Non-Intrusive Load Monitoring (NILM) using a LSTM with socio-economic parameters

By
Ahmed Taimoor
2018-NUST-MS-CS-8
00000274526

Supervisor
Dr. Asad Waqar Malik
Department of Computer Science

A thesis submitted in partial fulfillment of the requirements for the degree of Masters of Science in Computer Science (MS CS)

In
School of Electrical Engineering and Computer Science,
National University of Sciences and Technology (NUST),
Islamabad, Pakistan.

(April 2022)

Approval

It is certified that the contents and form of the thesis entitled "Non-Intrusive Load Monitoring (NILM)" submitted by AHMED TAIMOOR have been found satisfactory for the requirement of the degree

Advisor : Dr. Asad Waqar Malik

Signature: ____________________

Date: 28-Apr-2022

Committee Member 1:Dr. Safdar Abbas Khan

Signature: ____________________

Date: 28-Apr-2022

Committee Member 2:Dr. Omar Arif

Signature: ____________________

Date: 29-Apr-2022

Committee Member 3:Dr. Arsalan Ahmad

Signature: ____________________

Date: 25-Apr-2022

THESIS ACCEPTANCE CERTIFICATE

Certified that final copy of MS/MPhil thesis entitled "Non-Intrusive Load Monitoring (NILM) written by AHMED TAIMOOR, (Registration No 00000274526), of SEECS has been vetted by the undersigned, found complete in all respects as per NUST Statutes/Regulations, is free of plagiarism, errors and mistakes and is accepted as partial fulfillment for award of MS/M Phil degree. It is further certified that necessary amendments as pointed out by GEC members of the scholar have also been incorporated in the said thesis.

Signature: ____________________________

Name of Advisor: Dr. Asad Waqar Malik

Date: 28-Apr-2022

HoD/Associate Dean: ______________________

Date: ______________________

Signature (Dean/Principal): ______________________

Date: ______________________

Dedication

I dedicate this thesis to my parents, my colleagues, my Supervisor and all the teachers who have taught me, supported me and helped me in becoming a better human being.

Certificate of Originality

I hereby declare that this submission titled "Non-Intrusive Load Monitoring (NILM)" is my own work. To the best of my knowledge it contains no materials previously published or written by another person, nor material which to a substantial extent has been accepted for the award of any degree or diploma at NUST SEECS or at any other educational institute, except where due acknowledgement has been made in the thesis. Any contribution made to the research by others, with whom I have worked at NUST SEECS or elsewhere, is explicitly acknowledged in the thesis. I also declare that the intellectual content of this thesis is the product of my own work, except for the assistance from others in the project's design and conception or in style, presentation and linguistics, which has been acknowledged. I also verified the originality of contents through plagiarism software.

Student Name: AHMED TAIMOOR

Student Signature: Ahmed Taimoor

Certificate for Plagiarism

It is certified that PhD/M.Phil/MS Thesis Titled "Non-Intrusive Load Monitoring (NILM)" by AHMED TAIMOOR has been examined by us. We undertake the follows:

a. Thesis has significant new work/knowledge as compared already published or are under consideration to be published elsewhere. No sentence, equation, diagram, table, paragraph or section has been copied verbatim from previous work unless it is placed under quotation marks and duly referenced.
b. The work presented is original and own work of the author (i.e. there is no plagiarism). No ideas, processes, results or words of others have been presented as Author own work.
c. There is no fabrication of data or results which have been compiled/analyzed.
d. There is no falsification by manipulating research materials, equipment or processes, or changing or omitting data or results such that the research is not accurately represented in the research record.
e. The thesis has been checked using TURNITIN (copy of originality report attached) and found within limits as per HEC plagiarism Policy and instructions issued from time to time.

Name & Signature of Supervisor

Dr. Asad Waqar Malik

Signature : Asad

Acknowledgment

I would like to thank The Almighty Allah for blessing me with the courage, determination and strong willpower to achieve this milestone. I would like to thank my parents for their care, support and love. I would like to thank my teachers who taught me and made me a better part of this society. I am especially grateful to my supervisor; Dr. Asad Waqar Malik and by ex supervisor Dr. Fahad Javed for their guidance, support, faith and for giving me an opportunity to prove my skills. In the end, I would like to thank SEECS (School of Electrical Engineering and Computer Sciences) for helping me along the way and for giving me an opportunity to attend the best engineering institute in Pakistan.

Table of Contents

List of Figures

List of Tables

Abstract

Nonintrusive load monitoring (NILM) deconstructs aggregated electrical usage data into individual appliances. The dissemination of disaggregated data to customers raises consumer awareness and encourages them to save power, lowering CO_2 emissions to the environment. The performance of NILM systems has increased dramatically thanks to recent disaggregation methods. However, the capacity of these algorithms to generalize to various dwellings as well as the disaggregation of multi-state appliances remain significant obstacles. In this paper we propose an energy disaggregation approach by using socio-economic parameters with the aggregated data. The suggested approach helps in creating more accurate load profiles, which improves the accuracy and helps in better detection of the appliances. The proposed model outperforms state-of-the-art NILM techniques on the PRECON dataset. The mean absolute error reduces by percentage 5%- 10% on average across all appliances compared to the state-of-the-art. Thus, improving the detection of target appliance in the aggregate measurement.

Chapter 1

Introduction

1.1 Background and motivation

NILM, or load dis aggregation, is a technique for determining the operational status (on/off) and accurate power consumption profile of the individual electrical loads using just the aggregated consumption as input. Due to advances in the fields of deep learning and machine learning algorithms, this notion was initially suggested by Hart in 1992 [1], but it has been refined substantially over the previous decade. Many dis aggregation methods are utilized in the residential sector, however, there have been research on industrial [2], [3], As a non-intrusive mechanism, the strategies apply least amount of intrusion and have minimal impact on consumer privacy. Because the measurements are taken from the single location (aggregated load demand), the need to deploy additional equipment is eliminated, which would increase the installation's complexity and expense.

The provision of the access to the specific data of the appliance instead of the whole-house measurements has a variety of advantages for both house owners and energy companies. Consumers, for example, may better understand their energy use since they can see which appliances use the most energy. As a result, individuals will be better prepared to make energy-related decisions. The majority of customers are unaware of how much energy they use or their environmental effect. Increased awareness may result in more reasonable appliance usage. Consumers may choose to use less of their high-energy-consumption equipment and, in some situations, to replace inefficient appliances or using CO_2 emitting appliances more efficiently.

1.2 Problem Statement

According to the study "Net Zero by 2050" of International Energy Agency's (IEA), the energy sector is responsible for the major portion of the global greenhouse gas emissions [4] . Furthermore, as global energy demand grows faster than supply, existing power grid systems face major problems in terms of efficiency and dependability. The restructuring of electrical infrastructure is crucial on the way to zero CO_2 emissions by 2050. The combination of increased computing power and unique modelling and simulation capabilities allows for a seamless transformation from the conventional grids to the era of the smart grid [5]. The main objective of the smart grids is to integrate assets such as Distributed Energy Resources (DERs), hybrid electric vehicles (HEVs), and Energy Storage Systems as well as smart and efficient services, to unlock the flexibility potential that will allow for more efficient energy generation, distribution, and consumption. The use of Non-intrusive load monitoring (NILM) helps in achieving this goal by evaluating the consumption of individual appliances in a facility.

1.3 Prior State of the Art

For many years, Deep Neural Networks (DNN) techniques have been used to dis aggregate energy [6], [7]. These state-based methods are mostly utilized for low-frequency (less than 1 Hz) monitoring, which needs less expensive gear. Recurrent neural networks [6]- based methods, such as LSTM , [8], [7], or Gated Recurrent Unit [9], have mostly been proposed because they are ideally suited for 1D time series data. The authors of [10] present a Bayesian optimized bidirectional LSTM model for NILM that extends the RNN.

High-frequency data has more load features than low-frequency data. The harmonic approach was improved in [11] and [12]. [11] employed total harmonic distortion rate, power, and current harmonics as characteristics to construct similarity scores to achieve load detection. [12] proposed a method based on lower odd-numbered harmonics and bagging decision tree (FFT BDT), which included two processes: obtaining magnitude and phase at lower odd-numbered harmonics and recognizing loads using a bagging decision tree. For NILM, a voltage–current (V–I) image-based technique has been developed. [13]. The reconstructed picture of a V-I trajectory was employed as input data for a convolutions neural network (CNN) to categories appliances, especially resistive appliances, in the study. when compared to the other two approaches on the PLAID and IDOUC data sets, the proposed approach performs extremely well. [14] suggested a non-intrusive load detection system based on a two-stream convolutions neural network (CNN) with current time frequency feature fusion. To gather the frequency and time domain characteristics, a time series image coding approach was devised first. The load detection performance was then improved by using a dual-stream neural network integrating a gated recurrent unit (GRU) and a 2D-CNN. Finally, PLAID and IDOUC data sets were used to test it.

Traditional CNN's extract characteristics by just feeding data in one way. These networks are unable to capture data that varies over time, such as time-series data. Recurrent neural networks (RNN) and LSTM [15] were offered as solutions to this problem. The LSTM, which records time-series patterns through two states in each cell, is the most widely used recurrent model today.

1.4 Proposed Method

A typical system-level NILM setup refers to one time calibration period to learn appliance signatures and store them in database. Once the system learns these signatures it can identify the appliances based on those signatures when ever the switching event takes place [16]. These typical systems only use appliance signatures to train the model. In our approach we use socioeconomic parameters with the appliance signatures so this NILM setup can correctly identify those appliances based on those parameters, which can be seen in the results section that this NILM setup reduces the loss errors for correct appliance detection.

In our methodology, we focus on different scenarios where test houses are held out during the training process, for each scenario we used the PRECON data. In the first scenario we held out 5 houses and trained the modal on another house without using socio-economic parameters and calculated the average mean absolute errors. In the second scenario we again held the same 5 houses and trained the modal on the same other house but this time we used the socio-economic parameters by attaching them before the training so our modal could train based on these weights. In another scenario we also used 20% of the data of each house for training the modal and then validating using rest of the data, we named it "results on seen data and unseen data".

1.5 Contributions

The main contribution of this thesis includes the design of an approach to implement NILM in which socio-economic parameters are induced with the

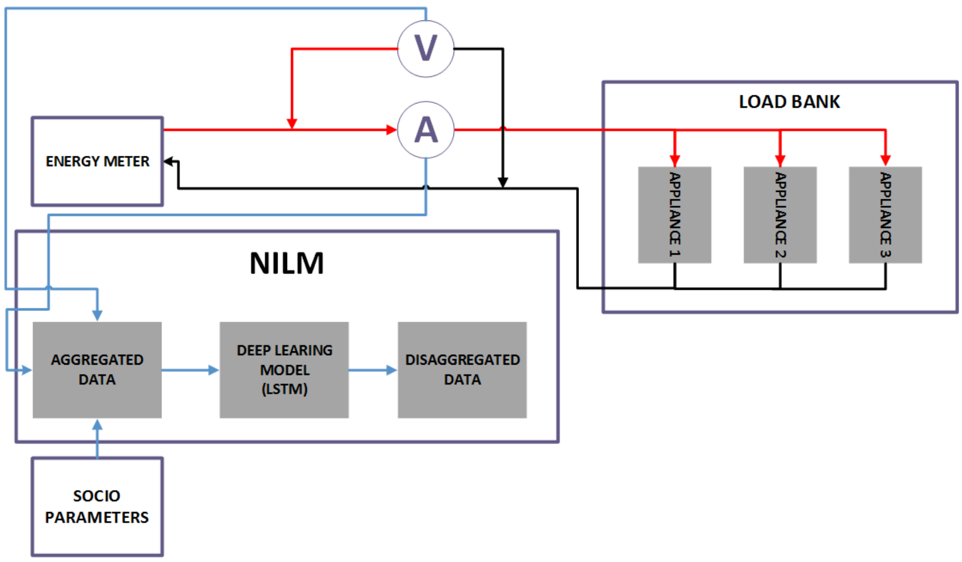

Figure 1.1: General Block Diagram of PHEV

data to make dis-aggregated load profile of each appliance more accurate, efficient and reliable. Moreover, to the best of my knowledge the proposed approach for NILM discussed in this thesis is new and it had not been covered in the literature before.

The key impacts of this research study are discussed as follows:

1. Induction of socio-economic parameters in the system of NILM.

2. Generation of data set by combining load profiles data and socio-economic parameters.

3. Generation of the predictions against the merged data set for load profile of each appliance.

1.6 Thesis Layout

In the Layout of this thesis, **chapter 2** includes the study of work done in the field of NILM techniques for their efficient working. **Chapter 3** covers the methodology and working of our NILM setup. In **chapter 4**, we will discuss our experiments which are done on the PRECON data set with and without socio-economic parameters. **Chapter 5** concludes the contributions and possible advancements in the future of this field of NILM techniques.

Chapter 2

Literature Work

2.1 Classification of Appliance Load monitoring systems

The challenge of optimization of energy consumption in residential and commercial buildings is becoming popular nowadays. According to a survey in the European Union, it has been noticed that energy consumption in the residential sector alone is above 30% and this will be doubled by the end of 2030 [17]. So, to optimize the use of energy in both residential and commercial buildings, the development of Appliance Load Monitoring (ALM) algorithms has been considered [18]. The purpose served by ALM is to perform detailed energy-sensing through IoT sensors and generate SOPs of energy consumption based on data received. The ALM is classified into the following two categories:

- Intrusive Load Monitoring (ILM)
- Non-Intrusive Load Monitoring (NILM)

2.2 Study of NILM methods

Several NILM methods are discussed in the literature, but they have limitations of not being efficient to control energy consumption in residential and commercial buildings. These techniques totally rely on the data obtained from the IoT sensors placed according to the requirements of the system used. Upcoming topics contain detailed discussions of these methods and their limitations.

2.2.1 Decision Tree based approach

In the article [19], the approach of the load decision tree algorithm is used for the analysis and identification of the appliance composition of the mixed electrical appliance group. Moreover, the establishment of the 0-1 programming model for the appliance status verification is carried out and the algorithm of Particle Swarm Optimization (PSO) is utilized to carry out the process of state identification of the appliance in the group of electrical appliances. The dataset used in this article is a customized dataset and the results obtained against this dataset are shown in the figure below.

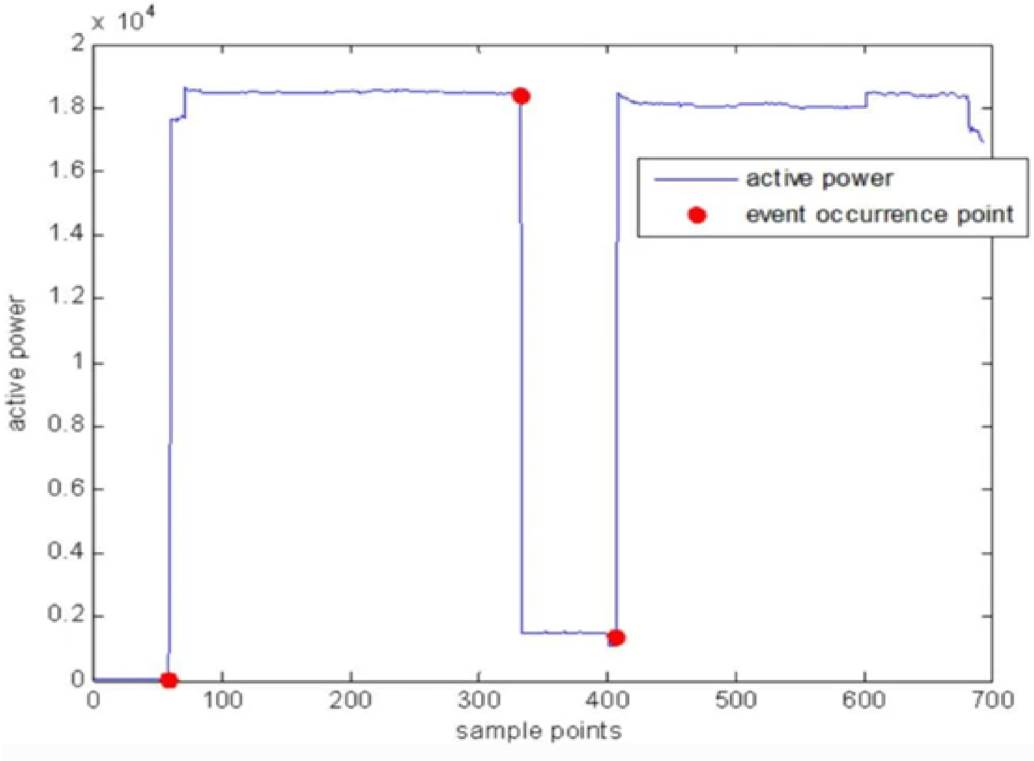

Figure 2.1: NILM system using decision tree

In the fig.2.1, red dots represent the events of the appliance. This decision tree based algorithm for NILM has higher accuracy, higher anti-interference and stringer identification ability.

2.2.2 Real time monitoring based approach

The article [20] emphasis on the implementation of an NILM algorithm that detect the change in state of target appliance, processes the transient response of the appliance and predict its power consumption profile in real time. this system is consist of a simple event detection algorithm, a convolutional neural network (CNN) classifier and a power estimation algorithm. the dataset used in this article to train the CNN model is BLUED dataset. Fig.2.2 shows the results obtained after using this approach for testing on single household appliance.

2.2.3 V-I trajectories with CNN based approach

In the article [21], the use of pixelated image of voltage and current trajectories is carried out as an input for the CNN which on the basis deep learning techniques classify the appliance used through the extraction of its key features. the datasets used in this article to carry out the deep learning process are PLAID and WHITED.

2.3 Conclusion of the literature review

The detailed review of the literature discussed in the above sections concludes that the discussed approaches in the literature, used to implement NILM for the optimization of power consumption profile of resdential and comercial buildings, totally rely on the use of datasets produced from the data coming from the IoT sensors (especially voltage and current sensors). This will make the predictions, obtained from the the process of disaggregation, inaccurate. To cater this problem, the integration of socio-economics parameters with the IoT sensors data is proposed. This will helps the process of disaggregation to make predictions, about the power consumption profile of individual appliance, more accurate and more efficient.

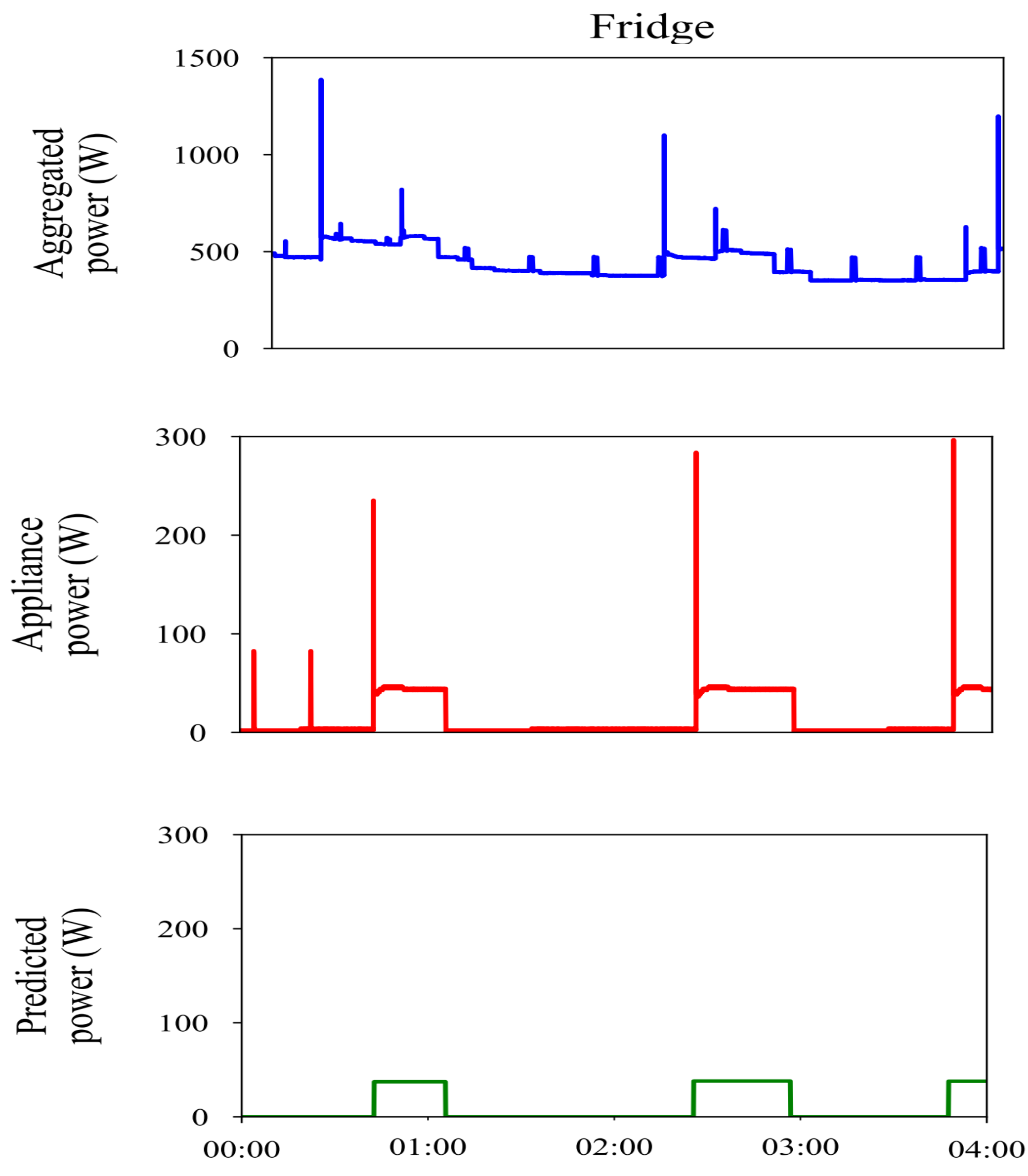

Figure 2.2: Power estimation for the selected appliances in real households. Time-series of (a) aggregated power, (b) actual target appliance power, (c) estimated power for fridge;

Chapter 3

Designing of NILM system

This section includes the process of designing the proposed NILM system. Fig. 1.1 represents the proposed NILM system which is divided into three parts: Data Acquisition, Deep learning model, and Train and Test sets. The detailed analysis of these different parts of our system is as follows:

3.1 Data Acquisition

The PRECON data set is used for our experiments, this reference data set is a set of 42 houses. Each house data was collected over the course of one year using smart meters. Aside from power usage statistics, various home attributes are also recorded in this data set. Some of these attributes are summarized as. Total number of people living in the house, their age groups, property area of the house etc. Other than this, all the other important electrical loads of the house are also recorded which include the number of LED lights, fans, washing machines, electric iron, tube-lights, electric heaters, water pumps, refrigerators, electronic devices, and water dispensers etc. Other typical data sets also categorized into following types

- **Type 1:** The devices that operate only on ON-OFF states. Examples are the lamp, toaster, etc.
- **Type 2:** The devices have more than two states of operation. These devices are also called Finite State Machine (FSM). Examples of this category of devices are washing machines, stove burners, etc.
- **Type 3:** The devices having characteristics of variable power consumption without fixing the number of states. These devices are also called Continuously Variable Devices (CVD). Examples of such devices are power drills and dimmer lights.

- **Type 4:** The devices that remain active throughout weeks or days and consume power at a constant rate. These devices are also classified as permanent consumer devices. Examples of such devices are telephone sets, smoke detectors, wifi routers etc.

3.2 Deep learning model

In our methodology, we focus on different scenarios where test houses are held out during the training process, for each scenario we used the PRECON data. In the first scenario we held out 5 houses and trained the modal on another house without using socio-economic parameters and calculated the average mean absolute errors. In the second scenario we again held the same 5 houses and trained the modal on the same other house but this time we used the socio-economic parameters by attaching them before the training so our modal could train based on these weights. In another scenario we also used 20% of the data of each house for training the modal and then validating using rest of the data, we named it “results on seen data and unseen data”.

We trained the proposed model through supervised learning. We used the optimizer 'Adam' with loss function 'Root Mean Square' with a dropout rate initialized at 0.3. The model contains six groups of layers with dimensions input, 64, 128, and 256 respectively with a dense layer attached at the end with this stacked LSTM architecture as showed in the fig 3.1 and Table 3.1.

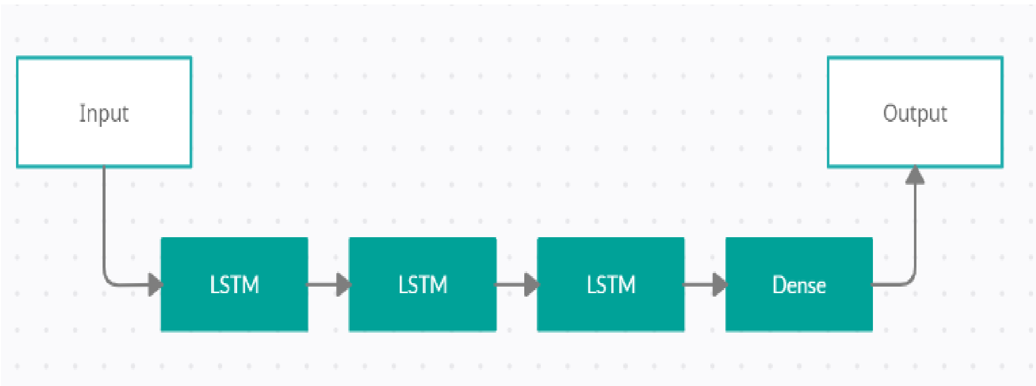

Figure 3.1: Stacked LSTM Architecture

Regardless of the appliance type all the hyper parameters are fixed and all the experiments are run on a maximum of 100 epochs using the batch size 60.

Table 3.1: NUMBER OF HIDDEN NEURONS AND OUTPUT SHAPE OF LSTM BLOCK

Name	Output Shape	Number of Parameter
LSTM layer	64	22,528
LSTM layer	128	98,816
LSTM layer	256	394,240
Dense	1	257

3.3 Training Set and Test Set

The house 4 in the PRECON data set is used as the train set. It contains Total Usage (Usage_kW), Air Conditioner in bedroom (AC_BR_kW), kitchen (Kitchen_kW), Air Conditioner in lounge room (AC_LR_kW), Air Conditioner in master bedroom (AC_MBR_kW) as shown in Table 3.2.

Table 3.2: OVERVIEW OF HOUSE 4 in PRECON DATA SET

Date and Time	Usage	AC-DR	Kitchen	...	AC-BR
2019-03-10 11:05:00	1.3084	0.008	0.5134	...	0.0268
2019-03-10 11:06:00	1.4813	0.009	0.5088	...	0.0268
....	....	....	....	...	....
2019-03-10 11:07:00	1.7421	0.008	0.5021	...	0.0270

With this data a metadata file is given which contains other attributes of the houses. The other attributes of house 4 are shown in Table 3.3

Table 3.3: METADATA OVERVIEW OF HOUSE 4 in PRECON DATA SET

Attributes	Value
Property Area sqft	5445.01
Number of people living	7
Total number of rooms	7
Number of electric heaters	0
Number of UPS	2
Number of fans	10
Number of refrigerators	3
...	...
Number of water pumps	1

This table shows metadata of house 4 in the precinct data set, it includes total number of people living in the house also differentiating between total number of children and adults, with this total number of appliances inside the house are given such as total number rod fans, total number of lights and total number refrigerators etc, geometric and other information has also been given such as property area in square feet of the house, if insulation is installed in the house and ceiling height etc.

Chapter 4

Experiments and Numerical Results

In this section, we conducted extensive experiments and made comparisons. We conducted experiments on two appliances, the kitchen (it includes aggregated readings of all the devices operating in the kitchen) and the air conditioner, of five different houses. After performing these experiments, we compared the results generated by using aggregated data that includes power consumption readings and socio-economic parameters with the results obtained by using data without socio-economic parameters. The model used in these experiments was initially trained on the individual power consumption profile of the appliances under consideration against the aggregated power consumption profile of the house. Then this model is trained by using a combination of aggregated power consumption profile and socio-economic parameters of the house against every appliance's power consumption profile. The upcoming sections will elaborate on the results obtained after performing these experiments.

4.1 Without socio-economic parameters

Table 4.1 shows our results obtained after testing the model, that was trained against the aggregated data without socio-economic parameters, by using aggregated power consumption profile of the house. The aggregated power consumption profile of house 1 is used to train the model while the aggregated power consumption profiles of other houses were used to test the model. Table 4.2 shows the mean absolute error for the same trained model. This is the approach used by the current typical NILM systems. We used these results obtained from our trained model were used as the benchmark against the other models that were trained using the combination of aggregated power consumption profile and the socio-economic parameters to see and generalize how these parameters effects the correct detection of the appliances.

Table 4.1: MEAN SQUARE ERROR WITHOUT SOCIO-ECONOMIC PARAMETERS

House No.	Kitchen	Air conditioner
House 1	0.0373	0.447
House 2	0.0171	0.1198
House 3	0.0088	0.1185
House 4	0.0274	0.2303
House 5	0.0322	0.0774

Table 4.2: MEAN ABSOLUTE ERROR WITHOUT SOCIO-ECONOMIC PARAMETERS

House No.	Kitchen	Air conditioner
House 1	0.1636	0.1143
House 2	0.0947	0.1259
House 3	0.0562	0.2547
House 4	0.0931	0.2360
House 5	0.1170	0.1392

4.1.1 Visualisations

Given below are the figures of real and predicted values for kitchen of different houses. The blue line shows the actual reading of the appliance and the red line shows the predicted reading of the house against the trained modal.

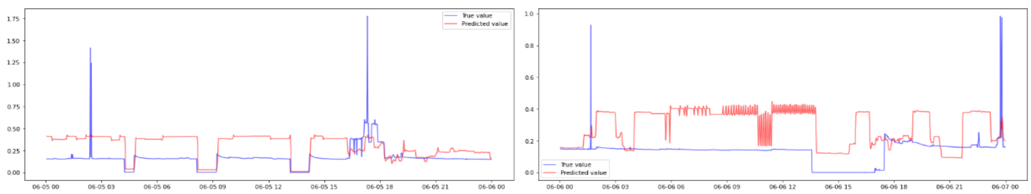

Figure 4.1: Real and Predicted values of house 2 for kitchen

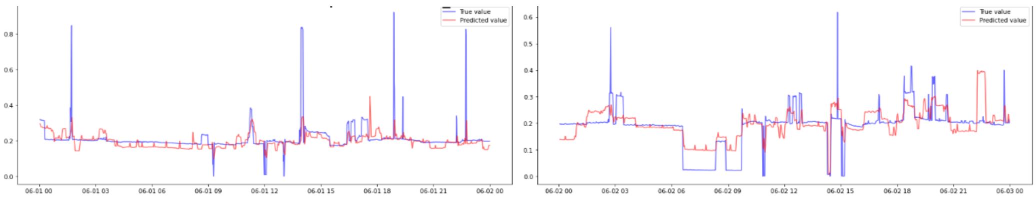

Figure 4.2: Real and Predicted values of house 3 for kitchen

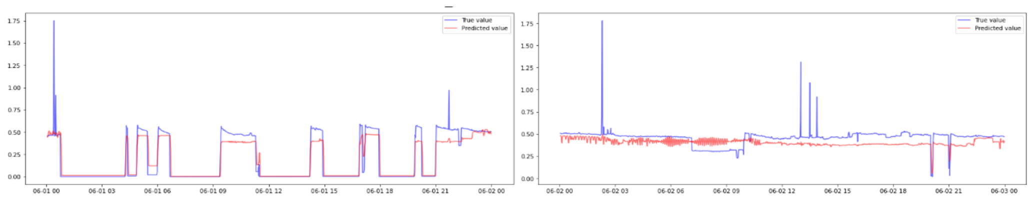

Figure 4.3: Real and Predicted values of house 4 for kitchen

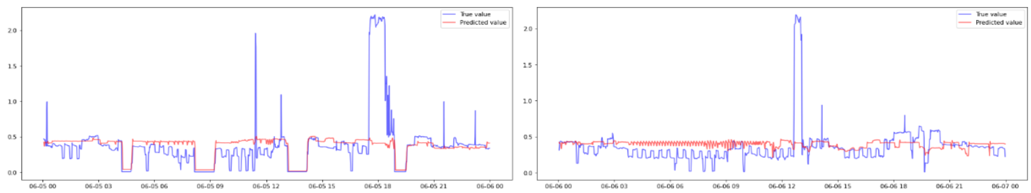

Figure 4.4: Real and Predicted values of house 5 for kitchen

4.2 With socio-economic unseen parameters

We trained the model using socio-economic parameters with the total usage of the house and we used those houses as test sets that have aggregated data completely unknown to the trained model. The model gave better results if the model was trained on some of the same parameters of the same house. Table 4.3 and Table 4.4 shows our mean square error and mean absolute errors with socio-parameters. The model is trained on House 1 total aggregated data and its socio-economic parameters, the rest of the 4 houses are used as test sets. The decrease in mean and absolute square errors can be seen in the results. Houses 1,2,3 showed better results with socio-economic parameters as for house 4,5 the mean square errors increased. This happened because the modal didn't go through with these unknown parameters and had no idea how to predict the test data correctly.

Table 4.3: MEAN SQUARE ERROR WITH SOCIO-ECONOMIC PARAMETERS (UNSEEN DATA)

House No.	Kitchen	Air conditioner
House 1	0.0309	0.0435
House 2	0.2531	0.1673
House 3	0.0600	0.0884
House 4	0.0382	0.2867
House 5	0.2720	0.0864

Table 4.4: MEAN ABSOLUTE ERROR WITH SOCIO-ECONOMIC PARAMETERS (UNSEEN DATA)

House No.	Kitchen	Air conditioner
House 1	0.1441	0.1174
House 2	0.4750	0.3440
House 3	0.2244	0.2726
House 4	0.1312	0.3242
House 5	0.4491	0.2383

4.2.1 Visualisations

Given below are the figures of real and predicted values for kitchen of different houses. The blue line shows the actual reading of the appliance and the red line shows the predicted reading of the house against the trained modal.

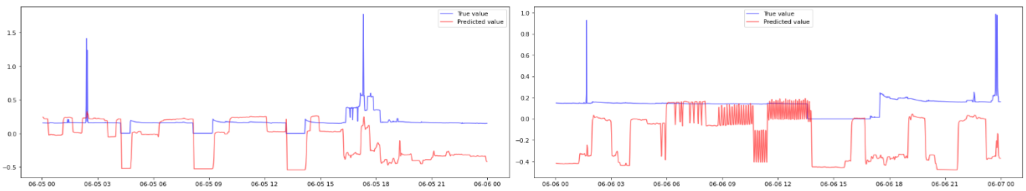

Figure 4.5: Real and Predicted values of house 2 for kitchen

4.3 With socio-economic seen parameters

In this scenario, we created a new train set which is trained on 20% of the aggregated data of each 5 house and the socio-economic parameters. This model gave the best results the mean square error and mean absolute error were less than both previous models. Table 4.5 and Table 4.6 show the mean

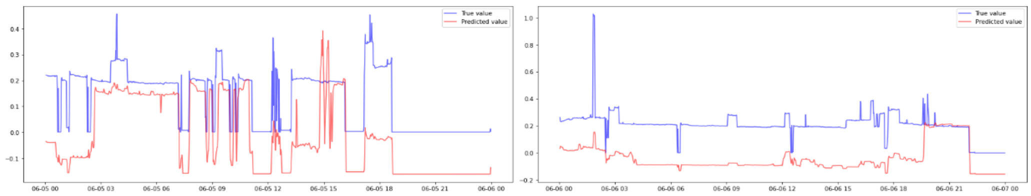

Figure 4.6: Real and Predicted values of house 3 for kitchen

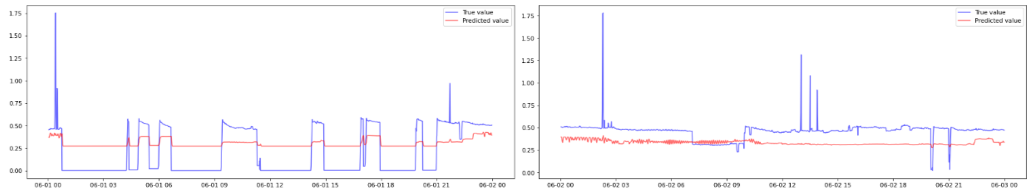

Figure 4.7: Real and Predicted values of house 4 for kitchen

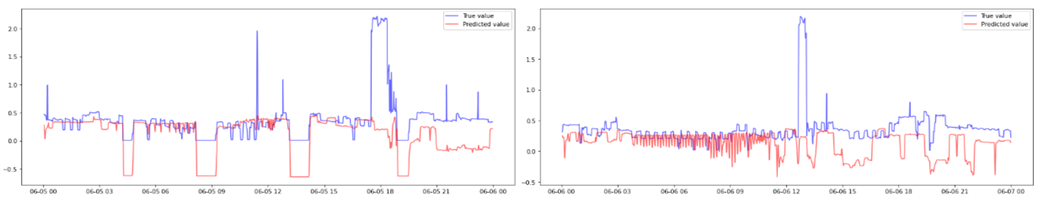

Figure 4.8: Real and Predicted values of house 5 for kitchen

square error and mean absolute error of the model which is trained using socio-economic parameters and where 20% of the data is used in the training set on each house. In short, if the modal has seen or has been trained on some of the readings of the house and the parameters it can correctly identify the appliances better than the modals trained without these parameters. It can be seen in the tables all 5 houses mean square error and mean absolute errors reduced drastically as compared to the previous models.

Table 4.5: MEAN SQUARE ERROR WITH SOCIO-ECONOMIC PARAMETERS (SEEN DATA)

House No.	**Kitchen**	**Air conditioner**
House 1	0.0335	0.0435
House 2	0.0078	0.0020
House 3	0.0018	0.0662
House 4	0.0354	0.0079
House 5	0.0372	0.0010

Table 4.6: MEAN ABSOLUTE ERROR WITH SOCIO-ECONOMIC PARAMETERS (SEEN DATA)

House No.	**Kitchen**	**Air conditioner**
House 1	0.1517	0.1174
House 2	0.0600	0.0205
House 3	0.0183	0.1879
House 4	0.1149	0.0278
House 5	0.1335	0.0174

4.3.1 Visualisations

Given below are the figures of real and predicted values for kitchen of different houses. The blue line shows the actual reading of the appliance and the red line shows the predicted reading of the house against the trained modal.

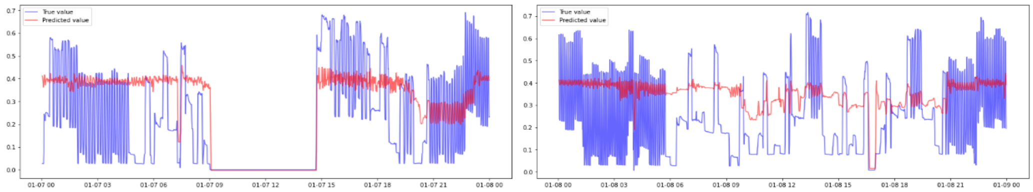

Figure 4.9: Real and Predicted values of house 1 for kitchen

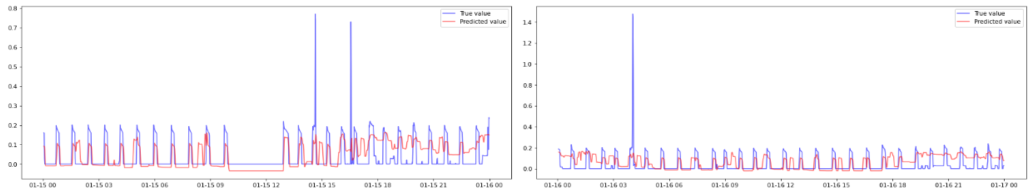

Figure 4.10: Real and Predicted values of house 2 for kitchen

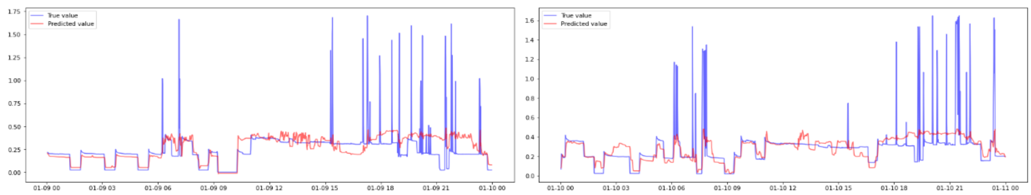

Figure 4.11: Real and Predicted values of house 4 for kitchen

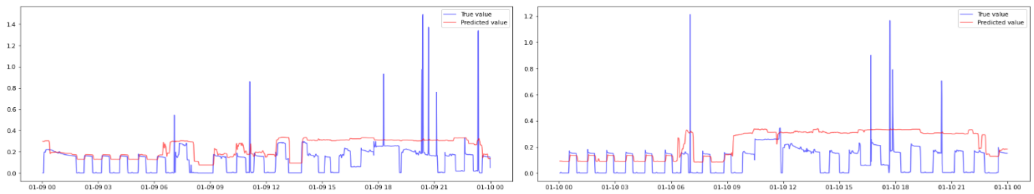

Figure 4.12: Real and Predicted values of house 5 for kitchen

4.4 COMPARISONS

In this section, we will compare our results of our models which were trained with and without socioeconomic parameters.

Table 4.7 shows mean square error of all three models for kitchen data set. It can be seen from the results that the model trained by using the socioeconomic parameters reduced the mean square error and improved the model in prediction of the appliance. House 1,2 and 3 shows the effectiveness of socioeconomic parameters on the Model trained with these parameters. The loss relatively decreased as compared to the model which was trained without these parameters.

Table 4.7: COMPARISON OF MEAN SQUARE ERROR OF ALL THREE MODELS ON KITCHEN

House No.	**Without parameters**	**With parameters (unseen data)**	**With parameters (seen data)**
House 1	0.0373	0.0309	0.0335
House 2	0.0171	0.2531	0.0078
House 3	0.0088	0.0600	0.0018
House 4	0.0274	0.0382	0.0354
House 5	0.0322	0.2720	0.0372

Table 4.8 shows mean square errors on air conditioner appliance. The model trained using the socio-economic parameters the mean square and mean absolute errors reduced drastically for all the five houses showing the effectiveness of using socioeconomic parameters.

4.4.1 Visualization

.

Table 4.8: COMPARISON OF MEAN SQUARE ERROR OF ALL THREE MODELS ON AIR CONDITIONER

House No.	**Without parameters**	**With parameters (unseen data)**	**With parameters (seen data)**
House 1	0.0447	0.0435	0.0435
House 2	0.1198	0.1673	0.0020
House 3	0.1185	0.0884	0.0662
House 4	0.2303	0.2867	0.0079
House 5	0.0774	0.0864	0.0010

Fig. 4.13 and Fig.4.14 show some graph visualization of true value and predicted value by the system of houses 2 and 4. The blue line shows the actual values of the appliance, and the red line shows the predicted values. The graphs show the data of two random days from the whole year.

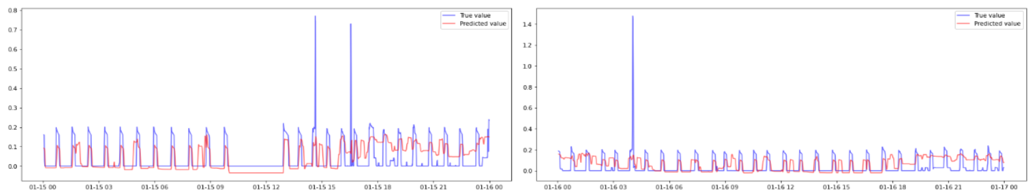

Figure 4.13: Real and Predicted values of house 2 for kitchen

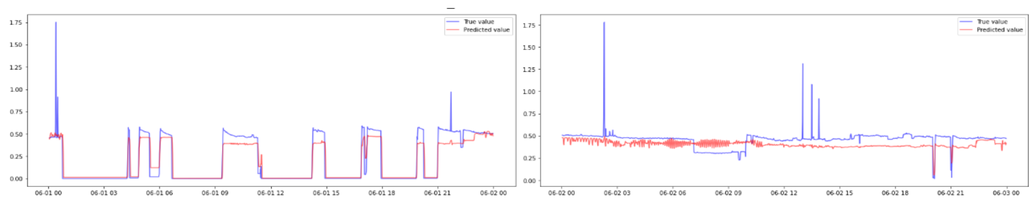

Figure 4.14: Real and Predicted values of house 4 for kitchen

Chapter 5

Conclusion and Future Work

5.1 Conclusion

The idea of NILM appears to have a prominent position as a future smart energy grid service, allowing users to gain control over their energy usage through enhanced awareness. This will help the users to better monitor and use their appliances more efficiently and effectively. This breakdown of energy use at the appliance level might also aid in the detection of abnormalities in equipment that are malfunctioning. The NILM techniques for the appliance desegregation also helps in addressing some serious challenges including inference, unsupervised learning, generation of generalized appliance models, and getting a better idea of the trends of the user power consumption.

This paper has proposed a non-intrusive load monitoring technique using socio-economic parameters on a stacked LSTM Algorithm and verified its effectiveness through PRECON data set. Compared with the traditional techniques the mentioned approach of using socioeconomic parameters with the house data reduces the losses of LSTM algorithm and showed better results in the detection of the appliances.

5.2 Future Work

In future work we would like to see the effectiveness of socio-economic parameters on other algorithms other than LSTM such as decision trees etc. We would also like to design a multi-headed network for this technique using time series data separately and socio-economic parameters as separate model and then concatenating both the outputs in the last layer. We believe this model can give even better results than our current dis aggregation solutions. We would also like to see how the environmental parameters such as weather conditions 'summer,winter' effects the usage of the electricity of the house and how these parameters can be used with NILM techniques or if they can also help in the detection of the appliances.

Bibliography

[1] G. W. Hart, "Nonintrusive appliance load monitoring," *Proceedings of the IEEE*, vol. 80, no. 12, pp. 1870–1891, 1992.

[2] E. Holmegaard and M. B. Kjærgaard, "Nilm in an industrial setting: A load characterization and algorithm evaluation," in *2016 IEEE International Conference on Smart Computing (SMARTCOMP)*. IEEE, 2016, pp. 1–8.

[3] P. B. Martins, J. G. Gomes, V. B. Nascimento, and A. R. de Freitas, "Application of a deep learning generative model to load disaggregation for industrial machinery power consumption monitoring," in *2018 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)*. IEEE, 2018, pp. 1–6.

[4] F. Kalinke, P. Bielski, S. Singh, E. Fouché, and K. Böhm, "An evaluation of nilm approaches on industrial energy-consumption data," in *Proceedings of the Twelfth ACM International Conference on Future Energy Systems*, 2021, pp. 239–243.

[5] R. W. Cox, P. L. Bennett, T. D. McKay, J. Paris, and S. B. Leeb, "Using the non-intrusive load monitor for shipboard supervisory control," in *2007 IEEE Electric Ship Technologies Symposium*. IEEE, 2007, pp. 523–530.

[6] J. Paris, Z. Remscrim, K. P. Douglas, S. B. Leeb, R. W. Cox, S. T. Galvin, S. G. Coe, J. R. Haag, and J. A. Goshorn, "Scalability of non-intrusive load monitoring for shipboard applications," 2009.

[7] IEA, "Net zero by 2050," https://www.iea.org/reports/net-zero-by-2050, 2021.

[8] K. Aurangzeb, S. Aslam, S. M. Mohsin, and M. Alhussein, "A fair pricing mechanism in smart grids for low energy consumption users," *IEEE Access*, vol. 9, pp. 22 035–22 044, 2021.

[9] J. Kelly and W. Knottenbelt, "Neural nilm: Deep neural networks applied to energy disaggregation," in *Proceedings of the 2nd ACM international conference on embedded systems for energy-efficient built environments*, 2015, pp. 55–64.

[10] L. Mauch and B. Yang, "A new approach for supervised power disaggregation by using a deep recurrent lstm network," in *2015 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*. IEEE, 2015, pp. 63–67.

[11] J. Kim, T.-T.-H. Le, and H. Kim, "Nonintrusive load monitoring based on advanced deep learning and novel signature," *Computational intelligence and neuroscience*, vol. 2017, 2017.

[12] O. Krystalakos, C. Nalmpantis, and D. Vrakas, "Sliding window approach for online energy disaggregation using artificial neural networks," in *Proceedings of the 10th Hellenic Conference on Artificial Intelligence*, 2018, pp. 1–6.

[13] M. Kaselimi, N. Doulamis, A. Doulamis, A. Voulodimos, and E. Protopapadakis, "Bayesian-optimized bidirectional lstm regression model for non-intrusive load monitoring," in *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2019, pp. 2747–2751.

[14] R. Doğan and E. Akarslan, "Investigation of electrical characteristics of residential light bulbs in load modelling studies with novel similarity score method," *IET Generation, Transmission & Distribution*, vol. 14, no. 23, pp. 5364–5371, 2020.

[15] D. Jia, Y. Li, Z. Du, J. Xu, and B. Yin, "Non-intrusive load identification using reconstructed voltage–current images," *IEEE Access*, vol. 9, pp. 77 349–77 358, 2021.

[16] K. Li, B. Yin, Z. Du, and Y. Sun, "A nonintrusive load identification model based on time-frequency features fusion," *IEEE Access*, vol. 9, pp. 1376–1387, 2020.

[17] L. D. S. J I Utley, "Domestic energy fact file 2008," https://www.bre.co.uk/filelibrary/pdf/rpts/Fact_File_2008.pdf, 2008, accessed 2008.

[18] A. Zoha, A. Gluhak, M. A. Imran, and S. Rajasegarar, "Non-intrusive load monitoring approaches for disaggregated energy sensing: A survey," *Sensors*, vol. 12, no. 12, pp. 16 838–16 866, 2012.

[19] J. Lin, X. Ding, D. Qu, and H. Li, "Non-intrusive load monitoring and decomposition method based on decision tree," *Journal of Mathematics in Industry*, vol. 10, no. 1, pp. 1–14, 2020.

[20] C. Athanasiadis, D. Doukas, T. Papadopoulos, and A. Chrysopoulos, "A scalable real-time non-intrusive load monitoring system for the estimation of household appliance power consumption," *Energies*, vol. 14, no. 3, p. 767, 2021.

[21] L. De Baets, J. Ruyssinck, C. Develder, T. Dhaene, and D. Deschrijver, "Appliance classification using vi trajectories and convolutional neural networks," *Energy and Buildings*, vol. 158, pp. 32–36, 2018.