



**Computer Science Department
Faculty of Computer and Information Sciences
Ain Shams University**

Data Mining Approach for Energy Management based on Internet of Things (IoT) Paradigm

Thesis submitted as a partial fulfillment of the requirements for the degree of
Master of Science in Computer and Information Sciences

By

Sarah Osama Anwar

Computer Science Department,
Faculty of Computer and Information Sciences,
Ain Shams University

Under Supervision of

Prof. Dr. Abdel-Badeeh Mohamed Salem

Professor of Computer Science,
Computer Science Department,
Faculty of Computer and Information Sciences,
Ain Shams University

Dr. Marco Alfonse

Lecturer of Computer Science,
Computer Science Department,
Faculty of Computer and Information Sciences,
Ain Shams University

March 2019
Cairo

Acknowledgment

I would like to thank everyone who has joined me in my thesis journey. I am blessed for having you. This work has been accomplished by your support.

First, I am so grateful to my supervisors for their continuous support. I have learnt a lot from your advices and your great knowledge. I would like to thank Prof. Abdel-Badeeh Salem for his wealthy knowledge and valuable feedback. I would like to thank my mentor Dr. Marco Alfonse for his endless patience and guidance.

I would like to thank my great family for their continuous support and prayers. I am grateful for everything you did for me. Special thanks to my great mother for helping me to find my way. Special thanks to my great father for believing in me. Special thanks to my amazing siblings for their unconditional love and support. Appreciating your immense presence a lot.

Finally, I would like to thank my supportive and lovely husband for his ultimate motivation. Thank you for listening to me and being my inspiration.

Abstract

With the rapid growth of population and the evolution of new technologies, energy demand is increasing expeditiously in the residential sector. Energy management solutions should be provided to save environmental resources and guarantee the efficient use of energy.

The Internet of Things (IoT) plays a crucial role in promoting the smart city development. Smart grid environment has been introduced with the advent of the IoT. The smart grid has enabled the transmission of electricity in a bidirectional way between dwellings and utilities since that dwellings can generate electricity by integrating with renewable energy sources such as solar panels and windmills.

Smart meters are considered the main component of the smart grid environment. Smart meters measure the consumed power frequently based on a specified time interval for each dwelling at appliance level. Infinite stream of consumption data is being generated from smart meters. Precious information can be obtained by analyzing these data.

Data mining techniques are extensively used for extracting valuable information from smart meters data. However, mining smart meter data is a challenging task with the infinite stream being generated with its granularity level.

Energy management solutions can be developed with the availability of smart meters data. Demand Response (DR) programs are developed for guiding home residents to save energy. The key success of DR programs is to respect home residents preferences and their usage behavior. Thus, Designing DR programs that are tailored based on their preferences to motivate them to respond to DR programs.

The proposed approach presented in this work extracts interesting usage patterns from smart meters data to help home residents understand their consumption usage. The First contribution of this work is to extract temporal associations of energy consumption at appliance level. The second contribution of this work is to extract association between appliances usage through mining temporal association rules. Moreover, applying temporal clustering technique for grouping appliances with similar usage time.

In this work, an incremental approach has been developed to process data generated from smart meters. Thereby, new associations are discovered without mining the whole data in addition to maintaining the old ones since that residents behavior is changing continuously with time.

The proposed approach extends the Utility-oriented Temporal Association Rules Mining (UTARM) algorithm to discover appliances usage preference at a time. The basic idea for extending the UTARM algorithm is that it merges two data mining tasks together: temporal association mining and utility oriented mining.

The proposed approach is applied on a real-world time series dataset to evaluate its results. The UK Domestic Appliance-Level Electricity (UK-DALE) dataset has been used in this work. The dataset holds consumption logs for five dwellings with different durations starting from November 2012 to April 2017 with time resolution of 6 seconds.

The results achieved succeeded to extract interesting usage patterns with respect to the 24-hours of the day. Appliance-time associations are extracted considering appliances usage priority as a utility factor and the hour as a temporal partitioning factor. In addition to, discovering appliance-appliance associations having similar usage pattern with respect to the 24-hours of the day.

List of Publications

- [1] S. Osama, M. Alfonse and A. M. Salem, "Intelligent Techniques for Smart Home Energy Management Based on Internet of Things (IoT) Paradigm," *Egyptian Computer Science Journal*, vol. 41, no. 3, pp. 33-43, 2017.
- [2] S. Osama, M. Alfonse and A. M. Salem, "A Study on Pattern Discovery of Smart Meter Data for Energy Efficiency," *International Journal Information Models and Analyses*", vol. 7, no. 2, pp. 14-24, 2018.
- [3] S. Osama, M. Alfonse and A. M. Salem, " Mining Temporal Patterns to Discover Inter-Appliance Associations using Smart Meter Data," *Big Data and Cognitive Computing*", MDPI, vol. 3, no. 2, 2019.
- [4] S. Osama, M. Alfonse and A. M. Salem, " An Efficient Algorithm for Extracting Appliance-Time Association using Smart Meter Data," *heliyon*", ELSEVIER. (accepted)

Table of Contents

Acknowledgment	II
Abstract	III
List of Publications	VI
Table of Contents	VII
List of Figures	IX
List of Algorithms	X
List of Tables	XI
List of Abbreviations	XII
Chapter 1. Introduction	2
1.1 Motivation	3
1.2 Objectives	3
1.3 Methodology	4
1.4 Contributions	5
1.5 Thesis Organization	5
Chapter 2. Data Mining for Energy Management	7
2.1 Energy Management in Residential Sector	7
2.2 The Role of Internet of Things (IoT) in Energy Management	11
2.2.1 Smart Grid	12
2.2.2 Smart Meters	13
2.2.3 Smart Meter Applications	14
2.3 The Role of Data Mining in Energy Management	18
2.3.1 Clustering techniques	20
2.3.2 Frequent Pattern Mining	21
2.3.3 Sequential Pattern Mining	21
2.3.4 Temporal Pattern Mining	22
2.4 Discussion	23
Chapter 3. Related work	25
3.1 Appliance-Appliance Association	26
3.1.1 Frequent Pattern Mining	26
3.1.2 Sequential Pattern Mining	33
3.2 Appliance-Time Association	36
3.3 Usage Pattern Mining	39
Chapter 4. Proposed Methodology for Extracting Appliance-Time Association	45
4.1 Dataset	45
4.2 Proposed Methodology	46
4.2.1 Generating Usage Matrix	46
4.2.2 Updating Utility Data	47

4.2.3	Extracting the Appliance-Time Association.....	51
4.3	Results and Discussion.....	56
4.4	Summary	65
Chapter 5.	Proposed Methodology for Extracting Appliance-Appliance Association.....	67
5.1	Temporal Association Rules	68
5.2	Temporal Clustering.....	71
5.3	Results and Discussion.....	74
5.3.1	Temporal Association Rules Results	74
5.3.2	Temporal Clustering Results	80
5.4	Evaluation.....	85
5.5	Summary	86
Chapter 6.	Summary, Conclusion and Future Work	88
6.1	Summary	88
6.2	Conclusion.....	90
6.3	Future work	91
References	93

List of Figures

Figure 2.1: Energy Management Solutions	8
Figure 2.2: Data Mining Techniques in Energy Management	19
Figure 4.1: Appliance-Time Association.....	52
Figure 4.2: House 1 Appliance-Time Association	57
Figure 4.3: House 2 Appliance-Time Association	59
Figure 4.4: House 3 Appliance-Time Association	60
Figure 4.5: House 4 Appliance-Time Association	62
Figure 4.6: House 5 Appliance-Time Association	63
Figure 5.1: Appliance-Appliance Association using Association Rules.....	69
Figure 5.2: Appliance-Appliance Association using Hierarchical Clustering	72
Figure 5.3: House 1 Appliance-Appliance Association	81
Figure 5.4: House 2 Appliance-Appliance Association	82
Figure 5.5: House 3 Appliance-Appliance Association	83
Figure 5.6: House 4 Appliance-Appliance Association	83
Figure 5.7: House 5 Appliance-Appliance Association	84
Figure 5.8: Runtime Analysis	85

List of Algorithms

Algorithm 4.1: Updating Utility Measures	53
Algorithm 4.2: Extracting Appliance-Time Association	54
Algorithm 5.1: Generating Appliances Association Rules	70

List of Tables

Table 2.1: Data Mining for Energy Management.....	23
Table 3.1: Frequent Pattern Mining Approaches/Algorithms	31
Table 3.2: Sequential Pattern Mining Approaches/Algorithms	36
Table 3.3: Appliance-Time Association Approaches/Algorithms	39
Table 3.4: Usage Pattern Mining Approaches/Algorithms	43
Table 4.1: The UK-DALE Dataset Specification	45
Table 4.2: Sample of House 1 Appliance-Time Association	58
Table 4.3: Sample of House 2 Appliance-Time Association	60
Table 4.4: Sample of House 3 Appliance-Time Association	61
Table 4.5: Sample of House 4 Appliance-Time Association	62
Table 4.6: Sample of House 5 Appliance-Time Association	64
Table 5.1: The Differences between Association Rules and Hierarchical Clustering	73
Table 5.2: Sample of House 1 Appliance-Appliance Association Rules	75
Table 5.3: Sample of House 2 Appliance-Appliance Association Rules	76
Table 5.4: Sample of House 3 Appliance-Appliance Association Rules	78
Table 5.5: Sample of House 4 Appliance-Appliance Association Rules	78
Table 5.6: Sample of House 5 Appliance-Appliance Association Rules	79

List of Abbreviations

ARMADA	Algorithm for Discovering Richer Relative Temporal Association Rules
CoPMiner	Correlation Pattern Mining
DCMiner	Dynamic Correlation Miner
DR	Demand Response
EU	External Utility
FP-tree	Frequent Pattern tree
FTU	Frequent Temporal Utility
FU	Frequency
HEMS	Home Energy Management System
IoT	Internet of Things
IU	Internal Utility
Kulc	Kulczynski Measure
LDA	Latent Dirichlet Allocation
M2M	Machine-to-Machine
PBuilder	Pattern Builder
SGSC	Smart Grid Smart City
SOM	Self Organising Maps
SPADE	Sequential PAttern Discovery using Equivalence classes
StrPMiner	Streaming Pattern Miner
TPUP	Time-slot Probability Usage Pattern
TWU	Transaction Weighted Utility
U	Utility
UK-DALE	UK Domestic Appliance-Level Electricity
UTARM	Utility-oriented Temporal Association Rules Mining

Chapter 1

Introduction

Chapter 1. Introduction

With the rapid growth of population, there is an expeditious rise in energy demand all over the globe [1]. Energy management solutions are needed to guarantee the efficiency of energy use, thereby maximizing productivity, minimizing energy costs, reducing environmental effects such as the emissions of carbon dioxide CO₂. Thereby, providing a cleaner environment and preserving natural resources.

Energy is consumed by mainly four sectors: residential, commercial, transportation and industrial. The proposed work focuses on energy management solutions in the residential sector.

Energy consumed in residential sector is related to the activities of home residents and their behavior of using their home appliances. The daily life of home residents follows some routines habit to some extent that makes it possible to study and understand their preferences. For example, microwave and coffee machine are used during breakfast activity while oven and kitchen lights are used during dinner activity. Understanding the behavior of home residents and respecting their preferences are essential factors for developing energy management solutions.

Extracting useful patterns from consumption data has a great benefit for both consumers and producers. For consumers by lowering their electricity bills and for producers by preventing system stress and supply shortage.

1.1 Motivation

In regard of the evolution of new technologies, energy demand is increasing expeditiously. Thus, increasing the pressure on power plants, utilities and natural resources. Moreover, traditional electric power systems are facing several challenges such as system stress conditions in peak hours causing supply limit and blackouts in large areas [1].

The residential sector is one of the main contributors in power consumption. The power consumed in residential sector has an increase around 300% in the United States from 1949 to 2011. The power consumed in 1949 was around 5,599 Quadrillion Btu of power energy and then in 2011 the consumption has been increased to be 21,619 Quadrillion Btu [2].

With the inflation rates of electricity prices, home residents begin to pay much attention to their energy consumption. However, they failed to save energy in their daily activities due to the lack of knowledge of their appliances usage consumption. Since the only measure they have is their electricity bills which do not help them to understand what behavior that they are doing to make it high or low.

1.2 Objectives

Recognizing and understanding home residents behavior and provisioning a continuous feedback about their consumption usage is one of the effective ways to save energy in the residential sector.