Non-Intrusive Air Leakage Detection in Residential Homes

Nilavra Pathak[‡], David Lachut[†], Nirmalya Roy[‡], Nilanjan Banerjee[†], Ryan Robucci[⁎]
[‡]Information Systems,[†]Computer Science, [⁎]Computer Engineering
University of Maryland, Baltimore County
nilavra1,dlachut1,nroy,nilanb,robucci@umbc.edu

ABSTRACT

Air leakages pose a major problem in both residential and commercial buildings. They increase the utility bill and result in excessive usage of Heating Ventilation and Air Conditioning (HVAC) systems, which impacts the environment and causes discomfort to residents. Repairing air leakages in a building is an expensive and time intensive task. Even detecting the leakages can require extensive professional testing. In this paper, we propose a method to identify the leaky homes from a set, provided their energy consumption data is accessible from residential smart meters. In the first phase, we employ a Non-Intrusive Load Monitoring (NILM) technique to disaggregate the HVAC data from total power consumption for several homes. We propose a recurrent neural network and a denoising autoencoder based approach to identify the 'ON' and 'OFF' cycles of the HVACs and their overall usages. We categorize the typical HVAC consumption of about 200 homes and any probable insulation and leakage problems using the Air Changes per Hour at 50 Pa (ACH_{50}) metric in the Dataport datasets. We perform our proposed NILM analysis on different granularities of smart meter data such as 1 min, 15 mins, and 1 hour to observe its effect on classifying the leaky homes. Our results show that disaggregation can be used to identify the residential air-conditioning, at 1 min granularity which in turn helps us to identify the leaky potential homes, with an accuracy of 86%.

CCS CONCEPTS

• **Computing methodologies → Neural networks**;

KEYWORDS

Air Leakage Detection, Non-Intrusive Load Monitoring , Deep Learning

ACM Reference Format:
Nilavra Pathak[‡], David Lachut[†], Nirmalya Roy[‡], Nilanjan Banerjee[†], Ryan Robucci[⁎]. 2018. Non-Intrusive Air Leakage Detection in Residential Homes. In *ICDCN '18: 19th International Conference on Distributed Computing and Networking, January 4–7, 2018, Varanasi, India.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3154273.3154345

ICDCN '18, January 4–7, 2018, Varanasi, India

ACM ISBN 978-1-4503-6372-3/18/01…$15.00
https://doi.org/10.1145/3154273.3154345

1 INTRODUCTION

Residential and commercial sectors consume the second highest amount of total energy usage in USA, and in residential sector about 40% of it goes to Heating Ventilation and Air Conditioning (HVAC) [1]. HVAC efficiency depends to a large degree on air-tightness, but also on several other factors such as building insulation, thermostat setpoint, etc. Air leakage in homes can result in temperature differentials, causing hot or cold regions, and can increase HVAC consumption. Air leaks in homes can occur through windows, doors, attics, and basements. Such leaks can cause HVAC to heat or cool a house for longer durations, increasing the utility bill. Uneven heating or cooling can be another effect of leakage which can cause discomfort to residents. A proper energy audit to check leakages, construction quality and other factors influencing the heat dynamics of the building, is expensive and labor intensive.

The blower door test is used to detect the amount of air leakage, and is measured as ACH_{50} or CFM_{50}. The former metric, ACH_{50}, is the more common unit used by blower door operators, which stands for Air Changes per Hour at 50 Pascals. Alternatively, CFM_{50} per square foot of building envelope is used as a metric where one CFM_{50} is a cubic foot per minute with a pressure difference of 50 Pascals between inside and outside. A required ACH_{50} value for a new home in the United States is between 3.0 and 5.0, with the exact value depending on the climate zone. There are multiple Climate zones in the United States as shown in Figure 1, and different climate zones have different requirements for building construction materials, infiltration rate, and insulation. For new buildings to be constructed the maximum requisite infiltration rate needs to satisfied. But for existing homes with ACH_{50} greater than 3 there is potential to repair the leakages and make them more air-tight. The blower door test typically costs around $450 and a more intensive audit may cost more.

Motivated by these problems, we look into quantitative means for non-intrusively determining potential homes where air leakage can be present in a large scale. We hypothesize that given certain building parameters and the HVAC usage, we can classify homes as "leaky" or "not-leaky". However, getting access to air-conditioning usage of homes is difficult. In general, utility providers have access to the smart meter data which is captured about 15 minute granularity. We propose to perform non-intrusive load monitoring (NILM) or energy disaggregation to gather the air-conditioning data, which can be used to generate features for the "leaky" home detection. Two important features which help the classification are the average 'ON' power consumption for AC and the duration for which it is in the 'ON' state. Our objective for disaggregation is to obtain these two factors.

The variance in air-conditioning power and usage from house to house makes it more challenging to detect these factors. Air conditioners come in different sizes and the usage patterns depends on the thermostat setting and house type. Therefore, finding a generalized approach to disaggregating energy data is a challenging task. We use the Dataport Dataset [16] which has a large collection of homes with aggregated and sub-metered data, and metadata regarding the home's ACH_{50} value , house size, thermostat setpoints which give us access to a rich dataset for out analytics. Previous work in [21] has shown a heuristic approach to finding residential air conditioning from the dataport dataset. However it works well for homes which has A/C 'ON' power wattage of 2.5 kW and more. In [15] building energy usage and metadata has been used to group homes and use a representative home's circuit level information to perform disaggregation in a neighborhood. A deep learning based approach has been proposed in [2] where three deep learning architectures were used to perform disaggregation with individualized appliance specific disaggregation model.

In our work we choose to perform disaggregation at 1 min, 15 mins and 60 mins granularity. Since we require only the air conditioning usage we take the deep learning approach of modeling one network to disaggregate A/C, and test the validity of the deep learning based disaggregation on a large number of similar homes. We construct two deep neural network architectures - LSTM and De-noising autoencoder. We apply the LSTM for 15 mins and 1 hour granular data and the denoising autoencoder on the 1 min granular data. Finally, using the disaggregated air-conditioning data, we check the classification results from the features obtained from disaggregated data. We make the following key contributions in this paper.

- We describe a leaky home classification approach by generating features from air-conditioning usage and building metadata information. We validate whether meaningful A/C usage features can be obtained from the A/C consumption and show that the usage and the power is necessary for the classification.
- We perform disaggregation on a large scale data to disaggregate air-conditioning usage from the aggregate energy consumption of about 70 homes with varying sizes over a period of 4 years.
- We propose two deep neural network architectures for disaggregating energy data of three different resolutions - 1mins, 15mins and 60mins, and compare their performance on disaggregation and leaky home classification.

This paper is structured as follows: In Section 2 we provide a related work explaining air leakage and the methods to identify air leakage in home, and also we provide a brief overview of the previous works in NILM domain. In Section 3 we show the data analytics on leaky home classification. In Section 4 we describe our approach to energy disaggregation designed based on two neural net architectures. In Section 5 we present the results of disaggregation and leaky home classification. Finally, we conclude in Section 6.

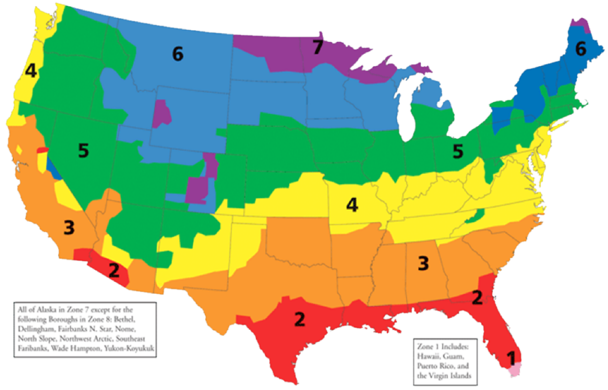


Figure 1: Climate Zone of USA

2 RELATED WORK

In this section, we first provide an overview of the traditional approaches to quantifying the air leakage in a house. We also discuss different structural and energy efficiency properties of a building which can be presumed as ground truths for categorizing a house as leaky or not. We present the methods used for air leakage detection and assessment through different tests. We briefly describe the different energy disaggregation methods in our context.

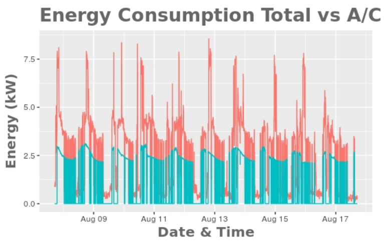
2.1 Air Leakage and Measurements

In this section, we explain some of the technical terms and methodology related with building air leakages. Air leakage occurs when outside air enters and conditioned air leaves the house through cracks and openings. During cold or windy weather, too much air may enter the house and, during warm or calm weather, too little. Also, a leaky house that allows moldy, dusty crawlspace or attic air to enter is not healthy, that can cause structural damages to the building. To find the air-leakages expensive tests are required. The blower door test provides an overall measure for leakage present in the house, and is measured by the metrics ACH_{50} or CFM_{50}. The blower door test comprises of a fan unit set into an exterior doorway, sealing the face of the fan, that measures baseline indoor and outdoor pressure differential. The idea behind creating the pressure differential is to force air out through the leaks. The amount of leakage is quantified by several metrics of which Air changes per hour at 50Pa (ACH_{50}) is the most common. It is represented as:

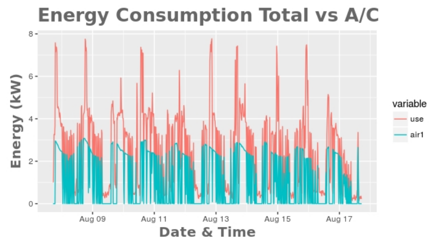
$$ACH_{50} = CFM_{50} * \frac{60}{Volume\_of\_Building} \tag{1}$$

where CFM_{50} is the airflow at 50 pascal (ft^3/minute) and Volume_of_Building is the volume of the building in cubic feet. The value of ACH_{50} is proportional to the leakage. The ideal value of ACH_{50} is less than equal to 3 in climate zones 3 and above, and less than equal to 5 for climate zones 1-2. A map of US climate zones is shown in Figure 1. The data that we use for our analysis is from a city in climate zone 2 and hence the acceptable ACH_{50} is 5 and a maximum allowed value for the house to be infiltration proof is 7.

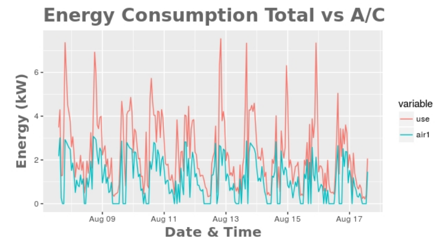
Once the ACH_{50} measurements are obtained for the house, and if the house has detected to have a poor ACH_{50} or CFM_{50} value, the house needs to be closely inspected for leakages. The task of



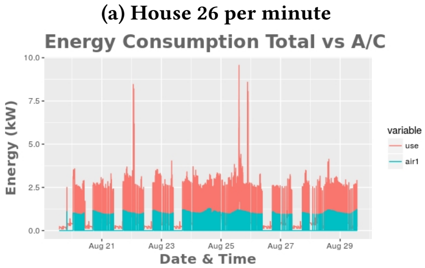
(a) House 26 per minute



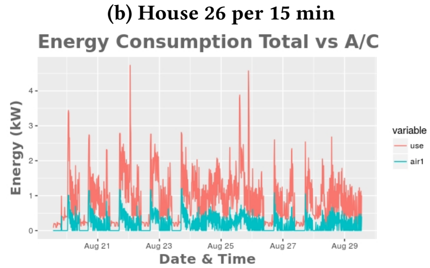
(b) House 26 per 15 min



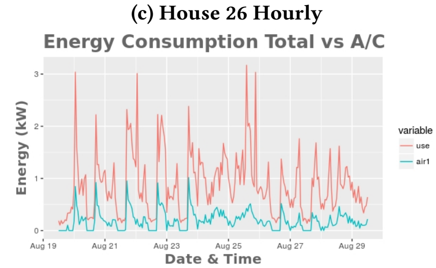
(c) House 26 Hourly



(d) House 94 per minute

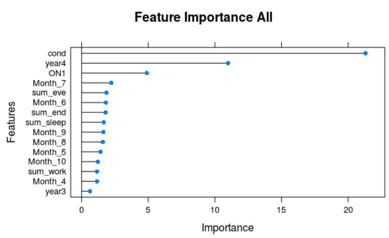


(e) House 94 per 15 min

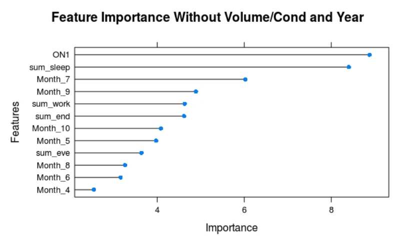


(f) House 94 Hourly

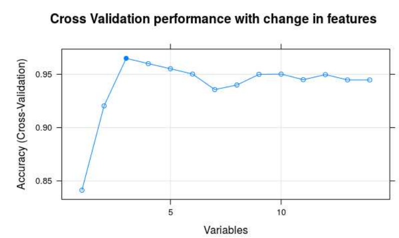
Figure 2: Pecan Street House Total vs A/C Usage



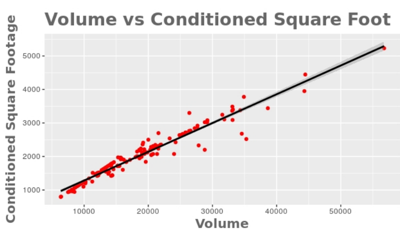
(a) Ranked Features



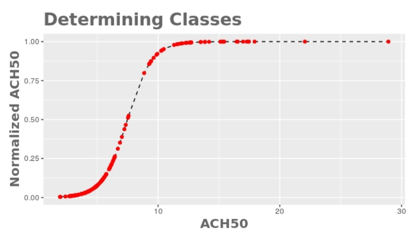
(b) Ranked Feature without Cond and Year



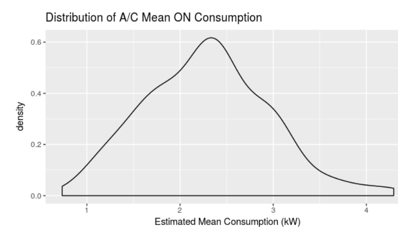
(c) Cross Validation Results for Different Features



(d) Volume vs Condition



(e) ACH_{50} Classes



(f) Distribution of A/C kW

Figure 3: Pecan Street Data Overview

finding the leakages is time consuming and expensive as it requires trained professionals to perform the inspections. In recent practice, Infrared (IR) Cameras are used to detect air-leakages, damp, leakage in ducts and insulation [4]. A review of the main areas of usage of IR cameras in buildings was provided in [27]. In [28, 29], spatial and temporal variations of thermographic data are quantitatively analyzed to provide reports on insulation deficiencies, air leakage detection and moisture content mapping.

2.2 Non-Intrusive Load Monitoring

Non-Intrusive Load Monitoring (NILM) or Energy disaggregation is the process of estimating the usage of individual appliances in a home from the aggregated energy consumption [3]. Two of the most common approaches are Combinatorial Optimization (CO) [3] and variants of Hidden Markov models (HMMs). A particular variant of HMM popularly used in disaggregation is Factorial Hidden Markov Model (FHMM) [6]. A benchmark result for CO and FHMM has been provided in [5] for a number of different datasets. FHMM is also suitable to perform disaggregation in scenarios with ambient sensors fusion [11] and Additive FHMM was part of an approach to reduce the number of needed energy sensors in renewable-powered homes by [30]. Although CO and FHMM are good approaches to disaggregation, yet they suffer from different drawbacks. In disaggregation nomenclature, supervised disaggregation assume that sub-metered appliance training data are available from the household in which disaggregation is to be performed. However, such an assumption does not quite hold when it comes to scale the model. [3, 22, 23]. One of the most challenging tasks is unsupervised disaggregation. In that approach, no prior knowledge of the appliances is assumed and the task is to identify the usage pattern of the appliances, and then profile and label them. The previous works on unsupervised disaggregation [10, 13] suffer from the drawback of the requirement for manually labeling the appliance data. A method for modeling the appliances were stated in [17], which can help label the appliances. More recent approaches try to address the labeling problems, however the scope of the problem is limited to small scale appliances. For large scale appliances, a third approach has been suggested in [18], where an exhaustive list of appliances are are profiled for disaggregation. The transferability of disaggregation is a particular challenge where the disaggregation method trained from one home is used to disaggregate the power consumption of other homes. Matrix factorization based approaches such as Non-Negative Matrix Factorization or Sparse Coding techniques have been proposed to perform transferable disaggregation in large scale [12, 23–25]. A deep network based approach to disaggregation has been suggested in [2], where three variants of deep neural network techniques were used to perform disaggregation - a recurrent neural network using a Long-Short Term Memory [7], a denoising auto-encoder [8] and a Regress Start Time, End Time and Power model. The Regress model helps estimate real-valued outputs, for example, the start time, the end time and mean power demand of the activation cycle of an appliance in the aggregate power signal. LSTM has also been used in other works like [19, 20], but most of the deep learning methods are applied to datasets which have granularity in seconds. Our choice for dataset is the Dataport dataset [16], which is a large repository of energy data and most granular data present therein is at 1 min level. We choose the homes where aggregated consumption and air-condition usage data are available along with some other metadata like building size, year of construction, thermostat settings etc. Since we have to determine the homes with leakages, we only choose to disaggregate air-conditioning data as that only is necessary for determining homes with potential leakages. Previous work for disaggregating residential air-conditioning from sub-hourly loads were performed in [21], where disaggregation has been performed across different granularities. The approach to disaggregation uses a heuristic algorithm where the authors consider A/C units which consume about 2.5kW of power. The proposed method is incapable of disaggregation in presence of other high loads and A/C units with lower power, such as 1kW, which easily get mixed up with other similarly sized loads. To circumvent the challenges we utilize a deep learning based approach to perform disaggregation and obtain the A/C unit power consumption and the 'ON' period. Another approach has been stated in [15], where a K-NN classification has been performed prior to energy disaggregation to form a neighborhood of homes for which there exist a candidate sub-metered house, whose data can be used to perform disaggregation. The methods described in [2] and [15] propose two different arguments for solving the NILM as a big data challenge. While the former hypothesizes that given enough data, deep neural networks can learn features for an arbitrarily complex appliance and perform disaggregation effectively, the later approach leverages the "higher order" relationship that exists between homes to disaggregate home's total energy. For our problem, we choose variants of deep neural network and disaggregate data from larger datasets to help classify the leaky homes.

3 LEAKY HOME CLASSIFICATION

Air leakage in a house is determined by ACH_{50}. As per the metadata available, a house is considered to be tight if the ACH_{50} is less than or equal to 7. To determine the classes we applied a sigmoid normalization where the transformed ACH_{50} is given by

$$Norm\_ACH_{50} = \frac{1}{1 + e^{-(ACH_{50} - \overline{ACH_{50}})}} \tag{2}$$

The relation between the normalized and original value of ACH_{50} has been shown in Figure 3e.It represents a distinct group of possible values ranging from 0 to 7.6 and 7.6 and above. We thus form two classes – "leaky" if ACH_{50} is more than 7.6 and "Not leaky" is less than that.

To build our classification module, we first identify the set of necessary features . Our hypothesis is that air-leakage will have some effect on the A/C usage, for which we validate by constructing features from available usage data. In Figure 2, the A/C usage for a house is shown for two different durations and considering a one minute data granularity. First we applied a continuous Hidden Markov Model for each home's A/C data to obtain the mean 'ON' power as a feature. We then assigned the entire A/C data to 1 when 'ON' cycle is observed and 0 when 'OFF' cycle is noted. This allowed us to obtain the duration of usage. Our assumption is that homes which require more cooling either have a low thermostat setting or have more leakage. Hence we calculate the monthly consumption and use it as a feature. Since the number of minutes is constant for a particular month, the total duration for which the A/C is 'ON' has an upper bound for every month. We selected the months April - October for our analytics as the A/C usage is more dominant during those months. We performed feature ranking by importance whose results are provided in Figure 3a. The most import features are Volume, Conditioned Square Foot, and the Year of construction. In Figure 3d, it can be seen that the volume and conditioned square foot are proportional and linearly related, hence we select conditioned square footage as a feature only as it is more readily available than the volume of the house. In addition, volume is often not documented for some houses in the dataset, which is an important factor to compute the missing ACH_{50} values. We applied a robust linear regression to estimate the missing volume values from the available conditioned square foot data and then used Equation 1 to compute ACH_{50} with the available Blower door test at 50 Pa values. We also computed the importance of the other features as shown in in Figure 3b. Figure 3c presents the result of our feature selection using Recursive Feature Elimination (RFE).

3.1 Dataset

We perform our experiments on approximately 202 homes from the publicly available Dataport data set [16] that have direct submetering infrastructure installed for one year and more, with available A/C data. Dataport dataset contains aggregate and appliance-level power consumption information for more than 700 homes in Texas, USA for up to 5 years and beyond. The sub-metered and household aggregate power data were collected at 1-minute resolution. We only selected the single-family homes in the city of Austin, Texas and found 159 homes contain data for 1 year or more with HVAC consumption. We also search for homes with some key metadata like size of house, number of floors, thermostat settings, and available ground truth for leakage i.e., ACH_{50}. Of the 159 homes available some have two air-conditioning compressor units, which we chose to ignore. In total there are around 70 unique homes with few having data for more than a year. We considered each home to be a datapoint that have data for a year, for example, House 94 has data in both 2013 and 2014 and we deem them as two separate homes, so that we can have a larger dataset for classification. Table 1 depicts this.

Table 1: Building Characteristics

No. of Homes	Construction Year	Avg Sq. ft.	ACH_{50} average
73	1800-1975	1529	10.9
26	1975-2000	1904	7.38
103	2000-2017	2141	4.42

3.2 Leakage Classification Results

We performed the classification on 202 homes, with a 70-30% split in training and test sets. The results for different features are given in Table 2. It can be observed that the 'ON' consumption mean has a significant effect and certain months are crucial to finding the usage patterns, which help directly to infer the homes with leakages. We chose multiple combinations of features and found that monthly usage information is important, which justifies the use of air conditioner usage data. Better analysis can be done with availability of a larger dataset, but this helps narrow down the homes which otherwise will require expensive energy audits. We applied Random Forest classification and observed that only power consumption and conditioned square footage is enough to classify with a F-measure of 0.73. The results of classification for different feature sets is given in Table 3.

Table 2: List of Features

Features	Suffixes
Monthly A/C Consumption in the ith month	Month_i (i = 4 ..10)
Mean A/C ON Power	ON
House conditioned square footage	Cond
House Volume	Vol
Thermostat Setpoints	T_i

4 NON INTRUSIVE LOAD MONITORING

In the previous section we found that the 'ON' consumption value and the duration of usage of air-conditioning are important for the classification. However it is unlikely that the utility providers will have access to individual air- conditioning data, which obviates the necessity for energy disaggregation. Typical disaggregation approaches attempt to figure out the individual appliance working patterns and average power consumption for all appliances. However, for air leakage detection we require only the air-conditioning data and hence partial disaggregation will suffice in our case. We employ two deep learning based approaches to perform this partial energy disaggregation where one network is well suited to disaggregate a distinct load. We also evaluate their performance across power consumption data with varying granularities. The Pecan Street Dataport dataset provides data with 1 min, 15 mins and 60 mins granularity. Our objective is to test for the robustness of disaggregation methods across different granularities and their effect on leaky home classification, primarily because typical smart meter data has 15 minute granularity. We implement two deep learning methods: i) a recurrent neural network based on Convolutional LSTM and ii) a denoising auto-encoder. Next we describe the architectures of the above two approaches.

4.1 Recurrent Network based approach

The concept behind the use of a recurrent neural network is to train in a regression like manner. At every time step, the network sees a single sample of aggregated power and outputs a single sample power data for the target appliance. In our case, since only air-conditioner needs to be estimated, we train one single recurrent neural network model. LSTM [7], has been used in earlier attempts to perform disaggregation [2, 19, 20] and proved to be successful when applied for data with second level granularity. We added a L1- regularizer for each of the hidden dense layers and performed disaggregation by providing one input at a time. We chose a convolutional LSTM with the following architecture for our experiments for the 15 minutes and 60 minutes data granularity. The structure of the chosen LSTM model is as follows.

- Input layer (1D input with different sequence length)
- 1D conv (filter size = 5, stride =1, number of filters = 16, activation function = relu, border mode = same)
- LSTM (N = 128, with return sequence, inner activation = relu)
- LSTM (N = 256, inner activation = relu)
- Dense (N = 256, with return sequence, inner activation = relu)
- Dense (N = 1, with return sequence, inner activation = linear)

4.2 Denoising autoencoder based approach

A Denoising autoencoder (DA) [8] was applied in [2] to perform disaggregation, considering the corruption as being the power demand from the other appliance swhile setting the appliance of interest without any corruption. In our early work in [23], we applied a variant of auto-encoder called sparse coding for disaggregation where we considered the appliance to be noise and attempted to remove that from the signal. The denoising autoencoder in [2] required the

Table 3: Classification Results of Leaky Homes

Features	F1 Score (Train, Test)
All	0.97 , 0.84
All - Year	0.94 , 0.69
All - Year,Cond =Months_i, ON	0.91, 0.71
Ranked Order 1: year, Cond, temp summer weekday workday, ON	0.96, 0.65
Ranked Order 2: year, Cond, temp summer weekday workday, ON, Month_7, Month_8	0.98 , 0.75
Ranked Order 3: year, Cond, temp summer weekday workday, temp summer sleeping hours, ON, Month_7, Month_8	0.95 , 0.75
Ranked Order 3: temp summer weekday workday, temp summer sleeping hours, ON, Month_7, Month_8 , Month_5	0.95 , 0.78
Ranked Order 4: temp summer weekday workday, temp summer sleeping hours, ON, Month_7, Month_8 , Month_5	0.92 , 0.72
A/C and Cond: ON, Month_4, Month_5 , Month_6, Month_7 , Month_8, Month_9 , Month_10, Cond	0.93 ,0.86

appliance duration length to be pre-determined and varied on an appliance to appliance basis. In our case we require disaggregation at 1 min granularity rather than 1 sec which complicates the task. Generally, A/C cycles are 'ON' for 5-20 minutes, for which 1 min granular data looks like noise rather than the actual signal. Our objective is to estimate when the A/C is 'ON' correctly, which then can be fed to estimate the magnitude of the 'ON' power. The target data is an 'ON' and 'OFF' data (ON:1, OFF:0), and by choosing the last layer to be sigmoid we ensure that the range is within [0,1] and finally use a K-means clustering for binarization. The architecture of denoising autoencoder that we applied is as follows.

- Input (length 720 min)
- Conv1D(filters = 8, strides = 3, activation = relu, padding = same)
- MaxPooling1D(pool_size = 2, padding = same)
- Conv1D(filters = 32, strides = 3, activation = relu, padding = same)
- MaxPooling1D(pool_size = 2, padding = same)
- Conv1D(filters = 32, strides = 3, activation = relu, padding = same)
- UpSampling1D(size = 2)
- Conv1D(filters = 32, strides = 3, activation = relu, padding = same)
- UpSampling1D(size = 2)
- Conv1D(filters = 1, strides = 3, activation = sigmoid, padding = same)

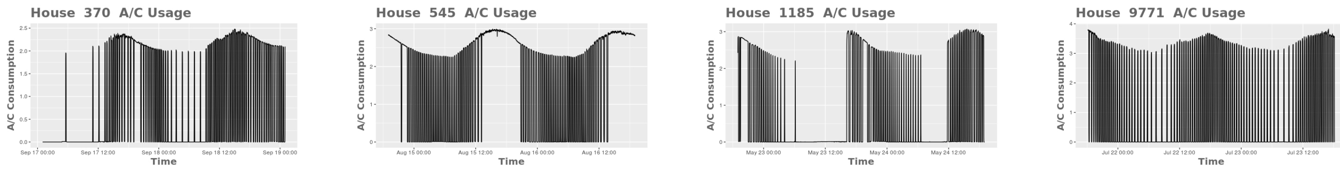
Post-Processing to obtain ON Power Consumption. The denoising auto-encoder provides an 'ON' and 'OFF' signal from the aggregated data but not the 'ON' power measure. In the post processing step we binarize the output to 1 and 0, and obtain the mean 'ON' power consumption. We do not estimate the mean 'OFF' power as it is negligible. To binarize the data we performed K-means clustering and assigned the points which have the greater centroid as 'ON' and the rest as 'OFF'. We selected the data from the months of July-September as the A/C is used more in those months, to compute the 'ON' power. The algorithm for computing the mean power consumption is given in Algorithm 1. Given the aggregated data and the 'ON'-'OFF' predicted data, we estimate the 'ON' power consumption. We only consider the instances when the A/C goes 'OFF' to 'ON' and collect the differences between the state change. When the A/C turns 'ON' it takes a minute to stabilize to the 'ON' power, so we take the difference of total use between the time 1 minute before and after the change. We store all the power changes and take the mean as the estimated 'ON' power consumption.

Algorithm 1 Mean ON Power Estimation

```
procedure MEAN ON POWER ESTIMATION (Input Use_i , Pred_i for ith appliance)
    new_use ← Use_i for months 7,8,9
    new_pred ← Pred_i for months 7,8,9
    start_state ← new_pred[1]
    for i = 2 : length(new_use) do
        if new_pred[i] = new_pred[i-1] then
            Continue
        else
            if (new_pred[i]!=new_pred[i-1]) && (new_pred[i-1]==1) then
                if new_pred[i+1]==new_pred[i] && (i<length(new_pred)) then
                    ON←ON∪(new_use[i+1]-new_use[i-1])
                end if
            end if
        end if
    end for
    ON_Power_Mean ← mean(ON[ON > 1])
end procedure
```

4.3 Transferability across homes

Transferability of disaggregation algorithms is a big issue when disaggregation model from one house is applied to another. This is due to the variation in appliances among the homes. So for scalability purpose and for practicality, the training and testing data should be from different homes. However, a model would not be able to learn when the test data is completely unknown. In Figure 4, a list of A/C usage across 4 homes are shown which have 42000 BTU (The British thermal unit (Btu or BTU) is a traditional unit of heat; it is defined as the amount of heat required to raise



(a) House 370 A/C Usage **(b) House 545 A/C Usage** **(c) House 1185 A/C Usage** **(d) House 9771 A/C Usage**

Figure 4: Differences in Consumption Patterns of the Same A/C

the temperature of one pound of water by one degree Fahrenheit. typically air-conditioners' capacity is measured in BTU). The A/C from homes 370 and 9771 are manufactured by Rheem and for 545 and 1185 by Carrier and they have the same model numbers. We note that although homes 370 and 9771 have same manufacturing company and the BTU of the A/C are the same, yet their kilo-wattage consumption is different. We inspected the UK-Dale dataset [26] used in [2] and observed that the Fridge and Kettle across the homes have more or less same wattage consumption. Due to large variation in the A/C consumption across the different homes we face the following challenges.

(1) Variation in consumption patterns: Same appliances can have different 'ON' power consumption, so if the training data does not have them then it becomes difficult to detect.
(2) A/C usage duration: The duration of A/C usage and its duty cycles can also vary even if they are of the same wattage and model. This depends on the temperature setpoint of the thermostat and can cause different duration of cycles in different homes. To address these challenges we take the following steps:
  (a) We merged two homes' data as a training data and tested on the other for LSTM. We segmented the homes in three parts depending on their square foot area and chose the home whose A/C power is closer to the mean of the subgroup for denoising autoencoder.
  (b) The duration of A/C is not pivotal for our LSTM energy disaggregation model since it requires an input dimension of size one, as the aggregated input is fed to it one instance at a time. For denoising autoencoder however the A/C duration is important and we chose windows of 720 mins with a 60 mins shift. Next we describe standardization of data.

4.4 Data Processing for disaggregation

The aggregated data is fed as input to the denoising autoencoder and the output is an equivalent 'ON' and 'OFF' pattern for the A/C. We first subtracted the mean of the aggregated use and divided by the standard deviation. For the LSTM, where we merged two homes' data for training, we standardized the data prior to merging. Targets are divided by a hand-coded maximum power demand for each appliance to put the target power demand into the range [0,1]. For the denoising auto-encoder we binarized the A/C consumption to 1 or 0 depending on when it is 'ON' or when it is 'OFF'. This is obtained by a simple k-means where the 'ON' and 'OFF' consumption centroids were obtained from the output of a Hidden Markov Model for the denoising auto-encoder. We applied a running batch normalization for each of the power signals. For the LSTM we divide the target data by a hand-coded maximum power demand for A/C to put them in a range of [0,1].

5 RESULTS OF DISAGGREGATION

We present the results of disaggregation in this section. First, we provide the different metrics used for the different disaggregation tasks. Then we provide the results of disaggregation for each granularity and its performance on leaky home classification using the disaggregated data and features.

5.1 Metrics for Disaggregation

We are actually interested to know two aspects of disaggregation, the mean power and to detect when the A/C is 'ON' and when it is 'OFF'. We use the following metrics:

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{3}$$

$$Absolute \quad percentage \quad error(APE) = \frac{|\bar{E} - E|}{(E) \times 100} \tag{4}$$

$$R - squared \quad measure = \frac{\sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2} \tag{5}$$

We estimated the mean 'ON' power consumption of the AC and its usage pattern. For the 1 hour and 15 minutes granular data, we chose the the R-squared measure to be the only metric as we attempt to predict the exact disaggregated consumption for these granularities. For the 1 min dataset we evaluated the disaggregation performance of the estimated 'ON' and 'OFF' pattern using F1 score as metric. We further estimated the mean 'ON' power consumption for which we used the absolute percentage error showed the variation in predicted mean power consumption from the original, as the mean 'ON' power is essential for classifying a home as "Leaky" or "Not-Leaky".

5.2 Hourly Data Disaggregation

The results for hourly data disaggregation are shown in Table 4. We employed the Convolutional LSTM for hourly data and the performance results are presented in terms of R-squared measure on the test data. Although hourly data is too coarse, such a granularity can be useful for energy analysis and leakage detection, particularly for providing feedback and finding patterns in energy peaks. The hourly disaggregation works well for heavy load appliances such as A/C in our case as shown in Table 4, and we can achieve a R-squared value of 0.78 for the best case. We used 12 homes as test cases and for each test home we selected two other homes' whose

Table 4: Disaggregation Results for LSTM

Granularity	House	R^2	Training	House	R^2	Training	House	R^2	Training
Hourly	2818	0.78	5275, 6990	1697	0.62	661,7731	1718	0.47	3039, 7731
Hourly	2575	0.40	661, 3009	545	0.66	1185, 2769	1642	0.52	624, 9654
Hourly	2094	0.76	1953, 2818	739	0.76	410,5275	1800	0.75	624, 2814
Hourly	1953	0.47	2818, 9926	1185	0.43	3039, 7731	1953	0.45	2829,4447
15 Minutes	2818	0.78	5275, 7731	1697	0.55	661,7731	1718	0.57	3039, 7731
15 Minutes	2575	0.43	661, 3009	545	0.64	1185, 2769	1642	0.52	624, 9654
15 Minutes	2094	0.75	1953, 2818	739	0.68	410, 5275	1800	0.59	624, 2814
15 Minutes	1953	0.55	2818, 9926	1185	0.55	3039, 7731	1953	0.59	2829,4447

AC consumption power is similar to the test home to create the training set.

5.3 Data Disaggregation for 15 minutes data

The results for 15 minutes data disaggregation is shown in Table 4. We also applied a convolutional LSTM with the same structure as that of hourly data, for 15 minute data granularity. We evaluated the performance results in terms of R-squared measure and the best result obtained is a R^2 measure of 0.78. The LSTM does not perform well on either of 1 hour or 15 minutes samples.

5.4 Data Disaggregation for 1 minute data

We only apply the denoising auto-encoder for the minute-wise data. The data with a one minute granularity is available across 5 years of which we chose 2013 - 2016 and split the data on a yearly basis. We have both the training and testing data from 2013 only, whereas for 2014 - 2016 we use datasets from 2014 for training and test them on 2014-2016. For each year we further segmented the data into small, medium and large sized homes where the the conditioned square footage are in the ranges of 0 - 1500, 1500 - 2500 and 2500-6000 sq. ft., respectively. We select a house whose overall consumption is high in the summer months. We use one home to train the auto-encoder and test on rest of the homes in the subgroup to obtain the 'ON' consumption pattern. We then use the aggregated power consumption and the disaggregated data to estimate the magnitude of mean 'ON' power. The results of disaggregation for 1 min is provided in Figure 5. We provide the Precision, Recall, F-measure and percentage error for the disaggregation results for 2013 - 2016 and overall data.

5.5 Leaky Home Classification using the Estimated AC Consumption

We perform the air-leakage classification using the disaggregated results. We apply the Random Forest classifier as before with the same features and check the applicability of disaggregated data for classification. We provided results where the features are derived from the original data and disaggregated data. We experimented with three cases, where the training and test sets are generated come from the original and disaggregated data, and showed that the disaggregated data can be properly used to classify the leaky homes, as shown in Table 5. We note that the features derived from the disaggregated result perform similar to that of the original data and hence we conclude that we can classify leaky homes with features derived from disaggregated A/C results. We presented some sample snippets of 'ON'-'OFF' prediction for A/C usage in Figure 6 from aggregated data of 4 homes from instances over 4 years.

Table 5: Comparison of Results for Air Leakage Classification with Disaggregated Results

Training	Testing	F1 (Train)	F1 (Test)
Original	Disaggregated	0.93	0.86
Disaggregated	Disaggregated	0.92	0.85
Original	Original	0.93	0.86

6 DISCUSSION

LSTM performs poorly on both the hourly data and 15 minutes data, however more experimentation can help us build a proper model for better disaggregation. In a coarse granularity the A/C use has some direct relationship with time of day and temperature, which can be utilized to get better disaggregation in a coarse level. The A/C usage patterns in 1 minute are picked up well by the denoising autoencoder, however it is difficult to predict the target power consumption with high accuracy. We believe blending different homes' data can mitigate this problem. We noticed that in the original data, the change in power consumption is not the same as of the change in power consumption in the aggregated data. The change in power for A/C sometimes is more than that in the aggregate. Various explanations can be postulated here such as the presence of phantom loads etc. We suspect some fault in the instrumentation can give rise to such errors. We observed that training the dataset on one year and testing disaggregation on another year is possible and does not cause much change as there is not much change in the power consumption patterns over the years.

Limitations: A major limitation is that, as of the time of writing, we could not gain access to adequate computational capacity to perform larger-scale disaggregation. This limits our use of extremely large scale data and testing LSTM for 1 minute granularity. Also we could not estimate the power consumption amplitude using the LSTM approach with precision. The result in that case is dependent on the training homes used in the data and probably a larger and diverse dataset can help improve disaggregation accuracy from coarse data.

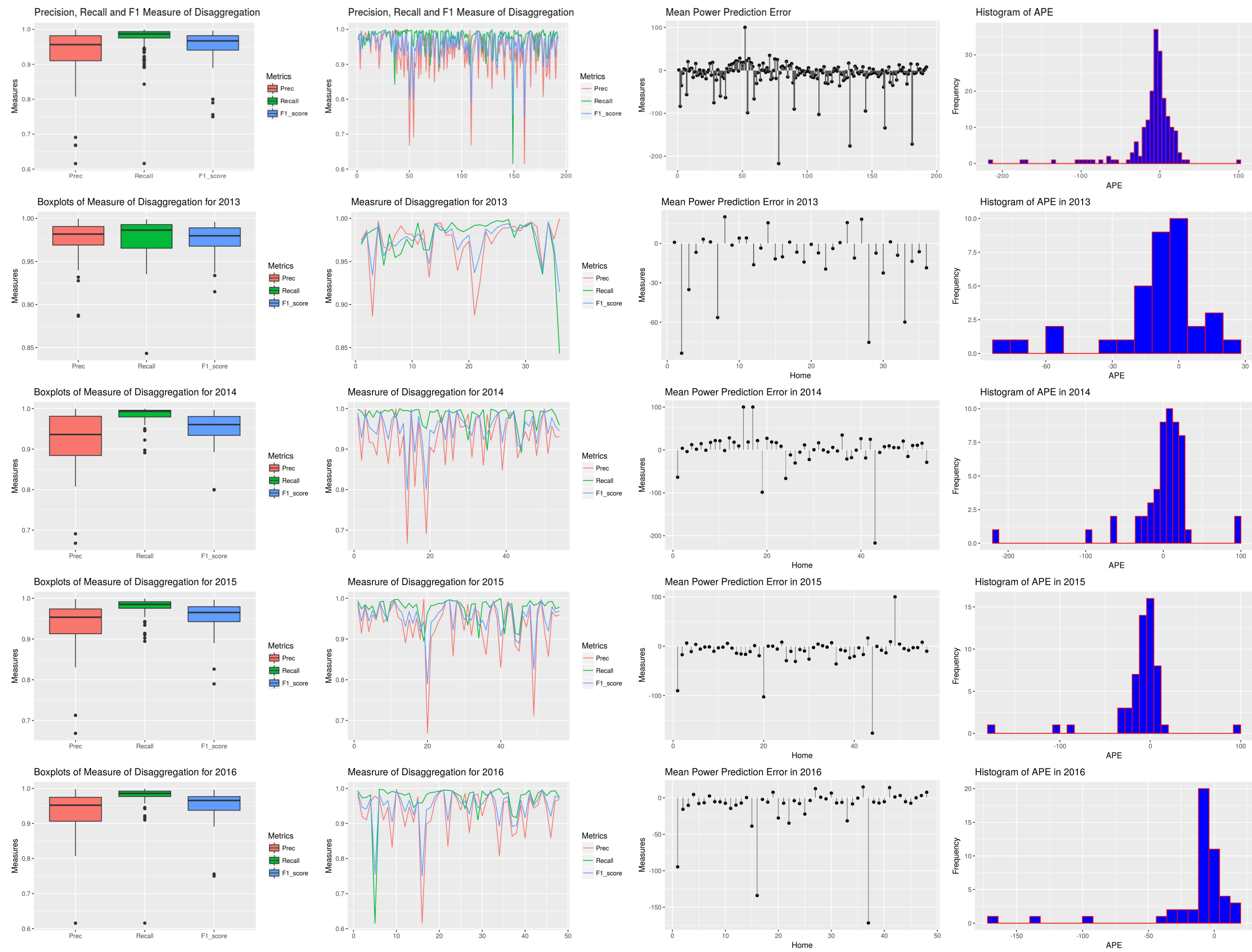


Figure 5: Disaggregation Performances across different homes. In this Figure the first column is the boxplots for Precision, Recall and F-measure for the disaggregation of the individual homes. The second column is the values of the disaggregation. The third column are the Absolute power prediction error and the final column gives the histogram of Absolute power error. The first row provides the result of all data. From second to fifth row we provide the results of disaggregation for the years 2013 - 2016.

Future Work: Our future goal is to devise a better neural network which can perform disaggregation for the higher granularity and can also help predict the power consumption. The main challenge is to obtain disaggregation at 15 minutes granularity which is available more ubiquitously from existing smart meters. We also would like to investigate the transferability of disaggregation when similar appliances with different power consumption characteristics are considered, and how large and diverse the dataset needs to be for coarse disaggregation.

7 CONCLUSION

In this paper we proposed a non intrusive data analytic approach for detecting potential homes with air leakages. We generated features from energy usage and some available metadata regarding the homes and showed the importance of non-intrusive load monitoring to obtain the A/C usage for detecting "leaky" homes. We investigated transferable approaches for energy disaggregation using a large scale publicly available dataset and applied two deep neural network algorithms for residential A/C disaggregation and found that disaggregation at 1 minute is possible using a denoising autoencoder. We investigated the performance of the different disaggregation methods across three granularities and noted that the de-noising auto-encoder can be used to estimate homes with different A/C usage characteristics. We further generated features for air leakage classification using the disaggregated data and obtained an F-measure of 0.85, which is not a significant change in performance from the original data. We thus conclude that the disaggregation at 1 minute level can be useful for air leakage detection and leaky home classification.

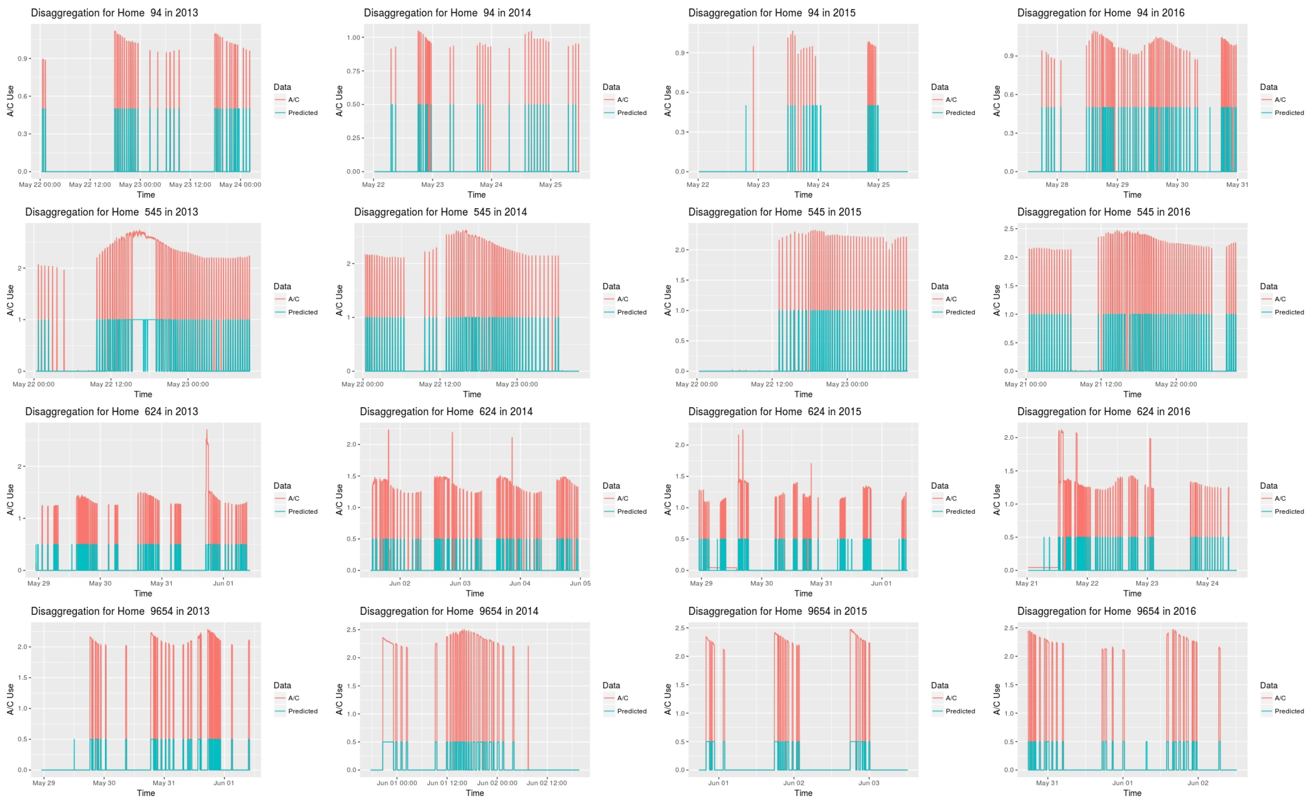


Figure 6: **Disaggregation Performances Snippets from Different Homes in Different Years. The rows are the different homes with dataid 94, 545, 624 and 9654. The columns are the years 2013, 2014, 2015 and 2016 respectively.**

8 ACKNOWLEDGEMENT

This work is supported by the NSF award # 1544687.

REFERENCES

[1] U.S. Energy Information Administration, Annual Energy Review 2011, 2012, 27-Sep-2012.

[2] Kelly, Jack, and William Knottenbelt. "Neural nilm: Deep neural networks applied to energy disaggregation." Proceedings of the 2nd ACM International Conference on Embedded Systems for Energy-Efficient Built Environments. ACM, 2015.

[3] G. W. Hart. "Nonintrusive appliance load monitoring." Proceedings of the IEEE, 80(12):1870–1891, Dec. 1992.

[4] David Lachut, Nilavra Pathak, Nilanjan Banerjee, Nirmalya Roy, Ryan Robucci. "Longitudinal Energy Waste Detection with Visualization." ACM Buildsys 2017.

[5] Nipun Batra , Jack Kelly , Oliver Parson , Haimonti Dutta , William Knottenbelt , Alex Rogers , Amarjeet Singh , Mani Srivastava, "NILMTK: an open source toolkit for non-intrusive load monitoring", Proceedings of the 5th international conference on Future energy systems, June 11-13, 2014, Cambridge, United Kingdom.

[6] J. Z. Kolter and M. J. Johnson. "REDD: A public data set for energy disaggregation research". In Workshop on Data Mining Applications in Sustainability (SIGKDD), San Diego, CA, volume 25, pages 59–62. Citeseer, 2011.

[7] Sepp Hochreiter , Jürgen Schmidhuber, "Long Short-Term Memory", Neural Computation, v.9 n.8, p.1735-1780, November 15, 1997.

[8] P. Vincent, H. Larochelle, Y. Bengio, and P.-A. Manzagol. Extracting and composing robust features with denoising autoencoders. In Proceedings of the 25th international conference on Machine learning, pages 1096-1103. ACM, 2008.

[9] Ghahramani, Zoubin, and Michael I. Jordan. "Factorial hidden Markov models." Advances in Neural Information Processing Systems. 1996.

[10] Kolter, J. Zico, and Tommi Jaakkola. "Approximate inference in additive factorial hmms with application to energy disaggregation." Artificial Intelligence and Statistics. 2012.

[11] Pathak, Nilavra, Md Abdullah Al Hafiz Khan, and Nirmalya Roy. "Acoustic based appliance state identifications for fine-grained energy analytics." Pervasive Computing and Communications (PerCom), 2015 IEEE International Conference on. IEEE, 2015.

[12] Kolter, J. Zico, Siddharth Batra, and Andrew Y. Ng. "Energy disaggregation via discriminative sparse coding." Advances in Neural Information Processing Systems. 2010.

[13] Kim, Hyungsul, et al. "Unsupervised disaggregation of low frequency power measurements." Proceedings of the 2011 SIAM International Conference on Data Mining. Society for Industrial and Applied Mathematics, 2011.

[14] Parson, Oliver, et al. "Non-Intrusive Load Monitoring Using Prior Models of General Appliance Types." AAAi. 2012.

[15] Batra, Nipun, Amarjeet Singh, and Kamin Whitehouse. "Neighbourhood NILM: A big-data approach to household energy disaggregation." arXiv preprint arXiv:1511.02900 (2015).

[16] http://www.pecanstreet.org/

[17] Iyengar, Srinivasan, David Irwin, and Prashant Shenoy. "Non-intrusive model derivation: automated modeling of residential electrical loads." Power (Watt) 500.1000 (2016): 1500.

[18] Lam, Hong Yin, G. S. K. Fung, and W. K. Lee. "A novel method to construct taxonomy electrical appliances based on load signatures." IEEE Transactions on Consumer Electronics 53.2 (2007).

[19] Mauch, Lukas, and Bin Yang. "A new approach for supervised power disaggregation by using a deep recurrent LSTM network." Signal and Information Processing (GlobalSIP), 2015 IEEE Global Conference on. IEEE, 2015.

[20] do Nascimento, Pedro Paulo Marques. "Applications of Deep Learning Techniques on NILM". Diss. Universidade Federal do Rio de Janeiro, 2016.

[21] Perez, Krystian X., et al. "Nonintrusive disaggregation of residential air-conditioning loads from sub-hourly smart meter data." Energy and Buildings 81 (2014): 316-325.

[22] Marchiori, Alan, et al. "Circuit-level load monitoring for household energy management." IEEE Pervasive Computing 10.1 (2011): 40-48.

[23] Pathak, Nilavra, Nirmalya Roy, and Animikh Biswas. "Iterative signal separation assisted energy disaggregation." Green Computing Conference and Sustainable Computing Conference (IGSC), 2015 Sixth International. IEEE, 2015.

[24] Henning Lange and Mario BergÃİs. 2016. BOLT: Energy Disaggregation by Online Binary Matrix Factorization of Current Waveforms. In Proceedings of the 3rd ACM International Conference on Systems for Energy-Efficient Built Environments (BuildSys '16). ACM, New York, NY, USA, 11-20.

[25] Batra, N., Wang, H., Singh, A., & Whitehouse, K. (2017, February). "Matrix Factorisation for Scalable Energy Breakdown". In AAAI (pp. 4467-4473).

[26] Jack Kelly and William Knottenbelt. The UK-DALE dataset, domestic appliance-level electricity demand and whole-house demand from five UK homes. Scientific Data 2, Article number:150007, 2015.

[27] Balaras, C. A., and A. A. Argiriou. "Infrared thermography for building diagnostics." Energy and buildings 34.2 (2002): 171-183.

[28] Grinzato, E., V. Vavilov, and T. Kauppinen. "Quantitative infrared thermography in buildings." Energy and Buildings 29.1 (1998): 1-9.

[29] Matthew Fox, David Coley, Steve Goodhew, Pieter De Wilde, Time-lapse thermography for building defect detection, Energy and Buildings, Volume 92, 2015, Pages 95-106.

[30] Lachut, et al. "Minimizing intrusiveness in home energy measurement." In BuildSys '12. Toronto, ON, Canada. p 56-63.