Cairo University

A NEW TECHNIQUE COMBINING SEMI-SUPERVISED AND ACTIVE LEARNING FOR NON-INTRUSIVE LOAD MONITORING

By

Ahmed Mohamed Fatouh Ahmed

A Thesis Submitted to the
Faculty of Engineering at Cairo University
in Partial Fulfillment of the
Requirements for the Degree of
MASTER OF SCIENCE
In
ELECTRONICS AND COMMUNICATIONS ENGINEERING

FACULTY OF ENGINEERING, CAIRO UNIVERSITY
GIZA, EGYPT
2019

A NEW TECHNIQUE COMBINING SEMI-SUPERVISED AND ACTIVE LEARNING FOR NON-INTRUSIVE LOAD MONITORING

By

Ahmed Mohamed Fatouh Ahmed

A Thesis Submitted to the
Faculty of Engineering at Cairo University
in Partial Fulfillment of the
Requirements for the Degree of
MASTER OF SCIENCE
In
ELECTRONICS AND COMMUNICATIONS ENGINEERING

Under the Supervision of
Omar Ahmed Ali Nasr
Assoc. Professor
Electronics and Communications Engineering Department
Faculty of Engineering, Cairo University

FACULTY OF ENGINEERING, CAIRO UNIVERSITY
GIZA, EGYPT
2019

A NEW TECHNIQUE COMBINING SEMI-SUPERVISED AND ACTIVE LEARNING FOR NON-INTRUSIVE LOAD MONITORING

By

Ahmed Mohamed Fatouh Ahmed

A Thesis Submitted to the
Faculty of Engineering at Cairo University
in Partial Fulfillment of the
Requirements for the Degree of
MASTER OF SCIENCE
In
ELECTRONICS AND COMMUNICATIONS ENGINEERING

Approved by the Examining Committee:

Assoc. Prof. Omar Ahmed Ali Nasr (Thesis Main Advisor)
Faculty of Engineering, Cairo University

Prof. Mohsen Abd El-Razek Rashwan (Internal Examiner)
Faculty of Engineering, Cairo University

Prof. Moustafa Mohamed Mohamed Eissa (External Examiner)
Faculty of Engineering, Helwan University

FACULTY OF ENGINEERING, CAIRO UNIVERSITY
GIZA, EGYPT
2019

Engineer : **Ahmed Mohamed Fatouh Ahmed**
Date of Birth : 27/5/1991
Nationality : Egyptian
E-mail : ahmed.fatouh100@gmail.com
Phone : +201111687493
Address : 9th Mubarak St., Ain Shams, Cairo
Registration Date : 1/3/2014
Awarding Date : / /2019
Degree : **Master of Science**

Department : **Electronics and Communications Department**

Supervisor : **Assoc. Prof. Omar Ahmed Ali Nasr**

Examiners :

Assoc. Prof. Omar Ahmed Ali Nasr (Thesis Main Advisor)
Faculty of Engineering, Cairo University

Prof. Mohsen Abd El-Razek Rashwan (Internal Examiner)
Faculty of Engineering, Cairo University

Prof. Moustafa Mohamed Mohamed Eissa (External Examiner)
Faculty of Engineering, Helwan University

Title of Thesis : A New Technique Combining Semi-supervised and Active Learning for Non-Intrusive Load Monitoring
Keywords : Non-Intrusive load monitoring (NILM), Semi-supervised learning, Active learning, Event classification

Summary:

The current work introduces a new technique that leverages both the semi-supervised and active learning together to the benefit of non-intrusive load monitoring (NILM), which is the procedure used to disaggregate the contributions of different appliances in a building. The main idea is that semi-supervised learning improves the results of active learning aiming to decrease the need to the user. Two different approaches were utilized, one used active and reactive power features and the other used current waveform harmonics to use them later in the machine learning model.

Disclaimer

I hereby declare that this thesis is my own original work and that no part of it has been submitted for a degree qualification at any other university or institute.

I further declare that I have appropriately acknowledged all sources used and have cited them in the references section.

Name: Date:

Signature:

Acknowledgements

I gratefully acknowledge the effort of my supervisor, Assoc. Prof. Omar A. Nasr, for his guidance and support along the journey of the thesis.

I would like to thank my parents for their patience as I had many troubles, when they supported me till the end.

I would like to thank anyone who had a positive impact on my life until this work is completed.

Table of Contents

List of Tables

List of Figures

Nomenclature

Latin letter symbols

C	Number of samples added per iteration	[--]
FN	False Negatives	[--]
FP	False Positives	[--]
H	Entropy	[--]
l	Maximum number of queries allowed	[--]
L	Labeled Dataset	[--]
P	Active Power	W
p	Probability	[--]
Q	Reactive Power	VAR
s	sample	[--]
t	Interval between different user queries	[--]
TN	True Negatives	[--]
TP	True Positives	[--]
U	Unlabeled dataset	[--]

Greek letter symbols

β	Weight for F-score measure

Abbreviations

BLUED	Building-Level fUlly-labeled dataset for Electricity Disaggregation
CO	Combinatorial Optimization
DNN	Deep Neural Network
DT	Decision Tree
EMI	Electromagnetic Interference
FHMM	Factorial Hidden Markov Model
HMM	Hidden Markov Model
ILM	Intrusive Load Monitoring
LSTM	Long Short-Term Memory
MF1	Mean F1
NILM	Non-Intrusive Load Monitoring
NN	Nearest Neighbor
RF	Random Forest
RNN	Recurrent Neural Network
SVM	Support Vector Machine

Abstract

Non-intrusive load monitoring (NILM) is that procedure used to disaggregate the contributions of different appliances in a building. Various techniques in the literature use supervised and unsupervised learning for disaggregation. Very few works utilize the semi-supervised or active approaches. The two approaches are used in case of training data scarcity. It is known that semi-supervised learning depends on the most certain samples to build the model. On the other, hand the active learning relies on most uncertain samples as it consults the user or the expert.

The current work introduces a new technique that leverages both the semi-supervised and active learning together, so it benefits from the merits of both learning models. The main idea is that semi-supervised learning improves the results of active learning aiming to decrease the need to the user. Two different approaches were utilized, one used active and reactive power features and the other used current waveform harmonics to use them later in the machine learning model.

Results show that when using the proposed technique, the number of queries can be reduced to fifth without losing more than 4% of accuracy. Also, the mixed technique reached 6% more accuracy than semi-supervised approach and up to 5% more accuracy than active only technique.

Chapter 1 : Introduction

Energy crisis and global warming are two major problems facing most of the economies in the current era, so new policies are introduced to mitigate the issues resulting from those hard problems. Increase in the demand for more energy makes clean sources of energy insufficient to cover different residential and industrial needs. This leads to make the use of fossil fuels still inevitable for scaling powerful economies regardless of the dangerous consequences on people and the planet. The combustion of more fossil fuels increases the volumes of greenhouse gases which are believed to be the major reason for climate change. So, it is apparent that there is a need to manage power consumption in a more efficient way and not only using non-conventional sources for power generation. Hence, if the demand of energy is reduced by efficient consumption management improves the status of energy crisis leading to more development for human kind.

Energy demand is currently erupting, and it is forecasted to grow further, on the other hand the need of lower the costs comes to the surface. This is crucial for residential, industrial and commercial sectors, so a feedback system with detailed consumption reports shows its effectiveness [1].

According to the UNEP Global Status Report [2], 31% of energy consumption is due to the buildings sector. Similar statistics [3] show that about 40% of US energy consumption is utilized by the buildings sector. Hence, the different trials to reduce the energy consumption will, for sure, have huge economic benefits. One of these trials, is the ability to monitor a building's consumption of electricity in a detailed way reaching the appliance level.

Appliance load monitoring is the technique used to monitor individual appliances electricity utilization. It can be categorized into intrusive load monitoring ILM and non-intrusive load monitoring NILM. In the ILM case, sensors are placed at each electrical appliance and measures the exact consumption of electricity. This method results in an accurate monitoring strategy, but it is expensive to apply to every device in a single household which makes it economically impossible. The other NILM technique flourished with the spread use of smart meters that it requires only measurement at the mains of a household leaving the need of huge ILM budget behind.

NILM appeared as an economic alternative for ILM technique, but it imposed new challenges such as how to detect the operation of different appliances from a single waveform and the accuracy that this results in. Different approaches used machine learning themes for tackling the issue, that are still in progress nowadays for better identification and detection of different appliance behaviors. Because of the economic value, NILM attracted researchers in other energy sectors such as water and gas [4].

NILM developed since firstly introduced by Hart [5] in the eighties when it was hard to progress more due to hardware limitations. Nowadays, the improvement in metering capabilities [6] and computing hardware led to more informative measurement and adoption of more sophisticated machine learning schemes like neural networks [7] and sophisticated hidden Markov models [8].

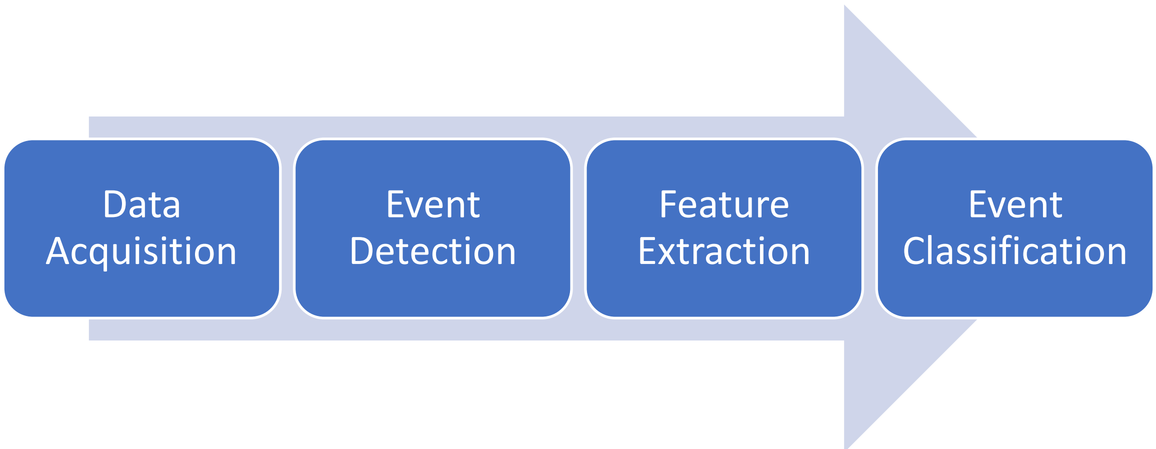

Figure 1.1: NILM process

NILM process is mainly composed of four stages as depicted in Figure 1.1. First, the electrical signal is sampled where the sampling frequency depends primarily on the measurement device [9]. Second, event, such as a device turned on or off, is detected using probabilistic or heuristic approaches [10]. Third, features describing this event are extracted, which may be steady or transient state features [11] and [12] depending on the sampling frequency used in the first step. Fourth, the event is classified to know which appliance it belongs to.

The classification stage is done through machine learning algorithms by training a classifier to get a hypothesized function that can classify different event classes. Two well-known types of machine learning are supervised and unsupervised learning, the first one is used for prediction and the other is used for description of the structure of data. The classification is one of the categories of supervised learning, but for the sake of training a classifier some input data with known labels must be provisioned.

In the case of NILM likewise most of data nowadays, there are some labeled datasets but also there is a massive low-cost amount of unlabeled data [13] produced by the measurement devices.

Hence, semi-supervised learning comes handy that it can exploit both labeled and unlabeled data which can lead to better results as shown in [14]. Another approach is to actively query the user to label the more informative unlabeled instances which is called active learning as conducted by [15] and [16].

The objective of the current thesis is to discuss the effect of merging both semi-supervised and active learning techniques and to show its ability to surpass the individual usage of any of the two techniques and also the supervised learning techniques in case of scarcity of training data.

The thesis is divided into five chapters. The current chapter provides an overview of the whole work. The second chapter contains the literature survey that has in-detail insights about the NILM from its early beginning, passing by different algorithms used till the current available datasets. The third chapter describes the proposed algorithm and BLUED dataset which is used entirely in the results. The fourth chapter describes different experimentation scenarios as well as the results. The final chapter concludes the work and lays some ideas for future.