THE LOAD LEVELING APPROACH TO REMOVING
APPLIANCE FEATURES FROM HOME ELECTRICITY
USAGE PROFILES

A Thesis in
Computer Science and Engineering
by
Stephen Elliot McLaughlin

© 2011 Stephen Elliot McLaughlin

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Master of Science

May 2011
The thesis of Stephen Elliot McLaughlin was reviewed and approved* by the following:

Patrick McDaniel  
Associate Professor of Computer Science and Engineering  
Thesis Adviser  

Trent Jaeger  
Associate Professor of Computer Science and Engineering  

Raj Acharya  
Professor of Computer Science and Engineering  
Head of the Department of Computer Science and Engineering  

*Signatures are on file in the Graduate School.
For the past twenty years, researchers have developed a class of algorithms that are capable of disaggregating a residential electric load into its set of contributing appliances. These algorithms, known as Non-Intrusive Load Monitors (NILMs), are now for the first time posing a threat to electricity consumer privacy. Fed by feature-rich electricity usage data from first generation smart electric meters, NILMs can reveal appliance usage patterns for every household in the developed world. In this work, we explore a new strategy for mitigating privacy loss due to NILM algorithms. Non-Intrusive Load Levelling (NILL) is introduced as a privacy preserving technique that places no trust in the utility to restrict the recording or processing of usage data. NILL relies on a stored supply of energy to both contribute to and draw from the net load exerted by a house. The resulting levelled load contains far fewer of the features that distinguish individual appliances. The primary result of this thesis is the simulation of a simple NILL algorithm against real-world load data collected from four homes in central Pennsylvania. The results show that over 90% of appliance features are removed from the load profile using a 50 Ampere-hour lead-acid battery for a stored energy supply.
# Table of Contents

List of Tables ................................................................. vi

List of Figures ................................................................. vii

Acknowledgments ............................................................... viii

Chapter 1. Introduction ..................................................... 1

Chapter 2. Background ...................................................... 5
  2.1 Load Profiles ........................................................... 5
  2.2 Non-Intrusive Load Monitoring ....................................... 7
  2.3 Energy Storage .......................................................... 7

Chapter 3. Non-Intrusive Load Leveling ................................. 9
  3.1 Run Time Control ....................................................... 10
  3.2 Determining Initial Target Loads .................................... 12

Chapter 4. Evaluation ........................................................ 14
  4.1 Source Data Collection ................................................ 14
  4.2 Full System Simulation ............................................... 17
    4.2.1 Simulation Procedure .......................................... 20
  4.3 Simulation Results .................................................... 21
  4.4 Countering NILM ....................................................... 22
  4.5 Residual Features ..................................................... 26
  4.6 Entropy Analysis ....................................................... 29
## List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Commodity smart meters.</td>
<td>5</td>
</tr>
<tr>
<td>4.1</td>
<td>Experimental data sources.</td>
<td>16</td>
</tr>
<tr>
<td>4.2</td>
<td>NILL feature reduction.</td>
<td>23</td>
</tr>
<tr>
<td>4.3</td>
<td>Feature reduction during NILL-effective ($RFM &lt; 0.1$) state in simulated environments.</td>
<td>25</td>
</tr>
<tr>
<td>4.4</td>
<td>Simulated time in NILL-effective state.</td>
<td>26</td>
</tr>
<tr>
<td>4.5</td>
<td>Residual Features.</td>
<td>26</td>
</tr>
<tr>
<td>4.6</td>
<td>Feature Exposure in the Low Recovery State.</td>
<td>28</td>
</tr>
<tr>
<td>4.7</td>
<td>Classification of residual features.</td>
<td>29</td>
</tr>
<tr>
<td>4.8</td>
<td>Empirical entropy.</td>
<td>30</td>
</tr>
<tr>
<td>5.1</td>
<td>NILL Cost Reduction per Month in Simulated Environments.</td>
<td>34</td>
</tr>
</tbody>
</table>
# List of Figures

1.1 Idealized non-invasive load leveling (NILL). ........................................ 3

2.1 Three-day load profile for large home taken in September 2010. ................. 6

3.1 Summary of the house and battery model used for NILL. .......................... 9

3.2 The transitions between the steady state target load and the high and low recover loads. ................................................................. 11

4.1 The TED 5000 MTU. ................................................................. 15

4.2 The TED 5000 gateway. ..................................................................... 15

4.3 Residence monthly, daily, and hourly profiles. ....................................... 17

4.4 The NILL Simulator. ......................................................................... 18

4.5 Example NILL and non-NILL load profiles. ......................................... 19

4.6 Relative feature mass over two day simulation. ..................................... 24

4.7 CDF of RFM as percentage of simulated time. ...................................... 27
Acknowledgments

For anyone who may find this work of some benefit, I should now make it clear that I am little more than a messenger. Thanks goes first to the unknown person who donated a pair of rollerblades to the Altoona Rescue Mission Thrift Store, without which I likely would never have had the Saturday afternoon roller rink conversation that led to my interest in pursuing education outside the realm of draftsmanship. Subsequently, thanks goes to Eric Sakmar who made the suggestion that I pursue advanced placement courses one day at the skating rink, to the Altoona Area High School students who filled up all other advanced placement courses except for computer science, and to Eugene Melcher, my computer science teacher of two years who was the only teacher to ever suggest that I pursue his course’s subject material as a career. Subsequent thanks goes to my parents for support throughout my undergraduate education (and life!), without which, I certainly would not have gotten far. Finally, I owe a debt of gratitude to my adviser Dr. Patrick McDaniel for opening numerous opportunities and to the members of the Systems and Internet Infrastructure Security lab for their guidance during my introduction to graduate education. It has been most enlightening.
Chapter 1

Introduction

Smart meters are being aggressively deployed in homes and businesses as part of a move to global smart grids [23]. This digitization of grid systems offers substantial benefits for society; increased efficiencies and information availability can enable cheaper and greener energy generation, less loss in energy storage and transmission, better fault isolation and recovery, and support for alternative energy sources, e.g., consumer generated wind and solar energy.

The move to digital grid control systems also introduces concerns about security [19, 24, 27]. The smart grid is a complex system of sensors, networks, and computing resources. Attacks against the smart-grid networks and computing elements can range from fraud, to denial of service, to privacy loss [27]. While some regulatory agencies have begun to explore security concerns, no comprehensive system has emerged to address these threats.

One area of particular concern is the loss of consumer privacy. Replacements for the antiquated in-home electromechanical meters, smart meters are embedded systems that use power and voltage sensors to collect and report load profiles. Load profiles are histories of energy usage collected at a configured granularity, e.g., seconds or minutes. While instrumental to managing energy use at the local and regional levels, such profiles are also sufficient to determine occupant behavior in residential settings [22, 21, 12]. Depicted here, this behavioral inference is made possible by a class of algorithms known as Non-Intrusive Load Monitoring (NILM):
To simplify, NILM algorithms decompose load profiles into composite appliance profiles based on known or learned signatures. For example, maps of discrete changes in energy use can be mapped directly to ON/OFF events associated with identifiable appliances. The profile represents a highly detailed map of appliance use and indirectly a surprisingly accurate model of human activity [22].

The concerns surrounding potential invasions of privacy via energy profiles appear to be more than hypothetical [16, 7]. Reuse of data by direct marketers, criminals, or law enforcement without prior approval or notification is often in conflict with privacy regulations, but may be occurring anyway [22]. Undercutting existing regulatory structures is a maze of often conflicting laws and court decisions relating to consumer privacy. For example, the 1939 Supreme Court United State v. Miller decision indicates that there is no reasonable expectation of privacy for information shared with third parties. State and regional agencies have built legal and regulatory structures to buttress privacy in the face of such decisions, but consumer rights remain, at best, murky.

The potential exposure of living conditions, occupancy, family routines, through energy profiles warrants vigilance [16]. This prompts the goal of this work: we aim to protect consumer privacy in the face of profile-exposing smart meters while acknowledging two practical constraints of the current grid environment:

- **Energy usage must always be accurately reported.** Any modification of the reported usage would undermine grid management, and introduce inaccuracies in billing and grid controls.

- **The privacy solution must not require any modification to the meters, appliances, grid control systems, or provider operation.** As they are not under control of the consumer, the meters and infrastructure must be assumed to be untrusted. Moreover, with millions of smart meters already installed and many more in deployment, any solution requiring new grid systems will not be logistically and economically viable [27].

\footnote{\$4.3 billion dollars has allocated by the U.S government for the smart grids, [28], with similar programs in progress in Asia and the EU.}
We address consumer privacy through *Non-Intrusive Load Leveling* (NILL), a novel technique to mitigate the invasion of consumer privacy via load profiling. Illustrated in Figure 1.1, NILL is conceptually simple; each consumer places a battery and control system between the smart meter and the circuit breaker of their residence. The load observed by the meter is “smoothed” by offsetting spikes and dips in usage by charging or discharging the battery. Hence, NILL removes the information content that allows for the construction of an appliance profile from a load profile. Here we make no assumptions about how an adversary may infer usage, and thus must strive to generate a constant load on the grid (thus providing zero information content). However, because of the physical limits and structure of electrical systems, this is a more challenging task than one might initially surmise. Note that NILL is currently not designed to mask longer term energy usage such as day/night diurnal energy patterns, but only the instantaneous energy transitions that expose minute-to-minute human behavior exploited by NILM algorithms. However, we do explore the challenges and countermeasures posed by future (currently undiscovered) NILM algorithms that use more sophisticated data mining techniques in Chapter 5.

NILL is an algorithm and control system that attempts to remove the fine-grained appliance signal represented by changes in the reported load. The control system directs the battery
charges and discharges (largely) to obscure energy usage. This is conceptually similar to queue delay perturbation countermeasures that prevent networking timing analysis [6]. By de-correlating both the timing and amplitude of ON/OFF events in the load profile, we remove the signal that NILM algorithms use to identify behavior.

Our adversarial model assumes that any party with access to smart meter data, whether directly (the utility) or indirectly (hackers or law enforcement), poses a privacy risk. All data generated by a smart meter is assumed to be archived and accessible to all adversary types arbitrarily far into the future. For example, the compromise of an electric utility’s database could disclose all measurements kept for a given residence. Adversaries do not have access to additional information beyond the load profile obtained from the electric meter, e.g. they cannot place or access additional sensors in a house.

The remainder of this paper identifies and evaluates a candidate NILL algorithm. A simulation of NILL is built on the widely-used SimPowerSystems [26] platform. We simulate four homes using energy profile data collected from real residential use. These experiments show that NILL exposed 1.1 and 5.9 identifiable appliance events per day. Such features reside amongst hundreds or thousands of battery-suppressed events, making reliable recovery of appliance profiles virtually impossible under current NILM. Further, we showed that such privacy can be achieved in the tested environments using only a moderately priced 50 amp-hour battery system—far smaller than the aggregate loads of protected residences.
Chapter 2

Background

Before delving into the details of a NILL system, we describe what information is made available by smart electric meters, namely load profiles, and how Non-Intrusive Load Monitors (NILMs) use this data to recovery appliance profiles for a home. While we review a number of NILM technologies from the literature, emphasis is placed on the specific type of NILM that is deemed most feasible for modern-era smart metering: steady state NILM. Because the battery used by NILL for energy storage is outside the scope of basic computer science knowledge, we have added a brief description of modern deep cycle batteries.

2.1 Load Profiles

<table>
<thead>
<tr>
<th>Epoch</th>
<th>Product(s)</th>
<th>Deployed</th>
</tr>
</thead>
<tbody>
<tr>
<td>monthly</td>
<td>electromechanical meters</td>
<td>N/A</td>
</tr>
<tr>
<td>(no data)</td>
<td>Sensus iCon [34]</td>
<td>7.6 million</td>
</tr>
<tr>
<td>15 min</td>
<td>Elster REX 2 [9]</td>
<td>4.0 million</td>
</tr>
<tr>
<td>5 min</td>
<td>Echelon NES echelon</td>
<td>4.4 million</td>
</tr>
<tr>
<td>1 min</td>
<td>Itron Centron [15]</td>
<td>14.4 million</td>
</tr>
<tr>
<td>1 s</td>
<td>TED 5000 [36]</td>
<td>(no data)</td>
</tr>
</tbody>
</table>

Conventional electromechanical watt-hour meters do not record load profiles, only net energy consumed over time. Thus they act as memoryless accumulators whose readouts are physically spinning dials. Energy use is measured by computing dial position changes since
Fig. 2.1. Three-day load profile for large home taken in September 2010. The drop-out box shows a higher resolution view of a television power ON and OFF.

the last reading (typically by a human meter reader once a month). In contrast, smart meters generate load profiles of energy use that are delivered to the provider at or near real time. A load profile is a record of the amount of energy consumed per epoch, typically stored by the meter in a sliding window of measurements. The level of detail in load profiles is useful for load forecasting and fraud detection [23]. Common low cost meters measure epochs at 15 minutes, but more sophisticated models can generate profiles at a second or lower granularity. Table 2.1 summarizes the capabilities of several market-leading meters with different capabilities.¹

The three-day load profile for a large 5-bedroom home is shown in Figure 2.1. A diurnal pattern is observable: peaks are felt in the morning, mid-day, and evening. The drop-out box shows an event occurring about 7pm on the 18th. A plasma television connected to a home theater system was turned on and then off about 5 seconds later using a master switch. The initial large spike represents the power-hungry television, followed by the theater receiver and speaker system powering on. The OFF event shows a symmetric decrease in power draw. NILM algorithms match these sister features (ON/OFF features of equal amplitude) against known appliance profiles to uncover in-residence behavior.

¹Note that TED identified in the table is not a smart meter, but a in-home device used to monitor energy usage (see Section 4.1).
2.2 Non-Intrusive Load Monitoring

NILM algorithms extract *appliance profiles* from load profiles. It is considered “non-intrusive” because it does this at the electric meter without instrumenting individual appliances. An appliance profile consists of the types of appliances and the times during which each is operational during the day. Load profiling techniques classify devices by the changes in steady state load caused by their being turned ON and OFF [18]. The approach is to decompose the load profile into a composite of individual appliances features, i.e., representative pairs of ON/OFF events. For example, periodic spikes in energy use are visible in homes with electric furnaces during cold weather. NILM algorithms can extract the expected furnace load from the load profile to expose other, possibly smaller, appliance loads. These techniques use appliance models and information learned about a residence over time to reconstruct behavior from a single aggregate signature. Such techniques have been shown to be highly accurate in practice [25, 21, 33, 22]. We discuss other classes of NILM that are not relevant to residential smart meters in Section 5.3.

2.3 Energy Storage

NILL requires one of a particular class of *deep-cycle* batteries. Deep cycle batteries are designed to be able to operate adequately during long cycles of charging and discharging without significantly reducing their lifetime. Such batteries are frequently used in recreational vehicles such as RVs and boats. There are several types of deep cycle batteries that support highly variable load profiles (at short timescales) present in home energy consumption. The Absorbed Glass Mat (AGM) battery (which has a lead-acid chemistry) has several properties that make it ideal for home use; they work well at extreme temperatures, have low internal resistance, can be charged at high voltages, and are designed to prevent leakage. To avoid sulfation (inability to hold charge due to crystallization of the lead sulfate), deep cycle batteries should not be allowed to discharge below 20% of their total capacity, and staying above 50% is optimal. When a battery
is to be charged beyond 90%, its charger should switch to a lower constant voltage than what
was used for previous charging [14].

For our evaluation of NILL, we model a 50 Ah\(^2\) lead-acid battery operating at a nominal
voltage of 120V. This is achievable by connecting typical 50 Ah sealed DC batteries, which
typically retail for approximately $100 [2], in series. One of the most common voltages for these
types of batteries is 12V, requiring 10 such batteries (approx. $1,000) to achieve the necessary
characteristics. We use a 60 ampere (A) maximum discharge current system as available in
modern home solar setups [1].

\(^2\)Ah stands for amp-hours, which is a measure of the battery charge capacity.
Chapter 3

Non-Intrusive Load Leveling

<table>
<thead>
<tr>
<th>Basic Definitions</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t$</td>
</tr>
<tr>
<td>$d(t)$</td>
</tr>
<tr>
<td>$u(t)$</td>
</tr>
<tr>
<td>$c(t)$</td>
</tr>
<tr>
<td>$b(t)$</td>
</tr>
<tr>
<td>$H$</td>
</tr>
<tr>
<td>$L$</td>
</tr>
<tr>
<td>$K_{SS}$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Relations</th>
</tr>
</thead>
<tbody>
<tr>
<td>$u(t) = d(t) + b(t)$</td>
</tr>
<tr>
<td>$c(t) = \int_{t_0}^{t} b(t) , dt + c(t_0) = K_{SS}[t - t_0] - \int_{t_0}^{t} d(t) , dt + c(t_0)$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>NILL Constraints</th>
</tr>
</thead>
<tbody>
<tr>
<td>$u(t) = K_{SS}$ for some constant $K_{SS}$</td>
</tr>
<tr>
<td>$L &lt; c(t) &lt; H$</td>
</tr>
</tbody>
</table>

Fig. 3.1. Summary of the house and battery model used for NILL.

The goal of a NILL system is to level the load profile to a constant *target load*, thus removing appliance features. To achieve this, NILL relies on the use of energy storage, e.g., a battery, to offset the power consumed by appliances. When an appliance turns ON, it will exert a load beyond the target load. Thus, NILL will discharge the battery to partially supply the load created by the appliance, maintaining the target load.\(^1\) Similarly, if an appliance enters the

\(^1\)The battery is only used to supply appliances in the house. It is never discharged back into the grid as is done in *net metering*. 
OFF state, the load profile will decrease below the target load. These opportunities are used to charge the battery while restoring the target load.

The NILL system presented here consists of two parts: a battery and a control system that regulates the battery’s charge and discharge based on the present load and battery state. The controller attempts to maintain a steady state target load $K_{SS}$, but will go into one of two special states $K_L$ or $K_H$ if the battery needs to recover from a low or high state of charge. This section describes the NILL runtime control system followed by the calculation of the initial parameters to bootstrap a NILL system.

### 3.1 Run Time Control

In a perfect NILL implementation, there would be no runtime control as the battery would have sufficient capacity for maintaining the target load. For any reasonably sized battery, there will be times when the state of charge is insufficient to maintain the target load under a heavy load. We call this a low recovery state because the battery’s SOC has become too low to maintain the target load. Similarly, in times of light load, the battery will draw from the utility to maintain the constant load. If however, the load remains light, eventually the battery will reach its maximum SOC. We call this a high recovery state.

We use the model shown in Figure 3.1 in describing the control system and bootstrapping phase in the next section. The model captures both the actual load profile of the house $d(t)$, as well as the load under the influence of NILL as perceived by the electric meter $u(t)$. The essence of NILL is described by the equation, $u(t) = d(t) + b(t)$, where $b$ is the battery’s rate of charge over time. If $b(t) > 0$, the battery is charging, otherwise $b(t) < 0$ and the battery is discharging. Finally, $c(t)$ is used to represent the battery’s state of charge (SOC). The NILL controller must maintain the target load and respond to low and high recovery states. The controller sits next

---

2 Some literature uses Depth of Discharge (DOD), the complementary quantity to SOC.
Fig. 3.2. The transitions between the steady state target load and the high and low recovery loads. In state $K_i$, the battery output is calculated using $K_i - d$.

to the battery at the service-panel or electric meter. A sensor placed on the same line as the meter is used to monitor $d$, and one on the battery to monitor $c$. Using these two parameters, the controller selects $b$ to maintain a constant $u$ within the battery’s operational constraints.

If the controller enters one of two recovery states, the target load will be altered to allow the battery to reach a safe state while still attempting to mask features present in the load profile. We denote the target load during steady state operation as $K_{SS}$, and the targets for the low and high recovery states as $K_L$ and $K_H$ respectively. Because entering a recovery state is a sign that the target load was either too high or too low to maintain, it is adjusted each time the controller transitions to a recovery state. Figure 3.2 illustrates the conditions under which the controller changes states and adjusts $K_{SS}$. While in state $K_i$, the battery is controlled using $b = K_i - d$.

In a low recovery state, a new target load $K_L$ is chosen to allow the battery to recharge while still hiding the majority of load events. Thus, $K_L$ is set equal to the battery’s maximum sustained charge amperage, masking all load events with amperages less than or equal to this maximum. To reduce the frequency of recovery states over time, the controller will adapt $K_{SS}$ each time a low or high recovery state occurs. This is done using the exponential weighted moving average of the instantaneous demand since the last recovery state at $t_r$: $K_{SS} \leftarrow \alpha \frac{D}{t - t_r} + (1 -$
Once the battery reaches 80% SOC, the system returns to steady state. The effects of low recovery states in our experimental results are shown in Section 4.4.

In a high recovery state, the battery is at its maximum SOC, and the load is below $K_{SS}$. Once in this state, the only choices are idling the battery, which allows all events to appear in the load profile, or discharging the battery. Because NILL’s goal is to cancel appliance level features, we choose the latter. The only question left is the choice of $K_H$. For this, the controller uses the most recent load samples to guess at a $K_H$ that will be just below the current load (by approximately 1 to 5 amperes). If this guess is not successful after the first few seconds, a more conservative guess is made. With $K_H < d$, the battery can discharge minimally while producing a flat area in the load profile. If the load increases by 5 or more amps, the system returns to steady state. An example of a low recovery state in our experimental results is shown in section 4.3.

3.2 Determining Initial Target Loads

A NILL system requires an initial value for $K_{SS}$ to bootstrap normal operation. A good target load is one that can be sustained with little variation over time. Target load selection is also useful for battery sizing, i.e., if there is no feasible target load for a given battery capacity, a larger battery should be used. For the remainder of this section, we refer to the initial $K_{SS}$ as simply $K$ to distinguish from the steady state target load during run time operation. Two constraints must be satisfied for the target load $K$ to be considered feasible. First, the utility observable profile should be leveled to $K$ at all times. Second, the battery charge and discharge required to achieve $u(t) = K$ must not cause the battery to exceed its safe capacity limits $L$ and $H$. Under these constraints, the equation for battery SOC can be rewritten in terms of $d$ and $K$

\[ \alpha)K_{SS}. \]

\[ \text{Note that this is not an EWMA over continuous time samples, but over discrete steady state periods.} \]
as shown in the RHS of the state of charge relation. This rewriting is used in the algorithm for finding a minimal target load.

**Algorithm 1** \textit{FindMinTargetLoad}

1: Given demand $d(t)$, start time $t_0$, end time $t$ and initial charge $c(t_0)$
2: $D \leftarrow \int_{t_0}^{t} d(t) \, dt$
3: $d_{max} \leftarrow \max_{[t_0,t]} d(t)$
4: Binary search $K \in [0, d_{max}]$
5: Check that $L \leq K[t - t_0] - D + c(t_0) \leq H$ over $[t_0, t]$
6: Output minimal satisfactory $K$

We use Algorithm 1 to find the minimal target load for a given battery capacity. The minimal feasible $K$ is chosen to put the least stress on the battery when coming on line. The input $d(t)$ is a sample of a load profile from the residence hosting the NILL installation. The output is the initial target $K$. In practice, we select $c(t_0) = 50\%$ SOC, $L = 20\%$ SOC, and $H = 90\%$ SOC to model the safe bounds on battery charge.
As one might imagine, the design space for NILL algorithms is quite large. The basic requirement for a NILL algorithm is that its resulting load profile does not contain sufficient information for a NILM to extract an appliance profile with high confidence. The parameters in this design space include the method of energy storage, the specification for the desired load profile, and the set of appliances to be suppressed. As this work represents the first inquiry into NILL systems, we choose to evaluate the basic algorithm presented in the previous chapter to test its efficacy at removing steady state appliance features from load profiles. We start using load profile data collected from four homes in Central Pennsylvania. These profiles are then replayed through a simulated home wiring that includes a NILL system complete with a lead-acid battery. The results show that 95%-98% of features usable by NILM algorithms are removed. We also examine the entropy of load profiles before and after NILL, and show a reduction in the upper bound of information available from load profiles under NILL.

4.1 Source Data Collection

The data used in the experiments was collected from devices installed in four homes in a cold weather region of the United States. A TED 5000 [36] measuring transmitting unit (MTU) device was installed in each monitored location and collected real power (kW) and voltage data (See Figures ?? and 4.2). TEDs passively measure power and voltage crossing the main circuit between the meter and the circuit panel. Energy readings are transmitted to a TED gateway over the house electrical circuits via power line communications (PLC). The gateway is connected via
Fig. 4.1. The TED 5000 MTU. The MTU measures power through the main circuit breaker in a house and broadcasts them over in-home power lines.

Fig. 4.2. The TED 5000 gateway. The gateway collects measurements via home power line communications from the MTU.
Table 4.1. Experimental data sources - energy usage traces collected from residences in spring of 2010

<table>
<thead>
<tr>
<th>Residence</th>
<th>( K_{SS} ) [kW]</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1</td>
<td>4.64 kW</td>
</tr>
<tr>
<td>H2</td>
<td>4.08 kW</td>
</tr>
<tr>
<td>A1</td>
<td>3.85 kW</td>
</tr>
<tr>
<td>T1</td>
<td>8.20 kW</td>
</tr>
</tbody>
</table>

a wired network to a personal computer, which can access readings via HTTP. The TEDs were polled at half-hour intervals to collect per-second load profiles.

The experiments in the following section use load profiles collected in the spring of 2010 in the four residences, as described in Table 4.1. We refer to the data sets as H1, H2, A2, and T1 throughout (homes 1 and 2, apartment 1 and townhouse 1, respectively). The data collection process introduced a small number of sample outages in which no data was collected. This was due to brief power cycles of the TED or lost communication between the TED collector and usage sensor. We repair these gaps by placing repeated samples of the constant average of the surrounding 100 seconds. The H1 data contained two such gaps (1 hour, 3 minutes), H2 contained one gap (19 minutes), A1 contained no outages, and T1 had 2 gaps (11 minutes, 13 minutes).

Given their relatively small size, these gaps have little influence on our experimental results.

Figure 4.3 illustrates energy use over different time scales for one residence, H1. The month profile highlights the relatively constant rate of use over time. Note that during the week of March 3rd, the usage drops off substantially. The occupants left for a spring break during this period, turned down the thermostat, and unplugged appliances throughout the house. The daily energy use exhibits similar diurnal patterns as described in Section 2.1. The periodic usage spikes observed in the day-scale data was the result of the home’s furnace turning ON and OFF blower motors to force heat in the gas-heated house. Finally, the hourly readings show the enormous
sensitivity of these devices to internal use; while spikes caused by starting dryers and heat pump blowers are very clear, small changes caused by appliances such as lamps are also visible.

### 4.2 Full System Simulation

The NILL system simulation answers the question of how effectively the battery and controller can remove appliance features from a metered load profile. To achieve a realistic battery model, we implemented the simulation in Simulink with the SimPowerSystems extension [26]. This was necessary, as using an oversimplified battery model, e.g. one with linear discharge characteristics, would lead to an inaccurate assessment of NILL’s capabilities.

Our simulation is shown in Figure 4.4. At the heart of the system is a circuit with three elements: a net load exerted by appliances, a voltage source for the utility, and a battery. The appliance load is supplied by demand data collected from the TEDs. It is injected into the system as a real valued signal, and converted to an electrical load by a variable resistor in the programmable load. (Like several other residential NILM studies [11, 4, 22], we focus on real power consumption.) The programmable load is supplied by the utility and potentially

Fig. 4.3. Residence monthly, daily, and hourly profiles.
Fig. 4.4. The NILL Simulator. Signals from the control algorithm (1) regulate the switching circuitry (2) that connects the battery (3) and load. The load profile data is inserted at (4) and exerted by a variable resistor (5). The sum of load and battery \( u(t) \) is measured by the meter (6), and rate and state of charge are measured at (7) and (8) respectively. A trace of the control signals (9) was used for debugging purposes.
the battery, which sit in parallel on the circuit. The part of the net load not supplied by
the battery will be supplied by the utility voltage source according to Kirchhoff’s circuit laws.
Battery charging is controlled by the switching connecting the battery to the rest of the circuit.
To avoid severe performance slowdowns in our simulation, we implement the battery switching
using variable resistance and voltage sources in place of simulating non-linear power electronics.
When the battery is discharging, a variable resistance is used to draw current from the battery,
and a current source supplies the load proportionally. This process is reversed during charging.

With regards to timing in the circuit, the load profile data is inserted at the rate of one
power sample per second, the same rate at which it was recorded. The metrology unit on the
utility meter decimates the signal to have 1/4 the resolution of the control unit. In reality, control
units with MHz clocks can have orders of magnitude higher resolution than solid state metering,
which can measure at rates as low as 10 kHz. Forcing sampling at these rates however would have
made the simulation run prohibitively slow. Finally, the memory block placed before the control
serves to add a one-step delay between the propagation of the sensor signals to the simulator,
and the propagation of the resulting control signals to the circuit.

![Graph](image.png)

Fig. 4.5. Top: The net load \(d\) consumed by T1 between Apr. 18 and Apr 30. Bottom: The
same load under the simulated NILL system with a 50 Ah battery \(u\).
The battery model used in the simulator was originally designed for SimPowerSystems in [37]. It maintains only a single state variable, the state of charge, as well as several other parameters to model voltage during the two phases of battery discharge: the exponential and nominal zones. The additional parameters, including internal resistance, are derived from the battery type, nominal voltage, and capacity. The model has been shown to accurately produce the original manufacturer voltage curves for lead acid batteries, as well as several other basic battery types.

The control unit receives two variables, the net demand and the battery’s state of charge. Based on the control model presented in section 3.1, these are used to calculate the charge state and rate of charge for the target $K_{SS}$. The control also checks the conditions for entering a recovery state and adjusts $K_{SS}$ using EWMA. The control system is implemented in just under 900 lines of Matlab code.

### 4.2.1 Simulation Procedure

We ran the simulator on each of the four load profile data sets. The runs for H2, A1, and T1 lasted for approximately 3 hours each, and H1 ran for 6 hours. The initial target load $K_{SS}$ for each simulation was calculated by running Algorithm 1 over one week of sample data occurring prior to the load profiles used in the simulation. The initial $K_{SS}$ for each residence is shown in Table 4.1. Though the initial values are not closely aligned with the longer term run-time steady states, they err on the side of caution due to the strict SOC constraint on line 5 of Algorithm 1. In each case, $K_{SS}$ was either close to the immediate load at the beginning of the trace, or led to a high recovery state with no leaked features. Along with the metered power under NILL $u(t)$, the SOC over time $c(t)$ was included in the output of each simulation.

The results report below were collected using a simulated 50 Ah battery with a 60 A maximum discharge. The $\alpha$ value for the EWMA used to adjust $K_{SS}$ was calibrated experimental to 0.5. Methods for choosing an optimal $\alpha$ and other parameters are covered in Section 5.2. The
lower bound on charge was chosen to be 20% SOC. The initial state of the battery was set to 80% SOC.

4.3 Simulation Results

A two-week example of the load profile witnessed by the electric meter ($u$) compared with the total load ($d$) is shown in Figure 4.5. From this example, several things are apparent. First, there are extended periods of hours to days during which no appliance features are visible in $u$. Second, the majority of features that do occur in $u$ have high amplitudes, and are among the largest present in $d$. Third, periods of light load in $d$ are accompanied by periods of very light or nearly zero load in $u$. And finally, the emergent shape of the load profile remains the same for both $d$ and $u$, and the area under the two curves is approximately the same. We now inspect each of these facets in detail.

Steady state loads comprise the majority of the NILL load profile. During these periods, the controller is able to maintain the target steady state load $K_{SS}$ without depleting or fully charging the battery. Thus, the better the choice of $K_{SS}$, the longer the steady state can be maintained. Following several high amplitude loads at the beginning of the trace, the steady state converges to durations of half a day or more by Apr. 22nd.

A number of large features appear in $u$ outside of the steady state. These features tend to appear at the peaks of comparatively steady high-amplitude loads. Examples of such events occur on Apr. 18th and 28th. These features, which are also some of the largest in $d$, either cause or occur during low recovery state. When the battery is recharging, the target load $K_L$ is set to allow the maximum rate of charge, causing the large yet steady load below the visible features.

After sustained periods of steady state operation or light loads in $d$, a sharp decrease will occur in $u$ as NILL enters a high recovery state. This happens when the battery has reached full charge while maintaining $K_{SS}$. The target load in this state $K_H$ is chosen to be just below
some of the recently sampled demand values, so as to allow the battery to discharge at a low rate. As can be seen, in the majority of cases, the choice of \( K_H \) is correct on the first try. On a few occasions, most notably after mid day on Apr. 18\(^{th} \), the initial guess is too high, and \( K_H \) is lowered by 0.6 kW to allow the battery to discharge. In several instances (such as just prior to Apr. 26\(^{th} \)) the NILL controller chose \( K_H \) to be slightly below zero, meaning that the battery is effectively discharging to the grid. This is only a concern if the meter can detect reverse power flows, in which case the NILL system can be configured to never set \( K_L \) below zero.

Having now covered how NILL reacts to loads, it should not be surprising that on the larger scale of days to weeks, the NILL graph looks similar to the original load. This is result of both the limited capacity batteries afforded by current technology, and typical weather and occupant behavioral patterns. While these larger trends reveal information about customers, most notably the likelihood of human presence in the house, they do not reveal the more fine-grained details such as how many occupants are at home or what their activities are. Thus, we now turn our analysis to quantifying NILL’s effectiveness in removing individual appliance features from load profiles.

### 4.4 Countering NILM

NILM algorithms exploit the information present in amplitude changes in load profiles. These “features” are indicative appliance ON/OFF events (and appliance-internal state changes that effect energy consumption, e.g., washing machine cycles [3]). Features exploited by NILM are represented by \((time, \text{amplitude})\). Here, amplitude is not the absolute metric of energy, but the relative change in energy use from the last sample. For example, a system exhibiting a 5-second load profile: \( \{t_0 : 0W, \ t_1 : 100W, \ t_2 : 200W, \ t_3 : 200W, \ t_4 : 200W, \ t_5 : 100W\} \) would yield features: \( \{(t_1,+100W), (t_2,+100W), (t_3,0W), (t_4,0W), (t_5,-100W)\} \). Samples with amplitude with no change are not features, and are ignored for analysis.
Table 4.2 shows the feature reduction for the simulated environments. The original TED profile data (Non-NILL) contains just over 1 million features on the 60 day profile of the large house H1, down to the 380 thousand features in the 29 day profile for the small townhouse T1. When the load profile is simulated in NILL-enabled residences (NILL), the number of features drops significantly (around 95% or more). This reduces the number of features from several hundred an hour to as few as 16 in the case of T1.

A key question asks if the residual features left after NILL are useful for the NILM analysis. Recall from Section 2 that NILM rely on pairs of symmetric ON/OFF “sister” features to infer appliance usage. We performed a sister matching algorithm that would match all events of greater than 10 W—about 1/6th the draw of a light bulb—as a pairing of positive charge increase and decrease (with a tolerance of 5 W). The matching algorithm linearly scans forward in the trace for the equal and opposite OFF event for each ON event. Table 4.2 shows that each simulated environment experienced 94% or greater reduction in identifiable sister features.
Raw feature reduction sketches the aggregate behavior of NILL, but only provides a crude estimate of privacy-preservation over time. A feature mass \((FM(w, D))\) is the number of non-zero features from over a discrete sample set \(D = (d_0, d_1, \ldots, d_w)\), e.g., the number of features over a given time window of size \(w\). In this work, the feature mass is the number of non-zero energy transitions occurring within a time-series interval sample \(D\) of length \(w\):

\[
FM(w, D) = \sum_{i=0}^{w} (d_i \neq 0)
\]

The relative feature mass (RFM) for an interval \(w\) is the ratio of the feature mass for two sample sets (in this case NILL over non-NILL) over \(w\), e.g.:

\[
RFM(w, D) = \frac{FM(w, D_{NILL})}{FM(w, D_{NON-NILL})}
\]

RFM acts as a measure of the relative feature density of the original and NILL data. Intuitively, as RFM approaches zero, there is very little signal relative to the original profile for a NILM algorithm to operate on. Figure 4.6 shows the RFM computed at ten minute intervals with a one hour sliding window, i.e., \(w = 1\) hour, over two days at A1 (11 days into the trace) and H1 (29 days into the trace).

The most visible characteristic in Figure 4.6 is the oscillation of RFM. Most of the trace shows the RFM at 1% or less, and at zero for some of the time (note that the Y axis is on a log scale). The reason for this is the nature of the battery system. When the battery is in the low recovery state, it is unable to suppress most appliance features, and they become visible in the
load profile. (We investigate the impact of features occurring during low recovery state in the next section). RFM increases in this off state, peaking in the case of H1 at almost 0.5 RFM for a few minutes. T1 has fewer and less pronounced increases in RFM. Note that the prolonged RFM of zero represents perfect privacy.

Table 4.3. Feature reduction during NILL-effective \((RFM < 0.1)\) state in simulated environments.

<table>
<thead>
<tr>
<th>Residence</th>
<th>Non-NILL</th>
<th>NILL</th>
<th>Change</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Total Features</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H1</td>
<td>879054</td>
<td>6808</td>
<td>-99.23%</td>
</tr>
<tr>
<td>H2</td>
<td>225088</td>
<td>1176</td>
<td>-99.48%</td>
</tr>
<tr>
<td>A1</td>
<td>354102</td>
<td>1508</td>
<td>-99.57%</td>
</tr>
<tr>
<td>T1</td>
<td>265400</td>
<td>1262</td>
<td>-99.52%</td>
</tr>
</tbody>
</table>

| **Features per hour** |       |      |        |
| H1        | 354    | 3    | -99.09%|
| H2        | 186    | 1    | -99.12%|
| A1        | 279    | 1    | -99.32%|
| T1        | 260    | 2    | -99.25%|

| **Sister feature pairs** |   |      |        |
| H1        | 291718 | 561  | -99.81%|
| H2        | 89558  | 118  | -99.87%|
| A1        | 148446 | 128  | -99.91%|
| T1        | 104012 | 161  | -99.99%|

To get a sense for the behavior of steady state NILL, consider the normal steady state. We arbitrarily define the system to be in a *NILL-effective* state when the RFM is less than 0.1—which represents 10% residual features resulting from NILL. Table 4.3 describes the features present during intervals when the system is NILL-effective. In all experiments the features, features per hour, and sister features are reduced by over 99%. Table 4.4 shows that this state is the norm—where between 65% to over 85% of the simulated time was in this state in the simulated environment. One might (incorrectly) attribute the success of in these instances to periods when the load is quiescent. Quite the contrary, the majority of total features in the original data are
Table 4.4. Simulated time in NILL-effective state.

<table>
<thead>
<tr>
<th>Residence</th>
<th>Non-NILL (sec)</th>
<th>NILL (sec)</th>
<th>% of sim.</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1</td>
<td>10527597</td>
<td>8928597</td>
<td>84.81%</td>
</tr>
<tr>
<td>H2</td>
<td>5183997</td>
<td>4345199</td>
<td>83.82%</td>
</tr>
<tr>
<td>A1</td>
<td>5356797</td>
<td>4570797</td>
<td>85.33%</td>
</tr>
<tr>
<td>T1</td>
<td>5009397</td>
<td>3255535</td>
<td>64.99%</td>
</tr>
</tbody>
</table>

Table 4.5. Residual Features.

<table>
<thead>
<tr>
<th>Residence</th>
<th>Features</th>
<th>Sisters</th>
<th>Residual Features (%)</th>
<th>Residual Sisters (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1</td>
<td>1047099</td>
<td>340986</td>
<td>35969 (3.4%)</td>
<td>5526 (1.6%)</td>
</tr>
<tr>
<td>H2</td>
<td>286960</td>
<td>110994</td>
<td>13230 (4.6%)</td>
<td>3112 (2.8%)</td>
</tr>
<tr>
<td>A1</td>
<td>430214</td>
<td>176540</td>
<td>15556 (3.6%)</td>
<td>3648 (2.0%)</td>
</tr>
<tr>
<td>T1</td>
<td>384847</td>
<td>147982</td>
<td>30861 (8.0%)</td>
<td>7640 (5.1%)</td>
</tr>
</tbody>
</table>

present during these periods. 83% of the original features for H1 are observed during a NILL effective state, 78.5% of H2, 82% of A1, and 90% of T1.

Figure 4.7 shows RFM as a CDF—the Y-axis indicates the simulated time in which the system encounters a RFM (X-axis) or less. For all traces the feature mass is zero for 65% to over 80% of the time. Trace time rises steadily until around an RFM of 50%, which is around 2% of the trace for all simulated environments.

4.5 Residual Features

A residual feature is any feature that appears in the original profile that also appears in the NILL trace, e.g., an energy transition that appears at (or about) the same time with similar amplitude (within few watts). Because they precisely reflect appliance activities, these are the only features that are of value to current NILM algorithms. Such features slip through the NILL control system and are observable by the adversary. Table 4.5 shows the residual features occurring in the simulated environments.
One reason for the presence of residual features is the low recovery state, hereafter called simply low state for brevity, during which the battery is recharging. Table 4.6 describes NILL behavior in low battery states. Here, we see only 24-56% of the residual features are present in the low state (which occurs less than 10% of the time). Thus, while low state charging explains some of the residual features, many are present in the also steady (non-charging) state.

One can partition residual features into three classes. The first class of features contains very small changes, e.g., less than 60 W, about the load of a typical light bulb. Such loads, the most common in load profiles, are driven by numerous causes including small lighting fixtures, appliance state changes, and small electronics requiring standby power. Because they reside in the band on the smaller end of the power spectrum, it would not be difficult for a NILL system to mimic their arrivals and departures using small variations in the battery charge and discharge rate. We thus do not consider them for the remainder of this analysis.

A second class of features contains large transitions caused by often short-lived heavy loads, e.g., kitchen microwave. The aggregate draw of the large loads combined with ongoing draws exceeds the target load ($K_L$) plus the maximum output of the battery (60 A for the simulated battery we tested). There is little NILL can do about these loads because it simply cannot safely supply the necessary current. The last class are legitimate loads that are not
Table 4.6. Feature Exposure in the Low Recovery State.

<table>
<thead>
<tr>
<th>Res.</th>
<th>Time in Low State (%)</th>
<th>Residual Features (%)</th>
<th>Residual Sisters (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1</td>
<td>891913 (8.5%)</td>
<td>18292 (50.8%)</td>
<td>3363 (60.8%)</td>
</tr>
<tr>
<td>H2</td>
<td>340685 (6.6%)</td>
<td>5248 (39.6%)</td>
<td>1414 (45.3%)</td>
</tr>
<tr>
<td>A1</td>
<td>435090 (8.1%)</td>
<td>8814 (56.6%)</td>
<td>2524 (69.1%)</td>
</tr>
<tr>
<td>T1</td>
<td>390279 (7.8%)</td>
<td>7456 (24.1%)</td>
<td>2172 (28.4%)</td>
</tr>
</tbody>
</table>

managed by NILL because of low state. Table 4.7 provides a breakdown of the residual features present in the simulated environment. Interestingly, 70-85% of the features and 92-98% of the sister features are of the small variety.

Regardless of the reason for their presence, the features usable by NILM to correctly uncover are the large and medium residual sister features. To summarize:

<table>
<thead>
<tr>
<th>Residence</th>
<th>Sister features</th>
<th>Per Day</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1</td>
<td>359 (0.11%)</td>
<td>5.9</td>
</tr>
<tr>
<td>H2</td>
<td>33 (0.03%)</td>
<td>1.1</td>
</tr>
<tr>
<td>A1</td>
<td>93 (0.05%)</td>
<td>3.1</td>
</tr>
<tr>
<td>T1</td>
<td>128 (0.09%)</td>
<td>4.4</td>
</tr>
</tbody>
</table>

Fewer than ten and as little as one true feature per day that are available to NILM algorithms in our simulated environment. This represents about 3 to 11 out of a thousand features exposed, a minuscule exposure of privacy.

One of the challenges a NILM algorithm might have in the presence of NILL is that it must identify which of the features out of the total trace are legitimate. For example, there are 473 true sister features hidden in 10,552 for H1. Unless there is a clear marker identifying a true feature in the timing/amplitude, then it would be virtually impossible to separate the real and synthetic transitions.
Table 4.7. Classification of residual features.

<table>
<thead>
<tr>
<th>Residence</th>
<th>Features</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Small (%)</td>
<td>Medium (%)</td>
<td>Large (%)</td>
</tr>
<tr>
<td>H1</td>
<td>25334 (70.4%)</td>
<td>6257 (17.4%)</td>
<td>4378 (12.1%)</td>
</tr>
<tr>
<td>H2</td>
<td>11319 (85.5%)</td>
<td>1193 (9.0%)</td>
<td>718 (5.4%)</td>
</tr>
<tr>
<td>A1</td>
<td>13139 (84.4%)</td>
<td>1249 (8.0%)</td>
<td>1168 (7.5%)</td>
</tr>
<tr>
<td>T1</td>
<td>25823 (83.7%)</td>
<td>4717 (15.3%)</td>
<td>321 (1.0%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Residence</th>
<th>Sister Features</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Small (%)</td>
<td>Medium (%)</td>
<td>Large (%)</td>
</tr>
<tr>
<td>H1</td>
<td>5167 (93.50%)</td>
<td>214 (3.87%)</td>
<td>145 (2.62%)</td>
</tr>
<tr>
<td>H2</td>
<td>3079 (98.94%)</td>
<td>21 (0.67%)</td>
<td>12 (0.39%)</td>
</tr>
<tr>
<td>A1</td>
<td>3555 (97.45%)</td>
<td>45 (1.23%)</td>
<td>48 (1.32%)</td>
</tr>
<tr>
<td>T1</td>
<td>7512 (98.32%)</td>
<td>118 (1.54%)</td>
<td>10 (0.13%)</td>
</tr>
</tbody>
</table>

Note that if someday a NILM algorithm were modified to recover privacy information with unmatched ON or OFF edges, NILL would still provide little information on which to base the algorithm. In the worst simulated case (H1), only 61,000 (of which 17% were residual) of the original 1047099 features were retained. Again, the limited number of features hidden in a huge body of battery-suppressed features would thwart existing approaches.

4.6 Entropy Analysis

In addition to comparing the number of features between the load seen by the battery and that seen by the utility, comparing the empirical entropy of two loads is also useful. As in the above discussion, distinguishing information about which appliance or type of appliance was just turned ON (or OFF) lies not in the $d$ and $u$ time series of aggregate load versus time but rather lies in the features (i.e., non-zero impulses) that. Define $\delta_d(t) = d(t) - d(t - 1)$ and $\delta_u(t) = u(t) - u(t - 1)$ as the time series of impulses generated by the appliances and seen by the utility, respectively. For each time series we define a probability mass function, $P_{\delta_d}$ and $P_{\delta_u}$, respectively, which takes a bin size $b$ as a parameter. For concreteness define the $i$th bin as being the range $[ib, (i + 1)b)$ in units of watts. $P_{\delta_d}(i|b)$ is just the fraction of the values of $\delta_d$ that are
Table 4.8. Empirical entropy.

<table>
<thead>
<tr>
<th>Residence</th>
<th>Non-NILL</th>
<th>NILL</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1</td>
<td>0.633267</td>
<td>0.054265</td>
<td>0.085690</td>
</tr>
<tr>
<td>H2</td>
<td>0.235150</td>
<td>0.029554</td>
<td>0.125681</td>
</tr>
<tr>
<td>A1</td>
<td>0.363826</td>
<td>0.038459</td>
<td>0.105701</td>
</tr>
<tr>
<td>T1</td>
<td>0.360635</td>
<td>0.019137</td>
<td>0.053067</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Non-NILL w/o zeroes</th>
<th>NILL w/o zeroes</th>
<th>Normalized Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1</td>
<td>5.663782</td>
<td>7.519812</td>
<td>0.078346</td>
</tr>
<tr>
<td>H2</td>
<td>4.516394</td>
<td>6.340144</td>
<td>0.101323</td>
</tr>
<tr>
<td>A1</td>
<td>4.468669</td>
<td>6.526919</td>
<td>0.039289</td>
</tr>
<tr>
<td>T1</td>
<td>5.114588</td>
<td>6.460732</td>
<td>0.084519</td>
</tr>
</tbody>
</table>

In bin $i$, $P_{\delta d}$ is defined similarly. Given an empirical probability mass function $P(i)$ we calculate the empirical entropy using base 2 in the standard way:

$$H(P) = -\frac{1}{|I|} \sum_{i \in I} P(i) \log_2 P(i),$$

where $I$ is the support of $P$, i.e., $I = \{i | P(i) > 0\}$.

Here, $H(P_{\delta d})$ can be interpreted as an upper bound on the number of bits of information a NILM algorithm could extract on average for each impulse sample in $\delta d$ (if it had access to $\delta d$, as it would in the case that our NILL hardware and software were not deployed). Likewise, $H(P_{\delta u})$ can be interpreted as upper bound on the number of bits of information a NILM algorithm could extract on average for each impulse sample in $\delta u$, i.e., after the signal has been smoothed by our NILL algorithm. Table 4.8 reports the empirical entropies for our four data sets, computed using a bin size of 1 W. The bin size was chosen this small to allow for the potential of higher entropy due to higher precision but not any smaller in order to mitigate the introduction of undue noise.

As noted above, the empirical entropy is an upper bound on the information extractable by a NILM algorithm from a time series. Strictly speaking, showing a gap between an upper bound on the information extractable from $d$ and an upper bound on the information extractable...
from $u$ does not prove that the NIL algorithm decreases the information content of $d$. That is, the actual information content extractable from $d$ may be below the upper bound on the information extractable from $u$. However, this would require that the information content of the two signals have significantly different characteristics. In the event that the sequences have similar characteristics, the ratio between the upper bound and the actual information extractable should be similar for each time series.

Note that if a time series is highly correlated, the empirical entropy will overestimate the information in the series. (For example, a series of 100 1’s followed by 100 2’s will have empirical entropy of 1 bit per time step although the series clearly does not contain 200 bits of information.) There are clearly examples of correlations in our time series. As mentioned above, when a home theatre system was turned on in H1 and H2, several features occurred in succession over a short period of time. In addition to such intra-appliance correlations, some home loads have quasi-periodic behavior. Examples of such loads include refrigerator compressors and home furnace systems.

To understand the extent of these time correlations we computed the autocorrelation function (Pearson variant) of $\delta d$ and $\delta u$. The autocorrelation on input $k$ of an $n$ element times series $s$ compares $s(i)$ with $s(i+k)$ for all $1 \leq i \leq n - k$. The autocorrelation is always at most 1 and at least -1, with values near 1 implying a high degree of correlation, values near -1 implying a high degree of anti-correlation, and values near 0 implying a high degree of independence. The autocorrelation of $\delta d$ had absolute value less than 0.02 for all values of the lag $k$. Thus, in spite of the fact that there are some correlated impulses in the data set, the vast majority of impulses appear to be uncorrelated. The non-zero impulses of $\delta u$ essentially constitute a subsequence of $\delta d$. (That is, if at time $t$, $\delta u(t)$ is non-zero then at the same time $\delta d(t)$ is non-zero as well. If fact, the size of the impulse is often the same in this case.) Since a subsequence of a sequence of independent events is itself independent, it follows that the impulses of $\delta u$ are largely uncorrelated.
as well. Indeed, the autocorrelation of the $\delta u$ time series was similarly small. This argues against the introduction of a systematic bias in empirical entropy of one series versus the other due to differences in time correlations between the two series.

In Table 4.8, the entropy is computed in two ways: one in which zero values of the respective time series are included when calculating a probability mass function for the time series, and one in which the zero values are excluded. Including the zero values has intuitive appeal as the fact that there are many more zeroes in the NILL time series does reflect that fact that the NILL time series conveys less information to a NILM algorithm. However, to ensure that the number of zeroes in the NILL time series were not biasing the results unduly, we also calculated the entropy of the various time series without the zeroes included. As the entropy represents the average amount of information per sample, and as there are more samples in the Non-NILL time series, we have included a ratio which normalizes for this fact. The numerator of the ratio represents the total information available in the NILL time series (i.e., the entropy times the number of terms in the NILL series without zeroes) and the denominator represents the total information available in the Non-NILL time series (i.e., the entropy times the number of terms in the Non-NILL series without zeroes).

Table 4.8 reports the empirical entropies for our four data sets. As can be seen, when including the zeroes of the time series, the ratio of the NILL to Non-NILL entropies varies from 0.53 on the low side for data set T1 to 0.126 on the high side for data set H2. When the zeroes are included, the normalized ratio varies from 0.039 on the low side for data set A1 to 0.101 on the high side for data set H2.
Chapter 5

Discussion

A number of issues demand discussion. First, beside the up-front costs of batteries, does NILL incur cost overhead on monthly electric bills? While this depends heavily on the type of electricity tariff in place, we show that for up-and-coming Time of Use tariffs, NILL does not create monthly overhead on electric bills. The second issue addressed is that of choosing the NILL parameters described in Chapter 3. Though the parameters chosen in our evaluation provide sufficiently good results, principles should be established for selecting these parameters in the general case. Finally, we discuss two types of NILM algorithms that may pose a threat to NILL. Adversarial NILMs attempt not only to deduce the appliances in the load profile, but also the behavior of the NILL algorithm itself in order to remove its affect on the load profile. The second type, load disaggregation NILMs attempt to approximate the set of appliances contributing to the steady state load at any given time. Assuming that the instantaneous value for the battery’s rate of charge is known, this may reveal significant information about appliance usage.

5.1 Energy Efficiency and Consumer Cost

One of the greatest challenges driving cost in the energy industry is anticipating and controlling load. The highest peak loads are observed daily between mid-morning and early evening (when most consumers are awake and active), and in the summer (when consumers likely run air conditioners). The cost of generating energy increases with demand; in the US Mid-Atlantic region, the price of wholesale generation increases from about 6¢/kWh to over $1/kWh. Thus, providers can be much more cost efficient and energy conserving when a constant demand is observed. NILL evens demand and therefore aids the providers in damping demand curves.
Table 5.1. NILL Cost Reduction per Month in Simulated Environments.

<table>
<thead>
<tr>
<th>Residence</th>
<th>Total kWh</th>
<th>Non-NILL</th>
<th>NIL</th>
<th>O&amp;R</th>
<th>Out</th>
<th>PG&amp;E</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1</td>
<td>9637.30</td>
<td>$427.10</td>
<td>$555.28</td>
<td>$1033.63</td>
<td>$418.15</td>
<td>$11.11 (2.09%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$8.94 (2.09%)</td>
<td></td>
<td></td>
<td></td>
<td>$1034.96</td>
</tr>
<tr>
<td>H2</td>
<td>1566.80</td>
<td>$48.14</td>
<td>$88.39</td>
<td>$144.39</td>
<td>$45.66</td>
<td>$84.61</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$2.49 (5.16%)</td>
<td></td>
<td></td>
<td></td>
<td>$138.22</td>
</tr>
<tr>
<td>A1</td>
<td>2052.95</td>
<td>$101.14</td>
<td>$130.19</td>
<td>$218.91</td>
<td>$97.72</td>
<td>$125.22</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$3.41 (3.37%)</td>
<td></td>
<td></td>
<td></td>
<td>$208.69</td>
</tr>
<tr>
<td>T1</td>
<td>2012.08</td>
<td>$105.56</td>
<td>$125.11</td>
<td>$263.65</td>
<td>$102.89</td>
<td>$121.39</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$2.67 (2.53%)</td>
<td></td>
<td></td>
<td></td>
<td>$256.73</td>
</tr>
</tbody>
</table>

A new cost structure being introduced into consumer energy markets is based on time of use (TOU) pricing [17]. TOU creates a cost schedule that is tied to the demand curve observed by the provider. Energy costs rise with expected demand. For example, Orange and Rockland power in New York charges 1.280¢/kWh for off-peak use (9:00pm-10:00am year round), and 7.17¢/kWh for peak use (10:00am-9:00pm) October through May. During June through September, peak charges increase to 7.17¢/kWh for shoulder peak use (10:00am-noon and 7:00pm-9:00pm) and 19.899¢/kWh for high peak use (noon-7:00pm) [30]. TOU is touted as a way to control costs by creating incentives for customers to use energy during non-peak hours.

NILL may positively or negatively impact consumer cost under a TOU schedule. Consider a simplified model of consumer cost in a NILL and non-NILL household. Assume that the total kWh usage for the home is \( U \), the percentage of use at peak is \( U_p \), and the provider costs for peak and off-peak are \( C_p \) and \( C_o \) per kWh, respectively. The total monthly bill for the home is:

\[
T = (U \cdot U_p) \cdot C_p + (U \cdot (1 - U_p)) \cdot C_o
\]

According to the U.S. Energy Information Administration, the average home uses 920 kWh per month [38]. Using this figure, we can bound on the cost or savings observable by an average residence under Orange and Rockland TOU cost schedule. In the extreme cases, the costs of using
NILL can increase by $24.84 (where the home uses no energy during peak hours) or decrease by as much as $29.35 (where the home uses no energy during off-peak hours). Based on usage statistics, the load of the average home will use slightly more energy during peak hours ($U_p > 50\%$), with more pronounced peak loads in the summer months (due to air conditioning).

Table 5.1 shows the energy costs in the simulated environments under three time-of-use pricing schemes including Orange and Rockland, Ontario Power Generation [10] (Ontario, Canada), and Pacific Gas and Electric [31] (California). The simulated environments observed cost savings ranging from 1.8% to 5.2%. This reflects the smoothing of load by the battery, where some amount of the peak load is shifted into the off-peak hours. Note that the price of energy is significantly higher in California than the other simulated environments, and thus the costs are about double that of the upper-mid Atlantic and Ontario regions. The costs savings of $2 to $10 may partially offset the costs of the NILL system.

Also contributing to the long-term cost of a NILL scheme is the cost of battery maintenance and replacement as a function of battery lifetime. A number of factors affect battery lifetime including operating temperature and discharge amperage. Here, we concentrate on the dominating contributor to battery wear, average depth of discharge (DOD). Under a NILL scheme that exercises a lead-acid battery to between 50 to 80% DOD, the battery can be expected to last between 500 to 1,000 cycles [8]. The actual lifetime of the battery in years is then a function of how frequently it is discharged to 80% DOD. Under the workloads examined in this paper, battery lifetime is expected to be between one and two years, which could be improved upon if the battery is only used when occupants are at home. This would not be a concern as most day-scale diurnal patterns are already visible under NILL. Additional measures to reduce the cycle frequency are the primary goal of future work for this project.
5.2 Optimizing NILL Parameters

The NILL system simulated in this paper relies on a number of important parameters for its operation. While some guidance has been given for parameter choice, such as the calculation of $K_{SS}$, others were chosen using heuristics and expert knowledge. In a full NILL implementation however, these parameters should be chosen to optimize the performance and lifetime of the system given the characteristics of the battery. Parameters of particular importance are the high and low recovery state target loads $K_H$ and $K_L$, the EWMA weight parameter $\alpha$, the bounds on safe battery capacity $L$ and $H$, and the capacity itself.

The $K_L$ value chosen in a low recovery state allows the battery to recharge quickly to just below a maximum SOC (80% in our experiments). For many lead acid batteries, optimal lifetime and performance can be achieved by recharging to between 50% and 80% SOC at a high rate, and then throttling back to charge to between 90% to 100% [14]. Thus, depending on battery characteristics, a lower secondary $K_L$ may be chosen to finish charging the battery at a lower rate. In our simulations, $K_H$ was chosen using the previous few data points from the net load. While this results in a low discharge from the battery, it is not guaranteed to be minimal. As lead-acid batteries tend to perform best at a high SOC, $K_H$ should be chosen to maximize efficient performance.

The EWMA was used to adapt $K_{SS}$ over time as the net load caused recovery states. The better the choice of $K_{SS}$, the longer a steady state may be maintained. While we achieved good results with $\alpha = 0.5$, this heuristic is not guaranteed to work under arbitrary conditions. Instead, a full implementation should consider techniques such as minimizing Mean Squared Error (MSE).

Arguably the most important parameter in a NILL system is the battery capacity. The capacity determines the majority of the battery price of the system as cost scales with Ah. But does larger capacity imply better privacy? In some of our preliminary experiments with batteries between 100 Ah and 200 Ah, we found that this is not necessarily the case. As was shown in
Section 4.4, the majority of relative feature mass is located at times when the system is recharging in a low recovery state. Thus, a larger capacity battery will not necessarily reduce the number of residual features if it spends the same average amount of time in recovery states as a smaller battery. This was the case for the NILL scheme presented in this paper, because while larger batteries enter fewer recovery states, they require more time for each recharge.

A limitation of the NILL battery is its maximum discharge amperage. High amperage short-lived loads often appear in the NILL load profile due to the battery’s inability to provide sufficient power. Though this fundamental limitation of energy versus power can be overcome by combining batteries in parallel, much research is currently being done in high-power ultracapacitors to complement traditional batteries in hybrid-electric vehicles. Future battery technologies [5] would also improve NILL’s ability to eliminate high amplitude residual features.

5.3 Power Signature Analysis

Power signature based NILMs attempt to identify appliances by unique signatures that appear in the load profile. Because of the small time scale on which these signatures occur, these schemes require strictly higher resolution measurement equipment than the feature-based method. Power signatures often consider both the real and reactive power generated in domestic environments. Appliance motors experience a turn on transient in power consumption while spinning up. This transient often functions as a unique signature for the device [35, 40]. Similarly, power signatures can be found in loads with a small capacitance such as fluorescent lighting and electronic devices [39]. This technique can be highly accurate, but requires equally accurate measurements with a sampling rate of at least 600 samples per second and a resolution of tenths of amps. In the most extreme case a signature-based NILM can characterize the line noise caused by different appliances during transitional phases using MHz sampling rates [32]. As modern residential smart meters simply do not support the sampling rate or computation needed for these techniques, we do not consider them in this preliminary work.
5.4 Load Disaggregation

One final type of NILM is fundamentally different from those previously described. Load disaggregation techniques do not search the net demand for features or signatures to indicate appliances. Instead, they consider the instantaneous demand as a sum of individual appliance loads. If the load of each appliance is known, the net demand can be disaggregated into the individual loads by solving a discrete knapsack problem. Though this technique was originally considered to be computationally impractical [13], it has recently been demonstrated as feasible for residential loads [20]. In the presence of NILL, the charge or discharge battery represents an unknown quantity that must be fit into the net load along with the known appliance loads. At one watt resolution, there are thousands of possible values for the battery quantity, many of which will have valid solutions to the knapsack fitting. Thus, we believe that load disaggregation based NILMs are not a significant threat to the privacy afforded by NILL.
Chapter 6

Conclusions

We have introduced a new class of privacy-preserving algorithms called non-intrusive load leveling (NILL). NILL blinds privacy-exposing NILM algorithms by removing the majority of useful energy use transitions sensed by recently introduced smart meters. Simulations of NILL over real usage profiles in four homes showed that between 1.1 and 5.9 meaningful events were exposed to NILM algorithms per day. Such features reside amongst hundreds or thousands of false events, making recovery of appliance profiles virtually impossible under current algorithms. Future efforts will expand on the analysis to explore tradeoffs between different battery systems and the integration with existing in-home alternative energy generation technology.
References


