Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 168 (2020) 210–216

Procedia Computer Science

www.elsevier.com/locate/procedia

Complex Adaptive Systems Conference with Theme:
Leveraging AI and Machine Learning for Societal Challenges, CAS 2019

Event-Driven System For Proficient Load Recognition by Interpreting the Smart Meter Data

Saeed Mian Qaisar*, Futoon Alsharif

Effat University, College of Engineering, Jeddah, 21478, KSA

Abstract

The technological advancements have evolved the deployment of smart meters. A fine-grained metering data collection and analysis is necessary to bring benefits to multiple smart grid stakeholders. The classical sensing mechanism is time-invariant. Therefore, it results in the collection, transmission, and processing of a large amount of unnecessary data. This work employs the event-driven sensing mechanism to achieve real-time data compression. Afterward, the novel adaptive rate techniques are employed for the data conditioning, segmentation, and extraction of features. The pertinent features regarding the appliances' consumption patterns are afterward used for their identification. It is realized by employing the mature Support Vector Machine and k-Nearest Neighbor classifiers. Results confirm a 3.4 times compression gain and the computational effectiveness of the suggested solution while securing 95.4% classification precision. It shows the benefits of integrating the proposed method in the realization of current energy efficiency services like enumerated consumption billing, effective load identification, and dynamic load management.

*Keywords:*Smart Meter Data; Consumption Pattern; Automatic Load Identification; Event-Driven sensing; Adaptive Rate Processing; Features Extraction; Machine Learning.

1. Introduction

A smart meter measures the consumed electricity and is able of controlling the supply remotely. The technological advancements have evolved the deployment of smart meters in place of the conventional ones [1], [2]. These meters are the vital elements of smart grids and are offering significant advantages to various stakeholders in terms of social, environmental, and economic constraints [3]. The

* Corresponding author. Tel.: +966122137849; fax: +966126377447
*E-mail address:*sqaisar@effatuniversity.edu.sa

1877-0509 © 2020 The Authors. Published by Elsevier B.V.

10.1016/j.procs.2020.02.267

broad installations of smart meters allow an enormous amount of data collection with a wanted granularity [4]. Automatic data acquisition, transmission, processing, and analysis are key factors behind the success of smart meters. The process is depicted with the help of Fig. 1.

Fig. 1. Components of the smart meter data intelligence chain [3]

A fine-grained metering data collection is necessary to bring practical benefits to the multiple smart grid stakeholders in terms of efficiency and sustainability [3], [4].

The metering data is processed and analyzed to extract its pertinent features. In [5] and [6], authors have defined the process of feature extraction. It is the process of establishing a set of features that can most meaningfully represent the information that is important for analysis and classification. There are many techniques for features extraction of appliances consumption patterns like Short–Time Fourier Transform (STFT), wavelet transform, and K–means algorithm. Besides, the power factor (PF) and the harmonic distortion in the consumption pattern can also be used to identify loads [7], [8], [20]. The extracted features are peak values, average values, Root Mean Square (RMS) values of the consumption and its harmonics.

The device pattern recognition is performed by using the extracted features. In [6], authors have defined pattern recognition as a study of how machines can observe, learn to distinguish, and make reasonable decisions about the pattern categories. In [7], authors have used pattern recognition for load disaggregation. The principle is to get the particular data of the appliance without intrusion. The device database includes load characteristics, used to perform the identification [7]. In [8], the authors used pattern recognition for leveraging smart meter data to disaggregate the total consumption of electricity. The pattern recognition is based on the classification. A method used to assign a description of a class to any given point in the feature space is called a classifier [6]. In this framework, a multitude of classification algorithms are used like k-Nearest Neighbour (KNN), Support Vector Machine (SVM), Artificial Neural Network (ANN), Random Forest, and Rotation Forest [8].

The classical data acquisition is time-invariant. The data is acquired at a Nyquist rate irrespective of its information rate and therefore results in the collection, transmission, processing, and analysis of a noticeably large amount of unnecessary data [9]. In this context, the event-driven sensing mechanism is used to achieve real-time data compression at the stage of data acquisition. In the next step, the original adaptive rate techniques are proposed for the data conditioning, segmentation, and extraction of features. It affirms a noteworthy compression and computational effectiveness of the designed solution compared to the traditional counterparts.

The following part of the paper consists of three main sections which are materials and methods, results, and conclusion.

2. Materials and Methods

Fig. 2 shows the principle of the proposed system. Following subsections describe different system modules.

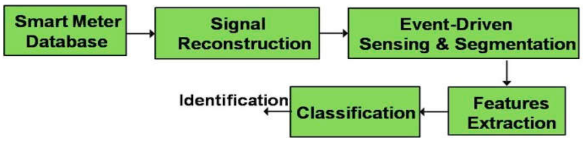

Fig. 2. The system blockdiagram

2.1. The Smart Meter Database

The database ACS-F2, which is composed of 15 categories of major home appliances, is used [10]. This database is composed of 225 appliances consumption parameters. Two acquisition sessions of one hour are carried out for each device [10]. The electric consumption parameters like real power, reactive power, RMS current, RMS voltage, frequency, and phase of voltage relative to current are recorded. These are acquired at the rate of 0.1Hz. The recordings are made in a disaggregated fashion. The electric consumption signature of one appliance is registered at a time with Plogg smart meter, designed by the Energy Optimizers Limited [18].

At first, the four categories of appliances are considered. It includes coffee machines, televisions, microwave ovens, and fridges & freezers. For each intended device, real power consumption is taken into account.

2.2. The Signal Reconstruction

The reconstruction process is the opposite of the sampling process, which is also known as interpolation [11]. The considered consumption parameters waveforms are up-sampled with a factor of 10000. It is performed to evaluate the event-driven sensing module.

The up-sampling is realized by using an assembly of four cascaded cubic-spline interpolators and anti-aliasing filters. After up-sampling,the incoming signal $y(t_n)$ is converted into its quasi analog version$\tilde{x}(t)$. The relationship between $y(t_n)$ and $\tilde{x}(t)$ can be mathematically presented by using Equation (1). Where U is the up-sampling factor.

$$\tilde{x}(t) = y(\tfrac{t_n}{U}). \tag{1}$$

2.3. The Event-Driven Sensing (EDS)

The Analog to Digital Converter (ADC) is required in contemporary smart meters. The classical ADCs are based on Nyquist's theory of sampling and processing [11]. The design parameters of the traditional ADCs are thus selected for the most unfavorable case [11]. In this way, in the case of arbitrary signals such as consumption parameters of the equipment, such ADCs are not effective [12]-[15]. The event-driven ADCs (EDADCS) are used in this context [16], [20]. They are based on the event-driven sensing (EDS) and can change their sampling frequency depending on incoming signal disparities [14], [15]. A sample is taken just after the input band-limited analog signal $\tilde{x}(t)$crosses one of the predefined thresholds. Therefore, samples are non-uniformly divided in time. The frequency of the taken samples is dependent on the $\tilde{x}(t)$variations [14], [15]. The process can be mathematically expressed by using Equation (2). Where, dt_n is the time distance between the current, t_n, and the previous, t_{n-1}, sampling instants. In the case of EDS, each sample is a pair, (x_n, t_n), of an amplitude x_n and a time t_n. x_n is equivalent to one of the predefined thresholds and t_n is clear from Equation (2).

$$t_n = t_{n-1} + dt_n . \tag{2}$$

The EDADC only acquires the related information while the rest of the signal is overlooked. Hence, a noticeable real-time reduction, compression gain, is attained in the acquired number of samples in comparison with the traditional counterparts. It further adds in the post-processing activity reduction and the system processing and power consumption efficiency [12]-[15].

2.4. The Event-Driven Segmentation

The segmentation is the process of dividing the signal into several fixed-length windows [11]. It allows analyzing and effectively extracting the pertinent parameters of the incoming signal. The output of EDS is non-uniformly placed in time. Therefore, it cannot be segmented by using the traditional windowing algorithms [11]-[13]. A novel Activity Selection Algorithm (ASA) is used in this framework [12], [13]. It segments the EDADC output in variable-length windows. The ASA is performed by utilizing the sampling non-uniformity, and it conserves the useful information, like consecutive sampling instants repartitioning, count of samples, etc. [15]. It allows the post adaptive rate features extraction and results in the extraction of pertinent classifiable features in the time domain.

2.5. The Features Extraction

The features to be classified are mined from each segment. Because of the event-driven acquisition and, the segments allow extracting important information about the signal frequency content in time-domain [17]. Therefore, in contrast to alternative solutions, based on frequency or time-frequency based analysis, the employed features extraction mechanism does not require the computationally complex frequency domain transformation and analysis operations [11]. The pertinent classifiable features are extracted by using the non-uniform data in time-domain.

For each intended appliance, real power consumption is considered. Each parameter waveform is reconstructed, acquired with the EDADC and segmented with the ASA. Afterward, four different features are extracted for each segment. Let i is indexing the i^{th} selected segment W^i. Then C^i, ΔA^i, A^i_{max}, and dt^i_{mean} are respectively the extracted number of threshold crossings, the peak-to-peak amplitude, the maximum amplitude and the average sampling step for W^i.

2.6. The Classification Techniques

The extracted features are used for the recognition of intended appliances. The k-Nearest Neighbor (KNN) and the Support Vector Machine (SVM) classifiers are used in this framework.

The KNN depends on the nearest feature space training examples [6]. It classifies the concerned object according to its majority vote by its neighbors [6]. The neighbors are derived from a series of objects known for the right classification [6]. These neighbors are considered as a training series for this algorithm. In this model, the learning is based on the storage, together with their class labels, of all training cases which match points in N-dimensional Euclidean space and are postponed until another example arrives [6]. The classification of the recent unlabeled query example or vector is performed by assigning the tag that is the most common among the closest k-training samples [6].

The SVM builds a hyperplane with N-dimension, which divides the information into two optimal groups [6]. It is extensively used to evaluate and learn from initial data, and then use it for classification and regression analysis, in a group of associated supervised learning techniques [6]. The standard SVM is a double class SVM. It takes input data from the two different possible classes and predicts the possibility for each class, making it a non-likely binary linear classifier among the two possible classes [6]. Given the collection of training samples in each class, the SVM learning algorithm constructs a model that takes new instances in one category or another [6]. The SVM classifiers are also used in the multi-class classification in which the classifiable data can belong to any class [6].

2.7. Evaluation Measures

2.7.1. Compression gain

In classical sensing, the incoming analog signal is acquired at a fixed rate. Therefore, the total number of samples, acquired for a considered time length L_T is straight forward to compute. Let F_{ref} be the selected sampling frequency for the acquisition of the smart-meter data.Then, the sampling step, the distance between two consecutive sampling instants, $T_{ref} = \frac{1}{F_{ref}}$ is unique in this case. Therefore, the acquired number of samples N for the considered $L_T$can be calculated by using Equation (3).

$$N = F_{ref} \times L_T \,. \tag{3}$$

For the event-driven sensing, the sampling frequency is not unique, and it adapts as a function of the input signal temporal variations [14], [15]. Therefore, for a considered time length L_T the acquired number of samples can be different and are a function of the EDADC resolution, the employed quantization scheme, and the signal characteristics [14]-[17]. Let N_{ED} is the number of samples obtained in the devised solution. Then the compression gain, G_{COMP}, can be calculated by using Equation (4).

$$G_{COMP} = \frac{N}{N_{ED}} \,. \tag{4}$$

2.7.2. Classification accuracy

The performance of the classification is evaluated in terms of precision. Equation (5) expresses the process mathematically. Where true negative (TN) and true positive (TP) classifications are accurate. An illustrative example of TP is the classification of a microwave instance as a microwave. A TN example is that a microwave instance is not categorized as either a refrigerator or a TV. A false positive (FP) occurs when it is fundamentally contrary when the outcome is not properly expected as positive. An instance of FP is the classification of microwave or fridge as television. If the result is not correctly recognized as adverse, a false negative (FN) occurs when it is a positive [6]. An example of FN is the failure to identify a microwave instance as a microwave.

$$Accuracy = \frac{T_P + T_N}{T_P + T_N + F_P + F_N} \times 100\% \,. \tag{5}$$

3. Results

To illustrate the interesting features of the proposed technique a case study is employed. In this framework, the intended appliance consumption parameters are taken from the ACS-F2 database [10]. At first, the four categories of appliances are considered. It includes coffee machines, televisions, microwave ovens, and fridges & freezers. Fifteen appliances from each category are considered. Therefore, in total 60 appliances are taken into account. Two acquisition sessions of one hour are carried out for each appliance. It results in 120 multi-dimensional time series of the considered appliances' electricity-related characteristics.

In the studied case, the real power consumption is considered from each multi-dimensional time series. It results in 120 instances for the studied appliances. The recordings are performed at a 0.1Hz sampling frequency. Afterward, each instance is up-sampled with a factor of 10000. In the classical case, a 12-Bit resolution A/D converter is used for acquiring the $\tilde{x}(t)$ [10]. However, in the devised solution the $\tilde{x}(t)$ is digitized with a 4-Bit resolution EDADC. It confirms a drastic diminishing of the suggested solution circuit complexity while comparing with the classical solution.

Examples of the considered real power consumption instances, acquired with the EDS mechanism are shown in Fig. 3.Fig. 3 displays the advantages of using EDS. It confirms that for a given resolution, M, and ΔV it adapts its sampling frequency following the signal temporal variations. Therefore, it acquires only the relevant signal information. It also diminishes the low amplitude noise by using its noise thresholding capability [12], [13] and enhances the precision of post features extraction and classification processes.

In this case, the EDADC output is segmented by using the ASA. The upper bound on the segment length is chosen as L_{ref} =5minutes [12], [13]. The ASA focuses only on the active signal parts and results in average 6 selected segments per incoming Instance. However, the number of samples per segment, N^i, can vary as a function of the $\tilde{x}(t)$ temporal variations [12], [13].

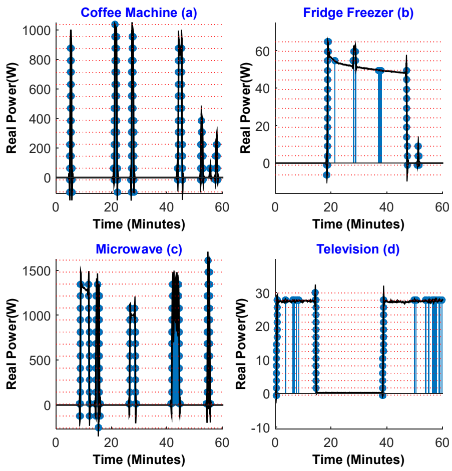

Fig. 3. The real power instances digitized with a 4-Bit resolution EDADC.

In the conventional case, $x(t)$is sampled constantly at 0.1Hz irrespective of its temporal variations. It renders an accumulation and processing of superfluous samples. In this case, L_{ref} =5minutes results in a fixed number of 30 samples per segment regardless of the signal variations. Consequently, it forces the post-processing modules to process insignificant data and as a result increases the system processing activity and power loss.

The obtained compression gains, calculated by using Equation (4), are summed up in Table 1. It depicts that the suggested solution attains an overall 2.1 times, 6.6 times, 1.6 times and 3.2 times compression gains respectively for the case of coffee machines, fridge, and freezers, microwaves and televisions. It results in an average compression gain of 3.4 times. It assures a noticeable decrease in the arithmetic complexity,and the power consumption of the suggested solution compare to the classical approach. Also, embedding the EDS in smart meters significantly decreases the activity of data storage and transmission compared to standard methods. It further augments the effectiveness of the designed strategy over the concurrent ones.

In the classical system, a reduced sampling rate and a greater quantization step may also be used. However, Firstly, it can lead to a reduction in the ADC performance in terms of the Signal to Noise Ratio (SNR), as the SNR output is inversely related to the quantum value used in the system[13], [19]. In the case of EDADC, the SNR performance is independent of the quantum value and is based on the sampling instants measuring counter clock frequency and resolution [13], [19]. Secondly, the uniform sampling-based approach will make it challenging to extract the relevant features of incoming instants in the time-domain. Such strategies, therefore, require the use of both time and frequency domain features extraction approaches and canincrease the computational complexity of the system and the processing load compared to the designed strategy.

Table 1. Summary of the Compression Gains

Appliances	Compression Gain	Average Compression Gain
Coffee Machines	2.1	3.4
Fridges and Freezers	6.6	
Microwaves	1.6	
Televisions	3.2	

In this case, 120 instances are considered for four considered categories of appliances. 70% of these Instances are used to prepare the reference templates and the remaining 30% is used for the testing purpose. To compensate for the limitation of dataset size the 10-folds cross-validation approach is employed [6].

The obtained percentage recognition accuracies are respectively summed up for the KNN and the SVM classifiers in Tables 2 and 3.

Table 2. Accuracy for the four-class appliances consumption pattern recognition (KNN)

Appliances	ClassificationAccuracy (%age)	Average ClassificationAccuracy (%age)
Coffee Machines	92.4	93
Fridges and Freezers	93.4	
Microwaves	92.5	
Televisions	93.8	

Table 3. Accuracy for the four-class appliances consumption pattern recognition (SVM)

Appliances	ClassificationAccuracy(%age)	Average ClassificationAccuracy (%age)
Coffee Machines	94.7	95.4
Fridges and Freezers	95.6	
Microwaves	94.9	
Televisions	96.3	

Table 2 depicts the obtained appliances' consumption signature recognition accuracies for the case of the KNN classifier. These are 92.4% for coffee machines, 93.4% for fridges and freezers, 92.5% for microwaves and 93.8% for televisions. The resulting average recognition accuracy is 93%.

The recognition accuracies for the case of SVM are summed up in Table 3. These are 94.7% for coffee machines, 95.6% for fridges and freezers, 94.9% for microwaves and 96.3% for televisions. The resulting average recognition accuracy is 95.4%.

Tables 2 and 3, show that for the studied case, the best average classification accuracy of 95.4% is obtained with the SVM method. The KNN follows with 93% accuracy. In this case, the SVM performs better because of its ability to avoid the most irrelevant nodes while classifying an instant under test. It concludes that the employed assembly of the EDADC, the ASA, the time-domain features extraction and the SVM provides the best appliances consumption patterns recognition.

4. Conclusion

A new method is proposed for automatic recognition of the major household appliances consumption pattern. It is based on the event-driven processing and the time-domain features extraction and classification. Unlike traditional tactics, it does not require the computationally complex frequency-domain feature extraction. It is demonstrated that the integration of EDADC and ASA significantly decreased the number of samples to process. A 3.4 times decrease in the collected number of samples is achieved over the classical approach. It confirms the devised system's drastic computational complexity reduction over classical counterparts.

It is demonstrated that the suggested approach attains an average appliance consumption pattern recognition accuracy of 95.4%. It confirms the interest of using the suggested solution in contemporary automatic dynamic load management and enumerated billing systems.

A future extension is to investigate the proposed approach while considering extended categories of appliances. The proposed system performance depends on the chosen system parameters like resolution, reference segment length, quantization scheme, parameter extraction, and the classification algorithms. Investigating the system performance for higher resolution, adaptive quantization schemes and other robust classifiers like Rotation Forest, Artificial Neural Networks, Random Forest, etc. is another prospect.

Acknowledgments

This paper is funded by Effat University, Jeddah, KSA. Authors are thankful to Dr. A. Subasi for the fruitful discussions. Authors are also thankful to anonymous reviewers for their valuable feedback.

References

[1] Darby, S. (2010). Smart metering: what potential for householder engagement?. Building Research & Information, 38(5), 442-457.

[2] Sun, Q., Li, H., Ma, Z., Wang, C., Campillo, J., Zhang, Q., ... & Guo, J. (2016). A comprehensive review of smart energy meters in intelligent energy networks. IEEE Internet of Things Journal, 3(4), 464-479.

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

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

[5] Mondini, V., Mangia, A. L., & Cappello, A. (2016). EEG-based BCI system using adaptive features extraction and classification procedures. Computational intelligence and neuroscience, 2016.

[6] Fu, K. S. (2019). Applications of pattern recognition. CRC press.

[7] Biansoongnern, S., & Plungklang, B. (2016). Non-Intrusive Appliances Load Monitoring (NILM) for Energy Conservation in Household with Low Sampling Rate. Procedia Computer Science, 86, 172-175. DOI:10.1016/j.procs.2016.05.049.

[8] Andreas Reinhardt, Paul Baumann, Daniel Burgstahler, Matthias Hollick, Hristo Chonov, Marc Werner, Ralf Steinmetz: On the Accuracy of Appliance Identification Based on Distributed Load Metering Data. Proceedings of the 2nd IFIP Conference on Sustainable Internet and ICT for Sustainability (SustainIT). October 2012.

[9] Verhelst, M., & Bahai, A. (2015). Where Analog Meets Digital: Analog-to-Information Conversion and Beyond. IEEE Solid-state circuits magazine, 7(3), 67-80.

[10] A. Ridi, Chr. Gisler, J. Hennebert. ACS-F2 – A New Database of Appliance Consumption Signatures, SoCPaR Conference, 2014.

[11] Ingle, V. K., & Proakis, J. G. (2016). Digital Signal Processing Using MATLAB: A Problem Solving Companion. Cengage Learning.

[12] Qaisar, S. M. (2018, July). A Computationally Efficient EEG Signals Segmentation and De-noising Based on an Adaptive Rate Acquisition and Processing. In 2018 IEEE 3rd International Conference on Signal and Image Processing (ICSIP) (pp. 182-186). IEEE.

[13] Qaisar, S. M., Fesquet, L., & Renaudin, M. (2014). Adaptive rate filtering a computationally efficient signal processing approach. Signal Processing, 94, 620-630.

[14] Qaisar, S. M., Dallet, D., Benjamin, S., Desprez, P., & Yahiaoui, R. (2013, May). Power-efficient analog to digital conversion for the Li-ion battery voltage monitoring and measurement. In 2013 IEEE International Instrumentation and Measurement Technology Conference (I2MTC) (pp. 1522-1525). IEEE.

[15] Qaisar, S. M., Yahiaoui, R., & Dominique, D. (2015, December). A smart power management system monitoring and measurement approach based on a signal driven data acquisition. In 2015 Saudi Arabia Smart Grid (SASG) (pp. 1-4). IEEE.

[16] Hou, Y., Qu, J., Tian, Z., Atef, M., Yousef, K., Lian, Y., & Wang, G. (2018). A 61-nW Level-Crossing ADC With Adaptive Sampling for Biomedical Applications. IEEE Transactions on Circuits and Systems II: Express Briefs, 66(1), 56-60.

[17] Moser, B. A. (2017). Similarity recovery from threshold-based sampling under general conditions. IEEE Transactions on Signal Processing, 65(17), 4645-4654.

[18] EOL (Energy Optimizers Limited), "ZigBee Smart EnergyMakes Electric Outlets Smart: Plogg smart meter plug bringsenergy-saving wireless intelligence to homes and buildings”, Business Wire, 2010.

[19] Qaisar, S. M., Yahiaoui, R., & Gharbi, T. (2013, September). An efficient signal acquisition with an adaptive rate A/D conversion. In 2013 IEEE International Conference on Circuits and Systems (ICCAS) (pp. 124-129). IEEE.

[20] Weiss, M., Helfenstein, A., Mattern, F., & Staake, T. (2012). Leveraging smart meter data to recognize home appliances. 2012 IEEE International Conference on Pervasive Computing and Communications.