). The individual's potential impact on the energy system became less important. Later, in the 1990s, the interest for energy feedback once more increased as the understanding of the global climate impact of burning fossil fuels grew.

Residential users generally state that they are interested in having more and better information about their energy usage (Chetty et al., 2008). However, the topic of providing useful and engaging energy-use information to improve the energy knowledge-level and enable effective decisions still needs further research and applied testing. Conclusions of smart-meter experiments caution that an introduction of more information has proven to be complex and not automatically lead to efficiency gains (Hargreaves et al., 2013; Darby, 2010a; Ehrhardt-Martinez et al., 2010).

Results of energy-use feedback research also depend on the specific context of that era. For example, Ehrhardt-Martinez et al. (2010) found that between the studies on energy-use feedback in 1970 and 1990, the results from the more recent, more slowly developing threat of climate change, show a consistently lower effect on energy usage than the former. This means that conclusions of feedback that might have been effective under certain circumstances, cannot be simply applied to the current context.

Over the same time span, between the 1970s to the present day, the general importance of information is becoming better understood. The recent fast development of ICT is the second main reason to revisit DSM. The previous informational void is replaced by energy interfaces capable of showing individual appliance's consumption (Grønhøj and Thøgersen, 2011), historical energy-use summaries (Wilhite et al., 1999) and predictions of the impact of alternative behaviors (Staats et al., 2004). Today the integration of renewable energy generation, storage solutions and energy efficiency depend on ICT to scale and become cost effective (EAC, 2008; Farhangi, 2010).

There are several research and commercial projects that continue to develop visualizations and flexible control of energy usage.[1] Investments in ICT for the energy system and DSM measures specifically, were estimated at between 2 and 3 billion Euro globally in 2009 and grew to 15 to 31 billion Euro in 2014 (Booth et al., 2010). In Canada and the US alone, the investments were already reported to be over 6 billion Euro in 2013 (CEE, 2014). One of the major reasons behind this investment is to decrease the lack of information by measuring and suppling energy information in a timely and flexible manner (Hart, 2008). In Germany it is mandatory for new buildings to install a digital energy meter that has the ability to supply near real-time information about the current household load (cf. Energiewirtschaftsge-

[1] For example: luciddesigngroup.com, tendrilinc.com, mysmartgrid.de, rrrevolve.ch, discovergy.com.

setz §21d Messsysteme). Sweden and Italy, among other countries, have already replaced all of their analog household electricity measurement devices. However, even with these resources, informing, engaging and sustaining residential users' interest with their energy-use information, it has proven to be a difficult and multifaceted issue. Two major public projects to gather and visualize energy-use data from Microsoft and Google have been retired due to lack of interest (Microsoft, 2011; Google, 2011).

Energy-use feedback can be given in several forms, ranging from market-based energy prices to artistic renderings (Faruqui and Sergici, 2009; Rodgers and Bartram, 2011). This thesis will focus on energy-use data as an information carrier and the impact of energy-use awareness for making effective energy usage decisions. This will be done by analyzing the potential of the current information system and its information content, which is currently being implemented at regional levels, and to explore the challenges and possibilities of further developing this information to improve the utility for users. The following section will derive research questions from related literature to clarify the current need for academic research.

1.1 Research Outline

The core research activity in this study concerns the provision of energy-use information, the analysis of the impact thereof on residential users and the further development of the information systems. The main questions pertaining to information-system research in the energy domain has been outlined by Watson et al. (2010) and encompasses the sensor network, flow network and sensitized objects and their stakeholders. This energy informatics' domain emphasizes the importance of effective solutions by focusing on "doing the right things" (Braungart et al., 2007, p.1342). Energy efficiency is important in order not to waste resources, however researchers are called to "do more than reducing the tempo of ecological destruction" (Watson et al., 2010, p.28). It remains to be seen how ICT can stimulate and support a more sustainable energy usage.

The first research question focuses the analysis on the new digital energy-monitors, which are being installed throughout energy systems worldwide. Specifically, the real-time aspect of the energy-use feedback that can be reached through a web-based system is evaluated, as this is the main feature provided to the residential users. The inquiry reflects the potential change from supporting energy awareness and its impact on energy usage. The question acknowledges the full range of outcomes from the potential of no change or even an increase in energy usage to an overall decrease in energy usage, which can come from becoming more efficient (using less energy in relation to the service provided) and making effective energy decisions (the ability to expend effort on actions with a high impact).

Research Question 1: Evaluation of smart-meter utility *What impact does online access to real-time energy-use information have on energy-use decisions in households?*

In order to analyze the coming energy feedback system, a sensor network that replicates the functionality of the current system is necessary. Additionally, how the information is being used has to be taken into account to understand user preferences and the need for technological development (Farhar and Fitzpatrick, 1989; Watson et al., 2010). Thus, in addition to information's impact on energy usage, the measurements and methodology also need to support an analysis of the user interaction with the given information, in order to describe information.

Based on the given requirements a sensor-network system and an analysis of residential users' interactions with the provided energy-use information is needed. By combining the information interaction aspect with the measured changes in the energy use, this study extends the current literature by evaluating the effectiveness of reported actions.

The second research question deepens the study of energy feedback in relation to the residential user. By evaluating and developing new informational functions based on the results from the first question, this extension focuses on increasing the users' utility from the provided information.

Research Question 2: User-centered energy feedback *Based on users' interaction and experiences with real-time energy feedback, what informational detail and content is required to align the energy feedback with their preferences?*

Augmenting the results from the initial research question with related literature clarifies the existing potential in current information systems. Based on the analysis of how users employed the information, it will also be more apparent how ICT should be implemented to support efficient and effective energy-use decisions. By integrating the evaluation of users' interaction patterns and challenges, a system that caters to their specific preferences can be developed (Wallenborn et al., 2011). This type of user-centric design has been suggested as critical for "the migration of electricity users to the demand response world" (Honebein et al., 2009, p.39), and provides a helpful perspective through which to evaluate the second research question.

Effective implementations of ICT mixes information on services and products with their related context (Ehrhardt-Martinez et al., 2010). However, instead of supporting the user, presenting too much information might instead overwhelm and confuse them (Malhotra, 1982). Such "information overload" has been shown to lead to less accurate and longer decision times, starting at levels of 10 different informational cues (Iselin, 1988). Access to a wealth of information has also been shown to decrease the overall satisfaction with the decisions made (Schwartz, 2004). Individual users have traditionally had limited access to information about their energy usage. However, as the access to and dissemination of information has been simplified through the development of ICT (through the installation of smart meters in residential buildings), the issue of information overload is becoming more prominent throughout society (Bawden and Robinson, 2008).

To mitigate this challenge, the presented information should be constrained (Schwartz, 2004). Obviously, removing alternatives is one way of limiting the presented choices. However, since individuals value their energy services differently at different times compared to other users, removing an alternative would risk loosing the access to a potential option. Alternatively, the information can be grouped to provide a constrained sub-category of choices. This alternative retains all the different options but can be sorted according to their energy impact. How the information should be grouped to provide enough detail to support satisfactory decisions yet limit information load requires further study, which is reflected in research question 3.

Research Question 3: User experience and information load *What level of information granularity and how many distinct cues lead novices to make high-impact decisions in few attempts and a short period of time?*

This question is specifically focused on how the presentation of information impacts the decision outcome. Information load can be evaluated based on its ability to lead to efficient (few decisions in a short time) and effective (high impact) decisions. By developing an interface experiment that explores the impact of grouping information on decision efficiency and effectiveness, recommendations for the presentation design can be given.

The following section contextualizes these main research questions into the study as a whole.

1.2 Structure

An outline of this study is presented in Figure 1.2. There are four major parts dividing the themes of the underlying chapters.

Part I encompasses the underlying theoretical and practical structure of this study. The first segment begins with Chapter 1 which motivates the topic of study, introduces the main research questions and explains the development of the work. Chapter 2 frames energy-use information research by explaining its characteristics. Specifically, the four influential features of information detail, feedback types, display methods and user attributes which cover the theoretical boundaries of this study, are explained consecutively. In Chapter 3, the opportunities, challenges and main research methods pertaining to the engineering and provision of energy-use information on a residential level is introduced.

Part II then develops and evaluates energy-use information on the currently considered household level. This part begins with Chapter 4 and the development of an experimental system for an aggregated household level that is easy to deploy and provides the necessary data to evaluate specifics about the user interaction. Chapter 5 implements this sensor system in the field. The resulting data on energy usage and users' interaction with the information is combined with interviews to reach conclusions on the specific patterns and missing functionality of an energy-use information system for residential users.

Part III then extends the information system to an individual user-level, based on the results from Part II. Chapter 6 explores two potential systems for appliance- and operational-level measurements which are required for a user-level system. The ability to scale the system for wide adoption is weighed against the level of information detail necessary. Appropriate methods for disaggregating appliances and operating modes are evaluated.

Chapter 7 then takes a design-science approach to define and develop a user-centered energy-use information system following international standards on user-centered design. The information system is conceptualized based on user requirements, and the design and the necessary supplementary information is generated and exemplified.

Chapter 8 considers how information diversity impacts users' ability to make informed and effective decisions efficiently or with little effort. This analysis is fundamental to a user-centered system that not only provides the expected information (established in Chapter 7), but also presents it in a way that facilitates effective decision making.

This thesis concludes with Part IV which highlights the study's contributions and opens questions for future research.

1.3 Research Development

Much of the work presented here has already been peer-reviewed and accepted for publication at international conferences. This section accounts for these publications and their relation to the given structure explained in the previous section. For clarity, each chapter also contains references to the specific related publications.

The sensor system exploration, development and evaluation was presented and published at the International Smart-Grid Technology (ISGT) conference in 2012 (Dalen and Weinhardt, 2012). The data processing and interface development, which used the measurements from the sensor hardware was then published at the 2012 International Conference on Advanced Collaborative Networks, Systems and

Applications (COLLA) (Dalén et al., 2012). The energy use monitoring system provides a portable sensor network platform that enables full access to the energy use and interface interaction parameters. Together, these two publications make up the fundamental system for the evaluation of current smart-meter technology and user interaction which is presented in Chapter 4.

The initial experimental framework was also included in the published COLLA article (Dalén et al., 2012). The final setup and experimental results were extended with a qualitative interview and published at the Erasmus Energy Forum 2014 (Dalén, 2014). Chapter 5 combines the experiment development in the former article with the evaluation and results from the final experiment detailed in the latter one.

The exploration of the extended sensor system is elaborated in two parts. First, a centralized sensor- and data-processing system to provide appliance-level information was evaluated and published at the biennial International Energy Conference (EnergyCon) 2014 (Dalén and Weinhardt, 2014). Second, the appliance level data was then further processed and disaggregated into the specific operational modes of specific appliances. This latter study is part of a larger, user-centered, design effort, which is currently under review at the Business and Information Systems Engineering (BISE) journal (Dalén and Krämer, 2014b). The research in these two studies use similar methods in their analysis and are combined in Chapter 6 to form the fundamental measurement infrastructure for Part III of this study.

The requirements and information system development is also included in the submitted article to the BISE journal (Dalén and Krämer, 2014b), and is fully elaborated on in Chapter 7.

Finally, the evaluation of the impact of informational load on residential users' ability to make decisions was conceptualized and published at the European Conference on Information Systems (ECIS) in 2013 (Dalén et al., 2013). The work in this article is further extended and evaluated in Chapter 8.

It should also be noted that a part of this work, specifically the experimental measurement and information system presented in Part II, has also been presented at the MFG foundation in order to fulfill the Karl-Steinbuch innovation scholarship requirements.

In summary, the research presented in this study covers a first design iteration of energy feedback for residential users. By analyzing how the current system is used, a user-centered development plan is detailed. This plan is then implemented and analyzed both in terms of its potential for providing utility to the user's decision-making process and in terms of how the information could be grouped so as not to overload and deter the user. This research contributes with a framework and new insights into the study of energy-use information for residential users, which exemplifies the challenges and potential of integrating information technology in this part of the energy system.

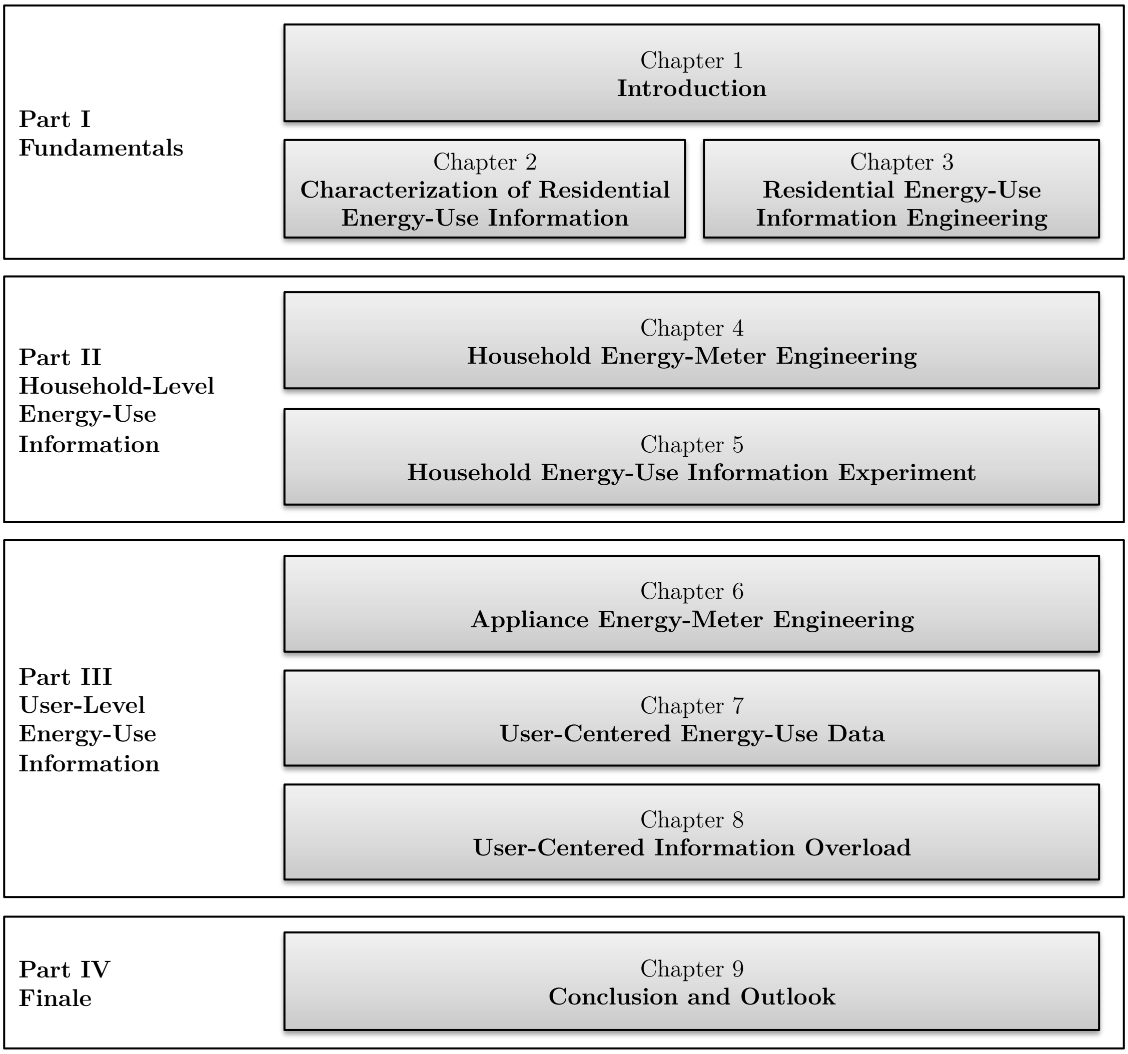

Figure 1.2: Structure of this thesis.

Chapter 2

Characterization of Residential Energy-Use Information

Energy-use metering in households is traditionally provided by the utility to settle the energy bill. Until recently the metering has been done by a mechanical system based on the magnetic flux from alternating current (AC) power. The meter was then read manually at a certain interval, for example once every year, and the flat monthly fee for the predicted energy usage was regulated based on the reading. Currently digital counterparts to the analog version are being considered, developed and distributed (ESMA, 2010). These new energy meters have been called "smart meters" or defined as part of an advanced meter infrastructure (AMI) (Darby, 2010a). While the name *smart meter* probably gives too much credit to a sensor that is providing energy-use data (Hauttekeete et al., 2010), the term will be used in this study to signify bidirectional information flow.

The main promise of the smart meter is to enable remote and frequent readings (Ellery E. Queen, 2011). This allows for direct billing and eliminates the need for manually reading and reporting meter values. For the utility the frequent readings also means that the energy can be charged for in a more dynamic way depending on time, critical events and available current generation capacity.

The cost of energy has traditionally been the main information carrier for energy-use information. However, with the developing information and communications technology a rich communications channel is opening up around energy, which previously has solely been a means to an end (Ayres and Warr, 2009). This particular data stream, from the meter to the user, is the primary focus throughout this study.

Such energy-use information can be used to make users aware of their households' energy need. This awareness is fundamental for making more knowledgable energy-use choices and, in turn, more sustainable actions (Mattle et al., 2011). Unfortunately, in practice unfamiliarity with the provided technical information, information overload, and a lack of means to interpret current energy use, make it difficult for the end user to develop a concrete plan of action. For example, a consumer survey issued by the German Federal Ministry of the Environment regarding environmental awareness in Germany revealed that 20 to 30% of the representative sample (n=2034) felt that a lack of transparency and facilitating support prevented even larger changes to the energy usage (e.g. by making more sustainable energy-use choices) (Kuckart et al., 2006). Socio-economic factors, such as age, gender, education level and income, showed no impact on the survey results, and all participants agreed that energy conservation should be actively pursued in everyday life.

The use of smart meter technology greatly facilitates the collection and exchange of information about private households' energy usage. As energy information become more pervasive and available over the internet, more sophisticated informational campaigns are possible. Several different information system designs are being evaluated for making traditional energy meter information available and more understandable (Darby, 2010c). For example, modern portable devices with internet access are improving the access to energy information (Weiss et al., 2010).

However, even though there is demand and new alternatives for providing energy-use information, recent field experiments supplying access to this type of information through web-based or dedicated displays have not been able to shown evidence of an altered energy usage (Thuvander et al., 2012; Hargreaves et al., 2013). Unfamiliarity with appliances and a lack of continuous support from manufacturers and installers often lead to wasteful operation (Darby, 2008). A recent review of advanced meter interfaces and energy information, therefore, concluded that the feedback has to be further developed to provide more “appropriate forms of interface feedback, narrative and support” in order to “reach diverse populations” (Darby, 2010c, p. 455). Consequently, it is evident that the design of new forms

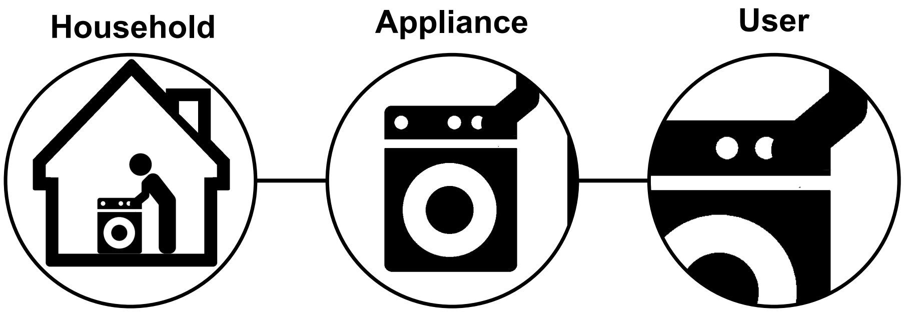

Figure 2.1: Overview of possible levels of energy-use information detail

of feedback to improve the efficiency and effectiveness of energy-use decisions is necessary to unlock the true potential of smart metering (Simmhan et al., 2011).

Energy-use information can take many forms depending on the context and its intended purpose. This chapter will present the characteristics of residential energy-use information. The focus will be on providing an understanding of the available parameters that exists for developing and changing this information to suit end-users' needs.

2.1 Level of Information Detail

Information on energy usage can be measured and given on different levels. Figure 2.1 shows the three levels, which will be in focus in this research study; household, appliance and user operation. Household level is the aggregated energy use from all the appliances and services connected with the residential energy supply. The appliance level provides energy use information of each appliance, irrespective of how the appliance is used, while the user level signifies energy use information of each operating mode of a specific appliance or service.

2.1.1 Household Level

Measuring energy use at the point where the mains supply the households' need for energy is convenient as it coincides with the utility's regulation of the energy usage of a specific household. Household-level information is also the most prominent choice for commercial and open-source projects that focus on supplying dwellers with energy-use information. This is also the level at which the new smart meter will be installed and replace the role of the analog meter.

A benefit from using household-level data is that it encompasses the two common energy-use behavior effects of rebound and spill-over. These effects are often missed when the data gathering of the study is too narrow to account for a general change (Abrahamse et al., 2005). The rebound effect describes a reaction that happens when an improvement in energy efficiency (or decrease in energy use) in one instance leads to an increase in energy use of another service. Thus, if only the heating system is measured and a lowered use of the main heating system leads to additional use of individual electrical heating comforters this rebound effect would go unnoticed (Nadel, 2012). Spill over effects has the opposite impact to the rebound effects. They are characterized as when an increase in efficiency in one particular service results in measures that also improve the efficiency of other energy services (Thø gersen and Ölander, 2003).

Because of the emerging prevalence of household smart-meters it is also appropriate to evaluate the impact of this level of information. The system used for providing household-level energy-use information will be presented in Chapter 4.

2.1.2 Appliance Level

The downside of household-level information is that it is disconnected from how users interact with their appliances. This separation makes it difficult to translate the household-level energy-use information into concrete actions. With more detailed information on specific appliances, users have been shown to change their behavior and lower overall standby power use and operate appliances more stringently (Ueno et al., 2006).

By combining smart-meter technology with (distributed) energy sensors, energy users receive disaggregated feedback on the energy consumption of every single appliance in their household (Froehlich et al., 2011). In principle, this information can be utilized to make users aware of their electricity consumption behavior and to induce more efficient energy consumption choices (Mattle et al., 2011).

Appliance-level feedback has been found to lead to an, on average, 9% decrease in energy use over one 3 month experiment (Ueno et al., 2006). A Danish study, similarly, concluded that their real-time, appliance-level, feedback successfully managed to motivate a change in energy use (Grønhøj and Thøgersen, 2011). This was confirmed both in their qualitative interviews and by comparing the experiment period to previous years. On average a 8.1% decrease in energy consumption was reported. How much influence the added detail of individual appliance's energy use had on the final result can unfortunately not be extracted from the results as no treatment was present in either study for the absence of this form of feedback.

The main critique against appliance level feedback is the cost and effort it takes to provide it (Darby, 2006). Active research of appliance recognition from one single location within apartments, which saves the amount of hardware sensors needed, might change this assumption. Here, aggregated appliance load profiles are measured, analyzed and traced back to the most probable source. By using modern statistical and data mining algorithms researchers have been able to accurately determine around 85% of the appliances (Pihala, 1998; Ruzzelli et al., 2010). This method minimizes hardware and installation costs by using one central spot for measuring and will be presented in more detail in Chapter 3 and evaluated in depth in Chapter 6.

2.1.3 User Level

The appliance-level measurements divide the household energy usage into its individual appliances. However, most appliances are an aggregation of several settings and operating choices for the user to make. Providing energy-use information on the level where users decide on specific settings and operating modes signifies the user information detail level. Previous researchers have found conclusive evidence that "an effective intervention must be customized to the population and situation being targeted", since generalized information has failed to evoke interest and make sense to energy users (McMakin et al., 2002, p.860). In a review of energy information experiments Fischer (2008) also concludes that solutions have to be flexible to cater to different target groups.

In order to move the research on energy information feedback forward, "... the experimental design must focus on individual households and appliance-specific and time-of-day usage" (Neenan, 2009, p.5). Generally, energy-use choices can be made at two different levels. First, a user could replace the appliance under consideration with a newer, more energy efficient one. As primarily older large electrical appliances are found in households, savings potential could exist in exchanging these for more efficient appliances. Household appliances continually improve their energy efficiency due to ongoing development, legal requirements and mandatory labeling (Taylor, 2007). Second, individuals can utilize user-level information to evaluate and change their energy usage with respect to specific devices. For example, instead of washing clothes at 60 °C, a consumer could utilize a 40 °C washing cycle. In this case, the behavioral change would have little impact on the wash result (Wagner, 2010), but would have a considerable impact on the energy usage over time.

Since disaggregating appliances from household measurements is similar to disaggregating operating modes from appliance measurements, the method used to gather the data will also be described in Chapter 6, in conjunction with the appliance disaggregation. An information system that integrates feedback on an appliance and user level will then be explored in Chapter 7.

2.2 Types of Feedback

Irrespective of the level of energy-use information detail (Section 2.1), there are a number of ways to process it. Depending on the feedback type given energy savings are referenced in the range between 4 and 15 percent (Neenan, 2009; Ehrhardt-Martinez et al., 2010; Darby, 2006). One main reason for this spread is the immediacy of the supplied information, shown in Figure 2.2, where frequencies at or above daily provide, on average, the highest savings. It should be noted that these experimental result's variance is large and the statistical strength is often weak, due to a low number of participants and few measurement points (Neenan, 2009; Ehrhardt-Martinez et al., 2010). Previous experiments of energy use feedback are reviewed in more depth in Chapter 5. The result from this review emphasizes this conclusion by showing that the impact of feedback has a large spread even when the studies have evaluated the same feedback types.

The fundamental feedback types that are related to research is indirect and direct feedback. The indirect feedback is given by stored information to provide an

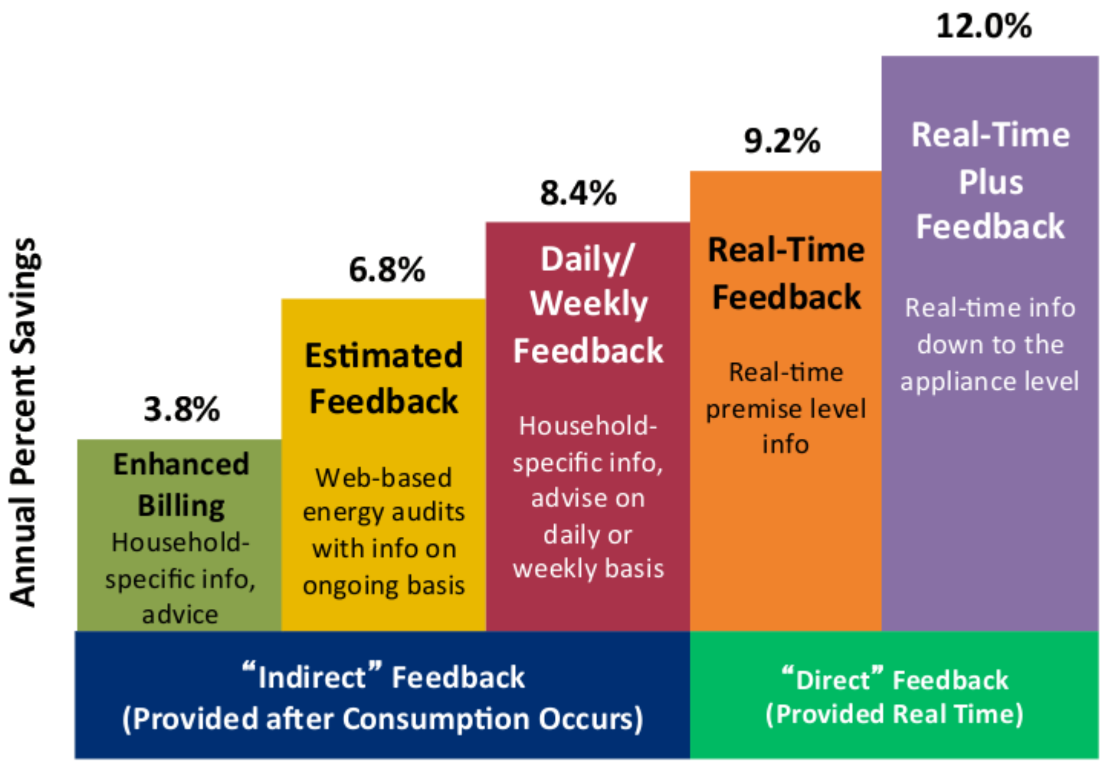

Figure 2.2: Feedback types and definitions, ordered by their comparable availability and cost to implement (Ehrhardt-Martinez et al., 2010)

overview over passed events. Direct feedback, in contrast, signifies the information that is directly linked to the energy use in real-time. While other feedback types have been evaluated, for example, how the information can be gamified (Brewer et al., 2011), presented in a goal oriented way (McCalley and Midden, 2002) or in a social context (Yim, 2011), they are all additional design considerations on top of the feedback types explored in this study. Feedback design options can help contextualize the information by changing the informational units from energy [kWh] to price (Ueno et al., 2006), by providing a normative comparison (Petersen et al., 2007), or a motivating goal (Van Houwelingen and Van Raaij, 1989). However, the complex interaction between feedback formats warrants separate discourses (Abrahamse et al., 2005), and this study will explore indirect and direct energy-use information.

2.2.1 Indirect

Storing energy-use data provides an essential part in the possibility to compare and learn from past energy usage levels and patterns, for example, a joint Norwegian and Finnish experiment extended the billing information to include graphical comparisons of energy consumption between the past and current year. The reception of this feedback was positive and led to an average saving in energy of 5-10% that lasted over the three year experiment period (Wilhite et al., 1999). Similar accounts can also be found from the late 70s and early 80s (Hayes and Cone, 1977, 1981). In these cases, energy usage is compared to previous weeks and months, and was found to have a significant positive effect on energy conservation.

Indirect forms of feedback tend to be better suited to help households understand the effects of changes in space heating, household energy use composition, and the effect of investments in new appliances and building shell upgrades (Darby, 2006). However, learning from past experiences has not only been shown to lead to lower energy use, but it has also been found to sediment a sense of limited influence, as the users find it increasingly challenging to find new ways to make their energy use more efficient (Froehlich, 2009; Hargreaves et al., 2010).

2.2.2 Direct

Continuously updated information on energy usage signifies direct feedback. This type of feedback is also sometimes called live or real-time feedback. In this study an information frequency below 15 minutes is chosen to correspond to direct feedback, as this is common in practice (ESMA, 2010; Seal and McGranaghan, 2010).

Direct forms of feedback (real-time and real-time plus in Figure 2.2) tend to be well suited for understanding energy savings associated with current uses on their own and in comparison to previous patterns of energy usage, which might be difficult to analyze indirectly (Darby, 2008). Direct feedback has been found to motivate a reduction in energy use (Grønhøj and Thøgersen, 2011), however, this effect is most often short lived and usually manifests itself in the beginning when users first gain access to their energy feedback (Hargreaves et al., 2013).

Previous experiments with direct energy feedback have shown a wide range of

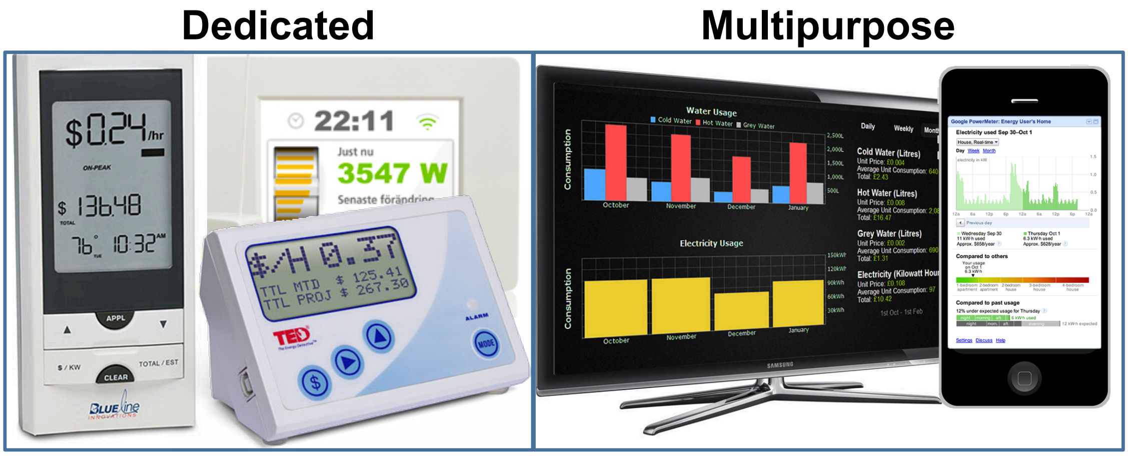

Figure 2.3: Examples of common types of energy-use display methods.

results, between non-significant (Allen and Janda, 2006; Thuvander et al., 2012) and 23% (Petersen et al., 2007). Due to the specific experimental designs and mixes of the feedback types given, it is difficult to pin-point which informational factor(s) is causing the effect (Fischer, 2008). This particular challenge will be further elaborated on in Chapter 3.

2.3 Display Methods

After processing the energy use data into indirect feedback, direct feedback or both it is transmitted to the household user. Generally display technologies can be broadly divided into dedicated and multipurpose interfaces. A dedicated interface is designed for a specific task and is not intended to be repurposed for other uses. The multipurpose interface is, however, designed to handle informational material from different sources. Figure 2.3 shows the "Blueline", "Eliq" and "The Energy Detective" as examples of common dedicated displays and the TV/PC display and the mobile phone are examples of multipurpose displays (Pierce et al., 2010; Thuvander et al., 2012; Parker et al., 2006; Raw and Ross, 2011; Houde et al., 2011).

This section will give some background of the available systems in both of these categories and point out some of their key strengths and weaknesses, which are fundamental for understanding the decisions made in this study.

2.3.1 Dedicated

In the energy-use feedback domain a dedicated display predominantly shows information that has to do with converting energy in some form. Additional services as telling the time or the weather is common, however, this information is secondary to the energy-use feedback. Most direct feedback experiments use dedicated displays that show current energy usage (Ehrhardt-Martinez et al., 2010).

Dedicated displays can either deliver the information based on text and symbols or provide it in an ambient way. Text, numbers and graphs have the ability to supply a specific and detailed message. Providing all available information is not always appropriate to residential households (Darby, 2008). Too much information detail can overload the individual. How to balance information load to support users to make accurate, timely and considerate decisions is a considerable challenge for designing interfaces. Chapter 8, therefore, examines this in more detail and evaluates how grouping information can assist the decision making process.

In contrast to displaying information textually or graphically, an ambient format has the benefit that users can be made aware of the feedback without having to focus their attention on an interface (Weiser and Brown, 1996). Ambient information is based on effects that can be processed and understood passively, as for example light, movement or sound. This design format has its roots in the calm technology field (Weiser and Brown, 1996).

Most use cases for ambient feedback are goal oriented, where users are guided by specific pre-determined targets. In a goal-oriented setting, the effect of ambient light information compared to factual numbers has indeed shown a stronger positive impact on users' behavior (Ham and Midden, 2010). In these circumstances it has been shown to support users in making more efficient decisions (Maan et al., 2010). However, although an ambient feedback form of energy-use information can be useful for comparing specific states, energy-use decisions also entail evaluating current appliances and behaviors on a more detailed level. In this case, where more options have to be parsed before taking an action, ambient feedback has only had a small impact (Faruqui and Sergici, 2009).

It should be noted that when users have grown accustomed to a common textual-based display, the given information has also begun to function in a similar way to ambient feedback devices. Even though the detailed textual information might become "backgrounded", the interface has been found to keep the awareness of the user passively (Hargreaves et al., 2013).

2.3.2 Multipurpose

Dedicated displays allow for direct access to a specific type of information but can be vulnerable to habituation due to their reliance on only one channel and one message (Hargreaves et al., 2013). The necessity for a specific display to access feedback can also narrow the targeted group significantly. Web-based, or other forms of information that can be received on a existing display platform, can be appropriate and an inexpensive way to reach a wide population who regularly use their computers to gather information. A pilot study of providing appliance information on a mobile phone was found to be useful by the focus group (Weiss, 2010). However, for populations where internet use is more sporadic, this form of feedback could involve a large change to the consumer's habits to be accessible. For these users, already known interfaces, as for example the TV set, have been suggested (Harrison, 2003). This issue is likely to diminish as more people become experienced with and have access to internet-based devices.

An example where direct feedback was delivered over a web interface is the study of Google employees' reaction to the *Google PowerMeter* interface (Houde et al., 2011). This study found that a web-based interface motivates energy conservation similar to dedicated displays. The overall conservation was calculated to be 5.7% among the participants. The persistence of these effects are, however, only significant during the first four weeks of the six months where feedback was given (Houde et al., 2011).

The potential scalability of using multipurpose displays make this format suitable for a wide household implementations. Chapter 4 will provide more details around the development of the web-based platform for the experimental study in this thesis. Chapter 5 will evaluate the usage and interview experiment participants about their interface preferences.

2.4 Energy User Characteristics

Recent advancements in information technology have made it possible to give individuals access to a similar information as was previously limited to experts and managers. This section will present the fundamental concept of the novice user, which this study defines as a residential users who previously have had little or no access to information about their energy-use related actions, and explain the implications of targeting this user segment.

2.4.1 Levels of Experience

It is well established that "information must be presented to consumers in formats that facilitate processing" (Bettman and Kakkar, 1977). However, in spite of great design efforts, novice end-users' interaction with informational displays has been shown to lead to sedimentation of existing habits instead of motivating a behavior change (Hargreaves et al., 2013). A more detailed understanding of how end-users particular circumstances affect the decision outcome is fundamental for advancing the current state of research (Shove, 2004). Fogg (2003) clarifies this statement by pointing out the different preferences users have for interaction when fueling their car, where they usually prefer a clear and minimally complex interface, compared to an entertainment situation like gaming or watching TV.

The context of the recipient of the energy-use information has shown to affect how it will be received (Abrahamse and Steg, 2009). Information systems geared towards experts can, for example, make presumptions about a certain domain knowledge or that they will train to understand the presented concepts (Green and Hughes, 1986). Furthermore, the incentives for experts and novices are different. The experts can be paid to learn and use information systems, where the consumer's interest and activity is predominantly self directed.

2.4.2 Information Processing Capacity

Beyond the differences in incentive structure between experts and novices, information design also has to take into consideration users constrained ability to process information. As a consequence of being exposed to new energy-use information, end users are expected to make more informed decisions. However, the challenge is increasingly becoming a question of the ability to "drink from the fire hose" (Waldrop, 1990), and find the *right* information to make efficient decisions. Even though information is accessible, an individual's rationality is bound by the limits of their information processing capacity (Simon, 1972). Furthermore, separating an individual's energy-use patterns from their everyday life to promote more rational decisions has proven to be very complicated (Röpke, 1999).

Feedback is necessary for understanding energy usage and for choosing an effective action to reduce it, but it is not always sufficient (Darby, 2008). The awareness and knowledgeable decision making do not necessarily increase with the amount and detail of information provided to the decision maker. An example from the energy field comes from Allen and Janda (2006) who found that adding features to an in-home display created confusion and could hinder the overall understanding.

These findings are well-known in the context of the cognitive load theory, which posits that human decision makers can easily be overloaded with information, yielding worse decisions (Jacoby, 1977). Information system research plays a central role in bridging this gap between the technological ability to increase information on the one hand, and balance the level of information detail for efficient decision making on the other hand (Watson et al., 2010). Although access to decision support tools was found to lower effort in processing a specific amount of information, it could not be shown that it provided the capability to process more information (Todd and Benbasat, 1992).

Previous research has found an inverse U-shaped relationship between task performance and information detail. Performance first improves with more information until a certain level when it starts to drop off as more information is introduced (Hwang and Lin, 1999). In particular minimizing the complexity, for example by combining and summarizing information, is recommended to make the decision process clearer for novice users (Hwang and Lin, 1999).

Chapter 8 will explore the possibility to categories information to facilitate decisions. A web-based interface is designed in different treatments to understand the general effects of information detail.

Chapter 3

Residential Energy-Use Information Engineering

Based on the characteristics of energy-use information, the analysis and development follow specific methods and have a number of constraints and challenges in common. This chapter will introduce the main engineering considerations, that are necessary to implement and evaluate energy-use feedback solutions in an iterative fashion.

3.1 Data Collection Methods

Reporting values with at least a 15-minute resolution, which the new Advanced Meter Infrastructure (AMI) proposes, is a great improvement from the yearly billing cycle that has been predominant for households in the energy domain. However, since standardization frameworks take consideration and time to be developed and adopted there are still disparity in functionality between different governments and utility offerings (European Commission, 2011). It is important to recognize that the traditional focus of metering only for billing might limit the design scope and stand in the way of future energy information analysis and feedback services. Commercial alternatives to gather energy-use feedback for end users also differ in their capacity. Everything from single appliance measurement devices like the "Kill-a-watt"[1] to sophisticated home sensor and automation systems are available.

[1] http://www.p3international.com/products/p4400.html

This section will present the current methods of collecting energy use information. Centralized household-level measurements and the main aspects of provisioning appliance- and user-level information through disaggregation and decentralized measurements related to this study will be explained.

3.1.1 Centralized Sensing

Sensing electrical current and voltage, in order to calculate power, can be done in a number of ways. A brief explanation and comparison of technologies ranging from looped current transformers and hall sensors to magnetoresistive material (that change properties depending on the surrounding electric field strength) is compiled by Asada et al. (2003).

Although contact-less magnetoresistive current sensors like the one employed by Patel et al. (2010) have obvious benefits by allowing less interaction with high current carrying power lines and might increase in popularity, the most prevalent method today to measure alternating current is with a split-core current transformer due to their simple and robust construction. The current-transformer sensor is used in commercial offers like *The Energy Detective*[2] and *Current Cost*[3] and open source projects like the *Open Energy Monitor*[4] and *Flukso*[5].

An alternative to gathering energy-use data from the magnetic field would be to attach a sub-meter to the already existing analog or digital household meter. An optical sensor could, for example, track a visible mark or notification light on the existing meter, which is also pursued commercially[6]. The sub-meter measurements are very accurate and perfectly fit the billed energy usage. However, this approach is limited to the information calculated and provided by the meter designers. No raw data on the electrical current and voltage is available. Subsequently, other dimensions of power (i.e. its reactive or harmonic components) are not available. Furthermore, in times of low energy use, time lags between measurements and the optical monitored notifications can become large.

[2] http://www.theenergydetective.com/
[3] http://www.currentcost.com/
[4] http://openenergymonitor.org/
[5] http://www.flukso.net/
[6] http://www.wattcher.nl/

3.1.2 Disaggregation

Centralized sensing by default measures the aggregated household's energy usage. In order to extract energy-use information on individual appliances this aggregated signal can be disaggegated with a number of methods, which will be introduced in this section.

Methods for disaggregating total load and energy usage information on specific appliances is increasing in popularity and has also been the focus of full doctoral dissertations, for example Pihala (1998) and Parson (2014). Belkin's recent programming competition, where any interested programmer could submit strategies to distinguish individual appliances from a household-level meter reading and win $25,000 show that the interest in disaggregating household energy-use information goes beyond a purely academic one. These systems parse the household-level energy-use data and use algorithms to recognize individual appliances.

George Hart et al. designed the first system for disaggregating appliances with minimal hardware in the mid 80s. By attaching a monitor to the existing mechanical meter they managed to describe appliance signatures in the two dimensions of real and reactive power. By using these two power entities, steady-state appliances with stable levels of power load could be found. As this method only needed one location for measurement and used software for recognizing specific appliances it could be implemented with a minimal set of resources (Hart, 1992).

Hart (1992) focused ON/OFF switches of appliances was to track their state change. This worked well for steady state appliances with a couple of possible states, like toasters and razors. However, for multiple or variable state devices more advanced methods of measurement, clustering and filtering are necessary (Laughman et al., 2003).

Machine-learning methods like artificial neural networks (ANN) and Bayesian filters have also been used to successfully distinguish a limited number of appliances. In this case features of appliances are learned by machines in a supervised manner with the Bayesian method or potentially unsupervised with ANN. The power levels are then compared to probabilities of certain appliance combinations. The single or combination of appliances most closely related to the training data will then be offered as feedback. The limitation of this approach is, however, that as more appliances and their different states are added to the probability calculations the computational load gets overwhelming. Fifteen different appliances are reported as the feasible limit with ANN (Ruzzelli et al., 2010), and - while not clearly stated - the Bayesian filter trial was limited to seventeen different combinations of nine appliances (Lin et al., 2009).

Another development in disaggregation techniques uses a high bandwidth current and voltage sensor at a building's central fuse box to measure transients and harmonic distortion in high resolution. Then, by using a digital signal processor, the range of appliances that can reliably be recognized was extended (Norford and Leeb, 1996). For example, appliances with a highly variable load such as, motors or compressors, emit transients at startup and can be monitored by combining the transients with the information of the complex power load (Norford and Leeb, 1996).

Extremely high frequency transients have also been successfully explored by Patel et al. This team of researchers used transients and noise are also potential dimensions that can be used to distinguish between appliances. By monitoring frequencies between 60 Hz and 100 MHz they managed to locate with high accuracy several household appliances, both larger and smaller ones (e.g. dishwashers and laptop adapters) respectively (Patel et al., 2007). This method has the ability to measure from electrical outlets, as it is based on the transients and noise, which travel throughout the interconnected electrical network on each phase. This system is affected by the location of the sensor in relation to what is sensed. This means that the system would have to be recalibrated if it or the measured appliances changed location. There is also no information about the current or power draw in transients and noise, so to gather information about the energy usage of individual appliances another current sensing monitor must be added to the setup (Froehlich et al., 2011).

Even though high frequency measurements provide many characteristics that lend themselves well to appliance recognition and differentiation, the required hardware for these laboratory measurements is still prohibitively costly to be employed on a large household scale (Zeifman and Roth, 2011). High frequency signal processing is by its very nature computationally expensive. To use this method efficiently this thesis explores how the frequency is related to the disaggregation accuracy of appliances is evaluated in Chapter 6.

3.1.3 Decentralized Sensing

Another alternative to provide appliance-level information is to measure directly at the appliance or at every outlet. The hardware for these measurements can be made less sophisticated in comparison to its centralized, high frequency sensor, counterpart. This is because little or no signal processing is needed to measure specific appliances. However, since more sensors are needed this system is often regarded as more costly (Laughman et al., 2003).

In order to design experiments to evaluate operation behavior on energy usage, the measurement and feedback must support a user level of information detail. Appliance disaggregation research has provided stable results in laboratory environments and in specific buildings. However, the information from central appliance disaggregation is still not reliable for in depth analyzes of how an appliance is being operated (Zeifman and Roth, 2011). For this reason, plug-level, decentralized sensors in every electrical outlet or in every appliance has been envisioned (Zou et al., 2011).

Measuring the energy usage at every electrical outlet provides highly robust and clear data in terms of the appliances attached to this outlet. It also opens up for other services like remote control and automation of individual appliances, which could not be supported by only a central installation. The method to measure single appliances follow the same procedure as the initial phase of centralized appliance disaggregation, and will be further explained in Chapter 6.

3.2 Challenges and Constraints

Developing the access to energy-use information to improve the general knowledge and lay the foundation for future services on a residential scale is a challenging task. First of all, the functionality have to provide enough information to support services that can help energy users become more active and sustainable in their energy usage. Second, the scale of targeting residential households limits what can be installed cost effectively. This section will introduce the theoretical and practical challenges and constraints to implementing energy-use feedback in the residential sector.

3.2.1 Theoretical

How energy-use information is used in order to accomplish a specific goal has received little research attention (Bartlett and Toms, 2005). This is a complex problem, which was already evident in the eighties as this quote explains: “In general, we know that increases in fuel prices, limited availability of some fuels, slower growth in economic activity and in population, and perhaps government conservation programs all contributed to changes in energy use. However, we are unable to accurately quantify the influence of each determinant on energy use changes” (Hirst, 1980, p.868).

Distinguishing what contributions the different forms of feedback have in households and what factors motivate desirable change are still unclear. User interaction is treated as a black box where experiments are only analyzed at the end of the interaction regarding its overall success “without providing insight into the reasons why” (Abrahamse et al., 2005, p.283).

Experimental results from giving direct energy-use feedback, thus, have contradictory results, which cannot be explained by current analyses. By measuring the times of information interaction, Nunes et al. (2011) approached the relationship between information interaction and energy usage. The study found that active participation is essential for impact, with a significant correlation between the frequencies with which real-time energy monitors were used, and a diminished use of energy. The authors concluded, however, that more research is needed to fully understand how the information was used to result in these changes.

A couple of studies that provide quantitative insight into the effect of information use come from Hargreaves et al. (2010, 2013). Through extensive interviews, the authors found that some actions were taken in direct response to the information shown on the display while others seemed to be based on an accumulated amount of information (Hargreaves et al., 2013). However, whether or not this information affected the energy usage is not clear, as only an initial behavior change could be measured.

Our understanding of how information is used is the bottleneck, and this understanding directly impacts how effective energy interaction systems should be developed (Watson et al., 2010). In order to proceed from the current state of inconclusive evidence with regard to information access and its impact, experiments need to be developed that follow user interaction more closely (Darby, 2010c). This thesis answers this call and presents the necessary monitoring system for continuously measuring interface interaction and energy usage in Chapter 4. Then, in Chapter 5, the system is used in the field and the correlation between interface interaction and energy usage is evaluated.

3.2.2 Practical

The fundamental practical constraint for smart-meter technology intended for households is the system cost for a regional or national roll out in relation to the utility it produces. The large scale of the system is necessary since the residential sector is comprised of many small entities, in contrast to the industrial sector where fewer large users have a comparable energy usage. The system has to balance information detail and the cost to provide the hardware for sensing, processing, storing and presenting the information (Zoha et al., 2012). Specific energy-use data on an individual appliance level, where every single device is recognized, is believed to be intuitive and able to increase energy-use awareness (Darby, 2010a). However, appliance specific information often incur a premium cost (Zoha et al., 2012).

At the end-user level the cost constraints are more severe compared to solutions for larger businesses with customized appliances. So far the evaluations of central appliance recognition that have been carried out in laboratory settings or field trials have been very successful at finding appliances (Norford et al., 1999). This scope has suited the purposes of keeping track of specific appliances with a large power load, like air-conditioning units (Hadden, 1999). However, as the scope of these appliance recognition tools are applied at the end-user level, as a potential part of the smart grid, new considerations must be taken into account. For example, differences between building types, their energy distribution system and the interaction with already installed appliances must be considered. Common building dynamics have to be understood and evaluated to get a robust information system.

So far it has been considered too expensive to scale this system to a smart grid level because of the extensive hardware needs and intrusive installation methods normally necessary (Darby, 2006). However, a better understanding of how data sample resolution impacts the results of appliance disaggregation is needed, as it directly influences the necessary hardware sophistication level. Two systems to provide appliance level data will, therefore, be evaluated in Chapter 6. The already known disaggregation techniques will be evaluated in terms of their dependency on frequent measurements. Additionally a new electrical outlet level system will be tested to provide, even more detailed energy-use information, on an user operation level.

3.3 Methods of Analysis

The research conducted in this thesis draw from Information Systems (IS) research methodologies and specifically Green IS, Energy Informatics and Design Science. This section will introduce these domains, the main framework they provide and how this affects the research development.

3.3.1 Green IS and Energy Informatics

Information-systems research has an important role to fill in the development of decision support systems for households by providing rigorous methods with which iterative research can be developed (Brocke et al., 2013). By fusing IS research with Green IT, Green IS and Energy Informatics are posed to help develop more sustainable business practices and products through focused reporting and analysis of their environmental impacts (Boudreau et al., 2008).

Historically, as the information technology (IT) industry expanded during a time of heightened sensitivity of global environmental issues, the impact of IT has been scrutinized in terms of their environmental impact. Green IT research is analyzing ways of improving the use, design, manufacturing and disposal of information technology to make its environmental impact smaller (Murugesan, 2008). The use of IT to improve sustainability and awareness has grown out of this research. This latter focus, which is directly related to the system proposed in this study, is receiving more attention under the banners of Green IS (Butler, 2011; Elliot, 2011) and Energy Informatics (Watson et al., 2010; Goebel et al., 2013).

Both Green IS and Energy Informatics emphasize the importance and challenge of focusing on the system instead of the technology and the complexities in the development of, for example, an energy-use information system. However, the scope is different between these two research directions. Green IS focus on the development of information systems to enable sustainable business processes while Energy Informatics analyse how information systems can be designed to increase the efficiency of energy usage (Watson and Boudreau, 2011).

3.3.2 Design Science

The situation of integrating information and communications technology (ICT) in the development of new artifacts is not limited to the IS domain. However, IS research has of late become recognized for its unique position to investigate and improve an ICT artifact's ability to solve "real-world business problems" (Hevner and Chatterjee, 2010, p.9). By combining insights from natural science, how individuals, organizations and technology interaction can be understood with the problem solving aspect of engineering, an IT artifact can be developed to improve the effectiveness and efficiency of an organization (Hevner et al., 2004).

A clear and structured framework is necessary to make the design process a stepping-stone for further research. By adhering to the proposed analytical framework, as proposed by design science (Hevner et al., 2004), the potential for cumulative development of the artifact is strengthened (Gregor and Jones, 2007).

The design science structure was originally ordered along the nominal process shown in Figure 3.1 (Hevner et al., 2004). However, in subsequent developments of the theory the possible entry point was recognized to be dependent on the specific problem, its context and who initiated the research (Peffers et al., 2007).

The four different entry points can be centered on: the problem, the objective, the design and development and the client. A problem centered approach is initiated by

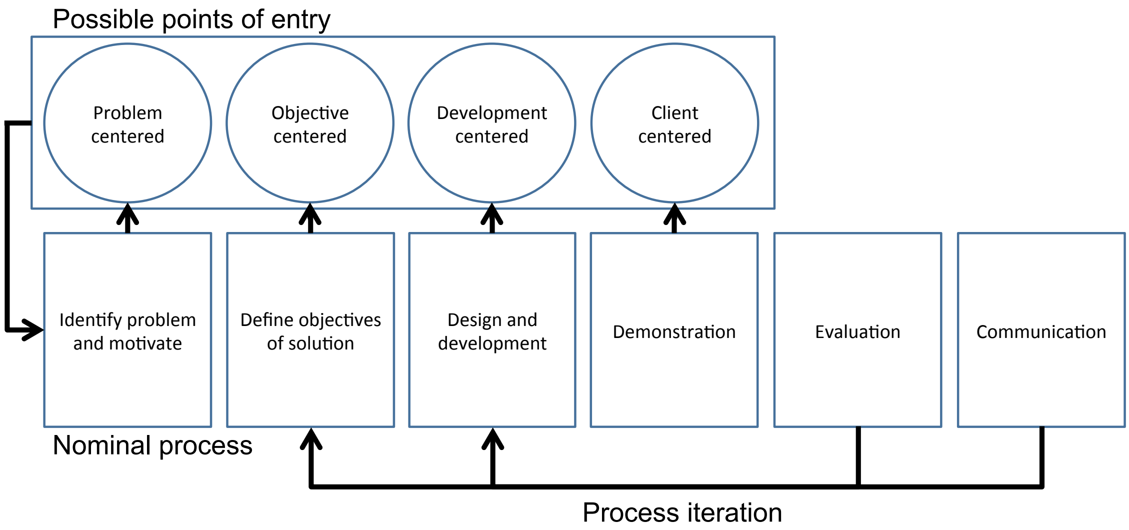

Figure 3.1: Design science research model with the different entry points and their relation to the nominal process, based on Peffers et al. (2007, p.54)

clearly defining the problem and motivating the need for a solution. An objective centered solution is based on a given problem, however, the solution has not been formalized. By clearly defining the solution scope the design science research can be evaluated. A design and development focus is when the solution is given. This could be when developments in engineering provide new technologies that can be used in IS solutions. Finally, a client-centered design iteration is based on current design implementation to which users react. Based on the input from these users, new design objectives and requirements can be developed.

Since the smart-meter infrastructure already exists it is appropriate to first evaluate users experiences with related versions before moving on to the next design iteration. In line with this analytical process, the development and evaluation of the system related to the current smart-meter specification is presented in Chapter 4 and 5, while the client initiated system, which resulted from this research, is developed and evaluated in Chapter 6 through 8.

Part II

Household-Level Energy-Use Information

Chapter 4

Household Energy-Meter Engineering

4.1 Problem Definition

The experimental evidence is inconclusive regarding the impact of energy-use information. In Chapter 5, Section 5.2.1 experimental projects, which provide direct access to the household's aggregated energy use, are compiled and show results that vary between non-significant to 23% energy savings (Ehrhardt-Martinez et al., 2010; Darby et al., 2011; Thuvander et al., 2012). Due to the many variables in energy-use information experiments it is difficult to deduce if and what treatment variable has influenced the participants to change or not (Abrahamse et al., 2005). Researchers of information systems (IS) science are therefore compelled to move their focus from the question *if* users are affected, to the question *how* they are affected (Watson et al., 2010; Hargreaves et al., 2013).

As explained in Chapter 3, Section 3.2, a major challenge is to control the experimental variables. This problem is also present on the measurement level. Full access to the energy-use measurements are necessary in order to follow the participants energy usage and information interaction. The experimental system should also have similar capabilities as the system from the utility or other meter provider. Since the currently quoted capability of smart-meter technology has a data polling frequency of between 7 seconds and 2 minutes (Ehrhardt-Martinez et al., 2010), the developed energy-use information sensor and feedback system has to provide access to the raw measurement data in a comparable frequency. Finally, the experimental information system also has to be portable and installable by non-intrusive means for practical and cost reasons of a temporary field trial.

In light of these requirements, this chapter aims to analyze the development of such an experimental platform and evaluate how aspects outside the control of the energy meter can influence the measurements. The engineering development to measure, process, store and present household-level energy-use and interaction information will be detailed. The measurement platform will be evaluated between buildings and analyze the influence of a building's specific wiring on the measurements. The results of this study has been previously published by Dalen and Weinhardt (2012). The resulting platform for processing and presenting the measured information was subsequently published by Dalén et al. (2012).

4.2 System Design

In line with the practical requirement of system scalability, which was introduced in Chapter 3, Section 3.2.2, the goal with our platform was to strike a balance between measurement frequency, information detail and creating a platform that could be implemented on any smart-meter even with a rudimentary microcontroller. In this first household-level design iteration we will focus on measurements at 1 Hz and slower, which is comparable to the current smart-meter capabilities (Ehrhardt-Martinez et al., 2010).

Being a temporary experimental platform, intrusive or labor intensive measurement installations, which are available to commercial meter providers, will not be evaluated in this study. The main difference between the commercial solution and their open-source counter parts is the proprietary nature of the former. This in turn limits the access to the underlying data, as it is normally given in an aggregated format. As this project is focused on testing how different aggregations and formats of energy information will influence the receiver, access to the raw data is necessary. The resulting system was, therefore, based on the Arduino platform and the Atmel Atmega 328 microcontroller[1]. The well-documented open-source nature

[1] http://arduino.cc/

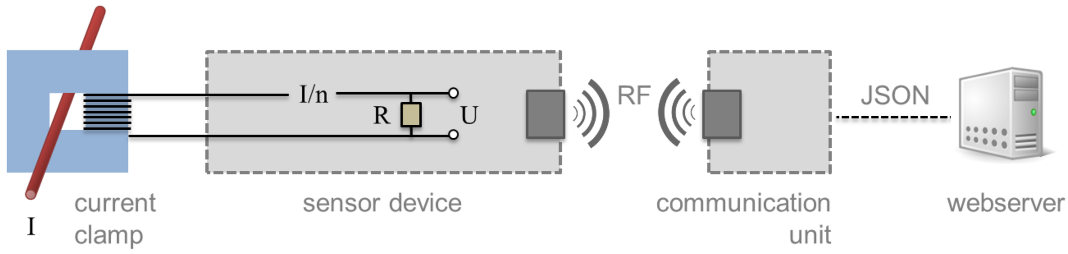

Figure 4.1: Measurement System Overview: power line (single phase), current clamp sensor, sensor device, communication unit, web server

of this specific hardware allowed for explicit control of every aspect of the sensing and processing the data before transmitting it to the main server.

The system was built using a battery-driven sensor unit with three current transformers (CTs). The sensor unit calculates power values from the current measurements and sends these values to a communication device via a radio link. The values where then transmitted to the central database, which supplied the web-interface with its information. The system is depicted in Figure 4.1 and shows the sensor unit with its current transformer, the radio transmission link and the communication unit that transmits the values to the database. This section will provide more detail on the physical and informational properties of the designed system.

4.2.1 Physical Properties

The measurement technique used in this study utilizes the magnetic field that is established around a wire when there is a change in the electric current. Since the magnetic field is proportional to the current in a long straight wire (Ampere's law) it can be used as a proxy for the flow of electrons and subsequently provide a less intrusive measurement than the alternative, in-line measurement option.

The German electrical power grid provides alternating current on 3 phases, at 50 Hz and with a voltage of 230V. At the fusebox each phase corresponds to one of the main wires. The ferrite split-core CTs, depicted in Figure 4.2, with $n = 2000$ turns per meter and a maximum current restriction of $I_{max} = 100A$ were attached

Figure 4.2: Ferrite split-core current transformer (CT) with $n = 2000$ turns and a maximum current restriction of $I_{max} = 100A$.

to the main wires. The magnetic field the main wires is then channeled in an iron ring, around which another wire is coiled. The magnetic field induces an electric current in the coiled wire. This secondary current is lead through a resistor, over which the voltage drop is measured. By using Ohm's law, which states that the current is the quotient of the voltage divided by the resistance ($I = U/R$), the secondary current (I_2) can be calculated. Supposing that the current transformer can be approximated as an ideal transformer ($I_1/I_2 = n_2/n_1$), the primary current (I_1) can be deduced since the turn densities (n) of the CTs are known. Finally, the power usage is calculated by multiplying the primary current with the known primary voltage ($P = I \cdot U$).

The goal with the measurement system was to strike a balance between the information detail (measurement frequency) and creating a platform that could be implemented on any smart meter. By using widely available technology with limited data processing capability the chances of wider use or even embedding the technology into smart meters is assumed to be higher. The open source platform Arduino is a general prototype platform with a simple interface which uses a derivative of C++ for coding. The same platform has successfully been used in other similar projects referenced in Chapter 3, Section 3.1.

The Atmega 328 micro-controller is fed by a voltage of $3.3V$, which restricts the incoming signal to the on-chip analog-to-digital converter (ADC) to a voltage up to this value. This means that the incoming current from the CTs must be limited, in order to be correctly registered and not to put the integrated circuit at risk of

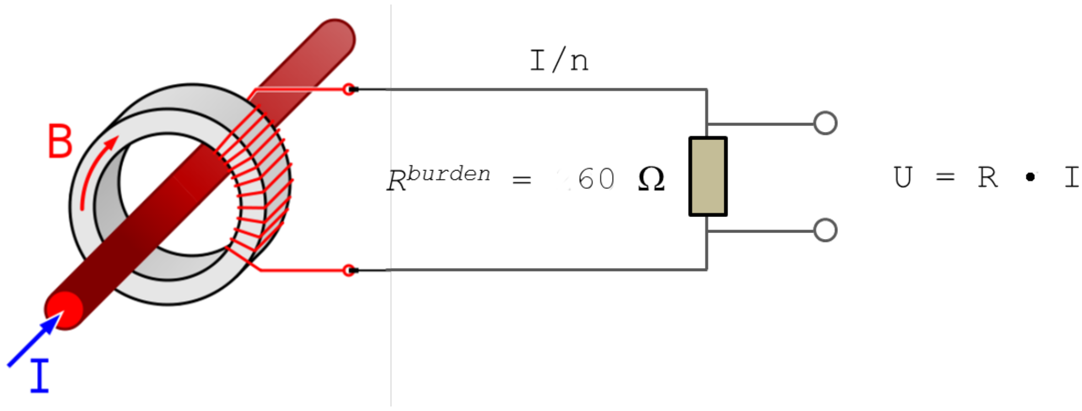

Figure 4.3: Physical Working Principle of power measurement by induction.

over voltage. However, by allowing large currents to correspond to the upper limit of the ADC the resolution for detecting a change in current suffers. The upper detection limit must therefore be designed with a minimal buffer. Since most of the fuses in German household's have a trigger threshold of 50 A or less, we chose the on-board burden resistor accordingly, shown in Figure 4.3. By using Ohm's law we deduce the burden resistance to be 66Ω, and with a 60Ω-resistor we leave room for a $5A$ overshoot without sacrificing too much resolution.

The micro controller processes the voltage information from the CTs and the burden resistor on a 10-bit scale (2^{10}), which results in a resolution of 8.7 VA/bit. The alternating current and voltage is sampled at 2000 samples per second and the real power is attained by multiplying the instantaneous current and voltage measurements according to Equation 4.1 and 4.2. Apparent power is calculated with the root-mean-square (RMS) value of the sampled values (see Equation 4.3 and 4.4). The apparent power-vector is normally expressed in its base units of real and reactive power, as pictured in Figure 4.4, the imaginary power unit (reactive power), was subsequently calculated according to the Pythagorean theorem.

$$u_i \cdot i_i = p_i \tag{4.1}$$

$$\sum_{i=1}^{n} \frac{p_i}{n} = P_{real} \tag{4.2}$$

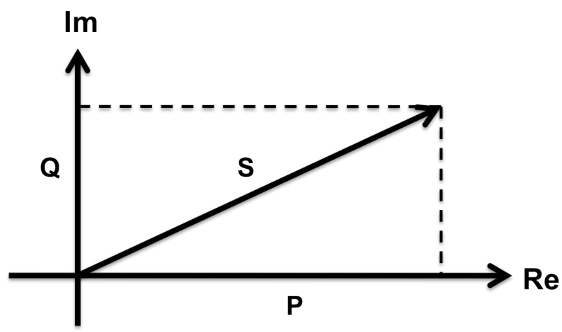

Figure 4.4: Apparent power (S) as the vector sum of the real (P) and reactive power (Q)

$$\begin{aligned} \sqrt{\frac{\sum_{i=1}^{n} u_i^2}{n}} &= U_{rms} \\ \sqrt{\frac{\sum_{i=1}^{n} i_i^2}{n}} &= I_{rms} \end{aligned} \tag{4.3}$$

$$U_{rms} \cdot I_{rms} = P_{rms} = S \tag{4.4}$$

4.2.2 Software Implementation

Two micro-controller boards were used to measure and send the energy-use data to the web-server. The measurement device is shown in its white casing in the system overview, (Figure 4.5). In an attempt to balance battery power with close to real-time values, energy-use readings from the metering devices were communicated once every eight seconds, since the radio communication used the most power by a large margin. This data was then aggregated into 5, 15 and 30-minute intervals to be displayed in the different historic view modes. The energy sensor and the communication unit communicate via radio at 868 MHz frequency. The second micro controller, shown in its black case in Figure 4.5, is connected to the router of the household via ethernet and transmits the data objects to the central server using JSON elements exemplified here below.

The presented system will be evaluated in two ways in this thesis. The first

```
{
  ''id'':''sensorID'',
  ''ct1'':''powerReading1'',
  ''ct2'':''powerReading2'',
  ''ct3'':''powerReading3'',
  ''batteryStatus'':''batteryReading''
}
```

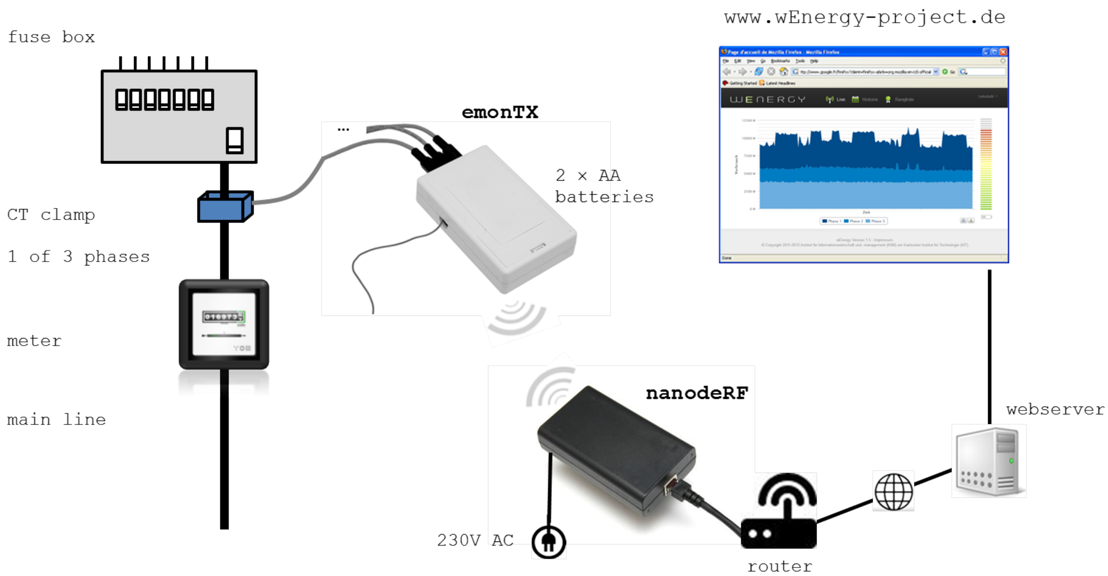

Figure 4.5: Household measurement, processing and presentation energy-use information system overview.

evaluation, which is presented in this chapter, focuses on the sensor hardware. Specifically, the energy sensor will be tested in different households in order to evaluate how it integrates with the household and explore the influence of household wiring on the measured data. The evaluation contribute to the understanding of how robust energy-use measurements are, and provide insight into the challenges of creating a system to recognize individual appliances between buildings. The second evaluation focus on the user interaction and is presented in Chapter 5. The main contribution from this evaluation is the increased understanding of how users interact with energy-use feedback that is given through a web-based interface.

4.3 Experimental Setup

To understand the transferability of the measurement system between buildings, how the individual household's influence the measurements has to be understood. This section will provide the details of an experiment that analyzes a building's influence on current and voltage measurements from a central point in the building. A specific method is devised to isolate the external affects of the building's electrical system characteristics. By using a fixed set of appliances between the houses most potential measurement artifacts can be attributed to the building.

4.3.1 Measurement Schedule

To isolate the building's electricity network the same seven appliances with a range of different power draws and different ratios between real and reactive power were measured in four buildings with different characteristics. The appliances where grouped into three groups and a specific running schedule was devised for them and cycled three times (as pictured in Table 4.1). The first group consisted of three kitchen appliances, a microwave oven, a hand mixer and an electric kettle. Half a liter of water was first warmed in the microwave for one minute, then mixed for 20 seconds, and finally brought to a boil in the electric kettle. The second group consisted of a fluorescent and an energy-saving light bulb. They where switched on for 30 seconds in consecutive order. Finally, the third group of appliances consisted of an iron and a table-fan. The iron was plugged in and left until warm and after unplugging the iron the fan was left running for 30 seconds.

The buildings studied were three apartment buildings and one office. Base-case measurements were measured directly at the appliances to have a common point of reference. Both the base-case and the apartment building one (B. One) have no other active appliances attached to the same electrical phase between the point of measurement and the measured appliance. Apartment building two and three (B. Two and B. Three) are both smaller apartments with similar attached loads (computers and fans). Finally, building four (B. Four) is the office building. The loads there are homogenous and consist mostly of computers, screens, fluorescent lights and other office equipment.

Table 4.1: Measurement schedule for energy-use sensor evaluation

Appliance Group	Type of Appliance	Running Time	
Kitchen	Microwave oven	1 min	x3
	Hand mixer	20 s	
	Electric kettle	Until boiling	
Lights	Fluorescent	30 s	x3
	Energy saver	30 s	
Other	Iron	Until warm	x3
	Table fan	30 s	

4.3.2 Parameters for Evaluation

The correlation between the appliances in different buildings can be understood by comparing the mean and variance of the individual device's power load given in Table 4.2. For example, the mean power load from an appliance which has been turned on can be compared between the appliances, while the variance is used to estimate whether the appliance measurements fluctuate more or less depending on the building's wiring and attached appliances on that phase. This could impact the load characteristics if the wires have are inductive or conductive loads relative to the measured appliance.

4.4 Results

The measurement micro-controller has been installed in four different buildings to test the sensors and evaluate the impact of a building's electrical infrastructure and secondary appliances' influence on the measurements. To evaluate the building's impact on the measurements the specific appliances are divided into the ones with a predominantly real power load and appliances with a significant reactive component.

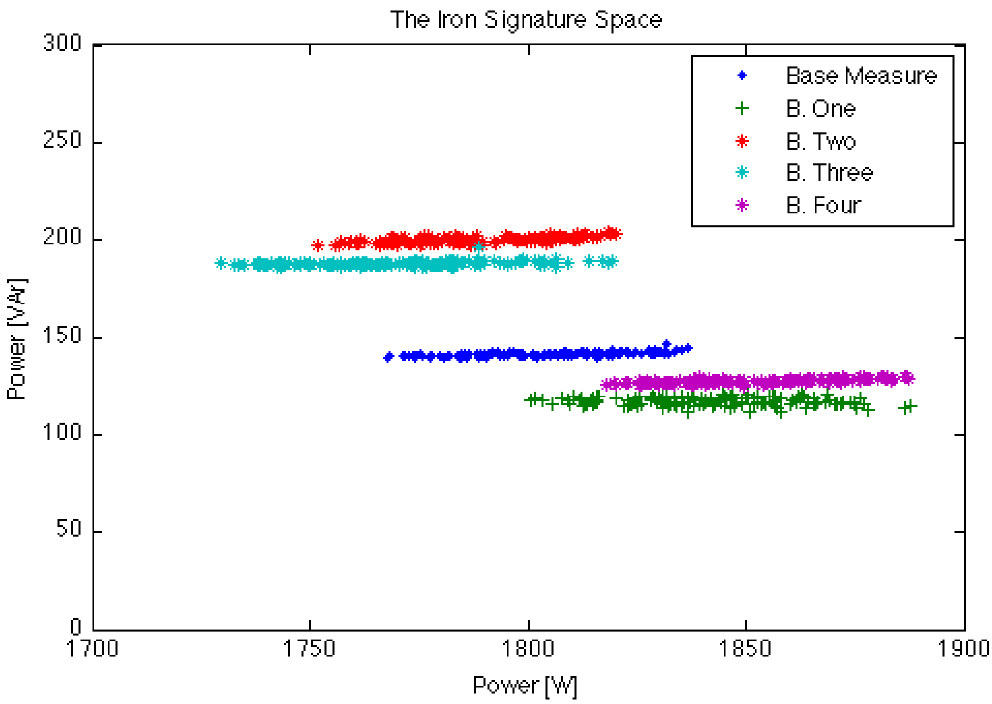

Figure 4.6: Real and reactive power measurements from the iron

4.4.1 Real Power Appliances

Appliances with a high power factor, (i.e. where the real power is dominant) give highly regular load clusters with small variations between the different buildings in relation to the overall load. A really clear example of this is the iron (shown in Figure 4.6), which only shows a slight variation in reactive power between buildings. The load is almost purely resistive and the different building's wiring and other active appliances only add a small reactive component.

4.4.2 Reactive Power Appliances

In contrast to the real power appliances, the influence of the appliances' amount of reactive power load is clear even from this small sample of measured buildings. Table 4.2 shows an excerpt of the mean and variance statistics for the iron, fan and fluorescent light measurements from the different buildings. Large differences from the mean in the reactive power between the appliances shows that the buildings

Table 4.2: Mean and variance of real (P) and reactive (Q) power statistics of the fluorescent light, the iron and the fan measurements while running.

Appliance	Statistic	Power	Base	B. One	B. Two	B. Three	B. Four
Flourescent light	Mean	P	6.99	9.12	4.20	3.09	3.78
		Q	9.81	11.47	10.58	10.42	31.59
	Var.	P	0.10	0.15	5.83	3.92	4.08
		Q	0.03	0.99	0.49	0.80	0.31
Iron	Mean	P	1801	1843	1786	1849	1760
		Q	141	117	200	127	188
	Var.	P	322	400	299	338	449
		Q	1	5	2	1	2
Fan	Mean	P	20.73	23.63	19.31	18.43	18.07
		Q	25.33	28.01	29.55	30.90	13.19
	Var.	P	0.42	0.56	4.57	5.60	6.32
		Q	0.84	0.75	0.27	0.35	0.23

influence the power load. The larger variance difference between devices shows that there is an interaction with charging and discharging components (i.e. capacitors or inductors on the same phase).

Appliances with a power load dominated by reactive power are influenced more by the appliances on the same phase. By interacting with other loads these predominantly reactive power appliances are characterized by a highly fluctuating real power. These effects can clearly be seen in the fluorescent light and the fan measurements visualized in Figures 4.7 and 4.8. Since these two different loads are inductive, the available capacitive power from surrounding appliances changes the amount of real power that has to be supplied to compensate for the reactive load. This effect can be likened to the storage of capacitors that larger industries install to balance the inductive loads of large or multiple motors. However, where motor windings and capacitor banks can be scaled to balance each other, the effects of fast switching inductive loads like the fluorescent light are more variable. This effect can clearly be seen in the office (B. Four) where the fan, which is a typical inductive load with its motor windings, draws a much smaller amount of reactive power from the source, since the demand for capacitive power is met by other active appliances

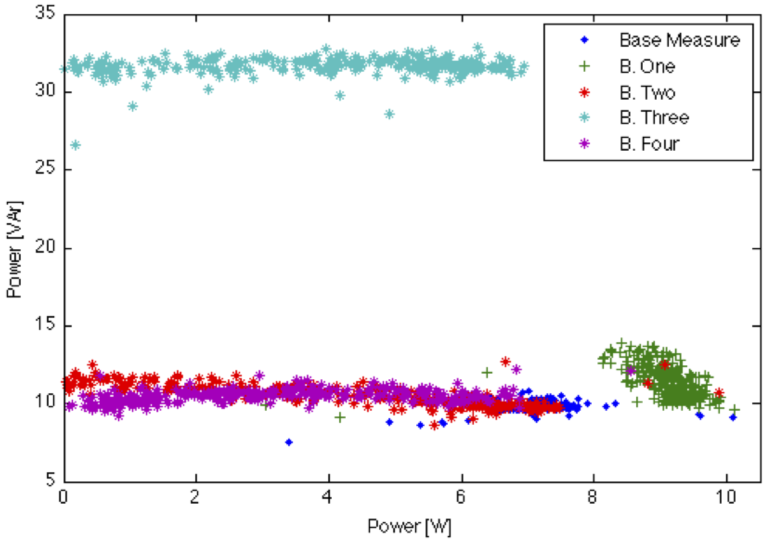

Figure 4.7: Real and reactive power measurements from the fluorescent light.

on that phase.

4.5 Discussion

In this study a central system for aggregated load measurements is developed to provide continuous data on energy-use and user interaction. However, the evaluation in this chapter focuses solely on the sensor measurements and not the user interaction, which will be fully evaluated in Chapter 5.

The goal of the measurement system was to design a portable, easy to install and only require an off-the-shelf micro controller to assure that the system could be easily implemented with current smart-meter technologies. The technologies to implement this system reflects these requirements. First of all, the CTs sensors clips around the wire and can be installed without making any changes to the existing infrastructure. Second, the micro controller is limited in terms of processing power to drive the radio and web link and transmit the data to the database. As this sensing and transmission link already exists in the current smart meter, the system proposed in this study adds some programmable space for calculations and analyzes in addition to the capabilities of the current smart-meter technologies.

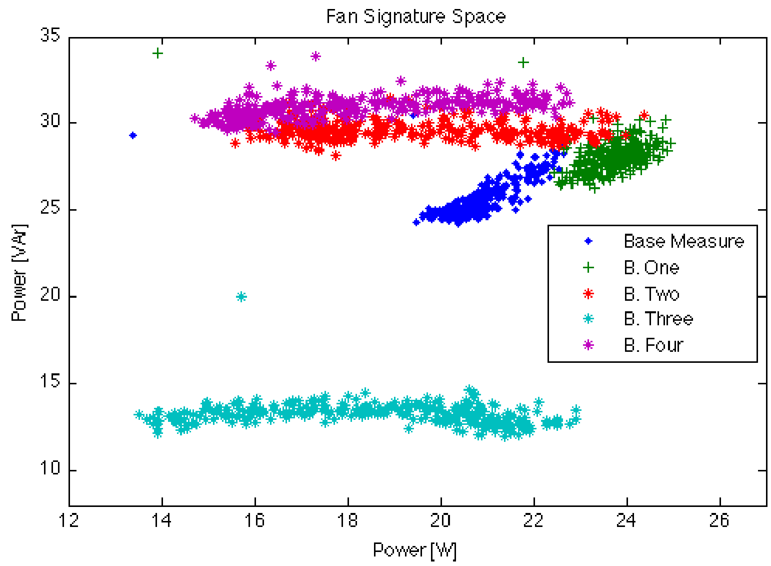

Figure 4.8: Real and reactive power measurements from the table fan.

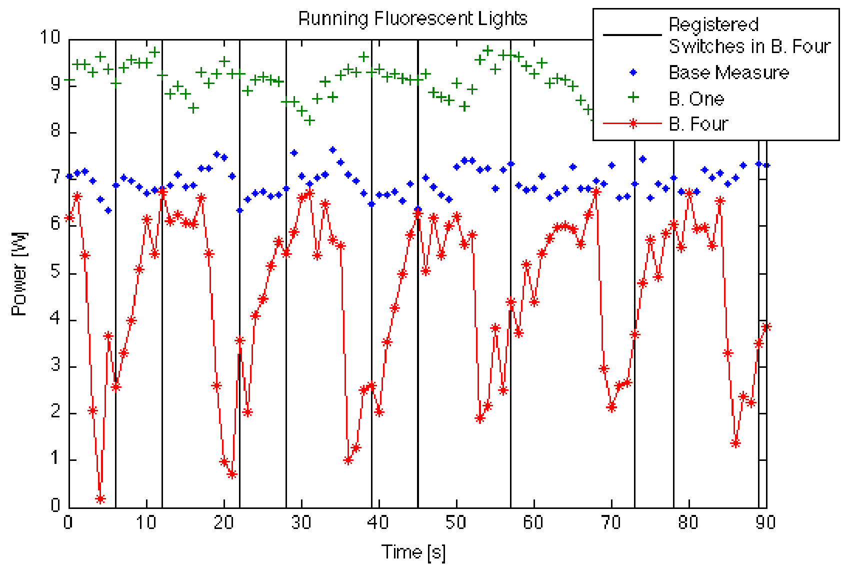

Figure 4.9: Time series profile of the fluorescent light with registered likelihood changes marked

This research, thus, exemplifies a way to update smart meters that were rolled out early without a full set of features for supporting the potential of modern energy use information (European Commission, 2011). Leaving some space for developers to create additional functionality that is strictly separated from the logic for the billing could be a way to rapidly evolve the functionality of smart meters.

The measurement system was evaluated in different buildings. By implementing the measurement system in four different buildings and logging the real and reactive power load from the same seven appliances, a method of isolating and comparing the effects of the buildings' electrical wiring and the already installed appliances between the buildings is introduced. Based on the results from this study, the conclusion can be drawn that household appliances' reactive power is influenced by the electrical infrastructure and secondary appliances. Most of the differences are due to the reactive power levels between buildings. This means that by reporting active power values many of the artifacts that are incurred by the buildings can be managed.

In sum, this chapter provides a straightforward example of how a system can be implemented to provide frequent household energy-use information and logging capabilities. In the next chapter the combined measurements of the user interaction and energy-use measurements are implemented in the field and further evaluated.

Chapter 5

Household Energy-Use Information Experiment

5.1 Problem Definition

Smart metering is being rolled out and users are seeing an increase in utility efforts to make the traditional energy-meter information more transparent. The European Union's directive on energy efficiency states that supporting individuals to save energy could have a major influence on market-wide adoption (European Commission, 2012). However, many business cases for smart-meter technologies in Europe rely on end-user's participation (Darby, 2010b). The European Smart Meter Alliance (ESMA, 2010, p.51) similarly concluded that "[e]stimates of likely benefits are needed to justify and design investments in demand response and smart metering."

Generally, consumers are interested in having more and better information about their energy usage (Chetty et al., 2008). However, how this information should be designed and presented remains uncertain (Kaufmann et al., 2013). Enticing and sustaining already busy individuals with direct access to their meter data has proven to be difficult. Two of the largest information driven businesses, Google and Microsoft, have both pulled the plug on their energy feedback projects due to a lack of interest.[1]

The necessary development of advanced meter infrastructure (AMI), where en-

[1] google.com/powermeter, microsoft-hohm.com

ergy use information is digitized and communicated bidirectional between the utility and the user, is still limited (Ipakchi and Albuyeh, 2009). So far, research has evaluated different types of energy-use feedback (Ehrhardt-Martinez et al., 2010; Alahmad et al., 2012), and measured information interaction related to the overall change in energy usage (Nunes et al., 2011). However, since these general results have little explanatory power on what information detail contributed to the impact (Abrahamse et al., 2005), research need to go beyond the question *if* users are affected by supportive systems and ask the question *how* they are affected (Watson et al., 2010; Darby, 2010c).

This chapter focuses on the research gap of the relationship between users' interaction with energy-use information and their subsequent energy related actions. The aim is to better understand the correlation between the information use and behavior change. By developing an experiment where users are supplied with direct and historic feedback through a web interface (enabled by the non-invasive measurement platform presented in Chapter 4), a realistic platform that is directly related to the current capabilities of smart meters, as defined in Chapter 2 and the European Union's energy directive on energy efficiency (European Commission, 2011), is provided. The experiment allows logging of both interface interaction and energy usage that can then be correlated to changes in appliance operation.

The presented material is based on the published articles (Dalén et al., 2012) and (Dalén, 2014) and begin with the hypothesis development based on work related to direct energy-use feedback and information use from experiments. The hypotheses then guide the experiment design and methodology. The experimental setup is explained, followed by the results section. The chapter concludes with a discussion of the findings and the main contributions from this research.

5.2 Related Work

Field experiments of smart meters are abundant. This section will provide a general introduction to these trials, their potential effects on energy efficiency and knowledge and current gaps in the research. The main focus will be on the experiments of live and historic feedback, which is related to the current smart-meter services and the research forcus of this thesis.

5.2.1 Related Experimental Results

Due to the rapid expansion of the capabilities of information and communications technology (ICT), new venues for inquiry are opening up and experiments on the impact of supplying energy-use information are taking place around the globe (Ehrhardt-Martinez et al., 2010). These studies are not only research endeavors but also commonly industry ventures to meet customer demand for more customized solutions to explore new potential markets. Based on the result from these studies, general predictions are made about the impact of smart-meter technology and residential energy-use information on saving energy (Darby, 2006; Seal and McGranaghan, 2010; ESMA, 2010).

However, "[d]espite several years of claims for smart metering, actual implementation at the household level is in its infancy and there is little hard evidence yet on what AMI can actually achieve."(Darby, 2010b, p.454). In context of the studies' importance for future expectations of the impact of smart-meter information, it is worrying that the reporting from the trials are rudimentary (Fischer, 2008). For example, comparing the starting point and the end point of an experiment provides little information about a trend's development. Table 5.1 show an overview of a subsection of the experimental studies reviewed by Ehrhardt-Martinez et al. (2010) and (Raw and Ross, 2011; Thuvander et al., 2012), which are related to this study. Specifically those studies which analyzed direct and indirect energy-use feedback (from daily and more frequent) have been evaluated.

The reviewed studies show that the results vary considerably, even between studies that use similar feedback technologies. These previous experimental results suggest that a deeper understanding of how information is used is necessary to develop smart meters functionality further (Ehrhardt-Martinez et al., 2010). In this chapter a mixed methods approach is pursued to evaluate this gap in the research.

Table 5.1: Real-time energy-use information experiments overview

Study	Year	Feedback Tech.	Treat. (Contr.)	Feedback Freq.	Length	Reported Saving
(Bittle et al., 1979)	1976	Cards	15(15)	1/Day	2.5 months	4%
(Seligman et al., 1978)	1977	Cards, Light	10-20(10-20)	3-5/week	2 months	10.5-15.7%
(Winett et al., 1982)	1980	Cards, Video, Goal	12-16(19-20)	1/Day	2 months	17%
(Hutton et al., 1986)	1981	Dedicated	275(251)	1/Hour	12 months	NS-5%
(Van Houwelingen and Van Raaij, 1989)	1984	Dedicated	50(55)	1/Day	12 months	12.3%
(Dobson and Griffin, 1992)	1992	PC	25(75)	1/Hour	2 months	12.9%
(Haakana et al., 1997)	1997	Video	23(26)	1/Month	17-21 months	11-16%
(Brandon and Lewis, 1999)	1999	PC	20(20)	1/Month	9 months	4.31%
(Petersen et al., 2007)	2005	Web, Competit.	2(16)	1/Day	2 weeks	23%
(Allen and Janda, 2006)	2006	Dedicated (TED)	10(60)	1/Sec	3 months	NS
(Sipe and Castor, 2009)	2008	Dedicated (Blueline)	210(691)	1/Hour	6 months	NS
(Parker et al., 2010)	2008	Dedicated (TED)	15(2 Million)	1/Sec	2 years	NS
(Thuvander et al., 2012)	2011	Dedicated (Eliq)	13(19)	1/Sec	3 months	NS

5.2.2 Mixed Methods Approach

Most experiments of energy-use feedback are evaluated separately on a quantitative or qualitative basis. In the former analysis this is done in terms of how much energy that has been saved by the introduction of the information, and in the latter users are commonly asked about strengths and weaknesses with the current platform. The challenge with relying on a qualitative measure alone is that the source of a positive, neutral or negative treatment response is difficult to pin-point. Strict qualitative studies on the other hand, are sensitive to biased answers and therefore also suseptible to missing the causal direction of the analyzed information (Kaplan and Duchon, 1988).

This problem, of understanding causality and controlling experimental variables is introduced in Chapter 3, Section 3.2.1 and is a complex issue to handle. So far, the qualitative and quantitative analyses have provided answers to unrelated questions, even though they are thematically similar. Information-systems research has a unique position to cover multiple methods through its multidisciplinary roots (Venkatesh et al., 2013), which can mitigate the challenge of evaluating how information is used. The ability to combine qualitative and quantitative analytical methods is convenient for evaluating user's interaction with energy-use information as the actions transpire at the cross-section between the individual's incentives and the measurable impact thereof.

This study, therefore, explores the power of a mixed method approach to confirm or reject the measured energy usage and interface interaction with the participants own statements. This method can help to triangulate the reasons for a measured outcome, which is fundamental for developing energy-use information systems further (Watson et al., 2012).

5.2.3 Hypothesis Development

Considering the conclusions of the previous energy use information experiments, specific hypotheses are formed to structure this analysis. First, reviews of energy information recommend direct feedback as a way to elicit significant energy savings (Ehrhardt-Martinez et al., 2010; Fischer, 2008). Although recent experiments contradict this view (Darby et al., 2011; Thuvander et al., 2012), this study will test for a significant positive effect of providing direct access to energy information in line with the previous studies.

H1: Providing direct access to household level energy-use information will lower overall energy usage in comparison to the time before having this access.

Second, to further position the experimental results in comparison to previous studies, the overall treatment effect is evaluated over discrete months. Based on previous research we expect that there will be a novelty effect of accessing the information, which wears off as the experiment progresses (van Dam et al., 2010; Wilhite and Ling, 1995). Thus we presume that the time directly after the treatment start will be the time with the lowest energy use.

H2: The month directly after the treatment begins will correlate with the lowest overall energy usage.

In the third hypothesis, the interface interaction is used to perform a detailed evaluation of the direct and indirect effects of information use. In contrast to H1 and H2, which tests whether direct access to real-time energy feedback is effective overall, Hypothesis 3 (H3) test the correlation between individuals' interface use and their energy usage.

This chapter focuses particularly on evaluating the relationship between user interaction with information and how this information then is applied. Whether the interface interaction can be directly correlated to a change in energy usage is tested. Related qualitative analyzes in previous trials, of how real-time energy-use information is used, concluded that energy-use knowledge accumulates over time (signifying an indirect use) and to evaluate and control unusual events in the moment (signifying a direct use) (Hargreaves et al., 2013). From this no direct presumption can be made of whether a direct or indirect relation is predominant. However, similar to H1 and H2, this hypothesis will test for a positive treatment effect, which in this case means an active and direct curtailment of the energy usage. To explore the effect a window of five hours after the interaction is analyzed for changes in the energy usage.

H3: Interface interaction has a significant and negative correlation with a below average energy usage.

As mentioned in Section 5.2.2, this study's goal is to combine the quantitative results with a qualitative interview to gain a new perspective and a new understanding of energy information interaction. The effects will, therefore, be mapped onto the categories of interview responses of the participants to evaluate certain patterns.

5.3 Experimental Design

To test the hypotheses a pilot experiment is developed based on the portable, smart-meter measurement system presented in Chapter 4. The experiment is a

Figure 5.1: Overview of the experimental phases with real-time energy-use feedback (2012)

quasi-experiment with an interupted time-series structure (Shadish et al., 2002). This means that the experiment was carried out in two stages: the calibration phase and the treatment phase (shown in Figure 5.1).

In the initial baseline phase, between mid July and mid August, the metering devices were installed and the energy usage was monitored. There was no feedback to the households in this phase and the website functionality was limited to basic project information. The calibration phase served to establish a base level of energy usage. The treatment phase lasted for three months, from mid-August to mid-November 2012. During this phase, the participants were able to access all website functionalities, which included the real-time and historic energy-use of their households.

5.3.1 Participants

Fifty student households, of comparable communal living arrangement, voluntarily signed up to have the measurement system installed. The interested individuals were then asked to take photos of their fuseboxes and meters and provide additional information about their households (size in m^2, floor level, number of inhabitants, gas or electric stove).

The fuseboxes were inspected to prove that a non-invasive installation would be possible. Twenty five out of the fifty households were suitable and out-fitted with the sensor and communication unit, presented in Chapter 4. The sensors were distributed with its associated materials (e.g. cables, batteries and instructions) to each household. Figure 5.2 show an example of a fusebox with sufficient room for the installation (left) and another example where an installation was not possible

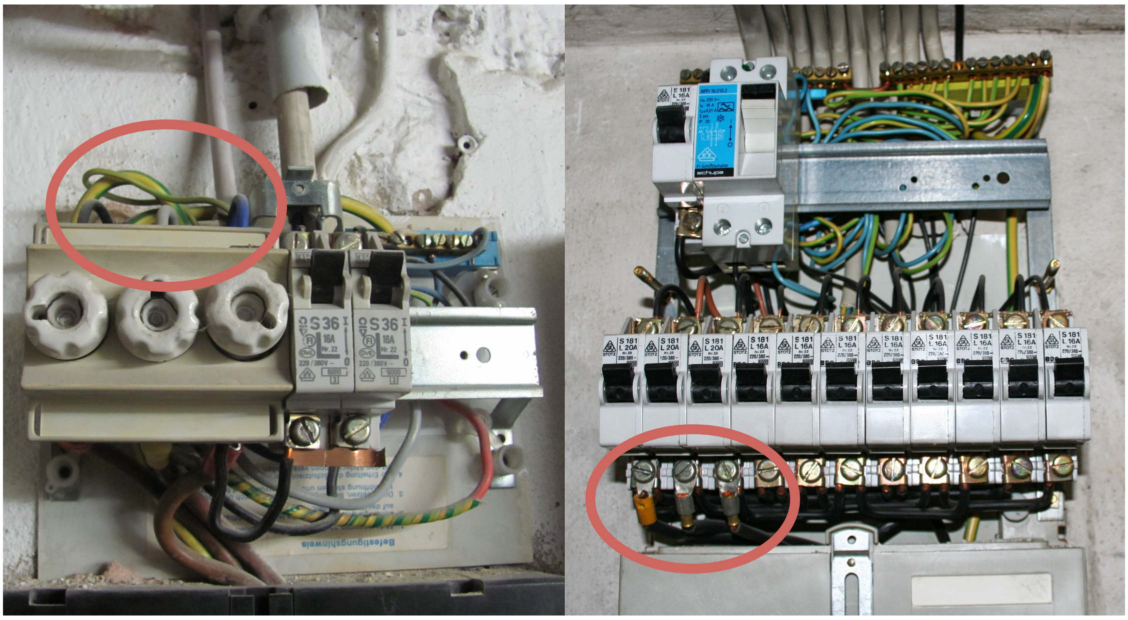

Figure 5.2: Two examples of fuseboxes showing the clearance to the main wires marked with red circles. The left example has enough room for the current transformers while the right one is too restricted.

(right). At some households extra hardware as routers had to be supplied in order for the system to work. Two households changed inhabitants during the project and were omitted from the experiment.

The whole installation process took between 30 minutes to an hour. Much time was spent explaining the system, energy-use information in general and the participants' reasons for participating in the project. The final functionality of the system was confirmed by accessing the admin interface on site. The support following the installation process was directed at keeping the sensors online and answering participants questions. This entailed debugging faulty soldering joints, electrical components, network setups and customizing the software to work with all the different existing routers in the households. Regular maintenance work such as replacing batteries and compiling support material were also handled in regular intervals.

5.3.2 Feedback

The basic functionality of the energy-use information system was based on a battery-driven sensor unit with three current transformers. The sensor unit sent calculated power values to a communication device via a radio link. The values where then pushed to the central database, which supplied the web-interface with its information. Further details regarding the hardware is presented in Chapter 4.

Web-based feedback has the advantage of being relatively inexpensive to install, and is scalable since it can use existing infrastructure (Darby, 2010c). Another benefit of web-based feedback from a research perspective is that the platform is shared between the user and information supplier. In addition to energy-use measurements, user-interaction data of the web based front-end was collected and logged for further analysis. In contrast to dedicated displays - where the display use is most often assumed - web-based feedback offers the opportunity to analyze the actual use of the display technology. This provides a quantitative method of exploring how information usage compares to a change in energy usage, which is not common in current energy-feedback research (Wood and Newborough, 2003). The study by Nunes et al. (2011) is an obvious exception where interaction and motion sensors around the feedback system is used to evaluate this aspect.

The modular web-based experimental framework is, therefore, designed around a database driven web site based on MySQL and AJAX. The incoming JSON data where timestamped and the ‘‘sensorID’’ was matched to pre-configured user accounts. The web-based design allows for multiple devices with internet access to view the information. The data was polled by an AJAX script and updated the dynamic rendering with new values.

The database driven web site allows the front-end developers to write and implement new features based on the available data in the database. There are currently two views implemented, the live and the historic view. Participants could log on to their household account with a username and password. The landing page after login is the live view where the power load is continuously displayed as data becomes available in a moving graph. Most German households are connected to the power grid through three main phases and these were kept in the interface to provide additional granularity. By separating the phases they could also be dis-

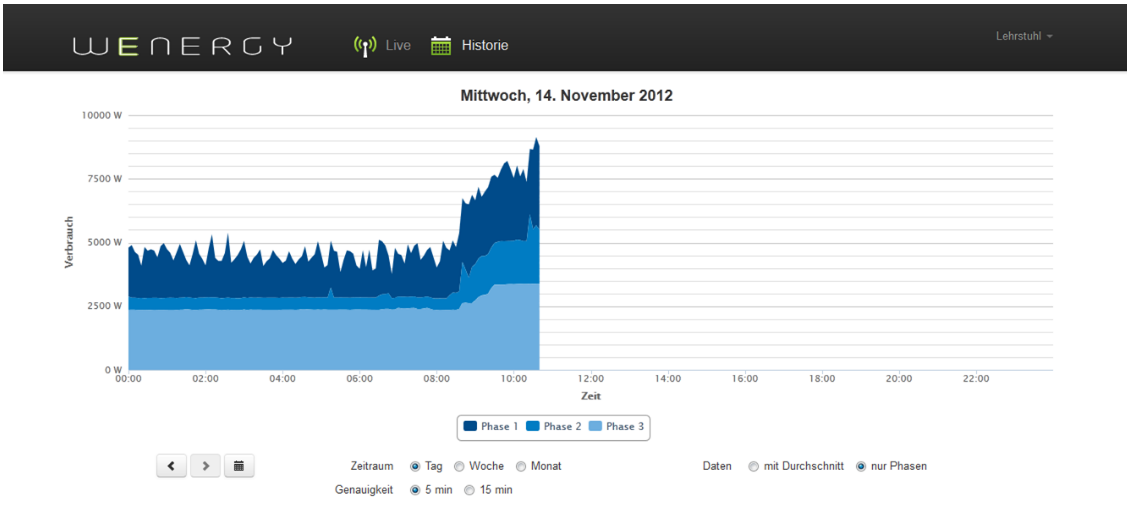

Figure 5.3: Web-based live feedback view based on three phase measurements

played individually or as an aggregated load in a stacked graph. Figure 5.3 shows this direct interface.

Every eight seconds, a new value is send to the server, and the chart moves to the left. This creates the feeling of live monitoring such as for medical devices. Additionally, a gradient progress-bar indicates the level of current energy use from green (low) to red (high). The gradient bar of the moving average power usage is based on two previous weeks, which makes the gradient account for seasonal effects and gradual energy-use changes made by the household inhabitants.

The energy-use feedback format that is normally followed by direct feedback is the aggregate historic view of these values. Past energy-use data is stored for further evaluation, which allows the energy user to focus on specific times in their energy usage. The actual energy use of the respective period is indicated in half transparent blue, and the historic average of the user's energy usage for the period is plotted unobtrusively in gray in the background, see Figure 5.4.

In the alternative, historic view, previous energy usage can be reviewed on a daily, weekly or monthly basis. Specific week days are compared with the same previous weekdays, weeks are compared to previous weeks and so forth. This allows for a direct reference to previous patterns without complicating the interface considerably.

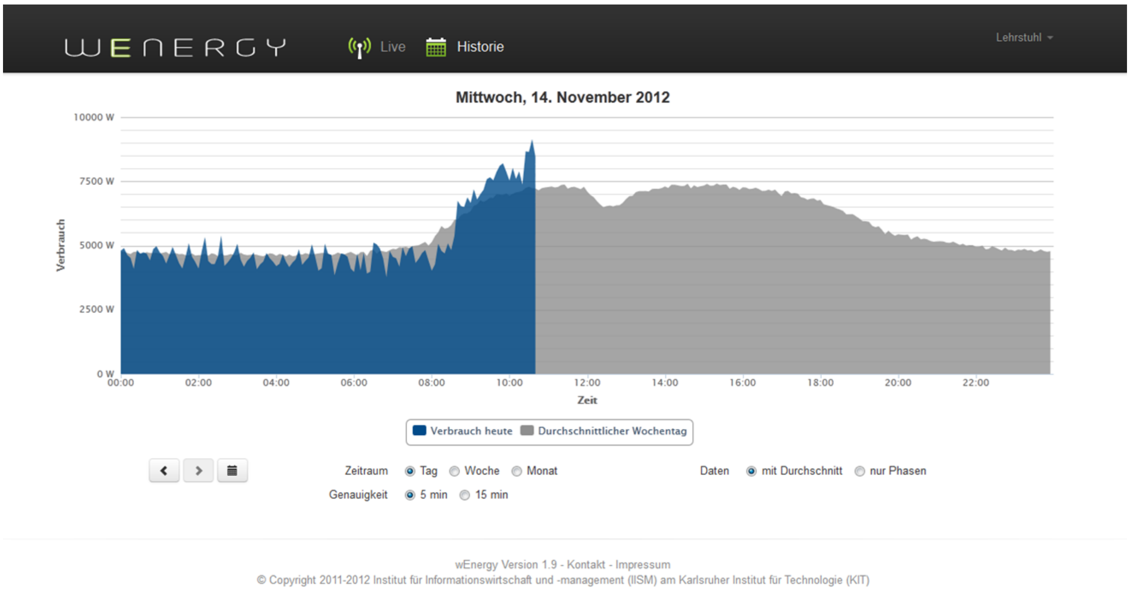

Figure 5.4: Web-based historic feedback view with current aggregated household power usage displayed on top of previous power-use data

Finally, the administrator interface allows complete control of the access to the different feedback views. An example of the control interface can be viewed in Figure 5.5. This granular control allows for a dynamic way of doing experiments and the possibility for “sandboxing” new features to a limited amount of participants before a general roll-out.

5.3.3 Experiment Data

Several types of data were collected for the experiment. Energy-use readings from the metering devices were recorded once every eight seconds. This data was then aggregated into 5, 15 and 30-minute intervals to display in the different historic view modes. In addition to energy-use data, usage of the web-based front-end was also collected. The energy-use data was logged and aggregated to 15-minute values before being statistically analyzed.

Of the 23 participants that stayed in the project until the end, 4 had to be omitted from the analysis due to sensor and communication outages and limited availability for support. This resulted in 19 households with stable hardware that

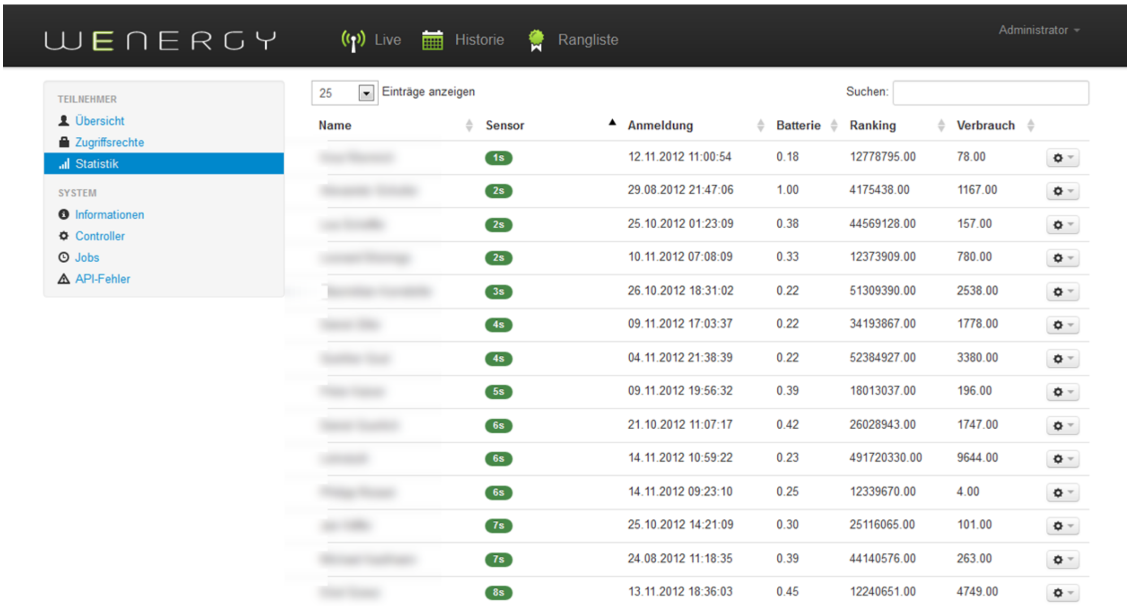

Figure 5.5: Administrator view showing the user management and access tools

were evaluated to test the hypotheses. Of these households, two had five residents, four households had four inhabitants, eight had three, and the final five apartments had two people living in them. This resulted in an average of 3.16 persons per household.

In addition, control group energy-use data were supplied by the local utility from 50 random households from the same location. These control group readings are unrelated to the experiment. The real-time energy-use treatment effects are analyzed both within the experiment (pre- and post treatment) and against the control-group data. Since the demographic of the latter is different from the student exclusive experiment no direct comparisons can be made besides supporting the evaluation of seasonal effects (weather and temperature).

5.3.4 Interview

The interview used the dramaturgical model suggested by Myers and Newman (2007). A semi-structured list of questions were used to allow for the participants to report their experiences freely (Gläser and Laudel, 2010), while covering the areas of interest; energy-use awareness, flexible energy usage and energy-use information privacy. The interviews involved 13 of the 19 households that participated in the full experiment and took place in the homes of the users to provide a comfortable setting. The 6 remaining households were unavailable during the time period for the interviews. Open descriptive questions were used to initiate and close the interview session.

The recorded interviews were coded by labeling specific phenomena in accordance with grounded theory (Strauss and Corbin, 1998). To build theory around the utility of energy-use information the phenomena were analyzed for common themes (Urquhart et al., 2009). By writing and sorting memos of these diverse subjects seven categories were established, which defined the participant's processes in regards to the energy-use information. These seven categories were:

- General energy-use practices
- Perceived energy-use savings
- Energy-use awareness
- Energy-use information practices
- Privacy concerns
- Display allowances and limitations
- Ability to shift energy usage in time

Every category was divided into a scale of how the answers fitted to the topic. For example, the users own estimated energy usage was scaled in three levels - low, normal and high - while the users reported strategies for saving energy were labeled in: *"no saving"* (no interest), *"no saving"* (no potential found), *"saving from behavioral change"* (less often used or more often turned off), *"base-load saving"* (not leaving appliances in idle or stand-by mode) and *"appliance exchange"*.

5.4 Results

This section will present the results of the quantitative and qualitative evaluation combined with the mixed method analysis of the field experiment. The seasonality of energy usage in the experiment region will be presented first to give a general understanding of the energy usage during the experiment time-frame. Thereafter, a more detailed analysis will be made on the within-subject experiment focusing on the effect of feedback access. Finally, the statements from the interview will be presented and compared against the changes in energy usage, both in relation to the experiment months and directly to the use of the web-based feedback.

5.4.1 General Seasonal Effects

Seasonal effects are outside the control of the experiment but can have a great influence on the energy used. For example, in a warmer period, there is less need for electric heating, however, this decrease in energy usage might be offset by a need for electric cooling.

The seasonal effects influence on energy usage is analyzed by comparing the control group against the experiment group. Figure 5.6 show the general trends of average daily power use of the control group (blue) and the experiment group (red). The figure uses the loess fitting procedure that uses a local selection of data points to fit the given line within a 0.95 confidence interval.

A seasonal trend is visible between the two groups, which share the same geographical location but differ in their origin. A comparable decrease in energy use appears before the treatment starts (indicated with vertical black line at day 46), between the two data sources, followed by a similar increase after the treatment start date.

A more detailed regression analysis is presented in order to analyze temporal and household composition effects on energy usage. For this analysis the data is structured and labeled in accordance with the factors of interest. An observation is the average power use of a certain household during a certain 15 minute time block. The factors are represented by dummy variables indicating whether it is present or not. For example, the binary dummy variable $Feedback$, indicates whether energy-use feedback was available (1) or not (0) in the respective time slot.

The first regression includes the $Feedback$ and $Control$ variables as predictors of the response in the levels of power load. Table 5.2 presents the statistical results from this analysis. On average the households in the experiment and control groups use 290.5 watts. This load amounts to an average energy use of 2.630 kWh per year and household, which is comparable to the German average of 3.413 kWh per year and household (Enerdata, 2012).

The tables provide the estimated power use (Estimate), standard error (Std.

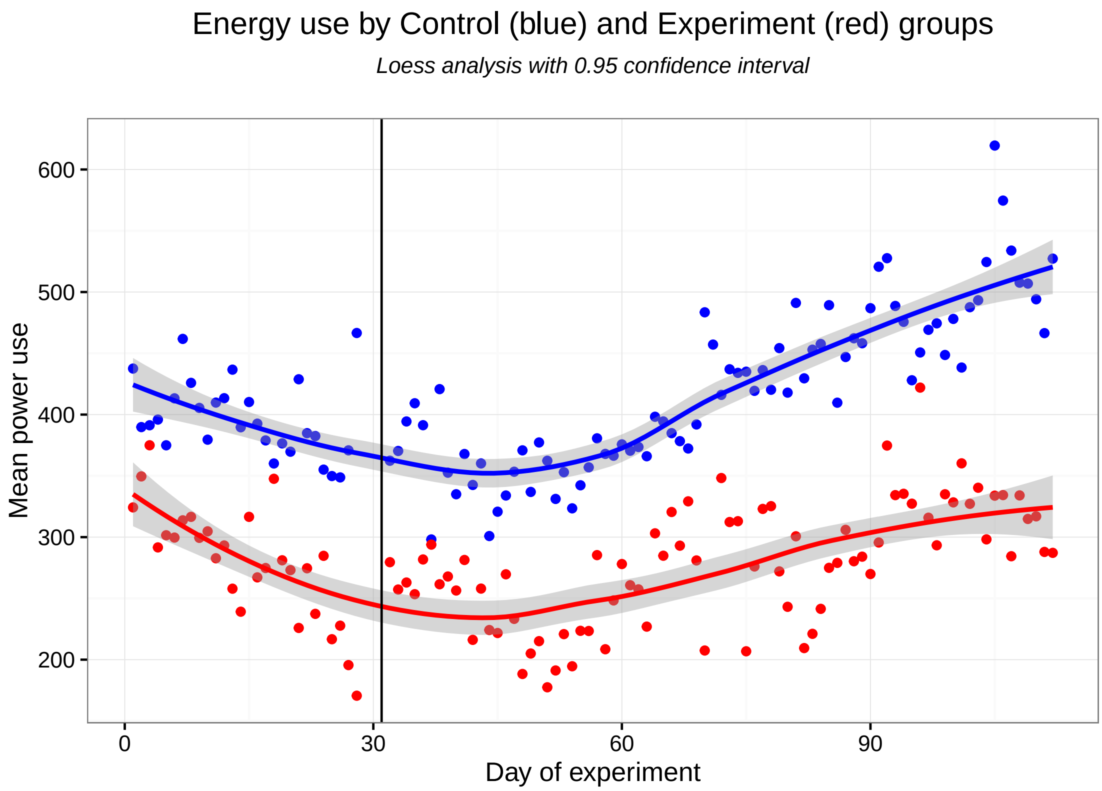

Figure 5.6: Energy usage of the control group (blue) and the experiment participants (red) over the entire experiment. Black line indicates the start of the treatment.

Table 5.2: Multiple linear regression with the power usage as response of *Feedback* and comparison group *Control*

	Estimate	***Std. Error***	***T***	***p***
(Intercept)	290.500	2.169	133.944	<0.001
Feedback [True]	-10.347	2.758	-3.752	<0.001
Control [True]	128.920	2.484	51.908	<0.001
Observations		298387		
R^2 **adjusted**		0.0185		

Error), T-statistic (T), and associated p-value (p). The negative estimate from feedback means that, on average, the period when feedback was given occurred when the power use was below average (approximately 10%). The control variable show that the control group uses almost 130 Watt more than the experiment group. These relations are also apparent in the previous figure (5.6) that presented the relation between the experiment and feedback groups.

The sample size is 298385 observations, where missing values exists in the combination of household and point in time there is data. The low $R^2 adjusted$ (1.9%) confirms that the two datasets are similar, both before and after access to the energy-use feedback was given.

5.4.2 Treatment Effects

To test the first hypothesis of a general treatment effect, a paired t-test was carried out. This test reflects the within-subject experimental design and the measurements of each subject before and after they had access to the interface. However, contrary to the stated hypothesis, no significant effect from facilitating direct access in real-time to individual's energy usage could be found between the pretreatment period (M=272.86, SD=412.24) and the treatment period (M=281.25, SD=405.14); $t(18) = -2.29$, $p = 0.03$. This suggests that the participants, on average, did not use the feedback to save energy in comparison to when they did not have real-time access to their energy-use data.

To explore if the participants changed their behavior during the experiment, the data was grouped by month. A one-way within-subject ANOVA was used to evaluate if there existed any significant differences between the energy use in the months of the experiment. The result from this analysis is pictured in Figure 5.7. Similar to the overall treatment analysis, the effect was also found to be non-significant over all four months [$F(3) = 1.96$, $p = 0.13$]. The second hypothesis that there would be a larger treatment effect directly after access was given compared to later months could not be confirmed. This suggests that the participants did not change their energy usage significantly in any of the experimental months.

To analyze the impact of the treatment in more depth, a subsequent multiple linear regression was done that focus on the treatment group. First the poten-

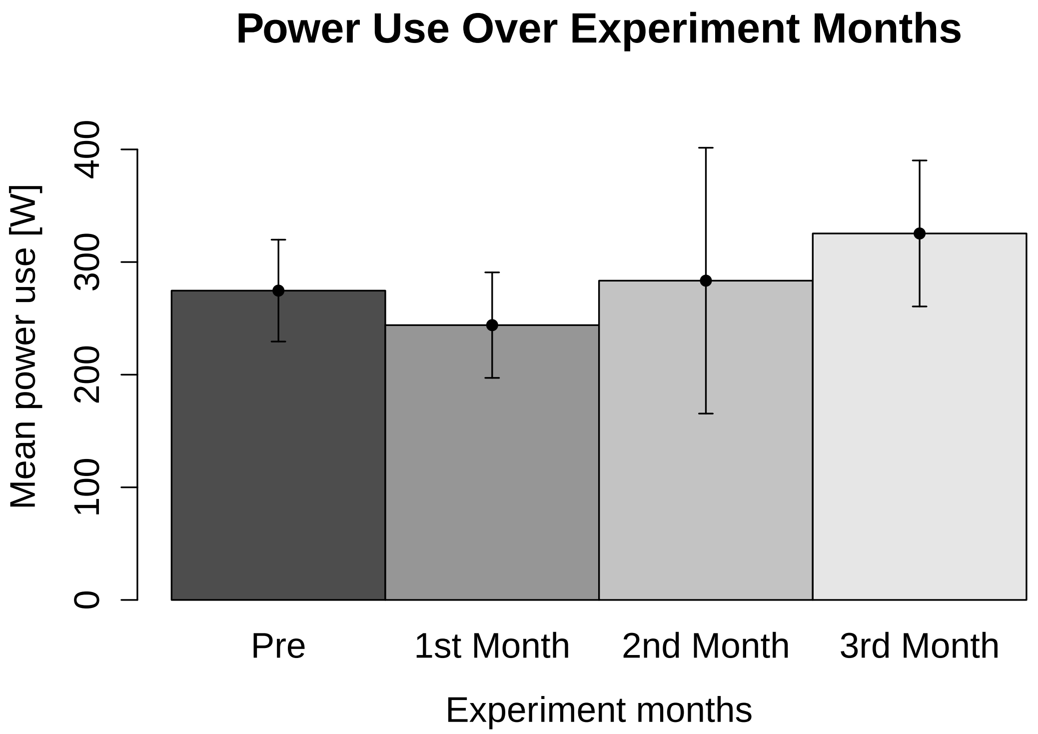

Figure 5.7: Treatment persistence over experiment months with upper and lower 95% confidence limits marked.

tial impact from specific households were explored followed by focusing on the number of residents. The data is structured with every row containing information on the number of residents and a series of dummy variables in order to control for household ($household_i, i \in 1, ..., 19$), the day of the week ($weekday_j, j \in Mo, Tu, We, Th, Fr, Sa, Su$), and the hour of the day ($hour_k, k \in 0, ..., 23$). Note that one dummy of each series becomes obsolete, since n - 1 binary variables specify the entire set.

Table 5.3 displays the results of the multiple linear regression that explains the respondent variable (power usage) in relation to the feedback predictor when weekday, hour of day and number of residents are taken into consideration. For clarity the dummy factors for weekday and hour of day are not shown in the table. Instead, Figure 5.8 indicates the estimates of these terms and their confidence interval (the full analysis is shown in Appendix A).

Table 5.3: Multiple linear regression with the power load as response to Feedback, Number of Residents, weekday and hour of day.

	Estimate	***Std. Error***	***T***	***p***
(Intercept)	104.163	7.468	13.947	<0.001
Feedback [True]	13.233	2.302	5.749	<0.001
N Residents	60.104	1.313	45.788	<0.001
Controlled for weekday		Yes		
Controlled for hour of day		Yes		
Observations		134091		
R^2 **adjusted**		0.063		

In contrast to the previous analysis, when both the control and the experiment group was analyzed, the feedback period shows a slight increase of power use when the focus is solely on the experiment group. The regression also shows that each additional resident increases the power use with 60 watts on average. The sample size for the experiment group is 134091 observations. The explanatory power of this analysis is slightly above 6%.

A second regression analysis considers the impact of the households, which makes the variable residents obsolete, since it is fixed for every household. The result is shown in Table 5.4. All effects prove to be robust and remain constant, in terms of sign, direction and magnitude between the impact of the number of residents and the households. The effects of feedback and household are statistically significant to the 1 percent level. The dummy coefficients for weekday and hour of day remain almost identical (cf. Appendices A.1 and A.2).

5.4.3 State-space Model Analysis

The previous ANOVA show that the treatment group has not changed their energy usage in a significant way. However, this analysis disregard the comparison to the general pattern of the general public at that time. A potential scenario that the ANOVA do not consider is that the overall energy-use in the area rise while the treatment group remain the same. In this scenario, a lack of change within the treatment group could in reality be a change in behavior compared to surrounding

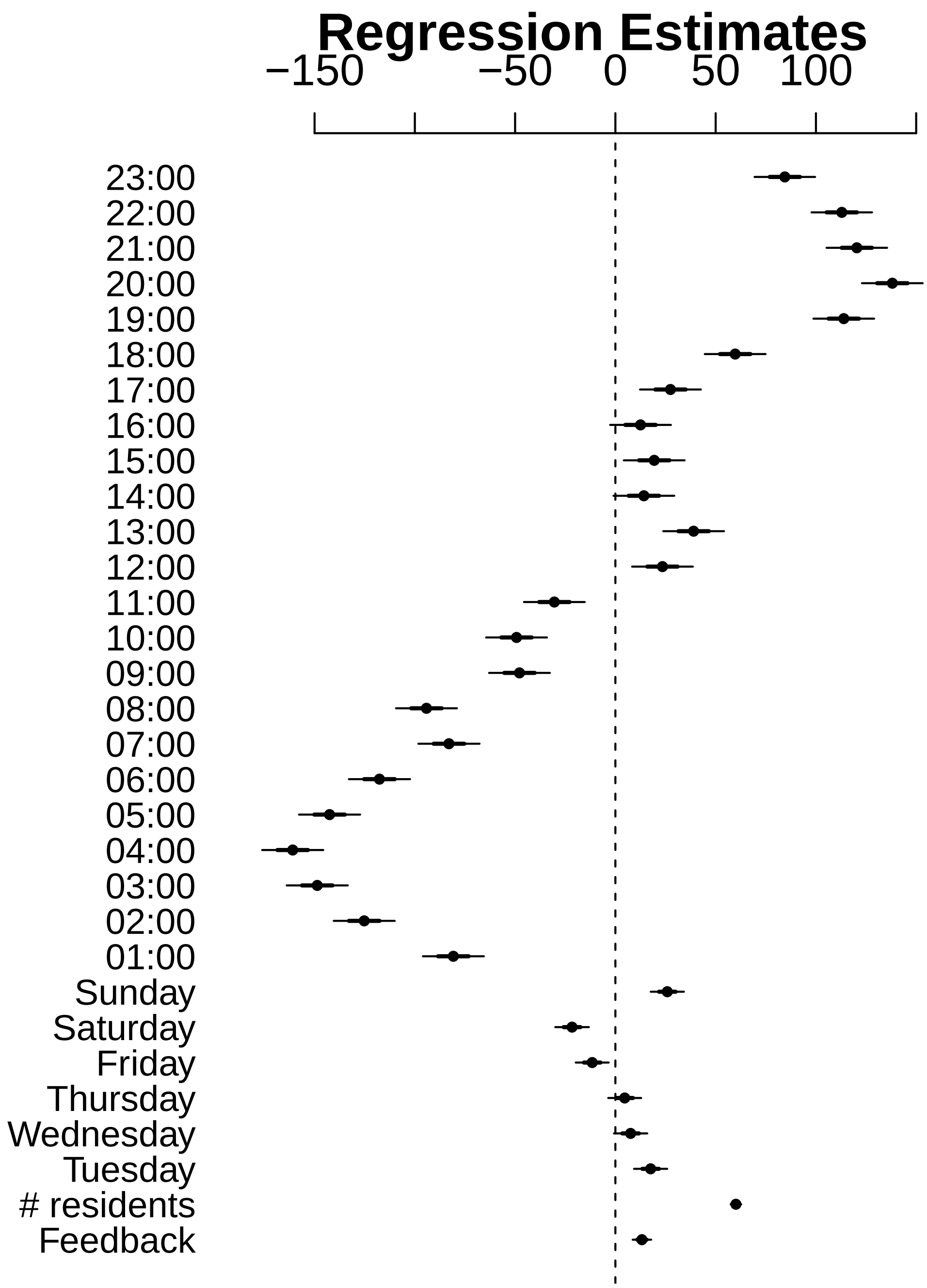

Figure 5.8: Coefficient ranges of a power load regression analysis

Table 5.4: Multiple linear regression with the power load as response to Feedback, Household, weekday and hour of day.

	Estimate	***Std. Error***	***T***	***p***
(Intercept)	404.307	9.173	44.078	<0.001
Feedback [True]	22.142	2.302	9.618	<0.001
Controlled for household		Yes		
Controlled for weekday		Yes		
Controlled for hour of day		Yes		
Observations		134091		
R^2 **adjusted**		0.114		

users.

To analyze this scenario a state-space model method has been suggested by Brodersen et al. (2015). This model trains a state-space model algorithm to match the pattern of the pattern during the baseline period while no treatment effect is expected and the relation between treatment and comparison group to be the strongest. During this time the comparison group and the treatment group have similar ability to influence their energy usage in terms of technology. It is important to remember, however, that the treatment group in this study is likely to be more homogeneous than the comparison group since they are all university students while the latter was randomly picked from the same region by the local utility.

The trained algorithm then estimates the likely energy usage during the ensuing treatment period if no treatment had been given. By comparing the actual pattern of the treatment group against this estimated "counterfactual" the treatment effect can be evaluated by seeing if they diverge or not.

Figure 5.9 shows the result of this analysis and give a strong indication that the effect of giving access to energy-use information is insignificant. This statement is based on that the actual energy usage (black line) is mostly contained within the 95% confidence interval (blue area) both before and after the treatment began (dashed vertical line). The post-treatment area probability of *p: 0.17263* emphasize this claim and the lack of a statistical significant treatment effect.

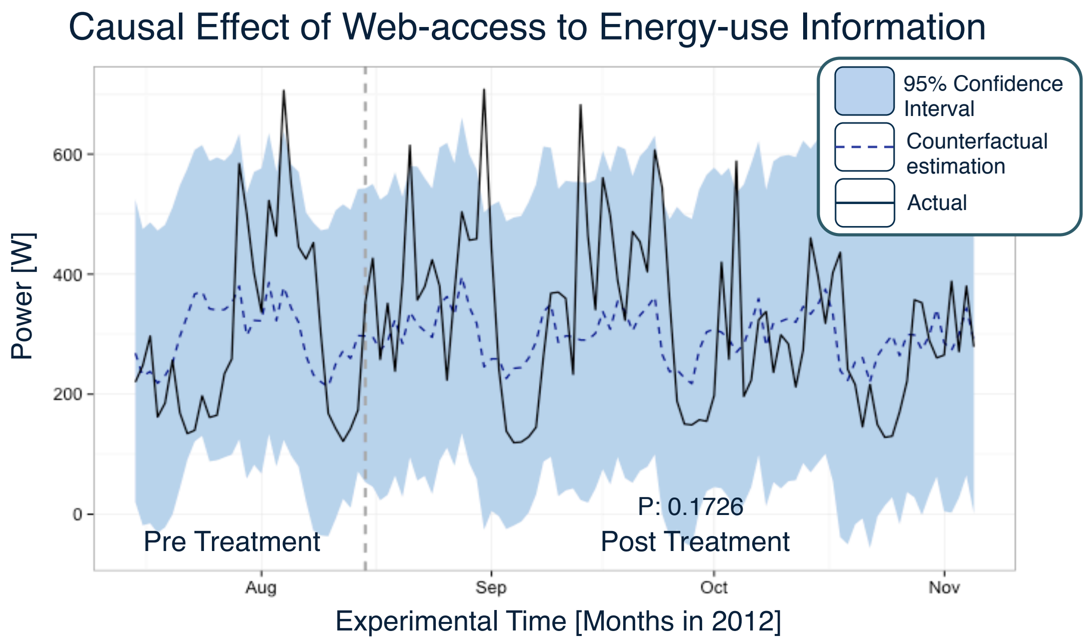

Figure 5.9: State-space model analysis of actual energy-use pattern divergence against calculated counterfactual energy usage.

5.4.4 Direct Effect of Interface Interaction

To analyze how the information was used in correlation to the use of energy, the two time-series of interface interaction and power use needs to be compared. For this purpose the hours after an interaction was marked with dummy variables $\{1, 2, 3, ...\}$. The 15 minute block in which the interaction fell was also marked with 1. In the analysis five hours were used to get a sense of the development after an interaction without it becoming too detached from the actual event. The influence from other, non-controllable factors are assumed to become greater as time passes.

Table 5.5 summarizes the results from the interaction analysis, which accounts for each household as they are individually linked to the interaction with the web site. The results show that only the second hour is significantly different to the first hour after the interaction. The second hour after the activity on the energy-feedback web-site the households use 80 watts less power on average when all other factors are held constant. The following hours, from the third to the fifth after the

Table 5.5: Multiple linear regression with the power load as a response to 2-5 Hours prior interaction and household, weekday and hour of day.

	Estimate	***Std. Error***	***T***	***p***
(Intercept)	406.921	17.266	23.567	<0.001
2 Hours	-80.896	25.108	-3.222	<0.01
3 Hours	-32.307	26.685	-1.211	0.226
4 Hours	-18.479	27.495	-0.672	0.552
5 Hours	46.606	28.938	1.611	0.107
Controlled for household		Yes		
Controlled for weekday		Yes		
Controlled for hour of day		Yes		
Observations		82933		
R^2 **adjusted**		0.134		

energy-use interaction, there is no significant effect in relation to the interaction. The full regression analysis table can be found in Appendix A.3.

5.4.5 In-depth Interviews

The interview results have been divided into the seven categories that became apparent in the analysis of the interviews (see Section 5.3.4). A summary of the interview results can be found in the Appendix B. This section will give a brief overview of how the results are divided over the participants and the broader relations between specific household inhabitants answers and their energy usage.

In the interview 11 out of the 13 users commented on their perceived energy use. The users appreciations were evenly split between frugal energy-use and a normal one, with 5 users in each group. One user related that their household was to be considered as a having high energy usage. This last household proved to be accurate in their estimate as they showed above average increase in energy usage over the experiment. Three households also accurately predicted that they were below average in energy usage, at least during the measurement period.

Prior knowledge about household energy-use was also a factor of the interviews. Five users said their prior knowledge was low and another five claimed to have a basic knowledge. One household related that they had almost no prior knowledge while another household said that they were highly knowledgeable. One final user made no comments on this topic. Interestingly, the four users who increased their energy usage the least were the once who viewed their knowledge level to be low or insignificant.

The usefulness of a web-based smart meter was considered to have high utility by most participants (7 out of 13). The remaining 6 participating households were evenly split between medium utility and low utility. A high utility correlated most with households who also had the least increase in energy usage over the experiment. However, there was one divergent example where one participant related that their household saw little benefit from the streaming energy meter, and still managed to keep their energy use on a low level throughout the experiment.

In the interviews there were four discernible energy-use strategies in relation to the given feedback. The "no strategy" group, where no changes had been made, was followed by 3 participating households. Another group who said they did not change their current behavior related that they had already done all they could do and saw no additional potential with the web-based energy-use feedback. This "no-potential" group was made up by two of the participating households. One of the largest groups (5 out of 13) claimed that they saved energy because of raised awareness and increased knowledge about what effective changes could be. Some of the individuals in this "effective operation change" strategy group also used the strategy of lowering base load by turning off appliances in standby mode. This final "base-load" strategy group were also followed by 5 households, out of which 2 households also followed the effective change in appliance operation. The users who claimed to have learned were the most effective operation change could be made correlated with the lowest increase in energy use (3 out of 4 energy conserving users). However, two users with the same strategy also had a medium and high increase in energy usage, which weakens this specific correlation.

All except two participants commented on their perception of having a dedicated display instead or in addition to the web-based interface focused on here in the experiment. One user criticized the introduction of a dedicated display with the argument that it would create tension between the inhabitants to have this constant reminder. However, most users were positive to the idea. Seven out of the thirteen stated one or two of the following benefits; that the dedicated display would keep them more aware since they would not have to log into anything, that abnormal situations would be more easily detected by having the information present and that it would work as a stepping stone for using the web-based feedback (where a more detailed analysis would be supported).

The ability to shift loads were considered to be small by nine households. One household related that they were flexible with certain devices, while the last four households could see themselves shift loads if the controlling could be automated. The answers were evenly split over the users, regardless of how they changed their energy use over the experiment period.

The final category of replies were related to the users' perceived concern with privacy and data protection. With a slight margin (5 out of 13) participants had considerable issues with the amount of information that could be possible to extract from a household's energy-use data. At the other end of the spectrum four participants explained that they had no concerns about sharing information about their energy usage. In contrast to the first group they felt that it did not give too much personal information and that other social services already had more personal knowledge. Similar to the response category about perceived ability to shift loads, the privacy concerns were evenly distributed over the all the inhabitants.

5.5 Discussion

Academic research on real-life energy-use behavior that closely follow end-users information interaction is sparse (Abrahamse et al., 2005). Lab experiments put participants outside their natural environment, which is crucial for studying energy-use behavior since it is inherently linked to the own home, particular habits and routines. The experiment presented here addresses this gap.

In this study we have demonstrated the functionality of a scalable framework to gather and provide aggregated energy-use feedback. Information and communications technology and embedded systems have clearly developed to the point that providing high-resolution data does not necessarily have to be expensive. By studying a sparse implementation of energy-use feedback with a mixed method approach, this study suggests aspects of how real-time feedback could have been used, which, to the author's knowledge, have not been commented upon by previous research.

This experiment made use of open-source hardware to provide control of the data acquisition and transfer of data. An energy-use information system was developed that allowed for real-time energy feedback at a frequency of one value every eight seconds. The calculated power value was provided graphically to households through a web-interface. The system received wide interest, and showed potential for further development with more information features. In order to provide accountability between real-time energy feedback and its influence on household dweller's use of energy this experiment provide a sparse feedback on real-time direct and historic values was given.

The general results from this experiment showed no significant decrease in energy use. This result is contrary to both Hypothesis 1, which expected an overall saving from the feedback, and Hypothesis 2, which assumed an increased saving the first month of the treatment period. As previously mentioned, these results are, however, similar to other recent experimental findings (Darby et al., 2011; Thuvander et al., 2012).

To rule out a savings effect more strongly a controlled experiment with randomized treatment and control groups would be necessary. This type of study was outside of the scope of this exploratory experiment. The comparison group used in this experiment share the same geographical and climate conditions, however, nothing can be said about their demographic similarities as this was not reported by the local utility. However, based on the analytical results presented here it is unlikely that energy-use information would motivate a significant treatment effect - all other parameters kept constant.A more recent randomized control study emphasize this claim, by reaching the same conclusion that feedback alone had no significant effect on energy usage compared to the randomized control Schultz et al. (2015).

In addition to randomized trials, running the experiment over multiple years also provides some potential comparisons with previous energy-use patterns. In the experiment presented here the interaction quickly peaked and then sharply declined over the first month until stabilizing with a few interactions per week or month depending on the specific household. Focusing on managing attrition is a whole research field in its own right. Related to this, 8 explore in an online experiment how interface design can accommodate efficient complex decisions that are constrained in reward and time.

Even though the interview provided statements of specific information usage, it was not always clear from the answers at what time the interface was used. However, by analyzing the interaction log patterns from using the information emerge. Hypothesis 3 could not be rejected as the regression analysis (Table 5.5) resulted in a significant decrease in power load two hours after activity with the online energy feedback was registered. More analysis of this result is needed to confirm this effect as there are many factors at play and the explanatory power (R^2 adjusted) is limited to slightly more than 13%.

The experimental setup provides a straightforward example of how a system can be implemented to provide highly detailed user information and logging capabilities. The experiment was intended to introduce a sparse amount of information features in order to provide accountability between the results and the real-time feedback given. By iteratively adding informational features or using different treatment groups, the individual information formats can be evaluated in more detail.

The study explores a quantitative metric with which new theories and hopefully a better understanding can be developed. Specifically, how the statistical tools of multiple linear regression can complement the qualitative interview that is predominantly used today. In this study, this allowed for a detailed evaluation of the moment of information interaction. This point in time is particularly important to understand, as it is also the juncture were the researcher or information provider has the ability to communicate with the user.

The mixed method approach, which combined the quantitative multiple linear regression analysis with qualitative interviews, provided rich information to help explain the patterns found by the treatment. Only relating one side of this analysis would have left open questions. For example, what the incentives were for a change that had only been measured or whether or not the interviewers behaved as they were saying. This study thus provides a practical example of how to combine research methods to support Kaplan and Duchon (1988, p.583) statement that *"[n]o one method can provide the richness that information systems as a discipline,* *needs for further advance."* This analysis gives valuable insight into some of the potential distinguishable traits of the households that changed their energy use compared with the households who remained on the same increasing trajectory as the control group.

The results show that a low or frugal perceived current energy-use level correlated with the participants who also increased their energy usage the least. One participant related that they *"... had always been careful and always shut off their lights".* The opposite correlation was found for participants who considered that their energy use was average or higher related that they *"... didn't really think about it"* or were *"... rather wasteful"* with their energy usage.

Even more interesting is that the participants who's energy usage increased the least also perceived that their knowledge of energy usage was the most limited. The combination of judging that the current energy usage is low with little previous knowledge of energy usage in households distinctly defines the users who used the least amount of energy over the experiment. Future experiments over multiple years would be needed to evaluate this correlation as these households might have had similar energy-use patterns in previous years. The size and number of inhabitants could also influence the impact of the energy-use feedback. For example, fewer inhabitants means that each has a bigger effect on the household's total energy usage. In contrast, several inhabitants might drown the effect of a few active participants.

The third response category suggests that learning what impact each appliance have on the energy usage will assist users to limit their energy-use. Similar to a limited initial knowledge and a perceived frugal energy usage, learning about what appliance is efficient to target also correlates with the households who increased their energy usage the least. The users who learned about appliances impact on their household's energy usage stated *"[w]e tried to see what the appliances looked on the [online] graph"* or they *"... looked at what happened when [they] turned off and on lights and the stove and so on".*

The same group of users also considered that the web-based access to their energy-use data had a high utility. However, this statement was also common among users who increased their energy usage more during the experiment. Participants who focused on shutting off standby operating modes showed a limited correlation with the actual overall change in energy usage.

The final three interview response categories, which have a more indirect relation to the energy-use feedback, did not show a direct correlation to the participants' energy-use behavior. They are included in this discussion as a potential input for future research and energy-use feedback design decisions. A dedicated display was sought after by most participants. This would save them from logging in and provide an indirect alarm when the energy-use would reach abnormal usage levels. Similarly, most participants related that they would be willing to shift loads in time if it did not impact their normal routines. For example, washing clothes at night or automating the control to shift certain appliances was frequently mentioned as acceptable operational changes.

Users were also asked to talk about their views on data privacy in relation to energy-use feedback. About one third of the users were not concerned about their privacy or that someone with access to the energy-use data stream could learn about their decisions and whereabouts. Indicative for this stance is this quote *"... doesn't bother me, we have a pretty unregular daily routine"*. One third thought it depended on *"... how detailed the transmitted data is"* and whether or not it would outweigh the benefits from saving energy. The last one third had considerable concerns about their privacy and wanted certification that it would be properly secured and that reselling energy-use information would be prohibited.

From the correlations presented it seems that direct access to energy-use feedback will support users with rudimentary knowledge about energy usage and have an internal drive to take advantage of the supplied information. Long term studies with direct energy-use feedback is proposed to improve the reliance in these theories. More advanced feedback that have the ability to grow as the knowledge levels develops and inspire more knowledgeable users is proposed to counter the main challenges of engaging participants (higher level of knowledge and the isolation of energy savers). Much can be learned from existing energy-use feedback systems and in Chapter 7 the design of a more adaptive system that focus on these challenges will be presented.

This study confirms that real-time energy-use information in itself has little overall impact on individuals energy usage (Darby et al., 2011; Thuvander et al., 2012). Thus, it is questionable if the current smart-meter roll out, which could allow for near real-time feedback (15 minute level), will be able to get users to change their habits as previously believed (Ehrhardt-Martinez et al., 2010). A wider and more engaging variety of information has been suggested to elicit a stronger response (Fischer, 2008), and should therefore be further evaluated.

This discussion highlight several potential ways of providing energy use information to support sustainable energy decisions. By interviewing users about their energy use and measuring this use, this chapter shows that some information leads to energy saving in its current form while other alternatives need to be more clearly presented. Instead of analyzing peoples behavior in terms of the saved energy, the results presented here show the different ways and reasons why users interact with energy-use information.

The results from evaluating how real-time energy-use information is used have practical ramifications for the development of energy information. In order to support the end-user to get this information more easily an enhanced design should be developed. Ambient feedback is one option, as this allows for temporary events, like sudden high energy usage, to be unobtrusive and effectively communicated (Maan et al., 2010). Another more invasive option, which was suggested in one of the interviews, would be to provide a service that could push specific event notifications or energy-use summaries to the user by email, a tweet, or text message. This would relieve the user of having to log in and parse the interface for a particular point of interest that is easily extracted.

Based on these results the second part of this research study will first analyze a measuring platform to support a more personal energy-use information service in Chapter 6 and provide a first iteration of the informational services' design in Chapter 7. Inevitably, using Green IS to provide support for more informed energy use will offer a clearer incentive to the user to use energy efficiently. In this context it is important to remember that a *"[s]imple design that doesn't overwhelm the consumer with too much information but does require consumer participation is key"* (Ehrhardt-Martinez et al., 2010, p.35). To understand the balance of information detail and information overload, Chapter 8 analyses how information feedback can be categorized to enable users to find local optima.

Part III

User-Level Energy-Use Information

Chapter 6

Appliance Energy-Meter Engineering

6.1 Problem Definition

The previous part of this thesis analyzed how the change from analog energy-meters, which have to be read manually, to digital “smart-meters” would affect a certain group of residential users. The change in measurement frequency that this change is substantial, going from once every year to at least a 15-minute resolution (ESMA, 2010). However, as was clear from the presented field experiment and its related literature, the result of more up-to-date information show little or now impact on a household’s energy usage (Darby, 2010a; Ehrhardt-Martinez et al., 2010). As discussed in Chapter 5, users are better able to use energy more efficiently if they understand their energy usage on an appliance or user level, which is not possible with household-level energy-use data.

Measuring more often with a higher frequency or by using distributed sensors provides even more detail of a given signal. This way, more subtle changes can be detected and evaluated, which users already tried to extract from the aggregated energy-use information. To separate single appliances among noise and several other appliances with the central measuring approach, research laboratories are currently analyzing values on a resolution of up to 100 million data points per second (Patel et al., 2007). By measuring with such high frequency, Switched Mode Power Supplies (SMPS) that alternate the power output over 10 thousand times per second, can be detected. SMPS are interesting as they are becoming common in consumer devices and by measuring their noise characteristics an additional signature is provided. However, at these sample-rates more sophisticated and costly hardware is needed. For example Nunes et al. (2011) and Patel et al. (2007) both use personal computers in their disaggregation experiments, which would be prohibitively costly to scale to a general household level (Zeifman and Roth, 2011).

The second alternative, with distributed sensors, which was introduced in Chapter 3, Section 3.1.3, appliance information is given with minimal interference. By using similar methods to appliance disaggregation, it is therefore possible to go further and also disaggregate appliance operation patterns.

To scale both of these solutions it is necessary to understand the measurement accuracy requirements for the disaggregation result. The main goal of this chapter is, therefore, to analyze how measurement sample-rate influences the ability to disaggregate appliances and operation modes with the commonly used appliance load features in previous literature. Since most previous methods of disaggregation deal with appliance-level disaggregation, this will also be the main focus of this chapter. Aggregated household energy-use data is more noisy and complex to analyze, which means that the description of these methods will also be applicable for disaggregating operating modes from appliance-level measurements.

This chapter is based on the published article Dalén and Weinhardt (2014) and the submitted and revised article Dalén and Krämer (2014b). The related work in the next section gives a brief introduction to the topic of appliance disaggregation and then describes some of the general concepts. Based on these concepts, an experiment model will be developed in MATLAB to parse and evaluate the laboratory energy data. After presenting the results from the analysis, the findings will then be discussed as well as their potential ramifications on future smart-meter developments.

6.2 System Design Properties

As introduced in Chapter 3, Section 3.1.2, Hart (1992) first implemented a platform for distinguishing patterns of distinct loads and appliances in households in the late

Figure 6.1: General overview of an event-based framework for household measurements and the following appliance and operation mode disaggregation steps.

eighties. The patented system used embedded electronics to measure and analyze the load patterns at a central point (the electrical mains power line). Since then, several alternatives with the same goal of disaggregating single load features and individual appliances have been presented.

The frameworks used by the different proposed systems all follow a general pattern (Du et al., 2010). Figure 6.1 gives a graphical overview of the general process of sensing, extracting and recognizing appliance loads and operating modes. First, the load is measured over the cycles of alternating current and voltage, which are digitized for further processing. This step will be introduced in Section 6.2.1 and cover the current range of methods of sensing appliances. Second, the data is then analyzed for changes that signify a load event. This is, for example, when an appliance is turned on or off, or when the appliance changes its state automatically. The event detection will be explained further in Section 6.2.2. Third, when an event has been detected, the data around this point is evaluated to extract and compare specific load features. A range of known appliance features will be presented in Section 6.2.3 to provide some background for the design of this study's analysis. Fourth, the calculated features are then compared to known appliance features to single out individual appliances in Section 6.2.4. Finally, in Section 6.2.5, the systems related to eliciting energy-use behavior and recognizing appliance operating settings based will be presented.

6.2.1 Data Acquisition

The flow of alternating current (AC) electricity can be sensed either directly inline with the electricity flow or indirectly through its magnetic fields, as explained in Chapter 4, Section 4.2.1. The number of different loads that could potentially be included depends on where in the tree-like electrical circuit structure of a household the AC is measured. To monitor a whole household either a few sensors could be given a central location or multiple sensors need to be distributed throughout the house (Gupta et al., 2010).

Multipoint sensing is normally located at the connection between the fixed household outlets and the appliances. University campus laboratory facilities are normally equipped with this kind of sensors to collect measurements that represent the individual appliances (Reiner et al., 2010). The data from the distributed sensors is well suited as training data for evaluating different methods of central appliance load disaggregation. This study will use a publicly available dataset from such a facility in Pittsburg, USA (Anderson et al., 2012). The use of a public dataset reflects this study's goal to enable a more cumulative research development, as other research teams can more easily verify, compare and extend their analyses based on this openly available dataset.

Measuring at several points throughout a house is more invasive and more expensive than the option of using a central measuring point. The potential of applying more sophisticated software instead of more hardware is a major motivation for focusing on a central solution (Zeifman, 2012). A single-point sensor is most often placed at the electrical mains, by the fusebox or electricity meter. However, depending on the features that are measured, other placements are also possible (Patel et al., 2007). While the central location of single point sensing is often convenient, the measurements can be influenced by other connected loads or the electrical wire infrastructure (Gupta et al., 2010; Dalen and Weinhardt, 2012).

Apart from the amount of hardware necessary for the system, the factor that impacts the cost, and thus the scalability of an appliance load disaggregation system, is the data sample-rate with which the electricity flow is measured. The sample-rate is generally divided into a macro and a micro level (Zeifman and Roth, 2011). In the macro category of disaggregation studies the sample-rate range from 1 data point per second (1Hz) (Hart, 1992), to 1 data point every 16 seconds (0.0625Hz) (Farinaccio and Zmeureanu, 1999). There is also research on disaggregating data on a 15-minute basis (Baranski and Voss, 2003). However, this is done "offline" by parsing the data multiple times and is therefore not suitable for direct feedback implementations, which this study focuses on.

At the micro frequency level, with more frequent measuring than once per second, additional information about the measured appliances' load characteristics is provided. The studies in this category range from a data sample-rate of a few kHz (Saitoh et al., 2008; Laughman et al., 2003) to 100MHz (Patel et al., 2007). The publicly available dataset that this study is based on has a sample-rate of 12kHz (Anderson et al., 2012). This data thus allows for an exploration of both highly detailed features and, by downscaling the data, comparative data for lower frequencies.

6.2.2 Event Detection

The event detection is an optional step depending on the subsequent method of analyzing the data. For example, an artificial neural network method analyzes the most likely combination of appliances for each evaluated dataset (Srinivasan et al., 2006). This method does not wait until a change has occurred and will therefore simply confirm that the same appliances are running if no change has been made. With this method no assumptions need to be made about individual appliance events. The downside of using non-event based methods, is that all combinations of potential appliances are needed for the analysis. This makes it challenging for the non-event approach to scale beyond 20 individual appliances (Zeifman, 2012).

The "building level fully labeled event based dataset" (BLUED) used in this study is based on individual appliance events. The event detection allows for focused analysis around certain points of interest, signified by the events (Norford and Leeb, 1996). Being able to parse the data in defined sample windows has the added benefit of limiting the number of appliance "finger prints" that need to be stored, since theoretically no combinations of appliances will be evaluated. Naturally frequently occurring events from a combination of appliances can also be stored for recognition.

This study will not use an event detection step as the primary focus is on the feature matching step and as the event detection has already been well described by the authors behind the BLUED (Anderson et al., 2012). However, a brief explanation of their design is included here for completeness. Their event detector method is based on the well-established General Likelihood Ratio (GLR), which can detect changes in signals as they pass through the filter "online" (Basseville and Nikiforov, 1993). The online aspect is important for creating a responsive system. The specific event detection design follows the GLR implementation described by Luo et al. (2002), with the addition of a power change threshold proposed by Berges et al. (2011).

The GLR works on a sample set of incoming data. This sample set is divided into two smaller sample windows around the data point that is being analyzed for a potential event. First, a threshold change in power use of 30W is analyzed through subtracting the two sample windows. The 30W limit was considered by Berges et al. (2011) to be a good balance between detecting events while removing much of the falsely labeled noise as events. The same threshold was later used by Anderson et al. (2012) when parsing the BLUED. Second, a likelihood ratio is computed on the data that is above the threshold of power change, or otherwise set to zero. Third, the results from the likelihood ratio test are summed up into a test statistic. In the fourth step, the data point with the highest test statistic receives a vote, signifying the most probably event point. Finally the whole sample set is shifted to include new data points, which through the same procedure adds another vote to the most likely event point in the set. When a certain point has received enough votes, the sample set of data for that event is extracted and the specific features are calculated. The current methods for the feature extraction will be reviewed in the next section.

6.2.3 Feature Extraction

The measured voltage and current has several layers of information or features that reflect certain characteristics of the appliances (Liang et al., 2010a; Ito et al., 2004). In this section, the different features that have been used in related research will be presented. These features will be used in this study's matching step.

The most well-known feature of any appliance is its real power, which is the average of the product of multiplying the supplied instantaneous voltage and current. Purely resistive loads, for example electrical heating elements or incandescent lights, use real power. By evaluating the real power of an appliance operating modes it can be singled out if it has a stable and large power use. However, many appliances vary and overlap in this feature, which makes it necessary to provide more dimensions to the solution space (Laughman et al., 2003). For modern appliances with rectifiers and pulse-width-modulation or appliances with motors, the load creates a mismatch between the current and voltage sine waves. This results in a reactive power. By combining the real power (P) and the reactive power (Q), each appliance can be placed in two dimensions. Having two distinguishing characteristics assists in detecting unique appliances. The so-called PQ feature was, until recently, frequently the only feature used in appliance disaggregation research (Hart, 1992; Berges et al., 2009; Cole and Albicki, 1998). One common critique of the PQ feature is that it requires distinct power levels to recognize appliances' state changes. Dimmable lights or computers, which ramp the CPU up or down depending on current demand, are hard to detect with the PQ feature (Laughman et al., 2003). Furthermore, the variability of some appliances' normal operation overshadows a whole range of devices. For example, an office lamp, which use little power is difficult to detect with the PQ feature when it is running in parallel to a kettle, which has a variable steady-state power use (Weiss et al., 2012).

One approach to distinguishing appliances that overlap within the PQ feature was suggested by Norford and Leeb (1996). These researchers found that even though the startup of different appliances reached similar power levels, the shape of how the appliances transitioned between the states was different. For example, a fluorescent light and an induction motor can both reach 400W power surge at the moment they were switched on. However, the power transient of the motor falls slowly, due to inertia, while the fluorescent light's power transient falls more abruptly. The so called "Switching Transient Waveform" (STW) (Norford and Leeb, 1996), was used by Berges et al. (2011) to disaggregate appliances with a correct classification result of 82%.

While the STW feature focuses on appliance characteristics at times of state change, appliances can also be identified from their signal form at a steady-state as well. Both the Instantaneous Power Waveform (IPW) and the Instantaneous Admittance Waveform (IAW) are examples of these kinds of features. The IPW uses the instantaneous power over time to describe the shape, while the IAW uses the instantaneous admittance (current divided by voltage) over time. Both features have been used quite successfully with between 70 and 80% appliance detection accuracy (Liang et al., 2010b).

The activity of an appliance can also be described without using voltage. For example, the "Current Waveform" (CW) feature is describing the altered sinusoidal curve of the current measurement. The CW feature has received mixed reception by the research community: one team claims it to be a distinct appliance feature (Lee et al., 2004), while another criticizes it for being too similar across different appliances (Zeifman and Roth, 2011). In disaggregation tests the CW feature has been equally successful as the IPW and IAW features, with around 75% detection rate (Liang et al., 2010b). At much higher measurement frequencies and by limiting the tested appliances to seven distinct ones 97% were found using the CW feature (Suzuki et al., 2008). However, in the same study comparisons with 15 appliances were made and subsequently the detection rate fell to a little more than 60%. Current specific metrics neglect the voltage and evaluate the peak current (Ipeak) and the root-mean-square current (Irms) features (Saitoh et al., 2008). These two features were reported to have a detection rate of 80% among 94 appliances. However, the experimental design and whether or not appliances had overlapping times of operation was not reported.

The final feature reviewed is increasing in popularity and is based on the harmonics (H) of the current measurements. The harmonics are multiples of the fundamental frequency (either 50Hz or 60Hz depending on the country) and are directly related to an appliance operation, particularly for semiconductors and power converters (Arrillaga and Watson, 2003). Fast Fourier analysis can analyze an appliance's harmonics in the frequency domain and with its information many overlapping appliances in the traditional PQ feature can now be separated Laughman et al. (2003). Harmonics have also been used as a separate feature successfully. For example Srinivasan et al. (2006) managed to detect 60-100% of the 8 appliances they were experimenting with.

The harmonics feature is limited by the available measurement frequency, as each higher harmonic requires the fundamental frequency multiplied by an additional integer. Furthermore, to recreate a signal, at least double the measurement bandwidth is needed according to the sampling theorem. This means that the measurement frequency must be at least double that of the harmonic level in focus. In practice this is not a problem as harmonics above the 11th are in most cases neglected (Zeifman and Roth, 2011). The 11th harmonic level could thus theoretically be evaluated at a sample-rate beginning at 1320Hz (1100Hz) for a fundamental frequency of 60Hz (50Hz).

Features that will not be included in this study's evaluation are the electrical noise feature (Patel et al., 2007), electromagnetic interference feature (Gupta et al., 2010) or features that are based solely on voltage (Ruzzelli et al., 2010). The first two features (electrical noise and interference) are not considered in this paper since the necessary measurement frequency to detect them calls for sophisticated and costly hardware (Zeifman, 2012). Additionally, the first method is path sensitive, which means that appliance noise is different depending on where the appliance is installed in the building (Patel et al., 2007). Thus, if the appliance changes location, its outputted noise also changes. The second feature has a narrow focus and is specific to appliances with a switch mode power supply (SMPS). These SMPS are becoming more prevalent but still exclude several household appliances (Gupta et al., 2010). Pure voltage features in buildings have only been suggested in a study by Ruzzelli et al. (2010) are not commonly studied in research. Cox et al. (2007) suggested that voltage could be used as a feature in confined spaces like ships, however this environment is different to most building where voltage is actively managed in the network and kept within narrow tolerance limits.

6.2.4 Appliance Identification

Based on the features derived from power and current the process of identifying individual appliances can begin. In this section the current theoretical concepts of appliance matching and how they relate to presented features will be reviewed. A brief overview of the identification algorithms will be given last.

Hart (1992) defined three broad classes of appliances: binary state appliances, finite-state appliances and infinite-state appliances. The binary state appliances are, for example, normal lights that are simply turned on and off. Depending on the power levels between which these appliances switch, and on what other appliances are running in parallel, binary state appliances can be detected with all the features presented in Section 6.2.3.

The finite-state appliance can be thought of as a collection of binary state appliances that work consecutively or in parallel to produce the expected service. A washing machine is a good example that contains a binary state valve to let water into the machine, a binary state heating coil to warm the water, a motor capable at running at different binary states for the wash and spin cycles and finally a binary state pump to eject the used water. For an appliance identifier, all the different possible states that the appliance can assume must be known, which is similar to any binary state appliance. However, since many finite-state appliances follow a predetermined path through a given program, the search space for the next event can be made relatively narrow (Hart, 1992).

Infinite-state appliances lack distinct states and are therefore not detectable by the normal event filter that evaluates sudden changes. Even if an infinite-state appliance would be detected, it would be challenging to build a database with which to match the recognized state.

A fourth class of single state appliances was later added by Baranski and Voss (2003). These types of appliances use a constant power and are never switched off. For example, aquarium pumps, electrical mains connected clocks and smoke detectors would fall into this category. These single state appliances do not produce any events, as they do not change states. To recognize this kind of appliances other non-event based methods that use the aggregated data for the analysis like neural networks must be used (Srinivasan et al., 2006).

The process of identifying an appliance from a feature or a set of features follows either an optimization method or a pattern recognition method (Zoha et al., 2012). The optimization method uses a certain algorithm to minimize the difference of already-known appliance features and the measured ones. Commonly used optimization algorithms are the least square (Liang et al., 2010a; Beckel et al., 2012), dynamic programming (Baranski and Voss, 2004) and integer quadratic programming (Suzuki et al., 2008). The pattern recognition, in contrast, commonly use a training set of several combined appliances that is then compared against the measured aggregated signal of active appliances according to a specific approach. Currently used approaches are the Naïve Bayes (Berges et al., 2011; Chahine et al., 2011) k-nearest neighbor (Berges et al., 2011; Weiss et al., 2012) and artificial neural networks (ANN) (Liang et al., 2010b; Ruzzelli et al., 2010).

This study will focus on the optimization method of least squares as it lends itself well to event-based detection, which is scalable over several appliances. This last criterion is important for a solution that has to support several households and a large set of different appliance states (Zeifman, 2012).

6.2.5 User Operation Identification

User behavior and the impact of choosing certain modes of operation has been the research focus of some specific appliances. For example, McCalley and Midden (2002) measures washing machines to evaluate how goal setting can influence user behavior. Furthermore, energy advice for specific appliances like refrigerators, TVs and computers is evaluated in a serious game setting (Gamberini et al., 2012).

In these studies single outlet sensors have been used to measure the appliance's energy use, while the setting has been noted by the experimenter. There are discussions of how specific user behaviors could be extracted from aggregated household data, and even how user information can be anonymized from such algorithms (Efthymiou and Kalogridis, 2010). However, central disaggregation techniques are currently not reliably generalizable beyond laboratory environments (Liang et al., 2010a; Carrie Armel et al., 2013), which is why this study will provide user-level energy-use information through outlet sensors. This way, the appliance event is bypassed in order to focus on the appliance-setting disaggregation. When the central measurement technique is successful, this approach could save the need for sensitizing single objects.

Manually marking every chosen appliance settings, as the previous studies did, was considered to labor intensive and not be appropriate for a energy use information system for households. Therefore, key distinguishing patterns of an appliance's different operating modes were defined and then compared against the running appliance.

The appliances' settings were divided into length and height patterns. A length pattern is significant for appliances that regulates its operation by switching on its specific load for a specific amount of time, depending on the chosen setting. A washing machine, which is used to test the system in this study, is an example of an appliance with length distinguishable patterns. In contrast, the height pattern manifests itself in appliances that control their load at different levels. A hairdryer, for example, uses separate heating coils as voltage dividers to increase or decrease the voltage, and therefore have distinguishable load levels.

6.3 Experiment Model

This section will detail the implemented system to evaluate the impact of data sample-rate on disaggregation accuracy. First the dataset will be introduced followed by an explanation of the different treatments. Second, the experimental framework, which is based on the reviewed literature, will be presented.

6.3.1 Public Energy Use Dataset

So far, most published implementations of appliance recognition have been highly successful, with between 60 and 95% rate of detection (Zoha et al., 2012). However, this research has been done with specific datasets that have been exclusive to each research team. This has made it difficult to quantify the value of the additional hardware and software solutions between the studies (Zoha et al., 2012). Therefore, large datasets from detailed measurements have been made public in recent years. Most notably are the REDD and the BLUED dataset that provide open access and thus the ability to verify approaches between studies and research teams (Anderson et al., 2012; Kolter and Johnson, 2011).

The BLUED is an analogue version of the Reference Energy Disaggregation Dataset (REDD) that was made publically available by Kolter and Johnson (2011). The current and voltage measurements are gathered at the circuit panel in the single-family residential building. The current transformer sensors were attached to the two phases A and B, which supply the household with energy. These two power lines physically divide the measured appliances in two datasets that is also kept in the evaluation. The BLUED, as the name implies, is additionally labeled according to the appliances' events, which was made possible by plug-level and environmental sensors. This setup is shown in Figure 6.2 and allows for a ground truth for the event based analyses. Thirty-one different appliances are recorded (8 on phase A and 23 on phase B) and span a power usage between 15W (kitchen light) to 1600W (hair dryer). The data was accumulated over a week at a sample-rate of 12kHz and take up 307 GB of computer storage space when uncompressed. The appliances were labeled through a system of environmental sensors and circuit level meters. 95% of the appliances were thus labeled, while the final 5% were labeled as "unknown" (Anderson et al., 2012).

To keep the integrity of the data, the treatments of different sample-rate levels is scaled in integer steps by removing data points from the dataset. Thus, the specific treatments that will be in focus are 12kHz, 6kHz, 4kHz, and so on down to 100Hz. This means that the frequency will be more detailed as the time in between data points becomes longer. Other potential methods of creating a more even down sampling would be to use a statistical average over a number of data points and interpolate the desired data point. Interpolation was not chosen because of the artifacts that this would introduce.

6.3.2 Experimental Design

The steps of a disaggregation framework were introduced in the related works (Section 6.2) and are visualized in Figure 6.1. Based on this framework, the specific setup for this experiment will be detailed. In general, the focus has been on creating a system that can handle a variety of sample frequencies and the algorithmic solutions have been chosen to support an online disaggregation and embeddable hardware. However, it should be noted that the simulation in this study uses a normal PC to parse the BLUED and the appliance specific dataset offline to make the evaluation process more efficient.

As previously mentioned the step of the event detection will not be considered here. Anderson et al. (2012) results (which is also the group behind the BLUED) show 49-78% event detection accuracy when focusing on the number of appliances and 74-80% detection accuracy when focusing on the total detected power. The second step, of feature extraction, was based on the BLUED labeled event data.

The disaggregation accuracy was evaluated based on the different features presented in the related work (Section 6.2.3) and the specific sample-rate treatment.

The comparison between the measured features and the ones stored as known in a database was done using a 4-fold cross validation. This means that one part of the dataset was used as completely known labeled data used to create the load signature database (LSDB). The other three parts were treated as unknown trial data, which were matched against the LSDB. These two steps in the analysis are shown in Figure 6.3.

To match a feature against the known labeled features the least residue approach is used for features detailing brief events like the PQ, Imax or Ipeak, while the features based on a shape are evaluated through cross-correlation. The least residue method is an optimization algorithm that minimizes the squared difference (residual) between the features of the found event against known appliance features. The cross-correlation on the other hand compares two time-series by running one of them over the other. This way, an exact timing of the feature between the two time-series is not necessary for matching the waveform, as every phase shift in the given range will be evaluated (Figueiredo et al., 2011).

In the previous literature several metrics for determining disaggregation accuracy has been used. Hart (1992) and (Anderson et al., 2012), for example, used a disaggregation definition based on the total amount of power that was correctly found. Although this metric is beneficial for developing methods for extracting the appliances with the highest power usage, it might not correlate with the appliances that the user would be willing to operate in a different and more sustainable way. This study, therefore, follows Liang et al. (2010b) method, which was recommended in a review of current disaggregation methods by (Zeifman and Roth, 2011). This accuracy metric is calculated as the factor between the total number of disaggregated appliances and the total number of labeled events.

6.3.3 User Behavior Data

To evaluate certain energy behavior through appliance settings certain key section of the appliance load over a running cycle has to be extracted. Following the related studies that also focused on user energy behaviors (McCalley and Midden, 2002; Gamberini et al., 2012), outlet sensors were used for this analysis. This dataset is similar to the ground truth data in the BLUED project (Anderson et al., 2012), and could become an extension of the traditional appliance disaggregation procedure presented above.

The operation mode disaggregation algorithm is based on a simple form of the Kirchhoff's law state model that was envisioned by Hart (1992). The basic premise is that every appliance returns to the original state through a program of a finite number of states. This model allowed us to analyze when a program was over, where the appliances power load went back to the starting point, and to compare the duration and peaks of specific states of different programs to distinguish between them.

An appliance cycling counter process the return to the initial off state and a minimum predetermined length of time, as the mark of a finished cycle. In the case of the washing machine, a 5 Watt power use change from the initial steady-state was used as the threshold to register the start and stop of the appliance, in order not to start measuring at spurious moments. The cycle start and stop naturally also framed the total amount of energy used.

6.4 Results

6.4.1 Appliance Disaggregation

By benchmarking the results against a related study, which also has evaluated several different methods against each other, a sense of the implementation of the algorithms can be captured. It is natural that the results deviate as the underlying data of this study is based on the publicly available BLUED, while it previously has been done on exclusive research team data (Liang et al., 2010b). In Table 6.1 the results can be seen from this benchmark. This study uses a lower detection limit of 30W instead of Liang et al. (2010b) 100W. A comparable higher limit is therefore provided for direct comparison and shows that the current study's implementation of the feature matching is comparable to the results from previous literature.

The main goal of the study is to analyze the impact of data sample-rate on disaggregation accuracy. Here below are the results from scaling the sample-rate down. The average disaggregation accuracy of the single-family household's Phase A show

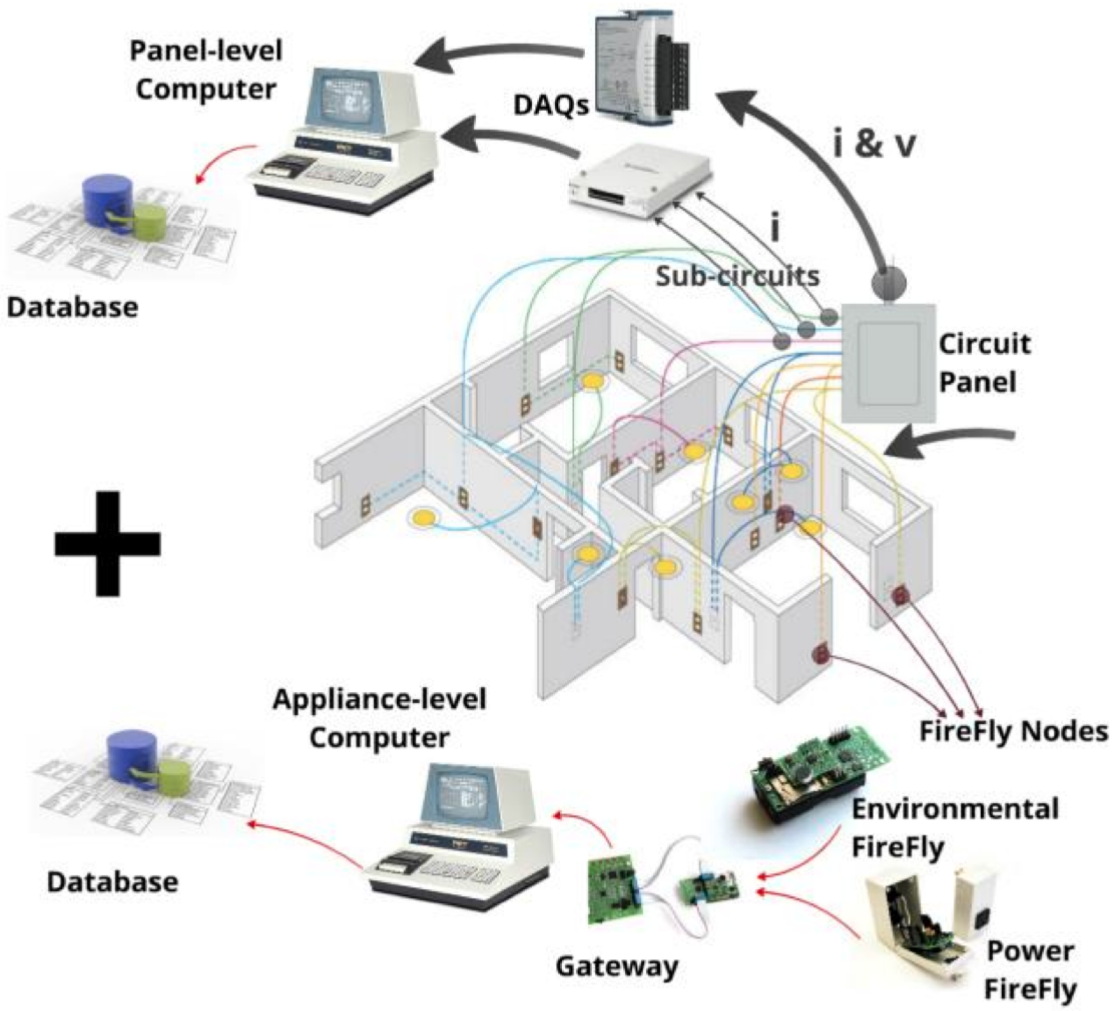

Figure 6.2: BLUED setup for data and ground truth collection (Anderson et al., 2012).

Table 6.1: Load disaggregation accuracy comparison between this study and the results of Liang et al. (2010b).

	Threshold	P >30 W		P >100 W		P >100 W
		Appliance disaggregation accuracy				
	Phase	A	B	A	B	Liang
Features	PQ	86.51%	60.82%	89.80%	81.52%	75-80%
	CW	72.57%	42.62%	88.16%	48.60%	80-85%
	IAW	61.67%	18.13%	80.62%	36.58%	75-80%
	IPW	69.66%	29.69%	83.68%	48.34%	75-80%
	H	81.71%	72.78%	90.03%	83.36%	65-70%

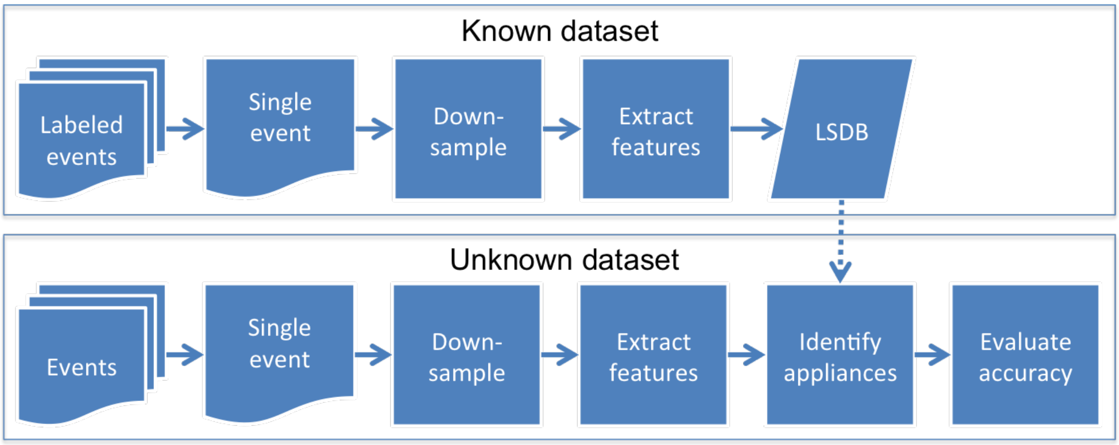

Figure 6.3: Four-fold cross-validation datasets, with one part known to build load signature database (LSDB) and three unknown parts for evaluating the appliance disaggregation accuracy.

a quite consistent detection-rate between 12kHz and 100Hz, which is visualized in Figure 6.4. The standard error over the different features and validation runs are between 0.2 and 3.3%.

The PQ feature provides the best results and stays above 85% for sample-rates above 1.5kHz. The harmonics feature show similar results for the higher sample-rates but proves to be more sensitive to a lower sample-rate as it drops from a 80% to a 70% detection accuracy when the sample-rate is changed from 4kHz to 3kHz. The features based on current (Ipeak and Imax) and waveforms (IAW, IPW, CW) all show no or only a slight decline in detection accuracy between 70% and 60%. Phase B is more sensitive to downscaling the data sample-rates than Phase A, which can be seen in Figure 6.5. It also has a higher standard error from 1.2 to 6.6% over the validation runs.

In Phase B, the harmonics feature showed the best disaggregation result for the highest sample-rate. This feature, however, drops off rapidly and falls below the PQ feature after 4 kHz. The PQ feature is more stable at appliance detection over different sample-rates. However, below 1.5kHz it also drops substantially. Similar to Phase A, the current (Imax and Ipeak) and waveform (STW, CW, IAW and IPW) features performed worse in terms of disaggregation accuracy and were also more stable over the data rate downscaling than the PQ and H features. Overall

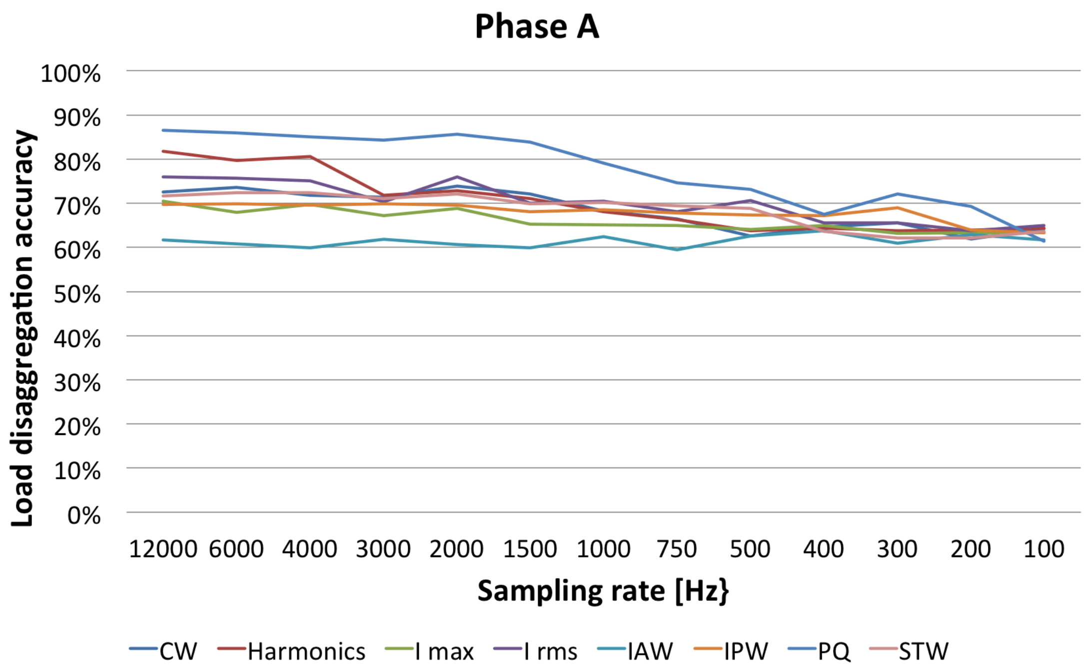

Figure 6.4: Disaggregation accuracy for tested features over different sampling rates on Phase A

these features drop at most 20% of detection accuracy (from 40% to 20% for CW) between 12kHz and 100Hz sample-rate.

6.4.2 User Operation Mode Disaggregation

In the operation mode analysis, the washing machine program load signature was determined to be the initial heating period in the beginning of the program, which is the initial energy spike period marked in Figure 6.6. This duration parameter could successfully sort all 55 of the 60 °C and 40 °C programs tested in this study.

6.5 Discussion

The disaggregation of individual appliances has developed much over the last few years. More precise measurement technology is used (Gupta et al., 2010), and several sophisticated methods have been applied (Liang et al., 2010b). However,

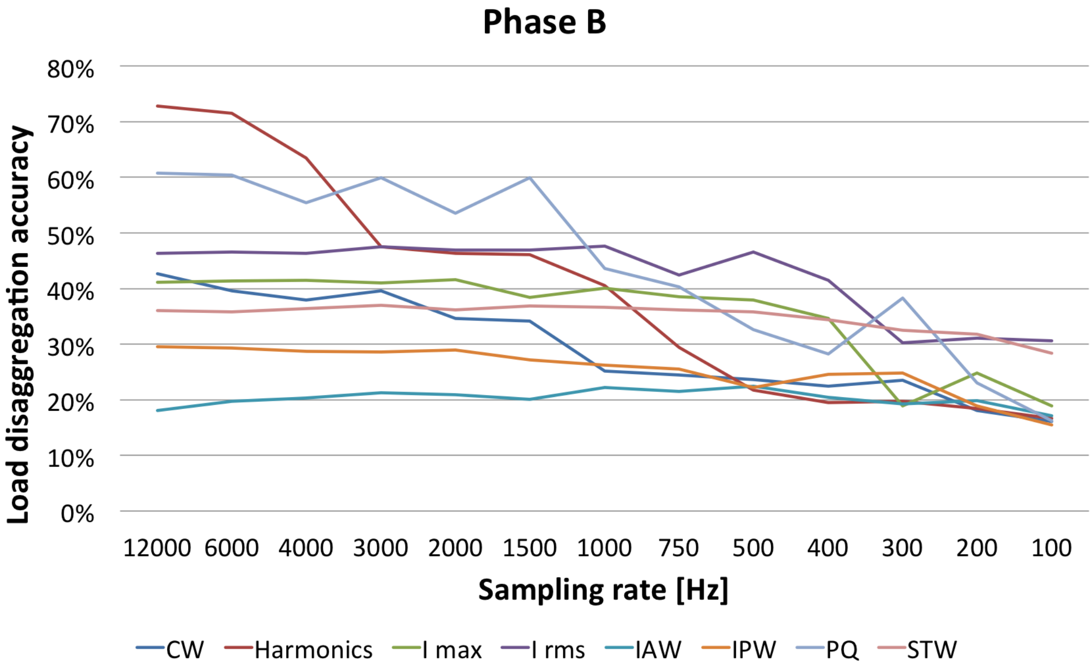

Figure 6.5: Disaggregation accuracy for tested features over different sampling rates on Phase B.

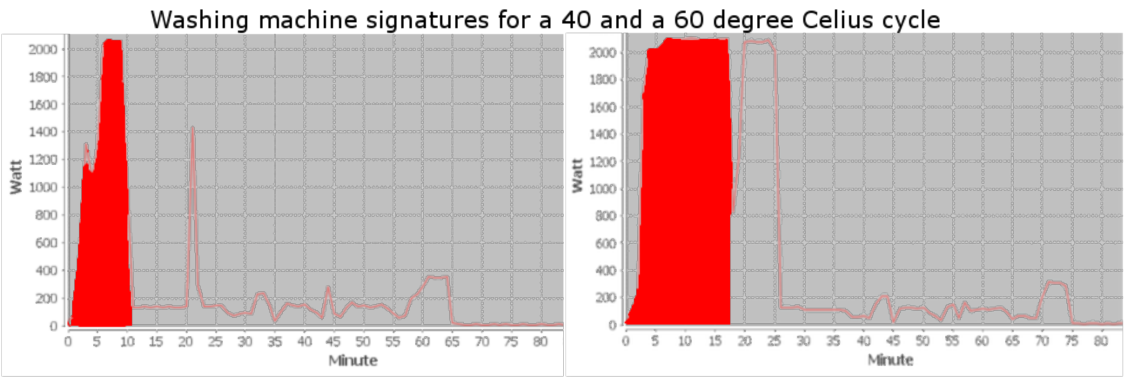

Figure 6.6: Disaggregated states for a washing machine, showing the heating power surges used for setting determination marked. With the area below the graph representing the energy used during a certain time-span

beyond the complexities of singling out appliances among similar ones, scaling these solutions for general use in households has not been evaluated.

This chapter presented a centralized and a decentralized system to disaggregate appliances and appliance's operating modes respectively. The disaggregation algorithms were trained and tested for accuracy, which proved trivial for decentralized operating mode identification, even at measurements of 1Hz. This result is directly related to the distributed sensors and their clear measurement signal. The influence of data sample-rate on disaggregation accuracy, which is one of the main parameters for choosing the necessary sensor hardware is, therefore, primarily discussed in this section with a focus on appliance disaggregation.

The detection accuracy of the appliance disaggregation was negatively affected by lowering the sample-rate, as expected. Phase A proved more robust than Phase B. The more stable impact over different sample-rates is due to the underlying appliances supplied by the two phases. First of all, the number of appliances are fewer, 8 compared to Phase B's 23 individual appliances. Second, one of Phase A's 8 appliances is dominant in terms of events (the refrigerator), which is also recognizable with all tested features.

Overall, the PQ and the harmonics features showed the best results, the higher the data sample-rates were, on both phases in the BLUE Dataset. These results are comparable to the ones found by (Liang et al., 2010b), who compared different features under a given sample-rate. Contrary to our findings however, Liang et al. (2010b) found that the waveform features (CW, IAW, IPW) were superior to the harmonics feature. This could be explained by the difference in the underlying data or the parameterization of the feature extraction algorithms. The difference in used data has been an inhibiting factor for comparing results between studies and this is why this study has elected to use the publicly available dataset provided by (Anderson et al., 2012).

An important result from this study is that the features proved to be robust over quite a range of down sampling. The PQ feature, for example, started its decline around 1.5kHz and the harmonics feature was negligibly affected by the lower sample-rate until 4kHz on Phase A and 6kHz on Phase B. This implies that disaggregation systems can be designed in the range between 1.5-6kHz and still be expected to accurately detect appliances comparable to other similar studies Liang et al. (2010b).

The most stable over all data samples were the STW and IAW features. Improving the accuracy of these features by using different parameters or alternative analysis methods than the cross-correlation used in this study would be interesting points for further research.

The user-level or appliance-operation-mode disaggregation was analyzed with a dataset of washing-machine measurements. The distinct heating period proved to be trivial to recognize. The result of this method will, however, depend on the appliance in use. For example, appliances with more granular settings like a stove top might have to be clustered into specific settings. Furthermore, fast switching appliances might be more difficult to define by length and height, as these can remain the same when measuring below the switching frequency. Measuring more often would thus handle this but it would negate the benefit of simpler hardware for decentralized sensing.

Research of information on user's operation of appliances is so far limited. The authors McCalley and Midden (2002), who explored energy behavior information pertaining to washing clothes, is perhaps the most well known exception (Abrahamse et al., 2005). Literature on operating mode disaggregation has only been found for commercial air condition installments (Luo et al., 2002). Compiling a database of appliances and their operating modes is, therefore, a first step to develop disaggregation of general appliances' operating modes further. In the next Chapter 7, such a database is exemplified.

Although the techniques for disaggregating appliances and their operating mode are similar, the appliance-level and user-level evaluations used different measurement technologies and data. The research presented here provide the first considerations for building a disaggregation system for appliance- and user-operation data. Future research could develop the algorithms for user-behavior disaggregation further based on aggregated household data and methods for sharing these recognizable features to limit the need for manual labeling.

Chapter 7

User-Centered Energy-Use Data

7.1 Problem Definition

The use of smart-meter technology greatly facilitates the collection and exchange of information about private households' energy usage. In principle, this information could be used to make energy users aware of their household's electricity use patterns and, thereby, induce more sustainable energy use choices (Mattle et al., 2011). Unfortunately, in practice unfamiliarity with the provided technical information, information overload, and lack of a means to interpret current electricity usage make it difficult for the end user to develop a concrete plan of action.

As discussed in Chapter 5, direct household level energy-use information is not enough to elicit a significant and persistent energy-saving behavior over a general population. Specific traits like few household inhabitants that focus on learning what appliance is most effective to target defined the participants who increased their energy usage the least in the experiment. More tailored information was concluded to be necessary to support users with different backgrounds and levels of knowledge and interest. After an evaluation of the challenges that the experimental participants have with adopting and using the current energy-use information design, Wallenborn et al. (2011), came to a similar conclusion, and recommend researchers to integrate the users in the design process. Furthermore, embracing a user-centric design has been suggested as critical for *"the migration of electricity users to the demand response world"* (Honebein et al., 2009, p. 39). However, to the best of the authors knowledge, no study has specifically designed energy-use information based on a user-centered paradigm.

The *objective* of this chapter is therefore to propose the design of an energy-use feedback system that provides personalized information on effective energy-use choices. The research presented in this chapter has been submitted for publication in the Business and Information Systems Engineering (BISE) journal (Dalén and Krämer, 2014b).

The design science methodology is applied to identify the requirements from the user's perspective for such a system and to develop the main design principles, which will be the basis for developing a first instantiation of the system (Hevner et al., 2004; Peffers et al., 2007). This methodology framework assists research of design theory that is prescriptive, practical and a basis for action (Baskerville and Pries-Heje, 2010). Additionally, the structure of design science also strengthens the potential for cumulative development of the artifact (Gregor and Jones, 2007).

The remainder of this chapter is structured as follows. First, design requirements for supporting sustainable action will be identified. Next, the design principles that detail the practical implementation of the requirements are described. Subsequently, the function of the energy-use information system will be demonstrated and evaluated. In closing, the artifact's implications, the study's limitations and future directions will be discussed.

7.2 Design Requirements

Developing a user-centered design is a well defined process, which involves the user throughout. The standardization of the processes has gone through multiple iterations and is currently covered under the international standard ISO 9241-210 "Human centered design for interactive systems" (ISO, 2010). In this study we have chosen to use the more common term "user-centered" to describe this process, however, we do recognize the need to integrate the wider population affected by the interactive system implicated by the "human" label used in the standard.

The six core principles of user-centric design are that: 1) the design is based upon an explicit understanding of users, tasks and environments, 2) users are involved throughout the process, 3) the design is driven and refined by user-centered evaluation, 4) the process is iterative, 5) the design addresses the whole user experience and 6) the design team includes multidisciplinary skills and perspectives. Design development is a continuous process and this study focuses on the first iterative step. More specifically, this study will use users' input from several months experience with energy use feedback technology and lead to a new set of design directives.

7.2.1 Design Process Entry Point

As explained in the fundamentals Chapter 3, Section 3.3.2, the entry point for initiating an design iteration of an IT artifact is given from the specific situation, available knowledge and available engineering utilities (Hevner and Chatterjee, 2010; Peffers et al., 2007).

To adhere to the principle of user-involvement in the design process and in order to get a complete picture of the user's situation and utility of energy use information, this study's requirements are based on experiments and qualitative interviews of users' experiences from having direct access to their energy use information. Thus, by basing this design approach on user's statements, which is fundamental in user-centered design, this study follows, what Peffers et al. (2007) call a *"client initiated project"*, where the design and development of a proof-of-concept instantiation is based on studies of users own interaction with previous artifact versions.

7.2.2 Client Initiated Requirements

One source of user inputs is the semi-structured interview with 13 participants who took part in the field experiment analyzing the use of web-based energy use information. The information was provided as direct and indirect power-use feedback over three months. The interview questions covered the topics of energy awareness, flexible energy-use and privacy. For a more detailed description of the interview process see Chapter 5, Section 5.3.4.

While the interview results were used specifically in comparison to the measured qualitative result in Chapter 5, this chapter will use the responses more broadly to define the design path for a user-centered energy-use information system. The final step of performing an interview analysis based on grounded theory is to compare the interview results in light of related literature and experiments to find commonalities and differences that can further the knowledge base (Strauss and Corbin, 1998). This action correlates well with the practice of developing design science requirements, which also should be applicable beyond the individual situation but rather to the whole class of artifacts (Gregor and Jones, 2007). In this study this amounts to studying related experiences with energy-use information and comparing qualitative findings to them.

The first major process the users in the experiment of household level energy-use information undertook was most often to learn or confirm the levels of energy usage of different appliances. For example, one user, in the household information level experiment, explained his evaluation process where he turned off the electricity supply to all rooms except one from the fusebox, and then tested the appliances of interest in this temporary laboratory. Furthermore, 5 from the 13 interviewed participants reported that the power use information had been used to support decisions to turn off unused appliances. Similar accounts of active user analysis of individual appliances were reported by Hargreaves et al. (2013) who interviewed 11 participants from UK who got access to energy-use information over 18 months through dedicated displays. Schwartz et al. (2013) also highlight the need for providing information that can be used to compare appliances based on interviews with users from a German living-lab study.

Requirement 1: Based on this evidence, providing access to individual appliances' energy-use is a requirement. This information should allow for learning what type of appliance provides the basis for the most effective intervention. The feedback should also support decisions on the most efficient appliance to change to within the same appliance type.

The requirement that directly follows is based on the importance of direct access to continuous energy use information. The high frequency, with which the power-use data was delivered (once every eight seconds) in the previously related smart-meter study (Chapter 5), allowed for a range of different information uses. The evaluation of individual appliance's characteristics, explained above, is only one example. Previous experiments also confirm that continuous provisioning of energy-use information is fundamental for a future information system (Ehrhardt-Martinez et al., 2010). It should be noted, however, that even though a heightened awareness and knowledge level has been achieved, direct feedback has not always been shown to lead to energy savings, even though users have gained an appreciable understanding of their energy usage (Darby et al., 2011; Hargreaves et al., 2013; Thuvander et al., 2012).

Requirement 2: Direct, highly granular energy-use information is a fundamental requirement for improving awareness, allowing for individual explorations and more advanced processing of information.

The second major usage pattern of energy-use feedback in the smart-meter study, besides the focus on individual appliances, was to evaluate the impact of changes in operating behavior. Five out of thirteen participants reported that they had tried to operate certain appliances with a large impact more efficiently, for example, by only running the washing machine when full or by cooking food collectively. Learning and changing operating behavior was defining factors for the three participants who increased their energy usage the least. Related to changing operating habits, many participants also commented that the impacts of certain changes were difficult to estimate with the current information and that certain tasks had to be performed regardless of time of day or electricity pricing. Hargreaves et al. (2013) similarly found that users would only go so far in changing their energy-use patterns since the impact was expected to be negligible or outside the preferred comfort zone. In contrast, Costanza et al. (2012), found that enabling activity level energy-use information did improve the understanding and could aid users' decisions.

Another change in behavior pattern was prompted when the users, in the household information level study, analyzed the power use when no electricity services where in active use. These participants were surprised to learn the significant power use of some of the stand-by enabled appliances, which influenced them to disconnect these appliances from the electricity source when not in use. Similar strategies (e.g. disconnecting and gathering appliances on power outlet strips) were also found in the living-lab study by Schwartz et al. (2013). Hargreaves et al. (2013) related a slightly different conclusion, their users reported that the base energy usage was quickly understood as the norm, and it is not reported whether it prompted any changes in appliance operation.

Requirement 3: To facilitate an analysis of how different modes of operating individual appliances impact the energy usage, the supplied information has to include the energy usage of different operating behaviors. This requires a more detailed information level than would be necessary for Requirement 1, which focuses on recognizing whole individual appliances, irrespective of the operation mode.

The feedback format, which was limited to current power load in the household level experiment, was also commented on by the participants. Specifically, more processing of the given information to provide it in well known units were suggested. Since electricity is purchased based on how much power is used over time, a few participants wanted to have the ability to evaluate how much energy had been used. Another user went further and reported that the monetary unit on the energy bill had a greater motivating impact to save energy than the current experimental feedback. These statements are also echoed by other experiments, where a more personal language was recommended to design future energy-use information systems (Thuvander et al., 2012; Schwartz et al., 2013).

Requirement 4: To present energy-use information in a motivating and understandable way, units that are commonly known to users have to be used.

In another experiment with dedicated displays, monetary feedback was however criticized for being *"unimpressive"* (Wallenborn et al., 2011, p. 151). The small short term gains that were displayed with the direct feedback did not motivate the users to change their energy use. A related informational issue, which was voiced in the interviews, was the uncertainty of the impact or worth of exchanging an appliance or operating behavior for another. For example, the debate whether LED lights would be profitable over the current compact fluorescent or in how many years a more efficient washing machine would pay back the investment over the old one, could not be satisfactorily answered with the provided information. Similar accounts were also evident in the study by Hargreaves et al. (2013), where the energy-use information also failed to provide a convincing argument of an action's value over a longer term.

Requirement 5: Further processing of the energy-use information should allow for long-term comparisons in order for users to make more substantiated decisions, both with respect to exchanging appliances as well as with respect to changing operating behavior.

In terms of display technology the participants, in this thesis' smart-meter experiment, reported contradictory needs. For example, on one side some users would prefer a dedicated display to remind them of the current energy use, while others saw a potential risk for conflict if this information was always visible. Both the potential reminder and the risk of conflict from using a dedicated display was also voiced by the participants in the studies by Wallenborn et al. (2011) and Hargreaves et al. (2013), who used different dedicated display technologies. On one hand the information was backgrounded and continued to remind users of their actions passively, however on the other hand, this reminder was, in some cases, experienced as "nagging", which could be further amplified by inhabitants. Due to the inconclusive outcome, whether the web-based or dedicated displays are preferred, these statements need further testing before becoming design requirements.

Another topic that also has contrary views are social comparisons and competitions. This information format is in some studies proposed to engage users and help motivate energy savings (Petersen et al., 2007), while other only reported negligible results (Abrahamse et al., 2005). Similar to the specific types of displays, the social comparative format also needs more analysis before potentially becoming a design requirement.

7.3 User-Centered Energy-Use Information

Before we proceed to the design of a user-centered energy-use information system we first survey current systems for energy-use or appliance information from a user-centered perspective. The related information systems will then provide a background for the design principles, which are necessary to fulfill the stated requirements.

7.3.1 Existing Systems

A number of different systems have already been devised in order to influence energy users to make more informed and efficient energy-use choices. In the following, we will review mass-media campaigns, home audits, energy-use feedback, appliance labeling and home automation, which are the main forms of energy-use information for end-users at present.

Media campaigns: Using traditional media outlets like newspapers, radio and TV emphasizes normative and cultural-cognitive pressures; however, few studies have found any significant impact from these campaigns. The lack of convenience of these forms of communication in terms of relevance and interactivity could be improved by introducing modern multimedia technology (Midden et al., 2007). However, the inherent monologue format of media campaigns make their information predominantly speak to the already concerned citizens (Staats et al., 1996). Although media campaigns can be successful in raising awareness about new developments, the targeted audience is generic, by design. This limits the sense of causal relationships between the forecast outcomes and the applicable information to the individuals who are already primed to pro-environmental behavior.

Labeling: Appliance labeling schemes have existed since the late seventies in order to promote new technologies and consumer adoption. Through a continuous adaptation process of the label methodology, the use of appliance energy labeling spread from Canada to the USA, and from there to Australia and Europe. But in spite of the labeling schemes' pervasiveness, no extensive change in purchase behavior could be determined in the studies made with this focus (Menanteau and Colombier, 1997). Although labeling provides a straightforward way of comparing appliances, the consumer is limited to evaluating the appliances that are available at the chosen retailers (Menanteau and Colombier, 1997). Further purchase options would have to be researched and found by the consumers themselves.

What is even more problematic is that the labels merely provide an indication of the general direction of energy efficiency that an appliance change would lead to. There is currently no follow up on how the appliance performs in a normal household throughout its operational life. Hence, the labeling lacks causal information which links different operating behaviors to the resulting energy consumption.

Home audits: Individual visits of energy efficiency experts to households are usually focused on evaluating large-scale house remodeling, such as the addition of insulation and changes of ventilation systems or window panes, due to the high salary costs involved. In theory, home audits could be highly user-centered as each review is done on the individual's premise. However, the audit information system is depended on the auditer's ability to supply a message that the user will understand. The costs of the system also prohibits revisits and further dialog. The success and effectiveness of these home audit programs is inconclusive (Magat et al., 1986). One reason could be the lack of understanding of the advised actions given in a home audit. For example, a review of home energy audit programs found that customers chose combinations of household energy improvement options that would lead to less than half of the energy cost savings possible (Magat et al., 1986).

Energy-use feedback: Feedback on energy usage is available to consumers in the form of either outlet sensors (distributed sensing), which report the energy usage of each appliance attached to the monitored electrical plug, or through household electricity meters. While both outlet sensors and whole household electricity meters make continuous information available, the breakdown of information is naturally more granular in the case of the outlet device. Possible interactivity with the available energy-use feedback devices is, however, limited (Hargreaves et al., 2013). Furthermore, structuring the data from these monitoring technologies to understand the impact of certain behaviors or appliances used requires significant effort on the part of the consumer (Fitzpatrick and Smith, 2009).

By not providing a course of action the information lack applicability and can cause frustration (Strengers, 2011). Although it is technically possible to process and give end-users customized energy-use information and advice that is generated from distributed energy sensing, it is so far a neglected opportunity (Marvin et al., 1999; Fitzpatrick and Smith, 2009).

Home automation: A related energy-use information system to energy-use feedback, which also include some form of automated operation control of appliances, is called home automation. This system and the flexible appliance control it provides has proven to help users cope with more complex information like frequently varying tariffs (Paetz et al., 2011). However, due to the high investment cost of such systems, its applicability has thus far been limited. Although a home automation technology arguably allows for highly customizable solutions, the system envisioned in this study focuses primarily on the information design. Including the aspects of more advanced energy control is therefore out of scope for this design iteration.

An overview of the relation between the existing energy-use information systems and the requirements can be seen in Table 7.1. First of all, mass-media campaigns and labeling provide information of differences within a certain appliance type segment. Both mass-media campaigns and labeling also present their information in well known units that the user can quickly parse either through a star ranking system or through monetary information. Second, comparison between appliance types and between operating behavior is possible with energy-use feedback through its continuous information flow. Third, home audits have the ability to provide user-centered energy-use information in almost all categories. Home audits can, therefore, function as an inspiration on how information can be processed and combined to suit individual users. The critical deficiency of the home audit system is its lack of a continuous information flow, which allows for up-to-date information that can follow situational and technological developments. The home automation system is not listed because the different instances vary considerably in terms of features and design.

Table 7.1: Comparison of current decision support frameworks for energy efficiency and appliance-specific design requirements.

Energy-Use Information System	**Requirements**					
	Appliance		*Behavior*	*Feedback Format*		
	Within type	Between types	Within type	Known units	Amortisation	Continuous
Mass-media campaigns	●	○	○	●	○	○
Labeling	●	○	○	●	○	○
Home audits	●	●	●	●	●	○
Energy-use feedback	○	●	●	○	○	●

7.3.2 Development Directions

The potential of implementing feedback on energy use through the use of ICT should not be underestimated. So far, it has only been compared to current systems. However, Green IS research show that these solutions can go well beyond the traditional decision support systems through technology that allows for accurate, frequent and useful information, and allow for new possibilities (Melville, 2010). Two critical strengths of ICT based energy information are its ability to combine information from separate but related sources and the ability to continuously update the given information in response to changes. These abilities will be incorporated with the existing energy information system's strengths, in terms of the user-centric requirements identified above.

The system proposed in this study does not claim to be superior to the other current energy information systems in each category. Each format has its particular strengths, for example, the face-to-face interaction of home audits is arguably more personal than other forms of communication and mass-media's ability to provide compelling and comprehensible explanations can be more powerful for introducing new services. However, the proposed system's ability to cover a broad range of key user-centric requirements, which go beyond the current instances, with the support of Green IS research is what drives its potential to be a transformative power in the energy system (Brocke et al., 2013).

7.4 Design Principles

In this section design principles are developed for the user-centered energy use information system that ensure that all of the identified requirements can be met. Our design is based on several sources of data, which, through processing, are combined to advise the individual energy user about the current appliance and operation behavior options. Figure 7.1 shows an overview of the planned energy use information system's modules and information flows.

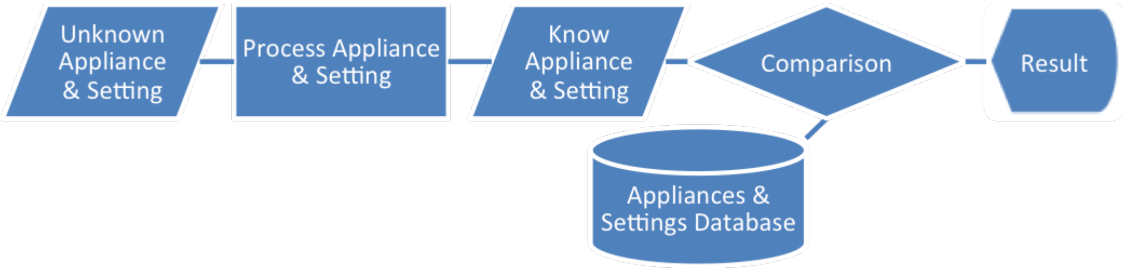

Figure 7.1: Schematic design of the proposed support system

7.4.1 Load Monitoring and State Differentiation

The necessary appliance-level information can either be measured at the supply of the energy flow network by distributed sensors or by disaggregating load data at a junction point (central sensing). In Chapter 6 both systems are presented and evaluated with regards to their accuracy of recognizing appliances over different sample rates and specific operational settings. It was found that the overall appliance recognition was stable down to 1.5 kHz measurements at 85% recognition accuracy. However, since Requirement 3 specifically calls for user-level information, reliable measurements and categorizations of individual appliances and their settings are necessary. Therefore, user-operational-level data was collected with plug level sensors in order to capture specific appliance operation features.

Design principle 1: Collect continuous energy-use data at the appliance level through electrical-outlet sensors.

As a consequence of gathering energy use data continuously, labeling the user operation manually for every new event is not considered an option. An automated post-processing was therefore designed and is described in Chapter 6, Section 6.3.3

Design principle 2: Provide information on appliances' states (e.g., "on" or "off") and operating settings (e.g., wash program) through automatic post-processing of the disaggregated energy use data.

7.4.2 Integration of Public Appliance Databases

An external source of data that provides information, which is both useful for users and specific to each appliance, are publicly available appliance databases. These databases, which entail energy use data for a wide range of appliances, are well suited to perform general comparisons within and between appliances types. For example, the appliance data supplied from these sources could be used in a purchase situation to compare some legacy devices and most new appliances available on the market. As users change their appliances or behavior this can be reflected in the feedback, as long as the individual appliances are available.

Unfortunately, these databases lack the depth necessary for the state and program matching that were explained above. "Wiki"-type projects for gathering and sharing information and appliance data like the "PowerPedia" (Weiss et al., 2012), are promising approaches in this regard and might eventually be the only feasible way of creating a global database of appliances and their operating modes. By introducing such a database, general appliance learning is reduced to the instance when the appliance is first measured. In order to demonstrate the value of appliance and behavior change, a prototype database with appliances that included the necessary appliance state and operational mode signatures information was designed in this study.

Design principle 3: The appliance exchange and behavior change information is based on comparisons between internal information (appliance-level measurements) and external information (appliance databases).

7.4.3 Providing Known Information Units

Traditionally, the energy efficiency of different appliances is compared through efficiency labels. However, this procedure lacks a relationship between the exchange of an appliance or change in operational behavior and the value thereof. Different operating modes of appliances make it complicated to compare their energy consumption in daily use directly. The labeling schemes handle this challenge by abstracting from the operating options by devising a normalized testing procedure that can be applied across a certain appliance type. However, this assumes a generalized behavior that might not be applicable to every use case. Moreover, the final grade is abstracted from the individual's operating habits, and does not retain any of the information about the cost of operating appliances.

In order to provide a causal relationship between actions and their effects on energy efficiency, this design study proposes to give decision support in the form of monetary savings. There are several advantages to this approach. First, in-line with Requirement 4 money is a known unit of value, and therefore has been reported to be accepted and well understood by a wide range of users (Fitzpatrick and Smith, 2009). Moreover, it does not require technical knowledge and thus, it will also help to reduce the cognitive burden on the decision maker.

Second, feedback in the form of monetary savings is likely to be more motivating than other forms of feedback. This argument is strengthened by Kamb et al. (1998), who found that a monetary incentive outperformed other units of equal value. In contrast, altruistic feedback, directed at the goodwill of users, show little or no effect on direct behavior change (McMakin et al., 2002; Ritchie and McDougall, 1985).

Third, a final benefit of using money as the feedback unit is the ability to compare different types of appliances between each other. This comparison is not possible with the current form of energy labels as they are tied to a type of appliance. As the overall energy saving is the main aim of exchanging a certain appliance, being able to compare different appliance types would allow the end consumer to exchange the appliance with the highest potential effect, irrespective of the appliance type.

Design principle 4: Provide energy use information support in the form of monetary savings, both in the short term and long term through amortisation calculations.

Moreover, according to the previously identified requirements, the decision support information shall be comprehensive and include not only the value of appliance alternatives, but also the value of behavioral alternatives.

Design principle 5: Provide support both with respect to appliance alternatives as well as with respect to appliance operation behavior alternatives.

7.5 Demonstration

The focus of this study lies in detailing the first iterative step in creating a user-centered energy use information system based on users' experiences with live energy use feedback. This demonstration will show a proof-of-concept of how the design principles can be executed. This study uses well known household appliances to exemplify the data gathering, processing and potential presentation approaches.

7.5.1 Setup

Following Principle 1, to allow for detailed appliance level information, a power outlet energy-logging device was used to gather data. The current and voltage is calculated at the relative high frequency of 1Hz by the micro controller into real power, in accordance with Principle 2. The information post-processing of operating modes, comparison and visualization was then handled by a custom-made Java program.

Altogether, a washing machine, a dryer and two refrigerators were equipped with the energy-logging device. These appliances were chosen based on their common occurrence and relative large energy consumption in households. To demonstrate the use of external information of alternative appliances, as predicated by Principle 3, related appliance energy use data for the comparison database was gathered from the publicly available appliance benchmark information portal EcoTopTen[1]. This website provides information on the most efficient household appliances currently available in the German market, and provides a relevant comparison for the analyzed appliances in this study.

7.5.2 Processing of Appliance and Operation Information

To provide the necessary data for calculating monetary savings and amortisation (Principle 4) and alternative appliances and operational choices (Principle 5), the logged energy consumption data was processed for three parameters; number of completed cycles (N^{cycle}), type of operation mode (X) and the appliance's energy consumption per cycle ($E^{appliance}$). By combining the number of cycles and type of operation modes (e.g., 60°C and 40°C for a washing machine) with the energy use measurements, the required variable of energy use of a certain operation behavior ($E^{behavior}$) for the evaluation were determined. This initial processing also provides the fundamental data that will be combined to serve the other requirements of

[1]http://www.ecotopten.de

appliance comparisons and the information presentation formatting. In the Results Section 6.4.2 in Chapter 6 this processing was found to sort all individually measured events into the correct setting.

In our demonstration of the applicability of the energy use information system, with respect to a behavioral change, this study will focus on the generation of feedback for the monitored washing machine. There are two reasons for this choice. First, washing machines' motors and heating blocks have a direct relation to the energy used. This is not necessarily true for all household appliances. Refrigerators, for example, are dependent on the temperature setting, ambient temperature, frequency and duration of door openings and closings and the thermal capacity of the content, which can only be read from the electricity consumption indirectly. Measuring in- and out-side temperatures could alleviate this specific problem but is out of the scope of this design demonstration. Second, the operation of washing machines allows for a straightforward evaluation of behavioral changes. An example of a possible alternative operating behavior was collected by reviewing research on washing machine and detergent technology. The recent improvements show insignificant differences in washing quality irrespective of the temperatures used (Rüdenauer et al., 2006). This result provides a host of possible operational changes that do not necessarily have an impact on the quality of the service provided. Finally, washing machines are also often targeted by current energy efficiency support campaigns and provide a good basis for evaluating the requirements in comparison to the reviewed energy-use information systems.

7.5.3 Attaching Value to Change

The parameters from the data processing were then combined to calculate the monetary value for both alternative appliances (Principle 4) and operating habits (Principle 5). The known unit and amortisation requirements were catered to by processing the energy usage from a unit of service in monetary terms to appreciate the yearly gain and the potential payback time of a full appliance change.

With respect to the replacement of an appliance, the annual savings potential ($M_y^{appliance}$) can be calculated by the difference between the amount of energy for the current appliance (E_{cur}) and the alternative appliance (E_{alt}) in performing a specific unit of service (e.g., one cycle or one hour of time), multiplied by the number of the units of service per year (N_y) and the current price of electric energy (C_{kWh}), as shown in Equation 7.1:

$$M_y^{appliance} = C_{kWh} \cdot (E_{cur}^{appliance} - E_{alt}^{appliance}) \cdot N_y \tag{7.1}$$

Some appliances involve a greater investment and are predicted to be running for several years. For these particular appliances, the interviewed users mentioned that it would be appropriate to take into consideration the period in which the appliance would earn back the investment. The simple payback method is the most common indicator for evaluating the profitability of investments or projects. Although it only provides a rough indication of the financial prospect, it is an estimate of how long the money will be tied up in an investment (White, 1993). Equation 7.2 details shows that the yearly amortisation (A_y) is the quotient from dividing the purchase cost of the alternative ($C_{alt}^{appliance}$) by the annual savings potential of the alternative appliance ($M_{alt}^{appliance}$) appliance is calculated.

$$A_y = \frac{C_{alt}^{appliance}}{M_y^{appliance}} = \frac{C_{alt}^{appliance}}{C_{kWh} \cdot (E_{cur}^{appliance} - E_{alt}^{appliance}) \cdot N_y} \tag{7.2}$$

This information further improves the basis for exchanging an appliance without demanding more sources of data to be collected. The problem formulation's simplicity also promotes a general understanding of the feedback given from the data.

Behavior change has statistically been more important for improving efficiency than what the rising cost of energy could accomplish in the same time (Frieden and Baker, 1983). Supporting the evaluation of the value of energy-use behavior is therefore important. It has also been shown that changing behaviors to improve efficiency is more successful for saving energy, both in the short and long term, and is easier to implement on a large scale than curtailing behaviors (Ritchie & McDougall, 1985).

Adjusting the temperature setting on a washing machine is an example of a potential operating behavior change. A switch from 60°C (140 F) to 40°C (104 F) has, for example, shown to have little impact on the resulting cleanliness of clothes but will impact the energy used of about 84kJ/kg water or about 230Wh energy for 10 liters of water (Rüdenauer et al., 2006)[2] .

The monetary value of changing operating behavior within the same appliance (i.e., the washing machine) was calculated for the current measured appliance and two newer alternatives. The number of cycles was normalized to be comparable to the yearly consumption base line of the analyzed appliances in the EcoTopTen database. The calculation to evaluate the behavior change in terms of monetary savings is shown in Equation 7.3. The annual savings ($M_y^{appliance}$) is calculated, similar to the savings from appliance change (Equation 7.1), by multiplying the current electricity price (C_{kWh}) with the number of yearly cycles (N_y) and the change in electricity consumption due to the behavior change ($E_{cur}^{behavior} - E_{alt}^{behavior}$). The factor X is a factor to vary the grade of operating behavior change between 0 and 100%. This variable was implemented to allow an evaluation of partial behavior changes.

$$M_y^{appliance} = C_{kWh} \cdot (E_{cur}^{behavior} - E_{alt}^{behavior}) \cdot N_y \cdot X \tag{7.3}$$

7.6 Results

This section will evaluate how well the design requirements are fulfilled in comparison to existing systems given the results from the demonstration. The evaluation will follow the "informed argument" approach, which means that the systems utility will be argued based on existing knowledge (Hevner et al., 2004). As explained in Chapter 3, Section 3.3.2, this evaluation was exemplified as appropriate when the design process is initiated by the client (Peffers et al., 2007), as is the case in this study.

7.6.1 Exchanging Appliances

Evaluations of appliance exchange often fail to use the current situation as the baseline for buying a new machine. Mass media campaigns and labeling rely on

[2]Specific heat capacity of water ($4186J/(kg \cdot deg)$) · temperature difference ($60deg - 40deg = 20deg$) · water mass ($10liter \approx 10kg$) · Wh per joule ($1/3600(Wh)/J$).

a statistical reference situations and operating habits, which lead to a purchase decision based on comparing newer models against each other on standardized grounds, which are more or less similar to the user's situation. Energy-use feedback and home audits, on the other hand, start with the current situation and then follow up with recommendations for the future. Home audits are, however, a snapshot process and lack support for continuous information updates. Energy-use feedback that is given direct to users has this continuous information component, however, the information leaves most of the processing to the individual user. The feedback implementation shown in this study uses the current situation as benchmark, similar to energy-use feedback and home audits.

The feedback design in this study uses the current situation as benchmark through continuous measurements, similar to energy-use feedback. Additionally, analogous to the expert advice of home audits, more in-depth information is added by the database of comparable appliances. *Requirement 1* is thus followed since the system provides the ability to compare appliances within a certain type.

In accordance with *Requirement 2*, the continuous information flow of the proposed system has a better chance to provide relevant information as technology develops and alternative services improves. It is important to note that this comparison only takes into consideration the electricity used, while other efficiency improvements - such as less water consumption for the washing machine - might also be an important reason to exchange appliances. Similar up-to-date information can also be delivered on specific appliances from external sources like mass-media campaigns and labeling schemes.

All information is given in a monetary unit as it is generally well known and understood in accordance with *Requirement 4*. Based on this monetary information *Requirement 5* is followed by providing support for long term decisions in the form of amortisation results.

When comparing the expected operational life of common household appliances (Gutberlet, 2008) to the calculated payback time (Table 7.2), it is immediately apparent that replacing an appliance is, in many cases, not cost effective. For example, there is only a marginal chance of getting a return on investment when exchanging the Medion refrigerator. This result is congruent with design *Requirement 5* by providing the necessary long term perspective to support informed decisions.

Table 7.2: Results showing the value of exchanging a washing machine, a dryer and a refrigerator to a current top alternative of comparable size and operational setting. The calculations are made based on the 2012 electricity price 0.26€/kWh.

Appliance Type	Current Appliance	Alternative Appliance	Yearly Savings	Payback Time	Operational Life
Washing Machine	Miele W 844	Bauknecht WA 634	18.6€	21.4 years	12.2 years
Clothes Dryer	Miele T 470	Bosch WTW841	37.4€	15.2 years	12.2 years
Refrigerator	Medion 5964	Bosch KT116	34.7€	22.5 years	14.6 years

7.6.2 Changing Operating Behavior

When a different behavior is possible without lowering the service quality or level of comfort, the change is limited only by a lack of information and knowledge (Hargreaves et al., 2013). This limitation is however real. New research and technological improvements have to compete with trained patterns and habits. In line with *Requirement 3*, the filtered operational settings from the continuous measurements (*Requirement 2*) are combined with the data of operation modes to give user behavior evaluation support. This customization is not possible with mass-media campaigns or labeling schemes, which rely on average user patterns. By using the current personal operational habits a more realistic energy and cost performance can be expected.

By providing a known unit of presentation (*Requirement 4*) for possible behavior changes, the proposed feedback implementation provides a basis for a more personalized way to compare the choice of different behaviors than mass-media campaigns, labeling and energy-use feedback. To support users in making long-term decisions, as *Requirement 5* stipulates, the monetary value of exchanging appliances was estimated in years through an amortisation algorithm. This support takes into consideration the user's current appliances and available alternatives. Of the reviewed

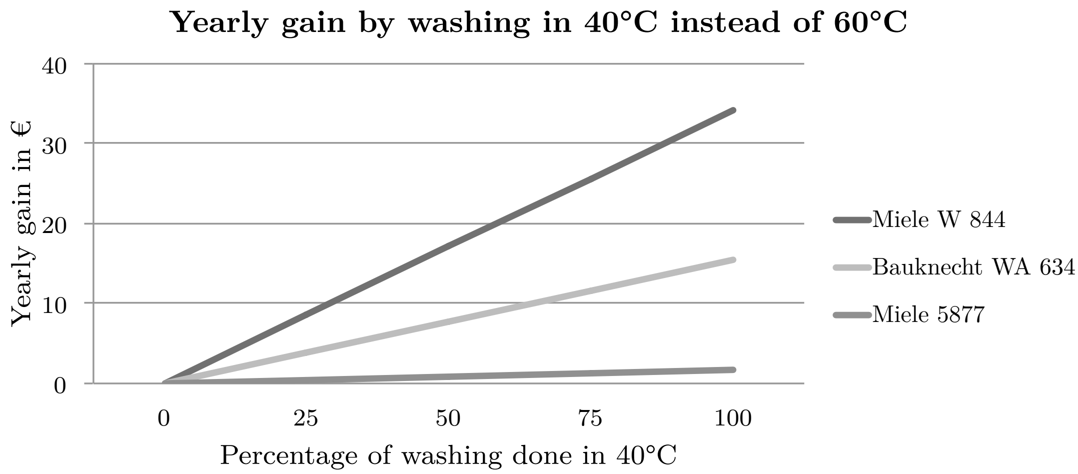

Figure 7.2: Results showing the value of changing the washing machine operating behavior from 60°C (140 F) to 40°C (104 F).

energy-use information systems, only home audits could make similar judgments as they also has access to the individual's appliances.

The operating behavior change feedback complements the appliance exchange by showing the user alternatives to improve the cost and energy efficiency when the replacement of an existing appliance is not appropriate. By comparing the two levels of feedback it is also possible to determine where the effort should be put. For example, a behavior change could be of equal or higher annual value than replacing a washing machine, depending on the amount of behavior change. Evidently, behavioral changes are generally more relevant for those appliances that are less energy efficient (as can be seen in 7.2). Thus, the effect of changing the behavior is directly dependent on what appliance is currently in use.

7.7 Discussion

Green IS and Energy Informatics research clearly has an interesting and complex challenge to tackle in improving information design for energy efficiency. This study analyzed how new energy-use feedback opportunities could be provided for private households by incorporating design science and information systems research with the recent developments in information technology. The objective of our design effort was to create a user-centered Green IS energy-use information system to promote energy efficient choices in private household. This objective is developed through involving users who are experienced with current instantiations of energy-use information systems. Based on qualitative interviews from three different experimental studies, design requirements were defined and operationalized through the design science methodology. This study thus contributes with requirements for designing solutions for the energy-use information class of systems as well as a defined artifact that lends itself well to continuous iterations. Key points of the design will be highlighted followed by a discussion of the study's limitations and potential future directions.

First, a key design principle of the proposed energy use information system is to provide both behavioral as well as appliance alternatives. In the example of the washing machine, the payback time for exchanging the older Miele W 844 for a more modern appliance was long. The same amount of savings were shown, through the behavior analysis, to be within reach by changing operating habits from washing in only 60°C to washing 75% of the time in 40°C. This result confirms that being able to compare appliances based on the same metrics is the key for efficient decisions both in terms of energy and economic aspects.

Second, in order to evaluate user's operational behavior continuous information is needed to evaluate appliance settings and operational habits. The proposed system's ability to grow along with how the user changes operation behavior, through its continual information flow also opens up new venues for feedback development. Different stages in life would have the possibility of being met by different messages. Thus, by analyzing users over years and categorizing them depending on actions, a form of curriculum could be developed for providing energy knowledge to user groups in steps over time. Furthermore, user rebound effects, where users balance their saved energy by using more elsewhere (Abrahamse et al., 2005), could also be targeted by following the user over time.

Third, by supporting a combination of household measurements and external information sophisticated reports, which were previously done by experts in home audits, can now be delivered continuously to the user. This study has shown great potential for implementing a shared database for appliance energy-use data. Current publicly available databases support decisions between new larger household appliances. However, it is still not possible to compare the current situation and most often not even the currently owned appliances with newer ones. With a shared appliance database with appliance setting differentiation, current appliances operated in specific ways can be used as the benchmark as was demonstrated in this study. The appliance efficiency could then be followed over its operational life, in contrast to the initial evaluation done today. Continuing the development of this database should, therefore, be prioritized to develop user-centered energy-use feedback further.

Fourth, to improve the information feedback, money was used as the comparison unit. Monetary units are easily comprehensible and in this vein it is straightforward to provide users with a direct link to the loss or gain that the energy efficiency choices lead to. Yearly gains and amortization period directly demonstrated the expected payback time in the current demonstration. The monetary unit also allowed for evaluating operating behavior changes and comparing appliances of different types.

In order to focus on a clear presentation of the conceptualization and design of the envisioned decision support system, we deliberately chose a simple approach to conduct the underlying economic evaluation of different appliances and behavioral alternatives. Obviously, several improvements are feasible here. For example, the payback period could incorporate an appropriate discount factor, and possibly also a forecast on future energy costs. Moreover, it would also be feasible to provide monetary information on the outcomes that can be achieved by replacing appliances and changing the usage behavior. In this context, it might be worthwhile to extend the comparison engine in our system by a collaborative-filtering based recommender system that could disseminate best practices of similar households.

Finally, it is important to remember that, *"feedback does not have to be complex to be effective"* (Darby, 2008, p. 506). There is a clear risk that the ability to add more information to a system might finally make it more complex. Therefore, before implementing this design for another quantitative and qualitative evaluation, the more fundamental concern of information overload should be evaluated.

Due to the potential informational richness of an energy information system based on Green IS, research exploring how to balance the information for accurate and timely decisions is becoming more important. Finding successful combinations of interaction technology, design, content and setting while closely monitoring the different channels of information is a challenging endeavor but also the most rewarding one by contributing cumulative knowledge that can lead to a more sustainable future. This proposed strict separation and measurement of information design and interface design is necessary to build cumulative knowledge of how the adoptions process is influenced (Bhattacherjee and Sanford, 2006). By combining the result from how an appropriate informational load should be designed with results from energy display design research (Anderson and White, 2009), another set of interface requirements can be tested in the next design iteration. This topic, of information overload and the design parameters to limit it, will therefore be evaluated in the next Chapter 8.

Chapter 8

User-Centered Information Overload

8.1 Problem Definition

Information and communications technology (ICT) is permeating ever greater aspects of our lives. With its help, access to information knows virtually no bounds. Sophisticated information services are now offered to each person with a portable computational device. This access can provide contextual details that previously were exclusively available to experts and managers. With the added detail, the specific value and cost of a decision can be more fully evaluated. For example, the upcoming informational services for end-users that is enabled through the Advanced Meter Infrastructure (AMI) in the energy domain, is one domain where users can now quantify their appliances and habits in greater detail and make more knowledgeable choices. Chapter 6 and 7 furthermore presented a user-centered energy-use information system, which provides highly processed information on an appliance and user operation level.

As explained in Chapter 2, Section 2.4.1, giving novices all this information might actually have the opposite effects to ones expected. Conclusions of smart-meter experiments also caution that an introduction of more information will not automatically lead to more efficient energy decisions by private households, but rather tend to make the decision task more complex (Darby, 2006). More recent research even suggests that a lack of spare time or technical interest is already enough to potentially preclude the potential efficiency gains that stem from the availability of more detailed energy-use information (Fischer, 2008).

End users have such a number of choices to improve the value or limit the cost incurred by appliances that they would stand a much better chance of handling the information overload if the energy-use feedback was preprocessed before being analyzed (Kempton and Layne, 1994). More precisely, a filtering of too detailed information is needed to guide decision makers in their use of increasingly complex information. “Else, we run the risk of the data being ignored, obviating many of the advantages of the smart grid” (Simmhan et al., 2011, p.6).

This chapter will focus on combinatorial problems where the user have to parse the available information and choose limited sets, and explore how much information detail should be presented to support effective decisions by novice end-users. To analyze this problem we have developed an experiment to test and compare different levels of information detail.

This research has been previously presented at the ECIS conference (Dalén et al., 2013) and is submitted to the BISE journal (Dalén and Krämer, 2014a). The remainder of this chapter is organized as follows. First decision support systems for end-users are contrasted against more traditional management decision support to understand the peculiarities with acquiring information on an individual end-user level. Second, a theoretical framework is introduced to help structure the evaluation of Decision Support Systems (DSS) for the end-user. The presented theoretical framework and previous related research provide the basis for the research hypotheses. Then, the design of an online decision task experiment is presented that enables us to study the impact of different levels of information load on decision quality. The study’s results are reported and then conclude with a discussion of the findings and avenues for future research.

8.2 Theoretical Framework

The research of decision performance and information processing has so far had little impact on the decision situation for novice end-users. For example, with respect to DSS for energy use, Allen and Janda (2006) found that adding features to an in-home energy display created confusion and could hinder the overall understanding. Strengers (2011, p.2141) further highlighted this problematic in her study on eco-feedback by stating: *"Worryingly, householders may react negatively if more 'bells and whistles' are added to in home displays to enhance and sustain consumption reductions."* No follow-up or suggested discourse to understand this problem in more detail was however given in the in-home energy display study. This is unfortunate since more research rigor could potentially help iteratively build systems that impact end-users to make more effective decisions (Melville, 2010), just as research have helped managers become more proficient at making complex decisions in their businesses (Shim et al., 2002).

Clearly, not every opportunity to maximize the amount of information given is appropriate. An example of how sparse information can successfully influence more conservative energy use is given by Ham and Midden (2010). In that study, an ambient light that changes color according to energy use provided a successful interface for goal oriented feedback. The effect of ambient light information compared to factual numbers had a stronger positive impact on users' behavior. However, it is still too early to generalize these findings to design DSS that focus on knowledge transfer in contrast to the goal-oriented processes it has been tested on.

The specific situation and its effect on the interaction can clearly not be dismissed. It is necessary to go further than gathering and supplying data to fully understand the novice end-user's relation to their information systems (Watson et al., 2012). Thus, by following an experimental approach we aim to contribute with a more detailed understanding of how a novice user is influenced by information diversity.

8.2.1 Expert vs. Novice Decision Support Systems

Previous research of information impact on decisions mainly focuses on managers and experts, who spend a considerable amount of time with a certain problem. Activities like strategic planning, management- and operational control lie at the foundation of these professional decision support systems (DSS) (Gorry and Scott Morton, 1971). Over time, DSS research has widened to support group decisions and has been customized for corporate executive decision making (Shim et al., 2002).

With the developments in information technology, measurement devices have spread to more peripheral areas and users. However, the metaphor of the novice end-user as a micro-resource manager has been criticized as this perspective could potentially mask certain practices in a normal household (Strengers, 2011). Besides the situational differences, another crucial difference between the target audiences and therefore the informational design lie in the characteristics of an expert and a novice end-user.

First of all, it has been shown that parsing information can be done in larger batches with improved retention rates by experts (Mackay et al., 1992). Novice end-users in contrast, proved to commonly miss the "larger picture" and instead focus on specific problem details. Second, Schenk et al. (1998) found evidence for a general lack of domain knowledge with novice end-users, which led to a less systematic course for problem solving. Third, by reducing the required processing of information, novice end-users have been shown to acquire the same amount of information as experts, however, by failing to map the informational pieces to one another decision accuracy still trailed behind the experts (Perry et al., 2012). Fourth, even though training has been proven to increase the ability to use and accept DSS (Nelson and Cheney, 1987), extensive training sessions for end-users in the energy domain is improbable considering the cost indifference and relative high premium on comfort that individuals are willing to carry (Hirst, 1980; Pierce et al., 2010). Research on expert DSS use, however, considers training to be a given prerequisite for implementing any information support system (Green and Hughes, 1986).

The main points that have to be taken into consideration when designing a DSS system for end-users instead of experts are summarized in Table 8.1. These distinct differences lead to the question; how should information be designed to support novice end-users to parse information? To make this question testable the next subsection will present a theoretical framework according to which DSS can be evaluated.

Table 8.1: Key differences between expert and novice end-users constraints for designing a DSS

DSS Considerations vs. Target Audience	Expert	Novice
Parsing	Batch	Piecewise
Domain knowledge	High	Low
Understanding	Conceptual	Individual details
Incentives	Salary	Mixed

8.2.2 Cost-Benefit Decision Performance Framework

To structure the research and guide the theory development, we use the cost-benefit theoretical framework (Beach and Mitchell, 1978). This framework "... is the most frequently used framework for explaining contingent decision behavior" (Payne et al., 1993, p.100), and therefore commonly used as the basis for explorations of DSS (Benbasat et al., 1987; Todd and Benbasat, 1994; Wang and Benbasat, 2009). The trade-off between minimizing the cost of a decision and maximizing the benefit thereof, influence how users interact with the DSS (Todd and Benbasat, 2000). The basic premise for the framework is that each decision strategy balances the associated cost of a decision strategy against the potential value of the outcome. Where cost can be interpreted as the effort of information acquisition and processing while the benefit is the decision accuracy and the speed with which the decision is taken. The task complexity is interpreted in the cost-benefit framework as the time pressure, number of alternatives or dimensions with which the information is presented (Payne, 1982).

In a combinatorial problem situation, where only a limited subset can be chosen and the information is diverse and has multiple dimensions, it is virtually impossible to find an optimal solution. However, by altering the number of dimensions for a given set of alternatives at a certain allotted time the relation between the independent variable and the dependent decision behavior can be evaluated.

Following this theoretical model, the goal of this study is to evaluate different informational diversity levels to understand at which level users are able to choose a strategy that allows them to reach an accurate decision outcome in as little time and with as little effort as possible, under the situational constraints of a novice end-user. This study aims to test what effects certain detailed changes to the information diversity will have on the decision strategy. In the following section related literature is presented to form the hypotheses to reach these goals.

8.3 Hypotheses Development

In this section results of information presentation on decision outcome are presented from literature on decision experiments. These findings are subsequently used to operationalize the theoretical framework. This will provide hypotheses and measurement methods to evaluate the informational details that balance the cost and benefit of decisions.

8.3.1 Decision Accuracy

Decision accuracy is not directly correlated with the amount and detail of information provided to the decision maker. Although access to decision support tools has been found to lower the effort of processing a specific amount of information, it could not be shown that it provided the capability to process more information (Todd and Benbasat, 1991). Novice end-users have shown a similar ability to acquire information to their expert peers when the information needed less processing. However, this ability did not improve the novices' solution accuracy (Perry et al., 2012). These results are contrary to what (Iselin, 1988) found, where inexperienced users managed to reach a level of accuracy on par with their more experienced seniors for more diverse and complex tasks. It should be noted that the inexperienced users in the latter experiment are not novices, as defined in this study, as they already had domain knowledge but simply lacked the years of training.

In contrast, minimizing the complexity, by combining and summarizing of information for example, is recommended to make the decision process clearer (Hwang and Lin, 1999). Experiments have found that not just the ability to parse information is limited but that having too many options can be confusing and lead to lower decision confidence than if fewer options had been given. For example, subjects were found to be more burdened and overloaded by choice in treatments with more than 20 different choices compared to treatments with only 6 choices in such disparate areas as selecting jams or essay topics (Iyengar and Lepper, 2000).

There is also a lower boundary for information detail for making accurate decisions. The current monthly or yearly feedback on energy consumption is a glaring example of the effects of restricting access to information detail. In fact, a three year study of the impact of energy bill information in Norway found that, increasing the billing frequency, including historic energy use comparisons or providing energy conservation tips led to a 10% average saving of energy (Wilhite and Ling, 1995).

To summarize, previous research has found an inverse U-shaped relationship between decision accuracy and information diversity. Performance first improves with more information until a certain level when it starts to drop off as more information is introduced (Hwang and Lin, 1999). Tipping points were found even for small increases in information dimensions, where more information load incurred decisions with lower quality (Hwang and Lin, 1999). Hence, we hypothesize that a similar, inverted U-shaped tendency exists between the information diversity provided and the solution accuracy. A common metric to measure the solution accuracy is to create a decision situation where a fixed ideal solution point can be compared against the participant's solution.

Hypothesis 1 (H1): The accuracy of the end-user's decisions is highest for intermediate levels of information task complexity.

8.3.2 Decision Time

Solution accuracy alone is not sufficient to judge the quality of a decision-making process (Vessey, 1994). The ability to consistently distinguish a good alternative, even in a limited time frame, is also an important criterion for feedback interfaces for end-users. Previous research is, however, inconclusive on the effects of lowering the task complexity through graphical representation on solution time.

One study comparing tabular and graphical information representation found, as predicted, that the graphical representation lowered the time to complete the given task when sufficient time was given (Benbasat and Dexter, 1986). A similar correlation was also found between more complex tasks with a higher diversity and longer solution times (Iselin, 1988). The influence of experience could however only be partially confirmed in the case of more complex tasks. A more surprising result was found by Speier and Morris (2003) where text-based presentation lead to improved solution times over its normally less complex graphical counterpart. The authors, however, argue that the visual novelty might have led to more exploration and thus longer solution times.

In light of this evidence's bias towards decreased solution time for decreased information diversity, we also hypothesize that the participants will find the solution with greater speed when the information diversity is lower. The solution time provide a natural metric of the timely progression of individual decisions towards an accumulated solution.

Hypothesis 2 (H2): The end-user decisions are made faster for more aggregated forms of information presentation.

8.3.3 Decision Processing Strategy

In addition to improving solution time, understandable and actionable information is predicted to lead to strategies that produce a higher decision benefit by improving its accuracy while minimizing the decision time and amount of processing effort needed (Todd and Benbasat, 1994).

The theory of *"Reflection-in-action"* proposed by Schön (1983) provides insight into the effects of how trial-and-error versus reflection can influence the effort to reach a solution. The fundamental premise of reflection is a behavior where the reasons of the outcome is analyzed and allowed to change the underlying behavior. By setting up an experiment with a decision process that emphasized reflection through increased feedback, computer science students improved their average decision accuracy and at the same time used fewer trials to accomplish the final solution (Edwards, 2004). Thus, by taking into account a trial-and-error behavior, based on how many individual decisions are taken to reach a solution, the effort to process and understand the information can be measured.

To mirror the result from changing the feedback environment on decision effort we hypothesize that the amount of tries to settle on a solution will be the inverse to the accuracy trajectory. Information that manages to balance task complexity will therefore allow for more reflection and lead to fewer number of individual decisions. The number of individual decisions taken to reach the solution will measure the effort of information acquisition and processing.

Hypothesis 3 (H3): A higher number of decision steps to reach a solution is necessary when the information is either too sparse or too detailed.

8.3.4 Learning

Understanding how learning influences decision benefits and costs is an important parameter in evaluating information interaction (Iselin, 1988). The experimental subject's familiarity with an experimental setting has shown to have a clear positive impact on the decision quality, both for structured (Iselin, 1988) as well as for unstructured tasks (Iselin, 1990). Similar conclusions are well established in the expert systems domain where the experience with a number of different problems and the learning of a specific problem is what defines experts from novice end-users (Hayes-Roth et al., 1983).

By running the experiment over multiple rounds the impact of the increased experience can be evaluated. Based on the previous findings, we expect that users will acquire experience with the problem and become more proficient in using the given information with every repetition at every task complexity level.

Hypothesis 4 (H4): At every level of information detail, decision accuracy increases while decision time and the number of decisions decreases over the number of rounds of the decision task.

8.4 Experimental Design

Field studies of how information and technology can alter behavior are generally criticized for using too many mixed sources of change (Fischer, 2008). A controlled setting is needed to learn how specific facets of information influence decisions. Laboratory research has long been requested to clarify the relation between applied causes and their effects in feedback for energy efficiency (Abrahamse et al., 2005). However, laboratories have two distinct drawbacks for a decision experiment. The first is that the participants are normally chosen from a common subject pool (e.g. students) and second, that they take place in a foreign environment. As the situational effects on feedback interfaces are important (Fogg, 2003), a laboratory experiment was dismissed.

Instead, an online experiment was developed to study the effects of changing information diversity for a diverse set of end-users in a familiar environment to them. To get a fundamental sense of how end-users are affected by information diversity it was decided that an abstract testing environment should be built. The problem abstraction allows for testing the hypotheses without priming the subjects and distorting the result. The devised decision framework provides internal validity (control) and allows a general set of participants to interact with the given structure.

8.4.1 Subjects and Design

The experiment design was based on the knapsack problem. This classic problem, where a hiker has to make careful decisions about the items to bring in a constrained knapsack, provide a clear decision goal of maximizing the value of the chosen items, while being restricted by a limited volume for the chosen items. Furthermore, the maturity of the problem provides a large set of algorithmic explorations. In this experiment, this meant that an optimal solution could be found in pseudo-polynomial time through reduction and sorting (Martello and Toth, 1990). This solution, in turn, supplied a fixed point for the comparison of participants and the hypotheses evaluation.

Changing the amount of information that the participant has about the optional items alters the informational diversity. We used the number of dimensions in which the information was represented to alter the task complexity in the experiment interface. In effect, the information of the value of each choice was gathered in more or fewer bins while the underlying choice set was the same for every participant and experimental round irrespective of treatment. Each bin value was represented by a shade of gray in the interface, while each item, representing a potential choice, was assigned a specific cost and value. A range of values was represented under each bin shade while the cost was represented by the item width in the interface; see Figure 8.1 and 8.2 for graphical examples of the abstract experimental interface.

The decision framework was published on Amazon's Mechanical Turk (MTurk) platform. This platform and its participants (or workers) have proven to produce

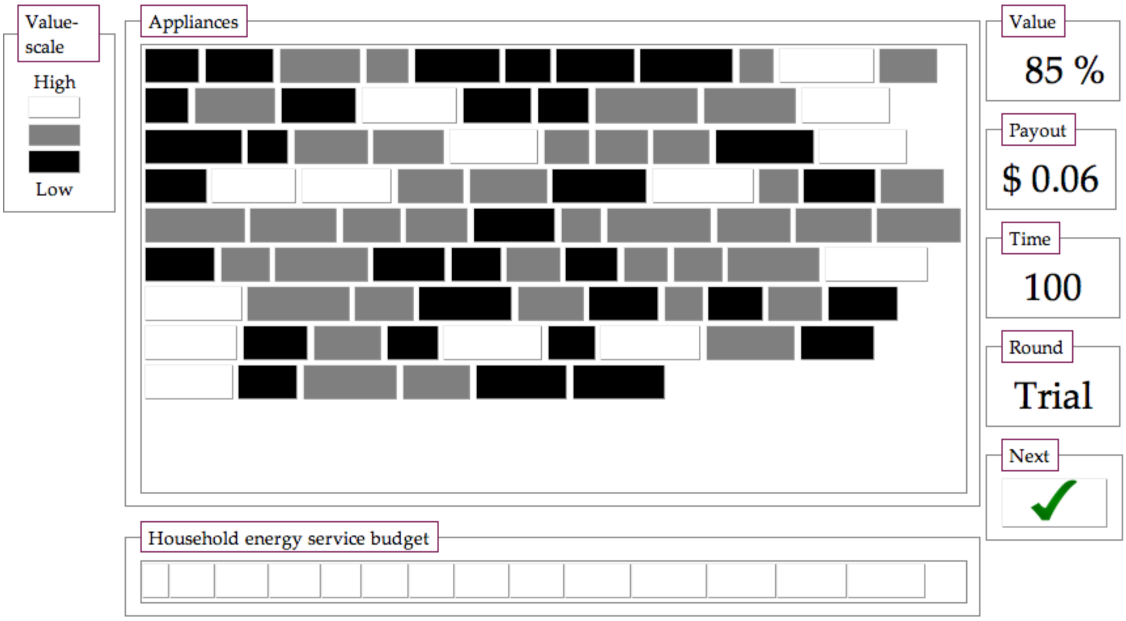

Figure 8.1: Interface for the knapsack problem showing the three-shades treatment condition.

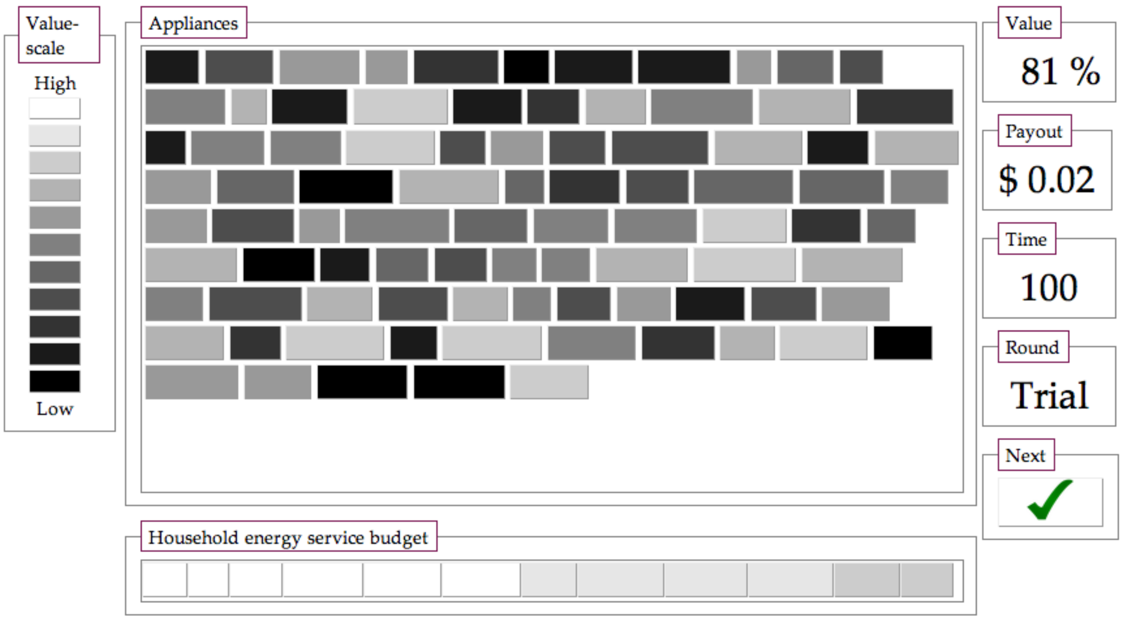

Figure 8.2: Interface for the knapsack problem showing the eleven-shades treatment condition.

reliable results that are consistent with what has been found in more traditional experimental settings (Goodman et al., 2013). MTurk participants have also been shown to better represent the overall population than in university campus experimental facilities (Berinsky et al., 2012). According to the explicit characteristics of domain knowledge and external incentives, as defined in the theoretical background, the MTurk workers fit the novice end-user well. Both the workers and novice end-users have no training with the current problem and are only given a small incremental incentive to interact with the interface. With the support of the MTurk platform, reliable and reproducible results can be achieved on the basis of a heterogeneous and representative sample that adhere to the study's predetermined requirements.

No country was restricted from participating in the online experiment. The main countries represented, based on the participant's IP address, were the USA and India, which is consistent with the findings of Ipeirotis (2010). Over the course of the study 400 workers were randomly assigned to the between and within subject experimental design at the time of registration.

It should be noted that while our results are meant to inform IS research on the interaction aspect of end-users and informational interfaces, our experimental interface is not considered to be a design effort in itself. The experimental interface deliberately abstracts from any specific choice context in order to eliminate distortions from priming subjects. Following the general character of the combinatorial problem situation, we have chosen an unframed design in order to gather results that are free of domain bias. Furthermore, in order not to push respondents towards a behavioral solution heuristic there is no explicit ordering of the items (e.g. along the cost or the value dimension).

8.4.2 Treatments

The treatments of the different information diversity levels were evenly split between the area of information overload at fifteen colors (Iselin 1990) and 3 colors that were assumed to signify a low level of information load. All together there were four treatments with three (*Color-3*), seven (*Color-7*), eleven (*Color-11*) and fifteen (*Color-15*) different colors.

The independent treatment variable is the information diversity value represented as the item shade between black and white. Therefore, a treatment with higher information diversity had more and smaller deltas for each bin of values than the treatment with a lower diversity. For example, 15 colors made the discrete steps between each shade smaller, and more closely represent the information about the underlying values, than in the 3, 7 or 11 color treatments. The color-coding allowed us to seamlessly manipulate the level of information diversity available to the subjects. By randomly assigning the treatment's conditions to the subjects, the study thus explores how the grouping of information impacts the users ability to solve the knapsack problem.

The box width, representing the choice cost, and a choice value was randomly drawn from a uniform distribution for every round in the same way for all the treatments. Multiple distributions were compared in terms of the difference between a greedy solution and an optimal one. The greedy solution followed by filling the knapsack with items in the order of the value/cost ratio until full. The trial and treatment rounds were chosen from instances with large differences between the greedy and optimal solution to increase the problem difficulty.

8.4.3 Procedure

Every user was greeted with an introduction to the knapsack problem, the goal and the payout structure. A trial round gave the participants the opportunity to familiarize themselves with the interface. During the trial round the users could play the game and learn more about the interface by hovering the pointer over specific parts of the interface and receiving information about the interface details without time restrictions. The legend of the item's bin values for the current treatment was shown to the left of the screen and followed a white to black color scale.

For each normal round a maximum time frame of 100 seconds was given. The remaining time was shown to the right in the interface. If the participants were satisfied with the solution before the time was up they could always click the "Next" button and move directly to the next round. The other alternative of letting participants wait for the time to run out before moving to the next round was pretested but was a common source of complaint (Dalén et al., 2013).

Figure 8.1 and 8.2 shows the decision task framework of the treatments with 3 and 11 levels of information diversity. The current value of chosen boxes is showed as a percentage as well as the current payout. By "clicking" the box again in the knapsack returns it to the top box frame with the available choice items. Each decision is time stamped along with the incremental change to the overall solution accuracy.

Every treatment was played over three rounds and the subjects were paid a small fixed payment with a performance based variable component. A subject's task was to select an item portfolio, which maximizes the sum of the items' values, given a size constraint and a limited time frame. To incentivize more participants to solve the problem earnestly, the starting payout was kept to the minimal 1 cent while every incremental percentage improvement above 80% was rewarded with an additional cent. Participants that only "clicked through" with a final result below 80% of a maximum 100% were excluded from the evaluation, as values below 80% were never encountered in the pretests of the experiment that produced a mean result of 93% (Dalén et al., 2013).

The level of decision accuracy was compared to the calculated optimum at two points during the game. Since the knapsack could be filled and emptied multiple times over each treatment it was determined to also evaluate the best overall solution of that round, as it did not necessarily coincide with the final measurement point at the end of the round. The best overall and final end points were also used to record the elapsed time to reach each solution and to test the within-subject development over the rounds. From the recording of the decisions made by the participants the number of decisions were also aggregated for the final solution.

8.5 Results

A statistical summary of the results from the experiment is shown in Table 8.2. *Color-7*, *11* and *15* are the treatment dummy variables, i.e. we test for differences in relation to the *Color-3* treatment. *LogRound* specifies the logarithmic of the round played and controls for the learning effect. The random effects in this model are assumed to be independently normally distributed with a variance estimated

Table 8.2: Mixed-effect regressions of treatments with dummy variables

Treatments	Best Result	Final Result	Best Time	Final Time	NumberOf Decisions
Color-7	-0.463	-0.633	1.378	1.617	-2.299
	(0.515)	(0.396)	(0.654)	(0.816)	(0.100)
Color-11	-0.185	-0.302	0.804	1.066	-3.513*
	(0.804)	(0.698)	(0.802)	(0.883)	(0.016)
Color-15	-2.008**	-2.256**	6.963*	19.71**	-0.163
	(0.006)	(0.003)	(0.028)	(0.006)	(0.910)
LogRound	1.753***	1.767***	-8.230***	-14.100**	1.717**
	(0.000)	(0.000)	(0.000)	(0.006)	(0.005)
N	573	573	573	573	573
Wald X^2	39.48	35.88	31.25	17.27	15.69
Log Likelihood	-1635.65	-1678.15	-2535.76	-3133.68	-2007.93
Note:				*p<0.1; **p<0.05; ***p<0.01	

through our regression.

8.5.1 Hypotheses Testing

Hypothesis H1 tested if participants would perform better over a mid-range of information granularity than at the outer extremes of less or more information diversity, and thus display an inverted U-shaped relationship of the accuracy. These results were evaluated both with respect to the overall best and final end solution points that were extracted. In order to control for the dependence in observations stemming from the same respondent, a two-level linear mixed-effect model was used to test for differences in the results between treatments.

The statistical fit of the best and final solutions only partially confirms the first hypotheses. There is a significant drop in accuracy only with respect to the *Color-15* treatment. That is, the subjects were able to reach a more accurate solution with *3*, *7* and *11* levels of information diversity. However we do not find statistically significant differences in accuracy between the *Color-3* and the *Color-7* or *Color-11* treatments. The benefit of decision accuracy would thus favor the three treatments with larger bin sizes and less information diversity.

Hypothesis H2 was aimed at testing whether participants reached a solution in less time with a mid range of information granularity, rather than at the extremes of the scale. Similar to the results on decision accuracy (H1), this study finds a significant effect only with respect to the *Color-15* treatment. Compared to the *Color-3* treatment, participants in the *Color-15* treatment needed significantly more time to reach a solution. This holds, both with respect to the best as well as the final solution. However, the hypothesized U-shaped relationship could not be confirmed with satisfying statistical rigor. All other treatments used a similar amount of time over the rounds. The benefit of decision speed would be similar to the accuracy, thus favor an interface with less than *15* information dimensions.

Counting the number of individual choices taken by the users then tested how many decisions were taken under the different information diversities. The statistical result from the analysis of the number of decisions over the whole round confirms H3. While *Color-15* and *Color-7* are not significantly different to *Color-3* with respect to the number of decisions, we do find that participants in the *Color-11* treatment use significantly fewer decisions to reach the final solution. In terms of the effort for information acquisition and processing the interface with *11* information dimensions enabled the strategy of least effort.

The fourth hypothesis H4 was aimed at testing whether within-subject effects were present over the different rounds. These effects are present when users become more familiar with a certain experiment. The effect for the best and final solution accuracy and number of decisions both followed a statistically significant, logarithmically improving, development over the experiment rounds. The statistical results of the solution time for the best and final points showed a significant decrease, while the numbers of decisions increased slightly over the rounds.

8.6 Discussion

In this study, users' ability to reach an accurate solution given different information diversity levels has been evaluated in a novel experimental environment. In particular, a decision task environment based on the knapsack problem was employed to mimic the complexity inherent to making effective decisions in a combinatorial problem setting with access to detailed information.

Based on the theory of information load (Hwang and Lin, 1999) and previous experiments in this field, the main hypothesis was that a midway between having too little information and being overloaded with information exists where users can make more accurate and faster decisions. The inverse U-shaped relationship for decision accuracy, the decreasing relationship of decision time and U-shaped relationship to the number of decisions in relation to information load were based on general experiments of information overload (Iselin, 1990; Edwards, 2004).

This study confirms that there is a significant overload effect on decision accuracy where the *Color-15* treatment led to a lower final result than the other treatments with less information diversity. More surprising was that no drop-off in decision accuracy could be found at the lower diversity end in the experiment, where the value was binned under fewer colors. Thus, the participants were apparently able to reach a comparable solution accuracy at all other levels of information diversity, all the way down to only three distinguishing colors.

The decision time showed a similar significant penalty for the treatment with the highest information diversity (*Color-15*). This confirmed Hypothesis 2 for the treatment with fifteen colors. Participants in this treatment took significantly longer time to reach their best and final solutions. However, no difference in solution time could be determined between the treatments with lower information diversity. It must therefore be concluded that information diversity has little to no impact on the time for novice end-users to reach a solution.

Based on the theoretical framework, the cost of making decisions was also investigated. Related theory led us to measure this parameter through the decision frequency in terms of the number of decisions taken over the whole round. By confirming Hypothesis 3 the number of decisions provide some more insight to the participant's behavior in the different treatments. The higher amounts of decisions taken in the *Color-3* and *Color-15* treatments, relative to the *Color-11* treatment, suggest that a "trial-and-error" approach was used in these former cases. The significantly lower numbers of decisions for eleven colors in contrast, suggests that users could absorb the information given and make more deliberate decisions.

The final main result showed a significant learning development over the three rounds of the experiment. Solution-accuracy improved while solution-time and the number of decisions was lowered logarithmically as subjects worked their way through the three rounds. This confirms previous research, which also found significant evidence for decision improvements with more experience (Iselin, 1990; Hayes-Roth et al., 1983).

By combining these results, the best compromise between cost and benefit of the solution was reached in the treatment where the information was fitted into 11 bins. In the *Color-11* treatment, the solutions were just as accurate and timely as in the *Color-3* and *Color-7* treatments, but since the *Color-11* treatment also led to the least amount of effort, as defined in the hypotheses development, this information diversity alternative shows the most potential.

The findings in this study have ramifications for theory and practice in the emerging field of novice end-user decision support systems. Our results are comparable to the findings of Iselin (1990), which also relate to fifteen informational cues as being in the overload area. This external reference provides some validation to the overload effects found in this study, even though the experiment designs were different. The overload effect above eleven colors is also further confirmed by the increase in time and relatively high number of decisions to reach the final solution.

This study's results provide insights on the information diversity's impact on end-user interface interaction. Here below three of the main resulting guidelines for future designs of effective information diversity in end-user DSS is presented.

First of all, a significant support for filtering and combining potential alternatives to protect against overloading users was evident. It is clear that providing access to all available logging data and the alternative options will inhibit the novice users from actually using the information for effective decision making. For example, support of more efficient appliances or operating behaviors from a seemingly endless database of alternatives, similar to what the energy use information system presented in Chapter 7 allows, could benefit from binning these alternatives based on their energy cost or service value. Based on our findings this would then provide a presentation that is easier to parse. This would allow novice users to understand what operating decisions can impact the energy bill the most while minimizing the impact on comfort, and thus support effective individual decisions.

Second, balancing information diversity to facilitate reflection and less trial-and-error behavior requires more consideration, as the optimal point to facilitate reflection and less trial-and-error behavior was affected both by too high information diversity (*Color-15*) as well as too low (*Color-3*). The fine-tuning of this parameter could be dependent on the decision context and needs to be further evaluated. For example, limiting the information diversity in terms of alternative appliances and operating behaviors too strictly could have adverse effects from supplying insufficient decision context and thus require the user to work through many suboptimal energy use choices. Instead, an energy interface should provide enough alternatives of appliances or operating habits to support users to effectively decide what combination suits them and provides them with the most value.

Third, it can be expected that novice users ability to parse information improve over the time they interact with the interface, as related in the theoretical background and confirmed by the hypothesis of learning (H4). The novice end-user DSS should thus be allowed to develop, in terms of information diversity, as users become more experienced with the subject. An effective energy DSS could potentially provide the information in a less processed format to a more expert end-user. This would allow for a more conceptual understanding of the information. However, more research of the path between the novice and the expert is needed to test this inferred relationship.

As information systems continue to evolve and make decentralized DSS more powerful we predict that research with this specific focus will become more important. Besides the practical implications, this research contributes to research with a novel experimental environment in which to evaluate interface interaction and information diversity's impact on novice end-user decision performance. The study also contributes to decision support methodology specifically by developing and testing a viable way of evaluating hypotheses empirically from a large sample size of heterogeneous subjects. The MTurk platform allowed us to focus on the end-user in a setting more familiar to the participants then a strict laboratory environment.

Part IV

Finale

Chapter 9

Conclusions

"There is a simple way to package information that, under the right circumstances, can make it irresistible. All you have to do is find it."

Malcolm Gladwell - The Tipping Point (2002)

MANY MARKETS HAVE BEEN TRANSFORMED by the recent developments of information and communications technology. These developments have also introduced new opportunities for making energy-use information widely-available throughout the energy system. The challenge of extending the information technology in the energy system is to match the intermittent renewable generation with more a informed and active demand.

The residential sector is particularly affected by these changes since it has had limited access to energy-use feedback. Gathering information about energy usage in buildings is the only way it can be understood and managed with consideration. Thus, to motivate the efficient usage of appliances and support effective decision-making, residential users have been provided with evermore convenient access to information about their energy usage.

However, it is currently unclear whether or not people inhabiting these kinds of buildings can be influenced to manage their energy usage with the information that is provided.

A market transformation is gradual and evolves over multiple iterations. The same development progression is likely in the energy system. This study, therefore, begins by assessing the impact of the energy metering and feedback technology, which is currently implemented. Specifically, the impact of real-time energy-use information is analyzed. Based on the results from this research, further developments which support the users according to their specific informational needs, are proposed and developed. Both the possible technological infrastructure and the further processing of information is critically evaluated and extended. By simplifying and aligning the information with the context of the user, this thesis' aim is to make energy-use feedback more compelling, or "irresistible", as Malcolm Gladwell puts it.

9.1 Contributions

This study focuses on the evaluation and further development of energy-use feedback in the residential sector. The current situation is analyzed in order to clarify the main features of the proposed smart-meter upgrade. Additionally, the complementary information necessary to evaluating the open research questions in the domain of energy-use feedback is defined.

Two challenges stand out in the review of experimental research that provides energy-use feedback. The first concerns the issue of interpreting what impact a specific informational features had on energy usage. The second is the significant variance in the reported results from providing more frequent energy-use feedback.

This study begins by evaluating the potential of the current smart-meter technology. By specifically targeting the provision of real-time energy-use information, the main contribution of the first part of this thesis is, thus, focused on understanding the utility of real-time energy use information in the residential sector. By asking *how* real-time energy feedback is used, the inquiry of this study is aligned with this goal, which is reflected in Research Question 1.

Research Question 1: Evaluation of smart-meter utility *What impact does online access to real-time energy-use information have on energy-use decisions in households?*

To evaluate the impact of implementing smart meters in the residential sector a comparable measurement platform that is portable, non-intrusive and can be easily deployed in the field is built. The design is based on bidirectional communication and real-time information, which are the two main characteristics of the coming smart meters (European Commission, 2011). By first developing and testing a prototype of the measurement system in four different buildings and logging the real and reactive power load from the same seven appliances, a method of isolating and comparing the effects of the buildings' electrical wiring and the already-installed appliances between the buildings is introduced in Chapter 4.

In addition, a method of logging users' interaction by recording their interaction with the real-time information is devised to provide quantitative insight into the use of the provided feedback. Finally, in-depth interviews with all participants provide a third dimension into the relation between energy and information use. By combining the presented approaches, this study exemplifies how the "mixed-method-approach" methodology can be applied to triangulate a certain phenomenon and contribute to a greater understanding of a socio-technical system.

The experimental study in Chapter 5, provides new insight into how the provided real-time information is used. Patterns in the participants' responses in the interview could often be tracked by the way the users altered their behavior. Of the 13 analyzed interviewees, all 4 who had the slowest increase of energy also claimed to have little or no previous knowledge of the challenges and potential of energy-use information. Similarly, frugal users also found that the access to energy-use information had provided a high utility. Finally, out of the 4 energy-use strategies ("no-strategy", "no-potential", "learning effective change" and "base-load change") most of the users who claimed to have learned effective operation changes belonged to the same group of energy-saving users. That all these traits resulted in a more careful increase in energy usage during the fall season of the experiment needs more specific analysis to be substantiated.

Based on the limited impact real-time energy-use feedback has shown on users' motivation and ability to make effective decisions and improve their energy efficiency, further developments of the system are appropriate. Instead of applying external pressure by gamifying the interaction and adding goal-directed or normative comparisons - tactics which has already shown promise (Houwelingen and

Raaij, 1989; McCalley and Midden, 2002; Loock et al., 2013) - this study continues to focus on the specific informational content. This focus has been chosen to explore the innate qualities of information, which are currently underrepresented in energy-use feedback research (Watson et al., 2010).

So far, targeting individual users has been prohibitively costly in terms of the limited impact each user can have. General campaigns with a wide appeal have instead been tested with limited success (Midden et al., 2007). However, this situation is changing as information technology is becoming more wide-spread and powerful. As new possibilities open up for more targeted energy feedback, new research questions become apparent.

Research Question 2: User-centered energy feedback *Based on users' interaction and experiences with real-time energy feedback, what informational detail and content is required to align the energy feedback with their preferences?*

The results from the real-time experiment developed to analyze Research Question 1 are combined with related literature to incorporate experiences from users with different demographical backgrounds and nationalities, and widen the application for the proposed solution. Based on this analysis, this study defines five specific requirements to cater to the needs of the analyzed users. Two requirements help define the necessary information granularity, while the other three pertain to the format and necessary complementary information.

The first two requirements, concerning how the necessary information granularity at the appliance operation level can be made available, are first explored in Chapter 6. Two different energy-use information systems, which use similar approaches, are evaluated. The first system uses frequent smart-meter data to disaggregate the necessary appliance-information. This analysis contributes to the knowledge of the necessary sensor hardware capability for reliable appliance-level information and a parameterization of the requirements for future scalable household disaggregation systems. Several different features from this data have been evaluated. The real and reactive feature (PQ) and the harmonics feature (H) are shown to be superior in our results, in terms of overall detection accuracy, while waveform features proved to be the most robust over the whole range of evaluated data sample-rates (100 Hz to 12 kHz). The PQ feature was slightly affected by the down-scaling of the sample frequency between 1,5 kHz and 12 kHz, while the H feature was more sensitive and dropped in accuracy between 4 kHz and 6 kHz depending on the complexity of the aggregated appliance signal.

By using a public dataset (BLUED) for the analysis (Anderson et al., 2012), this study also contributes with much needed disaggregation research that can be verified and continually developed by other research teams. This enables collaboration and exchange between research groups to spur on the development of a general and scalable energy-use information system that can support more knowledgeable and sustainable decisions.

Because of the amount of noise in the aggregated measurements, the second system instead uses distributed appliance measurements to disaggregate the specific operational settings of each appliance. Chapter 6, thus, evaluates how the same disaggregation techniques can be used on appliance level data to explore how the specific appliance is being operated and at what setting. This presents energy-use feedback on specific operational settings, which was found to be a common parameter that the users tried to evaluate.

The defined requirements for user-centered energy-use feedback were then compared to the capabilities and results of existing information system for residential energy use in Chapter 7. It was shown that currently available approaches – like media advertising, home audits, energy use feedback, and labeling schemes – could not meet all of the requirements. In particular, the ability to combine and process several sources of complimentary information clearly sets home audits and the proposed system apart. In contrast mass media campaigns, energy use feedback and labeling all lack the contextual information from either the user or the appliance manufacturer. In fact, there is no basis for why home audits would be less personalized than the Green IS artifact proposed here. However, due to the cost of performing a home audit, the value resulting from exchanging appliances or changing operating habits is often too small to warrant this analysis at an individual level. Furthermore, a home audit only provides a snapshot of the current state and alternatives. Through continuous measurements and repeated evaluations based on external data, this study instead provides the design for a "perpetual home audit."

This analysis is followed by evaluating appropriate design principles to adhere to the three informational format requirements. First, the analysis concludes that a greater understanding of the information, which is lacking in the presentation of power and energy, can be reached by using monetary units. Second, the monetary unit also provides the basis for supporting long-term decisions, which were frequently searched for by the users with access to real-time energy feedback. Third, a first prototype of a shared appliances and operating mode database is presented. This database provides direct access to alternatives at an appliance- and user level through complementary information from other appliances and operating behaviors.

The design of the developed system is finally analyzed in common appliance operation situations. The monetary units and the long-term decision perspective show direct promise, as it becomes clear that new purchases of more efficient appliances are generally less effective than changing the operating behavior with older appliances. Both the alternative appliance and operating behavior points of reference are supplied by the database information.

The final research direction in this study focuses on evaluating effective ways of presenting information. Although information is necessary in order to make knowledgeable decisions, too much information detail and too many cues to consider can also overload the user and lead to worse decision-making. Information technology is leveling the information-detail gap between experts and novices, however, the ability to parse and use the information depends upon the amount of domain experience. Chapter 8 specifically analyzes how the concept of information-overload affects the novice user, which is comparable to the context where the proposed development of the new energy-feedback system would be implemented. This specific presentation analysis is summarized in Research Question 3.

Research Question 3: User experience and information load *What level of information granularity and how many distinct cues lead novices to make high-impact decisions with the least attempts and in the shortest period of time?*

This research contributes with the decision strategy characteristics of a novice user compared to an expert one. In order to understand how novice users handle new decision-making situations with a large set of alternatives, and what impact the aggregation of the information details has on novice user's ability to make effective decisions, a combinatorial problem experiment is developed and evaluated. The experimental framework tests the novice user's ability to make effective decisions (i.e. decisions with great impact) efficiently (i.e. quickly and with few attempts) in relation to the distinct informational cues given.

The information diversity was given in four treatments with three, seven, eleven and fifteen diversity levels. The last level, with the information divided into fifteen categories, showed significant signs of overloading the user as it resulted in lower overall results and longer solution times. The treatment with the information divided into eleven levels was on par with three and seven, however, it resulted in significantly fewer decisions being needed to reach each solution. Therefore, aggregating the presented information for decisions in eleven or fewer levels is recommended for combinatorial problems directed at novices. Moreover, to support an efficient parsing of the presented information about the different possible options from an energy-use feedback system, the information should be grouped in between seven and eleven decision levels.

In summary, this study provides a sound basis for the design of a more user-centered energy feedback system and for how the information should be presented to be understood and useful for the novice user. This study contributes with specific recommendations for how to develop energy-use feedback to improve its utility in the residential sector, both on a technical and an applicational level.

In the following section open research questions and potential future research directions are outlined and discussed.

9.2 Open Research Questions and Future Work

Beyond the direct contributions of this research, the study has also uncovered further research directions. By critically examining the methods and results presented in this work, this section will propose alternatives of how to address the current limitations.

9.2.1 Energy-Use Sensor Technologies

Two methods to measure and process energy use with increasing levels of sophistication have been developed over the course of this study. The purpose of the different sensors has been to measure the aggregated household load (similar to the smart meter) and the individual appliance loads. The former system utilizes current-transformer sensors and is designed to be installed at a central location, as explained in Chapter 4. The latter uses inline measurements of the power flow, and is installed at the appliance level, as explained in Chapter 6. These two solutions are thus at different ends of the spectrum of measurement invasiveness and aggregation. There are also options in between these extremes that combine these two approaches. This section will examine the strengths and weaknesses of the chosen designs and discuss some of the main options.

The aggregated household-level measurements main strength is its complete coverage of a household's energy use from one point of installation. A central installation can more easily scale than multiple decentralized installations. An additional household-level sensor also supplies a benchmark to the household's energy meter and the subsequent billing of energy use. The major deficiency of this method is the difference in system boundaries between the measurements and the household residents. That is, residents use energy on an appliance level whereas the measurements provide feedback on a household level. This disconnect is also shown in the field experiment explained in Chapter 5 which focuses on how individuals interact with the information and their energy-based services. From the quantitative data and the qualitative interview concerning how the general energy-use information has been used, it is clear that an appliance- and user-operational-level detail is fundamental to providing the information that correlates with users' interaction behavior. This energy-use feedback format is important because the potential benefit of user-centered energy-use information has been quoted as the most likely form of feedback to transfer knowledge and motivate a more energy-saving behavior (Darby, 2010b).

The distributed appliance-level measurement system has the opposite attributes. By monitoring every appliance, the system boundaries are shared with the appliance operators, however, the system can be difficult to scale and install due to the many sensors needed. One alternative has already been analyzed in Chapter 6, and depends on more accurate and frequent measurements at a central location to disaggregate individual appliances with the aid of certain algorithms. This alternative's event-based disaggregation of appliances from central measurements relies on two consecutive steps to be successful. First, the recognition of an appliance state-change, and second, the correct labeling of this particular event. Currently the first event-recognition step is only able to detect a small number of regular events with high certainty; a signal containing more than 20 different appliances resulted in 50% missed events or falsely detected events (Anderson et al., 2012). Jin et al. (2011) has suggested the goodness-of-fit event detection method to be superior to the generalized log-likelihood ratio (GLR) method used by Anderson et al. (2012), which would be an interesting starting point for future research in this direction. The second labeling step has received much attention. However, future research, which intends to build cumulative knowledge, would benefit from benchmarking their approaches against a publicly available dataset similar to the setup in Chapter 6. But this alternative, of disaggregating appliances at a central location, was deemed too unreliable and lacks the necessary detail to be the foundation for the user-level energy-feedback system proposed in Chapter 7.

Combining central measurements with different forms of distributed sensors is another option that should be further explored. Since space heating and larger appliances still make up the majority of the energy used in a household (IEA, 2014), these particular services and appliances could be targeted by a limited number of intrusive sensors and contrasted against the central measurements. Since the distributed sensors most often measure inline with the power flow, additional functionality, as remote and/or automated control, could also be evaluated with this approach. By implementing a modular version of the system proposed in Chapter 7 in conjunction with central smart-meter data, a balance between household, appliance and operating details could be evaluated. Such a measurement system could be further extended by adding more sensors, and the impact of information about particular appliances or operating habits could then be evaluated.

9.2.2 Design Development

The information design presented in this thesis is based on user-centered requirements stipulated by users who have experience with a modern energy-use feedback system. In order to provide more general requirements for this class of artifacts, this study combined the interview responses with related experiments. However, depending on the technology used, the duration of the study, and participant demographics, different results have been reported (van Dam et al., 2010). This means that the current list of requirements should also be subject to an iterative development and not be regarded as final.

In this study it is shown how disaggregated energy-use data can be automatically collected, processed and presented in the form of monetary savings in order to support efficient energy-use decisions. To this end, the proposed energy-use information system calculates the value of currently-installed appliances and compares this against a database of appliances with previously uploaded energy-use data. While a complete appliance database as such is not yet available as this study attests to, much of the necessary information can be gathered through several publicly available sources.

It would require a community effort to additionally contribute the load profiles for various household appliances and their associated cycles and states. But when this information need only be logged once manually, to benefit all other users with the same appliance. By uploading more usage parameters to the database, the accuracy of the expected mean consumption of the appliances can be improved, which will also extend the usability of the platform for the whole community. A potential avenue for research would, thus, be to investigate how the development of such a database can be established. This entails research into how appliance signatures can be standardized, as well as how households can be incentivized to provide this information and regularly update the database when new appliances become available on the market.

In order to provide a clear conceptualization and design of the envisioned energy-use information system, a deliberately simple approach has been chosen to conduct the underlying economic evaluation of different appliances and behavioral alternatives. Several improvements are possible here. For example, the payback period could incorporate an appropriate discount factor, and also possibly forecast future energy costs. Moreover, it would also be feasible to provide monetary information on the outcomes that can be achieved by replacing appliances and changing the usage behavior. In this context, it might be worthwhile to extend the comparison engine in our system with a collaborative-filtering based recommender system that could disseminate the best practices of similar households (Elsner, 2011).

The proposed system is the result of the first design iteration in a design process. Input from intended users are once again needed to confirm and correct the course of this design effort. The evaluation can only show the decision support system's technical adherence to the given requirements. In particular, the technical evaluation cannot comment on whether or not the proposed information is comprehensible. As the objective of this study is to introduce and analyze the technical viability of a Green IS energy-feedback system, such a behavioral evaluation is beyond the scope of this study. One framework that has been suggested to structure this kind of research is the belief-action-outcome (BAO) framework, where each step towards a predetermined goal can be scrutinized (Melville, 2010). To evaluate the BAO adoption phases, the results of each step can then be contrasted with conclusions from other field trials or surveys of smart meters to help improve the direction of the energy usage.

Naturally, the possibilities for innovation in Green IS go well beyond using the information proposed in this study. For example, influencing attitudes of disillusionment - that each individual has a very small impact (Strengers, 2011) - could be targeted by integrating information about the power of many. Another target could be aimed at normative beliefs, which involve the perceived relation between oneself and other similar users. This belief could be addressed by providing comparisons between similar user and appliance groups (Petersen et al., 2007).

9.2.3 Experimental Approaches

The field experimental analysis described in Chapter 5 evaluates whether the treatment led to an overall change in behavior and how interface-use frequency correlates with energy-use. However, spurious effects that are not described by this data, for example, university terms and daily energy-use cycles, could also influence the correlations (Shumway and Stoffer, 2011). Multiple observations over several interactions have been combined to mediate these unrelated spurious effects as recommended by Cook et al. (1979).

Similarly, in the real-time energy feedback experiment, it cannot be assumed that every participant in each household had the same interest in the experiment. The different tenants could thus have distorted the effect in overall energy-use. Future research should assure that each participant has an individual login to better understand how individuals interact with the information. Additionally, it would have been helpful when analyzing the data if the location coordinates of the user's login were tracked. Simply knowing if the login was done remotely or locally would be a good start, as this would indicate whether or not the user was able to influence the current energy use directly.

The experiment duration should also ideally allow for an evaluation of seasonal effects. This study's focus on users' experience with real-time energy feedback has been found to be less affected by seasons, as the switch between active interaction with the feedback and a more "backgrounded" interaction has been shown to happen rapidly in similar experiments (Hargreaves et al., 2013). However, the results would have a wider application if they were collected over an extended period of time.

In this context, it should be noted that large and long term research studies require much planning and preparation as they are difficult to manage due to interdependencies and the different focuses of the people involved (Darby et al., 2011). Moreover, extensive hardware testing and action plans for measurement outages are very important, as unforeseen failures are common and affect the useful evaluation (Ehrhardt-Martinez et al., 2010; Darby et al., 2011; Hargreaves et al., 2013). Similar to these studies, the field experiment in this study also had hardware outages. The lapses in sent data were due to a lack of battery replacement and failing connections as the project wore on, as well as some participants' habits of disconnecting routers and network peripherals when they were not in active personal use. In future and more long-term studies, a more robust system that preferably uses mains power or a lower information resolution would improve the stability of these results. Furthermore, unforeseen, but highly likely outages in the information distribution system could be made more reliable by implementing an intermediate storage for the energy-use data.

Beyond the strictly real-time energy-use information, other studies within the same framework could also be anticipated. For example, communication among the members of a peer group can be introduced in order to foster peer effects like internal collaboration, e.g. the sharing of tips for efficient usage. Also, competition *between* different peer-groups appears to be a conceivable incentive (Brewer et al., 2011; Petersen et al., 2007). Eventually, the provision of applications for mobile devices would bring energy-use feedback even closer to the users and their daily routines (Weiss et al., 2009).

In contrast to field experiments, more specific design considerations can be handled more efficiently and with greater control in more focused design environments. Chapter 8 provides an example of how the impact of information overload can be studied more individually. Testing participants' reactions in an abstract environment, as the current experiment is set up, has been a good way to start building the theory in an emergent field.

In this study, Amazon's online "Mechanical Turk" platform was chosen as the experimental location in favor of a laboratory setup. The decision made it possible to attract participants with a diversity of backgrounds, which is uncommon in university laboratories. However, while the platform provides a heterogeneous pool of participants, the experimental design must take into account the fact that some subjects are only trying to get through the task quickly, without regard for understanding the task at hand. This is a well-known challenge, which is usually solved with comprehension questions (Oppenheimer et al., 2009). In hindsight, our implementation could have used a stricter implementation of flagging participant's gaming. We relied on monetary incentives for improving the result but this did not help participants that had misunderstood the game setup. More hands-on tutorials should be tested for future implementations.

By using an online platform that provides reliable results at a small cost in comparison to what university facilities cost to run, incremental changes can be tested and evaluated in detail. This opens up new venues to winnow out counteracting informational features before being tested in more expensive field trials. Controlled experiments for developing user experience is one example application and a method for evaluating user feedback for interaction development. In particular, further evaluation could be based on implementing specific changes to a population sample. This method is widely used on web platforms like Amazon to perform automated controlled experiments to optimize their design (Kohavi et al., 2007). Exploring this method with a "lean-research" approach could prove to be an effective process through which to explore the topics of information-systems research in general and energy informatics in particular.

Appendix A

Regression Analysis

Table A.1: Multiple linear regression analysis of power use (Watt) in experimental group as a function of feedback, Number of residents, Weekday and Hour of day

	Dependent variable:	
	Watt	Std. Error
feedback	13.233	(2.302)***
Nresidents	60.104	(1.313)***
dayOfWeek_Tuesday	17.575	(4.132)***
dayOfWeek_Wednesday	7.593	(4.147)*
dayOfWeek_Thursday	4.654	(4.110)
dayOfWeek_Friday	−11.550	(4.111)***
dayOfWeek_Saturday	−21.601	(4.189)***
dayOfWeek_Sunday	25.898	(4.132)***
hourOfDay01	−80.792	(7.607)***
hourOfDay02	−125.239	(7.606)***
hourOfDay03	−148.661	(7.615)***
hourOfDay04	−160.904	(7.617)***
hourOfDay05	−142.520	(7.618)***
hourOfDay06	−117.642	(7.624)***
hourOfDay07	−82.979	(7.621)***
hourOfDay08	−94.198	(7.588)***
hourOfDay09	−47.818	(7.578)***
hourOfDay10	−49.297	(7.582)***
hourOfDay11	−30.431	(7.576)***
hourOfDay12	23.479	(7.570)***
hourOfDay13	39.014	(7.564)***
hourOfDay14	14.210	(7.571)*
hourOfDay15	19.369	(7.573)**
hourOfDay16	12.538	(7.571)*
hourOfDay17	27.465	(7.582)***
hourOfDay18	59.745	(7.579)***
Note:	*p<0.1; **p<0.05; ***p<0.01	

Table A.1 – *Continued from previous page*

	Dependent variable:	
	Watt	Std. Error
hourOfDay19	113.891	(7.568)***
hourOfDay20	138.097	(7.566)***
hourOfDay21	120.397	(7.545)***
hourOfDay22	112.908	(7.544)***
hourOfDay23	84.498	(7.548)***
Constant	104.163	(7.468)***
Observations	134.091	
R^2	0.063	
Adjusted R^2	0.063	
Residual Std. Error	400.768	(df = 134059)
F Statistic	292.888***	(df = 31; 134059)
Note:	*p<0.1; **p<0.05; ***p<0.01	

Table A.2: Regression analysis of power use (Watt) in experimental group as a function of feedback, Household, Weekday and Hour of day

	Dependent variable:	
	Watt	Std. Error
feedback	22.142	(2.302)***
factor(House)B	93.546	(7.916)***
factor(House)C	−19.295	(8.322)**
factor(House)D	−105.710	(8.138)***
factor(House)E	−65.003	(7.988)***
factor(House)F	−1.252	(9.955)
factor(House)G	−90.981	(9.849)***
factor(House)H	−117.284	(8.818)***
factor(House)I	−220.272	(7.939)***
factor(House)J	−13.832	(8.164)*
factor(House)K	−28.176	(8.542)***
factor(House)L	−268.537	(8.483)***
factor(House)M	−177.698	(8.018)***
factor(House)N	−179.294	(7.944)***
factor(House)O	−245.561	(8.412)***
factor(House)P	−248.881	(7.964)***
factor(House)Q	−141.038	(10.605)***
factor(House)R	−193.943	(8.377)***
factor(House)S	−229.465	(8.645)***
dayOfWeekMonday	14.225	(4.002)***
dayOfWeekSaturday	−9.434	(3.996)**
dayOfWeekSunday	39.008	(3.945)***
dayOfWeekThursday	18.095	(3.918)***
dayOfWeekTuesday	32.310	(3.948)***
dayOfWeekWednesday	21.202	(3.956)***
hourOfDay01	−80.773	(7.400)***
hourOfDay02	−125.193	(7.399)***
hourOfDay03	−148.669	(7.407)***
hourOfDay04	−160.773	(7.410)***
hourOfDay05	−142.282	(7.411)***
hourOfDay06	−117.490	(7.417)***
hourOfDay07	−83.126	(7.414)***
hourOfDay08	−94.379	(7.382)***
hourOfDay09	−48.113	(7.372)***
hourOfDay10	−49.551	(7.376)***
hourOfDay11	−30.719	(7.370)***
hourOfDay12	23.440	(7.364)***
hourOfDay13	38.780	(7.359)***
hourOfDay14	13.933	(7.365)*
hourOfDay15	18.992	(7.367)***
hourOfDay16	12.135	(7.365)*
hourOfDay17	27.534	(7.376)***
Note:	*p<0.1; **p<0.05; ***p<0.01	

Table A.2 – *Continued from previous page*

	Dependent variable:	
	Watt	Std. Error
hourOfDay18	59.862	(7.373)***
hourOfDay19	113.849	(7.362)***
hourOfDay20	138.293	(7.360)***
hourOfDay21	120.496	(7.339)***
hourOfDay22	113.052	(7.339)***
hourOfDay23	84.828	(7.342)***
Constant	390.082	(9.120)***
Observations	134.091	
R^2	0.114	
Adjusted R^2	0.114	
Residual Std. Error	389.860	(df = 134042)
F Statistic	358.715***	(df = 48; 134042)
Note:	*p<0.1; **p<0.05; ***p<0.01	

Table A.3: Multiple linear regression with the power usage as a response to 2-5 Hours prior interaction and household, weekday and hour of day.

	Dependent variable:	
	Watt	Std. Error
factor(hour)2	−80.896	(25.108)***
factor(hour)3	−32.307	(26.685)
factor(hour)4	−18.479	(27.495)
factor(hour)5	46.606	(28.938)
factor(House)B	47.101	(13.003)***
factor(House)C	−42.049	(14.271)***
factor(House)D	−161.768	(14.063)***
factor(House)E	−147.745	(13.751)***
factor(House)F	28.038	(17.080)
factor(House)G	−146.942	(15.579)***
factor(House)H	−160.582	(16.015)***
factor(House)I	−257.319	(13.682)***
factor(House)J	21.571	(14.146)
factor(House)K	−96.017	(15.152)***
factor(House)L	−314.839	(13.933)***
factor(House)M	−225.132	(13.702)***
factor(House)N	−211.316	(13.710)***
factor(House)O	−295.358	(14.162)***
factor(House)P	−292.959	(13.596)***
factor(House)Q	−142.351	(16.965)***
factor(House)R	−242.045	(14.035)***
Note:	*p<0.1; **p<0.05; ***p<0.01	

Table A.3 – *Continued from previous page*

	Dependent variable:	
	Watt	Std. Error
factor(House)S	−270.906	(13.790)***
dayOfWeekMonday	8.780	(4.920)*
dayOfWeekSaturday	−1.056	(4.788)
dayOfWeekSunday	43.562	(4.847)***
dayOfWeekThursday	12.629	(4.706)***
dayOfWeekTuesday	27.720	(4.888)***
dayOfWeekWednesday	26.057	(4.748)***
hourOfDay01	−82.217	(8.988)***
hourOfDay02	−122.080	(8.988)***
hourOfDay03	−143.449	(9.005)***
hourOfDay04	−154.800	(9.009)***
hourOfDay05	−127.591	(9.010)***
hourOfDay06	−96.038	(9.021)***
hourOfDay07	−87.993	(9.026)***
hourOfDay08	−99.289	(9.028)***
hourOfDay09	−43.748	(9.026)***
hourOfDay10	−51.226	(9.042)***
hourOfDay11	−31.628	(9.052)***
hourOfDay12	17.200	(9.032)*
hourOfDay13	32.084	(9.020)***
hourOfDay14	12.953	(9.016)
hourOfDay15	27.437	(9.004)***
hourOfDay16	26.112	(9.001)***
hourOfDay17	51.728	(8.996)***
hourOfDay18	82.116	(9.000)***
hourOfDay19	133.628	(8.995)***
hourOfDay20	159.127	(8.992)***
hourOfDay21	133.508	(8.982)***
hourOfDay22	91.344	(8.980)***
hourOfDay23	79.226	(8.980)***
Constant	406.921	(17.266)***
Observations	82.933	
R^2	0.134	
Adjusted R^2	0.134	
Residual Std. Error	375.285	(df = 82879)
F Statistic	242.645***	(df = 53; 82879)
Note:	*p<0.1; **p<0.05; ***p<0.01	

Note: *p<0.1; **p<0.05; ***p<0.01

Appendix B

Interview Results

Table B.1: Interview results divided into seven categories based on keywords from the participants

House ID	Energy-use increase	Inhabitants	Size	Current energy-use	Energy-use knowledge	Web-based feedback usage
A	Mid	4	$130m^2$	No comment	No comment	Middle
B	Mid	3	$105m^2$	Netural	Fundamental	Active
C	High	4	$137m^2$	Neutral	Fundamental	Active
D	Low	3	$69m^2$	Careful	None	Latent
E	Mid	3	$80m^2$	Careful	Low	Middle
F	High	5	$139m^2$	Wasteful	Fundamental	Latent
G	Mid	3	$130m^2$	Neutral	Fundamental	Middle
I	Mid	4	$110m^2$	Careful	Expert	Active
L	Low	2	$62m^2$	Careful	Low	Active
M	Low	2	$82m^2$	Careful	Low	Active
N	Mid	3	$85m^2$	Neutral	Low	Active
P	Low	2	$52m^2$	No comment	Low	Active
R	Mid	2	$63m^2$	Neutral	Fundamental	Latent

House ID	Energy-use saving strategies	Dedicated display utility	Shift loads in time ability	Data privacy concerns
A	No potential	Neutral	Limited	Considerable
B	Turn off idle	Reduces obstacles	Limited	Some
C	Efficient operation	Induce conflict	Limited	Considerable
D	No effort	Reduces obstacles + Portal to web	Limited	None
E	Turn off idle	No comment	Limited	None
F	No effort	No comment	Automated control	Some
G	Turn off idle	Reduces obstacles	Automated control	None
I	No potential	Portal to web	Automated control	Considerable
L	Efficient operation + Turn off idle	Alarm function	Limited	Considerable
M	Efficient operation	Reduces obstacles	Limited	No comment
N	Efficient operation + Turn off idle	Alarm function	Automated control	Some
P	Efficient operation	Neutral	Limited	None
R	No effort	Neutral	Flexible	Considerable

Bibliography

Abrahamse, W. and L. Steg (2009, October). How do socio-demographic and psychological factors relate to households' direct and indirect energy use and savings? *Journal of Economic Psychology 30*(5), 711–720.

Abrahamse, W., L. Steg, C. Vlek, and T. Rothengatter (2005). A review of intervention studies aimed at household energy conservation. *Journal of Environmental Psychology 25*(3), 273–291.

AGEB (2014). Energieverbrauch in Deutschland im Jahr 2013. Technical report, Arbeitsgemeinschaft Energiebilanzen, Berlin, DE.

Alahmad, M. a., P. G. Wheeler, A. Schwer, J. Eiden, and A. Brumbaugh (2012, apr). A Comparative Study of Three Feedback Devices for Residential Real-Time Energy Monitoring. *IEEE Transactions on Industrial Electronics 59*(4), 2002–2013.

Allen, D. and K. Janda (2006). The effects of household characteristics and energy use consciousness on the effectivness of real time energy use feedback. In *American Council for an Energy-Efficient Economy 2006 Summer Study, 'Human and Social Dimensions of Energy Use'*, Washington, DC, US, pp. 1–12.

Anderson, K., A. Ocneanu, and D. Benitez (2012). BLUED: a fully labeled public dataset for Event-Based Non-Intrusive load monitoring research. In *Proceedings of the 2nd KDD Workshop on Data Mining Applications in Sustainability*, Beijing, China.

Anderson, K. D., M. E. Berges, A. Ocneanu, D. Benitez, and J. M. Moura (2012, October). Event detection for Non Intrusive load monitoring. In *IECON 2012 -*

38th Annual Conference on IEEE Industrial Electronics Society, pp. 3312–3317. IEEE.

Anderson, W. and V. White (2009). Exploring consumer preferences for home energy display functionality. *Centre for Sustainable Energy*, 1–31.

Arrillaga, J. and N. R. Watson (2003, September). *Power System Harmonics*. Chichester, UK: John Wiley & Sons, Ltd.

Asada, T., W. Odendaal, and J. van Wyk (2003). An overview of integratable current sensor technologies. In *38th IAS Annual Meeting on Conference Record of the Industry Applications Conference, 2003.*, Volume 2, pp. 1251–1258. IEEE.

Ayres, R. and B. Warr (2009). *The Economic Growth Engine: How Energy and Work Drive Material Prosperity*. Northampton: Edward Elgar Publishing Inc.

Baranski, M. and J. Voss (2003). Non-intrusive appliance load monitoring based on an optical sensor. In *2003 IEEE Bologna Power Tech Conference Proceedings,*, Volume 4, pp. 267–274. IEEE.

Baranski, M. and J. Voss (2004). Detecting Patterns of Appliances from Total Load Data Using a Dynamic Programming Approach. In *Fourth IEEE International Conference on Data Mining (ICDM'04)*, pp. 327–330. IEEE.

Bartlett, J. and E. Toms (2005). How is Information used? Applying task analysis to understanding information use. In *Proceedings of the 35th annual conference of the Canadian association for information science*.

Baskerville, R. and J. Pries-Heje (2010, September). Explanatory Design Theory. *Business & Information Systems Engineering 2*(5), 271–282.

Basseville, M. and I. V. Nikiforov (1993, March). *Detection of abrupt changes: theory and application*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc.

Bawden, D. and L. Robinson (2008, November). The dark side of information: overload, anxiety and other paradoxes and pathologies. *Journal of Information Science 35*(2), 180–191.

Beach, L. R. and T. R. Mitchell (1978, July). A Contingency Model for the Selection of Decision Strategies. *Academy of Management Review 3*(3), 439–449.

Beckel, C., W. Kleiminger, T. Staake, and S. Santini (2012, July). Improving device-level electricity consumption breakdowns in private households using ON/OFF events. *ACM SIGBED Review 9*(3), 32–38.

Benbasat, I. and A. S. Dexter (1986, March). An Investigation of the Effectiveness of Color and Graphical Information Presentation under Varying Time Constraints. *MIS Quarterly 10*(1), 59–83.

Benbasat, I., D. K. Goldstein, and M. Mead (1987, September). The Case Research Strategy in Studies of Information Systems. *MIS Quarterly 11*(3), 369–386.

Berges, M., E. Goldman, H. S. Matthews, and L. Soibelman (2009, June). Learning Systems for Electric Consumption of Buildings. In *Computing in Civil Engineering (2009)*, Reston, VA, pp. 1–10. American Society of Civil Engineers.

Berges, M., E. Goldman, H. S. Matthews, L. Soibelman, and K. Anderson (2011, November). User-Centered Nonintrusive Electricity Load Monitoring for Residential Buildings. *Journal of Computing in Civil Engineering 25*(6), 471–480.

Berinsky, a. J., G. a. Huber, and G. S. Lenz (2012, March). Evaluating Online Labor Markets for Experimental Research: Amazon.com's Mechanical Turk. *Political Analysis 20*(3), 351–368.

Bettman, J. R. and P. Kakkar (1977, March). Effects of Information Presentation Format on Consumer Information Acquisition Strategies. *Journal of Consumer Research 3*(4), 233–240.

Bhattacherjee, A. and C. Sanford (2006, December). Influence processes for information technology acceptance: an elaboration likelihood model. *MIS Quarterly 30*(4), 805–825.

Bittle, R., R. Valesano, and G. Thaler (1979). The effects of daily cost feedback on residential electricity consumption. *Behavior Modification 3*(2), 187–202.

Booth, A., N. Demirdoven, and H. Tai (2010). The smart grid opportunity for solution providers. Technical report, McKinsey, Atlanta.

Boudreau, M., A. Chen, and M. Huber (2008). Green IS: Building sustainable business practices. In R. T. Watson (Ed.), *Information Systems*. Athens, GA: Global Text Project.

Brandon, G. and A. Lewis (1999). Reducing household energy consumption: a qualitative and quantitative field study. *Journal of Environmental Psychology 19*(1), 75–85.

Braungart, M., W. McDonough, and A. Bollinger (2007, September). Cradle-to-cradle design: creating healthy emissions - a strategy for eco-effective product and system design. *Journal of Cleaner Production 15*(13-14), 1337–1348.

Brewer, R., G. Lee, Y. Xu, C. Desiato, M. Katchuck, and P. Johnson (2011). Lights Off. Game On. The Kukui Cup: A Dorm Energy Competition. In *Proceedings of the CHI 2011 Workshop on Everyday Practice and Sustainable HCI*, Vancouver, Canada.

Brillouin, L. (1962). *Science and information theory* (2nd ed.). New York, NY: Academic Press.

Brocke, J., P. Loos, S. Seidel, and R. T. Watson (2013, August). Green IS. *Business & Information Systems Engineering 5*(5), 295–297.

Brodersen, K. H., F. Gallusser, J. Koehler, N. Remy, and S. L. Scott (2015, mar). Inferring causal impact using Bayesian structural time-series models. *The Annals of Applied Statistics 9*(1), 247–274.

Butler, T. (2011, March). Compliance with institutional imperatives on environmental sustainability: Building theory on the role of Green IS. *The Journal of Strategic Information Systems 20*(1), 6–26.

Carrie Armel, K., A. Gupta, G. Shrimali, and A. Albert (2013, January). Is disaggregation the holy grail of energy efficiency? The case of electricity. *Energy Policy 52*, 213–234.

CEE (2014). 2013 State of the Efficiency Program Industry. Technical report, Consortium for Energy Efficiency, Boston, MA.

Chahine, K., K. E. K. Drissi, C. Pasquier, K. Kerroum, C. Faure, T. Jouannet, and M. Michou (2011, January). Electric Load Disaggregation in Smart Metering Using a Novel Feature Extraction Method and Supervised Classification. *Energy Procedia 6*, 627–632.

Chetty, M., D. Tran, and R. E. Grinter (2008, sep). Getting to green. In *Proceedings of the 10th international conference on Ubiquitous computing - UbiComp '08*, New York, New York, USA, pp. 242. ACM Press.

Cole, A. and A. Albicki (1998). Algorithm for nonintrusive identification of residential appliances. In *ISCAS '98. Proceedings of the 1998 IEEE International Symposium on Circuits and Systems (Cat. No.98CH36187)*, Volume 3, pp. 338–341. IEEE.

Cook, T., D. Campbell, and A. Day (1979). *Quasi-experimentation: Design & analysis issues for field settings* (1 ed.). Boston, MA, USA: Houghton Mifflin Co.

Costanza, E., S. D. Ramchurn, and N. R. Jennings (2012, September). Understanding domestic energy consumption through interactive visualisation. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing - UbiComp '12*, New York, New York, USA, pp. 216. ACM Press.

Cox, R. W., P. L. Bennett, T. D. McKay, J. Paris, and S. B. Leeb (2007, May). Using the Non-Intrusive Load Monitor for Shipboard Supervisory Control. In *2007 IEEE Electric Ship Technologies Symposium*, pp. 523–530. IEEE.

Dalén, A. (2014). Real-Time Energy Information for Real-Time Energy Use.

Dalén, A. and J. Krämer (2014a). Effective Information Load for Novice End-users - Experimental Evidence. *Business & Information Systems Engineering*.

Dalén, A. and J. Krämer (2014b). Towards a user-centered design for smart meter interfaces to support efficient energy-use choices: A design science approach. *BISE*.

Dalén, A., J. Krämer, and C. Weinhardt (2013). Determining The Optimal Level Of Information Granularity For Efficient Energy Consumption Decisions: Experimental Evidence. In *ECIS 2013 Research in Progress*, Utrecht, Netherlands.

Dalén, A., T. Teubner, and C. Weinhardt (2012, jun). wEnergy: A Field Experiment on Energy Consumption and Social Feedback. In *COLLA 2012, The Second International Conference on Advanced Collaborative Networks, Systems and Applications*, pp. 65–68.

Dalen, A. and C. Weinhardt (2012, jan). Fusebox dialects: Central appliance recognition for a smarter grid. In *2012 IEEE PES Innovative Smart Grid Technologies (ISGT)*, pp. 1–7. IEEE.

Dalén, A. and C. Weinhardt (2014). Evaluating the impact of data sample-rate on appliance disaggregation. In *EnergyCon IEEE International Energy Conference*, Dubrovnik, Croatia.

Darby, S. (2006). The effectiveness of feedback on energy consumption. *A Review for DEFRA of the Literature on Metering, Billing and direct Displays 486*.

Darby, S. (2008, October). Energy feedback in buildings: improving the infrastructure for demand reduction. *Building Research & Information 36*(5), 499–508.

Darby, S. (2010a). Literature review for the energy demand research project. *Change*, 1–30.

Darby, S. (2010b). Literature review for the energy demand research project. *Change*, 1–30.

Darby, S. (2010c). Smart metering: what potential for householder engagement? *Building Research & Information 38*(5), 442–457.

Darby, S., W. Anderson, and V. White (2011). Large-scale testing of new technology: some lessons from the UK smart metering and feedback trials. In *European Council for an Energy-Efficient Economy, Summer Study*, pp. 1–7.

Dobson, J. K. and J. A. Griffin (1992). Conservation effect of immediate electricity cost. feedback on residential consumption behavior. Technical report, Ontario Hydro.

Du, Y., L. Du, B. Lu, R. Harley, and T. Habetler (2010, September). A review of identification and monitoring methods for electric loads in commercial and residential buildings. In *2010 IEEE Energy Conversion Congress and Exposition*, pp. 4527–4533. IEEE.

EAC (2008). Smart Grid: Enabler of the New Energy Economy. Technical report, Electricity Advisory Committee.

Edwards, S. H. (2004, March). Using software testing to move students from trial-and-error to reflection-in-action. In *Proceedings of the 35th SIGCSE technical symposium on Computer science education - SIGCSE '04*, New York, NY, US, pp. 26–30. ACM Press.

Efthymiou, C. and G. Kalogridis (2010, October). Smart Grid Privacy via Anonymization of Smart Metering Data. In *2010 First IEEE International Conference on Smart Grid Communications*, pp. 238–243. IEEE.

Ehrhardt-Martinez, K., K. A. Donnelly, and S. Laitner (2010). Advanced metering initiatives and residential feedback programs: A meta-review for household electricity-saving opportunities. In *American Council for an Energy-Efficient Economy*, Washington, DC, US.

EIA (2014). Annual Energy Outlook 2014. Technical report, U.S. Energy Information Administration, Washington, DC, USA.

Ellery E. Queen (2011). Smart Meters and Smart Meter Systems : A Metering Industry Perspective. Technical report, Edison Electric Institute, Washington, DC, USA.

Elliot, S. (2011, March). Transdisciplinary perspectives on environmental sustainability: a resource base and framework for IT-enabled business transformation. *MIS Quarterly 35*(1), 197–236.

Elsner, H. (2011). *Managing the Evolution of Corporate Portals - A User-Centric Approach*. Ph. D. thesis, Fakultät für Wirtschaftswissenschaften, Karlsruhe Institute of Technogy (KIT).

Enerdata (2012). Energy Efficiency Indicators - Average electricity consumption per electrified household.

ESMA (2010). Annual Report on the Progress in Smart Metering. Technical report, Intelligent Energy Europe, Brussels.

European Commission (2011). Impact Assessment - Annex IX.

European Commission (2012). Directive 2012/27/EU on energy efficiency, amending Directives 2009/125/EC and 2010/30/EU and repealing Directives 2004/8/EC and 2006/32/EC.

Farhangi, H. (2010, January). The path of the smart grid. *IEEE Power and Energy Magazine 8*(1), 18–28.

Farhar, B. and C. Fitzpatrick (1989, January). Effects of feedback on residential electricity consumption: A literature review. Technical report, National Renewable Energy Laboratory (NREL), Golden, CO.

Farinaccio, L. and R. Zmeureanu (1999). Using a pattern recognition approach to disaggregate the total electricity consumption in a house into the major end-uses. *Energy and Buildings 30*(3), 245–259.

Faruqui, A. and S. Sergici (2009). BGE's Smart Energy Pricing Pilot Summer 2008 Impact Evaluation. Technical report, The Brattle Group Inc.

Figueiredo, M. B., A. de Almeida, and B. Ribeiro (2011, October). Wavelet Decomposition and Singular Spectrum Analysis for electrical signal denoising. In *2011 IEEE International Conference on Systems, Man, and Cybernetics*, pp. 3329–3334. IEEE.

Fischer, C. (2008, may). Feedback on household electricity consumption: a tool for saving energy? *Energy Efficiency 1*(1), 79–104.

Fitzpatrick, G. and G. Smith (2009, January). Technology-Enabled Feedback on Domestic Energy Consumption: Articulating a Set of Design Concerns. *IEEE Pervasive Computing 8*(1), 37–44.

Fogg, B. (2003). Computers as persuasive social actors. In *Persuasive Technology* (1 ed.)., pp. 89–120. Morgan Kaufmann.

Ford, A. (1997). System Dynamics and the Electric Power Industry. *System Dynamics Review 13*(1), 57–85.

Frieden, B. J. and K. Baker (1983). The Market Needs Help: the disappointing record of home energy conservation. *Journal of Policy Analysis and Management 2*(3), 432–448.

Froehlich, J. (2009). Promoting energy efficient behaviors in the home through feedback: The role of human-computer interaction. In *Proc. HCIC Workshop*, Volume 9.

Froehlich, J., E. Larson, S. Gupta, G. Cohn, M. Reynolds, and S. Patel (2011). Disaggregated end-use energy sensing for the smart grid. *Pervasive Computing, IEEE 10*(1), 28–39.

Gamberini, L., A. Spagnolli, N. Corradi, G. Jacucci, G. Tusa, T. Mikkola, L. Zamboni, and E. Hoggan (2012). Tailoring Feedback to Users' Actions in a Persuasive Game for Household Electricity Conservation. In *Persuasive Technology Design for Health and Safety*, Volume 7284 of *Lecture Notes in Computer Science*, pp. 100–111. Berlin, Heidelberg: Springer Berlin Heidelberg.

Gardner, G. T. and P. C. Stern (2002). *Environmental Problems and Human Behavior* (2 ed.). Boston, MA: Pearson Custom Publishing.

Gläser, J. and G. Laudel (2010). *Experteninterviews und qualitative Inhaltsanalyse* (4 ed.). Heidelberg, Germany: VS Verlag.

Goebel, C., H.-A. Jacobsen, V. Razo, C. Doblander, J. Rivera, J. Ilg, C. Flath, H. Schmeck, C. Weinhardt, D. Pathmaperuma, H.-J. Appelrath, M. Sonnenschein, S. Lehnhoff, O. Kramer, T. Staake, E. Fleisch, D. Neumann, J. Strüker,

K. Erek, R. Zarnekow, H. Ziekow, and J. Lässig (2013, December). Energy Informatics. *Business & Information Systems Engineering 6*(1), 25–31.

Goodman, J. K., C. E. Cryder, and A. Cheema (2013, July). Data Collection in a Flat World: The Strengths and Weaknesses of Mechanical Turk Samples. *Journal of Behavioral Decision Making 26*(3), 213–224.

Google (2011). Update on Google Health and Google PowerMeter. Available online at http://googleblog.blogspot.com/2011/06/update-on-google-health-and-google.html, Last accessed: 2015-04-07.

Gorry, G. A. and M. S. Scott Morton (1971). A Framework for Management Information Systems,. *Sloan Management Review 13*, 55–70.

Green, G. I. and C. T. Hughes (1986, October). Effects of decision support training and cognitive style on decision process attributes. *Journal of Management Information Systems 3*(2), 83–93.

Gregor, S. and D. Jones (2007). The anatomy of a design theory. *Journal of the Association of Information Systems 8*(5), 312–335.

Grønhøj, A. and J. Thøgersen (2011). Feedback on household electricity consumption: learning and social influence processes. *International Journal of Consumer Studies 35*(2), 138–145.

Gupta, S., M. S. Reynolds, and S. N. Patel (2010, September). ElectriSense: single-point sensing using EMI for electrical event detection and classification in the home. In *Proceedings of the 12th ACM international conference on Ubiquitous computing - Ubicomp '10*, New York, New York, USA, pp. 139–148. ACM Press.

Gutberlet, K.-L. (2008). Energieeffizienz im Haushalt.

Haakana, M., L. Sillanpaeae, and M. Talsi (1997). The effect of feedback and focused advice on household energy consumption. In *1997 ECEEE summer study conference*, Spindleruv Mlyn, Czech Republic.

Hadden, S. (1999). Commercial Nonintrusive Appliance Load Monitor: Beta-Test Results. Technical report, EPRI, Palo Alto, CA.

Ham, J. and C. Midden (2010). Ambient Persuasive Technology Needs Little Cognitive Effort: The Differential Effects of Cognitive Load on Lighting Feedback versus Factual Feedback. In *Persuasive Technology*, pp. 132–142. Springer Berlin Heidelberg.

Hargreaves, T., M. Nye, and J. Burgess (2010). Making energy visible: A qualitative field study of how householders interact with feedback from smart energy monitors. *Energy Policy 38*(10), 6111–6119.

Hargreaves, T., M. Nye, and J. Burgess (2013, January). Keeping energy visible? Exploring how householders interact with feedback from smart energy monitors in the longer term. *Energy Policy 52*, 126–134.

Harrison, B. (2003). Household Electricity Usage Display and Intelligent Metering Using Current Digital Service System Reception Platforms (Set Top Box). In *Proceedings, Third International Conference on Energy Efficiency in Domestic Appliances and Lighting*, Turin, Italy.

Hart, D. G. (2008, July). Using AMI to realize the Smart Grid. In *2008 IEEE Power and Energy Society General Meeting - Conversion and Delivery of Electrical Energy in the 21st Century*, pp. 1–2. IEEE.

Hart, G. (1992). Nonintrusive appliance load monitoring. *Proceedings of the IEEE 80*(12), 1870–1891.

Hauttekeete, L., J. Stragier, W. Haerick, and L. De Marez (2010). Smart, smarter, smartest…the consumer meets the smart electrical grid. In *Telecommunications Internet and Media Techno Economics (CTTE), 2010 9th Conference on*, pp. 1–6. IEEE.

Hayes, S. and J. Cone (1977). Reducing residential electrical energy use: Payments, information, and feedback. *Journal of Applied Behavior Analysis 10*(3), 425.

Hayes, S. C. and J. D. Cone (1981). Reduction of residential consumption of electricity through simple monthly feedback. *Journal of Applied Behavior Analysis 14*(1), 81–88.

Hayes-Roth, F., D. A. Waterman, and D. B. Lenat (1983). *Building expert systems.* Massachusetts, USA: Addison-Wesley.

Hevner, A. and S. Chatterjee (2010). *Design Research in Information Systems*, Volume 22 of *Integrated Series in Information Systems.* Boston, MA: Springer US.

Hevner, A. R., S. T. March, J. Park, and S. Ram (2004). Design science in information systems research. *MIS quarterly 28*(1), 75–105.

Hirst, E. (1980). Review of Data Related to Energy Use in Residential and Commercial Buildings. *Management Science 26*(9), 857–870.

Honebein, P. C., R. F. Cammarano, and K. A. Donnelly (2009). Will smart meters ripen or rot? Five first principles for embracing customers as co-creators of value. *The Electricity Journal 22*(5), 39–44.

Houde, S., A. Todd, A. Sudarshan, J. Flora, and K. Armel (2011). Real-time feedback and electricity consumption: A field experiment assessing the potential for savings and persistence. *Submitted, July.*

Houwelingen, J. V. and W. V. Raaij (1989). The Effect of Goal-Setting and Daily Electronic Feedback on In-Home Energy Use. *Journal of Consumer Research 16*(1), 98–105.

Hutton, R. B., G. A. Mauser, P. Filiatrault, and O. T. Ahtola (1986). Effects Consumer Behavior : Approach of Knowledge A Feedback on Consumption Experimental. *Journal of Consumer Research 13*(3), 327–336.

Hwang, M. and J. Lin (1999). Information dimension information overload and decision quality. *Journal of Information Science 25*(3), 213–218.

IEA (2013). Medium-Term Renewable Energy Market Report. Technical report, International Energy Agency, Paris, France.

IEA (2014). Energy Efficiency Indicators: Essentials for Policy Making. Technical report, International Energy Agency, Paris, France.

International Energy Agency (2012). IEA Response System for Oil Supply Emergencies. Technical report, IEA, Paris.

Ipakchi, A. and F. Albuyeh (2009, mar). Grid of the future. *IEEE Power and Energy Magazine 7*(2), 52–62.

Ipeirotis, P. G. (2010, March). Demographics of Mechanical Turk.

Iselin, E. R. (1988, January). The effects of information load and information diversity on decision quality in a structured decision task. *Accounting, Organizations and Society 13*(2), 147–164.

Iselin, E. R. (1990, January). Information, uncertainty and managerial decision quality: an experimental investigation. *Journal of Information Science 16*(4), 239–248.

ISO (2010). Ergonomics of human-system interaction – Part 210: Human-centred design for interactive systems.

Ito, M., R. Uda, S. Ichimura, K. Tago, T. Hoshi, and Y. Matsushita (2004). A method of appliance detection based on features of power waveform. In *CCECE 2003 - Canadian Conference on Electrical and Computer Engineering. Toward a Caring and Humane Technology (Cat. No.03CH37436)*, pp. 291–294. IEEE Comput. Soc.

Iyengar, S. S. and M. R. Lepper (2000). When choice is demotivating: Can one desire too much of a good thing? *Journal of Personality and Social Psychology 79*(6), 995–1006.

Jacoby, J. (1977, November). Information Load and Decision Quality: Some Contested Issues. *Journal of Marketing Research 14*(4), 569.

Jin, Y., E. Tebekaemi, M. Berges, and L. Soibelman (2011, May). Robust adaptive event detection in non-intrusive load monitoring for energy aware smart facilities. In *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 4340–4343. IEEE.

Kamb, M. L., F. Rhodes, T. Hoxworth, J. Rogers, A. Lentz, C. Kent, R. MacGowen, and T. A. Peterman (1998). What about money? Effect of small monetary incentives on enrollment, retention, and motivation to change behaviour in an HIV/STD prevention counselling intervention. The Project RESPECT Study Group. *Sexually transmitted infections 74*(4), 253–255.

Kaplan, B. and D. Duchon (1988, dec). Combining Qualitative and Quantitative Methods in Information Systems Research: A Case Study. *MIS Quarterly 12*(4), 571.

Kaufmann, S., K. Künzel, and M. Loock (2013, feb). Customer value of smart metering: Explorative evidence from a choice-based conjoint study in Switzerland. *Energy Policy 53*, 229–239.

Kempton, W. and L. Layne (1994). The consumer's energy analysis environment. *Energy Policy 22*(10), 857–866.

Kohavi, R., R. M. Henne, and D. Sommerfield (2007, August). Practical guide to controlled experiments on the web. In *Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining - KDD '07*, New York, New York, USA, pp. 959. ACM Press.

Kolter, J. and M. Johnson (2011). REDD: A public data set for energy disaggregation research. In *proceedings of the SustKDD workshop on Data Mining Applications in Sustainability*, pp. 1–6.

Kuckart, U., S. Rädiker, and A. Rheingans-Heintze (2006). Umweltbewusstsein in Deutschland 2006 - Ergebnisse einer repräsentativen Bevölkerungsumfrage. Technical report, Bundesministerium für Umwelt, Naturschutz und Reaktorsicherheit (BMU).

Laughman, C., R. Cox, S. Shaw, S. Leeb, L. Norford, and P. Armstrong (2003, March). Power signature analysis. *IEEE Power and Energy Magazine 1*(2), 56–63.

Lee, W., G. Fung, H. Lam, F. Chan, and M. Lucente (2004). Exploration on load signatures. In *International Conference on Electrical Engineering (ICEE)*, Sapporo, Japan, pp. 1–5.

Liang, J., S. K. K. Ng, G. Kendall, and J. W. M. Cheng (2010a, April). Load Signature Study-Part I: Basic Concept, Structure, and Methodology. *IEEE Transactions on Power Delivery 25*(2), 551–560.

Liang, J., S. K. K. Ng, G. Kendall, and J. W. M. Cheng (2010b, April). Load Signature Study-Part II: Disaggregation Framework, Simulation, and Applications. *IEEE Transactions on Power Delivery 25*(2), 561–569.

Lin, G., S. Lee, and J. Hsu (2009). Sensing from the panel: Applying the power meters for appliance recognition. In *Proceedings of the 14th Conference on Artificial Intelligence and Applications*.

Loock, C.-M., T. Staake, and F. Thiesse (2013). Motivating Energy-Efficient Behavior with Green IS: An Investigation of Goal Setting and the Role of Defaults.

Luo, D., L. Norford, S. Leeb, and S. Shaw (2002). Monitoring HVAC equipment electrical loads from a centralized location- methods and field test results. *ASHRAE Transactions 108*(1), 841–857.

Maan, S., B. Merkus, J. Ham, and C. Midden (2010, dec). Making it not too obvious: the effect of ambient light feedback on space heating energy consumption. *Energy Efficiency 4*(2), 175–183.

Mackay, J. M., S. H. Barr, and M. G. Kletke (1992, May). An Empirical Investigation of the Effects of Decision Aids on Problem-Solving Processes. *Decision Sciences 23*(3), 648–672.

Magat, W. A., J. W. Payne, and P. F. Brucato Jr (1986). How important is information format? An experimental study of home energy audit programs. *Journal of Policy Analysis and Management 6*(1), 20–34.

Malhotra, N. K. (1982, March). Information Load and Consumer Decision Making. *Journal of Consumer Research 8*(4), 419.

Martello, S. and P. Toth (1990). An exact algorithm for large unbounded knapsack problems. *Operations Research Letters 9*, 15–20.

Marvin, S., H. Chappells, and S. Guy (1999). Pathways of smart metering development: shaping environmental innovation. *Computers, Environment and Urban Systems 23*(2), 109–126.

Mattle, P., M. Aigner, E. Schmautzer, and L. Fickert (2011). Smart Plug - Konzept für ein intelligentes Lastmanagementsystem für Einzelverbraucher. In *Energieversorgung 2011: Märkte um des Marktes Willen?*, pp. 96–97. Institut für Energiesysteme und Elektrische Antriebe der TU Wien.

McCalley, L. T. and C. J. H. Midden (2002). Energy conservation through product-integrated feedback: The roles of goal-setting and social orientation. *Journal of Economic Psychology 23*(5), 589–603.

McDougall, G. H. G., J. D. Claxton, J. R. B. Ritchie, and C. D. Anderson (1981, December). Consumer Energy Research: A Review. *Journal of Consumer Research 8*(3), 343.

McMakin, A. H., E. L. Malone, and R. E. Lundgren (2002). Motivating residents to conserve energy without financial incentives. *Environment and Behavior 34*(6), 848–863.

Melville, N. P. (2010, March). Information systems innovation for environmental sustainability. *MIS Quarterly 34*(1), 1–21.

Menanteau, P. and M. Colombier (1997). Energy Efficiency Labelling for Appliances: a New Use of an Old Instrument in the French Region 'Nord/Pas de Calais'. In *The Energy Efficiency Challenge: Proceedings of the 1997 ECEEE Summer Study.*

Microsoft (2011). Microsoft Hohm Service Discontinuation. Available online at: http://tdworld.com/smart-energy-consumer/microsoft-hohm-service-be-discontinued-2012, Last accessed: 2015-04-07.

Midden, C. J. H., F. G. Kaiser, and L. Teddy McCalley (2007, March). Technology's Four Roles in Understanding Individuals' Conservation of Natural Resources. *Journal of Social Issues 63*(1), 155–174.

Murugesan, S. (2008). Harnessing Green IT: Principles and Practices. *IT Professional 10*(1), 24–33.

Myers, M. D. and M. Newman (2007, jan). The qualitative interview in IS research: Examining the craft. *Information and Organization 17*(1), 2–26.

Nadel, S. (2012). The Rebound Effect: Large or Small? Technical report, American Council For An Energy Efficient Economy.

Neenan, B. (2009, February). Residential electricity use feedback: A research synthesis and economic framework. Final Report 1016844, Electric Power Research Institute, 3420 Hillview Avenue, Palo Alto, California 94304-1338.

Nelson, R. R. and P. H. Cheney (1987, December). Training End Users: An Exploratory Study. *MIS Quarterly 11*(4), 547–559.

Norford, L., S. Leeb, D. Luo, and S. Shaw (1999). Advanced electrical load monitoring: A wealth of information at low cost. Technical report, Massachusetts Institute of Technology, Massachusetts, NE.

Norford, L. K. and S. B. Leeb (1996). Non-intrusive electrical load monitoring in commercial buildings based on steady-state and transient load-detection algorithms. *Energy and Buildings 24*(1), 51–64.

Nunes, N. J., L. Pereira, F. Quintal, and M. Berges (2011). Deploying and evaluating the effectiveness of energy eco-feedback through a low-cost NILM solution. In *Persuasive'2011*, Columbus, OH USA.

Oppenheimer, D. M., T. Meyvis, and N. Davidenko (2009, July). Instructional manipulation checks: Detecting satisficing to increase statistical power. *Journal of Experimental Social Psychology 45*(4), 867–872.

Paetz, A.-g., B. Becker, W. Fichtner, H. Schmeck, and F. D. Methods (2011). Shifting electricity demand with smart home technologies - an experimental study on user acceptance. In *30th USAEE/IAEE North American Conference Online Proceedings*, Volume 1, pp. 1–19.

Palensky, P. and D. Dietrich (2011, August). Demand Side Management: Demand Response, Intelligent Energy Systems, and Smart Loads. *IEEE Transactions on Industrial Informatics 7*(3), 381–388.

Parker, D., D. Hoak, J. Cummings, and F. Solar (2010). Pilot Evaluation of Energy Savings and Persistence from Residential Energy Demand Feedback Devices in a Hot Climate. In *2010 ACEEE Summer Study on Energy Efficiency in Buildings*, Washington, USA, pp. 245–259. ACEEE Summer Study on Energy Efficiency in Buildings.

Parker, D., D. Hoak, A. Meier, and R. Brown (2006). How much energy are we using? Potential of residential energy demand feedback devices. In *Proceedings of the 2006 Summer Study on Energy Efficiency in Buildings*, Pacific Grove, CA.

Parson, O. (2014). *Unsupervised Training Methods for Non-intrusive Appliance Load Monitoring from Smart Meter Data.* Forthcoming, University of Southampton.

Patel, S., S. Gupta, and M. Reynolds (2010). The design and evaluation of an end-user-deployable, whole house, contactless power consumption sensor. In *Proceedings of the 28th international conference on Human factors in computing systems*, pp. 2471–2480. ACM.

Patel, S. N., T. Robertson, J. A. Kientz, M. S. Reynolds, and G. D. Abowd (2007). At the Flick of a Switch: Detecting and Classifying Unique Electrical Events on the Residential Power Line. In *UbiComp 2007: Ubiquitous Computing: 9th International Conference, UbiComp 2007, Innsbruck, Austria, September 16-19, 2007. Proceedings*, Volume 4717, pp. 271. Springer.

Payne, J. W. (1982). Contingent decision behavior. *Psychological Bulletin 92*(2), 382–402.

Payne, J. W., J. R. Bettman, and E. J. Johnson (1993). *The adaptive decision maker.* New York, NY, US: Cambridge University Press.

Peffers, K., T. Tuunanen, M. A. Rothenberger, and S. Chatterjee (2007). A design science research methodology for information systems research. *Journal of Management Information Systems 24*(3), 45–77.

Perry, N. C., M. W. Wiggins, M. Childs, and G. Fogarty (2012). Can reduced processing decision support interfaces improve the decision-making of less-experienced incident commanders? *Decision Support Systems 52*(2), 497–504.

Petersen, J. E., V. Shunturov, K. Janda, G. Platt, and K. Weinberger (2007). Dormitory residents reduce electricity consumption when exposed to real-time visual feedback and incentives. *International Journal of Sustainability in Higher Education 8*(1), 16–33.

Pierce, J., C. Fan, D. Lomas, G. Marcu, and E. Paulos (2010). Some consideration on the (in)effectiveness of residential energy feedback systems. In *Proceedings of the 8th ACM Conference on Designing Interactive Systems - DIS '10*, New York, NY, US, pp. 244–247. ACM Press.

Pihala, H. (1998). *Non-intrusive appliance load monitoring system based on a modern kWh-meter* (356 ed.). Espoo, Finland: VTT Publications.

Raw, G. and D. Ross (2011, August). Energy Demand Research Project: final analysis. Technical report, Office of Gas and Electric Markets, London, UK.

Reiner, U., M. Kahl, T. Leibfried, F. Allerding, and H. Schmeck (2010). MeRegioMobil-Labor - Entwicklungsumgebung für zukünftige Smart-homes. *Proceedings of the VDE Kongress Leipzig*.

Ritchie, J. R. B. and G. H. G. McDougall (1985). Designing and marketing consumer energy conservation policies and programs: Implications from a decade of research. *Journal of Public Policy & Marketing 4*(1), 14–32.

Rodgers, J. and L. Bartram (2011, December). Exploring ambient and artistic visualization for residential energy use feedback. *IEEE transactions on visualization and computer graphics 17*(12), 2489–97.

Röpke, I. (1999). The dynamics of willingness to consume. *Ecological Economics 28*(3), 399–420.

Rüdenauer, I., U. Eberle, and R. Grieß hammer (2006). Ökobilanz und lebenszykluskostenrechnung wäschewaschen. *Vergleich des Waschens bei durchschnittlichen Temperaturen mit Waschen bei niedrigen Temperaturen. Freiburg, Hamburg: {Ö}ko-Institut e. V.*

Ruzzelli, A. G., C. Nicolas, A. Schoofs, and G. M. P. O'Hare (2010, June). Real-Time Recognition and Profiling of Appliances through a Single Electricity Sensor. In *2010 7th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks (SECON)*, pp. 1–9. IEEE.

Saitoh, T., Y. Aota, T. Osaki, R. Konishi, and K. Sugahara (2008). Current sensor based non-intrusive appliance recognition for intelligent outlet. In *The 23rd International Technical Conference on Circuits/Systems, Computers and Communications*, Yamaguchi, Japan, pp. 349–352.

Schenk, K. D., N. P. Vitalari, and K. S. Davis (1998, June). Differences between novice and expert systems analysts: what do we know and what do we do? *Journal of Management Information Systems 15*(1), 9–50.

Schön, D. A. (1983). *The Reflective Practitioner.* New York, NY, US: Basic Books.

Schultz, P. W., M. Estrada, J. Schmitt, R. Sokoloski, and N. Silva-Send (2015, sep). Using in-home displays to provide smart meter feedback about household electricity consumption: A randomized control trial comparing kilowatts, cost, and social norms. *Energy 90*, 351–358.

Schwartz, B. (2004). *The paradox of choice: Why more is less.* New York, NY, US: HarperCollins.

Schwartz, T., S. Denef, G. Stevens, L. Ramirez, and V. Wulf (2013, April). Cultivating energy literacy. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '13*, New York, New York, USA, pp. 1193. ACM Press.

Seal, B. and M. McGranaghan (2010). Accuracy of digital electricity meters. Technical report, Electric Power Research Institute, Palo Alto, CA.

Seligman, C., J. M. Darley, and L. J. Becker (1978, apr). Behavioral approaches to residential energy conservation. *Energy and Buildings 1*(3), 325–337.

Shadish, W. R., T. Cook, and D. Campbell (2002). *Experimental and quasi-experimental designs for generalized causal inference.* Boston, MA, US: Houghton Mifflin Co.

Shaw, R., M. Attree, T. Jackson, and M. Kay (2009, August). The value of reducing distribution losses by domestic load-shifting: a network perspective. *Energy Policy 37*(8), 3159–3167.

Shim, J., M. Warkentin, J. F. Courtney, D. J. Power, R. Sharda, and C. Carlsson (2002). Past, present, and future of decision support technology. *Decision Support Systems 33*(2), 111–126.

Shove, E. (2004). Changing human behaviour and lifestyle: A challenge for sustainable consumption? In L. Reisch and I. Rø pke (Eds.), *The Ecological Economics of Consumption*, Chapter 6. Cheltenham, UK: Edward Elgar Publishing Inc.

Shumway, R. H. and D. S. Stoffer (2011). *Time Series Analysis and Its Applications.* Springer Texts in Statistics. New York, NY: Springer New York.

Simmhan, Y., S. Aman, B. Cao, M. Giakkoupis, A. Kumbhare, Q. Zhou, D. Paul, C. Fern, A. Sharma, and V. Prasanna (2011). An informatics approach to demand response optimization in smart grids. *NATURAL GAS 31*, 60.

Simon, H. (1972). Theories of bounded rationality. *Decision and organization 1*(1), 161–176.

Sipe, B. and S. Castor (2009). The Net impact of Home Energy Feedback Devices. In *Energy Program Evaluation Conference*, Portland, OR, pp. 341–351.

Socolow, R. H. (1978, April). The twin rivers program on energy conservation in housing: Highlights and conclusions. *Energy and Buildings 1*(3), 207–242.

Sonderegger, R. C. (1978, April). Movers and stayers: The resident's contribution to variation across houses in energy consumption for space heating. *Energy and Buildings 1*(3), 313–324.

Speier, C. and M. Morris (2003). The influence of query interface design on decision-making performance. *Mis Quarterly 27*(3), 397–423.

Srinivasan, D., W. Ng, and A. Liew (2006, January). Neural-Network-Based Signature Recognition for Harmonic Source Identification. *IEEE Transactions on Power Delivery 21*(1), 398–405.

Staats, H., P. Harland, and H. Wilke (2004). Effecting durable change a team approach to improve environmental behavior in the household. *Environment and Behavior 36*(3), 341–367.

Staats, H. J., A. P. Wit, and C. Y. H. Midden (1996). Communicating the greenhouse effect to the public: Evaluation of a mass media campaign from a social dilemma perspective. *Journal of Environmental Management 46*(2), 189–203.

Strauss, A. and J. Corbin (1998). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (2 ed.). Thousand Oaks, CA, USA: Sage Publications.

Strbac, G. (2008, December). Demand side management: Benefits and challenges. *Energy Policy 36*(12), 4419–4426.

Strengers, Y. (2011). Designing eco-feedback systems for everyday life. In *Proceedings of the 2011 annual conference on Human factors in computing systems - CHI '11*, New York, NY, US, pp. 2135–2144. ACM Press.

Suzuki, K., S. Inagaki, T. Suzuki, H. Nakamura, and K. Ito (2008, August). Nonintrusive appliance load monitoring based on integer programming. In *2008 SICE Annual Conference*, pp. 2742–2747. IEEE.

Sweeney, J. L. (1993). Chapter 17 Economic theory of depletable resources: An introduction. In Allen V. Kneese and J. L. Sweeney (Eds.), *Handbook of Natural Resource and Energy Economics* (1 ed.), Volume 3, pp. 759–854. Elsevier B.V.

Taylor, P. (2007). Energy Use in the New Millennium - Trends in IEA Countries. Technical report, International Energy Agency (IEA).

Thø gersen, J. and F. Ölander (2003, September). Spillover of environment-friendly consumer behaviour. *Journal of Environmental Psychology 23*(3), 225–236.

Thuvander, L., P. Meiling, K. Andersson, and A. Nilsson (2012). Energivisualisering via display: Förändras beteendet när hyresgästerna har möjlighet att följa sin elförbrukning?

Todd, P. and I. Benbasat (1991). An experimental investigation of the impact of computer based decision aids on decision making strategies. *Information Systems Research 2*(2), 87–115.

Todd, P. and I. Benbasat (1992, September). The Use of Information in Decision Making: An Experimental Investigation of the Impact of Computer-Based Decision Aids. *MIS Quarterly 16*(3), 373.

Todd, P. and I. Benbasat (1994). The Influence of Decision Aids on Choice Strategies: An Experimental Analysis of the Role of Cognitive Effort. *Organizational Behavior and Human Decision Processes 60*(1), 36–74.

Todd, P. and I. Benbasat (2000, January). Inducing compensatory information processing through decision aids that facilitate effort reduction: an experimental assessment. *Journal of Behavioral Decision Making 13*(1), 91–106.

Tribus, M. and E. Mcirvine (1971). Energy and Information. *Scientific American 225*(3), 179 – 188.

Ueno, T., F. Sano, O. Saeki, and K. Tsuji (2006, February). Effectiveness of an energy-consumption information system on energy savings in residential houses based on monitored data. *Applied Energy 83*(2), 166–183.

Urquhart, C., H. Lehmann, and M. D. Myers (2009, may). Putting the âĂŸtheory' back into grounded theory: guidelines for grounded theory studies in information systems. *Information Systems Journal 20*(4), 357–381.

van Dam, S. S., C. a. Bakker, and J. D. M. van Hal (2010, oct). Home energy monitors: impact over the medium-term. *Building Research & Information 38*(5), 458–469.

Van Houwelingen, J. and W. Van Raaij (1989). The effect of goal-setting and daily electronic feedback on in-home energy use. *Journal of consumer research*, 98–105.

Venkatesh, V., S. Brown, and H. Bala (2013). Bridging the QualitativeâĂŞQuantitative Divide: Guidelines for Conducting Mixed Methods Research in Information Systems.

Vessey, I. (1994, August). The effect of information presentation on decision making: A cost-benefit analysis. *Information & Management 27*(2), 103–119.

Wagner, G. (2010). *Waschmittel Chemie, Umwelt, Nachhaltigkeit.* Weinheim: Wiley-VCH.

Waldrop, M. (1990, May). Learning to drink from a fire hose. *Science 248*(4956), 674–675.

Wallenborn, G., M. Orsini, and J. Vanhaverbeke (2011, March). Household appropriation of electricity monitors. *International Journal of Consumer Studies 35*(2), 146–152.

Wang, W. and I. Benbasat (2009, June). Interactive decision aids for consumer decision making in E-commerce: the influence of perceived strategy restrictiveness 1. *MIS Quarterly 33*(2), 293–320.

Watson, R. and M. Boudreau (2011). *Energy informatics.*

Watson, R., M.-C. Boudreau, and A. Chen (2010). Information Systems and Environmentally Sustainable Development: Energy Informatics and New Directions for the IS Community.

Watson, R. T., J. Howells, and M.-C. Boudreau (2012). *Green Business Process Management*. Berlin, Heidelberg: Springer Berlin Heidelberg.

Weiser, M. and J. Brown (1996). Designing calm technology. *PowerGrid Journal*, 1–5.

Weiss, M. (2010). emeter: Stromverbrauchsfeedback auf basis eines pervasive energy monitoring systems. In *Energieinformatik 2010*.

Weiss, M., A. Helfenstein, F. Mattern, and T. Staake (2012, March). Leveraging smart meter data to recognize home appliances. In *2012 IEEE International Conference on Pervasive Computing and Communications*, pp. 190–197. IEEE.

Weiss, M., C. Loock, and T. Staake (2010). Evaluating mobile phones as energy consumption feedback devices. *Mobile and Ubiquitous Systems: Computing, Networking, and Services*, 63–77.

Weiss, M., F. Mattern, T. Graml, T. Staake, and E. Fleisch (2009, nov). Handy feedback. In *Proceedings of the 8th International Conference on Mobile and Ubiquitous Multimedia - MUM '09*, New York, New York, USA, pp. 1–4. ACM Press.

Weiss, M., T. Staake, F. Mattern, and E. Fleisch (2012). PowerPedia: changing energy usage with the help of a community-based smartphone application. *Personal and Ubiquitous Computing*, 1–10.

White, A. L. (1993). Accounting for pollution prevention. *EPA Journal 19*, 23–25.

Wilhite, H., A. Hoivik, and J. G. Olsen (1999). Advances in the use of consumption feedback information in energy billing: The experiences of a norwegian energy utility. In *Proceedings from the ACEEE 1998 Summer Study on Energy Efficiency in Buildings*, Paris. European Council for an Energy Efficient Economy: ECEEE 1999 Summer Study on Energy Efficiency in Buildings.

Wilhite, H. and R. Ling (1995, jan). Measured energy savings from a more informative energy bill. *Energy and Buildings 22*(2), 145–155.

Winett, R. A., J. W. Hatcher, T. R. Fort, I. N. Leckliter, S. Q. Love, A. W. Riley, and J. F. Fishback (1982). The effects of videotape modeling and daily feedback on residential electricity conservation, home temperature and humidity, perceived comfort, and clothing worn: Winter and summer. *Journal of Applied Behavior Analysis 15*(3), 381–402.

Wood, G. and M. Newborough (2003, sep). Dynamic energy-consumption indicators for domestic appliances: environment, behaviour and design. *Energy and Buildings 35*(8), 821–841.

Yim, D. (2011). Tale of Two Green Communities: Energy Informatics and Social Competition on Energy Conservation Behavior. In *AMCIS 2011 Proceedings - All Submissions*.

Zeifman, M. (2012, February). Disaggregation of home energy display data using probabilistic approach. *IEEE Transactions on Consumer Electronics 58*(1), 23–31.

Zeifman, M. and K. Roth (2011, February). Nonintrusive appliance load monitoring: Review and outlook. *IEEE Transactions on Consumer Electronics 57*(1), 76–84.

Zoha, A., A. Gluhak, M. A. Imran, and S. Rajasegarar (2012, January). Non-intrusive load monitoring approaches for disaggregated energy sensing: a survey. *Sensors (Basel, Switzerland) 12*(12), 16838–16866.

Zou, Z., K.-J. Li, R. Li, and S. Wu (2011, January). Smart Home System Based on IPV6 and ZIGBEE Technology. *Procedia Engineering 15*, 1529–1533.