Stanford University Final Technical Report

Project Title: Large-Scale Energy Reductions through Sensors, Feedback, & Information Technology

Award: DE-AR0000018

<table>
<thead>
<tr>
<th>Award:</th>
<th>DE-AR0000018</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lead Recipient:</td>
<td>Stanford University</td>
</tr>
<tr>
<td>Project Title:</td>
<td>Large-Scale Energy Reductions through Sensors, Feedback, &amp; Information Technology</td>
</tr>
<tr>
<td>Program Director:</td>
<td>Dr. Timothy Heidel</td>
</tr>
</tbody>
</table>
| Principal Investigator: | Dr. James Sweeney (Phase II)  
Dr. Byron Reeves (Phase I) |
| Contract Administrator: | Nicole Pobuta (Stanford)  
Paul Basola (ARPA-e) |
| Date of Report: | March 1, 2017 |
| Reporting Period: | January 1, 2010 – December 1, 2016 |
ABSTRACT

Smart meters and related sensing technologies promise that energy information will play a key role in reducing energy use. However, poorly designed interactions with energy information, jeopardize these goals and reduce the efficacy of billion dollar utility infrastructure investments. The current problems are numerous: sensor information is complex and dull, incentives are inappropriate, interactions with energy information are poorly designed to modify behavior, and social context is ignored. These problems all involve the intersection of human behavior and technology.

The goal of this research initiative was to develop a comprehensive human-centered solution that leverages the widespread diffusion of energy sensors to significantly reduce and shift energy use. Our major innovation was the creation of a transformative system that combines behavioral techniques with human-centered design, computation, and technology to affect energy use behavior. The work involved a collaboration of Stanford University researchers and energy industry leaders to establish a new concentration in energy and human behavior. Our group conducted research, built systems, and tested solutions in the field.

Our initiative had three parts: (1) technology, including an extensible energy communications network to enable future innovation in home area networks; a software platform, to enable behavioral programs to be implemented at scale through a Living Laboratory; and algorithms to advance the areas of energy disaggregation, segmentation, and automation; (2) behavioral interventions to reduce and shift energy use, and (3) data evaluation and modeling approaches that applied economic and social network analysis techniques to data. The behavioral interventions included media (interaction design, social networking, games and feedback interfaces), incentive (behavioral economic programs) and community (schools, NGO’s, utility and social organizations).

Keywords: smart grid, smart meters, sensors, energy behavior, behavior, nudges, incentive, algorithm, disaggregation, demand side management, segmentation, communication protocol, community based program, girl scouts, game, browser application, facebook, twitter, diffusion, google, transport, insinc, appliance, ARPA-E, Stanford

Please use the following citation for this report:

ACKNOWLEDGEMENTS

This work was supported by the U.S. Advanced Research Projects Agency for Energy (ARPA-E) and the California Energy Commission Energy Research and Development Division. This report is the final report for the Advanced Research Projects Agency for Energy.

This work was conducted by Stanford University’s Center for Integrated Facility Engineering, Design for Change, H-Star Institute, Precourt Energy Efficiency Center and the Prevention Research Center. The information from this project contributes to Energy Research and Development Division’s Buildings End Use Energy Efficiency Program.

Stanford’s Precourt Energy Efficiency Center (PEEC) and Human Sciences and Technologies Advanced Research Institute (H-STAR) would like to acknowledge the work and support of the many individuals and organizations that contributed to the research and results of the Stanford Energy Behavior Initiative.

The support of the U.S. Advanced Research Projects Agency for Energy (ARPA-E) is gratefully acknowledged – Timothy Heidel, Mark Hartney, and Jonathan Burbaum, Program Directors; Scott Litzelman, Scientific Support Contractor. The support of the California Energy Commission’s Public Interest Energy Research program is also gratefully acknowledged - David Weightman, Contract Manager.

The support of our partners and collaborators is gratefully acknowledged: Bidgely Inc., Bonsai Development Corp., the City of Mountain View, DraftFCB, Google, Home Energy Analytics, Kuma Games, the Land Transport Authority of Singapore, National University of Singapore, People Power, PG&E, the Town of Hillsborough, University of Electro-Communications in Tokyo, and Venrock.

We also acknowledge tremendous intellectual and other support from colleagues and advisors too numerous to list here.

We thank the many talented researchers who contributed to this initiative:

Principal Investigators
• Byron Reeves
• James Sweeney

Project Director
• Carrie Armel

Project Manager
• June Flora

Core Investigators
• Banny Banerjee
• Tom Robinson
Other Faculty Investigators
• Hamid Aghajan
• Nicole Ardoin
• Martin Fischer
• Abby King
• Phil Levis
• Sam McClure
• Andrew Ng
• Ram Rajagopal
• Balaji Prabhakar
• Jeff Shrager
• Greg Walton
• John Weyant

Post Doctoral Fellows and Stanford Affiliates
• Eric Heckler
• Zico Kolter
• Luke Morton
• Hilary Schaffer-Boudet
• Annika Todd
• Gireesh Shrimali
• Dave Voelker

Graduate and Undergraduate Students

*For further information regarding this initiative please contact Carrie Armel* kcarmel@stanford.edu
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td></td>
<td>ii</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td></td>
<td>i</td>
</tr>
<tr>
<td>TABLE OF CONTENTS</td>
<td></td>
<td>i</td>
</tr>
<tr>
<td>EXECUTIVE SUMMARY</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>CHAPTER 1: Introduction</td>
<td></td>
<td>7</td>
</tr>
<tr>
<td>1.1</td>
<td>Innovativeness and Uniqueness of the Stanford Initiative</td>
<td>7</td>
</tr>
<tr>
<td>1.2</td>
<td>Defining Energy Technology to Include Human Behavior</td>
<td>8</td>
</tr>
<tr>
<td>1.2.1</td>
<td>‘Attention to’ Versus ‘Processing of’ Energy Information</td>
<td>8</td>
</tr>
<tr>
<td>1.2.2</td>
<td>Borrowing Lessons from Other Behavior Change Domains</td>
<td>9</td>
</tr>
<tr>
<td>1.3</td>
<td>Overview of the Stanford Initiative</td>
<td>9</td>
</tr>
<tr>
<td>1.3.1</td>
<td>Technology</td>
<td>10</td>
</tr>
<tr>
<td>1.3.2</td>
<td>Behavioral Interventions (“Programs”)</td>
<td>11</td>
</tr>
<tr>
<td>1.3.3</td>
<td>Evaluation and Modeling</td>
<td>12</td>
</tr>
<tr>
<td>1.3.4</td>
<td>The Projects</td>
<td>13</td>
</tr>
<tr>
<td>CHAPTER 2: Technology</td>
<td></td>
<td>17</td>
</tr>
<tr>
<td>2.1</td>
<td>Open Extensible Communication Network</td>
<td>17</td>
</tr>
<tr>
<td>2.2</td>
<td>Stanford Energy Services Platform (ESP)</td>
<td>20</td>
</tr>
<tr>
<td>2.3</td>
<td>Algorithms</td>
<td>22</td>
</tr>
<tr>
<td>2.3.1</td>
<td>Residential Energy Analytics: Segmentation, Targeting, Thermal Response, &amp; Building Energy Efficiency</td>
<td>22</td>
</tr>
<tr>
<td>2.3.2</td>
<td>Advanced Learning Automation</td>
<td>28</td>
</tr>
<tr>
<td>2.3.3</td>
<td>Disaggregation</td>
<td>30</td>
</tr>
<tr>
<td>2.3.3.1</td>
<td>Disaggregation Technical and Policy Survey Paper</td>
<td>30</td>
</tr>
<tr>
<td>2.3.3.2</td>
<td>Residential Energy Disaggregation Dataset (REDD)</td>
<td>34</td>
</tr>
<tr>
<td>2.3.3.3</td>
<td>Disaggregation Algorithms</td>
<td>37</td>
</tr>
<tr>
<td>2.4</td>
<td>Target Behaviors</td>
<td>40</td>
</tr>
<tr>
<td>2.4.1</td>
<td>Energy Behaviors Taxonomy</td>
<td>40</td>
</tr>
<tr>
<td>Chapter</td>
<td>Title</td>
<td>Pages</td>
</tr>
<tr>
<td>---------</td>
<td>-------</td>
<td>-------</td>
</tr>
<tr>
<td>2.4.2</td>
<td>Identifying Opportunities for Dramatic Energy Reductions in Residences</td>
<td>43</td>
</tr>
<tr>
<td>CHAPTER 3: Behavioral Interventions</td>
<td></td>
<td>46</td>
</tr>
<tr>
<td>3.1</td>
<td>Media Interventions</td>
<td>46</td>
</tr>
<tr>
<td>3.1.1</td>
<td>Multiplayer Online Game</td>
<td>46</td>
</tr>
<tr>
<td>3.1.2</td>
<td>Collective Action Feedback Interface</td>
<td>50</td>
</tr>
<tr>
<td>3.1.3</td>
<td>The Impact of Vivid Messages on Saving Behavior related to Hot Water Use</td>
<td>53</td>
</tr>
<tr>
<td>3.1.4</td>
<td>Motivationally Framed Facebook Energy Applications</td>
<td>57</td>
</tr>
<tr>
<td>3.2</td>
<td>Incentive Interventions</td>
<td>60</td>
</tr>
<tr>
<td>3.2.1</td>
<td>Nudges for Energy Efficiency through an online Appliance Calculator</td>
<td>60</td>
</tr>
<tr>
<td>3.2.2</td>
<td>Transportation Lottery</td>
<td>65</td>
</tr>
<tr>
<td>3.3</td>
<td>Community Based Interventions</td>
<td>67</td>
</tr>
<tr>
<td>3.3.1</td>
<td>Girl Scout “Girls Learning Energy and Environment” (GLEE) Program</td>
<td>67</td>
</tr>
<tr>
<td>3.4</td>
<td>Integrative Project</td>
<td>70</td>
</tr>
<tr>
<td>3.4.1</td>
<td>HowPower</td>
<td>70</td>
</tr>
<tr>
<td>CHAPTER 4: Evaluation and Modeling</td>
<td></td>
<td>74</td>
</tr>
<tr>
<td>4.1</td>
<td>Google Power Meter Evaluation</td>
<td>74</td>
</tr>
<tr>
<td>4.2</td>
<td>Social Media Analytics through Twitter Explorer</td>
<td>77</td>
</tr>
<tr>
<td>4.3</td>
<td>Diffusion Modeling of Behavioral Interventions</td>
<td>80</td>
</tr>
<tr>
<td>CHAPTER 5: Summary and Project Outputs</td>
<td></td>
<td>84</td>
</tr>
<tr>
<td>5.1</td>
<td>Project Outputs and Impacts</td>
<td>84</td>
</tr>
<tr>
<td>5.1.1</td>
<td>Summary</td>
<td>84</td>
</tr>
<tr>
<td>5.1.2</td>
<td>Additional Product Outputs</td>
<td>93</td>
</tr>
<tr>
<td>5.1.3</td>
<td>Potential Impacts</td>
<td>95</td>
</tr>
<tr>
<td>5.2</td>
<td>Conclusions</td>
<td>97</td>
</tr>
<tr>
<td>GLOSSARY</td>
<td></td>
<td>98</td>
</tr>
<tr>
<td>REFERENCES</td>
<td></td>
<td>99</td>
</tr>
</tbody>
</table>

**TABLE OF FIGURES**
TABLE OF FIGURES

Figure 1. The Stanford Engine. ................................................................. 10
Figure 2. Behavioral intervention categories for energy efficiency projects. .......... 12
Figure 3. Powernet hardware and high resolution data. ................................ 19
Figure 4. Energy Services Platform. ......................................................... 22
Figure 6. Visdom frequent load shapes. .................................................... 24
Figure 7. Visdom customer numbers. ....................................................... 25
Figure 8. Visdom comparison of thermal responses. ..................................... 26
Figure 9. Visdom energy efficient frontier, showing three different buildings on the frontier. 27
Figure 10. Automation activity recognition map. ........................................ 29
Figure 11. Disaggregation. ........................................................................ 31
Figure 12. REDD monitoring device collecting data from a circuit breaker panel. ... 36
Figure 13. Performance of different types of disaggregation algorithms on hourly data. 39
Figure 14. Energy Behavior Taxonomy. ..................................................... 43
Figure 15. Multiplayer online game landing page and action game. .................... 48
Figure 16. Collective action experiment website. ......................................... 51
Figure 17. Immersive virtual environment energy experiment conditions. ............ 56
Figure 18. Facebook applications, Kidogo and Power Tower. .......................... 59
Figure 19. Appliance Calculator browser application, accessed through Google Adwords. 62
Figure 20. Insinc transportation lottery website and game of chance. ................. 66
Figure 21. Girl Scouts Girls Learning Energy and Environment (GLEE) website and badges. 68
Figure 22. Google Powermeter evaluation. ................................................ 75
Figure 23. Twitter Explorer interface for data collection and analysis of energy tweets. 79
Figure 24. Visualization of illustrative results from diffusion modeling "experiments". 83

TABLE OF TABLES

Table 1. Overview of Projects. .................................................................. 14
Table 2. Uses of appliance specific feedback. .......................................... 30
Table 3. Summary of projects, partners, and product outputs. .................... 85
The Stanford Energy Behavior Initiative was created to forge effective ways of integrating behavioral science into smart grid technology in order to achieve meaningful energy savings. Phase I lasted from 2010-2013, and a subset of projects were extended into a Plus-Up Phase II period from 2014-2016. The initiative aimed to address two timely energy problems. First, significant low-cost energy reductions could be made in the residential and commercial sectors by individual and collective operational changes and energy efficiency investment decisions, but these savings have only partially been achieved. Second, billions of dollars are being spent to install smart meters, yet the energy saving and financial benefits of this infrastructure – without careful consideration of the human element – will not reach its full potential. By strategically marrying these problems, and leveraging now pervasive internet technology and behavioral science learnings, we believed that we could address these problems – specifically through a constellation of 20 projects that comprised a living laboratory.

The following product outputs were produced and results found, in each of the corresponding cluster research or program intervention areas noted in italics. Complementary workshops and other support efforts were also conducted. The investigators that are listed were the official project leads, but many projects had several key investigators.

Technical Platform

There are several aspects of the technical platform: hardware and a communications network; an Energy Services Platform (ESP) to streamline the creation of behavioral programs; and several types of algorithms to perform segmentation, automation, and disaggregation. We also compiled energy saving actions for use in recommendation engines, both recommendations that are typically recommended by energy and utility organizations currently, as well as innovative ones for example derived from other cultures and time periods.

1. Levis, in collaboration with a many others, helped create an open standard for TCP/IP in home area networks (HANs) as well as an open-source reference implementation of the standard for others to copy, extend, re-use, and improve. This open technology leverages the internet and will provide greater freedom in data collection, representation, storage, and communication between devices of different manufacturers, all leading to innovations and improvements in human interfaces to sensor-actuator networks. As a second deliverable, Levis et al. created an extensible and open-source sensing hardware platform for devices that monitor the power of individual electronics, such as computers and monitors.

2. Armel guided the effort to create the Stanford Energy Services Platform (ESP) in order to support the data collection, computational, and web presence needs of several of the behavioral interventions, as well as to benefit the future work of outside groups.

Segmentation Algorithms
3. Rajagopal and Fischer et al. developed algorithms to segment customers based on their energy consumption patterns over time using large residential and commercial smart meter data sets. They used this information to strategically and cost-effectively match customers with energy saving programs. They also developed algorithms to improve monitoring and rating of building energy efficiency in order to better assess energy efficiency programs.

Automation Algorithms

4. Aghajan et al. developed adaptive machine learning algorithms that utilized sensor data to improve lighting automation on the dimensions of both user preferences and energy savings.

Disaggregation Algorithms

5. Three projects on disaggregation, or the separation of a whole home electricity use into appliance-specific data to guide people on where they should take action, covered the scope of foundational work necessary to jumpstart significant interest and research in this space. Armel et al. wrote a comprehensive technical and policy oriented survey paper that has received tens of thousands downloads to date from the PEEC website.
6. Kolter et al. created a data set for disaggregation developers to train and test their algorithms; this data set has been extensively utilized.
7. Kolter developed disaggregation algorithms using sparse coding methods to advance the state of the art.

Target Actions

8. Boudet et al. identified approximately 250 energy saving actions, or behaviors, and created a taxonomy to support recommendation engines and behavioral programs within and outside of the Stanford Initiative.
9. Armel et al. collected energy saving actions from other cultures and throughout history as inspiration for modern day energy saving innovations, and their energy savings were quantified to provide an opportunity map for future design efforts. Many of these may require further development before incorporation into recommendation engines.

Behavioral Programs

Social-ecological models of behavior change, a dominant class of theories in public health, hold that the use of multiple types of programs or interventions are more effective than one, because they complement and reinforce one another. Several of the Initiative’s projects involved developing and testing media, incentive, and community-based program interventions to assess the effectiveness of these programs and the behavior change techniques embedded within them.

While some projects employed an approach typical in the field of public health or in utility pilots in which the effectiveness of multi-faceted programs was evaluated (e.g., a randomized controlled trial assessed the effectiveness of a program that incorporated multiple behavior change techniques versus a no treatment or status quo control), other projects followed an experimental psychology approach in which there were multiple experimental conditions that
were identical except for individual variables that were modified in each to investigate their specific impact).

Media

10. Reeves et al. with design and technical support from Kuma Games created the Power House online game that incorporated real world energy data into some of the game play, leveraged social competition, and retrained habits through reinforcement. Results of a laboratory and also a field trial suggest the game’s effectiveness in saving energy and energy related behaviors. In 2015, over 15,000 unique users played the game.

11. Walton et al. created a basic resource consumption feedback interface that included energy or water saving recommendations for two respective Northern California communities. He found that framing an individual’s efforts as part of the community effort could sometimes increase but other times decrease conservation efforts.

12. Bailenson et al. created an immersive virtual shower world and measured its impact on energy related water consumption behavior, with results suggesting that vivid visualizations of energy consumption (e.g., amount of coal instead of KWh) are more important than the personalization afforded by avatars.

13. Banerjee et al. created three Facebook applications to match the range of motivations exhibited by individuals: Power Tower is social in that it allows one to collaborate with others in a Tetras-like puzzle where pieces are granted based on multiple participants’ energy savings; Kidogo is affective in that it allows one to compute their real world energy savings and then microfinance individuals in developing countries based on this savings; and Powerbar is cognitive in that it primary displays energy energy feedback data. As an outgrowth of this work, an online course was developed and is offered through the Stanford Professional Development Program.

Incentive

14. McClure et al. created the Appliance Calculator application which has been used by over 60,000 people through Google adwords. Those exposed to behavioral “nudges” pursued appliances that used 10-20% fewer kWh than their counterparts in the control condition. By testing changes in the interface we found, contrary to expectations, that projecting out cost savings over time does not appear to prompt more energy efficient refrigerator browsing, whereas simply changing the default sort order to put the most efficient appliances on top does – this suggests that simply implementing the most effective behavior change techniques may be a more effective strategy than the traditional route of analyzing the underlying cause of a problem then trying to address it.

15. Prabhakar et al. created the Insinc lottery-like transportation incentive program, significantly shifting participants to off-peak travel or public transit, with an enrollment of over 15,000 Singaporeans to date with an average of 7.5% of trips shifted off peak.

Community

16. Robinson and Ardoin et al. created the Girls Learning Energy and Environment (GLEE) Girl Scout community based program; such programs can scale quickly by tapping into
established networks and providing close support from peers. In a pilot trial, Girl Scouts in troops randomly assigned to the residential energy program significantly increased their residential energy-saving behaviors by 49% following the intervention, and 27% persisting after more than seven months of follow-up, compared to controls. Their participation also facilitated behavior change in their parents. Regarding scaling activities, by end of the ARPA-e funding, 216 Girl Scout troop leaders in 30 states had registered for the GLEE online training course, and 72 leaders were working in or had worked through the course – potentially reaching nearly 900 girls and their families.

Integrative

17. Armel developed an integrative program - combining learnings from the other projects and additional high-impact approaches – focused on scaling feasibility, depth of savings, advanced EM&V capabilities, and marketplace sustainability. The core element of the project is the HowPower browser application that utilizes innovative approaches such as the use of a dynamic concierge; a focus on tips-to-action; an entertaining and frictionless experience; and triaging to determine which actions and products are right for the user. A five-month trial with a rudimentary version of the browser application garnered 15,000 unique users from Google Adwords.

Evaluation and Modeling

Three projects focused on developing methods for evaluating the effectiveness of energy programs, and modeling the effectiveness of interventions to guide future work.

18. Houde et al. applied an analysis derived from economic methods to quantify the effectiveness of Google Powermeter in saving energy. The approach could help inform approaches for legitimizing energy savings from behavioral programs within utilities and government agencies.
19. Russell et al. developed the Twitter Explorer which tracks all tweets containing any of ~150 energy related words, in order to track and map conversation changes pertinent to this topic across the internet’s online social network.
20. Shrager et al. developed a simulation tool that allows one to model the energy savings of behavioral interventions according to parameters such as time, behavioral technique used, and social network distance and type. This tool could serve as a foundation for developing similar but more sophisticated tools enhanced with additional parameters and empirical data that could eventually lessen time and cost of developing future interventions through predictive modeling.

In summary, the Stanford Energy Behavior Initiative achieved an impressive array of work spanning 20 projects overseen by thirteen investigators across ten departments and five schools, in collaboration numerous students and outside partner organizations. The Initiative was divided into software and analytics, behavioral programs, and evaluation tools projects. The central objective was to develop the components that would support a system aimed at utilizing smart meter and other sensor data, communication technologies, and behavioral approaches, to
achieve significant energy savings.

Many publications and product outputs were created from these projects; some outputs were scaled, and others could be used in applied technologies and programs for wider impact. Such demand-side management work has a wide array of benefits: reduced GHG emissions; reduced environmental impact; reduced system capacity requirements; increased energy security. Behavioral programs such as ours also provide the benefits of increased consumer appeal of energy efficiency actions, and potentially increased economic activity and improved smart grid efficiency and benefits. Depending on the program and user demographics, we would expect savings between the approximately 1.5% achieved by Opower’s mailer report and 20% estimated to be achievable over a population (Laitner; Stern; McKenzie).
A blank page is inserted to insure Chapter 1 starts on an odd number page. Blank pages are not labeled.
CHAPTER 1: Introduction

1.1 Innovativeness and Uniqueness of the Stanford Initiative

The most unique contribution of this initiative is to redefine energy technology to highlight the critical role of human behavior in reducing energy use. Our interdisciplinary group of Stanford researchers, the projects and innovations, and the resulting benefits, all have a common emphasis on the critical but frequently overlooked role of human behavior in reducing energy use. Without a complete consideration of human behavior, technical innovations will be at best incomplete and at worst misused or unused.

Here is a caricature of the current energy story that motivates our initiative:

“I use energy in my home but it’s invisible. I don’t consume it directly but only via things I want like light, heat and refrigeration. I rarely think about the energy I’m using, and most of my use is habitual and unconscious. The amount of energy I use is registered on a meter that’s out of sight, unintelligible, and read by someone else. I only get feedback about my energy use in the form of monthly bills that present complex data that are a month old, and are boring and impersonal. When information is provided to me about how and why I should change my behavior, it is also boring and impersonal and often not even applicable to my situation. Even when I understand it, I rarely act.”

This story is common throughout the United States. It describes what may be the worst possible system for promoting reduced energy use: complex, dated, boring and impersonal information presented via inaccessible and unintelligible devices that fail to engage or be personally relevant, and that describe behaviors mysteriously linked to a global problem that is personally distant and difficult to define.

Behavior change can fail at any point in the story, and it does. Consequently, we believe that systematically applying theory and methods from the behavioral sciences will produce greater benefits for energy related programs. A system of behavioral strategies to collectively target each step in the process of behavior change is needed to seize a significant opportunity to influence energy and climate issues in America. The system that we have worked on, made possible by feedback from new energy sensors, includes motivational interventions informed by the best new research about human behavior and were designed to make energy use and energy savings visible, immediate, compelling and even fun.

In the following pages, we redefine the energy efficiency problem by emphasizing its relationship to human behavior, noting different dimensions of the problem and also similarities to behavior problems in other social science arenas. Next, we provide additional background on the link between energy technology and human behavior, describe the four major parts of the Stanford Initiative, and then describe in detail each of the projects under those four parts.
1.2 Defining Energy Technology to Include Human Behavior

Energy technology is most often defined as devices that supply or improve energy efficiency. These devices most often succeed or fail independent of human action required to make them work (e.g., more efficient solar cells). Many devices, however, are critically dependent on human behavior to make them effective. This is true for the significant national investment in smart sensors that monitor residential energy use. Smart sensors (e.g., smart meters, plug load monitors and programmable controllable thermostats) provide information to people who are expected to monitor and understand the information provided, and to change behaviors accordingly. Consequently, human behavior, along with all of the cognitive, social and emotional constraints, must be considered a critical component of designing and deploying smart sensing systems.

Much more energy information will soon be available, accentuating the importance of human behavior in determining technology success. In addition to the first installations of smart meters and plug load monitors, wireless sensing technologies will be available for gas and hot water, and for transportation in the form of sensors that quantify miles per gallon, mode of transportation and number of trips. Clearly, the impact of energy information on behavior will play a critical role in energy efficiency in the next decade. Our project initially focused on smart meter and home area network (HAN) sensing technologies because of their impending rollouts. Our work, however, can be transferred to transportation, gas, and water sensors.

1.2.1 ‘Attention to’ Versus ‘Processing of’ Energy Information

The increased availability of sensing information creates two different information problems, both addressed in our initiative. Information processing theories differentiate selective attention to information (e.g., choosing to pay attention to smart meter data versus something else) from information processing (e.g., interpreting, remembering and changing attitudes and behavior based on the information selected).

We consider both problems. The increasing availability of energy information in already crowded information environments will make attention to energy information difficult. Consequently, our initiative considered how to increase attention by using techniques that increase engagement and interest in information (e.g., multiplayer games, incentive programs and social networks).

The increasing quantification of energy use provides a complimentary opportunity to influence information processing and retention. Substantial evidence in psychology shows that quantification of behavior promotes behavior change because it makes visible what is otherwise easy to forget or ignore, and because it informs people about whether their actions have effected change.

Furthermore, interventions can usefully depend on the quantification of behavior; for example, in incentive programs, energy markets, competitions, visualizations, games and social networking, automated appliance controls and behavior change guidance. Our initiative explicitly considered how the quantification of energy information can influence behavior.
1.2.2 Borrowing Lessons from Other Behavior Change Domains

Although many of our approaches are new to the field of energy efficiency, the consideration of human behavior for other social problems is well established. Large-scale behavior change programs, similar in scope to the projects in our initiative, have been shown to alter behavior in other areas with equally difficult information challenges, including health practices related to cardiovascular disease, smoking and drug use, community political participation, and sexual practices (Rice & Atkin, 2001; Singhal, Cody, Rogers, & Sabido, 2004). Indeed, the term “energy behavior change” comes from “health behavior change.” Psychologists attribute the success of these interventions to the application of proven behavioral principles such as engagement, modeling, and self-efficacy (Bandura, 1986, 1997). A distinguishing feature of our work was the collection of researchers with extensive expertise in other behavior change areas, and a commitment to borrow expertise from the larger behavior change literature that has already been completed with significant federal research investment.

1.3 Overview of the Stanford Initiative

We are at a unique point in history with enormous opportunity. On the one hand, wireless sensors - smart meters, home area networks, gas, transportation, and water sensors - that enable the quantification of energy usage are becoming pervasive. On the other hand, web enabled devices are also pervasive, meaning that web enabled computers and phones can deliver programs to help individuals reduce energy use.

The interfacing of these systems with an appropriate engine provides an opportunity for a Living Laboratory at a scale that has societal impact. Such an engine was funded by the Advanced Research Projects Agency for Energy, California Energy Commission, and Stanford University. This Stanford Engine is comprised of approximately 20 projects overseen by 15 faculty and spanning five schools, five centers, and ten departments ranging from computer science and electrical engineering, civil and environmental engineering, and economics, to psychology, communications, education, and behavioral epidemiology.

In summary, our initiative’s goal was this: To create comprehensive human-centered solutions that leveraged the widespread diffusion of energy sensors in order to significantly reduce and shift energy use. To address this goal, we did the following: (1) developed supporting technology, including an extensible energy communications network to enable future innovation in home area networks; a software platform, for use by Stanford and outside behavioral interventions; and algorithms to advance the areas of energy disaggregation, segmentation, and automation; (2) developed interventions in three categories that promote energy behavior change (media, incentive, and community based), and (3) developed data evaluation and modeling approaches that applied economic and social network analysis techniques to intervention data.

In the next sections, we review details of each of these three activities.

In the following pages we provide further descriptions of the initiative’s clusters. Table 1 lists of all the projects by cluster, followed by more detailed descriptions of all the projects in the rest of this report. The initiative mostly focused on residential buildings, but the aim is to extend into other related areas such as transportation, commercial buildings, and water.
1.3.1 Technology

The technology category of our project includes an open communications network, a software platform to support behavioral interventions, and analytics. Regarding the first, we expanded opportunities for flexibility and innovation in home area networks with an open and extensible communications protocol that can capitalize on unforeseen behavior change opportunities – such as the ability to provide energy feedback to users more quickly. An open technology will provide greater freedom in data collection, representation, and storage, leading to innovations and improvements in human interfaces to sensor-actuator networks.

Regarding the software platform, large interdisciplinary groups often result in a collection of thematically related projects that unfortunately do not utilize the technical work performed by one another. Our group, however, was motivated by the prospect that a common infrastructure will allow projects to be more influential, extensible and larger than they could be if conducted independently. The common infrastructure we developed included shared software, databases, and computing services that created technical economies in the conduct of large-scale field research. Through this platform, the effectiveness of interventions can be evaluated quickly, easily, inexpensively and at scale. This is possible for two reasons: (1) experimental manipulations can be generated and tracked automatically because they can be implemented and delivered via electronic media that all projects will share, and (2) objective measures of behavior change can be collected automatically by sensors, and aggregated in databases that are shared across projects.

Regarding the analytics, several types of algorithms, or computational problem solving procedures implemented by us in computer software, were developed to perform segmentation, disaggregation, and automation. Segmentation (of whole house electricity consumption data) algorithms were developed to group customers based on their energy consumption patterns using large residential and commercial smart meter data sets. This information can be used to strategically and cost-effectively match customers with energy saving and demand response programs. Three projects related to disaggregation, or the
separation of a whole home energy signal into appliance specific data to guide people on where they should take action; these included algorithm development, as well as a technology and policy review paper, and the collection of a dataset to speed work in this space. Adaptive machine learning algorithms were also developed that utilized sensor data to improve lighting automation on the dimensions of both user preferences and energy savings (Project 7).

There were additionally two projects that identified target actions for use in the interventions – through software or other recommendation systems. One identified and created a taxonomy of traditional energy saving actions or behaviors, based on characteristics such as cost, energy impact, who performs the action, and so forth. A second collected energy saving actions from other cultures and throughout history as inspiration for modern day energy saving innovations, and their energy savings were quantified to provide an opportunity map for design efforts and recommendation systems in the future.

1.3.2 Behavioral Interventions (“Programs”)

We developed and tested the effectiveness of several types of behavior change interventions. Figure 2 names these types and example topics in each. (Specific projects in each of the categories are reviewed in Table 1.) Given technology’s central role in this Initiative, it is depicted at the heart of the diagram below, and for the rest of this report Technology is broken out separately from the other types of interventions. However, note that the energy data produced by the sensors and technology projects is also used by the other three types of behavioral interventions. The arrows between the types of interventions below indicate that projects are mutually reinforcing and can both stand alone (e.g., a single test of one media strategy) and be combined in a coordinated system (e.g., different media strategies that incorporate policies). The general framework of multiple intervention types and their synergistic effects was adapted from the socio-ecological theory mentioned above which is widely used in public health.
1.3.3 Evaluation and Modeling

Beyond the evaluation of the aforementioned intervention projects, three projects focused on developing methods for evaluating the effectiveness of energy programs, and modeling the effectiveness of interventions to guide future work. One project used economic methods to quantify the effectiveness of Google Powermeter in saving energy; the approach could help inform approaches for legitimizing energy savings from behavioral programs within utilities and government agencies. A second project developed Twitter Explorer which collects and analyzes tweets, in order to track and map conversation changes pertinent to specified topics such as energy across the internet’s online social network. A third project developed a simulation tool that allows one to model the energy savings of behavioral interventions according to parameters such as time, behavioral technique used, and social network distance and type, with the goal of serving as a foundation for other tools that could eventually lessen time and cost of developing future interventions through predictive modeling.
1.3.4 The Projects

There are 19 projects summarized in Table 1 below, as well as the new integrative Project 20 that is underway with additional funding from ARPA-E. In the table we show the project number, its category (technology, intervention, evaluation and modeling), its name, deliverables, investigators and their Stanford University departments as well as partners. The rest of the report provide more details on each of the projects, and concludes with impacts and recommendations.
Table 1. Overview of Projects.

<table>
<thead>
<tr>
<th>Project # - Report Section</th>
<th>Project Name</th>
<th>Goals</th>
<th>Investigators</th>
</tr>
</thead>
<tbody>
<tr>
<td>Technology</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Open Extensible Communication Network</td>
<td>Development of an extensible HAN protocol to enable innovation</td>
<td>Levis, Kazandjieva</td>
</tr>
<tr>
<td>2</td>
<td>Stanford Energy Services Platform (ESP)</td>
<td>Database + analytics for experimentation at Stanford and beyond</td>
<td>Armel, Reeves</td>
</tr>
<tr>
<td>Algorithms</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Energy Consumption Forecasts</td>
<td>Segmentation of commercial and residential energy use data, models that forecast building energy consumption to guide utility interventions</td>
<td>Fischer, Rajagopal, Albert, Kavousian</td>
</tr>
<tr>
<td>4</td>
<td>Advanced Learning Automation</td>
<td>Software to customize HAN automation, specifically automation based on action recognition and individual preferences</td>
<td>Aghajan, Khalili, Chen</td>
</tr>
<tr>
<td>5</td>
<td>Disaggregation Technical and Policy Survey Paper</td>
<td>Assessment of the benefits of disaggregation, state of the art algorithms and performance, and smart meter suitability for this data</td>
<td>Armel, Gupta, Shrimali, Albert</td>
</tr>
<tr>
<td>6</td>
<td>Residential Energy Disaggregation Dataset (REDD)</td>
<td>Data set collected for developers to build and test disaggregation algorithms</td>
<td>Kolter, Chadwick, Armel, Flora</td>
</tr>
<tr>
<td>7</td>
<td>Disaggregation Algorithms</td>
<td>Development of algorithms to identify appliance level energy information from whole home data stream</td>
<td>Ng, Kolter</td>
</tr>
<tr>
<td>Target Behaviors</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Energy Behavior Taxonomy</td>
<td>Catalog of energy behaviors, and their impact and barriers</td>
<td>Flora, Boudet, Roumpani, Armel</td>
</tr>
<tr>
<td></td>
<td>Interventions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>----------------</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>9</td>
<td>Identification of Innovative Energy Behaviors</td>
<td>Identification of innovative energy reducing behaviors and their potential impact</td>
<td>Armel, Cornelius, Ardoin, Plano, Bridgeland, Morton, Chang, Allen</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Interventions</th>
<th>Media Interventions</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>Multiplayer Online Game</td>
<td>Online game utilizing team competition</td>
<td>Reeves, Cummings, Scarborough</td>
<td>Communication Kuma Games</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Interventions</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td>Collective Action Feedback Interface</td>
<td>Web application that helps consumers monitor goals and compare energy use</td>
<td>Walton, Sparkman, Clark, Paunesku, Armel, Luo, Flora</td>
<td>Psychology Steve and Lisa Schmidt &amp; City of Mountain View</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Interventions</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>Visual Metaphors in a Virtual Immersive Environment</td>
<td>Development and evaluation of metaphors to make energy vivid and personal</td>
<td>Bailenson, Bailey, Flora, Armel, Voelker, Reeves</td>
<td>Communications, PEEC DraftFCB</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Interventions</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>13</td>
<td>Motivationally Framed Facebook Applications</td>
<td>Web interfaces to motivate energy reductions</td>
<td>Banerjee, Flora, Sahoo, Bhansali, Greenspan, Khakwana, Liptsey-Rahe, Madres, Manley, Omer, Rajendra, Scalammini, Wong, Stehly, Voelker</td>
<td>Mechanical Engineering, H-STAR</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Interventions</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>14</td>
<td>Appliance Calculator</td>
<td>Information and framing tools for guiding the purchase of energy efficient appliances and electronics</td>
<td>McClure, Houde, Armel</td>
<td>Psychology, MS&amp;E, PEEC Sears</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Interventions</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>Transportation Lottery</td>
<td>Utility program that stretches the motivational value of monetary incentives</td>
<td>Prabhakar, Merugu, Pluntke, Gomes, D. Mandayam, Yue, Atikoglu, Albert, Fukumoto, Liu, Wischik, Rama</td>
<td>EE National University of Singapore, Land Transport Authority of Singapore</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Community Intervention</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>16</td>
<td>Girl Scout “GLEE” Program</td>
<td>Curricula that increase engagement with sensor data and diffuse sensor use and energy saving actions to families</td>
<td>Robinson, Ardoin, Boudet, Flora, Armel</td>
<td>Pediatrics, Education, H-STAR, PEEC Girl Scouts, People Power</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Evaluation &amp; Modeling</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
</table>

15
<table>
<thead>
<tr>
<th>17</th>
<th>Google Powermeter Evaluation</th>
<th>Field trial of Google PowerMeter impact using analysis tools from economics</th>
<th>Houde, Sudarshan, Todd, Flora, Armel</th>
<th>MS&amp;E, PEEC Google</th>
</tr>
</thead>
<tbody>
<tr>
<td>18</td>
<td>Social Media Analytics with Twitter Explorer</td>
<td>Analysis of the “conversation” about energy efficiency and behavior on the web &amp; analytical tool for platform</td>
<td>Russell, Rubens, Flora</td>
<td>H-STAR University of Electro-Communications, Tokyo</td>
</tr>
<tr>
<td>19</td>
<td>Diffusion Modeling of Behavioral Interventions</td>
<td>A simulation model to predict diffusion in behavioral interventions</td>
<td>Shrager</td>
<td>Symbolic Systems</td>
</tr>
</tbody>
</table>

**Integrative Project**

| 20 | Integrative Project | An integration of some of the most effective pieces of the initial projects, as well as new components such as narrative and added diffusion strategies | Armel | Habitable, Free Range Studios, possible others |

CS = Computer Science  
E-IPER = Emmett Interdisciplinary Program in Environment and Resources  
EE = Electrical Engineering  
FS = Freeman Spogli Institute for International Studies at Stanford  
H-STAR = Human Sciences and Technologies Advanced Research Institute  
ISB = Indian School of Business  
MS&E = Management Sciences and Engineering  
PEEC = Precourt Energy Efficiency Center
CHAPTER 2:
Technology

There are several aspects of the technical platform: hardware and a communications network; an Energy Services Platform (ESP) to streamline the creation of behavioral programs; and several types of algorithms to perform segmentation, automation, and disaggregation. Energy saving target actions were also collected for recommendation systems.

2.1 Open Extensible Communication Network

Investigators: Phil Levis, Maria Kazandjieva

Background

A Home Area Network (HAN) is a type of local area network that facilitates communication and interoperability among digital devices present inside or within the close vicinity of a home (through wired connections or wirelessly), such as "smart" appliances", home security systems, lights, and audio-visual consoles. Computers and smart phones receive data from and can sometimes control these devices through communication with the network. From an energy perspective, a HAN can enable control from a distance or can be programmable; it can be used to turn things on and off, reduce energy voltage, or put electronic devices in “sleep” mode to reduce energy waste when spaces are unoccupied, or devices are not needed to continually be in “active” mode. HANs can also “learn” use and occupancy patterns such that devices use energy only when necessary.

It is important to expand opportunities for flexibility and innovation in Home Area Networks (HANs) with an open and extensible communications protocol that will allow for unforeseen behavior change opportunities that can lead to energy use reduction or increased energy efficiency. For example, if research indicates that high frequency, real-time traces of energy use are effective at changing user behavior, having an open and extensible underlying technology will make it easier to provide such data. An open technology will also provide greater freedom in data collection, representation, storage, and communication between devices of different manufacturers, all leading to innovations and improvements in human interfaces to sensor-actuator networks.

Objectives

In order for this opportunity to be available for the next stage of research and real world applications, Professor Levis aimed to design and implement an open-source network architecture for home area networks. The first major deliverable was open-source code that could be used on a wide range of sensor hardware. The second major deliverable was an extensible and open-source sensing hardware platform for smart meters. Together, these two provide a complete hardware and software solution that users and companies can build on.
Methods - Deliverable 1, Open-Source Code

HANs have been plagued and limited by numerous, overlapping, closed commercial standards. As open-source implementations of closed protocols are of only limited use, Professor Levis participated in an effort to define an open Internet standard, called an IETF request for comments, or RFC. This RFC would provide an open and free definition of how HAN devices should interoperate in order to communicate data with the larger Internet. The protocol described in the RFC is called RPL (“ripple”), or Routing Protocol for Low Power and Lossy Networks. Professor Levis would follow this protocol standardization with supporting and managing an open-source implementation of RPL in the TinyOS operating system, an operating system written by Professor Levis and used by tens of thousands of people worldwide. Together, this means there is now an open standard for TCP/IP in HANs as well as an open-source reference implementation of the standard for others to copy, extend, re-use, and improve.

Outcomes - Deliverable 1, Open-Source Code

The major deliverables were completed. The RPL RFC was ratified in March of 2012. There are now numerous implementations and interoperability tests at IETF meetings. The TinyOS implementation of RPL supports almost all of its major features. Using TinyRPL, a developer can quickly build a large-scale, multi-hop wireless mesh that self-organizes to support TCP/IP to every device and self-heals in response to wireless signal changes, adding new devices, and devices failing. This open-source implementation has been used in several research papers as a basis for exploring ways to improve RPL and has also been used by several companies for prototyping new products.

Methods - Deliverable 2, Sensing Hardware

Professor Levis had proposed building an open hardware platform for smart meters that would allow scientists to access data at a scale or fidelity that is difficult to achieve with existing commercial products. The necessary lead-time for hardware development given the safety certifications needed for such a device turned out to be too long for the results to be useful to other scientists in the project who might have benefitted from the data. Our reasonable time estimates for the project indicated that Professor Levis’ team would have working sensors as data collection began to wind down. The team therefore decided to build a simple and easier power plug meter that would measure devices rather than whole buildings. This plug meter would use the TinyOS operating system and RPL protocol to establish and test its design.

Outcomes - Deliverable 2, Sensing Hardware

The team developed a wireless power plug meter that automatically joins a self-assembling, ad-hoc wireless mesh network to deliver data to collection points. They open-sourced the design, which has been used by several follow-on efforts by other groups as a basis for their own power plug designs. The team deployed a network of 200 such meters in the Gates Computer Science building at Stanford for over two years to obtain long-term, fine grained power draw
measurements of the building’s computing systems. This extensive data collection allowed the team to publish detailed data at a scale orders of magnitude than other, similar efforts, as well as establish the basic methodologies one should follow to measure computing energy. The lead student on this project has now graduated and has been working with several green computing companies in the Bay Area to write future energy standards for computing systems.

![PowerNet Hardware and High Resolution Data](image)

**Figure 3.** Powernet hardware and high resolution data. The foreground inset displays a wireless power plug meter that was built to automatically join a self-assembling, ad-hoc wireless mesh network; the deployed network of 200 meters allowed the team to publish detailed data at a scale orders of magnitude greater than other similar efforts (background), and is informing future energy standards for computing systems.

**Future work**

This work should have significant real world impact moving forward. These efforts helped establish the first Internet standard for HANs, which is being adopted by industrial consortia such as WirelessHART and ZigBee. The open source implementation of the protocol in TinyOS provides a starting point for its demonstration and improvement, both through research and engineering. The embedded wireless plug meters demonstrated low-cost ways to densely measure computing energy and has established techniques on how to collect and analyze such data. In summary, the open-source thrust of this work should facilitate real-world impact, as a ready-made hardware and software solution that can be easily extended will reduce the cost of entry for new companies and lower the bar for innovation.
2.2 Stanford Energy Services Platform (ESP)

Investigators: K. Carrie Armel, Byron Reeves, Jay Bartels, Andrew Davidson

Background

Many researchers are designing and testing feedback interfaces, as evidenced by proceedings from conferences such as the Behavior, Energy and Climate Change Conference (BECC) and the Computer Human Interaction Conference (CHI), and the dozens of publications on the topic over the past few decades (for reviews see Darby, 2006; Neenan & Robinson, 2009; Ehrhardt-Martinez & Donnelly, 2010). However, we found through interviews that researchers typically build the various software services needed from scratch and in a piecemeal manner. As a result, valuable time and resources are wasted, and often, researchers are confined to running very small scale experiments (i.e., 10-30 people per studies, only a couple studies per year). A platform could be of great benefit to researchers, utilities, and third party developers in this space. This kind of software is not available off the shelf, and the system has not been built in full robustness because while the benefit is high, the amount of effort for any one group to build the system is very large.

Objectives

The Stanford ARPA-E Energy Services Platform (ESP) is an attempt to address this need for feedback interface software that is available for multiple parties to use. The platform provides software services including data collection, data cleaning and storage (e.g., sensor, click, and survey data), analytics (e.g., baselining, comparison to other users, disaggregation), graphing, recommendation systems, participant registration and assignment, and front-end display and email notifications suitable for performing experimental manipulations.

Method

The ESP was coded by Bonsai Development Corporation, informed by the needs of Stanford ARPA-E Initiative projects. Five Stanford behavioral programs have or are using the ESP: PowerDown, Power House, Kidogo, Appliance Calculator, and Girls Learning Energy and Environment (GLEE).

Outcomes

The ESP platform has been created, including a back-end with data and service layers, and a front-end presentation layer (see Table 1). The software engineering design emphasizes modularity, extensibility, security, and scalability.

The data layer includes data collection, storage, and retrieval. There are currently five logical data stores, containing different types of data: (1) energy data (the system currently collects, via a web collector, hourly electric and gas smart meter data being stored by utilities), (2) project-specific data (e.g., experiment details, investigator names for data access), (3) participant data (e.g., survey input, utility username/password), (4) website/application activity, (5) external data (e.g., weather, regional specific information like fuel type penetration or socio-
demographic characteristics), and (5) recommendations, including energy-saving actions and also energy efficient appliance recommendations.

The services layer is anything that requires computation, logic, or analytics, such as baselining, graphing, or disaggregation; a determination of recommendations; or participant registration and assignment. Several services were developed to group data in ways that could motivate behavior change. For example, we compare a user’s energy usage to their baseline, or to a groups’ energy usage, through numbers and graphs. The Appliance Recommendation System uses user inputs on characteristics of their current appliance, what state they live in, and preferences for new appliance attributes, in concert with the databases described above, in order to provide recommendations on whether one should upgrade to a new appliance, and, if so, which one.

The front-end of the system is the presentation layer, and refers to the device (e.g., computer, ipad, handheld), application type or medium (e.g., text messaging, widgets, flash animation, social media), and graphical user interface (GUI) characteristics (e.g., style and layout, type of content (i.e., which widgets?), as well as the actual content (the specific text, images, sounds)). For research and evaluation purposes, it is important that users across a broad range of technical capabilities have flexibility and control in the presentation layer. This is important because researchers need to be able to make front-end changes to the applications easily and see the results, so they can continuously iterate, and to create variations on a “parent” web-page so they can easily create different conditions for their experiments. With the ESP developers can control the presentation using various technologies, depending on their skill level and desire for customization. These technologies include:

• Software Developers Kit (SDK): The platform has an API that can be made available for users with programming knowledge that allows them near complete flexibility in designing and developing applications (e.g., mobile, website, Facebook), while accessing databases or services on the platform.

• Mash-Ups: These are easier to use than an API, but allows the same access as the API, it is so suitable for those with limited programming knowledge.

• Widgets: the platform also allows for users with no programming knowledge to create applications, for example, by selecting and then dragging and dropping pre-made modules or “widgets” in systems such as iGoogle, Google sites, and iWeb.

• Off the shelf packages: Various presentation oriented packages, also known as management systems (CMS) such as Drupal or Joomla, can be interfaced with our platform.
<table>
<thead>
<tr>
<th>Data Collection, Storage, and Retrieval</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Energy</strong></td>
</tr>
<tr>
<td>- From web collector</td>
</tr>
<tr>
<td>- [From Bigeye’s disaggregation system]</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Services</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Analytic</strong></td>
</tr>
<tr>
<td>Analytics</td>
</tr>
<tr>
<td>- Baselining</td>
</tr>
<tr>
<td>- Aggregation over a community</td>
</tr>
<tr>
<td>- Disaggregation</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Presentation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Platform API</td>
</tr>
</tbody>
</table>

Figure 4. Energy Services Platform. Functional architecture. Brackets indicate components not currently or fully implemented.

### 2.3 Algorithms

#### 2.3.1 Residential Energy Analytics: Segmentation, Targeting, Thermal Response, & Building Energy Efficiency

Investigators: Ram Rajagopal, Martin Fisher, June A. Flora, Jung Suk Kwac, Adrian Albert, Amir Kavousian, Jeff Wong

Partners: PG&E, Google
Utility demand-side programs are designed towards achieving set target kWh reductions in consumption for a given time horizon. Demand response programs aim at short time scale reductions, such as 10 min to one hour, while energy efficiency programs aim at longer time scale baseload reductions such as over the period of weeks to years. Success of a program is measured by the yield in the program: the ratio between projected reductions and achieved reductions. Typical yields in current utility programs are in the low range of 10-30%. The primary reason is due to customer enrollment challenges. The population of consumers can be segmented according to how the program benefits them as shown in Figure 1(a). There are four groups of customers. The customers with large positive benefits usually enroll in the program, but are a small fraction of the customer population. Similarly, the customers with large but negative benefits (group C) usually do not enroll as expected. More importantly, customers with small positive benefits (group A) fail to enroll. Moreover, customers with small negative benefits (group B) who could also achieve positive benefits by performing small and inexpensive changes in consumption patterns also do not enroll. Groups A and B form the majority of the population and need to participate if a program aims at high yields. Our goals within the ARPAe projects was to develop and test scalable methods and metrics for customer segmentation; develop algorithms for customer targeting for demand response programs, and predicting energy consumption thermal response.

**Partners and Data**

We utilize an anonymized large utility provided data-set with more than 250,000 PG&E customers with at least one year worth of hourly smart meter data recordings. We also used TED collected data on over 1,000 Google employees.

This was project was continued during the Plus Up funding period.

**Methods and Outcomes**

---

Figure 5. Visdom program management goals, and opportunities.
The sum of our methods, we title *Behavior Analytics*. We begin by building a basic methodology for characterizing an electricity customer. The methodology decomposes a customer’s consumption into daily load shapes. Load shapes are analyzed in aggregate to obtain a small number of typical loads shapes that characterize the whole population. These typical load shapes can be then utilized to build behavioral models for customers. We use the load shapes to examine features such as variability, amount of kWh consumed, and thermal response. Finally, we develop innovative methods to quantify the energy efficiency of buildings.

We describe four projects examining methods and outcomes of customer’s consumption segmentation, targeting, thermal response, and quantification of energy efficiency in buildings (See sections A-D below).

**A. Load Shape Clustering and Customer Segmentation**

![Figure 6. Visdom frequent load shapes. Six most frequent load shapes from the dictionary and their average daily kWh using hourly data of 125K customers, August 2010 - July 2011.](image)

The wide-spread deployment of smart meters has made available concrete information about customer consumption patterns or *load shapes*, including the magnitude and timing of their electricity consumption. To populate a dictionary of *representative* load shapes for a large data set, a hierarchical clustering algorithm has been developed to separate shapes that are “close” to each other into different groups. This massive data analytics is scalable and reveals various key information about the data, including the most frequent load shapes that characterizes a customer’s lifestyle, the number of peaks, peak times and peak durations. The load shape dictionary is used to segment customers according to extracted features such as entropy of shape code which measures the amount of variability or stability in consumption. Segmentation strategies can be developed to target customers for specific applications such as demand response (DR). For example, to target households for an automated DR program, the focus should be on heavy use and stable customers, in the appropriate time-based segment. So
segmentation can be an effective filter to reduce the number of eligible customers. This is particularly useful when the customer size is very large in a full-scale deployment. Also, it will be easier to forecast consumption at an individual level for stable customers, and harder for unstable customers since variability is larger in their usage patterns.

**B. Customer Targeting Strategies for Improving Program Efficiency**

![Graphs showing Pmax vs N for different time blocks](image)

**Figure 7. Visdom customer numbers.**

Minimum numbers of customers needed for different confidence levels to achieve a reduction target of 1MWh at four 1-hour blocks of time in CA zone 13 (hot summer climate).

The idea of demand side load management has been employed to help reduce energy consumption at peak hours and provide demand curtailment during periods of shortfall in renewable generation. To achieve these objectives, utilities are rolling out different energy programs, such as demand response (DR), to small and medium-sized customers. Recruiting a customer into an energy program can have significant costs, including market enrollment, event notification, and very often, the setup for enabling hardware at the customer premise.

Moreover, the power curtailment potential across customers varies significantly. So for utilities to achieve high yields (for example, high energy savings) in these energy programs, effective strategies need to be designed to identify and target the right customers. For instance, one should not first target customers who will unlikely be consuming large amounts of energy at peak hours for a peak load reduction program. The availability of smart meter data allows a close examination of the customers' potential for an energy program, so that they are chosen adequately to balance the magnitude of response potential (reward) and the uncertainty in the prediction (risk). In this project, we propose a methodology to utilize a large data set with 200K smart meters to select the right customers for a given level of enrollment expenditure. The focus is on the Smart AC program that curtails power consumption by increasing the temperature setpoints. The method combines a customer response forecast model estimated from data with a stochastic discrete optimization program that selects customers in order to minimize risk given
a desired level of curtailment response. The algorithm gives a minimum number of customers needed to achieve a certain energy reduction target and a list of customers who can achieve this goal with a certain confidence level (probability). This is a scalable selection algorithm and has been validated on a large data set.

A key measure of success for an energy program is its ability to consistently achieve its energy saving goal. So a demand response (DR) program (with incentives), combined with a targeting mechanism and appropriate promotion messages for specific customer segments, built upon a number of DR experiments, needs to eventually induce "steady-state demand response behavior". Such steady-state behavior is necessary in helping utilities measure and quantify the amount of "reliable load reduction" from demand response. This also allows a better understanding of DR availability to identify key customer characteristics (such as demographics, or environmental orientation) that drive energy sustainability in each program.

C. Load Flexibility and Management of Thermal Sensitive Load

![Distribution of Thermal Regimes](image)

Figure 8. Visdom comparison of thermal responses.
Comparison of the effective thermal responses of two users with respect to temperature, where positive (resp., negative) kWh/F indicates energy sensitivity for cooling (resp., heating).

Thermal-sensitive energy consumption, such as that from a heat, ventilation and air conditioning (HVAC) unit, accounts for 25% of the residential electricity usage in the U.S. and is the single largest block of the electricity budget for most customers. There has been tremendous interest and effort in developing methodologies to decompose individual consumption into a thermal-sensitive component and a non-thermal-sensitive base load. Models and algorithms for stacking individual energy consumption in such a structure provide a useful mechanism for understanding the flexibility in managing customer load, and subsequently making effective demand response and energy efficiency intervention decisions. We have developed a model of temperature response that is based on "thermal regimes" for separating the thermal-sensitive load from the thermal-insensitive base load. These "thermal regimes" are consumers' unobserved decisions to use their heating or cooling appliances. With this description of consumption, a system operator could manage the thermal usage component over the course of
a planning horizon (24 hours) for a large population of consumers by communicating actions (thermostat temperature settings) to them such that the aggregate load follows a desired profile. This knowledge allows scarce operational and marketing budgets to be allocated to the right consumers when DR programs are executed. Moreover, the features computed from the temperature response model could be used to predict actual characteristics; for example, the presence of large appliances is best predicted by consumption magnitude features, and lifestyle is predicted by the rate at which consumers switch between different consumption regimes. Energy efficiency programs can also be designed to offer the right incentives, such as rebates for efficient appliances, to the consumers whose appliance stock and lifestyle patterns are likely to sustain most long-term energy savings.

D. Ranking Energy Efficiency of Buildings

![Efficiency Frontier Graph](image)

Figure 9. Visdom energy efficient frontier, showing three different buildings on the frontier.

Quantifying building energy efficiency is essential to developing and monitoring energy efficiency programs. Existing methods for energy efficiency ranking are time-consuming (thermodynamic models), based on industry averages as opposed to observed operational data (simulation models), non-adjustable to specific situation of a building (Energy Star), or too simplistic (energy intensity comparison). Frontier methods, on the other hand, quantify the energy efficiency of buildings by forming an efficient frontier (i.e., best-practice technology) and comparing all buildings against that frontier. The efficient frontier specifies the lowest energy consumption observed at any Level of Service, which is defined as the utility that the users receive from using energy. This is an indicator of the utility (service, output) that a building provides to its occupants. Stochastic Energy Efficiency Frontier (SEEF) is a performance-based method based on frontier methods for ranking building energy efficiency. It specifies the best practice combination of energy consumption and the level of service. It offers several improvements over existing frontier methods, such as recognizing the random nature of energy consumption and treating energy efficiency as a random variable. It creates a ranking process that identifies a probability distribution of building efficiency ranks, instead of declaring a
building is more efficient than another deterministically. The method also uses actual data to estimate uncertainty of efficiency scores.

2.3.2 Advanced Learning Automation

Investigators: Hamid Aghajan, Amir Hossein Khalili, Chen Wu, Louis Chen

Background

While encouraging user’s proactive involvement in reducing their energy use is an important goal, a complementary strategy looks at how automation can improve efficiency or reduce energy waste with minimal user input; for example, automatically shutting off lights or appliances that are not in use or adjusting the temperature control in response to outside weather conditions. This approach has the potential to reduce energy use with little direct user effort, circumventing persistence issues in user behavior. In order for these techniques to be effective, however, they need to consider behavior. The vast majority of current home automation systems operate without such consideration, using one of two insufficient paradigms: 1) either the systems operate using fixed rules that fail to account for individual differences or 2) they require that users specify the operation rules themselves which is time consuming, unintuitive and may be easily ignored.

Objectives

The first aim was to create a solution that is an integrated approach based on ubiquitous sensing of the environment, and algorithms that will predict user behavior and automatically adapt. The second goal was to evaluate the system through in vivo testing – that is, testing in a space that is being used for real work or living functions. Deliverables for this project included building and implementing a model using real world data that predicts user behavior. Predictions from the modeling system are integrated into existing user interfaces, to display predicted future behaviors in a visually accessible manner.

Product Description and Study Methods

The automation was accomplished by using machine learning algorithms that address both user models and decision-making. Modeling algorithms use sensor data collected from a variety of sources. Such a model could learn from observations such that, for example, when a user is sitting in the living room in the evening, the TV and lights in that room are likely to be on. And it could also learn the probability that the user will transition to a different situation (such as going to sleep). We can then apply decision-making algorithms in order to prescribe system changes that limit energy use; for example, the system could predict that computer use is typically low while the user is watching TV, and it could power down the computer in another room. The key element is that both the modeling and decision making processes are adaptive and will adjust to a user's behavior without requiring manual entry of preferences.

Outcomes
To achieve adaptivity in providing services to the user, the system was found to need to support two functions: 1) sense the activity and state of the user, and 2) customize service to the user’s profile. To achieve this, three functional modules were developed and described in published papers. In the first module, behavior analysis of the user in a home environment was achieved based on multi-camera vision processing. In the second module, a user profile was defined and hierarchical reinforcement learning was employed as a technique to learn the user profile dynamically. The third module is a decision maker which employs the user profile to control services to maximize user comfort and utility. An automated light and TV control implementation using a network of wireless switches was developed based on detecting the location of a user and his pose with a number of cameras. A web-based user interface was developed to capture the user’s input about the automation setting and build a context-aware user profile, which was used to adapt the setting according to the user’s preferences. For additional details, see publications at http://peec.stanford.edu/energybehavior/projects/HAN_automation.php.

Figure 10. Automation activity recognition map.
Activity recognition map and corresponding video coverage for the HAN learning and automation software.

Future Work

On further development effort, algorithms can be created to detect complex but unobvious schedule regularities and automatically control devices in a home area network (HAN), and the behavior of which can change over time to adapt to the underlying changes in the user’s behavior or preferences or with seasonal changes. Other algorithms can be developed to track actual and expected energy use of appliances and suggest when appliances or electronics should be repaired or replaced. On commercialization efforts, we can share our findings with
commercial entities involved in home automation services and collaborate with them on developing test cases for actual user deployment settings.

## 2.3.3 Disaggregation

### 2.3.3.1 Disaggregation Technical and Policy Survey Paper

 Investigators: K. Carrie Armel, Abhay Gupta, Gireesh Shrimali, Adrian Albert

### Background

Appliance-specific feedback may be the most effective type of energy feedback data for the purposes of reducing energy consumption. Appliance specific feedback can achieve this in numerous ways, as seen in Table 2.

Table 2. Uses of appliance specific feedback.

<table>
<thead>
<tr>
<th>Benefits</th>
<th>Domain</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consumer</td>
<td>Residential Energy Use</td>
<td>Greater energy reductions from this type of feedback</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(a) Automated personalized recommendations (through auto-commissioning, fault detection, elucidating behavioral patterns, analysis of when and what type of new appliance to purchase based on current use, etc.), (b) personalized recommendations allow for personalized information to reduce barriers to energy efficient actions (e.g., mapped recommendations on where to purchase recommended items); enabling of additional/enhanced behavioral techniques (feedback, competition, visualizations, markets, incentives, etc.)</td>
</tr>
<tr>
<td></td>
<td>Commercial Energy Use</td>
<td>Similar application to residential; large untapped savings here</td>
</tr>
<tr>
<td>Research and Development</td>
<td>Appliance Innovation</td>
<td>Better data to (a) redesign appliances for energy efficiency, (b) improved standards, and (c) back up appliance energy efficiency marketing</td>
</tr>
<tr>
<td></td>
<td>Building Research and Design</td>
<td>Improved building simulation models to increase design and operational efficiency (commissioning and auto-commissioning)</td>
</tr>
<tr>
<td>Utility and Policy</td>
<td>Segmentation for Energy Efficiency Marketing</td>
<td>Strategic, specific, energy efficiency marketing</td>
</tr>
<tr>
<td></td>
<td>Program Evaluation</td>
<td>(a) Improved objectivity, sensitivity, and causal inference in</td>
</tr>
</tbody>
</table>
Many approaches for supplying appliance-specific feedback are costly or effortful because they depend on hardware installations. Another approach is disaggregation - the use of algorithms to break down an aggregate or whole-home energy signal into its component appliance/electronic contributions.

**Objectives**
In this comprehensive survey paper, we explain how appliance level data affords numerous benefits, and why using the algorithms in conjunction with smart meters is the most cost-effective and scalable solution for getting this data. We review disaggregation algorithms and their requirements, and evaluate the extent to which smart meters can meet those requirements. Research, technology, and policy recommendations are also outlined.

Methods

In addition to reviewing relevant academic literature, white papers, and manufacturer’s manuals, we interviewed dozens of professionals across relevant sectors, and included electrical engineers with relevant expertise as co-authors. These approaches were particularly helpful in determining disaggregation algorithms data requirements as well as the type of data that is currently available through deployed smart meters.

Outcomes

The work reviewed suggests that there are compelling reasons to pursue disaggregation, and that it may be possible to leverage existing or future smart meters so that appliance specific information can provide benefits at scale. The following are several specific recommendations for moving forward, which are elaborated on in the paper.

The following research and development activities are suggested, as well as fiscal support for these:

1. Improve disaggregation algorithms, in order to improve robustness, accuracy, and number of appliances identified by the algorithms, while reducing frequency, processing, and training requirements, and develop high-yield data compression algorithms. These improvements will move algorithms closer towards in commercial grade (i.e., consumer acceptable) products, and will result in algorithms improvements that will enhance their ability to facilitate energy reductions under more diverse hardware conditions and while consuming less energy to operate.

2. Develop a common data set that captures variability over appliances as well as operating conditions. One such data set has been created at Stanford (see REDD project). Most of the algorithms that have been developed to date have utilized data from only a couple of buildings and so training and testing on this new data set should significantly improve the accuracy of the algorithms across a more representative and diverse housing stock. It would also be beneficial to: (a) Establish performance metrics, such as common definitions of accuracy to enable the comparison of algorithms, so that the effectiveness of different algorithms can be compared. (b) Organize a competition, as has been done previously with algorithm development in other domains.

3. Facilitate testing of compression and disaggregation algorithms on actual smart meters, to evaluate performance.

4. Perform key behavioral research: Identify popular use cases and their information requirements, as this has relevance to data handling and consumer display requirements.
The following steps should be taken to improve data on existing meters. Regarding firmware upgrades, these are similar to software updates, and can be performed remotely and can be appended to routine updates so as to minimize cost.

1. Upgrade firmware to make reactive power available in addition to real power. This allows algorithms to disaggregate more devices.

2. Upgrade firmware to support data compression. Transmitting events/transitions instead of raw load profiles could significantly improve the frequency of data available to HAN devices, as bandwidth is currently the bottleneck.

Regarding future smart meter hardware and firmware, we recommend the following:

1. Be capable of 10-15 kHz frequency, which would only cost $5-10 more, but would likely enable a jump in accuracy and the number of appliances recognized. Also, improve wattage granularity. These modifications could garner additional energy saving opportunities from additional identification of devices (electronics are the fastest growing energy consumer in the residential sector) and recommendations.

2. Be capable of supporting disaggregation inside smart meters through firmware upgrades; this would avoid AMI or HAN network constraints on the amount of data that can be transmitted and thus disaggregated (resulting in identification of fewer devices).

3. Add or replace 802.15.4 based radio (used by ZigBee) with 802.11 (WiFi or low power WiFi) so that meters can communicate directly with the broadband routers, rather than require additional hardware.

Additional policy recommendations include:

1. Disaggregation developers should contribute use case specifications and requirements to the standards process so that other forthcoming communications technologies are better suited for disaggregation.

2. Institute policies to ensure that utilities enable the HAN communication interface (example ZigBee radios in the meters deployed in CA) soon, at a minimum beginning with pilots.

3. Institute policies, such as rebates, to make HAN gateways (that enable consumers to get real time data from their smart meter) effectively free to consumers.

4. Institute policies to ensure that utilities select HAN devices during pilots that allow consumers to access or share their data with any third party.

5. Federal agencies and PUCs should demand improved transparency about meter specifications, and enable universities to test real meters to establish actual constraints.

6. Utilities and regulatory agencies should expediently approve guidelines for addressing privacy issues, if they have not already.

Future Work

Smart meters, given their widespread roll-outs, and ability to circumvent cost and effort barriers, offer an opportunity for quick, sweeping market penetration of sensing hardware
required for disaggregation in a relatively short time frame. Given the benefits, we anticipate researchers, policy makers, and manufacturers will work together to realize the application of disaggregation with smart meters. To date, significant interest is evidenced by tens of thousands of downloads of this survey paper. Further, this work could clear the way for similar energy disaggregation work on gas, water, and transport, and the work could have significant implications for demand response program development and outcomes in the commercial as well as residential sector.

2.3.3.2 Residential Energy Disaggregation Dataset (REDD)
Investigators: J. Zico Kolter, Sarah Jo Chadwick, Larsen Plano, K. Carrie Armel, June Flora

Background
The previous project outlined the importance of disaggregation, as well as other energy data analytics, and identified the need for further algorithm development. Despite its potential, there are a number of obstacles that have hampered academic work on energy data analytics. Data for energy domains is typically collected by companies or utilities themselves, often at substantial cost, and there are several factors that make sharing this data difficult: 1) energy data at the level of individual homes or buildings has intrinsic privacy concerns: connecting even moderate frequency power measurements to a specific home or individual has the potential to reveal substantial information about that person; 2) the value of sharing data freely with the entire body of academic researchers can be unclear, especially when the data itself has substantial business value; and 3) the volume of raw data for many systems can often be quite large, and it is not clear how to widely share the data in a manageable fashion. Due to these factors, academic studies on energy data have often followed a less-than-ideal pattern: a study or analysis is conducted using data gathered from by a third party, but the data is not released and cannot be obtained by researchers looking to build on the data. The specific line of work that originally spawned our work, energy disaggregation and non-intrusive load monitoring, has followed this rough trajectory: despite work in this area for over 20 years, we know of no independent studies prior to the initial release of REDD that used a common data set.

Objectives
The Reference Energy Disaggregation Data Set (REDD) project collects such a data set and standardizes the data collection process. REDD is a free and publicly available energy data set, with common evaluation metrics, to be used by researchers and algorithm developers to develop algorithms designed to separate an aggregate or whole-home energy signal into its component appliance/electronic contributions, as well as in the development of other energy related analytics.

Methods
Four weeks of data were acquired from approximately 40 homes in the Boston and San Francisco metropolitan. In the West Coast data collection, monitoring devices were installed in
up to seven homes for three to four weeks at a time; 30 homes were monitored over eighteen months. All of the below data was collected for 48 different circuit breakers, in 30 total homes, with the collection period for each home typically ranging from 2-4 weeks.

We developed and refined a multistep protocol for data collection that details recruitment, consultation, hardware specifications, equipment installation and removal, data evaluation, and a one-hour energy debrief with participants. The REDD protocol was streamlined to work around complex issues that arose.

The REDD monitoring devices were designed to record aggregate home power consumption signals at high frequency and granularity, as well as collect individual circuit-level and plug-level power consumption signals. More specifically, in each home we monitored:

1. Whole-home voltage and current monitored at high frequencies (16kHz), to record the actual AC waveforms of the aggregate electricity signals in the homes. Because the raw 16kHz data is quite large, and consists mainly of repetitions of identical waveforms, we compress the data temporally, and only report those times where the waveforms change (according to the criterion of a multiple changepoint detection algorithm).
2. Per-circuit electrical power monitored at medium frequency (approximately one measurement every three seconds). Whenever possible, we label the circuit with a human-readable description, and also identify some of the major loads present on the circuits.
3. Per-plug electrical power consumption, monitored at medium frequency (often once a second, though some homes have only once-a-minute monitoring) for about 20 select plug loads in the home. In all cases the appliance being monitored is labeled in the data set.

Taken together, such data gives a powerful snapshot as to what has happened regarding the energy in the house over this period: the whole-home signals provide high frequency data where devices can be identified the waveforms themselves, while the per-circuit and per-plug signals provide the “ground truth” labeling as to what was actually occurring the home.
Figure 12. REDD monitoring device collecting data from a circuit breaker panel.

**Outcomes**

The Reference Energy Disaggregation Data Set is available at: [http://redd.csail.mit.edu](http://redd.csail.mit.edu). Here you can download an initial version of the data set, containing several weeks of power data for 30 different homes, and high-frequency current/voltage data for the main power supply of two of these homes. The data itself and the hardware used to collect it are described more thoroughly in the Readme on the main page and in the papers: Kolter, J. Z., Chadwick, S.J., Armel, K.C.; REDD 2.0: The Expanded Reference Energy Disaggregation Data Set. (2013). In press; and J. Zico Kolter and Matthew J. Johnson. REDD: A public data set for energy disaggregation research. In proceedings of the SustKDD workshop on Data Mining Applications in Sustainability, 2011. [pdf]. Those wishing to use the dataset in academic work should cite this paper as the reference. Although the data set is freely available, for the time being we still ask those interested in downloading the data to email us (zkolter@cs.cmu.edu) to receive the username/password to download the data. See the available readme.txt file for a full description of the different downloads and their formats.

In addition to this work providing data for algorithm developers and testers, we anticipate that this database and protocol will allow professionals in other geographic regions and climate zones to collect and store data to facilitate the continued advancement of disaggregation algorithms. Of relevance to continued work is the lesson to date that even in homogenous suburban areas, home energy systems, appliance stock and consumption patterns are extremely diverse.

**Future Work**
This work was aimed at spawning the development of algorithms, more standardized testing of algorithms, and the collection of new datasets. Such a database could also provide a foundation for open competitions to seed innovation.

2.3.3.3 Disaggregation Algorithms

Investigators: Andrew Y. Ng, Siddartha Batra, Tommi Jaakkola (MIT), Matthew J. Johnson (MIT)

Background and Objective

The potential of energy disaggregation algorithms - computer language instructions that do specific calculations, to accurately determine electricity consumption for individual devices - holds promise for eliminating the need for costly equipment and installation processes typically required to monitor individual devices for power use. Algorithms that can accurately determine this level of discrete information have significance for potential application to utility demand response programs. Demand response programs seek to reduce demand on the electric grid to avoid the need to build new electricity capacity, and to avoid power blackouts, damage to transformers or disruption to power frequency. These types of events can be very costly and degrade the reliability and safety of the power system. However, algorithmic approaches to energy use disaggregation (a topic also referred to as non-intrusive load monitoring (Hart, 1992), have traditionally been very simple, and focused solely on just detecting “device state changes” in a power signal; unfortunately, such detection alone does not actually give a breakdown of power in homes, and is fundamentally limited to monitoring frequencies where such “events” are obvious: the methods would be unusable, for example, using just hourly data from smart meters. Thus, the algorithmic goal of this project has been to develop new algorithmic techniques that can breakdown energy more accurately than previous approaches, using data at a variety of different resolutions and times scales.

Methods

Formally, the algorithmic work we have pursued through this project approaches energy disaggregation as a source separation problem. That is, given an observed aggregate time signal:

\[ y_1, y_2, y_3, ..., y_T \]

the goal of energy disaggregation is to find a breakdown of this signal into multiple components

\[
\begin{array}{cccc}
  y_1^{(1)} & y_2^{(1)} & \cdots & y_T^{(1)} \\
  y_1^{(2)} & y_2^{(2)} & \cdots & y_T^{(2)} \\
  \vdots & \vdots & \ddots & \vdots \\
  y_1^{(N)} & y_2^{(N)} & \cdots & y_T^{(N)} \\
\end{array}
\]

where each \( y_t^{(i)} \) term denotes the energy consumed by appliance \( i \) (for example, a refrigerator, dryer, computer, etc) at time \( t \), with the additional constraint that at each time the power
consumption of all the appliances must add to the observed total aggregate power. Although mathematically straightforward this problem is rendered very difficult by the fact that, without significant restrictions, there are any number of appliance powers that could add to a given total. Thus, the algorithmic fundamentally focuses on two elements: specifying models of what appliances “typically” look like, as well as developing algorithms that can take such models and the aggregate signal, and predict the best breakdown of consumption.

Outcome 1: Energy Disaggregation via Sparse Coding

Our first study used a collection of about 10,000 individually monitored devices, with average power recorded each hour, and applied a method known as sparse coding to learn models for appliances and separate the signals. The basic principle is to look at a fixed period of time (in this study, we used a single week), and express each appliance’s entire energy trace over that week in terms of some linear combination of “basis functions”. That is

\[
\begin{bmatrix}
y_1^{(i)} \\
\vdots \\
y_T^{(i)}
\end{bmatrix} \approx \begin{bmatrix}
| & \cdots & | \\
b_1^{(i)} & \cdots & b_n^{(i)}
| & \cdots & |
\end{bmatrix} \begin{bmatrix}
a_1^{(i)} \\
\vdots \\
a_n^{(i)}
\end{bmatrix}
\]

where \(b_1^{(i)}\) are a set of basis functions that capture “typical” usage patterns for the \(i\)th device, and \(a_1^{(i)}\) are “activations” that specify which of these basis functions makes up any given signal. Given a collection of example usage patterns for an individual device, we can learn both the bases and activations using a method known as sparse coding (Lee et al., 2006); for this particular project we developed additional algorithmic extensions that tailor sparse coding to this source separation setting (Kolter et al, 2010).

Our study with this approach used data provided by Plugwise, a European manufacturer of plug monitors. We build models using appliance-level data from 413 homes, and then evaluated the learned models to separate appliances in the remaining 117 homes. Typical results of this method are shown in Figure 1, and on average the method is able to correctly assign 55% of the energy correctly into one of 10 device categories (evenly divided, so that random guessing would give about 15% accuracy). Furthermore, the method requires only hourly data, which can be readily obtained from smart meters, and thus provides a method for approximate disaggregation that can use the existing monitoring infrastructure.
Outcome 2: High-frequency disaggregation via hidden Markov models

Although the previous approach is promising in its ability to use the existing monitoring infrastructure, new sensing modalities offer the promise of much higher frequency data. Thus, our second project used the previously discussed REDD data set (Kolter and Johnson, 2011) which was collected in conjunction with the algorithmic development presented here, to disaggregate energy in a home using ~1Hz data. In particular, we used a model known as a factorial hidden Markov model (Ghahramani and Jordan, 1997), a graphical representation of which is shown in Figure 2, that captures a time varying process with several devices that can take on some discrete number of power states (e.g. on, off, standby). While not all devices have discrete power levels, several common appliances do, and the method is able to accurately model most of the devices in a common home. Our particular algorithmic contribution for this work was to develop an “inference” procedure (an algorithm that determines the state of appliances given an aggregate signal) that was many times more accurate and faster than existing approaches (Kolter and Jaakkola, 2012).

The main goal of this particular study was to evaluate the potential of higher-frequency data (in this case, 1Hz whole-home power) and compare to more traditional event based detection methods. We used data from an initial release of the REDD data (including 6 homes) to build models for appliances and then separate out the different end-uses. Figure 3 shows a typical example of our algorithm’s output, along with the output of an event-based approach. In total, our algorithm correctly assigns about 87% of the energy in a home (again, assigning to one of 10 categories of end-use), whereas the event-based approach assigns about 49% correctly. This both highlights the potential benefit to higher-frequency sampling, as well as demonstrates the advantage of using more advanced algorithms over the previous simple approaches.

Figure 13. Performance of different types of disaggregation algorithms on hourly data.

Future Work
Future/ongoing work for these approaches include: developing methods that can build models using only unsupervised information (i.e., aggregate data alone), rather than both aggregate and individual device level; combining small numbers homes monitored at high frequency with large data sets of smart meter data; integrating the approaches into deployed systems in building energy management solutions.

2.4 Target Behaviors

2.4.1 Energy Behaviors Taxonomy

Investigators: June Flora, Hilary Boudet, Carrie Armel, Maria Roumpani

Background

Energy actions or behaviors, such as purchasing a refrigerator, turning off lights, and cleaning a heating filter, are driven by various factors, and can be described along different dimensions. In trying to explain environmentally responsible behavior, Stern (2000) lists four major types of causal factors: (1) attitudinal, (2) contextual, (3) personal capabilities and (4) habit or routine. Numerous researchers — including Kaiser, Wolffing, and Fuher (1999); Barr, Gilg, and Ford (2005); Corraliza and Berenguer (2000); Oskamp (2000) and Shove (2010) — contend that psychosocial characteristics such as attitudes, beliefs, perceived norms, and perceived risk are not primary drivers of energy behaviors. Rather, they maintain it would be more useful to study the structural or personal capability causes of environmental behavior such as household location, frequency (or repetitiveness), skill required to conduct, and cost (Wilson & Dowlatiabadi, 2007, p. 182)

In this vein, different researchers have collected lists of energy behaviors and proposed different taxonomies or ways of classifying them according to subsets of contextual or personal capabilities dimensions. For example, a review of energy-reduction feedback interventions categorized energy behaviors by cost and frequency of performance (Ehrhardt-Martinez, Donnelly, & Laitner, 2010). Similarly, in reviewing how energy reduction behaviors were characterized in 28 studies, Karlin, Davis, Sanguinetti, Gamble, Kirkby, and Stokols (2012) distinguished curtailment behaviors from efficiency behaviors. Curtailment behaviors (Black, Stern, & Elworth, 1985) are generally habitual (Stern, 1992), low-cost, require little cognitive effort, can be undone, have lower savings impact, and can be performed by a wider range of actors (e.g., children). They amount to “usage-related adjustments” (Van Raaij & Verhallen, 1983b; Dillman, Rosa, & Dillman, 1983) that require minimal or no structural changes. Whereas curtailment behaviors are simple and routine, efficiency behaviors require structural changes, are infrequent and expensive, require conscious decisions, are more permanent, and require more time (Karlin et al., 2012). Such behaviors go by several names, including technology choices (Stern, 1992), conserving behaviors (Dillman et al., 1983), purchase-related behavior (Van Raaij & Verhallen, 1983b), and energy efficiency choices (Black et al., 1985).

Our work has similar goals to that body of literature; however, we collected a more comprehensive set of behaviors, identified ten different attributes or dimensions for classifying
the behaviors, developed a rigorous and graded rating classification scheme and methodology, and explored how the behaviors clustered together based on the ratings on these attributes.

Objectives

The first goal of the energy-reduction behaviors project was to develop a data set of searchable actions that directly reduce stationary residential energy consumption and their behavior change attributes, the latter which were derived in part from behavioral science theory (Flora et al. in preparation).

The second goal of the project was to use the data set created in the first phase of the project to develop a framework for the systematic study of behavior attributes and how they cluster (Boudet, Flora, Armel, 2016).

Methods for Behavior Set Selection

Our sample of energy efficient behaviors is an integration of lists of behaviors from public sources such as U.S. Department of Energy (DOE, 2012), Flexyourpower@CA.gov, and a comprehensive list produced by the city of Townsville, Australia (www.townsville.qld.gov.au/Pages/default.aspx) (full documentation of the source of every behavior is contained in the database). The original behavior set was a collection of energy saving behavior lists compiled from the above publicly accessible materials. Over 500 behaviors were in the original set. However, once redundancies were eliminated, inclusion and exclusion rules applied, large general actions were subdivided into smaller independent actions, the specificity of each behavior was standardized, and energy savings were identified or calculated, the behavior set was composed of 261 energy efficient behaviors (a 48% reduction).

Outcomes

In two papers, one in preparation another under review, we examined the behavior attributes, their distributions in the behavior set, and the results of a cluster analysis or a grouping of behaviors into homogeneous groups.

Frequency of occurrence. Many behaviors (38%) in the set are performed only once every three or more years. These include such behaviors as major appliance purchases and home weatherization. Another 15% of behaviors occur every one to three years, such as small appliance purchases and appliance maintenance. The second largest set of behaviors (20%) occurs with very high frequency (multiple times a day), such as turning off lights and unplugging/turning off computers and entertainment devices.

Required skill level. Some 52% of behaviors in the set require very little or no skill, such as closing shades or drapes to keep heat in or out, turning off lights, or installing a CFL. Medium skills — e.g., reading instructions and/or having tools — are required for 21% of behaviors. Another 28% of behaviors demand significant skill such that an expert may be needed to perform the behavior, e.g., to install insulation or high efficiency windows.
Household function. Actions that increase thermal comfort comprise 38% of the behavior set. These behaviors involve space heating and cooling, from installing “Energy Star” air conditioners to changing thermostat settings. Another 19% of behaviors have a housekeeping function (e.g., cleaning and maintenance). The smallest categories were behaviors associated with hygiene (i.e., showering) (8%) and outdoor recreation (i.e., pool maintenance) (2%).

Locus of decision. A third of behaviors (33%) were coded as having men or women as the locus of decision, meaning that two adults typically decide whether to adopt the behavior, such as large purchases or household organization and maintenance. Slightly fewer behaviors (31%) were actions for which primarily men make adoption decisions (e.g., insulating hot water heaters). For 14% of the behavior set, primarily women make adoption decisions (e.g., emptying/replacing vacuum cleaner filter bags regularly). Teens can perform 13% (e.g., unplugging the charger once a phone is charged), and young children can perform 9% of the behaviors (e.g., turning off lights).

Observability. Over a third of the behaviors (34%) are highly observable, both by household members and by outsiders and household members. Just under a third of the behaviors (31%) are observable by household members only. Interestingly, 35% of behaviors are invisible even to household members and observable only by the person who performs the behavior. Such invisible behaviors include, for example, adjusting the hot water heater temperature.

Home topography. The most common location for behaviors is in the shell of the house: 35% of the behaviors involve its walls, floors, ceiling, or roof. The second most common locations for behaviors are kitchen/dining areas (16%), multiple areas (e.g., lighting fixtures) (15%), and storage spaces (e.g., hot water heater closet) (14%).

Appliance topography. Most behaviors involve large electric appliances (33%) or no electrical devices (36%). The remainder of the behaviors involves small electrical appliances (10%), electronics (10%), lighting (9%), and craft and recreation (2%).

Energy savings. Almost half of the behaviors (46%) are projected to save more than 750 kWh/yr. About 20% of the behaviors have marginal energy savings of 1 to 25 kilowatt hours per year (kWh/yr). The remainder of the behaviors fall into the categories in between — 15% of the behaviors save 25-100 kWh/yr, 12% save 101-250 kWh/yr, and 10% save 250-750 kWh/yr.

Cost. More than half of behaviors cost no more than $20 — 43% cost under $5, and 11% cost $5 to $20. About 20% of the behaviors cost between $100 and $1,000, and 11% cost more than $1,000.
Behavior Clusters. Five clusters resulted from a K-means analysis. Each cluster was named as follows, based on the attribute means for the clusters and the behaviors represented in each cluster: Call an Expert (73 behaviors, 28% of the sample); Family Style (66 behaviors, 25%); Household Management (49 behaviors, 19%); Go Shopping (47 behaviors, 18%); and Behind the Scenes Work (25 behaviors, 10%).

Energy Action/Behavior Attributes

<table>
<thead>
<tr>
<th>Household Function</th>
<th>Appliance Topography</th>
<th>Home Topography</th>
<th>Locus of Decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thermal comfort</td>
<td>Large electrical/gas appl.</td>
<td>Shell/Envelope</td>
<td>Primarily Men</td>
</tr>
<tr>
<td>Food/Nourishment</td>
<td>Small electrical/gas appl.</td>
<td>Kitchen/dining</td>
<td>Primarily Women</td>
</tr>
<tr>
<td>Hygiene</td>
<td>Electrical Tools</td>
<td>Office</td>
<td>Either or Both</td>
</tr>
<tr>
<td>Lighting</td>
<td>Craft and Rec</td>
<td>Entertainment space</td>
<td>Adults &amp; Teenagers</td>
</tr>
<tr>
<td>Housekeeping</td>
<td>Electronics and computers</td>
<td>Bedroom</td>
<td>Adults &amp; Teenagers &amp; Kids</td>
</tr>
<tr>
<td>Outdoor recreation</td>
<td>Electrical Lighting</td>
<td>Storage spaces</td>
<td></td>
</tr>
<tr>
<td>Entertain. &amp; communic.</td>
<td>No appliance, device or tool</td>
<td>Multiple Locations</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Outside of home</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bathroom</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Behavioral Frequency</th>
<th>Skill Demand</th>
<th>Household Observability</th>
<th>Energy Savings</th>
<th>Fiscal Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>One-time</td>
<td>No Skill</td>
<td>To everyone</td>
<td>&lt; 250 kWh/year</td>
<td>&lt; $5</td>
</tr>
<tr>
<td>Irregular</td>
<td>Low</td>
<td>To adults</td>
<td>250-750 kWh/year</td>
<td>$5 - $20</td>
</tr>
<tr>
<td>Regular</td>
<td>Medium</td>
<td>Only to the person who performed it</td>
<td>&gt; 750 kWh/year</td>
<td>$20 - $100</td>
</tr>
<tr>
<td>Frequent (Weekly)</td>
<td>High</td>
<td></td>
<td></td>
<td>$100 - $1000</td>
</tr>
<tr>
<td>Very Freq. (Daily)</td>
<td></td>
<td></td>
<td></td>
<td>&gt; $1000</td>
</tr>
</tbody>
</table>

Figure 14. Energy Behavior Taxonomy.

2.4.2 Identifying Opportunities for Dramatic Energy Reductions in Residences

Background

The magnitude of energy savings achieved by any of our Initiative’s projects is limited by the technical potential of the practices and technologies that are currently feasible for widespread adoption in our society. A very high technical potential for savings is needed to meet proposed targets of recommendations by the Intergovernmental Panel on Climate Change (IPCC) (60% to
80% cuts in greenhouse gas (GHG) emissions below 1990 levels by 2050. Eighty percent reduction of GHG emissions below 1990 levels by 2050 is also the goal for the State of California established by AB 32 of 2006 (the Global Warming Solutions Act). The State of California also has established goals for new residential and commercial buildings to be zero-net energy by 2020 and 2030 respectively.

Energy use reduction in buildings is a critical component of meeting these goals and targets. Unfortunately, current estimates of cost-effective energy savings across existing residential and commercial buildings range between 15% and 35% in the United States (McKinsey & Company 2009; APS 2008; Hand et al. 2012). Low cost savings from operational changes include a range of options such as regular appliance maintenance and replacement, turning down the thermostat settings, turning off lights when spaces are unoccupied, and eliminating energy waste (e.g. unplugging electronic devices (Nair et al. 2010; Laitner, Ehrhardt-Martinez, and McKinney 2009; Gardner and Stern 2008; Parker et al. 2006; Black et al. 1985; Kempton et al. 1984).²

However, there have been cases of people achieving energy savings as high as 90% (Ninety Percent Reductions Group Website) primarily in less developed nations. Many non-western cultures in other parts of the world use or have used dramatically less energy to achieve daily living needs. Insights into the nature of these practices and low-technology approaches have potential application in western cultures for significant energy savings.

Objectives

The goal of this project was to compile a list of practices and technological insights, from other cultures, time periods, that have the potential for producing substantial residential building energy savings in our culture (if adapted to be culturally appropriate). We also quantified estimates of energy savings potential of these practices and technologies.

Methods

Combining approaches from the fields of anthropology and design, we collected data from secondary sources such as books and articles, and conducted in-depth interviews with twenty “extreme users” including energy experts, historians familiar with how people in the past have met their daily household needs, do-it-yourselfers aiming for steep energy reductions; and people from a variety of cultures, in particularly those from harsh climates. We augmented interview data with secondary research, drawing from cultural anthropology, history, and biology to collect examples of social and biological change and adaptation.

---

¹ In 1990 CO₂ levels were about 350ppm, which is considered the safe upper boundary; the 2012 figure is 394 ppm.

² Note that behavioral changes do matter and one residential electricity reduction study showed behavioral changes accounting for 59% of the energy savings, compared to 41% savings from structural changes (Wu 2012).
Outcomes

Our results include approximately 100 energy-saving options covering actions, products, and home adjustments organized by end uses: space heating and cooling, water heating, cooking, and refrigeration. The results were also organized according to the physical mechanism or property by which they operated: (1) eliminating energy waste; (2) insulation and sealing; (3) air flow and evaporative cooling; (4) reflection and shading; (5) absorption, storage and thermal mass; (6) alternative and latent energy; and (7) acclimatization and adaptation. For example, we found that cultures in hot dry climates achieved significant evaporative cooling effects by using porous drinking vessels (which sweat and evaporately cool), hanging moist cloth in well ventilated areas such as doorways or in wind towers, and using light colored loose clothing to facilitate air flow through the cloth which is moist with captured perspiration; they also used underground rooms to cool their bodies in the hot afternoons and preserve their food (root cellars are one example of this). Energy savings potential within each end use and principle are quantified. We also report barriers that interviewees identified at the individual and institutional level. This work offers an approach to identifying and prioritizing energy-saving options that are currently uncommonly applied in mainstream culture, with the intent of guiding further development such as redesign, overcoming barriers, and promoting widespread adoption.

Future Work

This work may be pro-actively shared with designers and policy makers in an attempt to disseminate the practices or principles identified that could be adapted for our culture to facilitate deeper energy savings. This approach may also be extendable beyond residential buildings in areas such as, for example, the food, transportation, and small and medium commercial building sectors.
CHAPTER 3: Behavioral Interventions

A dominant theory in public health holds that the use of multiple types of interventions (also referred to as programs or treatment approaches) is more effective than one, because they complement and reinforce one another. Several of the Initiative’s projects involved developing and evaluating media, incentive, and community-based program interventions. While some projects evaluated the effectiveness of multi-faceted programs which is a typical approach employed in the field of public health or in utility pilots, other projects followed an experimental psychology approach (i.e. using control and treatment groups) and systematically changed variables between conditions to investigate the impact of specific variables on energy behavior.

3.1 Media Interventions

3.1.1 Multiplayer Online Game

Investigators: J Byron Reeves, James K. Scarborough, James J. Cummings, Leo Yeykelis
Partners: Kuma Games, Inc.; Bonsai Corp.

Background

Energy information for consumers can be complex and uninteresting. Games offer a compelling new context for home energy information that may engage consumers and change behaviors. Multiplayer games, for example, are sprawling online communities where players interact with and compete against one another in real time within visually rich, three-dimensional virtual worlds that persist and evolve even while a player is away. These games may be the most engaging, sophisticated, and collaborative media ever to be applied to campaigns to change behavior in serious contexts (Reeves & Read, 2009). The audience is big, with as many as 400 million people worldwide that operate avatars in virtual environments (Gartner, 2007). And the audience is surprisingly diverse; for example, gamers average 33 years old and there are more of them in their 40’s or 50’s than in their teens, the majority of them have full time jobs and kids, and the gender ratio ranges from equal to 3:1 depending on the genre. Furthermore, people are coming to expect engagement in workplace settings as well – IBM and other corporations are beginning to incorporate game-like elements and virtual meetings into work tasks.

Objective

The objective was to leverage the popularity of online games to promote energy efficiency. The deliverable for this project was a website, which included a multiplayer game and supporting social media (such as facebook connect), that is suitable for use in experiments and deployment in utility smart meter trials. Empirical experiments on selected features of the media will guide future generation media.

Game Description and General Methods
Based on research showing the effectiveness of game elements used in serious contexts (Reeves & Read, 2009), we built a professional quality social game about energy use in a virtual home. We contracted with an experienced entertainment software development company, Kuma Games, Inc., to ensure we provided a professional and meaningful media experience to players. In the game, playable here https://www.freeenergygame.com/portal/, we link energy sensors in homes to multiplayer interactions that promote changes in energy use in the context of compelling play and community participation. More specifically, the game uses real world energy use data from smart meters and converts energy savings to rewards and advantages in the game. For example, one can challenge their friends to a “lights out night” and then see who won based on the actual energy consumption data. Embedded within the overall game is a smaller game. In one mini-game (depicted in the screen shot above), the user races around a virtual house over the course of several “days”, trying to achieve all the goals of the household members while turning appliances on and off so as to use the least amount of energy. In this manner, the game helps condition energy efficient behaviors in the player’s actual home by modeling and reinforcing those behaviors in the virtual setting. By speeding up time and providing feedback promptly in the form of points, the game helps players to more easily develop particular energy habits. Additional features considered in the overarching game include: (a) multi-player game play (permitting individuals as well as virtual and intact groups); (b) multi-period game play (accumulating across multiple play sessions over weeks or months); (c) launched via social networking sites (e.g., friend groups on Facebook that can encourage viral distribution); (d) episodic content (new challenges can be introduced regularly that match regional or seasonal energy goals and can be implemented through video vignettes); (e) competitive (use of point-based leader boards and energy data tracking that allow competition between existing real world groups e.g., companies, classrooms, geographic neighborhoods, community organizations); and (f) portable (game play, scoring and notifications can be tracked through social networks on mobile devices). The primary thesis is that an alignment of personal motivations (e.g., increased involvement encouraged by timely reinforcement, achievement recognition, and a sense of belonging), and community environmental goals (e.g., reduced electricity usage and time-shifted energy use) will result in sustainable behavior change that is personally rewarding as well as socially responsible.
Methods, Laboratory Study

In the laboratory experiment, 40 people were randomly assigned to play Power House or a similar but thematically different entertainment game (Diner Dash) for 30 minutes in a 10’ x 10’ office. Subjects were told that we were testing the popularity of a new game and would be asked to answer questions about how much they liked the game at the end of play. Fifteen minutes after play started, an experimenter entered the room and told subjects that she needed to leave the building before they would finish. The experimenter requested that subjects “close the office” when they were finished with the game and questionnaire. Before subjects entered the room to start playing, there were five appliances turned on in the office. There were two overhead lights controlled by a wall switch, a floor lamp, a desk lamp, and a computer and monitor. After subjects had finished playing the game and closed the office, we returned to the room to count how many appliances had been turned off after play ended and before subjects left the room.

Outcomes, Laboratory Study
Playing the game for 30 minutes resulted in significant increases in energy efficient behaviors after play ended, compared to playing the comparable non-energy focused game. In the energy game condition an average of 2.55 (out of 5) appliances were turned off when subjects left the room; in comparison, an average of .55 where turned off after playing the non-energy game. These show that the energy game was capable of inducing energy efficient behavior after only brief exposure to the game content. More, subjects reported no conscious connection between game play and the measured energy behaviors, suggesting that this behavior was primed through mere exposure to the game content.

**Methods, Field Study**

In the field study participants played the game in their homes over the course of one week to one month while their smart meter provided home energy consumption data for analysis. Participants in this study would typically play the game within a real social context. For example, while playing the game via Facebook, players were able to post in-game achievements and energy savings for their Facebook friends to see. Additionally, some participants would see a detailed energy consumption chart for the previous day while others see only their monthly energy bill. This allows researchers to determine if a higher frequency of feedback can motivate players to save more energy and develop stronger energy conservation habits. Participants also completed a home energy intention survey before and after playing the game to measure any impact on energy conservation intentions that playing the game might affect.

**Outcomes, Field Study**

Results from the field study suggest that participants use significantly less energy while they play the game. Energy use reduction was stronger for those participants who started with a lower energy baseline than those with a higher energy baseline. This difference might be explained by the presence of a high energy consuming device such as a pool or hot tub or by a higher number of occupants in the home. These results indicate that well designed entertainment software can make a significant difference in home energy sustainability.

**Dissemination Activities**

PowerHouse was one of the projects that continued into a Plus Up funding period. For dissemination activities, both the website and a Facebook application version were used. In 2015, 16,892 unique players played 21,221 game sessions through Google Adwords advertising, which was 168% of the goal of acquiring 10,000+ new players. However, regarding players sharing smart meter connections for game purposes, only 5.6% of the goal was achieved (only 56 players were willing to share their utility credentials; the cost of reaching 1,000 people willing to share their credentials was prohibitive). In winter of 2016, all Powerhouse traffic was organic (not paid), with approximately 30% coming from Facebook; 107 users played the game directly from the freenergygame.com Web site with an additional 46 playing from the Girl Scout Facebook page. 13 Facebook players attempted to register. Regarding business arrangements, without established scale, no major advertisers were interested in sponsoring or advertising in or around the game.
3.1.2 Collective Action Feedback Interface


Background

Increased home energy conservation may be accomplished by framing efforts to save energy as part of a community effort to provide residents with a motivating sense of togetherness when they perform energy saving behaviors. Past research found that the opportunity to participate in a collective endeavor can be a powerful source of motivation. For example, academic motivations and achievement increase when students feel socially connected to peers and teachers or when they can work together with others on a task (e.g., Furrer & Skinner, 2003; Goodenow, 1992; Roese, Midgley, & Urdan, 1996; Wentzel, 1997; Walton & Cohen, 2007). The importance of collective goals has not been specifically tested in the context of environmental behaviors, but a related phenomenon — the effects of descriptive norms — has been. Descriptive norms are perceptions people hold regarding which behaviors are commonplace. For example, people will more likely reuse hotel towels, reduce energy consumption, and keep petrified national park artifacts intact if they’re informed that others are also performing the behaviors than if they are given pro-environmental or monetary incentives. This is true even though people believe that the latter two appeals will be more effective (Goldstein, Cialdini & Griskevicius, 2008; Schultz and colleagues 2007). One reason norms may be effective is because they convey group intentions to individuals, and being a part of these collective actions provides a desirable and motivating sense of togetherness.

Objectives

Many studies present descriptive norms and clear group intentions together to movitative actions or changed behavior, making it difficult to tell if norms alone are helpful, and whether invoking a sense of community is more powerful motivator than basic descriptive norms. The following two studies separated descriptive norms and clear group intentions to determine if they have unique effects.

Methods, Study 1 Energy Upgrade Mountain View Experiment

This fist study involved providing energy related information and collecting data from over 800 residences (mostly single family homes, but also some duplexes, townhouses, apartments/condos) in Mountain View, California. Energy use data was collected for each residence that participated for a continuous six month period, sometime between March 2012 and July 2013 depending on when the resident signed up for the program. Participants were recruited, through a variety of methods, including displays and signup lists at community events, door-to-door canvassing, and sending flyers with energy bills as part of the Mountain View Energy Upgrade California program http://www.energyupgrademv.org/. The most successful method of recruitment was to include low budget flyers with the utility bills of all the residents of Mountain View. These flyers advertised the program, offered workshops for
learning about home energy reduction and also provided a free smart power strip as an incentive to sign up.

Figure 16. Collective action experiment website.

Participants viewed feedback about their energy consumption as well as energy saving recommendations and messaging in emails sent every other week, and they could also access this information through a website. Experimental conditions allowed for the comparison of the effects of providing energy use feedback and energy saving tips to households (through the differential framing of feedback data through comparison with others, pictures, and text-based messaging between versions of the emails/online interface) when saving energy was framed 1) as an intentional community effort done together (social togetherness condition; e.g., “We’re doing it together!…Mountain View residents have reduced their energy use by 3% in the last few years.”), 2) simply as a descriptive norm (norm condition, e.g., “Here’s a fact!…Mountain View residents have reduced their energy use by 3% in the last few years.”), or 3) without any normative information as a control (control condition). Additionally, this study compared the effects of presenting energy savings information in different ways by randomizing half of the participants in each of the three conditions above to either receive energy saving tips with or without a generalizable theme that readers may extend to other energy savings behaviors. For example, participants in the tips condition might see tips regarding plug loads, heating, and refrigerators
all in one email (e.g. “Unplug devices you never use, like an old VCR or fridge.”), while participants in the theme condition would also see three tips in one email but these would all relate to heating and an introductory comment on their common energy consuming mechanism would be included (e.g. “Off isn’t off”…off is still on. When you just press the off button on an electronic device, it may still be dripping electricity like a leaky faucet.”).

Outcomes, Study 1

We found that home energy reports in general reduced energy use, and thematic information about energy savings framed as a community effort in the social togetherness condition produced a statistically estimated energy use reduction of 7.7% (1 kWh/day) in home energy use when compared to the basic descriptive norm and control conditions after 6 months. This reduction was significantly better than those in thematic information with a descriptive norm only and thematic information under control conditions. Paradoxically, we also found that our control group given only tips (without themes) also reduced their energy use by 6.2%—significantly more than the social togetherness tips only condition, or descriptive norm and tip only condition. One possible explanation is that the togetherness condition provided sufficient motivation to read through the lengthier thematic information and implement the ideas presented to them, while those in the control tips only condition may be more motivated to read the contents of the email given its brevity (as this condition has the least text, allowing tips to be seen at a first glance).

Methods, Study 2 City of Hillsborough Water Use Experiment

The goal of Study 2 was similar to that in Study 1 - to measure how different types of framed messages would affect future consumption. In this study, water was targeted because reducing water consumption also reduces energy use through the embodied energy consumed to extract, treat, distribute, and as applicable to heat water. The study took place in the town of Hillsborough between May 2011 and May 2012 and included data from all city residents; over 10,000 people living in over 4000 households (there are no apartments or condominiums in Hillsborough). Residents were sent paper inserts with their monthly water bills through the postal system, and everyone also received a waterproof vinyl tag with irrigation information to hang on their outdoor faucet or elsewhere to serve as a reminder prompt. The experimental conditions and means of achieving these through data, pictorial, and text manipulations on the paper inserts and vinyl tag were similar to those used in Study 1, although the tips/themes question was not investigated. Instead, each of the three social togetherness, descriptive norm, or control groups – no message were divided further into three subgroups based on percentile ranking (highest third, middle third, lowest third) relative to the community (resulting in a 3 by 3 design: social togetherness /descriptive norm/control x highest/middle/lowest energy consumption). The community or norm based motivational text differed within each percentile rank. Water email alerts were also provided to the 286 households that signed up for this after the initial mailing. Participants were also assigned to fill out one of twelve different online surveys (3 by 2 by 2: social togetherness/descriptive norm/control x survey with writing/survey no writing x adult/child) which were in keeping with the experimental condition they were assigned to, and also had the added experimental manipulation testing whether the process of
writing about one's behavior increases behavioral change (through a paragraph response box on the survey), as it has in previous psychological research.

**Outcome of Study 2**

Researchers found that the interventions worked differently than expected. Households in the social togetherness condition used more water than those in the norm or control conditions. Interestingly, customers in the control condition used the least water, compared to both the norm and social condition customers. There are a few possible explanations for the unexpected results. It is plausible that political orientation moderated the effect given that conservation and sustainable behaviors have been greatly politicized, though this was tested and there is no empirical support of such an effect in this study. Instead, Hillsborough may not be enough of a close-knit community for the message of togetherness to be effective. Perhaps if the messages had been localized to neighborhoods rather than the whole town, it would have been effective. It is also possible that Hillsborough is a close-knit community, but residents already feel that they belong in the town, and therefore do not particularly feel the need to join in with the community. Additionally, in a recent paper, Hamedani, Markus, and Fu (2013) found that emphasizing interdependences undermined European American's motivation to learn about environmental sustainability, and led to decreased funding allocated to the cause (though this was not the case for Asian American participants). Therefore, it is possible that our message of community togetherness in fact backfired, and decreased residents' motivation to save water.

**Future Work**

Going forward, research on the effects of norms will focus on understanding when norms, particularly social norms invoking togetherness, are successful in motivating behavioral change, and when they are likely to cause negative reactions for intervention participants. There is potential for applying what was learned from this and similar studies to the design and nature of messages created for the purpose of recruiting participation in larger scale conservation programs.

3.1.3 **The Impact of Vivid Messages on Saving Behavior related to Hot Water Use**

Investigators: Jeremy N. Bailenson, Jakki Bailey, June Flora, K. Carrie Armel, Dave Voelker, Byron Reeves

**Background**

Virtual environments may offer a unique opportunity to facilitate cognition through embodied experiences that are personal and vivid. Immersive virtual environment technology (IVET) engages people in a three-dimensional (3D) virtual environment with a first person point of view, and it provides real-time multisensory feedback via visual, haptic, auditory, and olfactory cues. IVET also allows users to participate in actions that could not be accomplished in the real world; for example, passing time quickly, experiencing impossible physical spaces (e.g., geographically remote or fictitious), or experiencing behaviors that are novel, impossible or
undesirable (e.g., harmful to the self, others, or the environment). Allowing people to experience undesirable behaviors is especially pertinent to environmental behavior changes because it allows people to observe directly far-reaching negative outcomes associated with their actions (e.g., burning of coal, smog emitted, trees felled).

Independent of IVET, vividness and personalization have been particularly effective in promoting behavior change. Personalizing or customizing information increases attention and has improved the effectiveness of numerous public health interventions. Vivid messages are emotionally interesting and imagery-provoking in a sensory or spatial way; for example, auditors have been more successful in signing up homeowners for retrofits that reduce energy use when they vividly described the cumulative air leaks in a house as being the “size of a football” or a lack of insulation like having a “naked attic” (Gonzales, Aronson, & Costanzo, 1988). The impactful effects of personalization and vividness can be explained by the theory of embodied cognition (EC) which suggests that cognition is a grounded experience that occurs in relation to states of one’s body and perceptual simulation. If cognition is closely related to perceptual experiences, vivid messages that utilize strong imagery appeals, and personalized messages, may simulate sensory information in the brain that leads to changes in behaviors.

Although many researchers agree that pro-environmental interventions may be more effective if they contain vivid and personal elements, few interventions have successfully combined the two. Further, IVET has the capability to create vivid and personal interventions to a degree rarely seen in previous work. To date there has been only one study to our knowledge to use IVET to raise awareness on global warming, and it did not study changes in energy behavior (Zaalber & Midden, 2010).

Objectives

This study used IVET to investigate the impact of vivid and personal messages on energy use behavior specifically related to hot water use.

Methods

In order to select a vivid visualization or metaphor for energy consumption (e.g. CO2 balloons, energy vampires) in this study, we collected dozens of metaphors from public service announcements and online materials (e.g., a penguin bicycling to show how much energy was required to light a lightbulb), and also partnered with the marketing firm DraftFCB on an international competition among their offices to collect another approximately 100 ideas for visualizations. After reviewing them, we selected the metaphor of one actually eating the fuel they use in an activity to represent energy consumption, because it is a salient metaphor and IVET lends itself well to representing one’s self performing impossible or implausible actions.

During the experimental study, seventy participants were placed in a virtual shower for approximately six minutes – that is, they received 3D visual and audio cues, and were asked to move their hands over their limbs and head as though they were showering – in order to
simulate a shower experience. In addition, they received feedback about how much energy was used to heat and transport the water in their virtual shower. Participants in different experimental conditions received different feedback on the dimensions of message vividness (vivid or not vivid) and personalization (personal or not personal), as described below. Specifically, the four conditions included the following (see Figure 17). In the avatar-coal condition (top left) the treatment was both vivid and personal, and participants saw a 3D digital representation of him or herself called an avatar, standing outside of the shower window eating once piece coal for every fifteen seconds of virtual shower time. In the coal-only condition (top right) the treatment was vivid but not personal, and where the visual feedback was individual pieces of coal piling up on the table every fifteen seconds. The two non-vivid conditions provided feedback through the use of a counting ticker on a billboard sign shown outside the window (increased every 15 seconds). In the personal-sign condition (lower left) the treatment was not vivid but was personal treatment; the billboard sign used personal language when counting the number of pieces of coal consumed: “you have consumed 1 piece of coal.” Finally, the impersonal-sign condition (lower right cell) was the not personal and not vivid - participants saw a billboard sign that used impersonal language that used the passive voice to count the number of pieces of coal consumed: “1 piece of coal has been consumed.”
Figure 17. Immersive virtual environment energy experiment conditions. Each cell shows the participant’s view of the virtual shower. Participants were randomly assigned to one of four experimental conditions: (1) avatar-coal condition, (2) coal-only condition, (3) personal-sign condition, and (4) impersonal-sign condition.

**Outcomes**

Results showed a main effect of vividness. Before and after the IVET treatment, we measured the amount of hot water people used while washing their hands in a real sink placed near the laboratory. Participants exposed to vivid messages in the virtual shower experience generalized their learnings to hot water consumption in general and changed their behavior - they used less hot water when washing their hands compared to people exposed to non-vivid messages. There were no significant effects for the different levels of personalization and no interaction effects. The results suggest that new media-technology like IVET can leverage vivid sensory experiences to change environmental behavior.
Future Work

The value add of IVET should be explored; for example, how much does energy savings behavior decline (if at all) using less sophisticated but higher market penetration virtual technology like Microsoft’s Kinnect, or without using any IVET technology at all. Specific energy saving metaphors/visualizations in addition to those used here could also be tested to determine which are most effective in motivating energy saving behaviors. Additional benefits of IVET could also be explored for their effectiveness: speeding up climate impacts over time, experiencing impossible physical spaces (e.g., geographically remote or fictitious), or experiencing behaviors that are novel, impossible or undesirable (e.g., harmful to the self, others, or the environment).

3.1.4 Motivationally Framed Facebook Energy Applications

Investigators: Banny Banerjee, June Flora (Research Director), Team: Nishand Bhansali, Nicole Greenspan, Ollie Khakwana, Alexandra Liptsey-Rahe, Brett Madres, Ann Manley, Issra Omer, Nikhil Rajendra, Ansu Sahoo, Annie Scalamnini, Brian Wong, Shaun Stehly, Dave Voelker

Background

At the Stanford ChangeLabs, we created new methodologies for application to large real world audiences that combine principles of Design Thinking with those of Behavioral Sciences and Diffusion Theory [3, 4, 5, 7]. In this study we performed ethnographic research (research to explore the cultural processes and context within which energy is consumed), developed several Facebook energy applications, and performed associated research in part through using the ChangeLabs methodologies. Initial ethnographic research founded on the grounded theory approach [7], which uses collected data to develop hypotheses rather than vice versa, explored peoples’ motivations for engaging with their energy use. Building on our ethnographic findings and guided by the Comprehensive Behavior Determination Method [1], which holds that interventions that use multiple frames appropriately paired with peoples’ motivations are more effective at changing behavior than a single frame, we identified three key motivations for energy engagement: affective, cognitive, and social. We then used standardized measures of those concepts - need for positive affect [8, 9], need for cognition [10], and need for affiliation [11] - in our experiments on tests of images to be included in our applications, recommendations, and manipulations of application prototypes. This guided the final form of the Facebook applications.

Objectives

Our objective was to build, test, and diffuse Facebook energy reduction applications as well as stimulate a body of work that uses design and behavioral principles to advance innovations regarding energy reduction.

Methods
Early ethnographic work revealed that when asked about energy, consumers responded:

- “Comfort, convenience and peer comparison are important to us and our family.”
- “Our family deserves the best.”
- “Our family believes in volunteerism and public service.”
- “We try to model our values for our children.”
- “We are not sure how our small energy reduction efforts help the environment.”
- “I am not engaged with my energy use.”
- “Energy use and conservation is just no fun.”
- “Energy is inexpensive so why should I bother.”
- “We do not know what to do.”
- “We will buy new appliances when the time is right.”

Using the results from our ethnographic research, we developed three motivationally framed energy reduction applications, Cognitive, Social, and Affective (altruistic).

*Power Bar*, the cognitive app, is designed for people who are motivated by data about their home’s energy expenditure. The driving motivation underlying this is the “cognitive” frame, whereby, the behavior change mechanism is data about the home’s energy usage coupled with goal setting for energy savings and feedback regarding whether the target is likely to be met.

*Kidogo* is an application built around the motivation frame of “affect”. It works with the presumption that it is possible to map energy savings to some issue other than energy that the consumer might be more emotional about, such as global poverty. This enable creating a bridge between their high motivational level in one issue to a behavior in another, in this case energy. This application allows energy savings to be converted into the emotional satisfaction of having contributed towards micro-finance loads in developing countries, rather than the relatively insignificant amounts saved in the monetary terms.

*PowerTower*, the social app, is a tetris-like game in which individuals get blocks based on how much energy they’ve saved. If you do not have electricity input into the system then you get blocks from behavior commitments and reported changes. You can also create a team and compete with other teams – building your blocks into a tower to see who can get the highest. The premise behind this game is that rather than trying to get households to change their energy behaviors in radical and improbable ways, this game leverages the large number of people in a social network all of whom commit to achievable amounts, but the large number of people who might get involved due to network effects contribute to a large cumulative savings in energy.

In addition, we used formative empirical studies to test design components to validate aspects of each of the applications. We conducted 4 online pilot tests (with over 600 young adult community college students) examining Affective application images of social entrepreneurs, Cognitive graphic feedback types, and two studies examining the role of individual motivation orientation and behavior change potential after viewing video prototypes of the energy reduction applications.
Figure 18. Facebook applications, Kidogo and Power Tower.

Outcome

Findings from the formative empirical studies include: (more detail is available here [http://peec.stanford.edu/energybehavior/projects/facebook.php](http://peec.stanford.edu/energybehavior/projects/facebook.php))

(1) Affective image test: Sad (rather than happy) images of humans (rather than animals) were most effective for the Affective application (within subject design with N=67);

(2) Cognitive graph test: Bar graphs (rather than line or radial graphs) displaying either a single day of information or a comparison of two days were most liked and understood (mixed design with N=207);

(3) Affective orientation and affective and cognitive app prototype experiment: Individual positive affect motivation was associated with higher behavior intentions and self-efficacy (a main effect) and that affect orientation interacted positively with the affective application (one way between subjects design with N=158);

(4) Affective, Cognitive, and Affiliative orientation and three prototype apps experiment: We replicated the main effect of affective orientation on behavioral self-efficacy but found no effects of cognitive or affiliation orientation, while there were no behavior change intention change differences among the three applications; the affective app showed marginally significant greater behavior changes on easy behaviors and both the affective and social app were rated higher than the utility control (one way controlled between subjects design with N=224)
Using these results, we have developed applications for use in the Facebook environment to ensure an “ambient” presence. These applications are now functional and data is in process of being collected and analyzed with respect to assessing the impact on electricity consumption.

Also, given our weak results outcomes of association of individual orientation to application type, we are now in the process of examining the role of choice in application selection using a randomized controlled experiment. Our randomized experimental design had two levels; first participants were randomly assigned to an assigned or a choice condition. In the assigned condition, participants were randomly assigned to one of the three applications. In the choice condition, participants choose which app they want to use (based on a short description of the application) to view their energy information.

This design work has implications for the practical scalability of energy applications. When confronted with three applications, utilities, energy service providers or non-profits or other potential adopting organizations typically would choose one application. Yet, our applications were conceptualized as a motivational frame map, where users can be matched or choose how they want to engage with and change their behavior.

### Summary of Dissemination Activities

This was one of the projects that continued into a Plus Up funding period. As an outgrowth of Banerjee’s ARPAe funded research on motivationally framed energy reduction apps, the *Behaviorally Informed Design for Energy Conservation* Online course is offered through the Stanford Professional Development Program and is a course in the Energy Innovation and Emerging Technologies Professional Certificate Program. The course is in its third quarter of being offered. The course focuses on changing the behavior and practices of energy users which can be just as important as finding new sources of energy. The course also requires an understanding of community-based social marketing, psychology and behavioral economics for successful public action and support. Finally, the course covers strategies for designing and implementing effective behavior change programs for promoting environmental sustainability using innovation and design decision frameworks.

### 3.2 Incentive Interventions

#### 3.2.1 Nudges for Energy Efficiency through an online Appliance Calculator

**Investigators:** Sam McClure, Sebastien Houde, Carrie Armel, Samuel McClure  
**Partners:** Bonsai Corp.

**Background**

To encourage purchase of energy efficient appliances there are three primary types of policies in the US: minimum energy-efficiency standards, voluntary standards, and labeling. In the US, most appliances are required to have the EnergyGuide label, which was first introduced in 1979. The EnergyGuide label provides detailed information about energy costs. In 1992, the US Department of Energy introduced the ENERGY STAR program, a voluntary certification.
program that complements the EnergyGuide. The goal of the ENERGY STAR program is to facilitate the identification of the most energy efficient models and overcome the complexity inherent to energy information. The ENERGY STAR program is quite straightforward. For a given type of product, a threshold above the minimum energy efficiency is defined, if a product meets or exceeds this threshold, the product can earn the ENERGY STAR certification. The ENERGY STAR program has proven to be effective to influence consumers and firms (Houde, 2012, 2013). Some consumers, however, appear to trade-off energy efficiency with other attributes using the EnergyGuide label, while others do not pay attention to ENERGY STAR and EnergyGuide (Houde, 2012). These results suggest that there are opportunities to provide better energy information to consumers.

Furthermore, energy labels were designed with a focus on a shopping experience in brick-and-mortar stores. In the last decade, shopping habits have drastically changed. When it comes to appliances, online shopping is an important part of the shopping experience. In 2011, although only 8% of appliances were sold online, more than 38% of consumers that bought in-store said that online shopping influenced their decision (Traqline 2011). How energy information should be presented online, however, is unclear.

**Objectives**

The goal of this project is to investigate how and which type of energy information can nudge consumers to purchase energy efficient appliances when searching online for energy intensive durables, such as refrigerators. The nudges and frames that we use are guided by principles of behavioral economics, such Prospect Theory (Kahneman and Tversky, 1979) and intertemporal choice (Lowenstein and Prelec, 1992), as well as simple principles that have proven very effective to date, such as defaults (Levav et al., 2010).

**Methodology**

To study online purchasing behaviors, we developed an online appliance recommendation website (Figure 1). The website has three components. The first component allows users to learn about the electricity consumption and cost of the appliance (refrigerator) they currently own. This information is used to determine whether it is desirable for users to switch to a more energy efficient models. The second component allows users to search for a new appliance. The third component presents search results. We created different versions of the website to test how to present the search results to induce more energy efficient purchases. These different versions are discussed below.

For each user that comes to the website, we track key information about consumers’ preferences and the impact of framing. We track:

- Information about the refrigerator currently owned
- Search criteria
- Refrigerator models that were displayed to users
- Refrigerator models that were selected and saved to a “list” to compare
Our experimental outcomes consist of the average electricity consumption that each user browsed and saved to the list. Those are the best proxies we are able to observe in lieu of purchases. To be more precise, for each user we know the appliance models that were clicked on and saved to the list. For each of those appliance models, we also know their electricity consumption. We can then compute the average of the electricity consumption for all models a user was interested in. We report this average for different versions of the website. Versions with the lower averages are the ones that we consider the most effective to induce energy efficient purchases.

Figure 19. Appliance Calculator browser application, accessed through Google Adwords.

Results
**Experiment 1**

The first experiment compared two versions of the appliance recommendation website: a version with a strong emphasis on energy efficiency, and a version with minimal emphasis.

In the version with a strong emphasis on energy efficiency (treatment), products in the search results were displayed with a picture, a short description, price, and three pieces of energy information: kWh/year, lifetime electricity operating costs, and ENERGY STAR compliance. In addition, products were sorted in ascending order of electricity consumption, i.e., the most energy efficient product was always shown first. Finally, the search filters were ordered in a way that the electricity consumption range and ENERGY STAR compliance were shown first.

In the version with minimal emphasis on energy (control), search results only included the picture, a short description, and price information. Moreover, products were sorted in ascending order of price (cheaper product first), and the electricity consumption range and ENERGY STAR compliance search filters were placed at the bottom of the list in the Step 2.

Data from approximately 14,000 users were analyzed for this experiment. Users allocated to the control browsed refrigerator models that consumed 594 kWh/year on average, while consumers in the treatment browsed models that consumed 523 kWh/year. From a regression analysis that controls for month fixed effects, we found that the version with the energy focus led to a statistically significant decrease of 69 kWh/year. Using the second outcome variable, the average kWh/year saved to the list was 454 kWh/year in the control and 412 kWh/year in the treatment. The regression analysis suggests a decrease of 85 kWh/year for this outcome variable. In sum, the version of the website with a strong energy focus was successful in nudging consumers toward more energy efficient models.

**Experiment 2**

The goal of the second experiment was to disentangle the effects of some of the features of the website with a strong energy focus. The two main features to test were the addition of several metrics to measure energy efficiency, and the default sorting that presented the most energy efficient product first.

In this experiment, we created a second version of the website (referred as treatment 2) with a focus on energy with a subtle change; we altered the rank of the most energy efficient product, and pushed it further down the list of products at rank six. The two versions of the website with the energy focus were then identical except for this small change in the default rank of the six first products presented. Note that by moving the product ranked first to the sixth position we impact also the rank of the products in rank 2 to 6. This manipulation thus allowed us to isolate the effect of the first sixth ranks. More precisely, the experiment allowed us to determine whether showing a product at rank 2 vs. rank 1, 3 vs. 2, 4 vs. 3, 5 vs. 4, 6 vs. 5, and 1 vs. 6 would impact browsing behaviors.

Data from approximately 20,000 users were analyzed for this experiment. Comparing the two treatments to the version of the website without an energy focus (control), we found evidence that altering the rank had an important effect on browsing behaviors. Pushing the rank of the
most energy efficient refrigerator model down the list increased the average kWh/year browsed and saved to the list relative to the first treatment. Our number of users in each treatment scenario does not allow us to confirm whether this effect is statistically significant at a 5% level. For the first treatment (same treatment as experiment 1), the effect is significant at the 10% level.

These results suggest that rank has an important effect on whether consumers consider a product, and may have a bigger effect than energy information. The third experiment suggests that various pieces of energy information may have in fact a very small or even no effect.

**Experiment 3**

For this experiment, we created four versions of the website, each with a different piece of energy information, in an attempt to determine whether computing out future energy savings from an energy efficient appliance would address the first cost bias, or tendency to select an appliance based on its up-front cost rather than future energy savings. In one version, we presented search results with a picture, a short description, price and kWh/year. In a second version, we added the lifetime electricity costs. In a third version, we presented annual energy savings instead of lifetime of electricity costs. In the fourth version, we presented both lifetime electricity costs and annual energy savings, in addition of the other attributes (picture, short description, price, and kWh/year).

Data from approximately 30,000 users were analyzed for this experiment. The results for this experiment suggest that the four treatments led to similar outcomes. That is, they are not statistically significant from one another. This suggests that none of these pieces of information was more powerful than others in nudging consumers.

**Future Work**

Considering the three experiments together, these results have a somewhat profound implication for behavioral interventions. That is, in Experiment 3, an attempt to produce more energy efficient purchasing behaviors by using an intervention derived from an analysis of what the underlying problem was – the first cost bias – failed to produce an effect. That is, we attempted to address the tendency of people to make an appliance purchase decision based on the up-front cost of an appliance rather than future energy savings with implications from Prospect Theory and the intertemporal choice literature. In contrast, in Experiment 1, where the intervention was driven simply by using some of the previously demonstrated most effective behavior change techniques, we saw a change in average kWh of between 10-20% depending on the measure used. Thus, for quick and effective results, it may make sense to first try the most effective proven behavior change techniques to date.

Future research should test additional manipulations to clarify how to best design online appliance recommendation websites to encourage efficient purchases, and which types of behavioral techniques are most effective.
3.2.2 Transportation Lottery


Background

In many situations incentives motivate behavior change. However, in many public goods programs the amount of money available for such programs is constrained, and results in relatively small incentives per participant when divided across all participants. Evidence shows that small piece-rate monetary incentives can actually decrease desirable behaviors because the reward “crowds out” intrinsic motivations; for example, offering $7 per blood donation reduces the amount of blood donated (Mellstrom and Johannesson 2008). However, devising alternative incentive structures (also called mechanisms) using the same total pot of available funds can stretch the perceived value of the incentive. In particular, using a lottery-based system to pay out chunky prizes instead of lower, deterministic payouts can attract more participation in the incentive mechanism. This is especially true when the deterministic payouts are small; e.g., a recyclable refunds only $0.05, saving a kWh of energy only saves about $0.10, and an off-peak trips saves 5-10% fuel relative to traveling in the peak hour. In these cases the effort of taking the right action seems hardly worth the payoff. Lottery-like payment mechanisms are much more effective, exploiting the fact that in games with low stakes players are much more risk-seeking.

If such an incentive mechanism were developed, it could be applied in a variety of contexts: time shifting of electricity use in the home, recycling, step programs for increasing exercise, and beyond. In the project reported here, road traffic congestion was targeted, in part because collaborations were much more readily established with transportation related agencies than with electric utility companies. Traffic congestion is a serious issue in many cities around the world. It has worsened considerably in the past few years, causing an enormous wastage of time and fuel. For example, a study (Schrank and T. Lomax, 2005) of several urban areas in the U.S. reports that in 2005 an estimated 4.2 billion hours of time and 2.9 billion gallons of fuel were wasted due to congestion. This amounts to a total loss of about $78.2 billion, up from $73.1 billion in 2004. See the cited U.S. Dept of Energy (2006) and U.S. Dept. of Transport reports for other reports of the effects of congestion in the U.S. In urban areas, increased vehicular traffic has also led to severe pollution and parking problems.

Objectives

Prabhakar aimed to develop a sweepstakes or raffle-like incentive program that would stretch the value of monetary rewards so that the energy behavior change would be maximized for a given amount of money. He also aimed to develop a computational system to support the program, and demonstrate the program’s effectiveness.

Methods

The Insinc project (analogous to the Capri, Steptacular and Instant programs; http://scsn.stanford.edu/projects.php) was developed to incentivize commuters to travel at
uncongested times by giving them different numbers of credits (corresponding to cash) for shifting to off-peak travel, mode shifting (from private to public transit), or recommending a friend - as monitored through transportation sensors. Then individuals could choose to participate in a simple game of chance – that looked similar to chutes & ladders - to win a shot at a larger amount of money. Formative work on the project showed that adding a simple game of chance and social networking greatly improved engagement with the system.

Insinc was launched on January 10, 2012, as a six month research pilot by Stanford University and the National University of Singapore with the principal aim of shifting Singaporean rush hour commuters to off-peak times using Insinc. Commuters were invited to register for the program online. Partnerships with National University of Singapore and the Land Transport Authority of Singapore facilitated marketing and participant recruitment efforts, and they supplied incentive payments for the participants, though Prabhakar developed and supported the online system for the program.

![Insinc transportation lottery website and game of chance.](image)

**Outcomes**

INSINC enjoyed high word-of-mouth recruitment. In six months recruitment reached 21,000 users. We define “peak-shift” to be the change in percentage of peak trips made by a group of users after they signed up for Insinc. Overall, 7.5% of all Insinc trips were shifted off peak. Users who made regular peak-hour trips before joining Insinc shifted their peak trips by more than 11%. A p < 0.05 level of significance was observed for utilization of public transit and shifting...
time of use in the optimal direction for those users. Thus, it was shown that entering individuals into a lottery and compensating only a small number of lottery winners achieved significantly shorter commute times and reduced fuel consumption and congestion. Publications related to this work can be found here http://peec.stanford.edu/energybehavior/projects/transportationlottery.php.

Future Work

The software developed in part through this project, and general approach, are being applied in other settings and to address additional societal issues.

3.3 Community Based Interventions

3.3.1 Girl Scout “Girls Learning Energy and Environment” (GLEE) Program

Investigators: Thomas N. Robinson, Nicole Ardoin, Hilary Schaffer Boudet, June Flora, Carrie Armel, Manish Desai

Partner: Girl Scouts of Northern California

Background

Community-based programs have been widely used with success in public health, and they have also been used in the environmental domain. For example, in the Hood River weatherizing project, initially less than 10% of customers signed up for a voluntary program in response to traditional marketing communications, but this number increased to 85% of households in 2 years when the project switched to relying heavily on local residents, such as Citizen Advisory Councils, schools and churches (Cavanaugh, 1995). Key advantages of using community programs include the ripple effect from word of mouth, enhanced learning and mastery through direct experience or observation of others, and the ability to provide personalized messaging. Further, these approaches can be cost-effective by tapping into pre-existing diffusion channels and making the intervention highly structured and easy to replicate.

Objectives

Our goal was to apply behavioral theories and methods used in public health promotion to increase children’s energy-saving behaviors and reduce family/household energy use. We targeted youth for several reasons. First, attitudes and values start developing at an early age (Bryant & Hungerford, 1977) and are difficult to change once established (Asunta, 2003). Moreover, the earlier children embrace sustainable lifestyles, the longer they have to influence families, schools, and communities to embrace sustainable activities and policies (Leeming et al., 1997). Finally, empirical research shows that it is effective to target children when trying to influence family and household behaviors (Cornelius et al., 2013; Damarell et al., 2013; Robinson & Borzekowski, 2006).

Methods
We evaluated two curricula – one focused on behaviors Girl Scouts and their families could do to save energy in their residences and another focused on behaviors Girl Scouts and their families could do to save energy in their food and transportation choices – in a clustered randomized controlled trial from December 2010 to February 2012. We worked with the Girl Scouts of Northern California to recruit 30 fourth- and fifth-grade troops and their families in Santa Clara, San Mateo and Alameda counties in the San Francisco Bay Area to participate in our study. (This represents 4% of the total of 748 fourth- and fifth-grade troops in the area.) We used multiple recruitment strategies, including placing advertisements in regular electronic mailings from the Girl Scouts regional office to troop leaders as well as in-person solicitation at monthly troop leader meetings.

Figure 21. Girl Scouts Girls Learning Energy and Environment (GLEE) website and badges.
Troops were the unit of randomization. Fifteen troops were randomly assigned to the residential energy condition and fifteen troops were randomized to the food/transportation condition after completing baseline assessments. By contrasting the two curricula – one focused on reducing home energy use and the other on reducing food and transportation energy use – each curriculum served as an active control for the other. This created a true randomized controlled experimental trial, the strongest design for testing causality. The evaluation consisted of a baseline survey for all participants prior to the first troop meeting, followed by five sessions of the relevant curriculum, a posttest survey after the fifth session and a follow-up survey conducted between 2.5 and 10.7 months after the fifth session. All Girl Scouts received one of the experimental curricula, taught by a trained member of our research team with troop leaders present.

Troops ranged in size from 4 to 21 girls, with an average of 11 girls per troop. By troop, survey return rates at baseline averaged 98.7% for Girl Scouts and 88.5% for parents. At post-test, survey return rates averaged 94.9% for Girl Scouts and 91.4% for parents. At follow-up, survey return rates averaged 83.3% for Girl Scouts and 82.5% for parents. Multiple imputation techniques were used to handle missing data, for a total of 313 Girl Scout participants (149 residential; 164 food/transportation) and 318 parent participants (151 residential; 167 food/transportation). In addition to self-reported behavior change through the surveys, we also collected data on monthly electricity and gas usage from one-third of the parent participants; analysis of this data is currently underway.

**Outcomes**

To test these interventions, we conducted a cluster-randomized controlled trial (RCT) with 30 4th and 5th grade Girl Scout troops and their parents in Northern California (330 families). We found that Girl Scouts in troops randomly assigned to the residential energy intervention significantly increased their residential energy-saving behaviors by 49% following the intervention, and 27% persisting after more than seven months of follow-up, compared to controls. Parents also made significant behavior changes. Girl Scouts in troops randomly assigned to the food-and-transportation energy intervention significantly increased their food and-transportation energy-saving behaviors by 7%, compared to controls (Boudet et al. in 2016). These home energy behavior changes for girls and their families translate into estimated annualized total electricity reductions of 5% and gas reductions of 3%.

**Summary of Dissemination Activities**

GLEE was one of the projects that continued into a Plus Up funding period. At the end of the GLEE randomized controlled trial, we embarked on a dissemination process in which the translated the research form of GLEE into a program of practice aimed to retain the critical behavior change elements of the original research. To that end we:

- Developed a five-module online training program (a MOOC) to prepare Girl Scout troop leaders to teach GLEE to girls and engage parents,
Created an accompanying website (http://glee.stanford.edu) for girls, parents, and troop leaders, a Facebook page, and a monthly email newsletter.

Developed a contact list of all appropriate council staff contacts (such as STEM and Program staff) in 109 of the 112 US Girl Scout councils, all of whom we have had multiple email or phone contacts.

By end of the ARPA-e funding, 32 Girl Scout program staff and 216 troop leaders in 30 states had registered for the GLEE online course.

By end of the ARPAe funding, 72 leaders were working in or had worked through the course – potentially reaching nearly 900 girls and their families.

Stanford Staff participated in multiple scouting events in Oregon and California reaching hundreds of girls and tens of leaders personally to expose them to GLEE activities and the online course.

Embedded analytics reveal that our recruitment campaigns are working; during our recruitment campaigns and publication of our outcome paper we observed spikes in promotional video viewing, website views, and registration.

Additional funding has been received to support further scaling and sustainability of the GLEE program.

3.4 Integrative Project

3.4.1 HowPower

Investigator: Carrie Armel

Background

Utilities, cities, and others must meet increasingly aggressive energy reduction targets. For example, energy efficiency resource standards were enacted by three states in 2010 and this number increased to 24 in 2014. Like with the resource standards, energy efficiency will be a strategic choice in meeting other reduction targets; six unique climate studies all conclude that energy efficiency will be the most used and least-cost option for the US to reduce CO₂ (Wood 2015). This claim is consistent with its historical effectiveness; since 1973, U.S. energy and carbon intensity reduced most from energy efficiency (2%/yr, 57% total) compared to other efforts. Consumption reduced from what it would have been by over 3x the increase in all forms of domestic energy production (Sweeney 2016). Furthermore, near-term implementation means savings accumulate for years before game changing technologies are ready.

Unfortunately, there are currently no deep energy-saving, scalable, cost-effective residential or small and medium business (S&MB) energy efficiency programs. Such programs are needed to achieve adoption of the technologies and practices required to attain the aforementioned
targets. Energy behavior change programs are estimated to have the potential to achieve 20% savings over the population in the residential and S&MB sectors, using existing cost-effective technologies and practices, with limited effort and lifestyle changes (McKenzie, 2013; Gardner and Stern 2008; Laitner, Ehrhardt-Martinez, and McKinney 2009). In comparison, Opower, the current “gold standard” program in the space, achieves 1.5% energy savings over populations (Alcott, 2011). Such “state-of-the-art” behavioral efficiency programs, as well as traditional utility programs, have low customer engagement and limited energy savings because they are time consuming, effortful, and boring. Furthermore, most offerings are dated (e.g. mailer, email, rebate, informational web page) and rely on rudimentary behavioral science. They typically provide support of utility websites and mailers with graphs, tips, analytics, and neighbor comparisons. Other types of programs that could achieve deeper savings, such as traditional energy audits and ESCO programs, have not been cost-effective due to labor costs.

Many of the ARPA-e Energy Behavior Initiative projects were developed with the goal of addressing these issues. The Integrative Project aimed to incorporate the positive learnings from these initial projects, as well as additional impactful elements, into a new behavioral program.

**Objective**

This was a newly funded project during the Plus Up period.

The Integrative Project aimed to develop a new behavioral program capable of scale, depth of savings, advanced EM&V capabilities, and marketplace sustainability. Additional goals were to evaluate the program, and then disseminate it to utilities and third party vendors.

**Methods**

The core program – HowPower – was developed through significant research, and iterative development and testing. HowPower is a browser application or “web app” that works on any device, with no download needed. In contrast to most commercially available products in this space which focus on analytics and tips but neglect reach and action execution, this solution emphasizes:

- An entertaining and frictionless experience to enable **deep energy savings** per user. The focus is on getting the user from tips-to-action, instead of allowing user to drop off a cliff after receiving tips. The user is supported with a dynamic concierge; triaged to determine which actions and products are right for them; and has their barriers overcome for a frictionless experience (e.g., narrowing of choice set, pre-populated order forms). Furthermore, users are entertained with narrative (the most effective behavior technique in health fields) tied to the concierge, and surrealistic ‘surprise’ pictures (e.g., a house melting on a dashboard to indicate the magnitude of solar radiation absent shading). The use of entertainment stems from the insight that, although people

---

3 This refers to *any* energy efficiency actions, including adopting technologies, not just habits which is a definition commonly used by utilities.
complain they are too busy to take action, they average 3+ hours/day on entertainment media.

- Diverse social channels to garner **low-cost widespread adoption**. Examples include online social media (e.g., surreal energy pictures embedded in Pinterest, blogs, and content provider websites; viral YouTube videos), Google and Facebook advertising, and community based programs such as GLEE described above – which all link to the web app.

Together, deep and widespread savings produce **large aggregate savings**.

Furthermore, utilities need to verify the savings through evaluation, measurement, and verification (EM&V) requirements. Thus, the application was also designed to show stronger causality than other products. By linking web app data indicating when actions are completed with smart meter data, the application provides **proof of savings**.

Each screen was carefully designed, with iterative user testing, to ‘pull’ the user through the app. The experience flows as follows: (a) Users begin with a brief Intro Video where a concierge character orients and motivates them to enter (2) the Hub, with an energy end use (e.g., heating, cooling, electronics) carousel, and dynamic humorous concierge, whose body language and speech bubble change according to content the user is viewing.

The Hub leads to several discrete modules, each supporting a user in executing an energy saving action (which were selected for being high impact but low cost and effort); each module includes: (a) After choosing an action from the Hub, the Surprise Screen shows the user the “surprise” picture chosen on the hub in more detail with the corresponding explanation on why this action is so impactful. (b) Next, the Triage screen identifies the user’s use case to determine their best option for reducing energy on a specific end use (e.g., window film), and how we will overcome barriers for that option. (c) The Steps screen, modeled off recipe and how-to sites, breaks the action into brief steps illustrated with pictures and animated gifs. At the top, the user’s “to get” list links to, (d) The Recommendation Screen,

![Figure 22. HowPower browser application.](image)
where “Best for...” descriptors guide users in choosing a product from a short list, a pop-up asks for contact information to send them a reminder, and the item(s) are pre-populated in an Amazon cart for their purchase.

To validate HowPower, our team performed extensive user testing with hundreds of participants using in-home needfinding interviews, mechanical Turk surveys, in person and remote usability testing, Google Adword experimentation, and Google Analytics analysis of HowPower site patterns and activity.

Outcomes

From Dec 18 – April 14, 2016, with a preliminary mostly text version of the application, 15,000 unique users from Google Adwords search used the web app, with 6,000 of those viewing three or more pages. We anticipate greater adoption and depth of use with the enhanced version described here.

Future Work

The next step will be a validation trial to quantify depth of savings and uptake. Pecan Street in Texas has agreed to collaborate on a trial that would provide access to homes with high resolution energy data, and, separately, a major utility as well as a third party vendor that sells to utilities have also suggested collaborating on a trial.

Dissemination potential appears promising. Professionals in the venture capital community say there is a good business case for a behavioral program that can achieve modest depth of savings and scale. They also said that our pitch deck and revenue models are solid - the revenue approaches include: (1) Amazon affiliates program to obtain a commission for selling energy efficiency products; (2) lead generation for renewables and retrofit contractors; and (3) licensing of product to utilities and/or third party vendors who sell to them. Further, utilities and third party vendors (who would require less time and effort than going direct to utilities, given vendors already have multiple established utility relationships) have shown significant interest in the product, and want results from a trial in order to estimate savings. Companies in the space who are currently licensing products to utilities have also shown interest in partnering, particularly given that HowPower was designed to complement rather than compete with current offerings.
CHAPTER 4: Evaluation and Modeling

Three projects focused on developing methods for evaluating the effectiveness of energy programs, and modeling the effectiveness of interventions to guide future work.

4.1 Google Power Meter Evaluation

Investigators: Sébastien Houde, Annika Todd, Anant Sudarshan, June A. Flora, and K. Carrie Armel

Background

Readily available, easily accessible, real-time information delivered via technology is reported to produce important declines in residential energy consumption (Faruqui et al. 2010; Ehrhardt-Martinez et al. 2010). Designing interventions that use feedback technologies and rely primarily on information as a means of changing energy behaviors have been promoted as cost-effective policies (Fischer 2008; EPRI 2009) and possible alternatives to traditional price incentives (Allcott and Mullainathan 2010).

Estimates of the energy savings from feedback technologies vary widely, from none to as much as 20 percent (Faruqui et al. 2010; Ehrhardt-Martinez et al. 2010). There are three main factors at the source of this heterogeneity in outcomes. First, studies have employed different research designs. Second, the features of the feedback technology, such as timeliness, data display, interactivity, sociability, and controllability play a significant role in inducing energy reductions and have varied substantially across studies. Third, there is significant heterogeneity in the characteristics of the population of consumers participating in feedback interventions. Although several studies have looked at the impact of feedback technology, providing insights as to how study design, features of the technologies and characteristics of the people using them impact the energy savings estimates, several questions remain. To determine if feedback technologies are cost-effective measures to manage energy demand it is necessary to assess whether they provide persistent energy savings and how they change consumption profiles. Previous studies have remained silent on these questions due to limitation in study design and data available (EPRI 2009).

Objectives

The goal of this research was to provide an estimate of the potential for electricity savings for households that have access to real-time feedback technology, and to document how this technology changes consumption profiles and impacts the persistence of energy savings. The feedback technology, Google Powermeter, resembles the technologies being deployed by several utilities in the US and elsewhere.

Design and Methods

We used a randomized controlled trial to overcome issues of selection bias and to estimate treatment effects. Households participating in this study were recruited in collaboration with
Google, both in their California offices and with several offices across continental US. Employees (N=1743) from the company voluntarily enrolled their households for the study. As part of enrollment all participants were required to install The Energy Detective (TED) device (purchased by the company), complete an online survey and be randomly assigned to no-feedback (untreated control) or feedback (treatment) conditions. Only households in the feedback treatment condition were given access to the feedback technology initially. Households in the control condition were given access to the feedback technology after three months. The study took place between February, 2010 and October, 2010.

Figure 23. Google Powermeter evaluation.
Background inset shows Google PowerMeter display with graphs of electricity use data in 10 minute intervals, use compared to others, and current use compared to one’s past use. Foreground graph shows experimental trial results; energy use was reduced by around 6% when individuals were first exposed to the feedback. Persistence could likely be improved significantly through repeated exposure (Alcott and Rogers, 2013).

**Results**

**Summary.** Over the period of the field trial, March through October 2010, we found a statistically significant reduction in electricity use of 5.7 percent. However, an examination of persistence of effects over time shows that there is only a brief period of significant reductions in electricity consumption; by week four all statistically significant reductions have ended. In
examining time of day reductions, the largest reductions were observed initially at all times of the day but as time passes, morning and evening intervals show larger reductions. Evening reductions faded but morning reductions were sustained for eight weeks. However, the return to baseline in other day and evening periods cancelled out statistical significance in overall reductions. Thus, overall statistically significant reduction effects lasted for four weeks.

**Overall effects.** Ideally, our analytic models would be identified using electricity billing data for the months preceding the experiment; however, we did not have access to this data. Our models rely on the fact that after May 28, 2010 both the control group and the treatment group had access to the feedback technology. Under the assumption that the treatment effect is constant over time, when households in both groups have access to the technology, any differences in average consumption levels can be attributed to household specific fixed effects.

Results from the estimation of the fixed effects model show that the average treatment effect consists of a decrease in electricity use of 5.7 percent per hour (this works out to about 0.05kwh in absolute terms), significant at the 5 percent level.

**Time of Day effects.** An aspect of our primary research goal is to use our unique real time data to inform our understanding of time specific electricity use and reductions. We distinguish between periods of high and low household membership activity, which allows us to infer whether savings are attributable to habitual behavioral change (such as turning off lights) or to one time behaviors that are more structural in nature (such as installing energy efficient appliances or house insulation). Change in habitual behaviors should lead to reductions that are observable at periods of high occupancy while the latter class of actions should lead to reductions in the baseload levels of consumption.

The largest reductions in electricity consumption due to feedback occur during the morning and evening peak periods: between 5 am and 10 am, electricity consumption decreases by 12.2 percent in average and between 8 pm and 11 pm electricity consumption decreases by 8.2 percent on average. While energy savings during the middle of the day and night are insignificant, savings during the morning and evening peaks are large and significant. Savings occur at periods when household members occupy the house and engage in household functions, such as eating, entertainment, cleaning and household maintenance. Based on this finding, we argue that electricity use reduction during household activity is consistent with changes in energy behaviors that pertain to habits.

**Persistence at different times of day.** We find that in the first two weeks after having access to real time electricity feedback, electricity consumption decreases in all time periods. Starting at the third week, reductions during the day (10 am - 4 pm) and the night (11 pm - 4 am) start to fade away. In the long-run, only reductions during the morning and evening peak periods persisted.

**Future work**

This paper points out the challenges of conducting rigorous experimental work in the field; sufficient experimenter control, adequate funding, and expert staffing are all necessary for robust trials of feedback and energy consumption. In this paper we also discuss some of the
inherent challenges of this type of work; the heterogeneity of electricity consumption, the relatively low predictability of levels of that consumption, size of samples needed for detection of effects in the face of large heterogeneity, and necessity of data collected over periods of one year or more to adequately assess seasonal and weather effects – and, importantly, achieving persistence through interventions of this type. In addition to attempting to address these issues, future work can expand on analysis techniques to improve learnings from trials such as these, as well as to offer utilities and government entities trusted approaches to quantifying the impacts of behaviorally oriented programs.

4.2 Social Media Analytics through Twitter Explorer

Investigators: Martha Russell, Markus Strohmaier and Jan Pöschko, Technology University of Graz, Austria; Rafael Perez and Neil Rubens, University of Electro-Communications, Tokyo; June Flora, Jiafeng Yu and Marc A. Smith, mediaX at Stanford University

Background

The term social media describes the online tools and platforms that people use to share opinions, insights, experiences, and perspectives with each other. Social media can take many different forms, including text, images, audio, and video. Understanding conversations in online social media has the potential of providing program planners and communication campaign managers unique insights into individuals’ thoughts and verbal productions about energy efficiency and climate change. With the increasing adoption of social media (73 percent of teens and 72 percent of young adults (Lenhart et al., 2010)), new opportunities are available for studying the role of online conversation in persuasion.

In 140 characters or less, the concerns, interests and public narratives about energy efficiency and climate change can be identified through tweets. Twitter conversations also reflect the social aspects of information diffusion through conventions such as retweets and other conversational responses, through the membership and social distance of these conversations, and in overtime changes in frequency and form of the networks. With Twitter, a particularly popular type of social media that has proven relevant in a number of societal challenges and conversations recently, the social response to societal or national events, as well as to media coverage of these events, and persuasive communication campaigns can be observed. Twitter is both a medium and the message (Savage, 2011).

Objectives

Our goals were to track and analyze social media conversations related to energy efficiency and climate change in order to gain insights about consumer attitudes and behavior. By understanding the larger context of public sentiment about changing energy behavior, we sought to create insights on the context – cultural, economic, social – in which the Stanford Energy Behavior Initiative intervention projects were being conceived and implemented. Our
broad research question asked, “How can social media reflect consumer sentiment about energy and campaigns to reduce residential use of energy.”

Methods

Using an ecolinguistic taxonomy of terms to describe energy use opinions, energy efficiency behaviors, frames, metaphors, technologies, and also to determine standard sources of energy information such as the Department of Energy (DOE) and Environmental Protection Agency (EPA), we captured Tweets containing those terms, parsed identified elements of the communications, curated the data, and archived the data for access. The analysis of this ecolinguistic-filtered social media was aimed at understanding the frequency, context and potential persuasive influence of social media conversations about changing energy behaviors (Russell et al., 2011).

We collected data from the messages, users and content of the hashtags that occurred between September 3, 2010 and March 31, 2013. We operationalized conversations to be overtime mentions of predetermined ecolinguistic terms and co-occurrence of related hashtags in Twitter. Three techniques were used to cull discernable patterns from the large quantities of data and portray the data visually: content analysis of the full Tweets, network analysis of co-occurring hashtags, and semantic analysis of the co-occurring hashtags and their authors.

Data were acquired on a daily basis by utilizing the NodeXL Twitter Importer module (Smith et al., 2009), which captured the latest messages containing energy related keywords. The dataset was then parsed and analyzed by utilizing Hadoop Map Reduce distributed processing on the Amazon’s EC2 computing cloud. Data for these Tweets was then passed to other applications—Excel, NodeXL, Gephi and TwitterExplorer—for further analysis and visualization. TwitterExplorer (Russell et al. 2011), developed separately, was used to analyze and visualize the latent semantic structures embedded in energy-related conversations on Twitter. TwitterExplorer visualizes semantic relations between terms used in Twitter messages based on different aggregation and similarity measures; semantic similarity between terms is calculated based on co-occurrence of terms within messages (Tweets).

We analyzed samples of the 3+ billion filtered Tweets that used the ecolinguistic terms included in this study. Several analytical lenses were tested in this study: frequency, periodicity, valence, co-occurrence, and context. Through several methods, we were able to describe snapshots and detect changes over time in the conversation as well as identify conversation stimulating events, such as national policy, new technology launches, and media events.
Outcomes

The data, tools and initial analysis of this study represent first steps towards more refined analytical approaches that help understand the large scale conversations taking place on Twitter and elsewhere. This study demonstrated the feasibility of using data mining techniques to gather and analyze vast amounts of data from ongoing social media conversations and of analyzing the data for meaningful metrics that describe conversations about energy consumption behavior.
Our exploration confirmed that conversations about energy-related issues are, indeed, taking place in social media, specifically Twitter, and that these communications can be studied to better understand how to use technologically-enhanced word-of-mouth to stimulate user-generated persuasion. Using content analysis of full Tweets, network analysis of co-occurring hashtags, and semantic analysis of the co-occurring hashtags and their authors, this preliminary investigation identified descriptors, concerns, actions, and issues. We confirm that studying Twitter communications can provide actionable means for assessing engagement, identifying influencers, and identifying word-of-mouth communities that can accelerate change in energy efficiency behaviors.

Using network analysis of hashtags we analyzed and visualized contextual relationships of among the salient terms used in social conversations and identified several clusters of related issues, revealed by the co-occurrence of hashtags. We used social and linguistic structures of communication (repeat communications and varied types of Twitter communication conventions such as pictures, hashtags, retweets, and URLs) to analyze the establishment of self-organizing communities of consumers. These communities share many of the characteristics of issue publics, and further research on similarities and differences to other issue publics is needed in order to understand how to create, grow and sustain word-of-mouth persuasion for energy behavior change. We demonstrated tools that permit visualization of vast quantities of user-generated content about energy and sustainability.

**Future Work**

Based on our initial results, we recommend continued collection of data and development of analytical methods and tools that can: track public opinion related to energy consumption; analyze domain-specific, user generated content on social media platforms; identify and track indicators such as semantics and social roles; identify and explore patterns and disruptions; identify and benchmark grassroots resources such as author networks; characterize opportunities for resource transformation; and build semantic models to understand the aggregations of conversation streams.

To accelerate exploration of these important issues, we are making the Twitter Energy data available for other researchers. Against the urgency of climate change and the need to mobilize widespread changes in energy consumption, we encourage other researchers to join us in a research agenda that includes: analyzing and characterizing energy consumption behavior; tracking public opinion related to energy consumption; analyzing domain-specific, user generated content on social media platforms; identifying and tracking leading indicators of attitude and behavior change; and identifying patterns and disruptions that can accelerate change and provide alerts for emergency response.

### 4.3 Diffusion Modeling of Behavioral Interventions

**Investigator:** Jeff Shrager

**Background**
A central goal of the Stanford ARPA-e project is to achieve energy reductions by using feedback from sensors to inform decisions regarding which behavioral manipulations to deploy. Given this goal, it is important to be able to predict as accurately as possible, how different possible manipulations will move the needle toward this goal. Even if we cannot predict the exact amount of energy that might be saved under various manipulations, we might be able to at least rank-order the possible manipulations. Such a ranking, crossed with cost, would be used in choosing which manipulations to employ.

The most accurate way to conduct such analyses is, of course, through real, clinical-trial-like experiments, and this is the approach taken by some other Stanford ARPAe subprojects (e.g., the Girl Scout project). Unfortunately, real world experiments are extremely time consuming and costly, and can only be conducted on a small number of manipulations. Moreover, because one would assume that the efficacy any such manipulation will depend upon many features of the setting in which they are applied, one would have to run a huge number of experiments to tease out all of these interacting effects.

Although a simulation can never stand in fully for a clinical trial, in many fields they have proven extremely useful in getting a rapid (and very inexpensive) handle on the general response topology. For example, every modern biomedical clinical trial is computationally simulated before being actually run on real patients. These simulations provide very accurate estimates of the number of patients that one needs to run in order to get meaningful results in the real trial.

Objective

The goal of this project is to use computational simulation to estimate the differential efficacy of behavioral manipulations on energy usage applied in various environments. It aimed to specifically develop a general methodology to enable analysts to propose a novel manipulation, and then rapidly and inexpensively explore its potential efficacy under a variety of conditions, and, if desired, to rank these predictions against other proposed manipulations (or, used in another way, to choose in which real-world settings to deploy different manipulations because different manipulations may be predicted to have differential success in these).

Methods

Multi-agent simulation was used in this modeling effort. Muti-Agent simulation has a long history, beginning with Simula67, and leading to SmallTalk, and thence to all modern object-oriented programming languages. It has been extensively applied in many domains including military simulation (where gaming is very big), hardware (e.g., Verilog), socio-economics, and even in politics. Multi-agent simulation can model very complex systems with very complex dynamics in great detail, heterogeneous populations (without averaging), complex internal decision making algorithms, complex communications structures, and are generally relatively easy to develop. It is most powerful and useful when inter-individual interactions are high and the decision choice algorithms are influenced by changing local social parameters. Although multi-agent simulations are most interesting in these domains, they can be powerful in individual (non-social) domains as well, especially in this case where there is a tacit
communication via the community’s influence on the shared environment and on the shared
electrical grid. Also, the utility of an agent based model would in part be in communicating
analysis results and exploring consequences from the model – that is, agent based models are
probably much better than the way we currently communicate statistical outputs. (A
disadvantage of multi-agent simulations is that they are more difficult to utilize for abstraction
and optimization; math modeling / algebraic analysis is preferred in this case.)

In the multi-agent simulation, the following steps were carried out: 1. Create the model;
define the environment (e.g., weather, oil supply); define each agent’s possible actions and
decision algorithm; define the communications (i.e., ”social”) graph. 2. Parameterize the decision
algorithm, social graph, and environment: Set search ranges for target parameters (independent
variables, e.g., interpersonal communication graph structures; intervention manipulation
parameters, such as thresholds for interventions and intervention strengths; and so on); set as
many ”non-target” parameters as are available (e.g., default pricing of categories of energy;
population size and subcategories, and default energy usage averages for each population
subcategory); Guess at non-target parameters for which data is not available; randomize over all
others. 3. Run across search ranges of target parameters and gather data (dependent variables,
usually final energy utilization). 4. Interpret results using qualitative dynamics, graphical
visualizations, or fitting of the output data to algebraic models (possibly for algebraic analysis,
e.g., optimization).

Outcomes

Two models were developed. The “A” model (developed with S. White) was based upon
existing modeling technologies and was used to pilot the basic modeling theory and test the
hypothesis that very simple manipulations would produce observable differentials in energy
utilization. Based upon results from the “A” model experiment, we subsequently developed a
“B” model, called ESim, which was more general, and was built on a cloud-based modeling
infrastructure that would, in principle, enable other modelers to experiment with ours and
other similar models. Statistical results from these models suggest that approximately 10%
ergy savings could be expected on average, but that this depends critically upon the
communication model (i.e., social network structure and interaction density).
Figure 25. Visualization of illustrative results from diffusion modeling "experiments". The graph illustrates the impact of social distance and intervention exposure on energy consumption. “Clowned” is shorthand for any behavioral intervention.

**Future Work**

A model to guide program design and selection has not been possible in the past, but is now feasible due to our dramatically improved ability to quantify the effects of behavioral programs with sensor data, to an accumulated large body of empirical behavior change literature from which effect sizes can be drawn, and to sufficiently mature modeling approaches. The project reported here was ambitious in attempting this, and much work remains to be done to achieve the goals described here. The tool developed could be used to make predictions, which could then be tested in field trials. Parameters such as effect sizes for a wide variety of behavioral approaches, derived from existing research and evaluation studies, could be incorporated to improve the model. As more studies are behavioral program pilots are run, the true empirical responses to the corresponding programs can be incorporated into the model. A description of the work and a user’s guide is online, and the code (both for the model, and for the graphical user interface so that non-programmers can utilize the model to make predictions) is available for others to use; please contact the lead investigator for access jshrager@stanford.edu (it was taken down from the web due to server issues, but could be returned to operation).
CHAPTER 5:
Summary and Project Outputs

5.1 Project Outputs and Impacts

5.1.1 Summary
This initiative has met its objectives to develop an array of components needed in a system that utilizes smart meter and other sensor data, communication technologies, and behavioral approaches, to achieve significant energy savings. We accomplished major work on 20 projects tied to this goal.

A summary of project outputs is included in the pages that follow, covering:

- Website Featuring Project Work Results: https://peec.stanford.edu/research/behavior-initiative
- Partnerships/Collaborations and Networks
- Key Product Outputs and Results
- Invention Disclosures and Patent Applications
- Media, Awards, and Follow-On Funding
- Journal Articles and Papers (see final References section below; reference cited and those associated with funded work are separated under each project)
Table 3. Summary of projects, partners, and product outputs.

<table>
<thead>
<tr>
<th>Proj. #</th>
<th>Project Name</th>
<th>Investigators &amp; Partners</th>
<th>Potential Applications and Key Product Outputs (Publications for all projects can be found in the references at the end of this document and at <a href="https://peec.stanford.edu/research/behavior-initiative">https://peec.stanford.edu/research/behavior-initiative</a> )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Technology</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Communication Network</td>
<td>– Levis, Kazandjieva (CS)</td>
<td>Established home area network (HAN) Internet standard for use at scale; use the high granularity data from plug monitoring mesh network that was developed to inform future computing energy standards. First, helped establish the first Internet standard for home area networks (HANs), which is being adopted by industrial consortia such as WirelessHART and ZigBee, and is designed to support innovation. Second, developed a wireless power plug meter that automatically joins a self-assembling, ad-hoc wireless mesh network; the deployed network of 200 meters allowed the team to publish detailed data at a scale orders of magnitude greater than other similar efforts, and is informing future energy standards for computing systems</td>
</tr>
<tr>
<td>2</td>
<td>Stanford Energy Services Platform (ESP)</td>
<td>– Armel, Reeves (PEEC) – Bonsai Development Corp.</td>
<td>Improve the ease of implementation and evaluation of energy saving interventions that use sensor data via software platform. Software platform that includes data collection services, a database, analytics, and graphical user interface templates for behavioral program deployment and experimentation at Stanford and beyond</td>
</tr>
<tr>
<td>Algorithm</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Segmentation Algorithms</td>
<td>– Fischer, Rajagopal, Albert, Kavousian (CEE) – Google, PG&amp;E</td>
<td>Energy use and demand forecasting; demand response programs. Software to segment commercial and residential customers based on their smart meter data. Energy consumption patterns. This information can be strategically and cost-effectively used to target customers for energy savings; utility trial in planning phase</td>
</tr>
<tr>
<td>4</td>
<td>Learning and Automation</td>
<td>– Aghajan, Khalili, Chen (EE)</td>
<td>HAN energy efficiency (EE) and demand response (DR) automation. Software based on adaptive machine learning algorithms utilizing appliance and sensor data to improve TV and lighting automation on the dimensions of user activity, user preferences, and energy savings</td>
</tr>
<tr>
<td>5</td>
<td>Disaggregation Technical and Policy Survey Paper</td>
<td>– Armel (PEEC), Gupta, Shrimali, Albert – Bidgely, Venrock</td>
<td>Support the development of algorithms for improved demand side management (DSM). Comprehensive survey paper assessing the benefits of disaggregation (i.e., the statistical separation of the whole building energy signal into appliance level energy use data), overview of state of the art algorithms and their performance, and smart meter data suitability for these</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Residential Energy Disaggregation Dataset (REDD)</strong></td>
<td>Support the development of algorithms for improved DSM. Data set collected and made available for developers to improve, train, and test disaggregation algorithms; extensively used.</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Disaggregation Algorithms</strong></td>
<td>Support the development of algorithms for improved DSM. Disaggregation algorithms were developed using sparse coding methods to advance the state of the art.</td>
<td></td>
</tr>
<tr>
<td><strong>Target Behaviors</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Energy Behavior Taxonomy</strong></td>
<td>Energy saving action recommendations for use in software or programs. Database of 250 energy saving actions, their attributes and impact, and barriers. Implementation in Bidgely Inc.'s online recommendation system.</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Identification of Innovative Energy Behaviors</strong></td>
<td>Energy saving action recommendations for use in software or programs; development of new energy saving technologies. Opportunity map identifying energy reducing practices and technical insights from other cultures and time periods, quantifying their potential energy saving impacts across U.S. climate zones, if adapted for developed nations.</td>
<td></td>
</tr>
<tr>
<td><strong>Behavioral Interventions</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Media Interventions</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Online Game</strong></td>
<td>Energy savings through media programs. Online game utilizing real world energy data, social competition, and retraining of habits through reinforcement; laboratory and field studies suggested changes in energy saving behaviors and consumption.</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Social Norms</strong></td>
<td>Energy savings through media programs. Web application that helps consumers track energy use and receive tips; 800+ users, embedded experiments showed the effectiveness of thematically organized tips and collective action framing.</td>
<td></td>
</tr>
</tbody>
</table>
| 12 | Immersive Reality | – Bailenson, Bailey, Flora, Armel, Voelker, Reeves (Comm) – DraftFCB | Energy savings through media programs
Experimental evaluation of the utility of an immersive virtual environment in promoting energy saving behaviors, with results suggesting that vivid visualizations of energy consumption (e.g., amount of coal instead of KWh) may be more important than the personalization afforded by avatars. |
| 13 | Facebook Applications | – Banerjee, Flora, Sahoo, Bhansali, Greenspan, Khakwana, Liptsey-Rahe, Madres, Manley, Omer, Rajendra, Scalammini, Wong, Stehly, Voelker (ME/Design) | Energy savings through media programs
Three Facebook applications to motivate energy reductions and online experimental evaluations of these |

Incentive Interventions

| 14 | Appliance Calculator | – McClure, Houde, Armel (Psy) | Energy savings through nudge
Online appliance calculator application; 60,000 users via Google Ads, embedded experiments evaluated the effectiveness of information and framing tools for guiding the purchase of energy efficient appliances and electronics, e.g., selection of 10-20% more energy efficient refrigerators from default sort order (manipulations informed in part by a study with Sears) |
| 15 | Raffle Incentive | – Prabhakar, Merugu, Pluntke, Gomes, D. Mandayam, Yue, Atikoglu, Albert, Fukumoto, Liu, Wischik, Rama (EE) – National University of Singapore, Land Transport Authority of Singapore | Energy savings through incentives
Online software developed for a raffle-like incentive program to motivate energy savings; 21,000 users, ~10% of trips were shifted off peak to reduce congestion and associated fuel waste |

Community Intervention

| 16 | Community Program | – Robinson, Ardoin, Boudet, Flora, Armel (School of Med) | Energy savings through community based programs
Girl Scout “GLEE” curricula; 30 troop study, showed significant changes in self-reported home energy saving actions for both girls and their parents |
<table>
<thead>
<tr>
<th>#</th>
<th>Project Description</th>
<th>Participants</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>17</td>
<td>Google Powermeter Evaluation</td>
<td>Houde, Sudarshan, Todd, Flora, Armel (MS&amp;E) - Google</td>
<td>Benchmark for energy saving programs. Field trial and evaluation of Google PowerMeter impact using analysis tools from economics; 1000+ participants, 6% energy savings in the first two months (much longer persistence has been reported in other studies with different conditions). The study provides a benchmark, and illustrates experimental and analysis methods.</td>
</tr>
<tr>
<td>18</td>
<td>Twitter Explorer</td>
<td>Russell, Rubens, Flora (Comm) - University of Electro-Communications, Tokyo</td>
<td>Tracking uptake and changes in the online social conversation about energy. Twitter Explorer, a software tool for collecting all tweets containing any of ~150 energy or climate change words, as well as an analysis of the online social “conversation” about energy efficiency for over a year of data.</td>
</tr>
<tr>
<td>19</td>
<td>Diffusion Modeling</td>
<td>Shrager (Symbolic Systems)</td>
<td>Simulating impact of envisioned programs. Simulation tool to predict diffusion in behavioral interventions based on parameters such as time, behavioral technique used, and social network distance and type. More sophisticated tools could eventually lessen time and cost of developing interventions.</td>
</tr>
</tbody>
</table>

CS = Computer Science  
E-IPER = Emmett Interdisciplinary Program in Environment and Resources  
EE = Electrical Engineering  
FS = Freeman Spogli Institute for International Studies at Stanford  
H-STAR = Human Sciences and Technologies Advanced Research Institute  
ISB = Indian School of Business  
MS&E = Management Sciences and Engineering  
PEEC = Precourt Energy Efficiency Center
Elaboration of Key Product Outputs and Results

The following elaborates on the product outputs and results for each of the projects. Complementary workshops and other support efforts were also conducted.

TECHNOLOGY

1. Communication network. Levis, in collaboration with a many others, helped establish the first Internet standard for home area networks (HANs), which is being adopted by industrial consortia such as WirelessHART and ZigBee. Specifically, they created an open standard for TCP/IP in home area networks (HANs) as well as an open-source reference implementation of the standard for others to copy, extend, re-use, and improve. This technology will provide greater freedom in data collection, representation, storage, and communication between devices of different manufacturers, as well as lower the barriers to entry, all leading to innovations and improvements in human interfaces to sensor-actuator networks.

   As a second deliverable, this team developed a wireless power plug meter that automatically joins a self-assembling, ad-hoc wireless mesh network to deliver data to collection points (see Figure 1). The open-source design has been used by several follow-on efforts by other groups. Further, the deployed network of 200 such meters in the Stanford CS building for two years to obtain long-term, fine grained power draw measurements allowed the team to publish detailed data at a scale orders of magnitude greater than other, similar efforts, as well as establish the basic methodologies one should follow to measure computing energy. These results are being used by several green computing companies to write future energy standards for computing systems.

2. Stanford Energy Services Platform. This provided the computational backbone for our behavioral interventions and includes three layers: data collection and storage (user, energy, website activity, weather, property, and project data), services (analytics like baselining and peer comparison, registration, surveys, experimental condition assignment, alerts), and presentation (API, widgets, drupal modules). Unfortunately, given the rate of change of software, by the end of the project there was new more efficient software that could support upcoming behavioral interventions.

Algorithms

3. Segmentation and targeting algorithms. We utilized anonymized data from 250,000+ California utility customers with at least one years worth of hourly smart meter data, as well as 10 minute data collected on over 1,000 Google employees. The algorithms developed decompose a customer’s consumption into daily load shapes; load shapes are then analyzed in aggregate to obtain a small number of typical loads shapes that characterize the whole population. These shapes can be then utilized to build behavioral models for customers; examine features such as variability, amount of kWh consumed, and thermal response; and more effectively target customers for energy programs. Finally, we developed innovative methods to quantify the energy efficiency of buildings.
During Phase II the system was set up at PG&E for their commercial use, and other utility partners are in progress.

4. Learning algorithms to enhance automation. This project automated the activity of TVs and lights based on human activity, preferences, and energy savings criteria. Data was collected from the electronics and also low resolution cameras that monitored human activity, and machine learning algorithms were developed that incorporated user models and decision-making to predict user behavior, as well as user feedback for refinement. The models were tested with real world data.

Disaggregation. This is the statistical separation of the whole building energy signal into appliance level energy use data. There were three discrete projects:

5. A comprehensive review paper evaluating the benefits of disaggregation, algorithm requirements, and the ability of smart meters to meet these requirements - and which has had tens of thousands of downloads to date, and helped spawn an annual workshop.

6. Development of a high frequency data set, to aid algorithm development, testing, and benchmarking. Data was collected for three weeks from each of ~40 homes in Boston, MA and the Bay Area, CA, including 16 kHz whole home data, 3 sec circuit level data, and 1 min plug level data. Data available upon request, and will be publicly posted soon.

7. Algorithm development. Disaggregation algorithms using sparse coding methods were developed to advance the state of the art.

**Target Behaviors**

A list of target behaviors is useful for populating consumer facing recommendation systems with energy saving actions, and in guiding future work on the development of new energy saving actions. The two discrete projects included:

8. A database of 250 actions that directly reduce stationary residential energy use, and each action’s ratings on nine different attributes (e.g., energy savings, fiscal cost, frequency of the action, skill demand). Actions were implementation in Bidgely Inc.’s online recommendation system. This project also analyzed how the actions clustered based on the attributes so they could be more effectively “bundled” in behavior change programs.

9. A collection of energy saving actions from other cultures and throughout history as inspiration for modern day energy saving innovations; their potential energy savings across U.S. climate zones were quantified to provide an opportunity map for future design efforts. (work in progress)

**BEHAVIORAL INTERVENTIONS**
Media Interventions

10. Online game. This multiplayer game and supporting social media is suitable for use in experiments and deployment in utility smart meter trials. Power House incorporated real world energy data into some of the game play, leveraged social competition, and retrained habits through reinforcement. In a laboratory experiment and field study, participants respectively increased their short-term energy efficient behaviors (turning off devices) and used significantly less energy during the weeks they used the game. During Phase II, in 2015, over 15,000 unique users played the game.

11. Social norms. 800 self-selected residences received feedback about their energy consumption, energy saving tips, and normative framing in emails sent every other week. It was found that when energy-saving tips were organized thematically (e.g., all heat saving tips, rather than a random mix of tips), the collective-action frame (“We’re doing it together!”) led to significantly greater reductions than descriptive norms (“Residents here have reduced their energy use by x% this year.”) or the thematic recommendations alone.

12. Immersive reality. This work created an immersive virtual shower world and measured its impact on energy related hot water consumption behavior, with results suggesting that vivid visualizations of energy consumption (e.g., amount of coal instead of KWh) are more important than the personalization afforded by avatars.

13. Facebook applications. This project developed three Facebook applications to match the range of motivations exhibited by individuals: Power Tower is social in that it allows one to collaborate with others in a Tetras-like puzzle where pieces are granted based on multiple participants’ energy savings; Kidogo is affective in that it allows one to compute their real world energy savings and then microfinance individuals in developing countries based on these savings; and Powerbar is cognitive in that it primary displays energy feedback data. In Phase II, as an outgrowth of this work, an online course was developed and is offered through the Stanford Professional Development Program.

Policy Interventions

14. Appliance calculator. This application has been used by over 60,000 people via Google Ads. By testing changes in the interface we have found, contrary to expectations, that projecting out cost savings over time does not appear to prompt more energy efficient refrigerator browsing, whereas simply changing the default sort order to put the most efficient appliances on top reduced the average kWh consumption of items selected by 10-20% – this suggests that simply implementing the most effective behavior change
techniques may be a more effective strategy than a traditional route of analyzing the underlying cause of a problem then trying to address it.

15. Raffle incentive. The Insinc sweepstakes or raffle-like incentive program recruited 21,000 users in six months, with 7.5% of all Insinc trips shifting off peak to reduce congestion and associated fuel waste, 11% shifting by those previously making regular peak-hour trips, and significant shifting to the use of public transit. A computational platform was also developed (separate from project #2) to support this and similar programs.

Community Intervention

16. Community program. Community-based programs can be much more effective than traditional marketing communications when strategically tapping into existing social networks and providing close support from peers (e.g., 10 vs. 85% in the Hood River Project). Here we developed a five lesson Girl Scout program. In a pilot trial, Girl Scouts in troops randomly assigned to the residential energy program significantly increased their residential energy-saving behaviors by 49% following the intervention, and 27% persisting after more than seven months of follow-up, compared to controls. Their participation also facilitated behavior change in their parents. Regarding Phase II scaling activities, by end of the ARPA-e funding, 216 Girl Scout troop leaders in 30 states had registered for the GLEE online training course, and 72 leaders were working in or had worked through the course – potentially reaching nearly 900 girls and their families - , and the website had 14,000 pageviews.

Integrative Project

17. The integrative program combined learnings from the other projects and additional high-impact approaches, focusing on scaling feasibility, depth of savings, advanced EM&V capabilities, and marketplace sustainability. The core element of the project is the HowPower browser application that utilizes innovative approaches such as the use of a dynamic concierge; a focus on tips-to-action; an entertaining and frictionless experience; and triaging to determine which actions and products are right for the user. A five-month trial with a rudimentary version of the browser application garnered 15,000 unique users from Google Adwords. Additional goals are to evaluate the energy savings of the program, and then disseminate it to utilities and third party vendors.

DATA EVALUATION AND MODELING

18. Google Powermeter evaluation. Over 1000 individuals participated in a randomized controlled trial and showed an average of 6% energy savings in the first month or two of using the interface, which resembled many interfaces used on utility or other energy
feedback websites. The study thus provided a benchmark for the comparison of other interventions, as well as a rigorous example of experimental and analysis methods for such work. (A note regarding persistence - it may be substantially improved with periodic reminders, as reported by Hunt Alcott (NYU) regarding the Opower program.)

19. Twitter explorer. This project developed Twitter Explorer to mine and analyze data from ongoing Twitter and other social media conversations about energy. For a one year period, Tweets were collected if they contained ~150 energy related linguistic terms. Using content, network, and semantic analyses of Tweets and hashtags, we assessed engagement, identified influencers, and identified word-of-mouth communities. Further work can enable an understanding of how to create, grow, and sustain word-of-mouth acceleration of energy behavior change.

20. Diffusion modeling. This project developed a simulation methodology and tool that allows one to model the energy savings of behavioral interventions according to parameters such as time, behavioral technique used, and social network distance and type. This tool could serve as a foundation for developing similar but more sophisticated tools enhanced with additional parameters and empirical data that could eventually lessen time and cost of developing interventions through predictive modeling, or, used in another way, to choose in which real-world settings to deploy different manipulations.

5.1.2 Additional Product Outputs

<table>
<thead>
<tr>
<th>Invention Disclosure Title (Date)</th>
<th>Inventors</th>
<th>Patent Title</th>
<th>Patent Number</th>
<th>Patent Filing Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smart Meter Data-driven Targeting of Energy Programs (4/15/2013)</td>
<td>Ram Rajagopal and Adrian Albert</td>
<td>Customer energy consumption segmentation using time-series data</td>
<td>14/567,615</td>
<td>12/11/2014</td>
</tr>
<tr>
<td>Smart Meter Data-driven Targeting of Energy Programs (4/15/2013)</td>
<td>Ram Rajagopal and Adrian Albert</td>
<td>Data-driven targeting of energy programs using time-series data</td>
<td>14/567,648</td>
<td>12/11/2014</td>
</tr>
<tr>
<td>Thermal profiling of residential energy use (12/30/2013)</td>
<td>Ram Rajagopal, June Flora, and Jungsuk Kwac</td>
<td>Method and system for profiling and scheduling of thermal residential energy use for demand-side management programs</td>
<td>14/716,763</td>
<td>5/19/2015</td>
</tr>
</tbody>
</table>
Awards, Media, and Follow-On Funding

Awards
- HowPower, the Integrative Project, won a regional Energy Efficiency award from the CleanTech Open, the world’s largest cleantech accelerator.

Follow on funding
- GLEE received 2 years of additional Stanford funding to build a sustainability program for GLEE and expand online offerings to a GLEE water program.
- Pacific Gas and Electric paid Rajagopal’s associates to set up Visdom at their headquarters.
- Other projects extended their research in related directions.

Media
Some examples of media exposure included:
1. Several projects received significant exposure to relevant specialized communities rather than mainstream media. For example:
   - Learnings from the disaggregation review paper were incorporated in tech media, helped spawn an annual workshop on the topic, and resulted in tens of thousands of downloads of the paper with continued follow up contact by researchers and industry into 2017.
   - Visergy helped spawn the current Bits and Watts Initiative at Stanford, which has received significant follow on funding, publicity, and workshops for other smart grid algorithmic projects.
   - The open extensible communication network RPL helped establish the first Internet standard for HANs, which is being adopted by industrial consortia such as WirelessHART and ZigBee. Significant communication and convenings occurred within the community to make this happen.
   - The Insinc transportation program and related spin-off programs were publicized in multiple news sources in India, Singapore, and the U.S., which helped increase participation in the programs and also helped facilitate backing to spread the programs to additional sites.
2. In 2013 and 2014 the Stanford Report published stories on several of the Behavior Initiative projects. These stories are sometimes picked up by mainstream and specialty media, but were not tracked in this case.
3. GLEE did a media push in July 2016 following its publication in the journal Nature Energy. Some of the venues included:
   - https://psmag.com/mother-earths-secret-weapon-girl-scouts-679935fb146e#.v9i8o895g
5.1.3 Potential Impacts

Perhaps the most effective way to save energy is to simply not use it; this is in part because every unit of electricity not used (coined “negawatt”) avoids the consumption of three to four units of fuel at the power plant (Lovins 2010; Lovins and Browning 2000). Such reduced demand side use has a wide array of benefits: reduced GHG emissions; reduced environmental impact (from GHGs, pollution, etc.); reduced system capacity requirements (reducing the generation, transmission, and distribution investments required of utilities and IPPs to meet electricity demand); improved grid reliability (reduction in outages, etc.); increased energy security (reduction in vulnerability to long-term disruptions of energy supply through volatile fuel prices, energy imports, climate change impact on energy markets, and finite resource exhaustion). Behavioral programs such as ours also provide the benefits of increased consumer appeal of energy efficiency actions (making achieving energy savings easier, promoting their adoption and furthering the benefits they provide), and increased economic activity (for example, through an increase in retrofits, energy efficient appliance purchases), and potentially improved smart grid efficiency and benefits.

The majority of the Stanford Initiative’s work was in developing software, algorithms, programs, evaluation tools, and other supporting products, and testing these – precursors to scaling to large populations to achieve the widespread impacts described above. To date, there have been concrete energy saving impacts as described above, for example, from the Insinc, PowerMeter, and GLEE (as well as other suggestive savings from the appliance calculator, PowerHouse, and HowPower; and also impacts from other projects less directly tied to energy savings, like RPL and Visdom). A cautionary note to others pursuing work at the intersection of sensors, behavioral approaches, and energy is that it is difficult to get energy data at scale – it is difficult and timeconsuming to get from utilities, and individuals are mostly uninterested in accessing their own data or sharing it.

To provide an estimate of the potential energy savings if work of this type is scaled, we include
the following discussion. For an effective program that targets residential electricity use and space heating use, we can start with an estimate of lower bound savings up to 12%, and upper bound savings of 35%, with a middle ground estimate of 23%. These figures should then be tempered with the finding that large scale programs tend to achieve lower average energy savings per household than smaller programs, and most of the programs to date have been smaller, thus these figures may be biased upwards. The lower bound of achievable energy reductions is derived from over 50 studies showing that feedback about electricity use causes a reduction in use ranging from approximately zero to 20%, with the majority of studies showing savings between 7-16% (Darby 2006; Fischer 2008; EPRI 2009). These findings are derived from studies that used feedback as the sole or primary behavior change technique and greater savings may be possible by applying additional techniques (e.g., goal setting, game playing, framing effects, incentive structures, social marketing, and/or competition). It is also worth noting that in these studies, more frequent or appliance-specific feedback results in deeper savings, and savings tend to persist over two years according to a literature review by the Electric Power Research Institute (EPRI 2009) though questions about persistence continue and persistence may occur in some situations such as with periodically repeated programs but not in others (Alcott and Rogers, 2013). The upper bound estimate of 35% is derived from case studies (Parker et al. 2006) and a business analysis (McKinsey&Company 2009); there are also communities of households that have achieved much deeper savings, such as the 90% Reductions Group (Ninety Percent Reductions Group). The middle ground estimate of around 23% is the energy savings potential from behavioral programs, estimated by two independent research groups, that took into account technical potential from common energy saving actions as well as likely population penetration of these actions (Gardner and Stern 2008; Laitner, Ehrhardt-Martinez, and McKinney 2009). As mentioned, these figures should be tempered with the finding that large scale programs tend to achieve lower average energy savings per household than smaller programs, and most of the programs to date have been smaller.

It may be useful to compare these figures to the current widely utilized status quo energy behavior change savings program in the United States - Opower’s home energy report - which achieved approximately 2% residential electricity bill savings according to analysis of their first 17 experiments which included 600,000 households across the United States (Alcott, 2011). Another useful comparison is that with the residential, and also total, energy consumption of the United States and of California. In the U.S., about 23% of energy is consumed by residential buildings, 19% by commercial buildings, and 16% by passenger cars and light trucks (Energy Information Administration 2008). In addition, the U.S. Department of Agriculture (USDA) estimates that food-related energy use accounted for about 16% of the U.S. energy budget in 2007 (Canning et al. 2010). Greenhouse gas (GHG) emissions follow a similar pattern (EPA 2006; Vandenbergh et al. 2008). In California, residential buildings use 33% and 37% of the electricity and natural gas respectively, and in commercial buildings these these figures are 37% and 16%. Furthermore, a 15% reduction of total U.S. energy consumption is more than the total yearly

\[ \text{Strategies for savings in one can often transfer to the other due to similar types of technologies and functions in both types of buildings (e.g. the use of air-conditioning).} \]
energy consumption in Brazil or the UK, or the quantity of fossil fuels that would be saved and
GHG emissions reduced in the U.S. by a 25-fold increase in wind plus solar power, or a
doubling of nuclear power, based on 2007 figures (Energy Information Administration 2008;
Sweeney 2007).

5.2 Conclusions

The Stanford Energy Behavior Initiative, which ran from early 2010 through the summer of
2013, achieved an impressive array of work spanning 20 projects overseen by thirteen
investigators across ten departments and five schools, in collaboration numerous students and
outside partner organizations. The central objective of the Initiative was to develop the
components that would support a system aimed at utilizing smart meter and other sensor data,
communication technologies, and behavioral approaches, to achieve significant energy savings.
The effort focused on the stationary residential sector, though work is applicable to and has
explored transportation, water, and commercial applications. The Initiative can be divided into
software and analytics, behavioral programs, and evaluation tools projects. The ESP software
provided the technical backbone for the behavioral programs; the analytics were developed to
perform personalization, targeting, and other functions to improve uptake and effectiveness of
the behavioral programs; and evaluation tools were developed to help assess the impact of the
programs. Sensor data was used in a myriad of ways: for baselining, segmenting, and targeting;
to improve automation; to disaggregate energy use and provide personalized
recommendations; to provide users with feedback, comparison with others, and points towards
competition, donations, and incentives; to evaluate the effectiveness of energy saving programs;
and in other ways. Many products were created from these projects, with some scaling efforts
pursued for wider impact, and that could be used by other groups in the future; also, numerous
publications were produced describing the methods, findings, and deliverables.
### GLOSSARY

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>CS</td>
<td>Computer Science</td>
</tr>
<tr>
<td>DR</td>
<td>Demand Response</td>
</tr>
<tr>
<td>DSM</td>
<td>Demand Side Management</td>
</tr>
<tr>
<td>EE</td>
<td>Energy Efficiency</td>
</tr>
<tr>
<td>EE</td>
<td>Electrical Engineering (Dept. Affiliation in Tables)</td>
</tr>
<tr>
<td>E-IPER</td>
<td>Emmett Interdisciplinary Program in Environment and Resources</td>
</tr>
<tr>
<td>ESP</td>
<td>Energy Services Platform</td>
</tr>
<tr>
<td>FS</td>
<td>Freeman Spogli Institute for International Studies at Stanford</td>
</tr>
<tr>
<td>HAN</td>
<td>Home area network</td>
</tr>
<tr>
<td>HMD</td>
<td>Head mounted display</td>
</tr>
<tr>
<td>H-STAR</td>
<td>Human Sciences and Technologies Advanced Research Institute</td>
</tr>
<tr>
<td>IETF</td>
<td>Internet Engineering Task Force</td>
</tr>
<tr>
<td>ISB</td>
<td>Indian School of Business</td>
</tr>
<tr>
<td>IVET</td>
<td>Immersive virtual environment technology</td>
</tr>
<tr>
<td>MS&amp;E</td>
<td>Management Sciences and Engineering</td>
</tr>
<tr>
<td>PEEC</td>
<td>Precourt Energy Efficiency Center</td>
</tr>
<tr>
<td>REDD</td>
<td>Residential Disaggregation Dataset</td>
</tr>
<tr>
<td>RFC</td>
<td>Request for Comments</td>
</tr>
<tr>
<td>RPL (“ripple”)</td>
<td>Routing Protocol for Low-power and lossy</td>
</tr>
<tr>
<td>Smart Grid</td>
<td>Smart Grid is the thoughtful integration of intelligent technologies and innovative services that produce a more efficient, sustainable, economic, and secure electrical supply for California communities.</td>
</tr>
<tr>
<td>TCP/IP</td>
<td>Transmission Control Protocol/Internet Protocol</td>
</tr>
</tbody>
</table>
REFERENCES

Chapter 1


Chapter 2

2.1

References Cited

JeongGil Ko (Johns Hopkins University), Joakim Eriksson and Nicolas Tsiftes (SICS), Stephen Dawson-Haggerty (UC Berkeley), Jean-Philippe Vasseur and Mathilde Durvy (Cisco Systems), Andreas Terzis (Johns Hopkins University), Adam Dunkels (SICS), and David Culler (UC Berkeley). Industry: Beyond Interoperability - Pushing the Performance of Sensor Network IP Stacks

Jeonggil Ko, Stephen Dawson-Haggerty, Omprakash Gnawali, David Culler and Andreas Terzis. Evaluating the Performance of RPL and 6LoWPAN in TinyOS.

http://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=74B6210BEB1D1F106F9197754ECF4863?doi=10.1.1.221.2035&rep=rep1&type=pdf

References for Funded Work

http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6322264


### 2.2

References Cited


### 2.3.1

References for Funded Work


2.3.2

References for Funded Work


2.3.3.1

References for Funded Work


2.3.3.2

References for Funded Work


2.3.3.3

References Cited


References for Funded Work


2.4.1

References Cited


References for Funded Work


2.4.2

References Cited


Ninety Percent Reductions Group Website. http://groups.yahoo.com/group/90PercentReduction/ (last viewed 1/22/2013)


References for Funded Work

Chapter 3

3.1.1

References Cited


References for Funded Work


3.1.2

References Cited


References for Funded Work


3.1.3

References Cited


References for Funded Work


3.1.4
References Cited


References from Funded Work


3.2.1

References Cited


References from Funded Work


3.2.2

References Cited

D. Schrank and T. Lomax. The 2005 urban mobility report.


References for Funded Work


3.3.1

References Cited


References for Funded Work


3.4.1

References Cited


Wood. 2015.

**Chapter 4**

4.1

References Cited


Reference for Funded Work


4.2

References Cited


References for Funded Work


Chapter 5


Allcott, Hunt, and Todd Rogers. 2013. The Short-Run and Long-Run Effects of Behavioral


Ninety Percent Reductions Group Website. [http://groups.yahoo.com/group/90PercentReduction/](http://groups.yahoo.com/group/90PercentReduction/) (last viewed 1/22/2013)
