REPUBLIC OF TURKEY

ALTINBAS UNIVERSITY

Institute of Graduate Studies

Information Technologies

RESOLVING ENERGY CONSUMPTION ISSUES AND SPECTRUM ALLOCATION FOR FUTURE BROADBAND NETWORKS

Sinan najamaldeen azzah AZZAH

Master of Science

Supervisor

Asst. Prof. Dr. Muhammad ILYAS

Istanbul, 2021

RESOLVING ENERGY CONSUMPTION ISSUES AND SPECTRUM ALLOCATION FOR FUTURE BROADBAND NETWORKS

Sinan najamaldeen AZZAH

Information Technologies

Submitted to the Institute of Graduate Studies

in partial fulfillment of the requirements for the degree of

Master of Science

ALTINBAŞ UNIVERSITY

2021

The thesis titled "RESOLVING ENERGY CONSUMPTION ISSUES AND SPECTRUM ALLOCATION FOR FUTURE BROADBAND NETWORKS" prepared and presented by Sinan najamaldeen azzah AZZAH was accepted as a Master of Science Thesis in Electrical and Computer Engineering

____

Asst. Prof. Dr. Muhammad ILYAS

Supervisor

Thesis Defense Jury Members:

Asst. Prof. Dr. Muhammad ILYAS	Faculty of Engineering and Architecture, Altinbas University	__________________
Asst. Prof. Dr. Abdullahi Abdu IBRAHIM	Faculty of Engineering and Architecture, Altinbas University	__________________
Asst. Prof. Dr. Jawad RASHEED	Faculty of Engineering, Aydin University	__________________

I hereby declare that this thesis meets all format and submission requirements of a Master's thesis..

Accepted by Altınbaş University Institute of Graduate Studies on the following Data:

____/____/____

I hereby declare that all information data presented in this graduation project has been obtained in full accordance with academic rules and ethical conduct. I also declare all unoriginal materials and conclusions have been cited in the text and all references mentioned in the Reference List have been cited in the text, and vice versa as required by the abovementioned rules and conduct

Sinan Najamaldeen Azzah AZZAH

Signature

DEDICATION

I devote and pledge this research work to my supervisor who is salient for guiding me through whole research work as well as my family for always assisting me in my hard time.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my supervisor Asst. Prof. Dr. Muhammad ILYAS for guiding and helping me along the way in writing this dissertation. Discussing my progress, problems, and ideas with my supervisor Asst. Prof. Dr. Muhammad ILYAS a couple of times every week helped me tremendously in understanding the logic behind the research. It made me better realize the technical need for this research work.

ABSTRACT

RESOLVING ENERGY CONSUMPTION ISSUES AND SPECTRUM ALLOCATION FOR FUTURE BROADBAND NETWORKS

AZZAH, Sinan Najamaldeen Azzah,

M.Sc., Information Technologies, Altınbaş University,

Supervisor: Asst. Prof. Dr. Muhammad ILYAS

Date: December /2021

Pages: 61

The aim of this thesis is to resolve the energy consumption issues and determine the usage of energy or power with high spectrum allocation in future broadband networks with the help of clustering in data mining. The research work starts presenting an overview of the broadband network energy sector and the challenges it is facing; it is observed a change on the energy policies promoting the energy efficiency, encouraging an active role of the consumer, instructing them about the importance of the consumer behavior and also protecting consumer rights. Electricity is gaining room as energy source, its share will keep increasing constantly in the following decades. In this close future, broadband networks and smart meters' deployment will benefit both the utility and the consumer. In this environment, new services and new business appear, focusing on the energy management field and tools, they require specialization in fields such as, computer science, software development and data science. This research work has segmented the broadband networks according to the similarities of their electrical load profiles, using the proportion of energy usage per hour (%) as a common framework. The objective behind this energy consumption segmentation is to be able to provide personalized recommendations to each group in order to reduce their energy consumption and the associated costs, fostering energy efficiency measures and improving the consumer engagement. The desired segmentation is obtained by an iterative process, based on computational clusters calculation (using python programming language) and finalized by a post-clustering analysis applying visualization and statistical data mining technique to detect the energy consumption and reallocate them to a more appropriate group. The K-Means clustering technique was tested and compared, giving best prediction of accuracy 98.46% for all energy load profiles with high spectrum of 100GHz. The solution from the K-Means clustering is the one that better adapts to the segmentation sought, which is used as the base of the post-clustering stage to obtain the final energy consumption segmentation. Most of these methodologies use the absolute values in 100 kWh, as they were more focused on identify the users with higher energy savings potential. The energy consumption segmentation of the electricity consumers provides knowledge and a better understanding of the consumer. In this particular case, it allows to personalize energy savings recommendations according to the broadband networks specific characteristics; improving the consumer experience by being able to provide the adequate advices at the appropriate time, facts that increase the effectiveness of the energy efficiency advices' service for future broadband networks.

Keywords: Energy, Data Mining, K-Means, Broadband Networks

TABLE OF CONTENTS

Pages

LIST OF TABLES

Pages

LIST OF FIGURES

Pages

LIST OF ABBREVIATIONS

AMI	:	Advance Machine Infrastructure
API	:	Application Programming Interface
DER	:	Distributed Energy Resource
EU	:	European Union
G2V	:	Grid to Vehicle
GPRS	:	General Packet Radio Service
IEA	:	International Energy Agency
kWh	:	Kilowatt Hours
Mtoe	:	Millions of Tons of Oil Equivalent
SaaS	:	Software as a Service
TCP	:	Transmission Control Protocol
WAN	:	Wide Area Network

1. INTRODUCTION

1.1 INTRODUCTION

Energy has become a basic instrument for the well-being of the society, together with its increasing importance in the world; the energy sector is facing important challenges nowadays, aiming to cover all the new necessities and adapting the old standards to the new situation as mentioned in [1]. During the last decades there has been special concern on environmental issues, especially from the policy point of view. The celebration of three global environmental conferences in Stockholm (2015), Rio de Janeiro (2016) and Johannesburg (2017) prove so as mentioned in [2]. A thinking evolution throughout these meetings end up with the appearance of the Sustainable Development concept, which represents a more holistic approach where economic, social and environmental dimensions need to be treated together. Figure 1.1 illustrates the energy role in each Sustainable Development dimension.

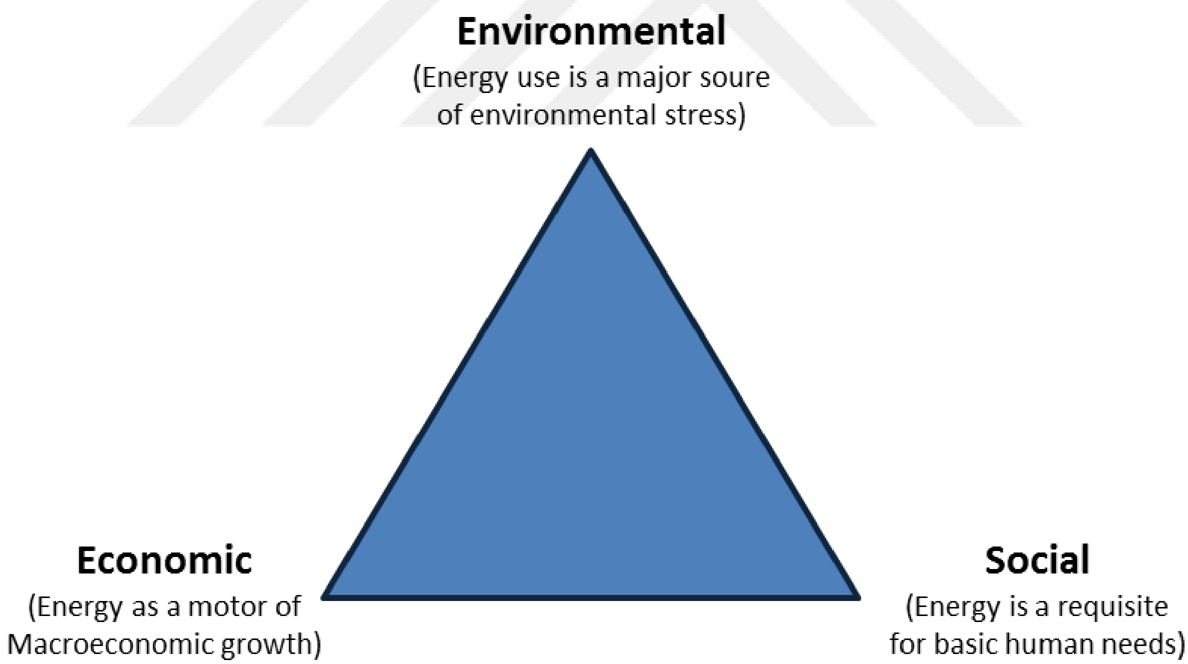

Figure 1.1: Energy and Sustainable Development linkages [3].

The extraction, process and use of the conventional energy sources, mainly fossil fuels such as oil, coal, natural gas; allowed, in the past, to have access to cheap and abundant energy contributing to the economic development, but also leaving behind major environmental impacts at a global, regional and local level by using these scarce and non-renewable resources as mentioned in [4].

The extraction of natural resources (mining) has significant impacts in the local ecosystems normal development; the process and use of energy implies atmospheric releases of greenhouse gases like CO2, CH4, N2O among others; leading to environmental damages: climate change, acid deposition, urban smog as mentioned in [5]. A way to address this problem is switching to renewable/clean energies, for instance photovoltaic, solar thermal, wind energy, biomass, hydropower… this kind of energies are characterized by a low or nearly zero environmental impacts during extraction, conversion and use, therefore contributing to the climate change mitigation., if and when this will be possible as mentioned in [6].

Energy is directly related to macroeconomic indicators; there is an explicit correlation between countries' energy demand and the size of their economies, measured by the countries' Gross Domestic Product (GDP) as it shows Figure 1.2. Principally, energy is consumed to feed the industrial processes that convert raw materials into final goods or services, creating added value.

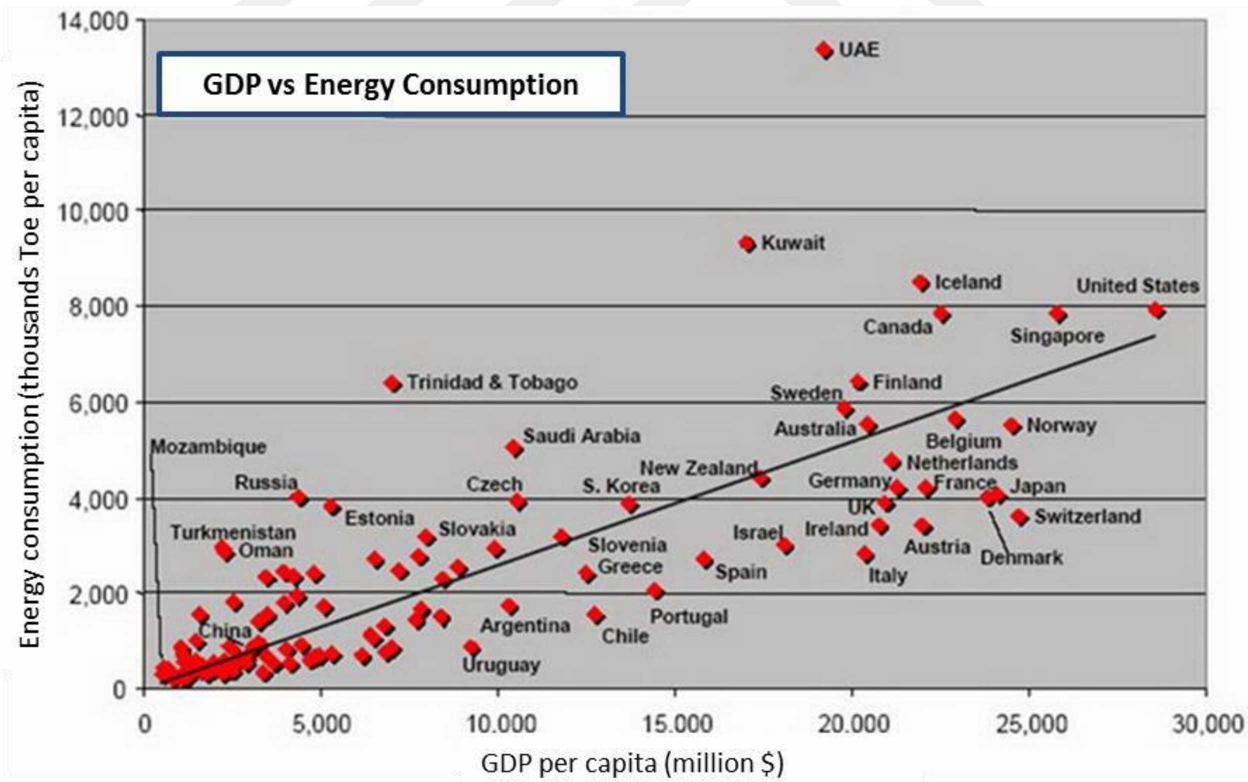

Figure 1.2: Correlation GDP vs Energy consumption. Energy Consumption per capita in thousands of Tons of equivalent oil (toe) and Gross Domestic Product per capita in million $ [7].

The previous relationship might seem obvious, although two much better indicators are:

Energy intensity, relating units of energy per units of GDP [tons of oil equivalent- toe/$], as shown in Figure 1.3 where general descending of the energy intensity is seen in the various economies presented, both developed and developing countries.

Energy Efficiency of an economy [$/toe], which is the energy intensity reverse. A clear classification and grouping of the top 40 economies can be found in [8], developed countries economies are placed in the high GDP per capita but low efficient, and most of the developing countries are also low efficient with a lower GDP per capita.

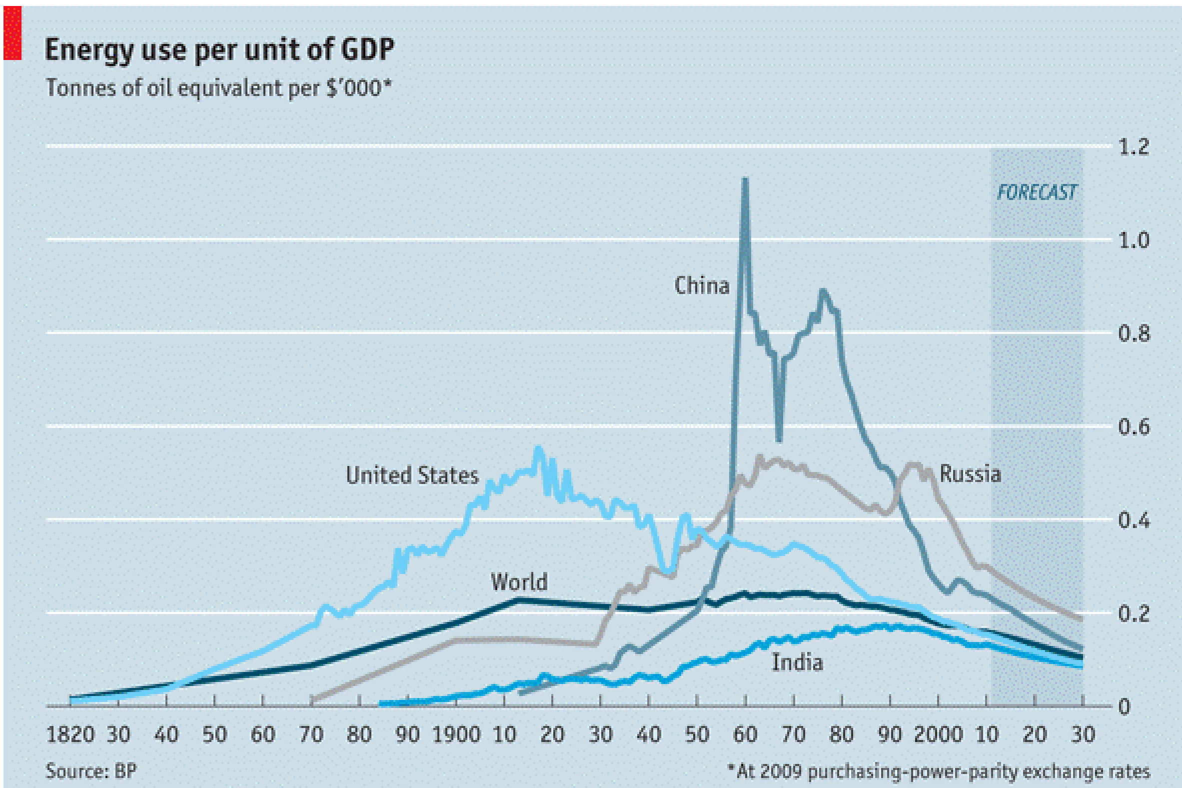

Figure 1.3: . Energy use per unit of GDP trend [9]

Not long ago, it has arisen the question of decoupling the energy use and economic growth, by using less energy to produce the same or a higher economic outcome. The example of Germany's Energy Transition states that it might be a room for improvement by means of energy efficiency as mentioned in [10]. The more complex connection of energy deals with the social dimension of the Sustainable Development, when meeting the basic human needs as mentioned in [11]. Energy is essential to fulfil these basic needs such as food, water, shelter, healthcare, education, unemployment…

Human Development Indicator (HDI) was created to measure and rank countries' human development in a common reference, this statistic parameter combines life expectancy, education and income indices to determine nation's social development as mentioned in [12]. The Human Development Reports from the United Nations Development Programmed recollect all countries' data referred to the HDI and many other indexes.

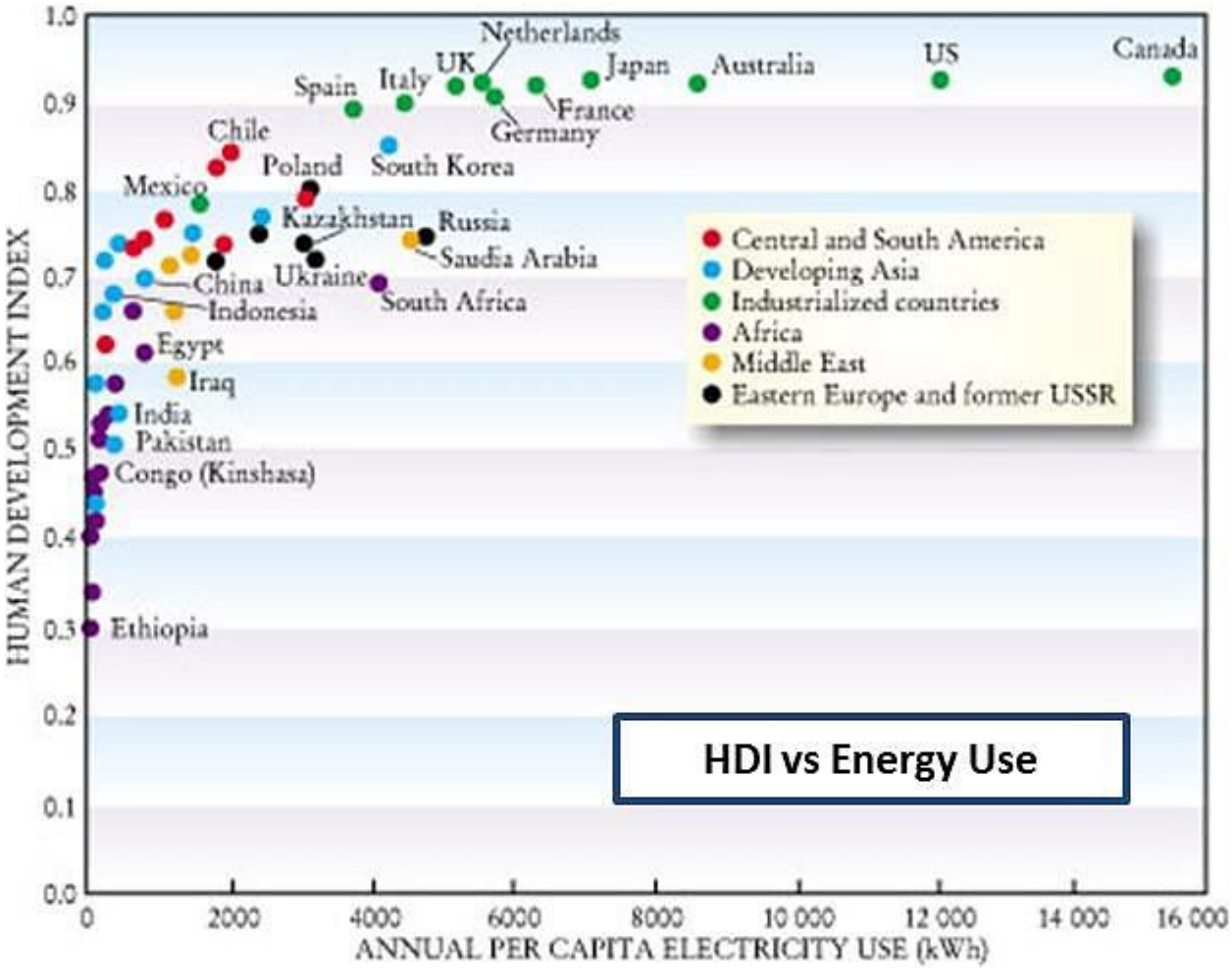

Figure 1.4: Energy use per country in kWh vs Human Development Index [13].

Some European articles and reports, such as (European Commission: Energy Supply and Demand, 2016) point out the strong relationship between social indicators and energy use, apart from the fact that more population growth more energy demand, developing countries consume much less energy than developed countries from household's energy requirements perspective as mentioned in [14]. A worrying fact related to the lack of access to energy services (electricity or cooking facilities), is the Energy Poverty. The International Energy Agency (IEA) corroborates that still 18% of the world population does not have access to electricity (IEA: Energy poverty, 2016) as mentioned in [15]. However, energy poverty is not only affecting the developing countries, the recent years pushed by the global economic crisis and the rise of energy prices, families from developed cannot face the costs of accessing to energy as mentioned in [16].

1.2 MOTIVATION

The energy demand is expected to grow significantly in the coming decades, up to doubling the current demand by 2035; this increase will be led by the developing and non-OECD countries mainly China and India. In the same direction the International Energy Agency, World Energy Outlook 2017, has made the projection by 2035 presented in Figure 1.5 which can be found in [17].

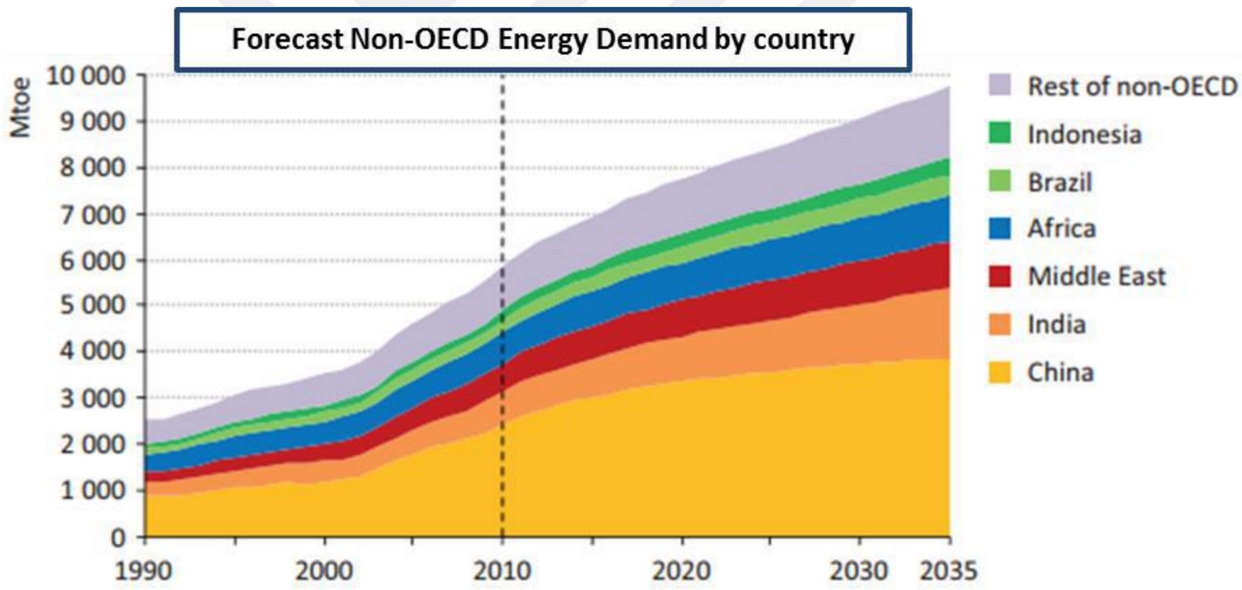

Figure 1.5: Forecast of Energy Demand by non-OECD countries [17].

Energy policies direct their focus to energy efficiency and to the integration of renewables starting an energy transition. A better use of the available energy and the use of clean energies will help to reduce the CO2 emissions leading to a decarburization of the economy as mentioned in [18]. The fully substitution of fossil fuels is yet far from reality, so the conventional primary energy sources will be still necessary to meet this demand increase, as seen in Figure 1.6.

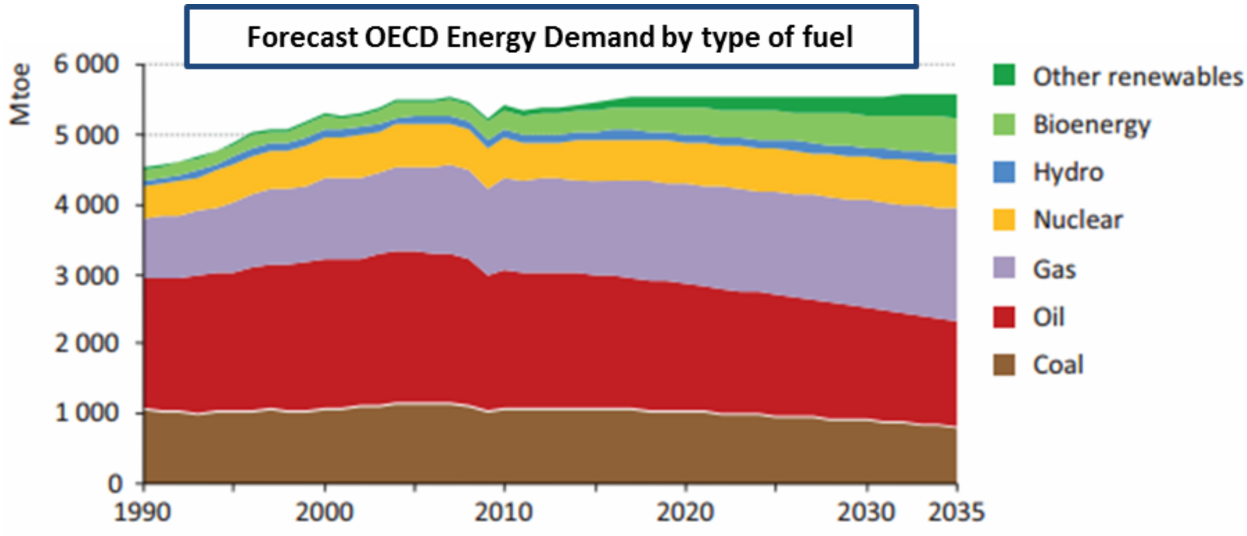

Figure 1.6: Forecast of Energy Demand on OECD countries by type of fuel [17].

Finally, the 2011 energy demand allocation by sector is presented in Figure 1.7, and it almost double the total energy consumed compared to the same figure by 1973; where the shares doesn't vary too significantly as mentioned in [19]. It allows having an idea to know where and what for this energy is consumed, helping to elaborate more adequate policies to each sector.

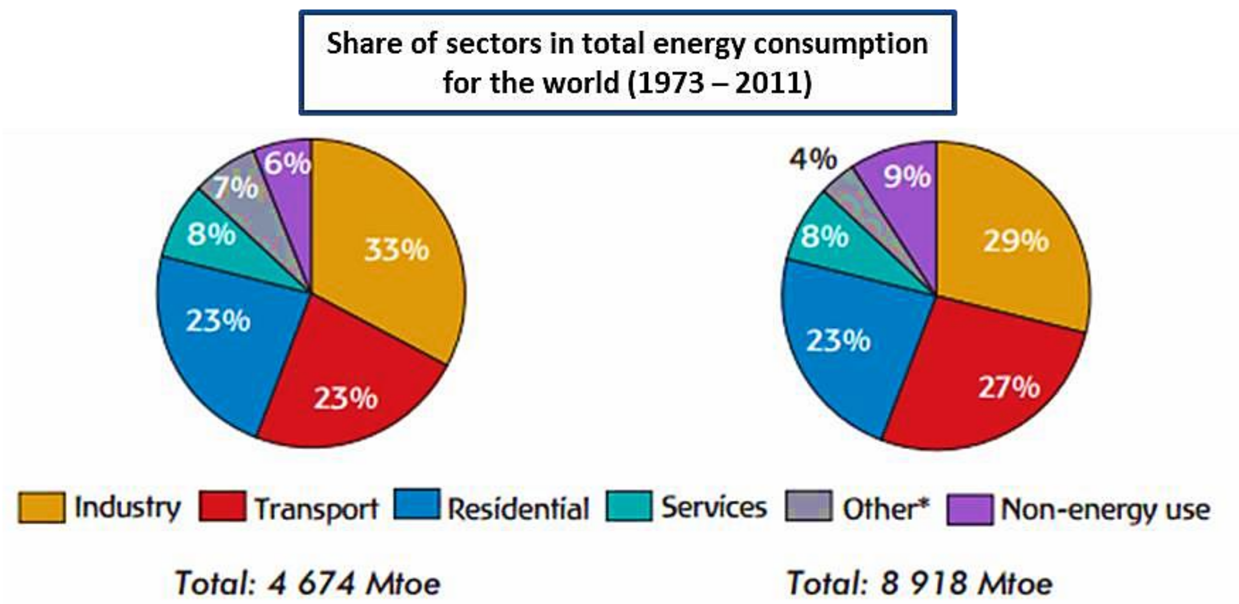

Figure 1.7: Comparison energy consumption by sector shares and total from 1973 (left) to 2011 (right) [19].

European Union countries are importing more than half of its energy for instance, around 90% of crude oil and 66% natural gas (European Commission: Imports and Secure Supplies, 2015); being the largest energy importer in the world, this fact has associated costs around 400 billion € per year as mentioned in [20].

1.3 PROBLEM STATEMENT

In that sense, the fragmented market and old energy infrastructure with insufficient countries interconnections, for example 12 state members are energy isolated islands (<10% interconnection) and 6 members depend of a single supplier, result to consumer's vulnerability and higher prices as mentioned in [21]. For instance, electricity prices in Europe are 30% higher, and gas prices can double than US ones; subtracting competitiveness to the European business and enterprises. A common market legislation and the market liberalization can help to solve this issue.

An appropriate integration is expected to save up to 40 billion € each year and also reduce the dependence of the external suppliers. Ensuring the energy supply to all EU citizens has become a problem, moreover after the economic crisis that left more than 10% of European population without being able to pay their energy bill (energy poverty).

In the actual EU framework, strategy targets are set by 2020 aiming for (compared to 1990):

20% GHG reduction (40% by 2030)

20% increase of renewable energy in EU mix (27% by 2030)

20% improve of energy efficiency (27% by 2030)

Energy efficiency and use of renewable energy technologies are the most important features to achieve the targets set. Also, improving these numbers is the core of the broadband network energy strategy by 2030 and 2050 to mitigate the Climate Change.

In addition to the Energy Efficiency Directive, there are three other directives devoted on broadband network's thermal and electrical demand, as broadband networks represent around 40% of the total energy demand in the world these directives are:

Energy performance of broadband networks directive: include broadband network energy certificates when selling and renting broadband networks; finance renovation of broadband network elements to a low energy needs, and all new broadband networks must be nearly-zero energy broadband networks.

Energy labelling directive: intending to help consumers to have more information in order to choose energy efficient products (network charges, network bandwidth consumption, energy consumption, links…)

Eco-design directive: directed on product manufacturers, regulating the requirements of manufacturers to establish the minimum energy efficiency standards for future broadband networks.

1.4 RESEARCH CONTRIBUTIONS

Energy efficiency has been gaining importance in the energy world's priorities, it is named as the "invisible fuel" stating that the best choice is not to waste energy. In regions like Europe with highly external energy supply dependences, an optimization at all stages in the energy chain is a must from both environmental and economic point of view. Rules and obligations set by the Energy Efficiency Directive are in that direction. The target of the international community is to reduce up to 20 % the energy consumption, achieving by 2020 an energy consumption lower than 1.474 Mtoe of primary energy or less than 1.078 Mtoe of final energy as mentioned in [22], but setting an objective adequate to each country characteristics.

This directive differentiates the energy efficiency in broadband network supply and energy efficiency in energy use with high spectrum allocation. The measures adopted by the international standards can be summarized:

Broadband network energy distributors and retailing must reduce annually 1.5% their energy sales with high spectrum allocation.

Annual energy efficient renovation of least broadband network bandwidth owned or occupied by energy consumption.

Incentive the broadband networks renovation, i.e. adding energy insulation, network links, broadband network routing; to improve their energy performance

Mandatory energy performance certificates when renting broadband networks with low energy consumption.

Set of minimum labels or standards for a range of products, such as high spectrum allocation and reducing minor issues in the broadband network.

Periodical energy audits for large companies and incentives for Small and Medium enterprises to undergo energy audits

Protecting consumer rights to receive comprehensive information, access to real-time and historically energy consumption and billing data to a better consumption management

Deployment of 200 million of electricity smart meter (72% of the total) by 2020.

The energy efficiency in energy supply is principally focused on the highly-efficiency cogeneration plants (electricity + spectrum allocation), implementation of district heating and cooling and integration of renewables. Together with the energy savings attained due to supplier's obligations and the application of the white certificates. A modernization on energy infrastructure (transformation, transmission and distribution) leads to the final deployment of the smart grids, reducing grid losses (accounted by almost 30%) derived from energy generation, transportation and distribution.

1.5 THESIS ORGANIZATION

In the first chapter, background and motivation, research question, methodology and outline are presented. The broadband network based on high spectrum allocation faces challenges which are presented with a specific emphasis in introduction of dissertation with dependencies of problem statement and research contribution. The chapter ends with a description of the characteristics of an organization of thesis.

The second chapter gives a basic description of the literature review on broadband network and energy consumption, beginning with a small historical note, a definition of the commercially available types of available solutions and ending with a description of the major drawbacks of not classifying the energy needs for broadband network and the definition of high spectrum allocative system for real time energy usage monitoring in large scale environment.

The third chapter presents the broadband network system with the methodology background; the concept of data mining technique (K-Means Clustering) with networking concepts is presented, followed by the methodology for intelligence based broadband network with spectrum allocation monitoring techniques, a characterization of the broadband network, and the method for energy consumption classification and reducing cost of broadband network energy monitoring system with internet of things by providing the cost analysis, the purpose of this chapter, go through a sensitivity analysis and are presented and discussed.

The forth chapter in terms of structure is very similar to the third chapter but for the result of dissertation with aspect of this work, that is the broadband network with energy consumption monitoring system using time related features on broadband network with high spectrum allocation. A typical framework outline is followed: first the concept of broadband network based smart phones results is described, then the objective and scope of the analysis are being found, followed by data collection for broadband network using the energy.

The fifth chapter provides the discussion of the results with an analysis and their respective discussion based on the limitations and comparison with other existing systems on broadband network with energy consumption with high spectrum allocation monitoring for energy usage.

The sixth chapter presents the main conclusions of this dissertation with a summary of the main findings, the consequences these results have in policy making, followed by the main limitations of this study and ends with leads for future work.

2. RELATED WORK

This thesis work is motivated by the goal of overcoming the challenges and difficulties that energy consumption units in broadband network faces with the help data mining techniques in real time with their structure and the environmental hostility surrounds them, that is by creating a better environment solution, a solution that will avoid the shortcomings of the traditional networks for energy consumption and controlling operations. For the sake of explaining the motivation of this thesis work, a detailed review of all these difficulties and constraints is discussed in depth in the sections below; furthermore, a comprehensive review of broadband network system and their advantages that this thesis work aims to overcome.

2.1 LITERATURE REVIEW

The energy consumption paradigm has changed in recent years, for the first time policies seem to be consumer oriented ensuring them to receive a good service and able to defend their rights as energy consumer. Among the list presented in [23], there are two points clearly related to the purpose of the present work of tracking and managing the energy use:

Consumer has to receive and have access to accurate information to their electricity and gas consumption and billing.

Consumer has to receive information on how much energy uses and how to use energy more efficiently. Also the benefits of using renewable energy equipment and vehicles.

Also, policies seek to give the final push to the infrastructure modernization, not only interconnecting the countries but also by embracing innovation and new technologies to complete this new energy infrastructure as mentioned in [24]. An important transformation is occurring in the power generation model, moving from a Centralized to Decentralized energy production, illustrated in Figure 2.1.

The traditional Centralized power production is constituted of few and big power stations, the consumers are far from the production sites, the network is unidirectional from producer to consumer, there is no flexibility on the energy demand, so consumers have a passive paper as mentioned in [25]. In contrast, with the new Distributed production is not only delivering a service also including other features that improve the whole process. The production is also close to the consumption points, thanks to the small and medium plants built mostly by renewable energy sources (PV, wind…) or new technologies like fuel cells, the "prosumer" term appears to define those who produce and consume their own energy, the possibility to store the energy to use when and where is necessary, thanks to a network that becomes bidirectional, prioritizes the demand by using more efficiently the resources, taking advantage of ICT and the big data to shave consumption peaks and engage the customer, who will have an active role as mentioned in [26].

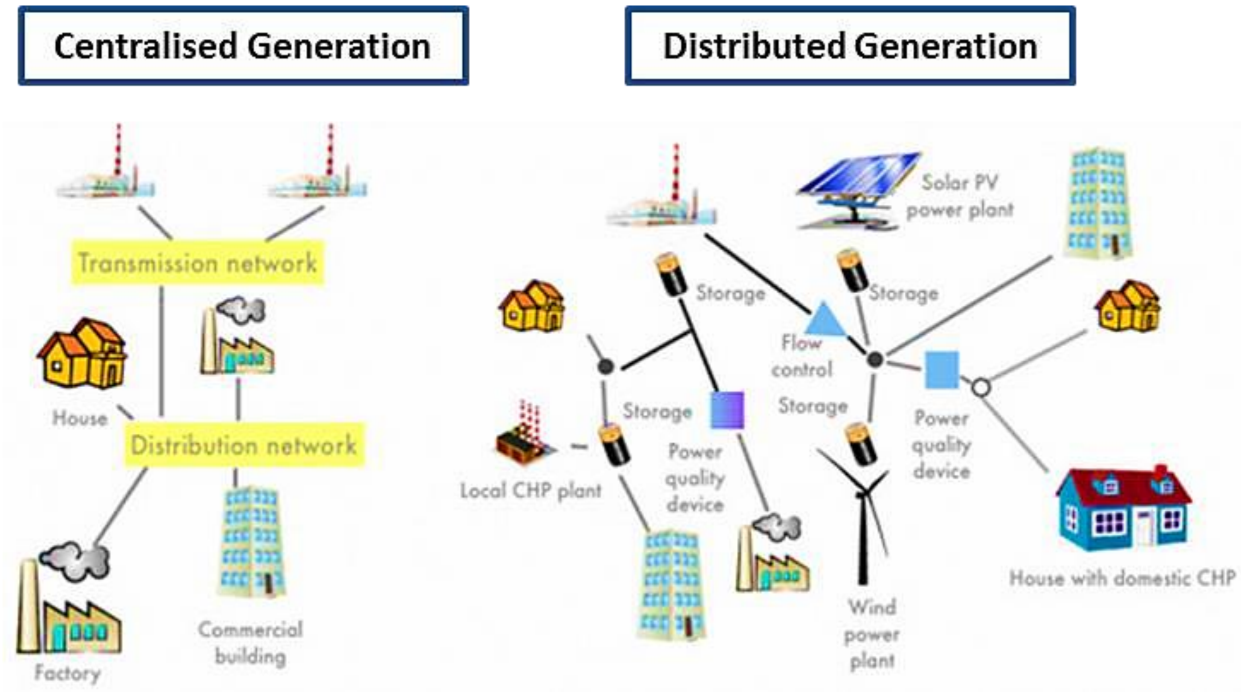

Figure 2.1: Conceptual model comparison between Centralized and Decentralized. [27].

In this new landscape, electricity is expected to be the world key energy, some studies, point that society is advancing towards an electricity world: devices, cars, heat pumps… where almost everything will be powered by electricity; in this sense smart grids are necessary to update the electricity network to the meet the actual situation as mentioned in [28].

2.2 USAGE OF ELECTRICITY

Electricity is called to be the fastest-growing component on the energy supply portfolio; the increase on electricity consumption is expected to achieve the 40% of the total energy market by 2020 and reaching an important share by 2050 as mentioned in [29]. This growth will be essentially led by the emerging economies and developing countries, as is shown in 2.2, while a modest increment in other developed economies will occur.

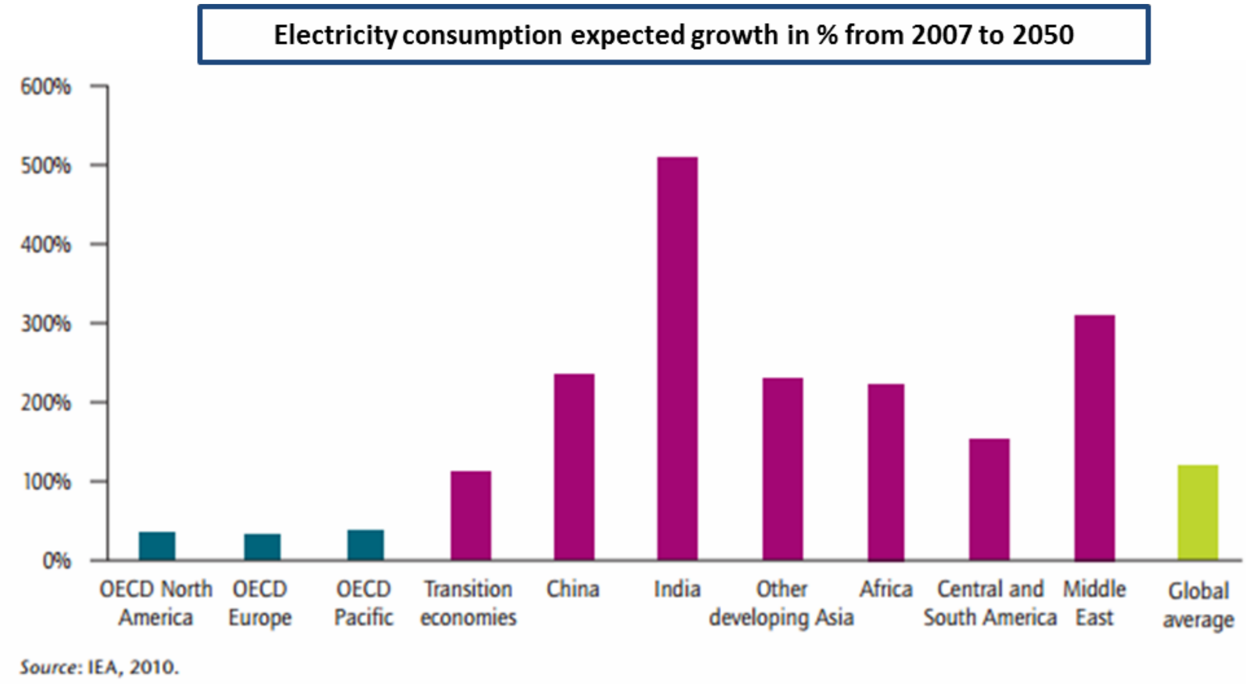

Figure 2.2: Electricity consumption expected % increase by regions [30].

This large growth on electricity usage is likely to arise in the:

Residential heating and cooling due to the use of heat pumps. Either individual at home heat pumps or larger district heating and cooling heat pumps.

The electrification of the transportation and mobility, due to vast deployment of Electric Vehicles (EV) and the Plug-in Hybrid Electric vehicles (PHEV)

Variable electricity generation, by using PV, wind, hydro, biomass, combined heat and power (CHP) technologies, decentralizing the production.

This new conjuncture will tackle the climate change to a CO2 emissions reduction, while covering the same needs moving from rich-carbon sources to low-carbon ones as mentioned in [31]. And maintain the economic and social development by using more sustainable systems; however, the electrical grid will have to be ready and able to host these new features.

2.3 SMART ELECTRICAL GRIDS FOR BROADBAND NETWORKS

The concept of smart grid has been widely used these recent years, but there is not a standard definition, it can be seen as the process of making the actual electrical grid smarter in a sense of modernization and adaptation to the new situation as mentioned in [32]. Smart grids are electricity networks that make use of the information and communication technologies (ICT) and are able to integrate other advanced technologies (distributed electricity generators, electricity storage, electric vehicles…); to maximize the efficiency, reliability and safety of the power grid, and minimize the costs and environmental impacts as mentioned in [33]. Allowing a flexible production and a consumer engagement, by providing information and tools to both electricity producers and consumers. As consumers will play an important role in the new electric grid landscape.

It that sense a difference can be made between the core of the smart grid that falls on the transmission and distribution of electricity and the peripheral ones that are the generators are consumers.

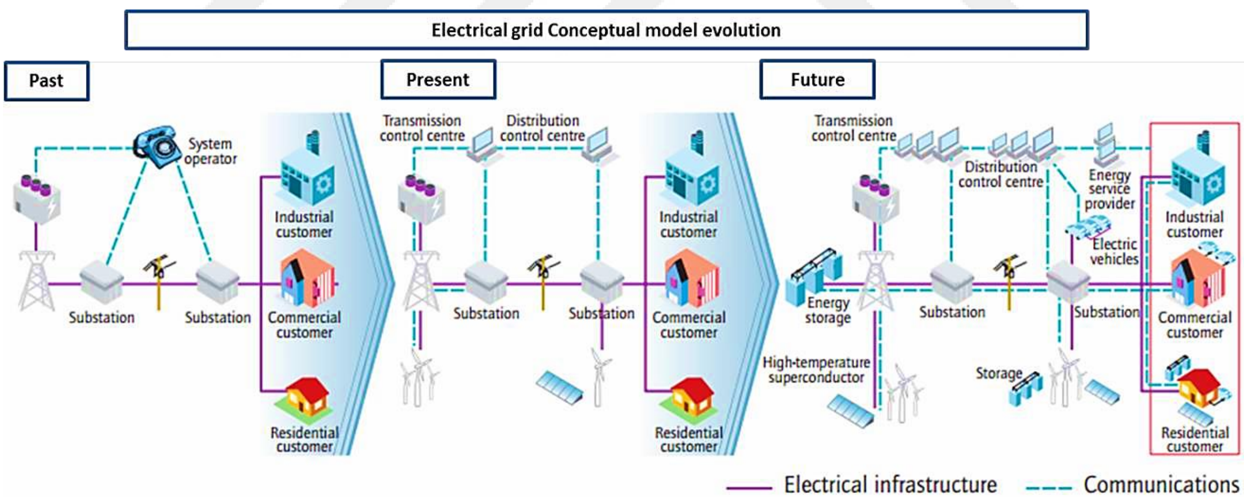

Figure 2.3: The smart electrical grid conceptual model evolution. Source [34]

Today internet of things makes our homes, offices and vehicles smarter, more measurable and this in turn makes our life more pleasing. Smart speakers like Amazon’s echo or Google home makes it very easy to set timers, get information on the go or be it to play music as mentioned in [35]. Through home security system, we can monitor our home from other places remotely and can see what is happening on the inside and outside or talk to visitors. Smart lighting systems can reduce unnecessary energy consumption as it stays off if there is no one in the house or smart air conditioners will turn on when it knows you are on your way to your home.

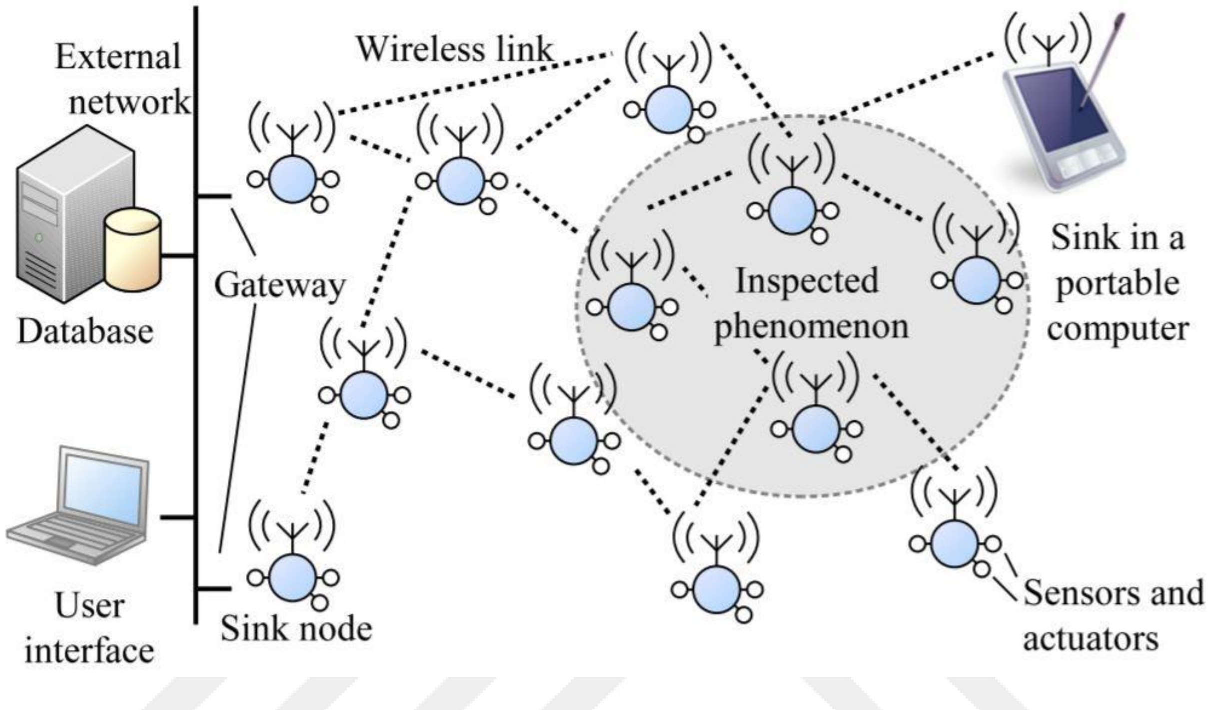

Figure 2.4: The linking of energy saving nodes in WSN network with external network and database [36].

2.4 ENERGY POLICIES AND ENERGY EFFICIENCY

In this section the main attributes of smart grid's performance and benefits are described, based the document from the Modern Grid Benefits.

2.4.1 Reliability, Quality and Security

The modern grid offers new means to control the transmission system, is able to monitor in real-time the grid conditions, performing self-assessments to detect, analyses, respond to, and if is necessary to, restore grid components as mentioned in [37].

Automatically diagnoses of the grid disturbances, and automated response to unexpected failures or attacks, by isolating problematic elements while preventing of so the rest of the area which is restored to normal operations as mentioned in [38]. The grid's faculty to carry out these selfhealing actions bring a better quality service as the interruption periods are reduced and so helping service providers to better manage the delivery infrastructure.

2.4.2 Ability to Accommodate all Types of Generation and Storage

Smart grids facilitate the connection and management of the distributed generation plants of different sizes and technologies as mentioned in [39]. From medium scale wind farms, to combined heat and power plant, or roof-top solar photovoltaic installations, the new grid allow to connect all the new generating plants, having the so called “plug and play” ability.
Also, the grid will be capable to include energy storage systems (batteries, hydrogen, compressed air, pumped hydro…) in order to optimize the distribution and reduce the price of the electricity. The grid would also manage the connection of the increasing number of electric vehicles (EV), which can also be a storage system.

2.4.3 Consumer Engagement

Consumers are called to play an important role in the new electric grid, they will have information and visibility to how much they consume and at which price in a real-time basis as mentioned in [40]. The increased interaction with the grid, the better options when choosing the energy supplier, the new forms of electricity pricing, such as time-of-use tariffs where energy costs more during the peak hours; are aiming to change the consumption patterns and behaviors bringing benefits by reducing the cost of the electricity and the environmental impact as mentioned in [41]. In a future, the grid itself would be able to reduce automatically consumer’s consumption when power is expensive or scarce. Up to the point, where utilities can ask the consumers to switch off some appliances but being compensated for doing so.

2.4.4 Optimize Assets Utilization and Operating Efficiency

A better visibility of the power flow on the network facilitates distributors to understand where losses occur and where to invest, cutting the costs of maintenance and reparation as the problem or inefficiency is specifically detected at the precise time as mentioned in [42].
The intelligence of the technologies used in the smart grids optimizes the performance of its assets. Increasing the operating performance when looking for the minimum cost system

operation, the control devices will automatically adjust to reduce losses while operating to an optimized capacity as close as possible to the full utilization of its assets.

2.4.5 Enable New Products, Service and Markets

Markets play a major role in the management of the grid, as they can control the energy, capacity, location, time and power quality as mentioned in [43]. The open-access (liberalization) market will provide market flexibility due to the deregulation and new competitive services will appear to offer better choices to the customer.

2.5 SMART GRID FUNCTIONS, CHARACTERISTICS AND BENEFITS

To have a clearer idea of a smart grid in this section there is an explanation based Smart grid conceptual model from the National Institute of Standards Technology which provides a description of the areas in which a smart grid can be divided as mentioned in [44]. There are seven differentiated domains as shown in Figure 2.5 include the bulk production, transmission, distribution, customer, markets, operators and service providers.

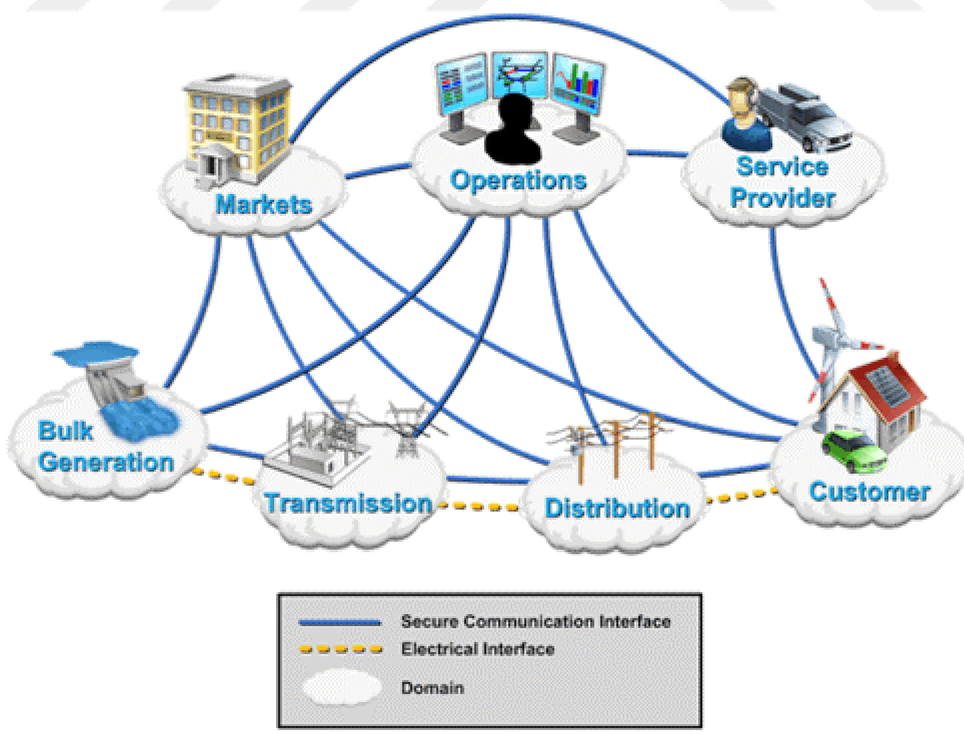

Figure 2.5: Conceptual illustration with all domains participating in a Smart Grid [44].

2.5.1 Bulk Generation

Bulk is referred to the large electricity production; this electricity can be generated from different types of energy and sizes of technologies, seen in Figure 2.6. Including:

Conventional non-renewable and non-variable power plants: nuclear fission, coal and gas plants

Renewable and non-variable: hydro, biomass, geothermal and pump storage

Renewable and variable: solar and wind

Once the electricity is generated it is connected to the transmission domain, and the bulk generation domain communicates to the Operations, Markets and Transmission domains as mentioned in [45]. The Information and Communication Technology (ICT) infrastructure let the grid to automatically reroute the power flow when some generators fail, also allows starting and stopping electricity production, allocating distributed production and storage devices to store energy for later distribution.

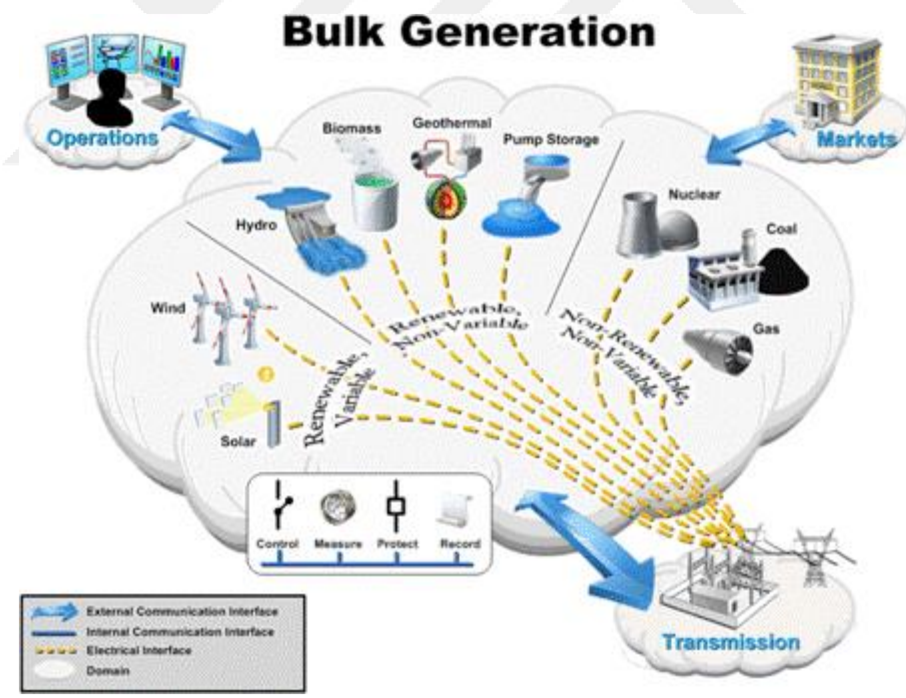

Figure 2.6: Deep look at the Bulk Generation domain [45].

2.5.2 Transmission

The transmission network operators have to balance the supply and the demand to maintain the stability of the grid as mentioned in [46]. Thus may contain distributed energy sources such as storage and peaking generation units with the capacity to manage and dispatch the energy when is needed. To do so, the Wide-area monitoring and control is necessary and it involves the generation and transmission areas, it is able to display, monitor and even optimize all power system components at real-time as mentioned in [47]. This technology gives a real-time view of the power system behavior, helps to avoid blackouts and facilitates integration of renewable energy sources. A set of equipment and systems are needed, these may include remote terminal units, substation meters, protection relays, power quality monitors, phasor measurement units.

A transmission enhancement is also needed; it is referred to the physical technologies to improve the transmission system, to name some:

- The Dynamic Line Rating (DLR) use sensors to monitor the carrying capability of a section of network, optimizing the transmission assets with risking overloads.
- Flexible AC transmission systems (FACTS) maximizes power transfer capability, improving transmission networks control.
- High voltage DC (HVDC) allows decreasing system losses when connecting remote wind or solar farms to large power areas.
- High-temperature superconductors (HTS): can improve the transmission performance by reducing its losses.

2.5.3 Distribution

Distribution domain connects electrically the Transmission domain and the Customer domain, also communicates with the Operations and Markets domains as mentioned in [48]. Important changes occur in electricity distribution from historically very little communications, hierarchical and unidirectional interfaces; to the nowadays two-directional distribution grids built with much more interconnection, extensive monitoring and control devices capable to allocate the distributed generation and storage, control the demand response and the load, improving its reliability.

The Distribution grid management uses active network management technologies on the distribution and sub-stations that improve the intelligence of the grid by allowing automation distribution processes by taking real-time information to identify the failure and reduce the interruption time of restoring the network.

Capacitors banks, sectionalizes, recloses, protection relays, distributed generators and storage devices, are devices included; capable to automatically monitor and control the voltage, compensate the reactive power. Systems used are regarding geographic information (GIS), distribution management (DMS), outage management (OMS) as mentioned in [49].

2.5.4 Customer

Customer is called to have an important and active role in the new smart grid, so its engagement is essential for the well-functioning of the smart grids. To do so, the grid system should provide the customers with the necessary instruments to develop this paper as mentioned in [50]. Customer domain is electrically connected to the Distribution domain; and communicates to several domains: Distribution, Operations, Markets, and Service Provider; see Figure 2.7.

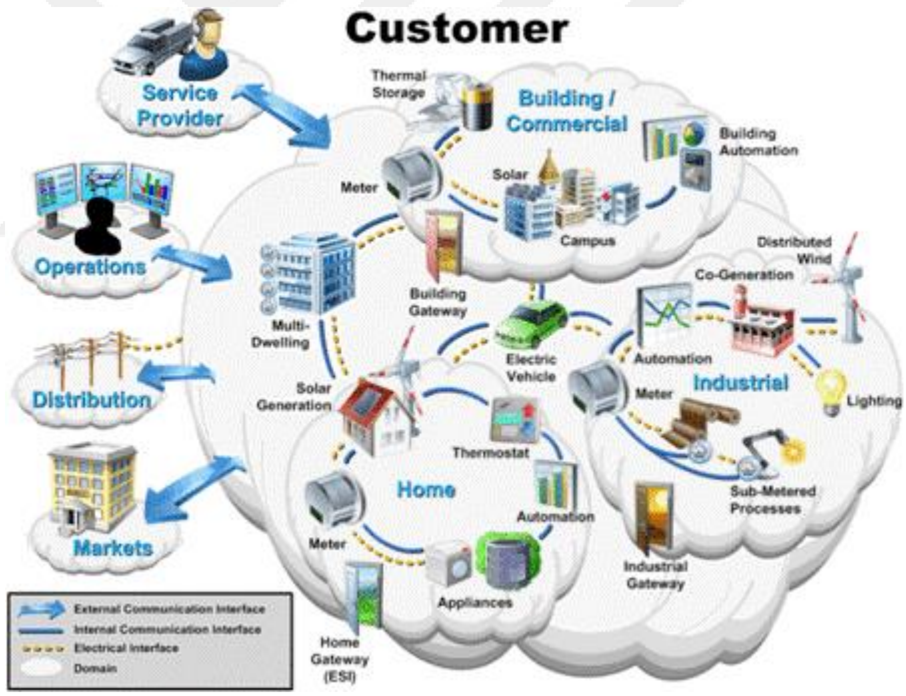

Figure 2.7: Deep look at the Customer domain [50].

A smart grid will enable the customer to manage not only the energy use but also the energy generation, providing control and information flow between the Customer and the other domains. There are three types of customers by size and specifications: industrial, commercial/broadband network and residential; however, all of them share a similar Advanced Metering Infrastructure, can allocate the electric vehicle charging infrastructure and dispose of similar customer side systems.

The advanced metering infrastructure (AMI) includes the use of different technologies, such as smart meters, communication technologies, in-home displays, gateways, severs… to a two-way flow of information providing customers and utilities with data on electricity price and consumption, including the time and amount of electricity consumed as mentioned in [51]. The Meter Data Management System (MDMS), within the AMI, enables functionalities like remote load control, monitoring and control of distributed generation, in-home display of customer usage and integration with broadband network management systems as mentioned in [52].

Electric vehicle charging infrastructure includes chargers, batteries, inverters; is expected to provide a capacity reserve and peak load shaving. With the possibility to schedule the smart charging (grid to vehicle-G2V) and discharging (vehicle to grid-V2G), it may suppose a positive effect into the billing by reducing the electricity costs and offering storage capacity for the surplus of energy when there is a low demand as mentioned in [53]. Implies a direct interaction with both AMI and costumer-side systems. Customer-side systems, refers to smart appliances, in-home displays, energy management systems and dashboards; that are used as a visual instrument to manage the electricity consumption at the residential, commercial and residential levels as mentioned in [54]. Making use of this technology offers the possibility to increase the energy efficiency and reduce the peak demand, together with the consumer manual demand response or even the automated and remotely controlled appliances response to prices.

2.5.5 Market

Markets are responsible to perform both pricing and balance supply and demand in the system. To do so interfaces with the generation, transmission, distribution and customer domains are crucial for this system matching, as seen in Figure 2.8. Due to the fact that generation not only occurs at a large scale (Bulk Generation domain), the increasing distributed energy resources (DER) will participate actively to the grid supply and balance as mentioned in [55]. The liberalization of the electricity market will help to customer to choose among different options, allowing to a fair trading and wholesaling electricity.

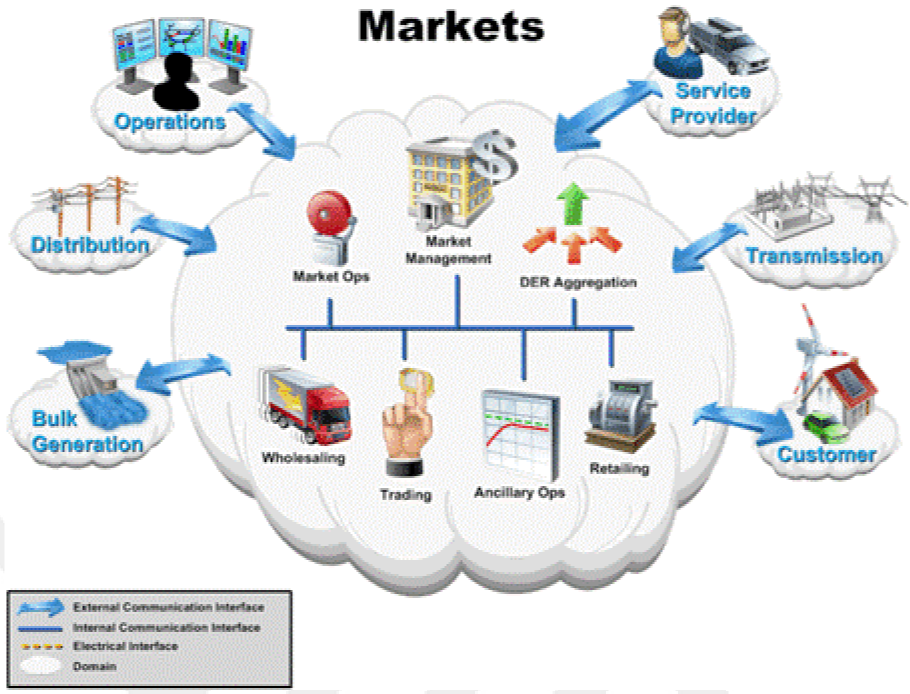

Figure 2.8: Deep look at the Market domain [56].

2.5.6 Operations

Operators perform the ongoing management functions, typical operations may include: network operation, network operation monitoring, network control, fault management, operation feedback analysis, operational statistics and reporting, real-time network calculation, dispatcher training as mentioned in [57]. They have the role to smooth the operation of the power system, many of these tasks are responsibility of the regulated utility such as the maintenance and construction, meter reading or security management; and some of them can be outsourced to other domains.

2.5.7 Service Provider

Service Provider includes the entities that provide services to electrical customers and utilities; performing support tasks to the business processes of power system producers, distributors and customers. The Service Provider communicates with the Markets, Operations and Customer domains as mentioned in [58]. Interfacing the Markets and Customer domains, Figure 2.9; will enable the appearance of innovative services that will bring new business opportunities in response to the market needs and requirements. These services may be performed by the electric service provider or by third-party in support to the new services necessity.

Figure 2.9: Deep look at the Servicer Provider domain [58].

These services range from the modernization and enhancement of the traditional services of billing and customer account management, to the appearance of new services like the home energy management or generation. The so-called Software as a Service (SaaS) will have a major impact in this area, as it is able to provide an easy and comprehensive front-end to the consumers with the information about electricity consumption and price.

2.6 CHARACTERISTICS OF INTELLIGENT HOME

Advanced metering infrastructure (AMI) systems enable detail and time-based data measurement, collection and transmittal to the grid actors as mentioned in [59]. Due to the installation of smart meters and whole AMI, the so called smart metering arises as one of the keys for the Smart grids implementation. And contribute to benefit the participants of the grid:

2.6.1 Anatomy of AMI

The Advanced Metering Infrastructure is composed by several elements from the inside of the customer broadband network until the utility. These components are: Home display units, Home Area Network, Smart meters, Wide Area Network, Data collection Head-end, Meter Data Management Systems (MDMS) and Data portal to stakeholder's systems as mentioned in [60].

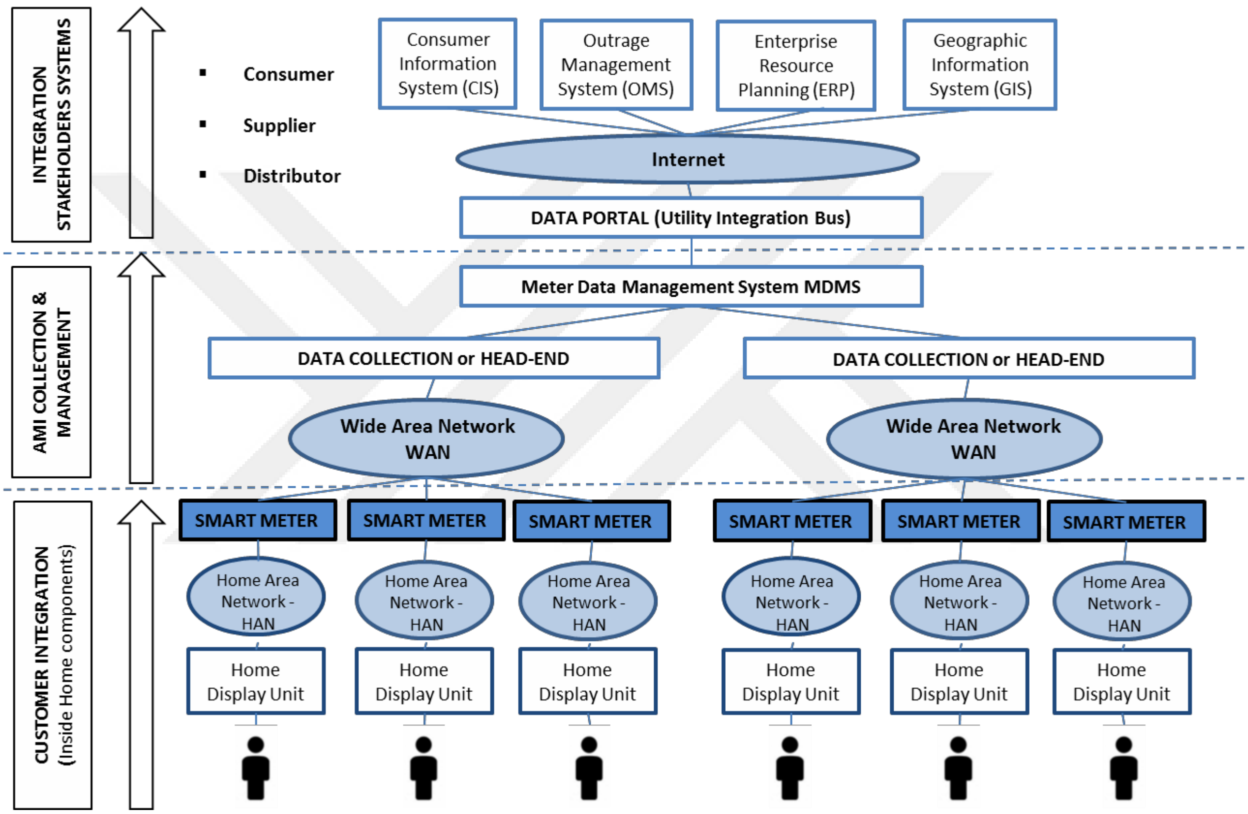

Figure 2.10: Advanced Metering Infrastructure diagram, from customer's home to stakeholder's services [59].

2.6.2 Home Display Units

Are the different interfaces with the consumer, where is displayed at almost real time the information about energy use (kWh), cost (€ or $) and even CO2 emissions. It can be different type of devices, specific equipment devices from a particular company that just has this functionality, or may be integrated in a computer through a website, an application in the smartphone or tablet as mentioned in [61]. These units offer the visualization of the data collected and emitted from the smart meter, in Figure 2.10 an example is shown.

Figure 2.11: Monitoring energy consumption in different supports [61].

2.6.3 Home or Local Area Network (HAN or LAN)

A Local Area Network (LAN) is a computer network interconnecting computers within limited areas, houses, schools, labs… and able to exchange data between them thanks to the gateways. The Home Area Network is a specific LAN that facilitates the communication between devices and appliances within a home limits.

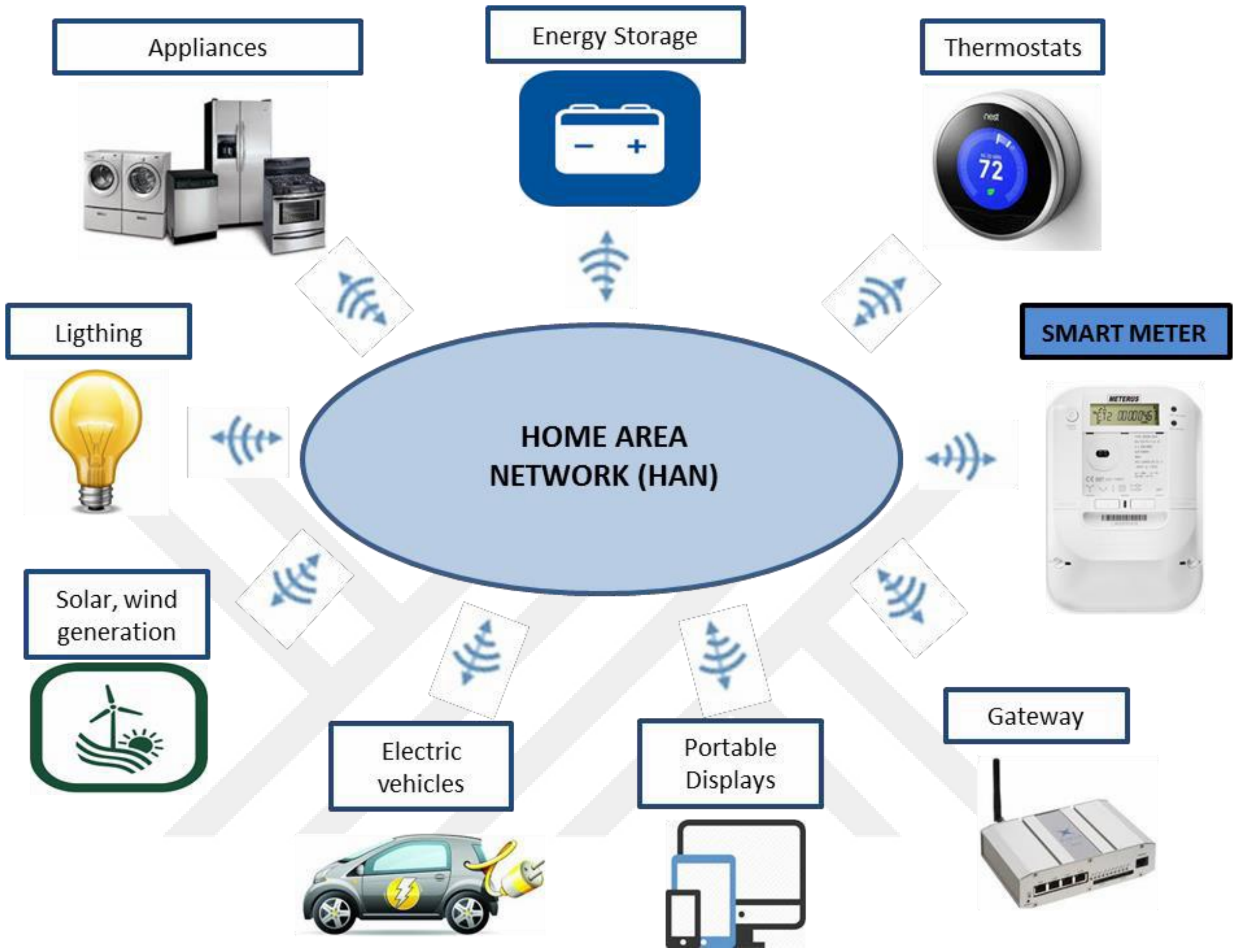

Figure 2.12: Conceptual model diagram of a Home Area Network (HAN) and features [62].

Figure 2.12 shows the smart devices (smart meters, computers, tablets), smart appliances (fridge, washing machines, washing, kitchen…), micro-generation from decentralized energy resources (DER) and others able to communicate inside the HAN as mentioned in [63]. The smart devices are already in our daily life, but in a future vision of a smart home all the appliances will be connected through different communication protocols, and thus, automatically controlled.

2.6.4 Smart Meters

Smart meter is the core of smart metering and AMI, their functionalities are far extended beyond the simple reading the energy consumer over a period of time, usually a month like the previous traditional meters do. The smart meter is capable to read the energy consumption remotely and with a much more reduced time interval, usually configured to hourly, half hour or quarter frequency as mentioned in [64]. And communicate this information to the utility and also to the consumer, due to its connectivity. However, smart meters are not exclusive of electricity, also there are water and gas smart meters available.

Figure 2.13: Smart Meter with Open Smart Grid Protocol [64].

2.6.5 Wide Area Network (WAN)

These networks are capable to cover a broader area, are the medium in which various groups of smart meters from different LANs communicate the data collected to the data concentrator or head end, as illustrated in Figure 2.14.

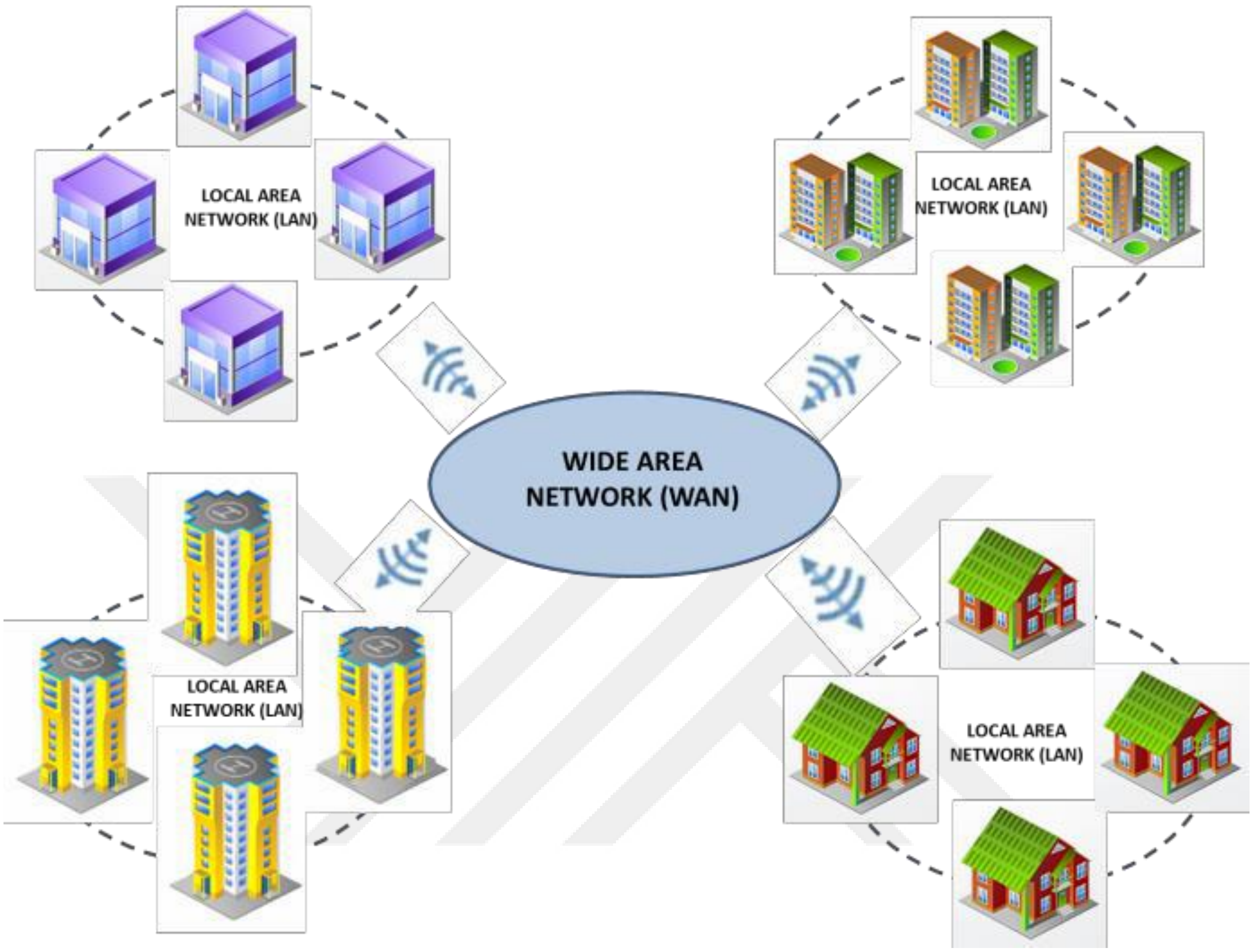

Figure 2.14: Wide Area Network (WAN) model, communicating and aggregating various smart meter measures from the Home and Local Area Networks

There are different ways to communicate the smart meter to other AMI components, to enumerate some communication technologies that may be appropriate:

Power line carrier (PLC): carries data across the power lines that deliver electricity

Broadband over Power Line (BPL): Also uses the power lines to communicate data, but requires concentrators and repeaters.

General Packet Radio Service (GPRS): uses the mobile's operators WAN to transfer data, but is not free as charges per amount of data transferred.

Radio: uses radio waves to communicate the smart meter readings

Wi-Fi: wireless connection to internet via an access point, it is low cost but needs high amount of power.

ZigBee: it is a low power and low cost, wireless protocol suitable for local coverage.

2.6.6 Data Collection System or Head End

The head end is the responsible to put together all data from a large number and disparate set of smart meters. To be able to do this the head-end system should use the same protocol as the smart meter as mentioned in [65]. A protocol is system of digital rules for data exchange within or between computers. Smart meters may use various types of protocols, such as Open Smart Grid Protocol (OSGP), Transmission Control Protocol and Internet Protocol (TCP/IP), IEC, ANSI.

2.6.7 Meter Data Management System (MDMS)

MDMS is a single repository capable to store massive quantity of smart meter readings, it gathers the data. Smart meters are able to record the interval consumption, so they produce a large amount of data, for example a reading per hour or every half hour can produce around 4.000-8.000 readings per year per single meter as mentioned in [66]. This is the whole process of data collection, communication and analysis within the MDMS infrastructure; upgrading its data value by turning it into information. Then, it reaches the customer or utility and gives them the opportunity to better manage either consumption or service thanks to the support that represents to have this data processed.

3. METHODOLOGY

The research is organized in various sections starting with a description and the philosophy of the research from which the electricity consumption data and house information is taken and used as input for the study; a remark is done in how consumer behavior can contribute to foster energy efficiency; then the objectives and goals are defined.

Moving to the more technical part, the procedure of cleaning the raw data is explained to have the processed data in order to start the data analysis sought. First, an exploration and visualization of the data was carried out to extract the first outcomes from the consumer's load profiles.

The core of this research focuses on the broadband network segmentation according to their electrical load profiles, various clustering methods are used and compared in order to find the most adequate consumer segmentation. Then, the house features and household characteristics are considered to find whether exists any relation between customer's segments and their house properties or not.

The data generated has increased exponentially this last decade, this data could contain valuable information but it needs to be explored. The electricity market is also living a data boom, for instance the electricity meter reading has moved from one read every month to host smart meters able to read electricity consumption every broadband network; this implies that each consumer will have around 100 measurements per year.

Store and manage this data is already a challenge itself, that is the reason why broadband networks and computer science have increased their presence in the energy sector; but also the possibility to add value to this massive amounts of data applying data mining and analytics to extract hidden information or at least be able to separate the useful one to the less interesting. Big data analytic tools are essential to add value to all this data, specifically in the energy sector could add value in terms of energy balance, energy efficiency or energy prediction; that could be profitable for both the consumer and the utility.

Consumer behavior and demand response can contribute significantly to foster energy efficiency in the residential sector pushing a behavioral change to boost energy savings. By analyzing different types of data, such as, consumers' load profiles, households' properties and householder's characteristics; some patterns could be found and knowledge could be discovered in order to guide the end-consumer not only how to reduce its energy consumption but also when is the best time to do so.

Specialized software programs and programming languages are used to perform such analysis involving large sets of data; for instance: customer segmentation by clustering them by load profiles similarities, finding correlation between electrical consumption and temperature, performing consumption data exploration and visualization or consumption prediction; among many others.

In the current research data analysis, the “Python” programming language is used, “Python” is a free open-source software that could be used for both experts and beginners who have the first contact with the programming languages as many on-line help material and tutorials are available, which facilitates the learning.

3.1 DATA ANALYSIS FOR ENERGY CONSUMPTION

The main objective is to discover knowledge from the electricity consumption data measured in residential broadband networks, in concordance to the idea to provide personalized energy efficiency recommendations to the consumers. We are aiming to stablish a methodology to automatically classify any new user into load pattern group by looking at the household and house characteristics.

This process consists in three phases described below, and graphically shown in Figure 3.1:

- Phase 1: From the data already existing, the current users are clustered/segmented into different groups according to the similarity of their electrical load profile. Similar load profiles mean that the consumer behavior and habits are almost identical.

- Phase 2: Once, the users are placed in one cluster or another, a further analysis is carried out in order to find the more likely properties related to the house and household of each group.

- Phase 3: Classify any new user to one of the existing segments by looking at the household properties and the house characteristics.

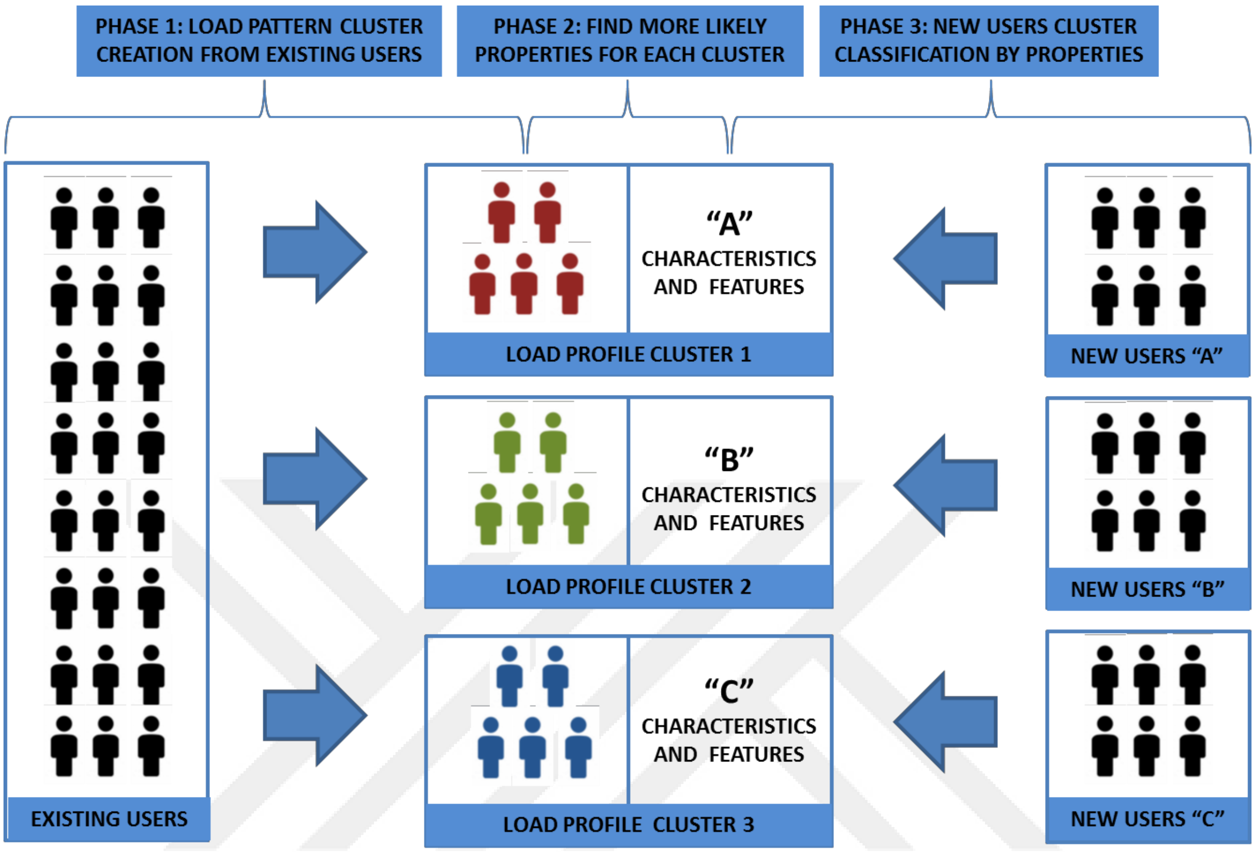

Figure 3.1: Diagram that describes the phases of the methodology applied in this study.

Also, the household and householder's characteristics were analyzed in order to find the most common attributes for the whole set of consumers and inside each group; looking for the existence of any pattern or relation when crosschecking load profiles groups and house characteristics. Likewise, additional actions were done in terms of exploring and visualizing data in order to be incorporated into the newsletter or monthly reports that are sent to the users periodically. And in parallel to provide to the municipality the starting point for deeper data analysis that is currently willing to perform.

3.2 DATASET DESCRIPTION

A dataset is composed of rows and columns, where, usually, each row represents an observation and each column a different variable. This data can be either qualitative (categorical) such as gender, country, city; or quantitative (numeric) such as electricity consumption, height, price, distance. Again and as said before depending on the type of data the approach used to the analysis may vary, because is no treated the same way numeric and categorical data. The size of this dataset is relatively manageable due to mainly three facts.

The number of sub-meter installed are a small amount accounting for different sub-broadband networks.

The measurements period is in most of the cases around year length

The data was aggregated hourly, fact that reduced the final number of observations and therefore the dataset's size

Electricity consumption data for broadband networks

Households and householder's information based on broadband network usage.

3.3 COMMUNICATION VIA BROADBAND NETWORK

Communication via broadband network solves many problems that directly affected the data quality, as if the broadband network is switched on, communication can be stablished interrupting the sending and reception of data. Then, the data for this time period is represented by zeros as it was no consumption; thus altering the dataset values which need to be eliminated in order to perform an accurate data analysis.

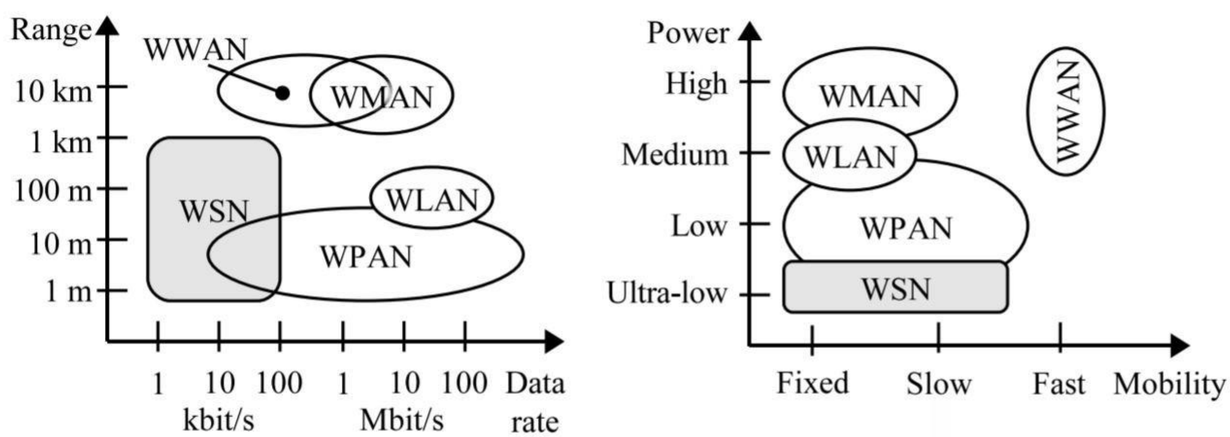

Figure 3.2: Diagram that describes the different networks with data transfer rate.

From the database server the Software as a Service (SaaS) company, in this case, access to the data stored and proceed to upload it into the web-app platform for the visualization of each user. Finally, the data used for the current analysis was obtained through an Application Programming Interface (API) configured by this research in order to:

First, aggregate all the data hourly. The sub metering equipment is able to measure the consumption in high-resolution frequencies, each 5-15 minutes according to the device. For the aim of this study, the hourly aggregation is enough and a good starting point if a further deeper study is sought in the future.

Second, all the household's consumption data are in the same file, fact that facilitates the analysis procedure as reading one file is enough to analyses the data.

Hence, by using the "python programming" that permits the access to the database server a ".csv" file is downloaded, this file is the input data regarding electrical consumption that will be used for the data analysis. This data is considered time series numerical as it records the electrical consumption and time of it.

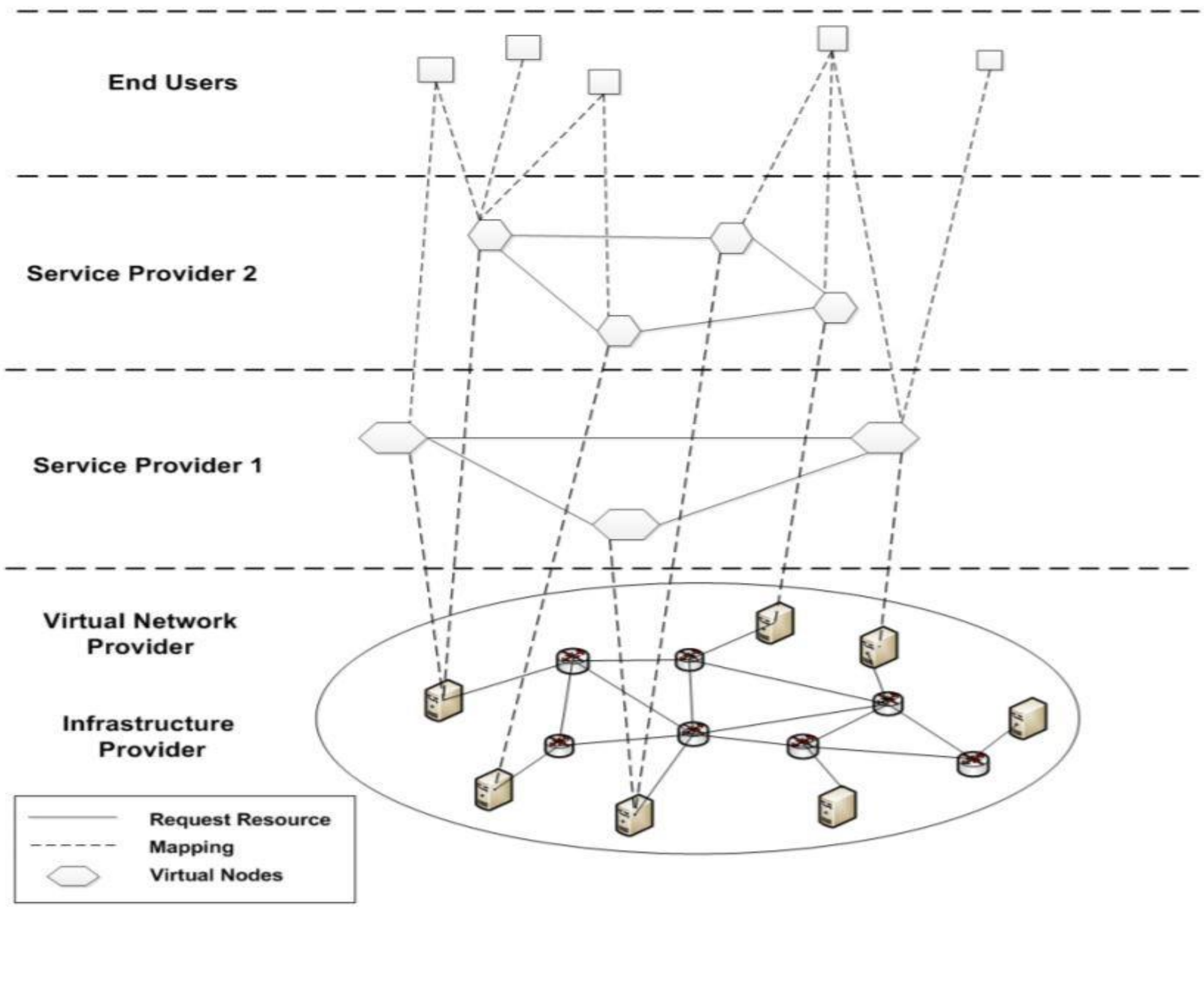

Figure 3.3: Diagram that describes the layers of broadband service provider network in this work.

The “.csv” files have the characteristic that are easy to read by software prepared to conduct a data analysis, for instance “MS Excel”, “Python” among others. “Python” is the programming language used to carry out the current analysis.

ENERGY CONSUMPTION IMPUTATION & CORRELATION

The data regarding the electricity consumption is the main one used in the later analysis; it is numerical data as its records the measurements of electrical consumption for every household participating into the research. The following Figure 3.4 summarizes graphically how the electrical consumption data used for the analysis is obtained:

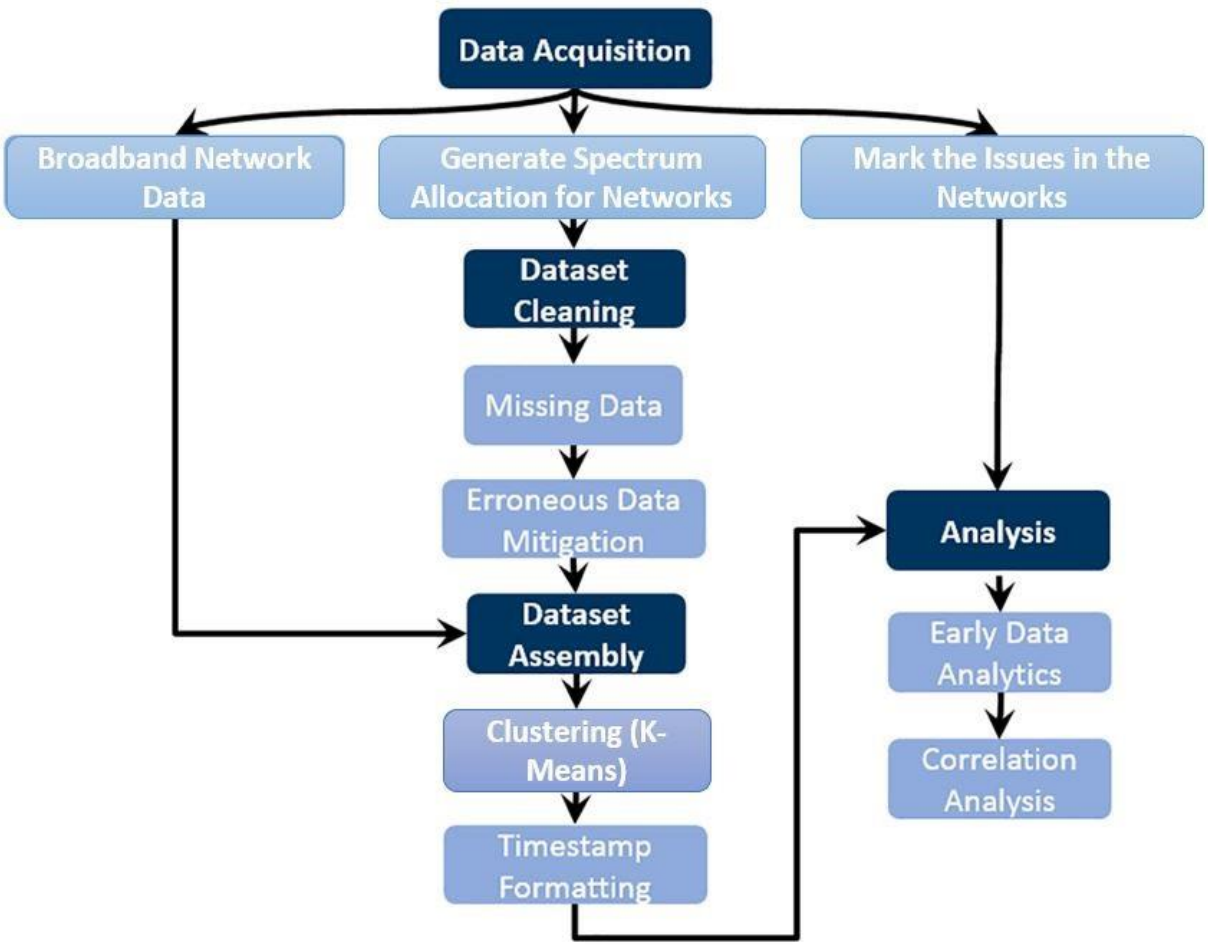

Figure 3.4: The flowchart of approach being followed.

Although this data is processed and presented in a tidy format, it needs to be cleaned. Cleaning the data means detect and delete all the possible elements that could affect to the well-being of the analysis; in this particular case:

First, we need to select the sub-meters that are currently being used in the research; as the column under the label contains both current and old sub-meters identifiers. Hence, this step will consist on selecting only the valid sub-meters that are currently working and have consistent data.

Second, from the selected sub-meters in the first step we should detect and delete those periods of data that contains “energy” as a measure. The “energy” values are quite common due to data communication system used is the Wi-Fi or router of the house, so when it is switched off no data is send and stored.

Third, detect and select those periods that present the so called “frozen” values. The “frozen” values are the repetition of the same value for consecutive observations; this is due to a hardware problem during the data communication which froze the measurement and repeats it over a time.

It is essential to do these three steps to have a valid and verified output of the analysis and to obtain more accurate results; otherwise the as the input data is somehow “corrupted” it may affect the analysis.

After the first step, the sub-meter’s dataset was reduced from the initial different sub-meters to multiple meters and accounts for a total of 7 cluster observations. The discarded sub-meters are due to facts that don’t add value to the analysis, such as:

No data was stored

All the data stored were “zeros”

Only few observations (less than a month) were stored

3.4 BROADBAND NETWORK DATA VISUALIZATION

Before being able to do the broadband network segmentation, knowing the data that will be analyzed is essential to get a sense of what is in the dataset, which particularities and problematic points could be found to deal with. Getting familiar with the electricity usage data and also with the “python programming” commands that will be used when treating this data.

When exploring the data, three aspects should be taken into account:

First, regards the dataset format as in most of the cases it needs to be reshaped into the best format to have the desired graphical output.

Second, the dates' treatment is a challenge when exploring the data, as time series data needs further and different approaches. Especially, if division between weekdays and weekends, or monthly distinction is needed.

Third, focus on the specificity of the electricity use data, studying which is the best path to extract the information or conclusions sought.

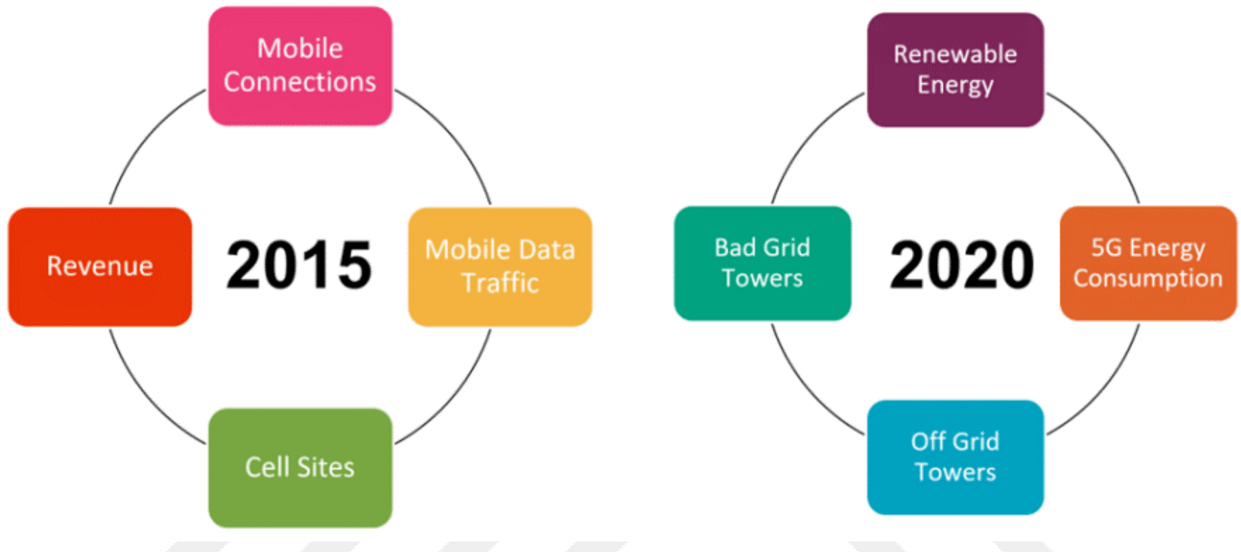

Figure 3.5: The badging of different between the broadband network from 2015-2020.

The research work is to organize and visualize all the data using different representations according to the specific goal of the visualization for energy consumption and spectrum allocation in the broadband network. It is interesting to provide a tool that allows the visual representation for each individual customer and satisfies all the purposes sought. This facilitates the comparison among consumers, permits to access easily to the individual electric consumption main characteristics that, for instance, facilitate a quick audit on its consumption and detect possible irregularities or problems.

Among these purposes are:

Differentiate the weekdays and weekends load profile, are compare them to the load profile output accounting all days not differentiating weekdays and weekends.

Differentiate the load profile for each day of the week, to know if among the weekdays the load profiles are similar or not.

Representation in absolute values, percentage of consumption per hour, accumulated in a period of time

Study if there is a consumption difference among months and seasons

Find the characteristic load profile per each month of the year

Be able to specify exact dates and plot the consumption at that period

Using different kinds of visualization graphs shapes (bars, points, lines, boxplots…) and colors to facilitate the information visualization

Facilitate the comparisons between users by using interactive graphs

3.5 CLUSTERING ALGORITHM FOR PREDICTION

The K-means solution adopted was to cluster the data in 7 groups by using the python programming language, but as said before the applying the algorithms could also be chosen. Due to the fact at this stage is not possible to stablish big differences among spectrum allocation of broadband network. The number of members assigned to each cluster, range from the minim different number of cluster for energy consumption data; so a distributed user's allocation is also achieved; avoiding clusters with 1 to 5 members which could mislead the analysis. Table 3.1 shows the within distance clusters where, as expected, the clusters with large number of networks present a higher within cluster distance.

Table 3.1: Number of members per each cluster defined by K-Means algorithm.

K-Means Algorithm on Energy Consumption Data							
Cluster number	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6	Cluster 7
Number of Networks	9	17	8	7	37	8	35

Cluster 1: there are two peaks, morning and evening both of them represents around 7% of the energy usage. It is similar to the cluster 6, although this cluster one presents a higher proportional values between peaks that ranges from 3-5%, it could mean that some activity is done for broadband networks.

Cluster 2: Presents a not sharp morning peak and a low consumption percentage during day of around 3% similar to the night hours' proportion; this means that there is no activity at home during the day. The peculiarity of this cluster is the there is a night peak at 00:00. The inhabitants of the house are staying longer at night, usually the young people or young couples are associated to that type of load profile.

Cluster 3: Presents a high morning peak, around 9% of the daily energy use is consumed in that peak that is around 8:00h. Then, there is no activity as the proportion of energy use in the afternoon decreases down to 3%, similar at the night percentage, and there is a small evening peak of around 5% at the dinner time.

Cluster 4: It presents two peaks prolonged in the morning (9:00) and in the afternoon (19:00), following the schedule of a small business or local, also the night consumption proportion is so low. So, it seems it is not a residential consumer.

Cluster 5: This cluster presents a flat profile, with the higher values during the day at lunch and dinner time. It could mean that people stays home during the day.

Cluster 6: Presents two sharp peaks, one in the morning and other in the evening, that reach almost 7% and 10% respectively. Then the energy usage during the rest of the day and during the night is so low, constant at 2.5%, so no activity during these periods; it means that the households are not home during the day; could correspond to working people with children.

Cluster 7: Presents a significant evening peak around 21:00h of almost 9%. There is also a small morning peak, almost no noticeable which does not reach the 5% of energy usage. During the days it also presents a continuous consumption of around 4- 5% until the evening peak.

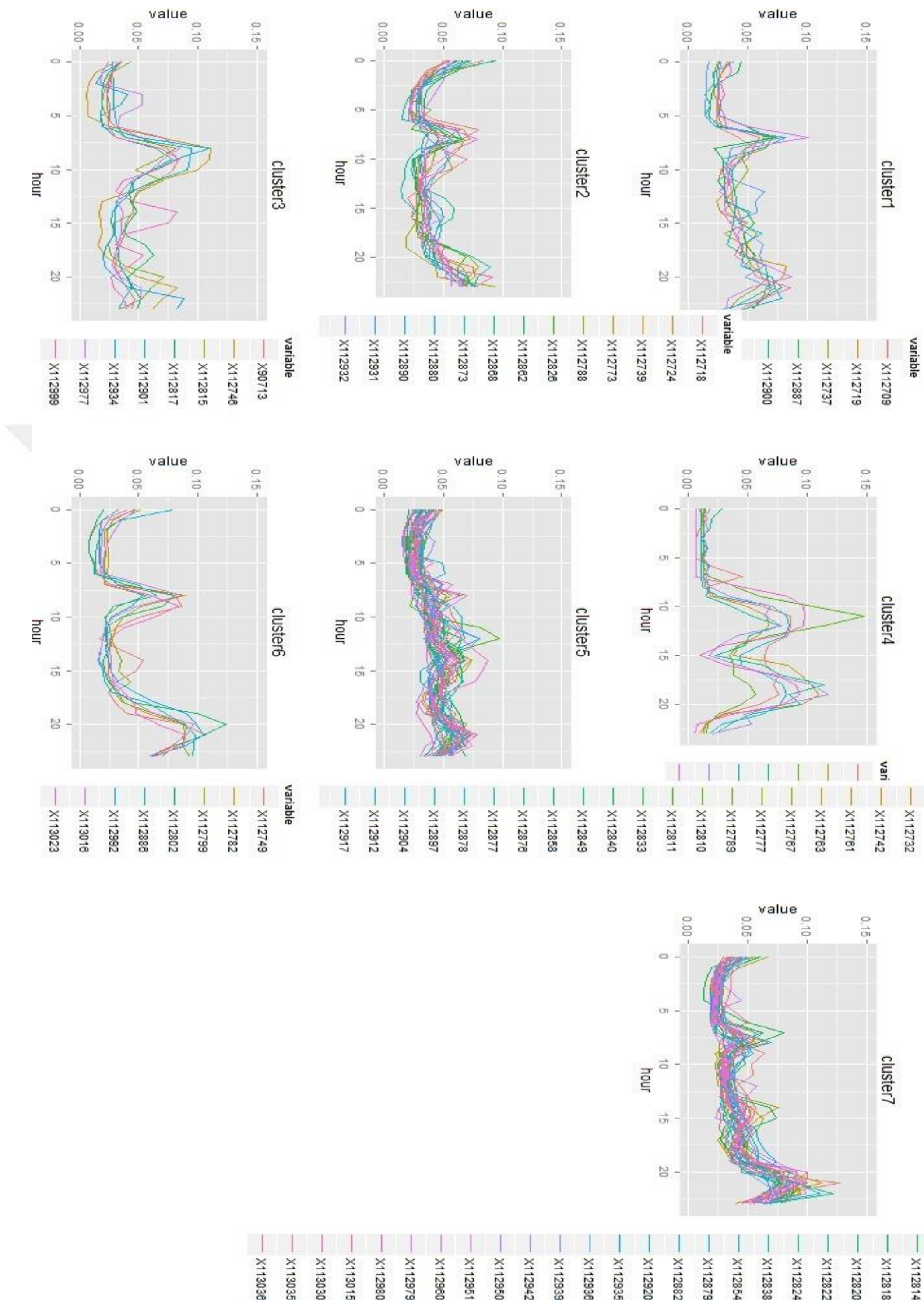

Figure 3.6: Plotting all the networks of each of the 7 cluster, to see the variability and check the clustering results from the k-means clustering

4. RESULTS

The electrical consumption is analyzed in different levels of detail. The procedure followed starts from analyzing the whole gross data and gradually moving towards a more detailed data analysis by segmenting and grouping the data for a more accurate study. Two kinds of values were used to plot the electrical consumption; the absolute units of power consumption [kWh] and the proportional hourly energy usage in percentages (%) were used. When analyzing each consumer individually the important values are the absolute units in terms of kWh; but to fairly compare various consumers' profiles is necessary to create a common normalized framework where to contrast the curves' shape; this can be done with the hourly percentages of energy consumption. As said before, depending on the analysis' goal one parameter or another would be used. So, although the possible approaches to the consumption data an endless, below different representation ways are shown that would be useful to know and better understand the consumer. This section uses various consumers for the different representations, in order to see variety of load profiles and also test the validity of the illustrations. Remark that, the scope of analyzing the features for each cluster would have been more adequate if a larger sample with better data quality and more specific features were available. So, the current process could not add any value to the research; but it is a procedure that could be followed when a larger dataset would be available.

The load pattern for each customer is obtained by considering all data available, without making distinction between weekdays and weekends, for the period selected. In the case presented in this work, the load profile in terms of hour proportion of energy usage of the broadband networks is shown; this user has the highest in the evening and another peak of consumption at lunch time. In this work, the load profile in terms of power of the broadband networks is displayed, and it can be seen another evening peak. These illustrations are useful to get a sense of the user's load profile in both absolute and percentage terms. The load profile represented was obtained by calculation the mean for each hour of the day input all the days available in the dataset. Among the multiple possible representations, the circular one illustrated in this work is another way to represent the load profile giving more emphasis on the hours that present the higher and lower consumption for spectrum allocation in broadband network.

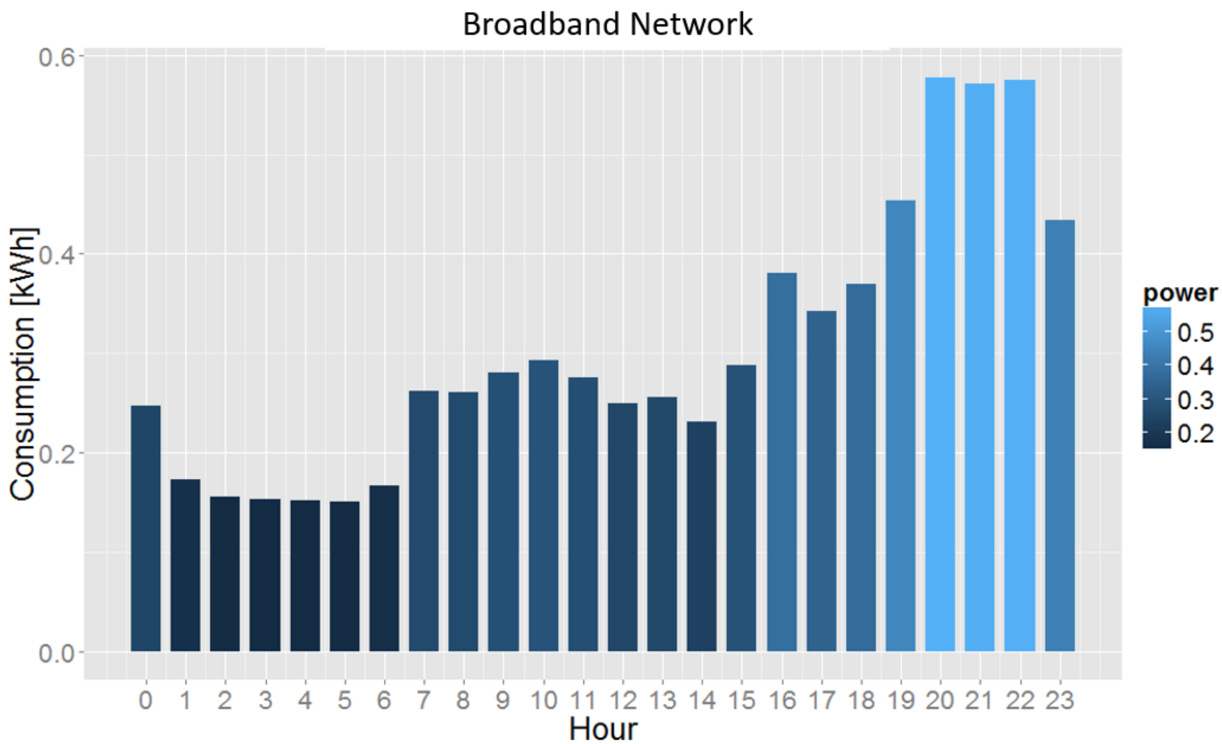

Figure 4.1: The broadband network consumption expressed in hourly absolute consumption units (kWh), in graph bars.

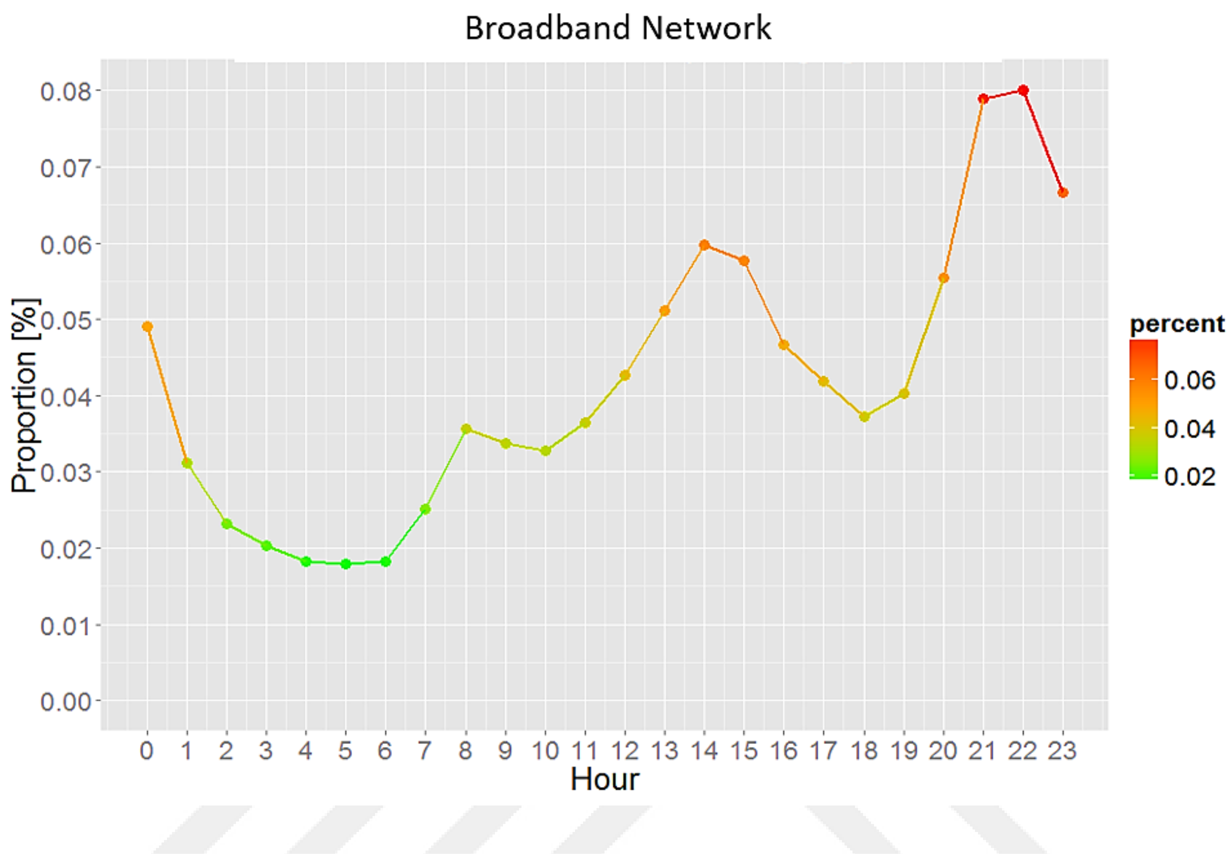

Figure 4.2: The energy transmission and consumption expressed in hourly proportion (%), represented in points and lines.

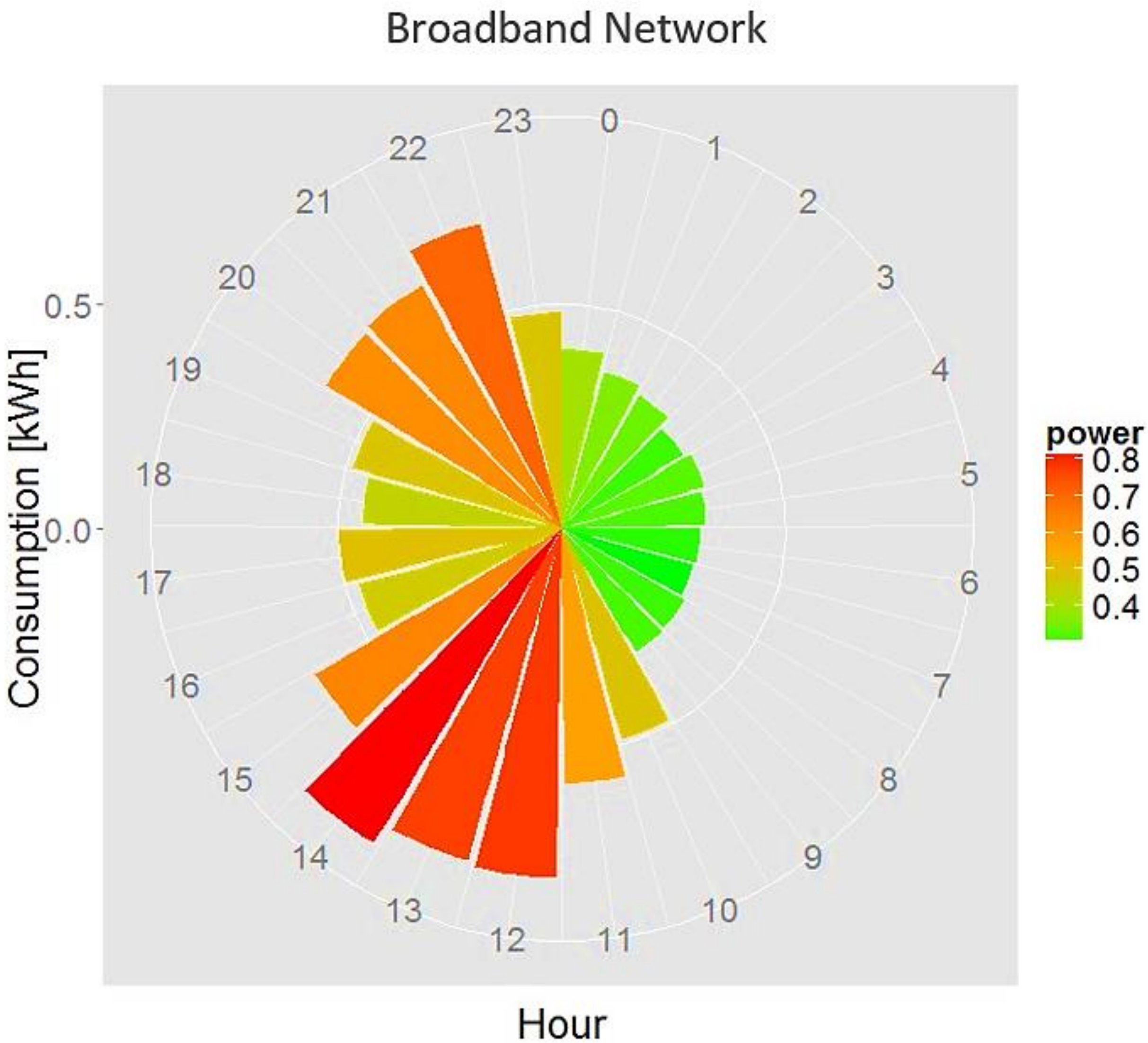

Figure 4.3: The load curve expressed in hourly absolute consumption units (kWh), represented in bars in a circular plot.

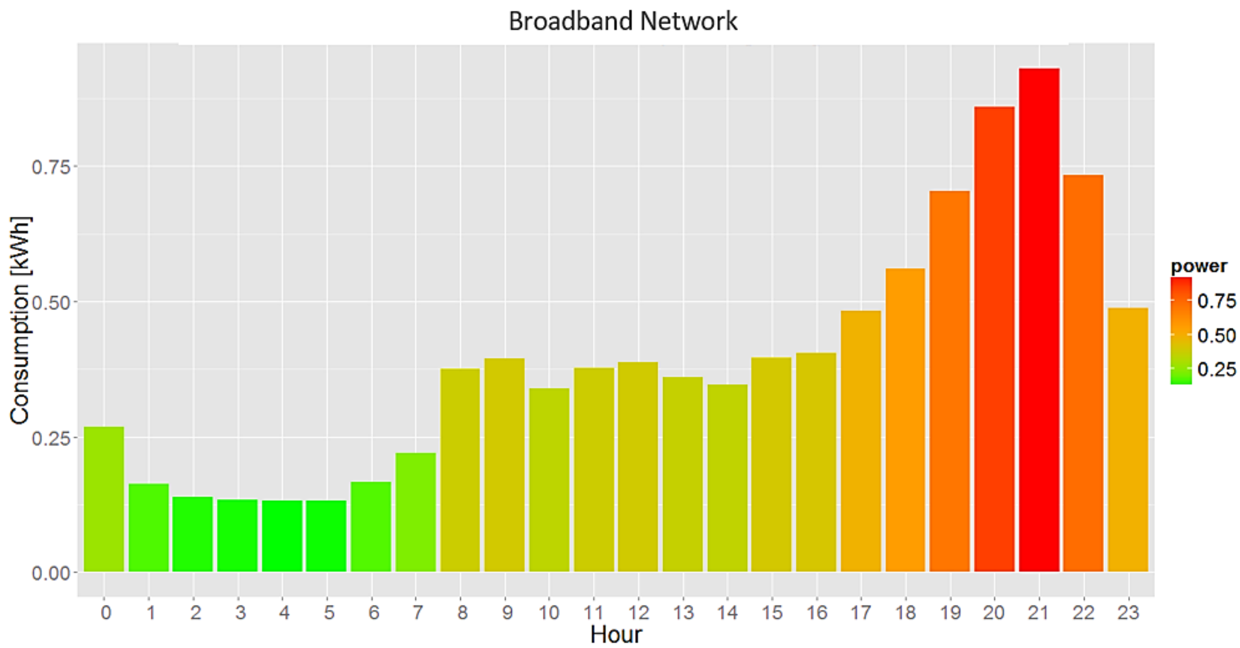

Figure 4.4: The broadband network energy load consumption expressed in hourly absolute consumption units (kWh), in graph bars.

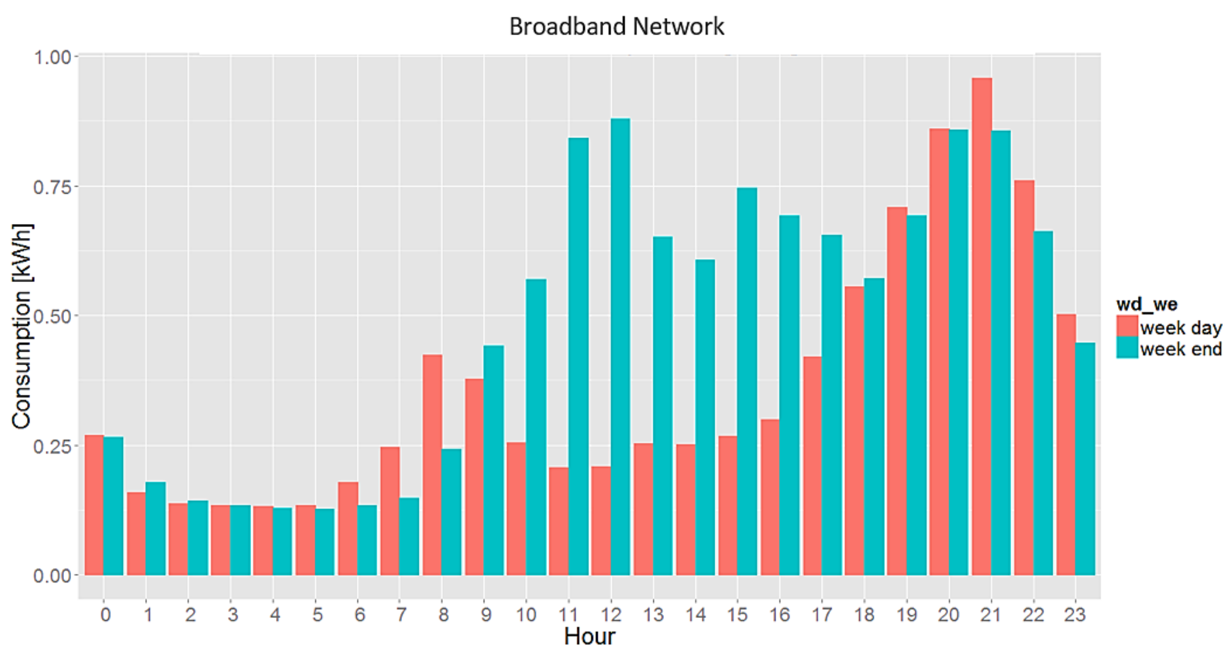

Figure 4.5: The prediction of energy consumption in weekdays (red) and weekends (blue) load curves expressed in hourly absolute consumption units (kWh) in graph bars in the same plot.

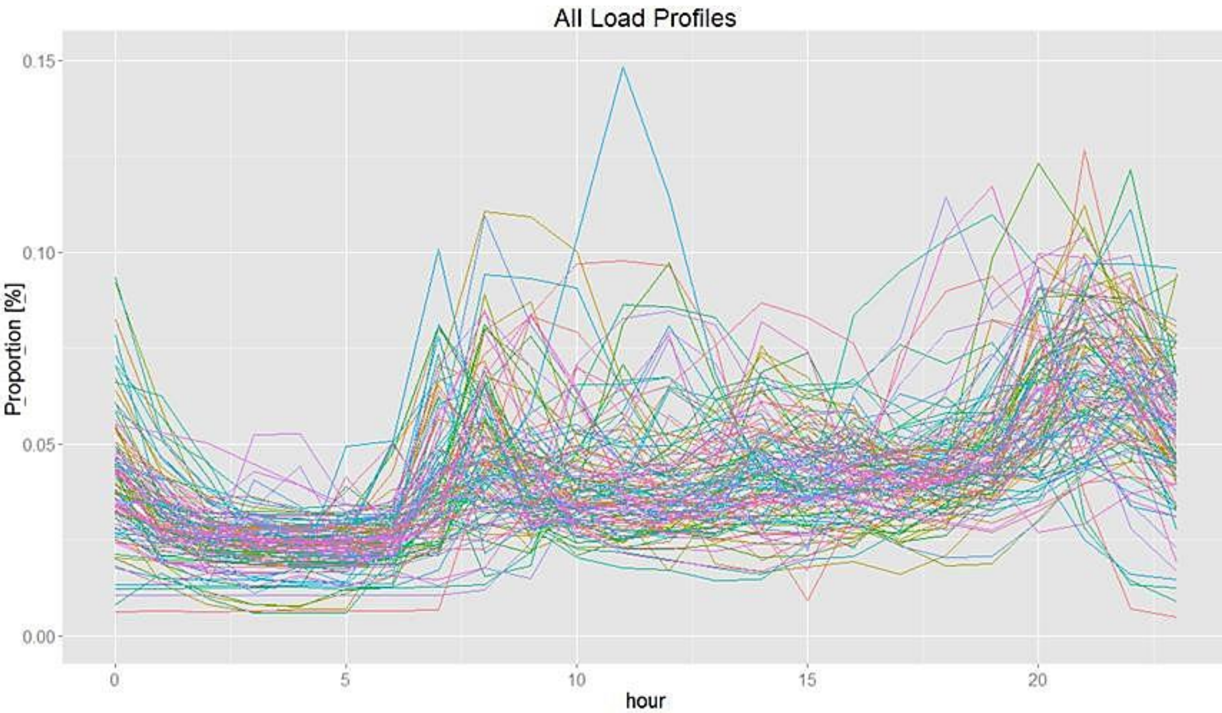

Figure 4.6: Visualization of all broadband network loads profiles in terms of energy usage proportion.

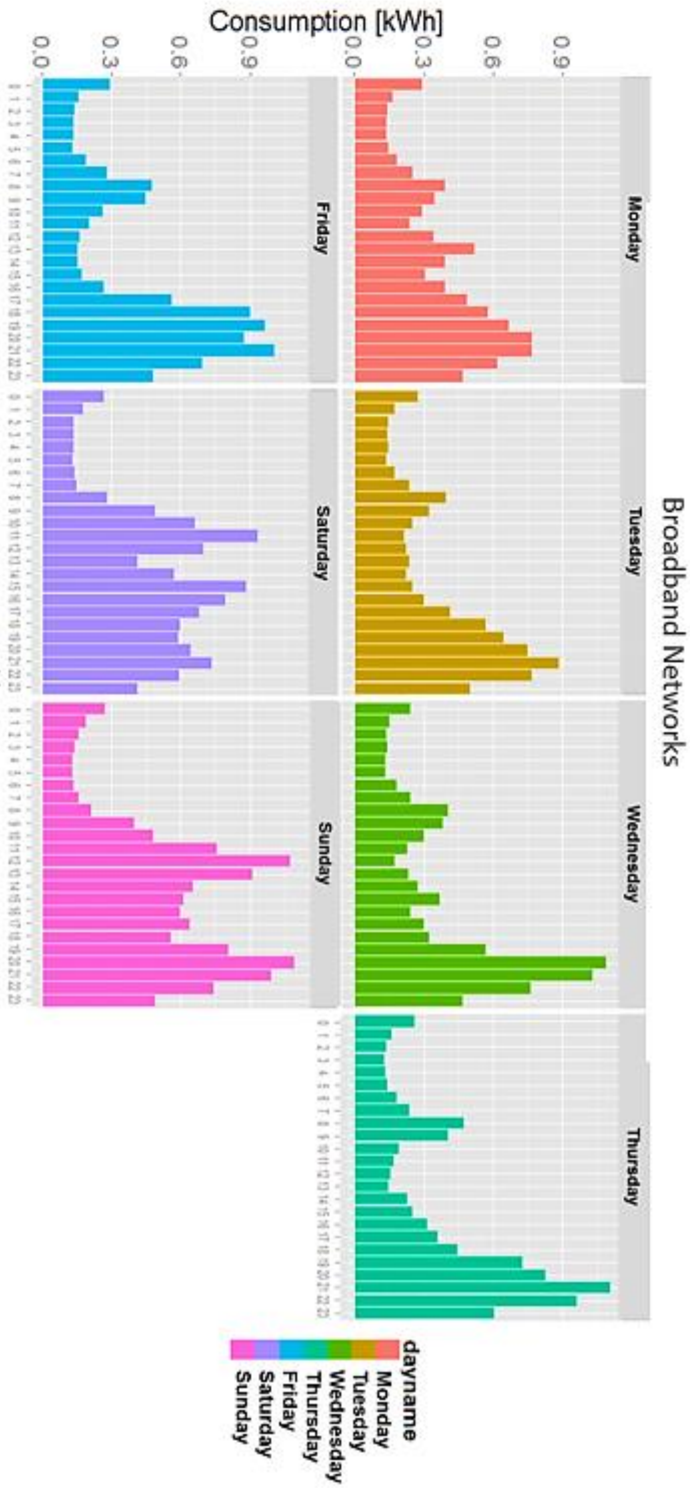

Figure 4.7: The broadband network in days of the week load curves expressed in hourly absolute consumption units (kWh) and represented in bars.

The aim of the cluster analysis is to group a set of observations based on one or various properties that make them similar to each other, hence the objects inside the same group will be more similar to any other object placed in other groups. The properties to group or cluster the observations are set by the analyst according to the aim of the grouping; this means that the same set of observations can be clustered in many different ways; depending on what is considered similar, it is a task of the analyst to define the appropriate similarity/distance metric.

In addition to that, there are so many clustering algorithms and techniques to group observations, each one may also lead to a different grouping result. The selection of the appropriate technique is essential to perform the desired clustering analysis; also the typology of the input data to be clustered is the main factor that will make choose one technique or another, taking into account that not every algorithm is valid for a specific problem. Finally, again the analyst should be able to create an easy visualization of the results and also interpret and communicate the finding to the easy understanding of the audience.

5. DISCUSSION

Energy management tools are not extended within the residential sector, since the energy consumed is so small compared to business and industries, the savings obtained don't pay-off the investment. However, there is room for providing such tools to residential consumers, for instance by creating a community as broadband networks are doing with the collaboration of public administrations and electricity utilities. This research is an example of a public administration pioneer initiative to engage the consumers and foster the energy efficiency among them, aiming to provide energy knowledge, understanding and guidance to the user in order to reduce its consumption on average up to 10%. Nonetheless, it was seen that providing the technical consumption data in kWh to the householders has limited influence and effect, so is necessary to go beyond and translate the technical information into call-to-action measures, guiding the user to smart energy choices (like a GPS), have a better response from the householder. Also the need of sub-metering devices to obtain the consumption data limits the scalability, as the cost associated is significant. To be able to reproduce this kind of research in a large-scale the access to the data from the utilities smart meters is necessary. So, the combination of the smart meter deployment and the big data analytics are called to play an important on the energy sector. The smart meters generate large amounts of raw data that need to be managed and, once analyzed can be converted into useful information that benefits both the utility and the consumer, as aim to improve the customer engagement and the quality of the service. Creating new business opportunities, mainly related to data science and data analysis in response to the market needs.

A brief description of each cluster features and possible correlation suppositions are presented below:

Cluster 1: big broadband networks, probably occupied by a couple of networks with little sub-networks or a couple with a pair of grown network links considered sub-network in the analysis. Not using electricity for heating and cooling as no air conditioning neither electric heating are installed. The power contracted is the highest one among all the clusters, which could be related to the broadband network spectrum.

Cluster 2: relative small networks, low power contracted below 5kW, it may be sub-network with no broadband networks as no links are in this cluster. The occupancy of small broadband networks could explain the peak at night.

Cluster 3: relative small and modern broadband networks with low spectrum value, the spectrum is allocated for whole network. No electric heating, but there are some electric appliances could be used with the help of broadband network.

Cluster 4: almost the same number of networks and sub-networks in this work, with a spectrum od lower than 100 GHz, built for high energy consumption networks. It could be families with small sub-networks as the number of networks in this cluster is higher than the other clusters.

Cluster 5: almost only adults are presents in this larger number than in other clusters (2-4 sub-broadband networks), no links, no spectrum issue, less electric consumption, different schedules.

Cluster 6: The new broadband networks could perform better in taking low energy, it may be occupied by links with small nodes of network. No electric consumption, but with air conditioning and electric consumption as the issues with the broadband network. The time of waking-up and bring the broadband network to the up-state could cause the peaks.

Cluster 7: There is no data available as the initial aim the research was devoted to broadband network. The technical audits and the on-line info-house are not yet adapted to high spectrum.

It was possible to detect any pattern or clear evidence that could relate some of the features analyzed to be the cause of a load profile in broadband network pattern or another. Issues, such as; the small dimension of the sample studied, the number of missing values ("NA") in it, and the fact the features available were not the most suitable (adding i.e. type of employment, people home during the day…); made not possible to reach the desired output when analyzing the features in each load profile cluster.

Most of the home appliances features are common to each cluster, with energy usage data in the broadband network, and not available data for the domestic hot water system. Also it does not exist any significant difference among the house properties and the sociological features in the different clusters, in order to be able to affirm that one or some specific features are the cause of the load profile pattern.

Remark that, the scope of analyzing the features for each cluster would have been more adequate if a larger sample with better data quality in broadband networks and more specific features were available in the result section of this research. So, the current process could not add any value to the research; but it is a procedure that could be followed when a larger dataset would be available.

6. CONCLUSION

Data quality is the key to perform a proper analysis and extract reliable conclusions from it. In this study, the dataset was not the optimal neither quantitatively nor qualitatively. A much larger dataset, for instance counting for thousands of users, would have been more adequate to the purpose of the study than the small sample studied of only users. In addition to that, the electrical consumption data from the sub-metering equipment presented some issues that difficult the analysis at some point and demanded a pre-clustering stage for treating and cleaning the data. As the communication between the sub-metering device and the server is done through the broadband network of the consumer, when it is switched-off the consumption data is lost and counted as high spectrum allocation like if it had been no consumption. So, when the possibility to access to the hourly smart meter consumption data becomes a reality, the current study purposes will find a suitable framework since the numbers of users could be much higher and the data quality should be almost perfect as are the measures from the electricity utility to bill their customers. The methodology used in this research is simplified compared to the procedures of the referenced papers, as larger datasets and more complex parameters are used. normalize the electricity consumption considering the weather and seasonal effects, although for the present study there was no temperature data available. The K-Means clustering technique was tested and compared, giving best prediction of accuracy 98.46% for all energy load profiles with high spectrum of 100GHz. The solution from the K-Means clustering is the one that better adapts to the segmentation sought, which is used as the base of the post-clustering stage to obtain the final energy consumption segmentation. This normalization is advisable as it mitigates the electricity consumption correlated to the temperature; however, this is more appropriate for aggregated daily consumption. But for instance, the electricity consumption is not normalized with the temperature. Most of these methodologies use the absolute values in 100 kWh, as they were more focused on identify the users with higher energy savings potential. But for the segmentation purpose, the decision to use percentages of energy usage fulfils the expectations sought. Providing consumers segments similar to the load profile archetypes from the power study basis of the current study (i.e. Evening peaks, double peaks, spectrum allocation, consumers). There are some similarities among the methodologies, mainly regarding the pre-clustering phase of cleaning the data, and also regarding the clustering phase as the K-Means algorithms.

6.1 FUTURE RECOMMANDATION

Consumer segmentation allows the utility to understand individual and groups of broadband network offering tailored services, as not all the spectrum allocations are equal, so each of them needs different personalized energy-savings recommendations at a different time of the day. By providing that, this will improve the effectiveness of energy efficiency programs and also it will have a positive effect on the energy reduction recommendations. As well as, improving the effectiveness of the demand response programs aiming to shift the time of energy usage to avoid peaks or reduce the peaks duration for broadband networks with high spectrum allocation.

We would use the broadband network to a larger and complete dataset using the deep learning, for example using the utility smart meter data. More advanced deep learning techniques could be used in order to improve some stages of the procedure, together with a deeper literature review regarding smart meter data treatment and segmentation for broadband network.

We will perform an actual analysis can incorporate additional figures or processes leading to even more accurate results; such as, the use of energy normalized electricity consumption data, use more features than only the 24h vector to cluster, eliminate outliers when finding the mean load profile of a user, perform the same study using the weekends profiles, also, apply other and more advanced deep learning techniques.

However, in future we would calculate the energy consumption for hourly data was enough for the purpose of clustering load profiles with high spectrum allocation, but if what is sought is to determine the cause of the curve shape and the peaks hourly data won't be enough. So that, it will make necessary to use high spectrum allocation through electricity consumption data and detailed appliances and features data

REFERENCES

[1]Horner, N.C.; Shehabi, A.; Azevedo, I.L. Known unknowns: Indirect energy effects of information and communication technology. Environ. Res. Lett. 2016, 11, 103001.

[2]Achachlouei, M.A. Exploring the Effects of ICT on Environmental Sustainability: From Life Cycle Assessment to Complex Systems Modelling; KTH Royal Institute of Technology: Stockholm, Sweden, 2015.

[3]Wu, G.; Yang, C.; Li, S.; Li, G.Y. Recent advances in energy-efficient networks and their application in 5G systems. IEEE Wirel. Commun. 2015, 22, 145–151.

[4]Buzzi, S.; Chih-Lin, I.; Klein, T.E.; Poor, H.V.; Yang, C.; Zappone, A. A survey of energy-efficient techniques for 5G networks and challenges ahead. IEEE J. Sel. Areas Commun. 2016, 34, 697–709.

[5]Ge, X.; Yang, J.; Gharavi, H.; Sun, Y. Energy Efficiency Challenges of 5G Small Cell Networks. IEEE Commun. Mag. 2017, 55, 184–191.

[6]Mämmelä, A.; Anttonen, A. Why Will Computing Power Need Particular Attention in Future Wireless Devices? IEEE Circuits Syst. Mag. 2017, 17, 12–26.

[7]Björnson, E.; Sanguinetti, L.; Hoydis, J.; Debbah, M. Optimal Design of Energy-Efficient Multi-User MIMO Systems: Is Massive MIMO the Answer? IEEE Trans. Wirel. Commun. 2015, 14, 3059–3075.

[8]Hischier, R.; Coroama, V.C.; Schien, D.; Achachlouei, M.A. Grey energy and environmental impacts of ICT hardware. In ICT Innovations for Sustainability. Advances in Intelligent Systems and Computing; Hilty, L.M., Aebischer, B., Eds.; Springer International Publishing: Basel, Switzerland, 2015; Volume 310, pp. 171–189.

[9]European Commission. Ecodesign Working Plan 2016–2019 COM (2016) 773 Final; European Commission: Brussels, Belgium, 2016.

[10] Mallick, S.; Bin Habib, A.-Z.S.; Ahmed, A.S.; Alam, S.S. Performance appraisal of Wireless Energy Harvesting in IoT. In Proceedings of the 2017 3rd International Conference on Electrical Information and Communication Technology (EICT), Khulna, Bangladesh, 7–9 December 2017.

[11] Amaro, J.P.; Cortesão, R.; Ferreira, F.J.T.E.; Landeck, J. Device and operation mechanism for non-beacon IEEE802.15.4/Zigbee nodes running on harvested energy. Ad Hoc Networks 2015, 26, 50–68.

[12] Bianchi, V.; Ciampolini, P.; De Munari, I. RSSI-Based Indoor Localization and Identification for ZigBee Wireless Sensor Networks in Smart Homes. IEEE Trans. Instrum. Meas. 2018, 1–10.

[13] Al-Sarawi, S.; Anbar, M.; Alieyan, K.; Alzubaidi, M. Internet of Things (IoT) communication protocols: Review. In Proceedings of the 2017 8th International Conference on Information Technology (ICIT), Amman, Jordan, 17–18 May 2017. 38. QCA4020|Qualcomm. Available online: https://www.qualcomm.com/products/qca4020

[14] Bassoli, M.; Bianchi, V.; De Munari, I.; Ciampolini, P. An IoT approach for an AAL Wi-Fi-based monitoring system. IEEE Trans. Instrum. Meas. 2017, 66, 3200–3209.

[15] Moraru, S.; Perniu, L.; Kristaly, D.M.; Ungureanu, D.A.; Sandu, F.; Mo, A.A. Proceedings of the 2017 25th Mediterranean Conference on Control and Automation (MED), Valletta, Malta, 3–6 July 2017.

[16] Kristaly, D.M.; Moraru, S.-A.; Stefan Petre, V.; Parvan, C.A.; Ungureanu, D.E.; Mosoi, A.A. A Solution for Mobile Computing in a Cloud Environment for Ambient Assisted Living. In Proceedings of the 2018 26th Mediterranean Conference on Control and Automation (MED), Zadar, Croatia, 19–22 June 2018.

[17] Lago, P.; Roncancio, C.; Jiménez-Guarín, C. Learning and managing context enriched behavior patterns in smart homes. Futur. Gener. Comput. Syst. 2019, 91, 191–205.

[18] Bassoli, M.; Bianchi, V.; De Munari, I.; Ciampolini, P. An unobtrusive Wi-Fi system for human monitoring. In Proceedings of the IEEE International Conference on Consumer Electronics, Berlin, Germany, 3–6 September 2017.

[19] Boztas, A.; Riethoven, A.R.J.; Roeloffs, M. Smart TV forensics: Digital traces on televisions. Digit. Investig. 2015, 12, 72–80.

[20] Kang, S.; Kim, S.; Kim, J. Forensic analysis for iot fitness trackers and its application. Peer-to-Peer Netw. Appl. 2020, 13, 564–573

[21] Li, S.; Choo, K.-K.; Sun, Q.; Buchanan, W.J.; Cao, J. IoT forensics: Amazon Echo as a use case. IEEE Internet Things J. 2019, 6, 6487–6497.

[22] Becirovic, S.; Mrdovic, S. Manual IoT Forensics of a Samsung Gear S3 Frontier Smartwatch. In Proceedings of the 2019 International Conference on Software, Telecommunications and Computer Networks (SoftCOM), Split, Croatia, 19–21 September 2019; pp. 1–5.

[23] Dorai, G.; Houshmand, S.; Baggili, I. I know what you did last summer: Your smart home Internet of Things and your iPhone forensically ratting you out. In Proceedings of the 13th International Conference on Availability, Reliability and Security, Hamburg, Germany, 27–30 August 2018.

[24] Awasthi, A.; Read, H.O.; Xynos, K.; Sutherland, I. Welcome pwn: Almond smart home hub forensics. Digit. Investig. 2018, 26, 38–46.

[25] Chung, H.; Park, J.; Lee, S. Digital forensic approaches for Amazon Alexa ecosystem. Digit. Investig. 2017, 22, 15–25.

[26] Orr, D.A.; Sanchez, L. Alexa, did you get that? Determining the evidentiary value of data stored by the Amazon R Echo. Digit. Investig. 2018, 24, 72–78.

[27] Fensel, A.; Tomic, D.K.; Koller, A. Contributing to appliances' energy efficiency with Internet of Things, smart data and user engagement. Future Gener. Comput. Syst. 2017, 76, 329–338.

[28] Goudbeek, A.; Choo, K.K.; Le-Khac, N.A. A Forensic Investigation Framework for Smart Home Environment. In Proceedings of the 2018 17th IEEE International Conference on Trust, Security and Privacy in Computing And Communications/12th IEEE International Conference on Big Data Science And Engineering (TrustCom/BigDataSE), New York, NY, USA, 1–3 August 2018; pp. 1446–1451.

[29] Alharbi, R.; Allen, W.H. Collection and Analysis of Digital Forensic Data from Devices in the Internet of Things. In Proceedings of the 2019 SoutheastCon, Huntsville, AL, USA, 11–14 April 2019.

[30] Islam, M.J.; Mahin, M.; Khatun, A.; Debnath, B.C.; Kabir, S. Digital Forensic Investigation Framework for Internet of Things (IoT): A Comprehensive Approach. In Proceedings of the 2019 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT), Dhaka, Bangladesh, 3–5 May 2019; pp. 1–6.

[31] Philomin, S.; Singh, A.; Ikuesan, A.; Venter, H. Digital Forensic Readiness Framework for Smart Homes. In Proceedings of the International Conference on Cyber Warfare and Security, Islamabad, Pakistan, 29 September–1 October 2020; pp. 627–638.

[32] MacDermott, A.; Baker, T.; Shi, Q. Iot forensics: Challenges for the ioa era. In Proceedings of the 2018 9th IFIP International Conference on New Technologies, Mobility and Security (NTMS), Paris, France, 26–28 February 2018; pp. 1–5.

[33] Aslan, J.; Mayers, K.; Koomey, J.G.; France, C. Electricity Intensity of Internet Data Transmission: Untangling the Estimates. J. Ind. Ecol. 2017, 1–14.

[34] Zulkipli, N.H.; Alenezi, A.; Wills, G.B. IoT Forensic: Bridging the Challenges in Digital Forensic and the Internet of Things. In Proceedings of the 2nd International Conference on Internet of Things, Big Data and Security (IoTBDS 2017), Porto, Portugal, 24–26 April 2017; pp. 315–324.

[35] Stoyanova, M.; Nikoloudakis, Y.; Panagiotakis, S.; Pallis, E.; Markakis, E.K. A Survey on the Internet of Things (IoT) Forensics: Challenges, Approaches and Open Issues. IEEE Commun. Surv. Tutor. 2020, 22, 1191–1221.

[36] Servida, F.; Casey, E. IoT forensic challenges and opportunities for digital traces. Digit. Investig. 2019, 28, 22–29.

[37] Kruger, J.-L.; Venter, H. Requirements for IoT Forensics. In Proceedings of the 2019 Conference on Next Generation Computing Applications (NextComp), Mauritius, 19–21 September 2019; pp. 1–7.

[38] Yaqoob, I.; Hashem, I.A.T.; Ahmed, A.; Kazmi, S.A.; Hong, C.S. Internet of things forensics: Recent advances, taxonomy, requirements, and open challenges. Future Gener. Comput. Syst. 2019, 92, 265–275.

[39] Yang, W.; Johnstone, M.N.; Sikos, L.F.; Wang, S. Security and Forensics in the Internet of Things: Research Advances and Challenges. In Proceedings of the 2020 Workshop on Emerging Technologies for Security in IoT (ETSecIoT), Sydney, Australia, 21–21 April 2020; pp. 12–17.

[40] Ikpehai, A.; Adebisi, B.; Rabie, K.M. Broadband PLC for Clustered Advanced Metering Infrastructure (AMI) Architecture. Energies 2016, 9, 569.

[41] Ahmed, M.A.; Kang, Y.C.; Kim, Y.-C. Communication Network Architectures for Smart-House with Renewable Energy Resources. Energies 2015, 8, 8716–8735.

[42] Bradac, Z.; Kaczmarczyk, V.; Fiedler, P. Optimal Scheduling of Domestic Appliances via MILP. Energies 2015, 8, 217–232.

[43] Collotta, M.; Pau, G. A Novel Energy Management Approach for Smart Homes Using Bluetooth Low Energy. IEEE J. Sel. Areas Commun. 2015, 33, 2988–2996.

[44] Pascual, J.; Sanchis, P.; Marroyo, L. Implementation and Control of a Residential Electrothermal Microgrid Based on Renewable Energies, a Hybrid Storage System and Demand Side Management. Energies 2014, 7, 210–237.

[45] Collotta, M.; Pau, G. A Solution Based on Bluetooth Low Energy for Smart Home Energy Management. Energies 2015, 8, 11916–11938.

[46] Collotta, M.; Pau, G. An Innovative Approach for forecasting of Energy Requirements to improve a Smart Home Management System based on BLE. IEEE Trans. Green Commun. Netw. 2017.

[47] Subbiah, R.; Pal, A.; Nordberg, E.K.; Marathe, A.; Marathe, M.V. Energy Demand Model for Residential Sector: A First Principles Approach. IEEE Trans. Sustain. Energy 2017.

[48] Javaid, N.; Javaid, S.; Abdul, W.; Ahmed, I.; Almogren, A.; Alamri, A.; Niaz, I.A. A Hybrid Genetic Wind Driven Heuristic Optimization Algorithm for Demand Side Management in Smart Grid. Energies 2017, 10, 319.

[49] Fan, W.; Liu, N.; Zhang, J. An Event-Triggered Online Energy Management Algorithm of Smart Home: Lyapunov Optimization Approach. Energies 2016, 9, 381.

[50] Fan, W.; Liu, N.; Zhang, J.; Lei, J. Online Air-Conditioning Energy Management under Coalitional Game Framework in Smart Community. Energies 2016, 9, 689.

[51] Ikpehai, A.; Adebisi, B.; Rabie, K.M.; Haggar, R.; Baker, M. Experimental Study of 6LoPLC for Home Energy Management Systems. Energies 2016, 9, 1046.

[52] Ahmed, M.S.; Mohamed, A.; Homod, R.Z.; Shareef, H. Hybrid LSA-ANN Based Home Energy Management Scheduling Controller for Residential Demand Response Strategy. Energies 2016, 9, 716.

[53] Nguyen, T.T.K.; Shimada, K.; Ochi, Y.; Matsumoto, T.; Matsugi, H.; Awata, T. An Experimental Study of the Impact of Dynamic Electricity Pricing on Consumer Behavior: An Analysis for a Remote Island in Japan. Energies 2016, 9, 1093.

[54] Moon, J.W.; Chung, M.H.; Song, H.; Lee, S.-Y. Performance of a Predictive Model for Calculating Ascent Time to a Target Temperature. Energies 2016, 9, 1090.

[55] Pitì, A.; Verticale, G.; Rottondi, C.; Capone, A.; Lo Schiavo, L. The Role of Smart Meters in Enabling Real-Time Energy Services for Households: The Italian Case. Energies 2017, 10, 199

[56] Niaz, M.T.; Imdad, F.; Kim, H.S. Power Consumption Efficiency Evaluation of Multi-User Full-Duplex Visible Light Communication Systems for Smart Home Technologies. Energies 2017, 10, 254.

[57] Alamin, Y.I.; Castilla, M.M.; Álvarez, J.D.; Ruano, A. An Economic Model-Based Predictive Control to Manage the Users' Thermal Comfort in a Broadband network. Energies 2017, 10, 321.

[58] Luthander, R.; Widén, J.; Nilsson, D.; Palm, J. Photovoltaic self-consumption in broadband networks: A review. Appl. Energy 2015, 142, 80–94.

[59] Pickering, B.; Choudhary, R. District energy system optimisation under uncertain demand: Handling data-driven stochastic profiles. Appl. Energy 2019, 236, 1138–1157.

[60] Onishi, V.C.; Antunes, C.H.; Fraga, E.S.; Cabezas, H. Stochastic optimization of trigeneration systems for decision-making under long-term uncertainty in energy demands and prices. Energy 2019, 175, 781–797.

[61] Afzali, S.F.; Cotton, J.S.; Mahalec, V. Urban community energy systems design under uncertainty for specified levels of carbon dioxide emissions. Appl. Energy 2020, 259, 114084.

[62] Zhang, B.; Hu, W.; Cao, D.; Huang, Q.; Chen, Z.; Blaabjerg, F. Deep reinforcement learning–based approach for optimizing energy conversion in integrated electrical and heating system with renewable energy. Energy Convers. Manag. 2019, 202, 112199.

[63] Gabrielli, P.; Fürer, F.; Mavromatidis, G.; Mazzotti, M. Robust and optimal design of multi-energy systems with seasonal storage through uncertainty analysis. Appl. Energy 2019, 238, 1192–1210.

[64] Testi, D.; Urbanucci, L.; Giola, C.; Schito, E.; Conti, P. Stochastic optimal integration of decentralized heat pumps in a smart thermal and electric micro-grid. Energy Convers. Manag. 2020, 210, 112734.

[65] Andrae, A.; Edler, T. On Global Electricity Usage of Communication Technology: Trends to 2030. Challenges 2015, 6, 117–157.

[66] Malmodin, J.; Lundén, D. The energy and carbon footprint of the ICT and E&M sector in Sweden 1990–2015 and beyond. In Proceedings of the International Conference on ICT for Sustainability (ICT4S), Amsterdam, Netherlands, 30–31 August 2016; pp. 209–218.