Simulating Household Electricity Consumption

Mutinta Mwansa, William Hurst, Carl Chalmers, Yuanyuan Shen, Casimiro A. Curbelo Montañez
Department of Computer Science
Liverpool John Moores University
Byrom Street
Liverpool, L3 3AF, UK
Email: M.Mwansa@2017.ljmu.ac.uk, {W.Hurst, C.Chalmers, Y.Shen, C.A.CurbeloMontanez}@ljmu.ac.uk

Abstract— **The smart grid is an advanced infrastructure that leverages communication technology, data analytics and cloud computing to control the distribution and consumption of energy. Smart grid systems include producers, consumers and actors, to ensure a resource saving and efficient electrical network. Within the smart grid, the smart meter records the consumption of electricity in private homes and businesses accurately. The data generated can be used to provide an insight into social demographics, household behaviour patterns, social clusters, general energy consumption patterns and a variety of value-added services. However, one of the biggest challenges for researchers in this area is the access to smart meter datasets. This is because real world datasets contain sensitive consumer information and, therefore, privacy is a key concern. Therefore, this paper focuses on simulating realistic data collected from the residential smart meter. As such, this paper presents a simulation of a home environment and the data produced. The validity of the data is justified through a visual comparison with a real-world smart meter dataset.**

***Keywords- Smart Meters; Profiling; Simulation; Visualisation*.**

I. INTRODUCTION

Smart meters are a core component of the smart grid. Typically, they reduce financial losses, operational costs and enable energy suppliers to forecast customer demand [1]. As a result, smart meters are being implemented on a global scale. Many countries such as the UK, USA, Australia and Italy are already advanced in their smart meter implementation. Additionally, Sweden is one of the first countries in Europe to carry out metering reform and large-scale smart meter roll out. Before the reform, electricity consumption data for small customers is typically read on a yearly basis and billing is estimated based on the previous year's consumption, instead of actual meter readings. Consumer demand for timely and correct billing is the main driver for smart meter deployment [2].The smart meter system is equipped with a large number of sensors and actuators placed in all parts of the grid to monitor and control the operational characteristics and behaviour. Based on the data collected from these sensors, smart meter entities and electricity suppliers (utility companies) are able to make more insightful and better decisions. For example, they are able to manage and optimise the electricity flows, forecasting users' demand for electricity and balancing the grid more efficiently; and even detect when there is abnormal energy usage in homes. The potential research implications of access to this data is significant. For example, considerable research has been implemented into the use of smart meter data for remote healthcare monitoring [3]. Whereas, other research, has focused primarily on load balancing to support the efficiency of the grid and resource allocation [4]. The remainder of the paper is as follows. A background research on smart meter systems is put forward in Section 2. Subsequently, the research aims and objectives and the methodology is discussed in Section 3. Section 3 also presents a sample of the data collected from our smart meter case study. Section 4 discusses the methodology and techniques used for profiling users. The paper is concluded in Section 5. In particular, this paper focuses on the smart meter and investigates the novel approaches for consumer profiling and for the consumers to monitor energy usage in real time.

II. BACKGROUND RESEARCH

A smart meter is an electronic device that records the consumption of energy with high accuracy. However, smart meter is part of the much wider Advanced Metering Infrastructure (AMI).

A. *The Advanced Metering Infrastructure (AMI)*

An AMI is comprised of systems and networks that receive data from smart meters; and it facilitates the bidirectional communication between the consumer and the rest of the smart grid stake holders. It reduces the traditional need for energy usage readings to be collected manually [7]. Therefore, the smart meter is able to communicate with a gateway through a Home Area Network (HAN), Wide Area Network (WAN) or a NAN, which is outlined as shown in Figure 1.

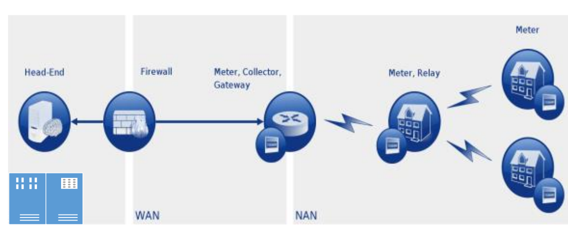

Figure 1. Advanced Metering Infrastructure

The HAN is housed inside the consumer premises and is made up of different devices, e.g., Meters, Thermostats, Electric storage devices, ZigBee transmitters. All of the acquired data is sent to the Meter Data Management System (MDMS), which is responsible for storing, managing and analysing the data [8]. The MDMS sits within the data and communications layer of the AMI. This component is an advanced software platform, which deploys data analytics while facilitating various AMI applications, including:

- Managing metered consumption data.
- Outage management.
- Demand and response.
- Remote connect / disconnect.
- Smart meter events and billing [9].

This information can be shared with consumers, partners, market operators and regulators. The Wide Area Network (WAN) handles the communication between the utility companies and the HAN. The Head-End System (HES), also known as the meter control system, is located within a metering company network.

B. *Machine Learning Techniques for Profiling*

The first step to profile behaviour from the data produced by smart meters is to model and understand the normal patterns. Therefore, the field of machine learning provides methodologies that are ideally suited to the task of extracting knowledge from these data. A parametric approach often used is linear regression, which predicts a real valued output based on one or more input values. Prediction of a single output variable from a single input variable is called “univariate linear regression” whereas “multivariate linear regression” indicates multiple features. This module is used to define a linear regression method, which trains a model using a labelled dataset. The trained model can then be used to make predictions. Regression analysis is usually the best option and the fastest method to analyse the consumption data of buildings [10]. Among the statistical approaches, regression techniques deserve attention due to:

- Relative ease to implement.
- Interpretability of the results.
- The requirement of less computational power than other statistical approaches (genetic algorithms, neural networks, support vectors machine).
- Satisfactory prediction ability.
- Increased availability of data through smart metering.

Linear regression is a statistical analysis method used to model the relationship between two variables via fitting a linear equation to observed data. This relationship can be identified between the independent (explanatory) and dependent (response) variables. The response must be continuous, whereas the independent variables may be either continuous, binary or categorical. Linear regression can be expressed as:

$$y = \beta_0 + \beta_1 x \ , \tag{1}$$

Where Y is the dependent variable, and X is the independent or explanatory variable and the betas (β_0 and β_1) are the coefficients that we need to identify to make predictions. This is demonstrated in Figure 2.

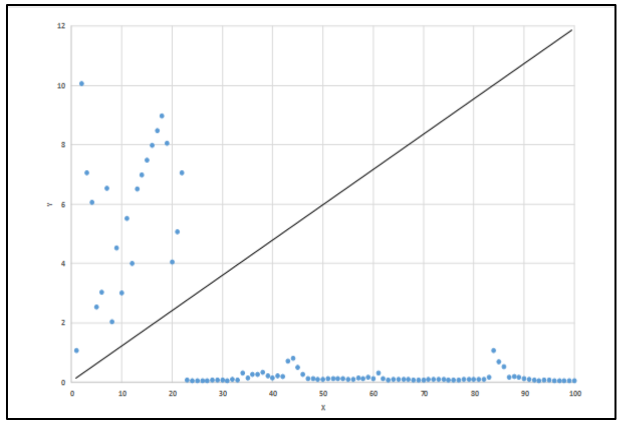

Figure 2. Simple Linear Regression

The line is modelled based on the linear equation shown in Figure 2 fitted to our data. Here, x and y are known variables from our data used to estimate the parameters β_0 and β_1 of the regression line. Typically, the parameters are estimated by minimising the least squares or the sum of squared errors. Therefore, for n observations, the linear regression model can be defined as

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i, \quad i = 1, 2, \ldots, n. \tag{2}$$

The random variable ε represents the error term in the model, a statistical term that represents random fluctuations and measurement errors among other factors out of our control [11].

Techniques, such as linear regression, can be used to predict future energy trends. For example, by using the smart meter data readings taken from one month, it would be possible to predict with a relative accuracy the expected consumption for the following month. In the following sub-section, a demonstration of the data that is collected from smart meters and how it can be analysed to model user behaviour is presented.

III. Case Study: Real-World Data

The electronic meters for electricity (smart meters) are undergoing an increasing deployment in private homes all over the world. As a consequence, an ever growing physical communication network, made up of millions of local meters, has been established, whose considerable advantages are so far in favour primarily, if not solely, of the energy distributors, since they are enabled at simplified, more efficient, and less costly transactions with the customers, e.g., for meter reading, billing, and energy supply administration. The detail and granularity of the data collected can be used in so many ways by utility companies, the future challenges faced is the issue of data storage and data management costs which prevent initiatives from becoming widely adopted. The amount of data produced by two million smart meter customers reaches upwards of 22 gigabytes per day [15]. Naturally, it is a significant challenge to manage this data; which may include the selection, deployment, monitoring, and analysis processes.

C. Data Description

Any real-time information processing usually requires cloud computing [16]. Any delay may cause a serious consequence in the whole system, which has to be avoided as much as possible. The dataset used in this research is comprised of one-month's energy readings from 5 different users. Table 1 demonstrates a sample of smart meter data collected over a period of one month (January) for a single home occupant. The general supply of energy used on a daily basis (the energy consumed) is measured in kilo watts per hour (KWH) and can be described as what is used to bill the customer. Table 1 shows an example of energy reading of an individual household meter. Data is collected over a 30 min time interval period and the "energy delivered" in KWH. The customer key is the primary key used to identify the consumer while the End Date Time highlights the time and date of the acquired reading. Both the general supply and off peak supply are recorded based on the specified tariff.

Table I. SMART METER DATA SAMPLE

CUSTOMER_KEY	End Date time	General Supply KWH	Off Peak KWH	Year
8410148	1/1/13 0:29	0.081	0	2013
8410148	1/1/13 0:59	0.079	0	2013
8410148	1/1/13 1:29	0.082	0	2013
8410148	1/1/13 1:59	0.085	0	2013
8410148	1/1/13 2:29	0.073	0	2013
8410148	1/1/13 2:59	0.07	0	2013
8410148	1/1/13 3:29	0.07	0	2013

As above, a sample of the dataset is presented in the visualisation in Figure 3. In this case, five user's energy consumption over a five hour period is displayed.

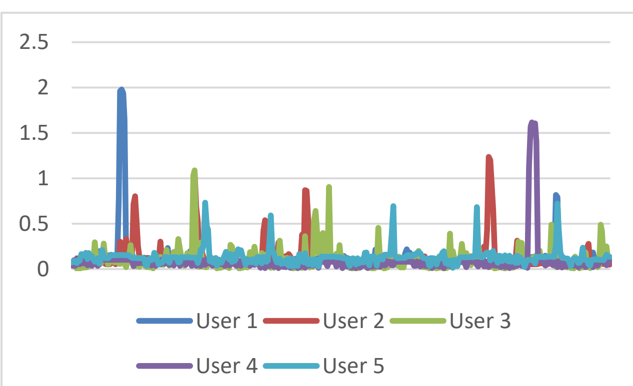

Figure 3. Visualisation of Data Sample

The y-axis displays the energy usage reading, while the x-axis displays the time in half-hours. As there are five hours, there are ten time stamps on the x-axis. The total dataset for the five users is plotted in a scatter matrix (Figure 4), which shows the correlation between their individual energy usage patterns.

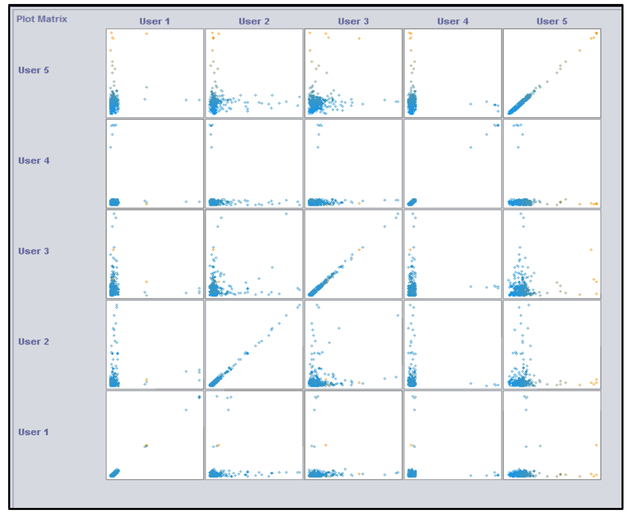

Figure 4. A generic pattern captured for 5 different users showing energy usage.

In order to visualise and analyse the total energy usage patterns over a much longer period, the smart meter data is loaded into a data model. The software used for this task is Microsoft Power BI [17]. Figure 5 presents an example of a much larger dataset that is comprised of seventy thousand household meter readings showing the energy usage and the behaviour trend over a period of 12 months. Here, the general distribution of energy readings highlights the energy consumptions levels for different households. This type of data visualisation could give suggestion to the number of occupants living in a given premise. Houses with increased energy usage are more likely to have an increased number of occupants or devices [12].

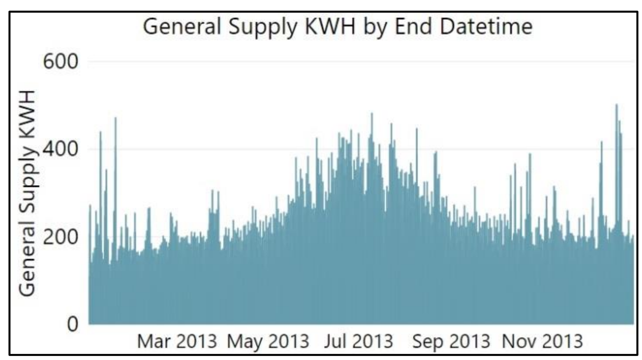

Figure 5. A generic pattern captured for 5 different users showing energy usage.

However, this information can be used by others, either maliciously or inadvertently to ascertain an insight into an individual's home life. For example, activities or occupancies of a home for specific periods can be determined. In a general, analysis of granular smart meter energy data could result in 1) invasion of privacy; 2) unwanted publicity and embarrassment (e.g., public disclosure of private facts of people's daily living lifestyles).

The security policies governing the reliability of the smart grid depend on appropriate connectivity protocols and the national institute of standards and technology being the reference model proposed [16]. Recognizing the urgent need for standards to support Smart Grid interoperability and security, NIST developed a three-phase plan. 1) Identify an initial set of standards that would promote the rapid development of the Smart Grid, 2) establish a robust framework for the sustaining development of many additional standards, and 3) establish the a framework for the conformity testing infrastructure that is needed.

III. CASE STUDY: SIMULATION DATA

As previously discussed, access to smart meter data is limited. In this section, a case study on the simulation of smart meter data is presented.

D. Simulationd Design.

Figure 6 displays a model of a simulated home. The home was designed to resemble a moderate family home, with a standard set of appliances.

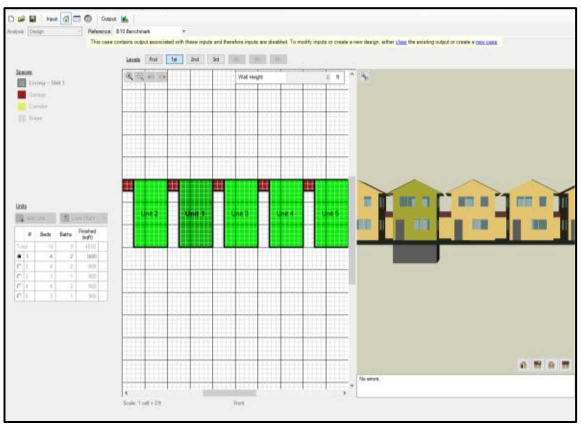

Figure 6. Simulated Home

The number and type of appliances present in the home are customised in Figure 7.

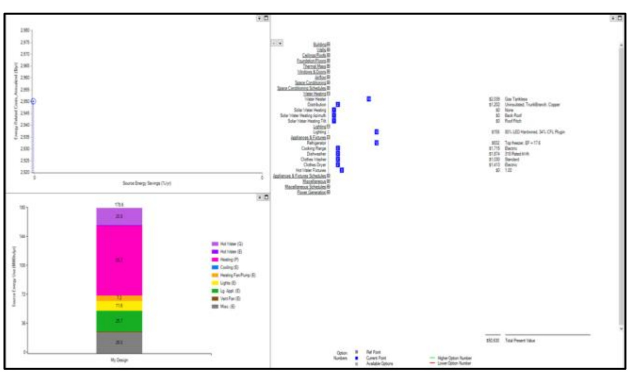

Figure 7. Home Customisation

The properties of the simulated home are detailed in Table II. Other properties, such as insulation, construction materials, weather sheathing and exterior finish are included in the simulation but omitted from Table II.

Table II. SIMULATED SMART HOME PROPERTIES

Input	Value
Project Type	Standard
Application Type	New Construction
Building Type	Multi-Family
Analysis Mode	Design
Reference Building	My Design
Sim Engine	EnergyPlus
Building: Finished Floor Area	4500
Building: Bedrooms	18
Building: Bathrooms	8

By running the simulation, data is constructed for a given simulation period. A visualisation of the energy readings from the simulated home is presented in Figure 8. The months are displayed on the x-axis and the total energy usage values for a given day are displayed on the y-axis.

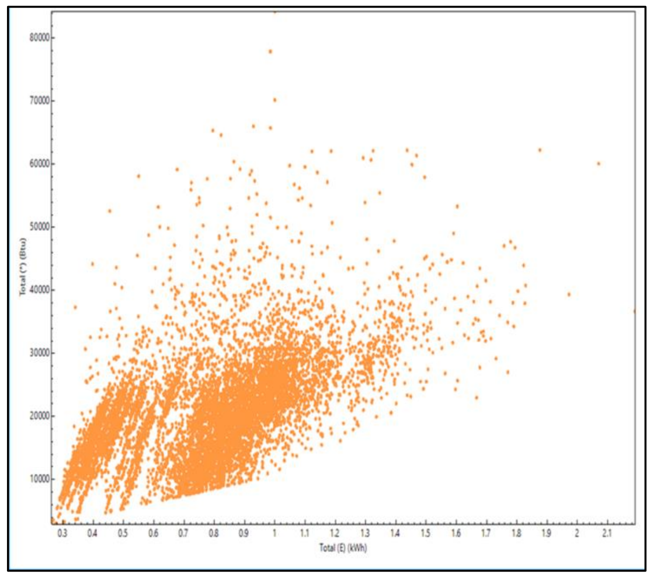

Figure 8. Simulation Data Visualisation Sample

In the following section, a discussion is presented on the use of real-world smart meter data in comparison to the simulated data generated by the artificial home.

IV. DISCUSSION

Figure 9 displays real-world energy readings from a single home. The home is picked at random from the data set, but it meets the following criteria:

- Data shows that there is an occupant in the premises
- They live in a standard house

Each morning demonstrates a sudden change in user behaviour. The energy usage, in KWH, is shown in the y-axis while the time the reading was taken is shown in the x-axis. The graphs indicate the time when the consumer becomes active in the morning. These activity start times vary depending on each user and readings are captured for

 ISBN: 978-1-61208-724-5

the whole 24 hour period. These types of behaviour can be attributed to the consumer's morning, afternoon and evening activities and is a key indicator for understanding and identifying alterations in routine.

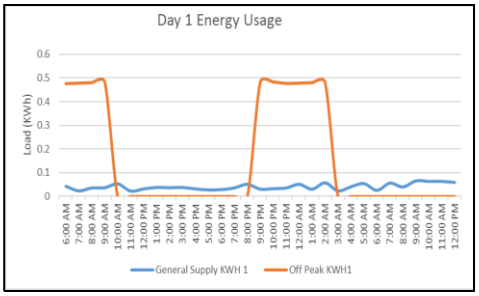

Figure 9a. Measured power load over a 24 hour period on day 1 showing occupant using energy mostly in the moning and late evenings.

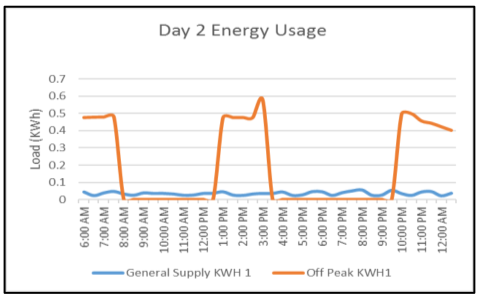

Figure 9b. Measured power load over a 24 hour period on day 2 showing occupant using energy in the morning, mid-afternoon and late evenings.

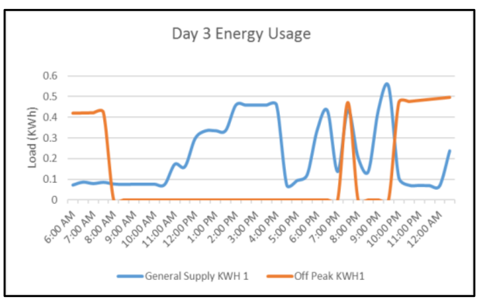

Figure 9c. Measured power load over a 24 hour period on day 3 showing occupant using energy throughout the day.

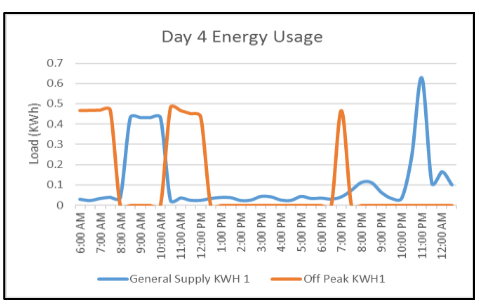

Figure 9d. Measured power load over a 24 hour period on day 4 showing occupant using energy in the morning and late evenings.

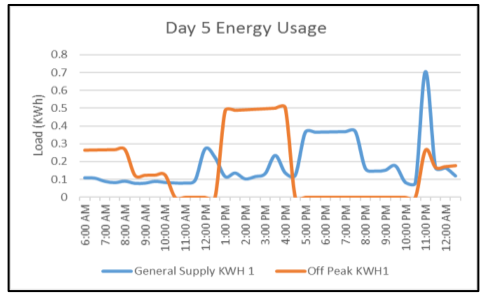

Figure 9e. Measured power load over a 24 hour period on day 3 showing occupant using energy throughout the day.

By the efficient analysis of the energy use data, different energy consumption patterns of different household are demonstrated below and corresponding energy use behavioural characteristics are identified. Along the x-axis is the hours of the day over a 24 hour period with the readings taken for each day for 5 consecutive days for each house. The y-axis values refer to the energy usage in kilowatts (KWh). Using this data, we will build up a pattern of expected behaviours and identify a trend. Figure 10 displays the simulated data home patterns over a 12-month period. The green line is the total energy usage, whereas the orange and the blue line depicts the individual energy readings for unit 1 and unit 2 in the simulation study. Similarly to the real-world data, on visual inspection, patterns in the energy consumption are apparent. For example, similar spikes in the energy use on a weekly basis, lower energy use in the summer months and higher energy use during winter.

Figure 10. Simulated Home Graph Visualisation

V. CONCLUSION

In this paper, several visualisations of electricity consumption patterns both from a real-world dataset and from a simulated smart home have been conducted. Smart meters produce considerable volumes of data, presenting the opportunities for utilities to enhance customer service, lower costs and improve energy efficiency; and for customers to reduce their bills and save energy. The availability of such information results in more informed consumers who can better self-manage their electricity usage, choose low energy saving appliances and thus contribute to the reduction of greenhouse gas emissions. Therefore, we conclude that the analytic techniques and methodology proposed in this paper can be practical and useful to benefit the consumers and utility companies as well as the governments. Our ambition is that our work will lead to a realistic simulated smart meter dataset that other researchers can use for investigations into home profiling.

This research project is funded in part by the EPRSC - EP/R020922/1. Owing to the ethical sensitive nature of this research, the data underlying this publication cannot be made openly available.

REFERENCES

[1] Skopik, F. and Smith, P.D. eds., 2015.Smart grid security: Innovative solutions for a modernized smart grid . Syngress.

[2] J.Zheng, D.W. Gao, and L.Lin, " Smart meters in smart grid: An overview. In Green Technologies," Conference IEEE, pp. 57-64, April, 2013.

[3] E.D. Knapp and J.T. Longil, " Industrial Network Security: Securing Critical Infrastructure Networks and Smart Grid, SCADA and other Industrial Control Systems," Chapter, 12, 2015.

[4] Espinoza, M., Joye, C., Belmans, R, and De Moor, B., 2005. Short-term load forecasting, profile identification, and customer segmentation: a methodology based on periodic time series. IEEE Transactions on Power Systems, 20(3), pp.1622-1630.

[5] Bedingfield, S., Alahakoon, D., Genegedera, H, and Chilamkurti, N., Multi-granular electricity consumer load profiling for smart homes using a scalable big data algorithm. Sustainable cities and society, 40, pp.611-624,2018.

[6] Virone G, et al. Behavioral patterns of older adults in assisted living. IEEE Trans Inf Technol Biomed 12(3), pp. 387–398,2008.

[7] M.Popa, Data Collecting from Smart Meters in an Advanced Metering Infrastructure, Proceedings of 15th International Conference on Intelligent Engineering Systems, pp. 137-142,2011.

[8] D. Niyato and P.Wang, "Cooperative transmission for meter data collection in smart grid," IEEE Communications Magazine, vol. 40, pp. 137-142, 2012.

[9] Y.T. Hoi, et al. "The Generic Design of a High-Traffic Advanced Metering Infrastructure Using ZigBee," vol. 10, pp. 836-844, 2014.

[10] Malekian, R et al. A novel smart ECO model for energy consumption optimization. Elektronika ir Elektrotechnika, 21(6), pp.75-80, 2015.

[11] A. C. Rencher and G. B. Schaalje, Linear Models in Statistics. Hoboken, NJ, USA: John Wiley & Sons, Inc., 2007.

[12] Bartusch C, Odlare M, Wallin F and Wester L. Exploring variance in residential electricity consumption: Household features and building properties. Applied Energy 2012;92:637–643

[13] World Health organisation on Ageing 2015, accessed 19 march 2018,<http://www.dw.com/en/aging-health-who-call-to-action-as-world-population-over-60-set-to-double-by-2050/a-18751636#>

[14] Chalmers, C., Hurst, W., Mackay, M., & Fergus, P. Smart Meter Profiling For Health Applications. In the Proceedings of the International Joint Conference on Neural Networks, July 2015

[15] Choo, K.K.R., 2011. The cyber threat landscape: Challenges and future research directions. Computers & Security, 30(8), pp.719-731.

[16] Advanced metering Security Threat Model (DRAFT); R Robinson, J McDonald,B Singletary, D Highfill,N Greenfield,M Gilmore

[17] T. Lachev, and E. Price, 2018. Applied Microsoft Power BI: Bring your data to life!. Prologika Press.

[18] Coalton Bennet and Darren Highfill. Networking AMI Smart Meters, IEEE Energy 2030, November 2008.

[19] Bera, S., Misra, S. and Rodrigues, J.J., Cloud computing applications for smart grid: A survey. IEEE Transactions on Parallel and Distributed Systems, 26(5), pp.1477-1494, 2012.

 ISBN: 978-1-61208-724-5