LocED: Location-aware Energy Disaggregation Framework

Akshay Uttama Nambi S. N. Antonio Reyes Lua R. Venkatesha Prasad

Embedded Software group, EEMCS
Delft University of Technology (TUDelft), The Netherlands
akshay.narashiman@tudelft.nl, a.reyeslua@student.tudelft.nl, r.r.venkateshaprasad@tudelft.nl

ABSTRACT

Providing detailed appliance level energy consumption information may lead consumers to understand their usage behavior and encourage them to optimize the energy usage. Non-intrusive load monitoring (NILM) or energy disaggregation aims to estimate appliance level energy consumption from the aggregate consumption data of households. NILM algorithms, proposed hitherto, are either centralized or do require high performance systems to derive appliance level data, owing to the computational complexity associated. This approach raises several issues related to scalability and privacy of consumer's data. In this paper, we present the *Location-aware Energy Disaggregation Framework (LocED)* that utilizes occupancy of users to derive accurate appliance level usage information. LocED framework limits the appliances considered for disaggregation based on the current location of occupants. Thus, LocED can provide real-time feedback on appliance level energy consumption and run on an embedded system locally at the household. We propose several accuracy metrics to study the performance of LocED. To test the robustness of LocED, we empirically evaluated it across multiple publicly available datasets. LocED has significantly high energy disaggregation accuracy while exponentially reducing the computational complexity. We also release our comprehensive dataset **DRED** (**D**utch **R**esidential **E**nergy **D**ataset) for public use, which measures electricity, occupancy and ambient parameters of the household.

Categories and Subject Descriptors

H.4 [**Information Systems Applications**]: Miscellaneous; I.5 [**Pattern Recognition**]: Applications

General Terms

Measurement, Performance Analysis, Algorithms

Keywords

NILM; energy disaggregation; localization; public dataset; smart metering

BuildSys'15, November 04–05, 2015, Seoul, South Korea
© 2015 ACM. ISBN 978-1-4503-3981-0/15/11 ...$15.00.
DOI: http://dx.doi.org/10.1145/2821650.2821659.

1. INTRODUCTION

Worldwide total energy consumption in residential and commercial buildings is estimated to be 30-40% of generation [1] and is expected to rise due to increased use of appliances and electronic devices. A significant part of this could be reduced with better real-time information of appliance level consumption statistics. With this information, users can be encouraged to change their behavior to save 5-15% of electricity usage [2, 3]. Several home automation systems are now available for providing feedback on energy usage. Such systems lack the ability to provide appliance level consumption feedback and personalized recommendations in real-time to the occupants. One of the most important benefits of appliance level usage information is providing automated personalized recommendations by identifying which appliances could most effectively reduce energy usage in a household. The recommendation system will be able to inform the occupants on potential savings by deferring usage of an appliance to the time of a day when the electricity price is lesser. Furthermore, fine-grained information can also be used to identify faulty or malfunctioning appliances that consume more energy than they should. Consequently, occupants know where the energy is being wasted. Several utility companies (or utilities) are now interested in providing appliance level consumption feedback as a service to their customers.

The most common way of obtaining appliance level information is by deploying sensors for each appliance. Such a deployment is intrusive, cumbersome to maintain and has high cost. Alternatively, recent home energy monitoring techniques have utilized *non-intrusive load monitoring* (NILM) algorithms that aim to break down a household's aggregate energy consumption into individual appliances [5]. NILM techniques are gaining popularity due to low cost sensors for measuring energy usage, large-scale smart meter deployments to obtain household's aggregate energy consumption and inference algorithms proposed for energy disaggregation [4, 5, 6].

There still exist several challenges preventing NILM techniques to be widely adopted in households: (i) Most of the proposed mechanisms consider only a subset of appliances – a few high energy consuming appliances – for disaggregation. This is due to the exponential computation complexity associated with the number of appliances, hence tractable only for a small number of appliances [7]. (ii) Several appliances with similar energy consumption profiles may exist and moreover, each appliance may have multiple states. Thus modeling and inferring accurately the states of appliances is not trivial. (iii) NILM is often performed in a centralized manner with third-party services or utilities having privacy-sensitive information of consumers. Commercially available NILM systems are required to send smart meter data to a cloud service for energy disaggregation (for example, Bidgely, PlotWatt). This approach raises several issues related to scalability and privacy. (iv) Lastly, only a few NILM systems manage to provide near real-time energy disaggregation. The ones that do provide, require detailed information of the household and its occupants, and generally utilize cloud services.

To this end, this paper presents *Location-aware energy disaggregation framework* (LocED) that utilizes user occupancy information and aggregated energy data to derive accurate appliance level information. The motivation for using location information is threefold. *First*, by utilizing location information of occupants, the NILM algorithms can reduce the number of potential appliances considered for energy disaggregation. *Second*, by reducing the state explosion, the processing power and storage capacity required for disaggregation are also reduced, making NILM algorithms tractable and implementable. *Third*, with the large-scale proliferation of smartphones and wearables, it is now possible to monitor location of the occupants (indoor room-level localization) in a non-intrusive and cost-effective manner. LocED performs energy disaggregation at the household on a low-cost embedded system such as Raspberry Pi, due to which consumers' privacy-sensitive data is stored and processed locally. This approach further is highly scalable and avoids sharing of privacy-sensitive information to the utilities.

The primary objective of this work is to develop a location-aware energy disaggregation framework that can: (i) provide real-time feedback on appliance level energy consumption; and (ii) lower the complexity of disaggregation algorithms and run on an embedded system locally at the household. We have also released our collected dataset – **DRED** (**D**utch **R**esidential **E**nergy **D**ataset) – that can be used to test the performance of disaggregation algorithms, derive appliance usage behavior and analyze demand response algorithms. Our deployment is currently live and the dataset will be constantly updated. The DRED dataset and the LocED framework is made publicly available[1] for the community to support additional analysis. The main contributions of this paper are:

- We propose a novel real-time location-aware energy disaggregation framework (LocED) to derive appliance level information with lesser computation complexity (Section 3).
- We provide our data set – *DRED* (Dutch Residential Energy Dataset) – that contains appliance level and aggregated energy data from a household. To the best of our knowledge, this is the first open-access, publicly available dataset from the Netherlands. The dataset also includes occupancy information and several ambient parameters (Section 4).
- We propose several accuracy metrics to determine the efficacy of LocED both at house level and at appliance level. LocED was empirically evaluated across several publicly available datasets (Section 5 & 6).

[1] `http://www.st.ewi.tudelft.nl/~akshay/dred/`

2. RELATED WORK

Several NILM algorithms have been proposed in the literature to derive fine-grained appliance level information. These algorithms rely on various techniques (supervised, semi-supervised or unsupervised) and also additional data [8]. We first provide details of the existing algorithms and then describe how our approach enhances the current state-of-the-art NILM algorithms.

NILM Techniques

Unsupervised NILM techniques use no prior knowledge of the appliances but often require appliances to be manually labeled and work on low frequency (i.e., 1 Hz) data. These techniques typically rely on accurate detection and modeling of the state change in the aggregate consumption data [5, 9, 10]. Several variants of factorial hidden markov models (FHMMs) to model the states of the appliances are proposed in [5, 9]. Furthermore, other machine learning approaches such as artificial neural networks (ANNs) and genetic algorithms are also used [10]. These approaches are computationally intensive and exact inference from models with large number of HMMs is intractable.

Supervised NILM techniques assume that ground truth appliance level data is available to train and develop appliance models prior to performing disaggregation. Hart's algorithm identifies step changes in the aggregate electricity consumption and matches them with the appliance signature database to learn the states of the appliance [4]. Other approaches employ both real and reactive power measurements for energy disaggregation [11]. These algorithms require extensive training on appliance level data to model the states accurately.

Semi-supervised NILM techniques avoid the need to intrusively install sensors for deriving appliance signatures [6, 12]. Nambi et al. [6] propose a semi-intrusive approach to determine the most optimal number of appliances to be monitored for accurate energy disaggregation. Parson et al. [12] utilize prior models of general appliance types, which are tuned to specific appliance instances using signatures extracted from the aggregate load. In general, due to computational complexity involved in training and inference, these algorithms require systems with high processing power for energy disaggregation and hence are not suitable for low power embedded systems.

Additional data considered in NILM

NILM algorithms also use different additional information (either energy related or contextual data) to simplify energy disaggregation and enhance its accuracy. Some algorithms rely only on real power consumption of the household [5, 4]. However, other algorithms require both real and reactive power for energy disaggregation [11]. Recent algorithms use information on how loads are distributed across different phases in a household [11, 13] or use transient and harmonic information with very high frequency sampling [14]. However, sampling at high frequency requires expensive hardware and determining appliance distribution across different phases is not trivial. Algorithms described in [15, 16] employ information provided by other sensors as additional input for energy disaggregation. Rowe et al. [15] propose an event detector to determine the state change by sensing the electromagnetic field (EMF) in the surrounding. Kim et al. [16] utilize ambient signals from inexpensive sensors placed near

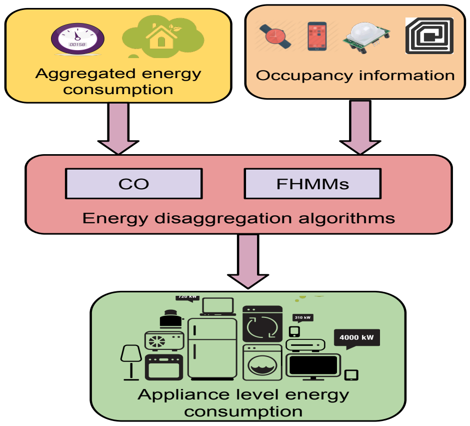

Figure 1: Location-aware energy disaggregation.

appliances to estimate power consumption. While the aforementioned approaches improve NILM accuracy, they also require additional deployment and maintenance of these sensors. Moreover, algorithms developed by using these additional data are generally constrained to a particular dataset or a household; consequently, making it nearly impossible to employ the algorithm with other publicly available datasets.

One of the major roadblocks in large-scale adoption of NILM algorithms is its scalability [8]. The proposed LocED framework utilizes a modified combinatorial optimization (CO) algorithm to reduce the computation complexity and accurately infer the states of the appliances. Our framework can be used with any dataset containing occupants room level location information, for example, Smart* [19] and iAWE [20] datasets collect occupants room-level location information using PIR sensors. Unlike the existing approaches [19, 20, 21], we do not deploy any additional sensors for deriving location information in our DRED dataset, but rather utilize WiFi/Bluetooth(BT) received signal strength (RSS) data from occupants smartphone/wearable to derive room-level location information. We show the efficacy of the proposed LocED framework by evaluating it across several publicly available datasets and our own dataset. The framework and the dataset is made open for the community for further analysis. To best of our knowledge (apart from Non-Intrusive Load Monitoring Toolkit (NILMTK) [7]), we are one of the firsts to validate and compare NILM algorithms across multiple datasets.

3. LOCATION-AWARE ENERGY DISAGGREGATION

In this section, we describe the usage of occupancy information to derive accurate appliance state information. Fig. 1 shows the block diagram of location-aware energy disaggregation.

3.1 User occupancy modeling

Occupancy information is generally used to develop efficient energy management systems for smart homes [23]. For example, occupancy information can be used to control the HVAC system efficiently or turn off appliances (lights) when user has left the room. We employ user occupancy information to improve NILM algorithms by considering only those appliances that are in the current user location for disaggregation. Several direct and indirect approaches have been proposed in the literature to derive user occupancy information [23]. Direct approaches employ low cost sensors such as passive infrared (PIR), reed switches, RFID tags to determine room-level occupancy information. Even-though these approaches are cost-effective, they are cumbersome to maintain and intrusive in residential settings.

In this work, we employ an indirect approach for deriving occupancy information with the help of smartphones/ wearables. Indirect approaches does not use additional hardware deployment, but rely on existing infrastructure for localization. Smartphones and wearables enable collection of received signal strength (RSS) from WiFi and/or Bluetooth (BT) radios in an indoor environment. In our DRED dataset (see Section. 4), we collected both Bluetooth (BT) and WiFi RSS information using occupants mobile phone to infer user location. To save battery and also to derive accurate location, a radio scan is performed only upon detection of a user movement (i.e., change in accelerometer data or step detection).

The data stream from a radio scan includes the list of all visible access points (APs) and their RSS values along with the timestamp information. In case of a WiFi scan, the list of APs indicate the access points from the neighboring houses, whereas the BT scan indicates the Bluetooth beacons available in the house. Currently there exist several Bluetooth enabled devices in a household such as laptops, mobile phones, speakers, etc. Furthermore, in the near-future most of the household appliances will be Bluetooth enabled [2]. Bluetooth enabled devices can now determine accurately indoor location information of the occupants. Classification techniques such as Bayesian, Support Vector Machines, K-nearest neighbor, decision trees, etc., have been proposed in the literature to derive room-level occupancy using RSS information. Our localization algorithm is based on Bayesian classification technique and has two phases *viz.,* training and testing phase as shown in Fig. 2. During the training phase, data is collected at each room to build a classifier model. In testing phase, a new data from the scan is evaluated using the classifier model built to obtain the room-level occupancy information.

For more details on our WiFi and BT localization algorithms see Sec.4 of [18]. The LocED framework is independent of the approaches used in obtaining location information.

[2]http://www.bluetooth.com/Pages/Smart-Home-Market.aspx

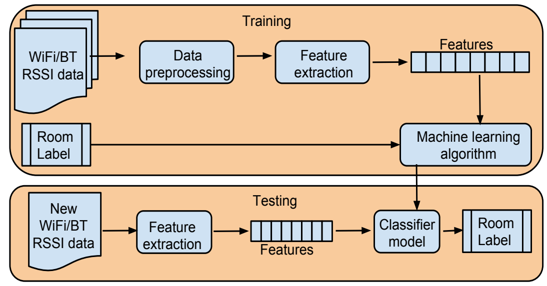

Figure 2: Indoor localization using WiFi/BT RSSI.

3.2 Aggregate energy consumption modeling

We provide a brief description of the CO algorithm for energy disaggregation [4] and then, propose a modified CO algorithm used in our LocED framework.

Combinatorial Optimization (CO): The goal of an energy disaggregation algorithm is to provide estimates of actual energy consumed by each appliance from the aggregate energy consumption data. Let $\hat{y}_t^{(n)}$ be the estimated energy consumed and $y_t^{(n)}$ be the actual energy demand of each appliance n at time t. $\overline{y}_t$ represent the aggregate energy reading of the household. The ground truth state of an appliance is represented by $x_t^{(n)} \in Z \geq 0$ and $\hat{x}_t^{(n)}$ represents the appliance state estimated by the disaggregation algorithm. CO finds the optimal combination of appliance states, which minimizes the difference between the sum of predicted appliance power and the observed aggregate power. It is given by,

$$\hat{x}_t^{(n)} = \arg\min_{\hat{x}_t^{(n)}} \left| \overline{y}_t - \sum_{n=1}^{N} \hat{y}_t^{(n)} \right| \quad (1)$$

where N is the set of all appliances in the household and t is the current time period. The predicted energy consumption of an appliance $\hat{y}_t^{(n)}$ is then mapped to the closest appliance state $x_t^{(n)}$. This approach requires an appliance model, which includes power consumption details for each state of the appliance. This is further used during inference to predict the current state of the appliance. The computational complexity of disaggregation for T time periods is $O(TS^N)$, where S is the number of appliance states and N is the set of all appliances.

CO algorithm has several drawbacks. Firstly, this optimization problem resembles subset sum problem and is NP-complete. Furthermore, the computation complexity in CO increases exponentially with the number of appliances. Secondly, this algorithm does not differentiate between appliances with similar power consumption and appliances with similar states. Third, this algorithm assumes all the appliances in the household are being monitored and assigns some portion of energy to appliances even if they are not currently used, resulting in low disaggregation accuracy.

3.3 LocED Framework

LocED framework includes preprocessing techniques that can simplify the NILM computation and improve energy disaggregation accuracy. LocED framework utilize aggregated energy data and occupants location information to derive accurate appliance level information.

We propose a modified CO algorithm to overcome some of the drawbacks of original CO. Our modified CO algorithm, constrains the number of appliances considered for disaggregation based on the current location of the occupants. This results in exponential reduction in state space for disaggregation. Furthermore, we employ a crowd-sourced generic appliance model from the power consumption database. For example, the power consumption database provides crowd-sourced information on maximum and idle power for a wide range of loads indexed by type, manufacturer, and model number[3]. This information can be obtained *a priori* based on the appliances in the household from the manufacturers datasheet or crowd-sourced data, thus eliminating appliance level energy modeling. Furthermore, our modified CO algorithm requires to know the number of appliances and their location in the household. This metadata information is collected once during the deployment and, except from a few appliances like vacuum cleaner, hair dryer, the location of the appliances is generally static. Fig. 3 shows an overview of the proposed LocED framework.

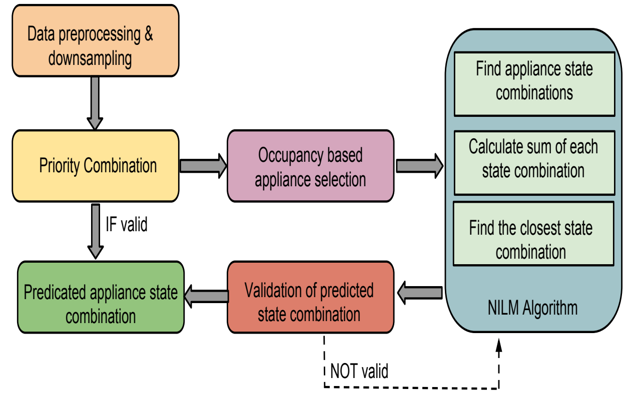

Figure 3: An overview of LocED Framework.

Data preprocessing and downsampling: Our framework can handle various data sampling rates and is designed to work with several datasets. In general, during data collection there might be gaps in the data due to sensor malfunction, network connectivity, etc. Hence, it is important to preprocess these gaps either by removing them or using statistical models such as smoothing, interpolation, forward filling, etc. Furthermore, different datasets include different sampling intervals typically from 1 second to 15 minutes. LocED applies a downsampling mechanism similar to NILMTK, to filter transients that occur due to high starting current of an appliance.

Priority combination: In original CO, at each time period the algorithm tries to find the set of appliances, which are closest to the current aggregated energy consumption. This may result in different set of appliances being used in each time period. For example, at time period 't', CO may determine appliance TV and microwave are being currently used and at time period '$t + 1$' it may select fan and microwave. This is due to the fact that TV and fan may have similar energy consumption profiles. This result would mean TV is switched ON in one minute and switched OFF the next minute and so on. Hence, it is necessary to preserve consistency in selection of appliances during consecutive state estimations. LocED defines a *priority combination* that is the set of appliances which are assumed to be currently running. This information can be retrieved from the last iteration of NILM algorithm. At each time period, LocED first evaluates the priority combination to check whether the sum of all appliances in the priority combination matches the current aggregated value. If the difference between the sum of priority combination and the aggregated energy is within a threshold δ, then the current priority combination is retained as the predicted set. LocED evaluates the following equation to determine whether the current priority combination of appliances are still valid or

[3]The Power consumption database. [Online] http://www.tpcdb.com/

not, $[|\overline{y}_t - \sum_{n=1}^{K} \hat{y}_t^{(n)}| \leq \delta]$, where K is the set of appliances present in the priority combination and δ is the variation threshold. The variation threshold parameter ensures small fluctuations in aggregate power has minimal effect. Since these fluctuations vary for different appliances based on their power rating, the δ value needs to be adaptive. The δ value can be obtained by analyzing the energy consumption profiles of the appliances. However, when the difference between current priority combination and aggregate consumption is greater than δ, LocED finds the new set of appliances that are used.

Occupancy based appliance selection: When the current priority combination does not match the aggregate energy consumption, LocED estimates the set of appliances that could be currently used. This stage identifies the set of appliances which are present in the current user location. For example, if the current location information of all occupants includes Kitchen and Living room, only appliances present in these locations are considered valid during that time period for energy disaggregation. In general, the appliances considered for evaluation at a particular time period include, (i) appliances present in the current location of the occupants; (ii) appliances that are already "ON"; (iii) appliances that are always "ON", these are autonomous appliances such as Refrigerator; and (iv) appliances that can be remotely controlled such as lights and other smart appliances. We refer to these appliances as "constrained set of appliances". LocED uses this constrained set for energy disaggregation rather than the complete set of appliances present in the household. If for a time period, there is no occupancy information available all appliances present in the household are considered for evaluation.

CO based NILM algorithm: In this work, we employ modified CO algorithm to find the optimal combination of appliance states. We calculate the sum of all possible state combinations from the constrained set and select the closest combination of appliances that match the aggregated energy consumption. The computational complexity of disaggregation for T time periods in LocED is $O(TS^{N_c})$, where S is the number of appliance states, N_c is the constrained set of appliances and $N_c \leq N$. This reduced computational complexity enables LocED to determine the state of appliances in real-time. As mentioned earlier, other NILM algorithms can be used at this stage to infer the state of the appliances from the constrained set. For example, in case of FHMMs the constrained list of appliances can be used during decoding the HMM state sequence.

Validation: We now validate the set of appliances predicted in the previous stage. Using occupancy based appliance selection, LocED ensures we do not turn "ON" an appliance when user is not present in that location. However, validation stage ensures not to turn "OFF" an already "ON" appliance when the appliance location is different than the current user location (except remotely controllable appliances). Moreover, this depends on the type of the appliance. In this work, we broadly classify the set of appliances into: (i) User dependent appliances – appliances that require user interaction to turn "OFF", for example, TV, fan, etc., and (ii) User independent appliances – appliances that can turn "OFF" themselves and require no user interaction, for example, microwave, washing machine, dishwasher, etc. If the set of appliances selected in the previous stage involves one or more user dependent appliances being turned "OFF" when the occupants location differs from the appliance location, validation stage eliminates this combination of appliances. LocED then selects the second closest combination from the previous stage and re-validates.

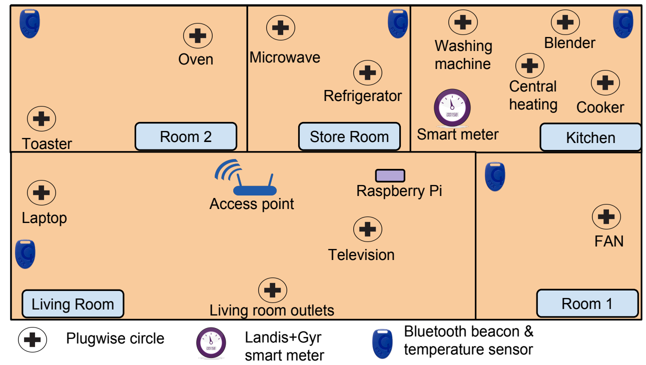

Figure 4: Deployment setup in DRED.

4. THE DRED DATASET

In this section, we describe the details of our live deployment and the sensor data collected from a household in the Netherlands. The dataset includes both appliance level and mains level energy consumption data. We currently release over 2 months of data to the research community. We refer to this dataset as **DRED** (**D**utch **R**esidential **E**nergy **D**ataset). Fig. 4 shows the layout of our deployment along with location of the sensors and appliances in the household.

4.1 Sensing infrastructure and data collection

Our live deployment consists of several sensors measuring electricity, occupancy and ambient parameters in a household. The objective of collecting the data was to test the performance of energy disaggregation algorithms, derive appliance usage behavior and analyze demand response algorithms. Similar to Smart* and iAWE, we decided to measure all possible parameters. The sensors were carefully installed to avoid any inconvenience for the occupants.

Electricity monitoring: We used off-the-shelf sensors to monitor energy consumption at 1 Hz sampling frequency.

(i) *Mains level*: We installed a smart electricity meter from Landis+Gyr E350 to measure the aggregate energy consumption information of a household. The data from the smart meter was retrieved using Plugwise Smile[4].

(ii) *Appliance level*: We used off-the-shelf smart plugs from Plugwise circle[5] to collect appliance level energy consumption data. 12 smart plugs were installed to monitor the appliances across the household, *viz.,* (1) Refrigerator, (2) Washing Machine, (3) Central Heating, (4) Microwave, (5) Oven, (6) Cooker, (7) Blender, (8) Toaster, (9) TV, (10) Fan, (11) Living room outlets, and (12) Laptop.
The plugs installed in the household communicate via Zigbee protocol by forming a mesh network. We use an open source library python-plugwise to query the data from the plugs at

[4]Smile:`https://www.plugwise.com/smile-p1`
[5]Circle:`https://www.plugwise.com/circle`

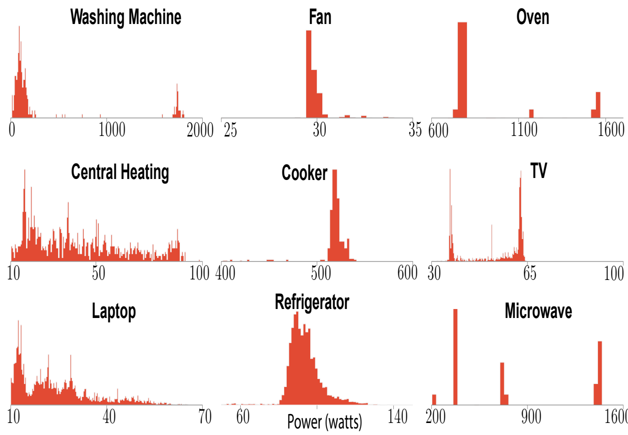

Figure 5: Histograms of appliance load profiles in DRED.

1 Hz frequency. A Raspberry Pi was deployed locally to generate periodic queries and to store the data. Furthermore, this data is also sent to a server for making it available for the research community. Fig. 5 shows the histograms of appliance power demand in our dataset. It can be seen that some appliances have multiple states (washing machine, oven) and others have only two states of operation (fridge, cooker).

Ambient monitoring: Apart from collecting energy related data, in our deployment we also collected room level indoor temperature, outside temperature, wind speed, precipitation and humidity. We deployed low-cost Bluetooth beacons from Gimbal[6] with in-built temperature sensor for each room and also one outside the house. These beacons have a battery lifetime of 4 to 5 months. A smartphone and smartwatch application in Android was developed to read the data from these beacons every 1 minute. The wind speed, precipitation and humidity data was collected from the publicly available Royal Netherlands Meteorological Institute (KNMI) website every hour[7].

Occupancy monitoring: In our deployment, we scan both visible WiFi access points and the Bluetooth beacons present in the household for indoor localization every 1 minute. This data is further used with different machine learning algorithms to determine the indoor room level location of occupants. The room level location inferred from the localization algorithm is also made available. For WiFi based localization, no additional infrastructure is deployed, however, for BT based localization we deployed the BT beacons, which could be further replaced by the smart Bluetooth enabled devices.

Household metadata: Our dataset also includes household metadata such as number of occupants, house layout, mapping between appliance and location. This metadata is generally useful for NILM algorithms. Further details on the metadata can be found in [17].

4.2 Dataset characteristics and comparison with existing datasets

It is important to compare and evaluate energy management algorithms across datasets from different countries, due to change in usage behavior of appliances. The Reference Energy Disaggregation Dataset (REDD) was the first publicly available dataset to test NILM algorithms [22]. This was followed by other datasets such as BLUED [24], SMART* [19], AMPds [25], iAWE [20], ECO [21], UK-Dale [26] and Pecan Street [27]. We describe how our DRED dataset extends the current publicly available datasets:

(i) In DRED, almost all appliances are monitored and has very constant baseline consumption. Baseline consumption includes appliances which are occasionally used (guest devices) or not monitored. Popular datasets such as REDD, Smart*, iAWE and ECO has very high and varying baseline consumption. This variation significantly hinders the performance of NILM algorithms.
(ii) DRED dataset has less than 5% dropout rate in energy data. Dropout rate indicates the missing data due to communication issue or sensor faults. Most of the other datasets have around 10-20% dropout rate apart from Smart*.
(iii) Our deployment is still live. The dataset released contains over 2 months of data and will be updated every month. Only ECO and UK-Dale have data greater than 100 days.
(iv) Even though ECO, Smart*, iAWE datasets include occupancy data, they have large gaps and missing data. However, DRED uses an indirect sensing approach for obtaining room-level occupancy information and has high data availability rate.

We believe that DRED dataset with the above-mentioned extensions compared to other datasets will be useful to the research community to validate NILM algorithms and analyze energy management algorithms. Furthermore, we provide HDF5 version of the DRED dataset for direct usage with NILMTK toolkit [17].

5. EVALUATION

5.1 Datasets

We provide performance evaluation results of the proposed framework across multiple datasets to support wide-adoption and also to validate our work. Our framework imports data from DRED dataset and also other popular publicly available datasets such as REDD (House 1), Smart* and iAWE. Hence, we show the performance results across four datasets collected in different countries.

Dataset Statistics: Each dataset includes data from different set of appliances and for varying time duration. In order to evaluate the performance of LocED across multiple datasets, it is necessary to understand the characteristics of each dataset. Fig. 6 shows the characteristics of the datasets. NILMTK already provides some basic functions such as mains availability, percentage of energy sub-metered and top-k appliance to analyze the dataset, however we further extend these functions in LocED.

Fig. 6(a) shows the percentage of total energy measured at the appliance level for all days in the dataset. Most of the datasets do not monitor all the appliances in the household, leading to large (sometimes more than 50%) unaccounted energy in the aggregated consumption data. Furthermore, the variation of this unaccounted energy data significantly reduces the accuracy of disaggregation algorithms. DRED has around 75% of energy sub metered and all other datasets have around 45% of energy measured at the appliance level. It is clear from the figure that DRED captures

[6]https://store.gimbal.com/collections/beacons/products/s10
[7]KNMI:`http://www.knmi.nl/climatology/daily_data/selection.cgi`

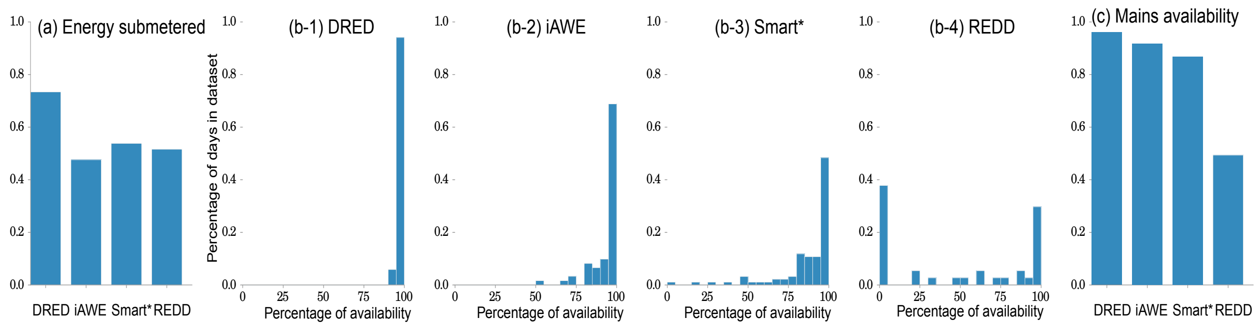

Figure 6: Data characteristics across different datasets.

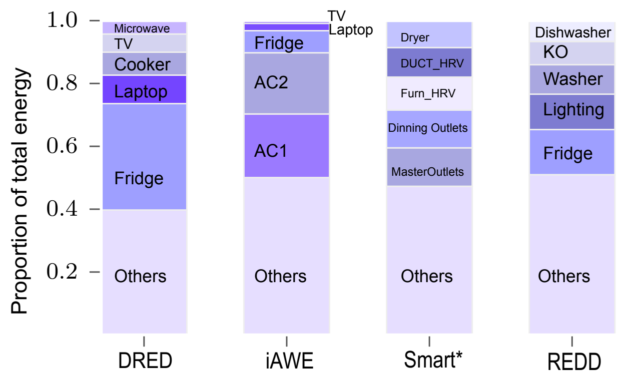

Figure 7: Energy consumed by top-5 appliances.

significant proportion of the energy consumed in the household, whereas, iAWE dataset has the lowest percentage of energy captured at the appliance level.

Another important statistic to be considered is the percentage of aggregated data available in the dataset. This is the ratio of the number of data points recorded over the total number of data points that can be collected in a day. Fig. 6(b) shows histogram of average aggregated data available throughout the dataset. The y-axis indicates the percentage of days and x-axis indicates the data availability rate. DRED has more than 90% of the aggregated data available throughout the data collection period (Fig. 6(b-1)). Other datasets have much lower data availability. This may be due to communication issues or malfunctioning of sensors deployed.

Similar to the previous statistic, Fig. 6(c) shows the average percentage of data availability at mains level across datasets. DRED dataset has around 95% data availability rate for all the appliances being monitored. The other datasets have 90%, 86% and 50% mains data availability rates for all the days. REDD has only 50% of aggregated data availability this is due to long gaps in data collection (can also be seen in Fig. 6(b-4)).

In general, only a few appliances constitute the majority of power consumed in a household. Hence, it is necessary to derive accurate information of these high power consuming appliances during energy disaggregation. Fig. 7 shows the proportion of energy consumed by top-5 appliances and other appliances present in the household across datasets. It is interesting to see the variation of top-5 appliances across datasets, indicating the varying preference of appliance usage in different countries. The top-5 appliances in DRED cover around 60% of total energy consumed.

Finally, since LocED relies on the occupancy information collected, it is important to find the occupancy data availability rate. The occupancy availability rate is the ratio of total number of occupancy data recorded over the total number of expected occupancy data. DRED, iAWE and Smart* has occupancy rate of 81%, 76% and 36% respectively. We further determine the relevant occupancy information that corresponds to the usage of appliances at that time period. For example, if an appliance currently being used is in living room and the occupancy data includes living room as one of the occupant's location, then, this occupancy data is considered as relevant. DRED and iAWE has 68% and 53% of valid occupancy information, whereas, Smart* has only about 10%.

5.2 Accuracy Metrics

Several accuracy metrics both at house level and at appliance level are considered for evaluation of LocED. Different metrics at house level are described below:

Fraction of total energy assigned correctly (FTE): It measures the fraction of energy correctly assigned to an appliance and is one of the common accuracy metrics for NILM algorithms [22, 7]. FTE is the overlap between the actual fraction of energy consumed by each appliance and the fraction of energy assigned to each appliance. It is defined as,

$$FTE = \sum_{n} min\left(\frac{\sum_{n} y_t^{(n)}}{\sum_{n,t} y_t^{(n)}}, \frac{\sum_{n} \hat{y}_t^{(n)}}{\sum_{n,t} \hat{y}_t^{(n)}}\right), \tag{2}$$

where $n \in \{1, .., N\}$ and N is the total number of appliances. Also $t \in \{1, .., T\}$ and T is the total time period considered.

Total disaggregation error (T_e): Total disaggregation error is the difference between the total energy consumed by all appliances and the actual energy consumed by the appliances, normalized by the total energy consumed. It is given by,

$$T_e = \frac{\sum_{n,t} |y_t^{(n)} - \hat{y}_t^{(n)}|}{\sum_{n,t} y_t^{(n)}} \tag{3}$$

We employed the functions provided in NILMTK for calculating FTE and T_e metrics.

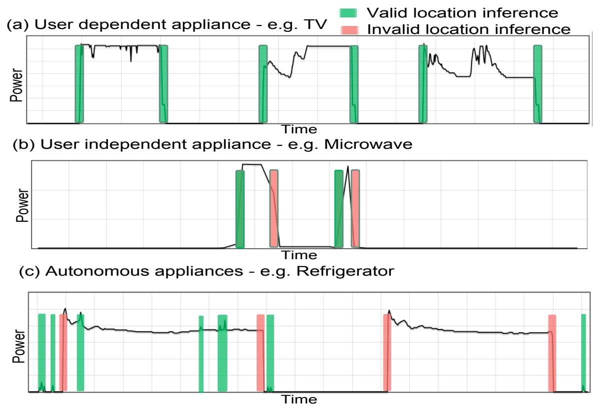

Figure 8: Location inference from REDD dataset.

Number of appliances identified correctly (J_a): Jaccard similarity coefficient is used to measure the similarity between the predicted set of appliances (J_a^p) and the actual set of appliances (J_a^a) used over a time period. J_a measures the percentage of appliances correctly identified by the disaggregation algorithm. It is given by,

$$J_a = \frac{|J_a^p \cap J_a^a|}{|J_a^p \cup J_a^a|} \quad (4)$$

Number of appliance states identified correctly (J_s): It measures the similarity between the predicted set of appliance states (J_s^p) and the actual set of appliance states (J_s^a). It is given by,

$$J_s = \frac{|J_s^p \cap J_s^a|}{|J_s^p \cup J_s^a|} \quad (5)$$

We now describe the set of metrics considered at the appliance level for evaluation.

Proportion error per appliance (P_e): It measures the difference between the proportion of the energy assigned to an appliance and the actual energy consumed by the same appliance. It is defined as,

$$P_e = \left| \sum_t y_t^{(n)} - \sum_t \hat{y}_t^{(n)} \right| \quad (6)$$

Normalized error per appliance (N_e): It measures the sum of the differences between the assigned energy and the actual energy consumed by the appliance, normalized by the total energy consumed by the appliance. It is given by,

$$N_e = \frac{\sum_t |y_t^{(n)} - \hat{y}_t^{(n)}|}{\sum_t y_t^{(n)}} \quad (7)$$

6. RESULTS

We compare LocED disaggregation results across various datasets. REDD dataset from MIT does not include occupancy data. To this end, we developed a module to infer user location information from the ground truth appliance level data in the REDD dataset.

Location inference from REDD dataset: To enable fair comparison across popular datasets, we infer user location with the help of appliance level data in REDD. LocED differentiates between user dependent and user independent appliances to accurately infer occupancy information. Fig. 8 shows the locations inferred based on the appliance energy consumption information. For a user dependent appliance, a user is present in that location when an appliance is being turned "ON" or "OFF" (see Fig. 8(a)). Similarly, for a user independent appliance, a user is present during the "ON" event but may or may not be present during the "OFF" event. Hence we label this as an invalid location as shown in Fig. 8(b). Furthermore, special consideration needs to be given for appliances such as Refrigerator, where occupancy information is valid only when a user opens/closes the door (Fig. 8(c)). We eliminate the compressor energy consumption and infer locations only when the refrigerator door is opened/closed. Please note that location information only when an appliance is being used will be available with the above mentioned inference procedure. For further details see Sec.6 of [18].

We now show the performance of LocED and original CO algorithm. To ensure fair comparison, both LocED and CO utilize the same appliance model from the crowd-sourced database as described in Section 3.3. Since the model and make of an appliance varies from one dataset to another due to the geo-location of data collected in these datasets, applying a generic model across all datasets is challenging. LocED uses a crowd-sourced appliance model from the power consumption database based on the manufacturer and model number of an appliance. In our evaluation, we used data obtained using direct sensing (PIR sensors) in Smart*, iAWE datasets and also data from indirect sensing in DRED dataset. Furthermore, we used an adaptive δ value for determining the priority combination. The values of δ was determined based on the appliance type. From our experimentation, we found that appliances with low power consumption have lower noise and smaller variation in their energy consumption and appliances with high power consumption have large variation due to the noise associated. Furthermore, the δ value can be also used to account for unmonitored appliances or guest appliances by modeling the historic household energy consumption.

Fig. 9 shows the disaggregation performance of CO and LocED across the house level accuracy metrics. We considered one week of data (the week with highest data availability rate) across all the four datasets. In general, FTE, J_a and J_s can vary between 0 and 1, and T_e can take any non-negative value. It can be seen that, LocED performs significantly better across all the datasets for all the metrics. LocED performs better than CO mainly due to two reasons, (i) LocED ensures that the predicted set of appliances does not vary significantly for consecutive time periods, thanks to *priority combination.* (ii) LocED constrains the number of appliances considered to disaggregate based on occupancy information ensuring similar appliances from a different location are not selected.

Fig. 9(a) shows that in DRED, LocED correctly assigns up to 80% of energy to all appliances, which is 40% more compared to CO. Furthermore, it determines more than 25% of correct appliances and states than original CO. Fig. 9(c) shows more than 30% improvement across all metrics for Smart* dataset and similar trends can be seen in iAWE and REDD datasets. LocED also has much lower T_e across all datasets compared to CO. Fig. 9(e),(f),(g),(h) show the disaggregation performance of CO and LocED for top-k (k=5)

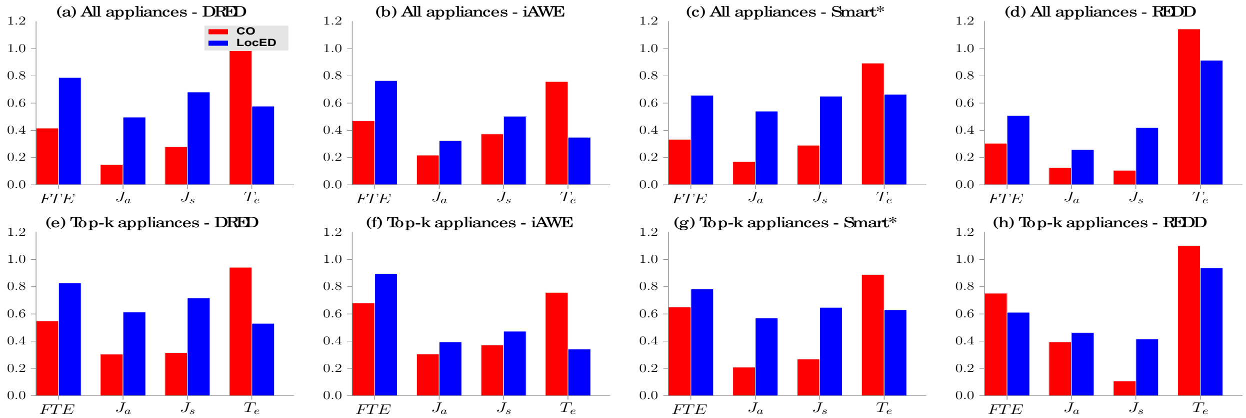

Figure 9: Disaggregation performance of CO and LocED across datasets (1 week).

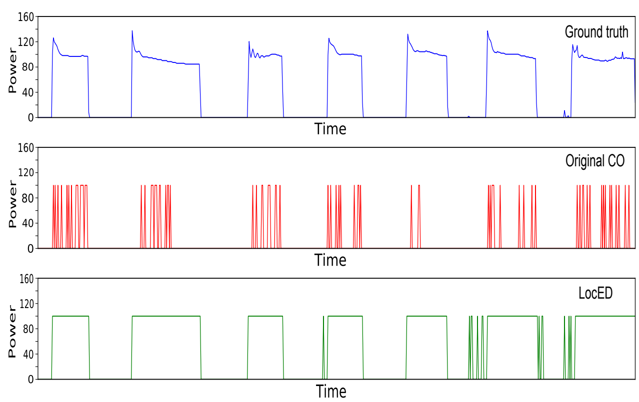

Figure 10: Original and disaggregated energy profile of refrigerator using CO and LocED.

Dataset	CO P_e	CO N_e	LocED P_e	LocED N_e
DRED	0.07	5.51	0.04	3.16
iAWE	0.10	12.66	0.09	7.70
Smart*	0.14	22.07	0.12	13.02
REDD	0.06	39.20	0.05	23.01

Table 1: Appliance level accuracy metrics for all appliances (all days).

Dataset	All Appliances FTE	J_a	J_s	T_e	Top-k Appliances FTE	J_a	J_s	T_e
DRED	30.5	28.7	36.6	-37.9	22.4	23.3	36.5	-41.2
iAWE	8.5	3.2	2.2	-7.9	14.8	3.3	4.5	-9.6
Smart*	29.3	28.3	28.6	-14.4	1.9	28.1	30.4	-18.6
REDD	11.4	12.7	27.6	-13.6	-22.1	5.3	27.3	-8.8

Table 2: Percentage increase in performance of LocED over CO (all days).

appliances. As mentioned previously, disaggregating accurately top energy consuming appliances would be very beneficial to reduce cost and manage energy efficiently. It can be seen that LocED correctly assigns upto 89% of energy to the top-k appliances in iAWE and around 80% in DRED and Smart* datasets. Furthermore, the number of appliances and states identified is also higher compared to original CO. In DRED and Smart* datasets the number of appliances and states determined is more than 30% compared to original CO.

Fig. 10 shows the original power consumed by the refrigerator (top) and the resulting disaggregation output using the original CO (middle) and LocED (bottom). CO has a interrupted load profile due to its sensitivity to small changes in aggregated power, however, LocED overcomes this with the help of priority combination and the δ parameter described in Section 3.3.

Table. 1 shows the appliance level accuracy metrics for all appliances across all days. In general, P_e and N_e can take any non-negative values. It can be seen that across all the datasets, P_e and N_e values for LocED are lower compared to CO; indicating better energy disaggregation for all appliances.

Table. 2 shows the percentage increase in disaggregation accuracy of LocED compared to CO for all the days across the datasets with all appliances and top-k appliances. It can be seen that FTE improvement of 30%, 9%, 30%, and 12% is obtained for all days considered in DRED, iAWE, Smart* and REDD datasets respectively. Similarly, number of appliances correctly identified improves over 30% for all days considered in DRED and Smart* datasets respectively. In DRED, FTE improvement of 22% was achieved for top-k appliances and number of appliances and states identified improved by 23% and 36% respectively. The negative T_e shows the percentage reduction in total error achieved by LocED. The FTE for top-k appliances in REDD dataset is lower for LocED. This is likely due to wrong inference of locations from appliance ground truth data.

In general, if the occupants are spread out across the building or if all the appliances are close to one another, then the benefits of using location information for disaggregation is less. However, in residential settings as seen from the above datasets these cases arise occasionally. In our evaluation, we showed that even with very less location information, LocED was still significantly able to improve the disaggregation accuracy. Furthermore, the framework proposed can include other contextual information such as room temperature, number of users, etc. to further improve energy disaggregation accuracy.

Finally, we also computed the average number of state combinations evaluated in each dataset by CO and LocED to disaggregate. Original CO has a fixed number of state combinations depending upon the number of appliances and its states. However, for LocED the number of appliances considered varies and is determined based on the constrained set of appliances. In iAWE and Smart* the average state combinations to be evaluated for disaggregating a value is 59049 and 8192 for CO and it is 162 and 60 for LocED. Similarly, in DRED 104976 combinations was evaluated by CO and LocED evaluated only 10 combinations on average. It can be seen that across all datasets the average number of state combinations evaluated by LocED is drastically reduced, consequently, decreasing the computation complexity for real-time disaggregation.

7. CONCLUSIONS AND FUTURE WORK

We proposed a *novel* location-aware energy disaggregation framework (LocED) to derive accurate appliance level data. We employed a modified CO algorithm to infer the state of the appliances accurately. We also presented a comprehensive dataset DRED that can be used to test the performance of energy disaggregation algorithms, derive appliance usage behavior and analyze demand response algorithms. We evaluated LocED across multiple publicly available datasets such as DRED, iAWE, Smart* and REDD. Our evaluation shows that around 80% disaggregation accuracy can be achieved for all appliances on DRED and iAWE datasets. Furthermore, up to 90% accuracy is achieved when only top-5 appliances are considered for disaggregation in DRED and iAWE. The number of correctly identified appliances and states are 61% and 68% in DRED using LocED.

Even with additional location information there are errors associated with disaggregation due to several factors. In most of the datasets, due to lack of knowledge on number of appliances and lack of monitoring of all appliances in the household, there exists a significant amount of unaccounted energy in the aggregate consumption. Only DRED dataset monitors almost all appliances and has a very low variation in baseline consumption. Moreover, the percentage of occupancy information available plays an important role in improving the accuracy. Only DRED and iAWE have more than 70% of occupancy data available. Furthermore, LocED uses a generic approximate model to find the states of an appliance. Accurate modeling of appliance states will further improve the disaggregation performance.

8. REFERENCES

[1] Energy outlook 2010. Energy Information Administration. [Online] http://www.eia.doe.gov/oiaf/ieo/index.html, 2010.

[2] Y Guo, M Jones, B Cowan, and R Beale. Take it personally: personal accountability and energy consumption in domestic households. *In CHI, 2013.*

[3] C. Fischer. Feedback on household electricity consumption: A tool for saving energy? *Energy Efficiency, 2008.*

[4] G. W. Hart. Nonintrusive appliance load monitoring. *Proc. of the IEEE, 1992.*

[5] J.Z.Kolter and T.Jaakkola. Approximate inference in additive factorial HMMs with application to energy disaggregation. *In Proc. AIS-TATS, 2012.*

[6] S. N. Akshay Uttama Nambi, T G. Papaioannou, D Chakraborty and K Aberer. Sustainable energy consumption monitoring in residential settings. *In Proc. IEEE INFOCOM 2013.*

[7] N. Batra, J. Kelly, O. Parson, H. Dutta, W. Knottenbelt, A. Rogers, A. Singh, and M. Srivastava. NILMTK: An open source toolkit for non-intrusive load monitoring. *In Proc. e-Energy. ACM, 2014.*

[8] M. Zeifman and K. Roth. Nonintrusive appliance load monitoring:Review and outlook. *IEEE Trans. on Consumer Electronics, 2011.*

[9] H. Kim, M. Marwah, M. Arlitt, G. Lyon, and J. Han. Unsupervised disaggregation of low frequency power measurements. *In Proc. SDM. SIAM, 2010.*

[10] M. Baranski and J. Voss. Genetic algorithm for pattern detection in NIALM systems. *In Proc. SMCS. IEEE, 2004.*

[11] M. Weiss, A. Helfenstein, F. Mattern, and T. Staake. Leveraging smart meter data to recognize home appliances. *In Proc. PerCom. IEEE, 2012.*

[12] O. Parson, S. Ghosh, M. Weal, and A. Rogers. Non-intrusive load monitoring using prior models of general appliance types. *In Proc. AAAI, 2012.*

[13] N. Batra, H. Dutta, A. Singh. INDiC: Improved Non-Intrusive Load Monitoring using load Division and Calibration. *In Proc. of ICMLA, 2013.*

[14] S. Gupta, M. S. Reynolds, and S. N. Patel. ElectriSense: Single-point sensing using EMI for electrical event detection and classification in the home. *In Proc. UbiComp. ACM, 2010.*

[15] A. Rowe, M. Berges, and R. Rajkumar. Contactless sensing of appliance state transitions through variations in electromagnetic fields. *In Proc. BuildSys. ACM, 2010.*

[16] Y. Kim, T. Schmid, Z. Charbiwala, and M. Srivastava. ViridiScope: Design and implementation of a fine grained power monitoring system for homes. *In UbiComp 2009.*

[17] Dutch Residential Energy Dataset. [Online] `http://www.st.ewi.tudelft.nl/~akshay/dred/`

[18] Akshay Uttama Nambi, Antonio Reyes, Venkatesha Prasad. *[Technical Report]* LocED: Loaction-aware Energy Disaggregation Framework [Online] `http://www.st.ewi.tudelft.nl/~akshay/papers/Tech_report_2015_01_july_TUD.pdf`.

[19] S. Barker, A. Mishra, D. Irwin, E. Cecchet, P. Shenoy, and J. Albrecht. Smart*: An open data set and tools for enabling research in sustainable homes. *In SustKDD, 2012.*

[20] N. Batra, M. Gulati, A. Singh, and M. B. Srivastava. It's Different: Insights into home energy consumption in India. *In Proc. Buildsys, 2013.*

[21] C. Beckel, W. Kleiminger, R. Cicchetti, T. Staake, S. Santini. The ECO Data Set and the Performance of Non-Intrusive Load Monitoring Algorithms. *In Proc. Buildsys 2014.*

[22] J. Z. Kolter, M. J. Johnson. REDD: A public data set for energy disaggregation research. *In SustKDD, 2011.*

[23] J. Lu, T. Sookoor, V. Srinivasan, G. Gao, B. Holben, J. Stankovic, E. Field, and K. Whitehouse. The smart thermostat: Using occupancy sensors to save energy in homes. *In Proc. SenSys, Nov. 2010.*

[24] K. Anderson, A. Ocneanu, D. Benitez, D. Carlson, A. Rowe, and M. Berges. BLUED: A fully labeled public dataset for Event-Based Non-Intrusive load monitoring research. *In Proc. SustKDD, 2012.*

[25] S. Makonin, F. Popowich, L. Bartram, B. Gill, and I. V. Bajic. AMPds: A Public Dataset for Load Disaggregation and Eco-Feedback Research. *In Proc. EPEC, 2013.*

[26] J. Kelly and W. Knottenbelt. UK-DALE: A dataset recording UK Domestic Appliance-Level Electricity demand and whole-house demand. *ArXiv, 2014.*

[27] C. Holcomb. Pecan Street Inc.: A Test-bed for NILM. *In Proc. NILM workshop, 2012.*