International Journal of Computing and Digital Systems
ISSN (2210-142X)
Int. J. Com. Dig. Sys. 11, No.1 (Feb-2022)
https://dx.doi.org/10.12785/ijcds/110170

A Taxonomy of Smart Meter Analytics:
Forecasting, Knowledge Discovery, and Power Management

Elham M. Eskandarnia and Hesham M. Al-Ammal

Department of Computer Science, University of Bahrain, Bahrain

Received 16 Aug. 2021, Revised 27 Oct. 2021, Accepted 28 Jan. 2022, Published 15 Feb. 2022

Abstract: The field of smart meter data analytic is a relatively young field that grew because of the wealth of data generated from the use of smart meters. This data has been used for several sustainability applications including power management and planning. This review paper aims at presenting a unifying taxonomy to classify the various domains of smart meter analytic and their underlying functions and techniques. The aim is to better understand the current research trends, approaches, and opportunities. The paper reviews the functions, applications, and techniques after examining a huge body of knowledge within the three main domains of the field, namely forecasting, knowledge discovery, and power management. In the forecasting domain, the functions are classified based on the scale and horizon. Other domains are divided into relevant functions, algorithms, and general techniques. The review is performed from the perspective of data science; emphasizing the data tasks such as data wrangling, analytics algorithms, and evaluation methods for each domain. A review of the various algorithms and techniques within each function is presented, and the paper concludes with a discussion of issues and research opportunities for smart meter data within the machine learning and data science fields
Keywords: Power management, smart meters, data analytics, smart grid, machine learning.

1. Introduction

The 21st century's economy grows increasingly dependent on energy that utilizes smart power transfer methods. A smart grid is an updated form of an electrical power system, which uses smart meters as one of its main components. This smart infrastructure is a requirement for sustainability, as it is the first building block for sustainable use of energy. Power management, forecasting, planning, and many other functions can only be implemented through the efficient use of this smart meter data.

Smart meters record the power used by consumers periodically and transmit it to the data collector through a Local Area Network (LAN). The collector retrieves the data and transmits it to the utility's central collection points. The large volume of raw data generated by smart meters needs to be stored, cleaned, and analyzed; in order to discover commercially usable knowledge and gain insight. This wealth of

information can be utilized by individuals, the energy industry, service providers, as well as regulators, and planners. The significant potential of the massive amounts of data collected by smart meters is heavily dependent on the capacity to store, manage, analyze, and extract valuable information from this data. Smart meter analytic refers to the conversion of the massive amounts of data flowing through different layers to actionable knowledge. Looking at the literature, there is a need for a comprehensive survey that covers all aspect of smart meter data analytics, only a few review papers cover all aspect of smart meter analysis none of the previous review papers covers all aspects of smart meter analytic [1]. This paper attempts to cover this gap by providing a broad overview of the current research spectrum while identifying future challenges for smart meter data analytics within a detailed taxonomy. The proposed taxonomy identifies the main domains within the field, as well as the various functions that are conducted within each domain. It includes a survey of the various algorithms and techniques that were proposed by the researchers to achieve these functions. This classification will enumerate the current research trends and provide a better understanding of the available opportunities for researchers.

2. Review Method and Contribution

A biblio-metric analysis of the field lead to a comprehensive review of the research related to smart meter data analytics. Figure 1 summarizes the number of

E-mail address: e.eskandar@bub.bh ; halammal@uob.edu.bh

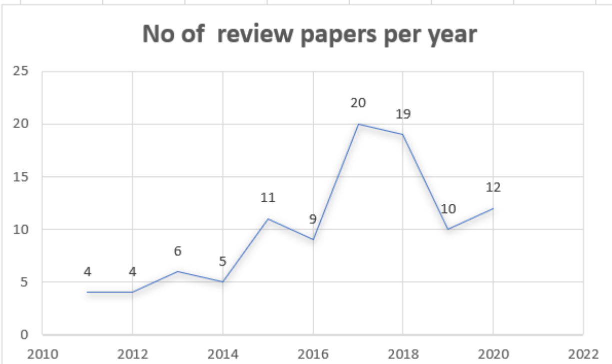

Figure 1. Summary of Review Papers Per Year

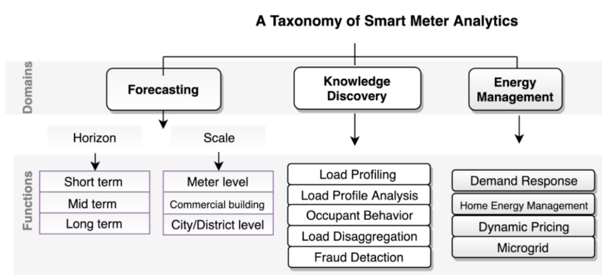

Figure 2. A Taxonomy of Smart Meter Analytics

review papers published annually. These review papers generally focus on one domain of smart meter analytics. A more comprehensive search was then conducted using keywords extracted from around 200 review papers. This analysis lead to a classification of the research into the taxonomy presented in the next section covering all domains of the field. Although several review papers were published in the field, they mainly focus on a specific topic. These papers can be categorized into two groups: reviews about special domains of smart meter analytics [2]–[5], and reviews about specific techniques used within these domain(s) [6]–[9]. Among these review papers only few attempt to cover all aspects of smart meter analysis [1], [10], [11]

To remedy this situation analysis of around 200 papers from the literature resulted in a clear taxonomy of the domains and functions available within the field. Figure 2 presents the proposed taxonomy that spans the field of smart meter data analytics.

Load forecasting refers to the prediction of future electricity demand in different horizons and scales. *Knowledge discovery* refers to new information extracted from raw databases of smart meter data by applying techniques such as machine learning, pattern recognition, classification, clustering, statistical analysis, etc. While *energy management* refers to the processing of smart grid data in order to deploy functions that are related to the management of grid resources, such as: improving energy efficiency, reducing the cost of energy production, response to demand fluctuation, the introduction of dynamic pricing, etc. In the following sections, we will discuss each of the domains and summarize related research.

3. Domain I: Forecasting

Forecasting refers to the prediction of future electricity demand in different horizons and scales. More exact forecasting on multiple scales, ranging from a single smart meter (the end-user) to an entire grid system, is increasingly necessary for today's competitive energy market. As a result, load forecasting is recognized as an essential component of the planning process in the electric sector. Furthermore, load forecasting is critical for customers to optimize their use of electrical energy. It can be done on a different scale (the size of the unit at which the load forecast is done) and on a different horizon (the time period of forecasting); The extended discussion of the load forecasting domain is provided by Iskandarnia et al. [12].

Load Forecasting Techniques

Considering the data science approach, the framework for forecasting will include the use of exploratory analysis and statistical methods to understand trends within the given data sets; and applying smoothing techniques in order to remove noise prior to applying the algorithms. Load forecasting is done based on this approach discussed in details [12].In general, this framework includes: Pre-Processing and Data Wrangling, forecasting Algorithms, and evaluation of forecasting techniques Statistical techniques along with machine learning techniques are widely used in load forecasting. [13], [14]

4. Domain II: Knowledge Discovery

The second domain of smart meter data analytics involves the extraction of new knowledge from data. Applications in this domain vary widely and can be implemented at the energy supplier's side or the consumer's side. The techniques vary from classical data mining to methods such as deep learning and meta-heuristics. The following sections review various functions within this domain.

A. Load Profiling

A *load profile* is a graph/curve that represents the variation of electrical usage over time. Power companies traditionally create aggregate load profiles and create *Typical Load Profiles* (TLP) and group their consumers and model each group by a *Standard Load Profile* (SLP). By employing the opportunities provided by smart meters, customers can now be classified by analyzing their load curves. Load profiling helps in grouping the consumers based on similarity in their consumption behavior, and subsequently, produces the typical load profile for each group. A load profiling problem can be formulated as an unsupervised machine learning task, and hence, it is completely data-driven.

It should be noted that often load profiling produces input for the other functions of the knowledge discovery domain, and furthermore, it can also be employed as input for grid management tasks. This highlights the significance of load profiling as a key tool for smart meter analytics. Load profiling applications are classified as follows:

- *Energy management task*: Load profiling can be used as input to electricity management applications such as demand response and security monitoring. Kwac et al. [15] applied load profiling to create a dictionary for consumer patterns. The TLP of each consumer was matched by representative load shapes in the dictionary, to target consumers for demand response. Load profiling can also be used for detecting unusual usage and initiating warnings for such usage [16]. It also can be used for other functions such as tariff setting, where consumers are first divided into different categories based on their individual information by a clustering algorithm; then, A pricing model is presented as a nonlinear programming problem from the perspective of a load service company, with the goal of minimizing the overall operating cost.
- *Cluster-based aggregate load forecasting*: Load profiling can improve the accuracy of forecasting if done beforehand. When load forecasting is performed on an aggregate scale, consumers with common load trends are identified, and forecasting can be extended to each group [17], [18].
- *Other tasks in Knowledge discovery*: Load Profiling improves the accuracy and efficiency of other tasks in this domain. For example, it feeds information to outlier detection algorithms and to improve fraud detection tasks [19]. It may also be used prior to the occupant and load analysis tasks. Consumer segmentation is usually accomplished using a combination of clustering and classification approaches. Following the grouping of users based on the similarities of their actions by load profiling, supervised methods (classification) can be used to extract further information for decision-making purposes [20].

Load Profiling Techniques

There are three approaches to load profiling: engineering, statistics, and data-driven. The first two approaches are out of scope for this paper. The data-driven clustering components are as follows: Cluster-driven load profiling attempts to group similar consumers into a single cluster based on similarities in consumption patterns. Based on the available literature, load profiling can be defined as clustering high dimensional data with no labels, and it is usually performed in three main stages:

- *Data preparation*: the dataset is prepossessed and dimension reduction is performed to deal with the problem of high dimensionality and compress the size of data. This stage aims to map raw data into a new feature space, where the generated data are easier to separate by existing classifiers. Traditional data transformation approaches include linear and non-linear transformations, such as Principal Component Analysis (PCA), Kernal, and spectral methods respectively. A review of feature reduction techniques used for load profiling is represented in [5].
- *Load curve clustering*: K-means, Fuzzy clustering (FCM), Hierarchical clustering process, and Self Organize Maps (SOM) are some of the traditional and most widely used clustering algorithms in load profiling.
- *Evaluation*: Smart meter readings are unlabeled and no structural knowledge about the data set is available. As a result, calculating the optimum number of clusters and evaluating the accuracy of clustering outcomes is a challenging task. The most popular method for determining the optimum number of clusters is to run the clustering algorithm several times while attempting various numbers of clusters, and then choose the number of clusters that produces the best results in a predefined evaluation criterion function. Clustering validity is usually evaluated with visualization, the goodness of fit as well as scalar measurements [21]. The clustering quality also can be assessed with clustering evaluation indexes such as the Silhouette Coefficient, and Davies-Bouldin Index.

Traditional load profiling methods, typically suffer from the curse of dimensionality and thus perform poorly on high-dimensional smart meter readings. Furthermore, on large-scale smart meters datasets, these approaches typically suffer from large computational complexity. This all leads to the use of more advantageous techniques in literature such as hybrid machine learning and deep learnings.

B. Load Profile Analysis

In general, this is an extended form of load profiling, where factors are combined with smart meter measurements to extract new knowledge about factors that affect the load shape. Ultimately, load profiles can be affected by exogenous factors (type of appliances, insulation) or context factors (occupancy, weather, seasons, holidays). some papers that investigate these factors are: building characteristics [22] socioeconomic factors [23], geographical location [24], appliance in use [25], regional [26], weather and climate conditions [16].

It is important to emphasize the similarities and differences between two functions of knowledge discovery: occupant behavior and load profile analysis. Even though both methods are used to discover knowledge about factors affecting the load curve, they vary largely due to the techniques applied to extract the knowledge. In load analyses, clustering and classification are used; while in occupant behavior, techniques such as motif discovery and association rules are usually adopted.

Two approaches for load analysis are implemented in the literature:

- *Feature extraction*: where methods like: PCA, deep learning, frequency analyzer and SAX are implemented as well as classification methods [24], [27].
- *Investigation of existence of a factor*: researchers try to reveal knowledge about some characteristics of the household from smart meter data [28]–[30].

Load Profile Analysis Techniques

as explained, smart meter readings are integrated with other relevant data sets such as building characteristics, weather information, etc. Data reduction methods such as PCA or deep learning techniques are used to extract the main factors. Since classification methods are widely used in this regard, a discussion of these methods is presented below.

Classification

For the classification of time series data, where the aim is to exploit the sequential nature of the data, two approaches exist: (i) existing algorithms are redesigned and modified to handle sequential data, (ii) the sequential data is transformed to fit the input requirements of existing algorithms.

The classification has pre-processing steps that are very similar to those performed for load profiling; so the reader is referred to the previous section. The major three classification classes include: *Distance-based classification* [31]; *Feature-based classification* using techniques such as SVD (Singular Value Decomposition), DFT (Discrete Fourier Transform), DWT (Discrete Wavelet Transform) and SVMs (Support Vector Machines) [32]; and *Model-based classification* using statistical models like Gaussian, Poisson, Markov Model and Hidden Markov Models [33].

C. Occupant Behavior

Smart meter data can be analyzed to discover the consumer's lifestyle. Different terms are used in the literature for this concept, including: pattern discovery, occupant behavior, lifestyle discovery, activity-based loading profiling, and routine discovery [24], [34]–[36]. This task is mostly performed by implementing rule-based algorithms that are used to extract useful information from the data set, and then to define appropriate IF-THEN rules for the system. The output of other domains of smart meter analytics can be used as input to this application.

Occupant Behavior Techniques

In the literature, various techniques were used for occupant behavior identification but all of them have two main steps: discovering frequent itemsets and generating association rules based on the frequent itemsets. for instance, Hidden Markov Models (HMMs) to [37] and Association rule mining (ARM) [34]. The goal of motif discovery is to find frequently recurring subsequences in a time series. A univariate time series is first segmented into distinct subsequences depending on user-specified window size. Certain distance measures are employed to assess the correlations between different pairs of subsequences. Motifs are discovered as sub-sequences with a significant degree of similarity.

D. Load Disaggregation

Load disaggregation is achieved by analyzing the integrated load of a household in order to discover the different appliances used by the occupant. Providing direct and real-time feedback to the end-users on the appliance's consumption may contribute to energy savings. Moreover, load disaggregation can be utilized to understand the lifestyle of the consumers [38].

There are two main approaches to load disaggregation [39]:

- Intrusive Appliance Load Monitoring (IALM): IALM can perfectly anticipate the operating state by employing sensors linked to every device in the house.
- Non-intrusive Appliance Load Monitoring (NILM): This method entails the practice of separating the overall electrical load of a household measured at a single location into separate appliance signals. signals [40].

Load Disaggregation Techniques

The NILM method consists of two steps: extraction and identification. The extraction phase transforms the smart meter readings into a collection of characteristics that are utilized for load recognition. Following categorization, each class will be allocated with an electrical signature. The detected events are connected with a loads signature database in order to determine the load state (on or off) and track the appliance usage over time. The load is identified by using machine learning techniques [27], [39].

E. Fraud Detection

Energy fraud detection is a crucial component of the smart grid. Energy fraudsters can manipulate smart meter data in two ways: they consume more energy while reporting less, or they report more energy for certain meters to obtain additional advantages from the utility. Physical tampering, such as illegal tapping, bypassing meters, firmware modification, and cyberattacks against the AMI infrastructure are used to manipulate smart meter reading. Non-technical loss (NTL) in power systems is mostly the result of electricity theft, billing problems, and unpaid bills. Machine learning and data mining have been widely utilized by researchers for comprehensive intelligent data analysis in order to recognize typical patterns of behavior so that deviations may be recognized as anomalies.

Fraud Detection Techniques

Machine learning and data mining have been widely utilized by researchers for comprehensive intelligent data analysis in order to recognize typical patterns of behavior so that deviations may be recognized as anomalies.

Fraud detection consists of the following stages [41]:

- *Pre-possessing:*
  Raw data goes through the pre-processing stage to extract features for the NILM algorithm. Since feature selection has a direct effect on algorithm output, feature selection algorithms can be used for defining an optimal set of features [20].
- *Machine learning:* Two main categories of machine learning are used in: supervised methods (for labeled data) including SVM, Artificial Neural Networks (ANN), Optimum Path Forrest (OPF), and The Na'ive Bayes Classifier; and Unsupervised methods (for unlabeled data) which include self-organizing maps and regression methods.
- *Evaluation:* Depending on the type of algorithm used, different metrics are used for assessing the performance of detection methods. Besides metrics that evaluate the performance of machine learning algorithms, other metrics are also defined to quantify the effects of the detection methods.

5. Domain III: Energy Management

Implementation of new power techniques which come with the smart grid such as renewable energy sources, energy storage systems, chargeable batteries for vehicles, etc.; not only brings many benefits but also creates some challenges for utilities. Effective energy management systems lead to an increase in grid sustainability and durability. It also supports smart grid functionality in various areas such as market control, integration of decentralized energy resources, cost control, and infrastructure construction. Different functions within the domain of energy management are discussed briefly in the following sections. Also, Table I provides a summary that can be used for further study of the topic of energy management, providing a classified set of references addressing the subject.

A. Demand Response

The aim of this unction is to lower the utility's total load by means of a reduction in the consumer's usage of electricity. There are two main approaches for demand response: (i) *incentive-based* where the utility can indirectly switch on/off certain appliances at a certain time at the consumer's premises, which carries some concerns about the consumer's privacy. (ii)*price-based* where customers adjust the electricity consumption in response to the price assigned by the suppliers. This is possible because of the use of a smart grid where two-way communication between the utility and consumer provides feedback to the consumer.

B. Home Energy Management Systems

A *Home Energy Management System (HEMS)*is capable of regulating household loads, scheduling appliances, and establishing the optimal load [42]. Today, HEMS must evaluate both energy generation (e.g., solar) and consumption at the same time in order to reduce energy costs [43].

HEMS is a scheduling problem that aims to produce optimal production and consumption schedules by considering multiple objectives such as reducing energy costs, environmental concerns, and consumer comfort.

C. Dynamic Pricing

Electricity is priced using a variety of tariffs, including basic tariffs, flat-rate tariffs, block rate tariffs, two-part tariffs, maximum demand tariffs, and power factor tariffs. Dynamic pricing was adopted to meet smart grid requirements such as fair pricing. [44]. The ultimate aim of dynamic pricing is to reduce the peak to the average consumption of energy using approaches like real-time pricing (RTP), critical peak pricing (CPP), time of use (ToU) rate, and extreme day pricing (EDP) [4].

D. Microgrids

With many regions of the world facing increasingly significant environmental challenges, the use of renewable energy sources such as solar and wind has become a prominent trend. Distributed generation (DG) is an efficient approach to utilizing renewable energy sources.

The term "microgrid" refers to the integration of renewable energy resources made feasible by storage capacity [45]. Although a complete definition of microgrids is still being contested in technical forums, a microgrid may be defined as a cluster of loads, distributed producing units, and an energy storage system. Microgrids

TABLE I. Articles on Energy Management Functions

Function		References
Demand Response	Price-Based	Ambreen et al., 2017 [47]
		Mahmud & Sant, 2017 [34]
	Incentive-Based	Shafie-khah et al., 2017 [35]
Home Energy Management Systems		Muratori & Rizzoni, 2016 [37]
		Tasdighi et al., 2014 [48]
Dynamic Pricing		Assaf, Osman, & Hassan, 2016 [49]
		Shuai, Hu, Peng, Tu, & Shen, 2017 [50]
Microgrid		Nafi et al., 2016 [4]
		Shuai et al., 2017 [50]

TABLE II. Articles Classified Based on Energy Management Techniques

Technique	References
Game Theory	Assaf et al., 2016 [49]
	Yu & Hong, 2017 [52]
Optimization	Mahmud & Sant, 2017 [34]
	Shafie-khah et al., 2017 [35]
	Tariq et al., 2017 [36]

are intended to increase the reliability of the electricity supply. A single point of contact connects each microgrid to the host power system at the distribution level. [46]. The microgrid effectively will integrate new trends into a power system and overcome their challenges.

Table I provides a summary that can be used for further study of the topic of energy management, providing a classified set of references addressing the subject.

Energy Management Techniques

There are various approaches to energy management such as peak clipping, load shifting, conservation, valley filling, load building, and flexible load. The objective of all these approaches is to convert stochastic demand to steady as much as possible [51].

There are many algorithms in the literature for solving multi-objective optimization problems in general; among them, the two most popular are game theory and optimization. Table II gives a summary of work conducted by these methods.

6. Data

Although in recent years the importance of smart meter analysis for sustainable development in smart cities was discussed and this type of analysis needs open access data set, only two open-access datasets are available for real-world smart meter data (http://www.ucd.ie/issda/,https://data.london.gov.uk/dataset/). smart meter dataset is an m×n matrix, where m is the number of reading spare time and is the number of households. London data set is big volume data and because of challenges come with big data many papers on different tasks used smaller data sets collected locally.

7. Challenges and opportunities

data volume and computational cost, combining multivariate data, and data velocity is the main challenge of smart meter analysis In the following sections, open research issues in detail are presented.

A. Volume

Potentially, smart meters produce a huge number of data points. However, current studies mostly concentrate on medium-sized datasets. Thus, computational techniques for handling and analyzing big data were rarely investigated in the literature. In order to address this issue, highly efficient algorithms and tools, such as distributed and parallel computing; as well as the deployment of computational platforms such as DASK, Hadoop's map-reduce, and Spark should be further investigated.

Manipulation of big data sets can be efficiently performed through one of the following approaches which have not been exploited fully within the smart meter data field:

1) *Cloud-based analysis* in environments such as Colab or Databricks which offers a Platform as a Service (PaaS) and Software as a Service (SaaS).
2) *High-performance computation* using hardware such as GPUs which can facilities highly efficient parallel computing but equally challenge practitioners in requiring redesign of existing algorithms and techniques.

Recently, data volume-related techniques for smart meter analysis have piqued the interest of researchers such as [53], who proposed a novel core-broker-client system architecture for big data analytics, capable of performing data storage, query, analysis, and visualization tasks on large data sets at a 20TB scale.

B. Velocity

Although many incremental and online machine learning algorithms have been proposed in other big data applications, existing work on smart meters does not adequately consider the potential of online algorithms. In the following, some relevant works in each domain will be described.

In the domain of knowledge discovery, Wang et al. [54] proposed an incremental algorithm that was implemented and evaluated on a synthesized data set. Khan et al. [55] present a clustering method to group different factories by applying clustering on stream data produced by smart meters.

In the domain of forecasting, Vrablecová et al. [56] proposed an online version of *support vector regression* (SVR) for load forecasting; while Bahrami et al. [57] present an online learning algorithm to address demand response tasks within the energy management domain.

C. Multivariate Data Analysis

Current studies mainly focus on the data produced by the smart meter. Integrating more external data, such as crowd-sourced data from the Internet, weather data, voltage, and current data, and special events data (i.e. calendar-based events such as holidays or weekdays) may reveal much more information and insight [58]. This aspect of data analytics clearly contains a number of challenges and is not yet mature enough due to the difficulty intrinsically found in data integration. Although visualization is a critical component of data analytics in data science and is performed prior to fitting models, making predictions, or drawing conclusions, visualizing high dimensional and multivariate data to highlight the critical components and discover hidden patterns or correlations among these data is a very rarely performed in the literature on smart meter data. [59].

8. Conclusion

This paper presented a review of cutting-edge research in the field of smart meter data analytics. The body of knowledge was structured using a new taxonomy that spans the research into three main domains: forecasting, knowledge discovery, and energy management. The three areas were reviewed by investigating various functions while discussing techniques and challenges.

The review showed the growing importance of data science elements such as data cleaning, data integration, and data transformation for achieving forecasting results. It also showed that load profiling in particular is an essential part of all domains of smart meter analytics. The growing importance of machine learning tasks such as feature extraction and dimensionality reduction, which leads to more cost-effective methods in terms of computational cost.

Recently, several researchers attempted the use of hybrid methods to improve the efficiency and reliability of different functions. There is still a lot of demand for investigating techniques and algorithms that efficiently utilized and integrate smart meter data with other sources to produce useful insight and knowledge. Recent advances in deep learning and big data only highlight new challenges and the need for more work in this area.

References

[1] Y. Wang, Q. Chen, T. Hong, C. Kang, Review of smart meter data analytics: Applications, methodologies, and challenges, IEEE Transactions on Smart Grid (2018).

[2] K. Amasyali, N. M. El-Gohary, A review of data-driven building energy consumption prediction studies, Renewable and Sustainable Energy Reviews 81 (2018) 1192–1205.

[3] S. Kakran, S. Chanana, Smart operations of smart grids integrated with distributed generation: A review, Renewable and Sustainable Energy Reviews 81 (2018) 524–535.

[4] N. S. Nafi, K. Ahmed, M. A. Gregory, M. Datta, A survey of smart grid architectures, applications, benefits and standardization, Journal of Network and Computer Applications 76 (2016) 23–36.

[5] Y. Wang, Q. Chen, C. Kang, M. Zhang, K. Wang, Y. Zhao, Load profiling and its application to demand response: A review, Tsinghua Science and Technology 20 (2) (2015) 117–129.

[6] C. Deb, F. Zhang, J. Yang, S. E. Lee, K. W. Shah, A review on time series forecasting techniques for building energy consumption, Renewable and Sustainable Energy Reviews 74 (2017-07-01) 902–924. doi:10.1016/j.rser.2017.02.085.

[7] P. Vrablecová, A. Bou Ezzeddine, V. Rozinajová, S. Šárik, A. K. Sangaiah, Smart grid load forecasting using online support vector regression, Computers and Electrical Engineering (2018). arXiv:S0045790617320645, doi:10.1016/j.compeleceng.2017.07.006.

[8] Y. Wei, X. Zhang, Y. Shi, L. Xia, S. Pan, J. Wu, M. Han, X. Zhao, A review of data-driven approaches for prediction and classification of building energy consumption, Renewable and Sustainable Energy Reviews 82 (2018) 1027–1047.

[9] K. Zor, O. Timur, A. Teke, A state-of-the-art review of artificial intelligence techniques for short-term electric load forecasting, in: Energy (IYCE), 2017 6th International Youth Conference on, IEEE, 2017, pp. 1–7.

[10] D. Alahakoon, X. Yu, Smart electricity meter data intelligence for future energy systems: A survey, IEEE Transactions on Industrial Informatics 12 (1) (2016) 425–436.

[11] B. Yildiz, J. I. Bilbao, J. Dore, A. B. Sproul, Recent advances in the analysis of residential electricity consumption and applications of smart meter data, Applied Energy 208 (2017-12-15) 402–427. doi:10.1016/j.apenergy.2017.10.014.

[12] E. Iskandarnia, H. Al-Amal, W. Al-Madany, Load forecasting in different scale and horizon-a review (2021).

[13] E. M. Eskandar, H. Alammal, W. Almadany, What prophet says about electrical consumption-forecasting techniques for big temporal data (2021).

[14] E. Eskandarnia, M. AlHammad, Predication of future energy consumption using sarimax (2021).

[15] J. Kwac, J. Flora, R. Rajagopal, Household energy consumption segmentation using hourly data, IEEE Transactions on Smart Grid 5 (1) (2014-01) 420–430. doi:10.1109/tsg.2013.2278477.

[16] A. Ozawa, R. Furusato, Y. Yoshida, Determining the relationship between a household's lifestyle and its electricity consumption in japan by analyzing measured electric load profiles, Energy and Buildings 119 (2016-05) 200–210. doi:10.1016/j.enbuild.2016.03.047.

[17] A. Al-Wakeel, J. Wu, K-means based cluster analysis of residential smart meter measurements, Energy Procedia 88 (2016-06-01) 754–760. doi:10.1016/j.egypro.2016.06.066.

[18] Y. Li, P. Guo, X. Li, Short-term load forecasting based on the analysis of user electricity behavior, Algorithms 9 (4)

(2016). arXiv:edselc.2-52.0-84999885171, doi:10.3390/a9040080.

[19] E. Villar-Rodriguez, J. Del Ser, I. Oregi, M. N. Bilbao, S. Gil-Lopez, Detection of non-technical losses in smart meter data based on load curve profiling and time series analysis, Energy 137 (2017-10-15) 118–128. doi:10.1016/j.energy.2017.07.008.

[20] D. Vercamer, B. Steurtewagen, D. Van den Poel, F. Vermeulen, Predicting consumer load profiles using commercial and open data, IEEE Transactions on Power Systems 31 (5) (2016) 3693–3701.

[21] S. Aghabozorgi, A. Seyed Shirkhorshidi, T. Ying Wah, Time-series clustering – a decade review, Information Systems 53 (2015-10-01) 16–38. doi:10.1016/j.is.2015.04.007.

[22] J. A. Fonseca, C. Miller, A. Schlueter, Unsupervised load shape clustering for urban building performance assessment, Energy Procedia 122 (2017-09-01) 229–234. doi:10.1016/j.egypro.2017.07.350.

[23] Z. Guo, K. Zhou, C. Zhang, X. Lu, W. Chen, S. Yang, Residential electricity consumption behavior: Influencing factors, related theories and intervention strategies, Renewable and Sustainable Energy Reviews 81 (2018) 399–412.

[24] R. R. Rathod, R. D. Garg, Regional electricity consumption analysis for consumers using data mining techniques and consumer meter reading data, International Journal of Electrical Power & Energy Systems 78 (2016) 368–374.

[25] L. Stankovic, V. Stankovic, J. Liao, C. Wilson, Measuring the energy intensity of domestic activities from smart meter data, Applied Energy 183 (2016-12) 1565–1580. doi:10.1016/j.apenergy.2016.09.087.

[26] L. Sun, K. Zhou, S. Yang, Regional difference of household electricity consumption: An empirical study of jiangsu, china, Journal of Cleaner Production 171 (2018-01-10) 1415–1428. doi:10.1016/j.jclepro.2017.10.123.

[27] M. Azaza, F. Wallin, Evaluation of classification methodologies and features selection from smart meter data, Energy Procedia 142 (2017) 2250–2256.

[28] X. Luo, T. Z. Hong, Y. X. Chen, M. A. Piette, Electric load shape benchmarking for small- and medium-sized commercial buildings, Applied Energy 204 (2017-10) 715–725. doi:10.1016/j.apenergy.2017.07.108.

[29] A. Albert, R. Rajagopal, Thermal profiling of residential energy use, IEEE Transactions on Power Systems 30 (2) (2015-03) 602–611. doi:10.1109/tpwrs.2014.2329485.

[30] K. Hopf, M. Sodenkamp, I. Kozlovkiy, T. Staake, Feature extraction and filtering for household classification based on smart electricity meter data, Computer Science-Research and Development 31 (3) (2016) 141–148.

[31] H. Yang, J. Zhang, J. Qiu, S. Zhang, M. Lai, Z. Y. Dong, A practical pricing approach to smart grid demand response based on load classification, IEEE Transactions on Smart Grid 9 (1) (2018) 179–190. doi:10.1109/TSG.2016.2547883.

[32] S. Pawar, B. Momin, Smart electricity meter data analytics: A brief review, in: IEEE Region 10 Symposium (TENSYMP), 2017, IEEE, 2017, pp. 1–5.

[33] K. Tornai, L. Kovács, A. Oláh, R. Drenyovszki, I. Pintér, D. Tisza, J. Levendovszky, Classification for consumption data in smart grid based on forecasting time series, Electric Power Systems Research 141 (2016) 191–201.

[34] A. A. Mahmud, P. Sant, Real-time price savings through price suggestions for the smart grid demand response model, in: Smart Grid and Cities Congress and Fair (ICSG), 2017 5th International Istanbul, IEEE, 2017, pp. 65–69.

[35] M. Shafie-khah, S. Javadi, P. Siano, J. Catalão, Optimal behavior of smart households facing with both price-based and incentive-based demand response programs, in: PowerTech, 2017 IEEE Manchester, IEEE, 2017, pp. 1–6.

[36] M. Tariq, A. Khalid, I. Ahmad, M. Khan, B. Zaheer, N. Javaid, Load scheduling in home energy management system using meta-heuristic techniques and critical peak pricing tariff, in: International Conference on P2P, Parallel, Grid, Cloud and Internet Computing, Springer, 2017, pp. 50–62.

[37] M. Muratori, G. Rizzoni, Residential demand response: Dynamic energy management and time-varying electricity pricing, IEEE Transactions on Power systems 31 (2) (2016) 1108–1117.

[38] S. S. Hosseini, K. Agbossou, S. Kelouwani, A. Cardenas, Non-intrusive load monitoring through home energy management systems: A comprehensive review, Renewable and Sustainable Energy Reviews 79 (2017-11-01) 1266–1274. doi:10.1016/j.rser.2017.05.096.

[39] N. F. Esa, M. P. Abdullah, M. Y. Hassan, A review disaggregation method in non-intrusive appliance load monitoring, Renewable and Sustainable Energy Reviews 66 (2016) 163–173.

[40] K. Basu, V. Debusschere, S. Bacha, U. Maulik, S. Bondyopadhyay, Nonintrusive load monitoring: A temporal multilabel classification approach, IEEE Transactions on industrial informatics 11 (1) (2015) 262–270.

[41] G. M. Messinis, N. D. Hatziargyriou, Review of non-technical loss detection methods, Electric Power Systems Research 158 (2018) 250–266.

[42] M. Rodriguez Fernandez, A. C. Garcia, I. G. Alonso, E. Z. Casanova, Using the big data generated by the smart home to improve energy efficiency management, Energy Efficiency 9 (1) (2016-01) 249–260. doi:10.1007/s12053-015-9361-3.

[43] J. Han, C.-S. Choi, W.-K. Park, I. Lee, S.-H. Kim, Smart home energy management system including renewable energy based on ZigBee and PLC, IEEE Transactions on Consumer Electronics 60 (2) (2014) 198–202.

[44] A. R. Khan, A. Mahmood, A. Safdar, Z. A. Khan, N. A. Khan, Load forecasting, dynamic pricing and DSM in smart grid: A review, Renewable and Sustainable Energy Reviews 54 (2016-02-01) 1311–1322. doi:10.1016/j.rser.2015.10.117.

[45] D. E. Olivares, A. Mehrizi-Sani, A. H. Etemadi, C. A. Cañizares, R. Iravani, M. Kazerani, A. H. Hajimiragha, O. Gomis-Bellmunt, M. Saeedifard, R. Palma-Behnke, Trends in microgrid control, IEEE Transactions on smart grid 5 (4) (2014) 1905–1919.

[46] H. Suyanto, R. Irawati, Study trends and challenges of the development of microgrids, in: 2017 15th International Conference on Quality in Research (QiR) : International Symposium on Electrical and Computer Engineering, 2017-07-24, pp. 383–387. doi:10.1109/QIR.2017.8168516.

[47] K. Ambreen, R. Khalid, R. Maroof, H. N. Khan, S. Asif, H. Iftikhar, Implementing critical peak pricing in home energy management using biography based optimization and genetic algorithm in smart grid, in: International Conference on Broad-

band and Wireless Computing, Communication and Applications, Springer, 2017, pp. 560–569.

[48] M. Tasdighi, H. Ghasemi, A. Rahimi-Kian, Residential microgrid scheduling based on smart meters data and temperature dependent thermal load modeling, IEEE Transactions on Smart Grid 5 (1) (2014-01) 349–357. doi:10.1109/tsg.2013.2261829.

[49] T. Assaf, A. H. Osman, M. Hassan, Fair autonomous energy consumption scheduling based on game-theoretic approach for the future smart grid, in: 2016 UKSim-AMSS 18th International Conference on Computer Modelling and Simulation (UKSim), 2016-04-06, pp. 235–239. doi:10.1109/UKSim.2016.67.

[50] Z. Shuai, Y. Hu, Y. Peng, C. Tu, Z. J. Shen, Dynamic stability analysis of synchronverter-dominated microgrid based on bifurcation theory, IEEE Transactions on Industrial Electronics (2017).

[51] B. P. Esther, K. S. Kumar, A survey on residential demand side management architecture, approaches, optimization models and methods, Renewable and Sustainable Energy Reviews 59 (2016) 342–351.

[52] M. Yu, S. H. Hong, Incentive-based demand response considering hierarchical electricity market: A stackelberg game approach, Applied Energy 203 (2017) 267–279.

[53] T. Wilcox, N. Jin, P. Flach, J. Thumim, A big data platform for smart meter data analytics, Computers in Industry 105 (2019) 250–259.

[54] J. Wang, R. Cardell-Oliver, W. Liu, An incremental algorithm for discovering routine behaviours from smart meter data, Knowledge-Based Systems 113 (2016) 61–74.

[55] I. Khan, J. Z. Huang, K. Ivanov, Incremental density-based ensemble clustering over evolving data streams, Neurocomputing 191 (2016) 34–43.

[56] P. Vrablecová, A. B. Ezzeddine, V. Rozinajová, S. Šárik, A. K. Sangaiah, Smart grid load forecasting using online support vector regression, Computers & Electrical Engineering 65 (2018) 102–117.

[57] S. Bahrami, V. W. Wong, J. Huang, An online learning algorithm for demand response in smart grid, IEEE Transactions on Smart Grid 9 (5) (2017) 4712–4725.

[58] Y. Wang, Q. Chen, T. Hong, C. Kang, Review of smart meter data analytics: Applications, methodologies, and challenges, IEEE Transactions on Smart Grid 10 (3) (2018) 3125–3148.

[59] R. J. Hyndman, X. A. Liu, P. Pinson, Visualizing big energy data: Solutions for this crucial component of data analysis, IEEE Power and Energy Magazine 16 (3) (2018) 18–25.

Elham M. Eskandarnia Elham Eskandarnia received her BSc in Software Engineering and her MSc in IT from the University of Bahrain in 2016 with Honors. Currently, she is enrolled in the Ph.D. program at the University of Bahrain. She is a member of the faculty of the college of computer and Engineering at the British University of Bahrain. Her research interests include data science, machine learning, and deep learning.

Dr. Hesham M. Al-Ammal received his BSc (honors) in Computer Engineering from King Fahd University of Petroleum and Minerals, Dhahran, Saudi Arabia in 1990. He received his MSc with honors in Computer Science from Louisiana State University, Baton Rouge in 1993; and his Ph.D. with honors in Computer Science from the University of Warwick, Coventry, the UK in 2000. He is a senior member of the Association for Computing Machinery. He served as Chair of the Department of Computer Science (2006 to 2010), Director of Quality Assurance Center (2009 to 2011), Dean of the College of IT, University of Bahrain, (2011 to 2017). His research interests include randomized algorithms, Markov chains, cyber-security, encryption, and data analytics.