



A Novel Approach for Detecting Anomalous Energy Consumption Based on Micro-Moments and Deep Neural Networks

Yassine Himeur¹ · Abdullah Alsalemi¹ · Faycal Bensaali¹ · Abbès Amira²

Received: 28 January 2020 / Accepted: 21 August 2020 / Published online: 25 September 2020
© The Author(s) 2020

Abstract

Nowadays, analyzing, detecting, and visualizing abnormal power consumption behavior of householders are among the principal challenges in identifying ways to reduce power consumption. This paper introduces a new solution to detect energy consumption anomalies based on extracting micro-moment features using a rule-based model. The latter is used to draw out load characteristics using daily intent-driven moments of user consumption actions. Besides micro-moment features extraction, we also experiment with a deep neural network architecture for efficient abnormality detection and classification. In the following, a novel anomaly visualization technique is introduced that is based on a scatter representation of the micro-moment classes, and hence providing consumers an easy solution to understand their abnormal behavior. Moreover, in order to validate the proposed system, a new energy consumption dataset at appliance level is also designed through a measurement campaign carried out at Qatar University Energy Lab, namely, Qatar University dataset. Experimental results on simulated and real datasets collected at two regions, which have extremely different climate conditions, confirm that the proposed deep micro-moment architecture outperforms other machine learning algorithms and can effectively detect anomalous patterns. For example, 99.58% accuracy and 97.85% F1 score have been achieved under Qatar University dataset. These promising results establish the efficacy of the proposed deep micro-moment solution for detecting abnormal energy consumption, promoting energy efficiency behaviors, and reducing wasted energy.

Keywords Energy consumption · Micro-moments · Deep neural network · Anomalies detection · Visualization · Energy efficiency

Introduction

Most of today's end-user behaviors provoke a high energy cost from keeping the room lights on to watching TV all

the day long, for instance. Undoubtedly, energy efficiency and green buildings have been getting an increased amount of attention in many countries in the last few years [1]. In short, it is the practice of minimizing energy usage without enduring a loss in quality. For many motives, end-users have to support the green energy industry from the first flip on light to the last push on the start button of computers [2, 3]. In fact, energy efficiency can enhance the manner buildings consume power in order to diminish detrimental effects on society, economy, and global environment [4].

In order to preserve what people deem to be a perfectly natural life; massive use of natural resources for building-up energy, pollution, and global warming effects should be reduced and optimized by energy users [5, 6]. For the sake of their environment and nature, they need for example to detach themselves from total reliance on fossil fuels or to cut-in demand to ensure fewer harmful carbon emissions in the atmosphere [7, 8]. Besides, hundreds of dollars could be saved on end-users' energy bills annually. For such financial-saving reasons, end-users must raise their

✉ Yassine Himeur
yassine.himeur@qu.edu.qa

Abdullah Alsalemi
a.alsalemi@qu.edu.qa

Faycal Bensaali
f.bensaali@qu.edu.qa

Abbès Amira
abbès.amira@dmu.ac.uk

¹ Department of Electrical Engineering, Qatar University, Doha, Qatar

² Institute of Artificial Intelligence, De Montfort University, Leicester, UK

willingness to sacrifice living comforts and to draw a red line under the undue and harmful energy wasting behaviors [9]. Forgetting to unplug the charger and leaving on standby unnecessary appliances are the most common contradicted behaviors to energy management among households [10, 11]. From a national perspective, conserving energy can aid to curtail the financial burden of oil imports and so to reduce the investment needed for the expansion of the electricity sector, especially in countries with high growth of the electricity demand [12, 13].

Recent studies demonstrate that roughly 10 to 40% of the electrical power can be conserved in domestic households if individuals' occupation profiles are embedded into building energy management systems [14]. Moreover, residential and commercial sectors consume up to 40% of the overall energy produced worldwide [15]. Furthermore, the consumption rate is predicted to be increased more in the upcoming years due to the improvement of living conditions and increasing use of appliances and electrical apparatus [16]. Overall, the conducted investigations illustrate that both the whole energy usage rate and plug loads are considerably related to the occupancy ratios [17, 18]. The presence/absence of individuals in a building can highly affect energy consumption rate since a big number of users keep appliances switched on for long time without their physical presence. This represents an anomalous behavior, especially for a set of appliances such as televisions, air conditioners, lights, laptops/desktops, and fans. End-users justified such abnormal behavior by laziness and lack of awareness [19, 20]. Consequently, energy efficiency potential can not be well invested without paving the way to detect and change abnormal power consumption behavior.

Detection of anomalies and abnormalities is a challenging field of research employed in many applications, including biomedical [21, 22], power generation [23], network traffic [24], cybersecurity [25], and energy consumption [26]. Detecting and analyzing abnormal energy usage patterns in real time can not only promote the process of energy saving, but it can help also tracing appliance failures through analyzing sudden and unexpected changes in energy usage [26]. Once the anomalous energy usage behavior is identified, the end-user will be notified, where he can correspondingly execute opportune power efficiency schemes. Furthermore, the proliferation of wireless sensors and sub-meters nowadays, surveying domestic consumption behaviors in buildings to detect abnormal usage becomes at hand. Hence, for global energy efficiency motives, a growing worldwide attention to put technologies accompanying the detection of anomalous consumption in use is being observed [27]. Supervised, semi-supervised and unsupervised anomaly detections are the three main techniques of identifying irregular energy patterns according to the availability or not of labeled data [26, 28].

Furthermore, since the advent of the Internet, mobile devices have become an indispensable part of our lives. Worldwide studies have shown that more Google searches now take place on mobile devices than on desktop computers [29]. This growing and strong reliance of today's consumers on smartphones is logical since they enable them not only to chat and share ideas with their friends or acquaintances, but also to shop, pay bills, reserve plane tickets, and even remotely control their own homes and cars right from anywhere. So, consumer's behavior is now altered by such dependency to meet their needs immediately and on the fly. Consumer's behavior on mobile devices varies widely within different situations and moments. In relation to that, early in 2015, the term of "micro-moments" was coined by Google [30]. The latter refers to instances or cases when a consumer spontaneously relies on a device (generally a mobile) to address an immediate need. Micro-moments are game changers for both consumers and brands. Google categorized micro-moments as follows: <I-want-to-know>, <I-want-to-go>, <I-want-to-do>, and <I-want-to-buy>. Inspired by this idea, recent works exploring the use of micro-moments to monitor, analyze, and understand energy consumption behavior have been presented [31–33].

In recent years, the promising progress in developing machine learning algorithms has substantiated their efficacy in various research fields. Starting from the fact that learning-based techniques may be useful to identify rare occurrences or observations (outliers) that arouse suspicion, their deployment for anomaly detection for several applications becomes prevalent, e.g., intrusion detection [34], fraud detection [35], abnormality detection in video surveillance [36], turbine combustor anomaly detection [23]. Moreover, machine learning algorithms can be combined with adequate feature extraction approaches to enhance their efficiency. Consequently, a well-trained algorithm with in-depth anomaly types will boost the anomaly detection effectiveness considerably while maintaining an acceptable cost and complexity [37].

Difficulties and Challenges

Detecting anomalous power consumption suffers from various issues, difficulties, and challenges that hinder the progress in this field. These problems are mainly domain-specific and they can be summarized as follows: (i) there is often no obvious definition of normal and anomalous energy consumption; (ii) there is no clear frontier between normal and anomalous behaviors; (iii) missing unified metrics deployed for performance assessment; and (iv) there is missing of ground-truth data, each group of researchers uses their own data. Thus, this results in a major challenge for reproducing the outputs of existing solutions; therefore,

it is impossible to check which frameworks present the state-of-the-art in various scenarios.

Specifically, one of the major obstacles to develop and evaluate anomaly detection techniques is the lack of labeled ground-truth datasets. Simply put, there is a lack of works that discuss how to label energy consumption observations into normal or abnormal, and, to be even more precise, which kind of abnormalities they are. To that end, we propose a novel approach for labeling energy consumption occurrences using occupancy patterns, power consumption footprints, and the micro-moments paradigm.

In addition, it is difficult to find what are the best classification schemes that can be deployed to automatically classify energy consumption observations and further what are the appropriate evaluation metrics. Consequently, we propose in this framework a deep neural network (DNN) algorithm to identify anomalies; its performance is compared with that of various machine learning classifiers. Moreover, a set of metrics is selected to assess their performance and conduct a comprehensive comparison.

Paper Contributions

As explained in the aforementioned section, the lack of ground-truth data hinders developing and testing novel techniques because of the absence of an explicit methodology to test their performance and efficiency. Aiming at addressing the challenges encountered in the validation of anomaly detection solutions, our paper provides a set of original contributions that can be summarized as follows:

- Inspired by the micro-moment paradigm, a novel approach to detect abnormal energy consumption is introduced in which, a definition of normal and abnormal power consumption is introduced and power consumption observations are divided into five major classes, denoted as follows: “good usage,” “turn on a device,” “turn off a device,” “excessive consumption,” and “consumption while outside.”
- In addition to detecting excessive power consumption anomalies, a novel form of abnormal consumption is detected, which is mainly related to the consumption while end-users are absent (i.e., outside the household/office). In contrast to existing frameworks, which can only detect anomalies related to excessive/low consumption using traditional outlier detection methods.
- A novel rule-based algorithm to label power consumption datasets is proposed based on the micro-moment paradigm as a feature extraction module. Practically, the proposed system exploits the micro-moments of power consumption as the features representing relevant moments, especially when users make decisions and share preferences to discriminate between various user consumption moments. This was possible by

analyzing energy consumption signals and occupancy patterns collected through different sensors deployed in buildings with reference to each appliance’s active consumption rate, its maximum operation time, and its standby consumption level. Specifically, observations of time-series consumption are then categorized into the aforementioned classes after undergoing a pre-processing step, which includes data cleaning and resampling.

- A DNN-based anomaly classifier is implemented to automatically identify abnormal consumption classes generated via the micro-moments procedure. Furthermore, its performance has been validated and compared with that of various machine learning classifiers using a K-fold cross-validation.
- An anomaly detection dataset based on real data with its ground-truth labels is generated based on an experimental campaign performed at Qatar University Energy Lab, namely Qatar University dataset (QUD), in order to test and validate the proposed anomaly detection system. In addition, two other existing datasets, namely Dutch Residential Energy Dataset (DRED) [38] and Power Consumption Simulated Dataset (PCSiD) [39], are also considered in this study. Accordingly, their ground-truths are also generated and used to investigate the performance of the proposed solution.
- Finally, by using the proposed anomaly detection scheme, promising performance has been obtained, especially on the real datasets, i.e., QUD and DRED. Various metrics have been used to assess the performance of the proposed technique, such as the accuracy, F1 score, confusion matrix, and under area receiver operating characteristic (AUROC) curves.

The rest of this work is arranged as follows. Section “[Related Works](#)” summarizes related works and outlines their limitations. The proposed approach for the detection of abnormal energy consumption is described in “[Proposed Methodology](#).” In “[Experimental Results](#),” the new designed dataset named QUD is explained together with an overview of other databases from the literature review. Then, the performance of the proposed system is evaluated using different evaluation criteria. In “[Conclusion](#),” conclusions derived from this work are resumed and future directions are drawn.

Related Works

Developing efficient approaches for abnormal energy consumption and energy saving necessitates two main steps. First, creating new challenging datasets that fit user

consumption behavior and collect other data sources that can influence energy usage, such as consumer occupancy patterns, indoor ambient conditions, and outdoor weather information. Second, elaborating powerful algorithms to handle large amounts of data. In this context, this section summarizes the most important works that address these challenges.

Energy Consumption Datasets

Investigating the first step, a set of realistic and simulated datasets and repositories have been created to help researchers test their algorithms. However, most of them focused only on collecting electrical power consumption. Furthermore, each database has its specific application ranging from energy disaggregation, non-intrusive load monitoring (NILM), and appliance recognition to energy saving. Unfortunately, there is no particular dataset dedicated only for power consumption anomaly detection. Simply put, there is a lack of publicly available annotated power consumption datasets to train/learn anomaly detection algorithms, in which power consumption variables are clearly labeled as normal or anomalous. Furthermore, predecessor works mainly focus on using unsupervised anomaly detection techniques that have been implemented in other fields (e.g., fraud detection, cyber-intrusion detection, video surveillance, medical anomaly detection), which do not require annotated datasets. In addition, for the case of anomaly detection in energy consumption, there are other particular anomalies which can not be detected using conventional unsupervised anomaly detection algorithms, such as the consumption while end-users are outside. In this context, the need to define the specific anomalies that should be detected in energy consumption is of paramount importance and further the necessity to produce annotated datasets that can be used to train/learn supervised algorithms is a current hot topic.

In [40], Reference Energy Disaggregation Dataset (REDD) was introduced where power consumption of six households in MA, USA, was monitored. The aggregated electricity consumption was measured at a sampling rate of 15 KHz for a period of 119 days. In [41, 42], two widely used open-access databases were proposed, called the Almanac of Minutely Power Dataset versions 1 and 2 (AMPds1 and AMPds2). They include energy usage footprints for each individual appliance collected at 1-min sampling spans. In both versions, data were assembled from inspecting only one home. The main difference between them is that the AMPds1 monitored energy usage for 1 year, while in AMPds2, data were accumulated for 2 years. Furthermore, aggregated power consumption data were collected at a frequency of 1 Hz as well. In [43], authors launched the rainforest automation

energy (RAE) dataset that was collected through monitoring two different households in Vancouver, Canada. Energy consumption profiles were gathered at a sampling frequency of 1 Hz for the aggregated circuits and device level. Moreover, outdoor and indoor ambient conditions were also picked up, including temperature, wind speed, and humidity. In [38], the DRED, namely Dutch Residential Energy Dataset, is introduced. It captures energy usage samples, occupancy patterns, and environmental data of one household in the Netherlands. Sensor units have been installed to measure aggregated energy consumption and appliance-level electricity usage. In fact, 12 different domestic appliances were sub-metered at sampling intervals of 1 min while 1-Hz sampling rate was used to gather aggregated consumption. In [44], the REFIT electrical load measurements database was developed, which encompassed power usage fingerprints collected from 20 houses at the grouped circuits and individual appliances. The power consumption profiles were obtained from 20 households in the UK at 8-s sampling intervals. Data about the number of occupants in each house were also provided. In [45], the UK domestic appliance-level electricity (UK-DALE) dataset was proposed. It outlined current and voltage patterns of three different domestic buildings at 16-KHz frequency samples and monitored two other dwellings at a frequency of 1 Hz. In addition, consumption profiles of individual devices from five other households were collected at a sampling rate of 6 s for various collection periods varying from 39 to 655 days. And besides, in [39], the PCSiD (<http://em3.i-know.org/datasets/>) was proposed, which was conceived based on data generation of hourly consumption profiles for a period of 2 years at a device level. Device's manufacturer specifications were used to define power consumption patterns of each appliance in watts. Additionally, occupancy patterns were then generated employing some occupancy rules that represented the percentage of the occupancy per day for every room in the household.

As we have mentioned above, the main issue with these datasets is the lack of ground truth references about abnormal and normal consumption. Specifically, there is no dataset in the state-of-the-art that includes such information. In this regard, developing strategies to label observed energy consumption footprints as normal or abnormal is a must in order to detect and report anomalous behaviors.

Detection of Energy Consumption Anomalies

Machine Learning–Based Techniques

In this section, we go beyond a retrospective review of machine learning and its application for detecting abnormal energy consumption. Aiming at building robust machine

algorithms to promote energy saving, a set of works has been solicited in literature. In [46], the authors propose an unsupervised scheme to detect anomalous energy consumption footprints for comprehensive visual analytics of power data. In that regard, electricity consumption anomalies are caught according to either a classification-based technique or a time-weighted forecasting. Moreover, anomaly detection score outputs are visualized based on different visualization tools guiding the user in extracting relevant time points and proposing solutions to correct abnormal behavior. Unfortunately, more in-depth analysis is desired since this approach did not consider external knowledge, such as familiar events affecting the power usage data or ambient and weather conditions. In [27], an identification method that detects abnormal energy usage behavior is developed. This technique pinpoints anomalous activities in two steps: consumption forecasting and abnormal identification. Overall, real power consumption is forecasted from daily domestic actions requiring electricity using combined auto-regressive integrated moving average and neural networks (ARIMA-NN). Then, anomalies are then captured through estimating the difference between forecasted and realistic consumption fingerprints according to a two-sigma regulation for a period of 17 weeks.

Considering the fact that building energy consumption takes about 41% of the total energy consumption in China, Ma and Zhang [47] propose a real-time detection method of abnormal building energy consumption. They connect fractal correlation dimension (FCD) and proper orthogonal decomposition linear stochastic estimation (POD-LSE) to finally generate a higher correctness rate. In their conclusion, Ma et al. indicate that the validity of their method depends on a proper threshold of the FCD deviation. Thus, they suggest integrating other intelligent algorithms. In [48], abnormal events in a school building are appointed by capturing anomalous patterns occurred in electrical consumption. Accordingly, Cui and Wang have explored a hybrid system that merges the polynomial regression and Gaussian model to detect abnormal consumption observations. By reference to this design, anomalous consumption events are spotted, visualized, and submitted to a power management firm, and hence helping their engineers analyzing and comprehending load usage behavior. However, this model shows some limitations, especially for the reason that it can only pinpoint abnormal usage in consumption profiles of the weekend days. Furthermore, when trying to use the Gaussian kernel to track down anomalous patterns in weekly data, a high false positive rate is induced, leading to an overfitting issue.

In addition, with the proliferation of DNN and its deployment in different fields and applications, a set of works has been proposed in the literature to assess its applicability for anomaly detection [49, 50]. In this context,

deep learning-based abnormality detection solutions receive an increasing interest [51]. In [23], Yan proposes a deep anomaly detection to identify gas turbine combustor anomalies based on two principal stages: (i) it uses a DNN for learning characteristic representations extracted from multivariate time-series sensor records; and (ii) it deploys a one-class classification for modeling normal variables in the training feature set and helps in identifying anomalies through capturing the variables that do not fall into the normal class. In [34], a deep learning approach based on an autoencoder architecture is proposed for intrusion detection. It relies on using statistical data analysis with a deep autoencoder-based long short-term memory (LSTM) for extracting optimal, robust, and highly correlated characteristics. In [52], non-technical anomalies and electricity theft detection are performed using a hybrid DNN. Specifically, a deep architecture based on both LSTM and multilayer perceptron (MLP) network is proposed. To learn the characteristics of the abnormalities, raw data are fed into the first module to analyze historic data while the second architecture is responsible for integrating non-sequential variables, including its contracted power and/or geographic coordinates. In [36], a deep learning solution to detect abnormalities in crowded scenes for video surveillance applications is proposed. In this context, a convolutional architecture based on autoencoder and LSTM is used to identify local anomaly events via learning raw image sequences and detected edges.

Statistical Techniques

Abnormal energy consumption based on statistical models extracts statistical features of the normal usage for establishing a norm profile and utilizes statistical experiences for measuring the divergence of the monitoring consumption patterns from the norm value [53].

In [54], a temperature-based approach called the DET-Toa is proposed by Lin and Claridge to provide a better fault detection process. This system distinguishes an anomalous energy consumption behavior through assessing the deviation between collected and simulated energy usage patterns. Next, by defining an empirical threshold using the standard deviation of the residuals, an abnormal consumption or a fault is detected. Unfortunately, this scheme witnesses the weakness of using limited assumed cases, and thereby the authors suggest more field tests to fortify the performance of the process. In [55], by assimilating contextual characteristics and historical sensor data using a sliding window framework, Araya et al. detect anomalous behavior. This framework introduces a classification system of anomalous consumption patterns using sliding windows. This system ascertains anomalous consumption footprints by making use of overlapping sliding windows.

Moreover, for the sake of enhancing the abnormal detection accuracy, the ensemble anomaly detection framework is developed as well, which is a general arrangement. It fuses various classification approaches according to a voting process. The performance of this scheme is assessed on a real dataset delivered by Powersmiths (Brampton), Ontario, Canada. Authors in [56] develop a supervised learning and statistical-based anomalous power consumption detection system, and use a Lambda scheme that is based on both in-memory distributed computing algorithms, Spark and Spark streaming. Moving forward, a real time anomalous identification is achieved by analyzing scalable live patterns in addition to an iterative process, which helps refreshing consumption signatures from realistic databases. Table 1 resumes the characteristics of anomaly detection-related works discussed in this section.

Despite the fact that only a few numbers of works have been proposed in literature to tackle the abnormal energy consumption, most of the existing techniques are hard to implement and have time-consuming architectures. By contrast, in our case, we propose a simple yet effective solution to detect anomalous consumption through analyzing the electricity footprints with a novel micro-moments-based DNN algorithm. Furthermore, the performance of the proposed deep micro-moments system is assessed on both simulated and real-world datasets. Experimental datasets considered in the validation phase, named QUD and DRED,

are collected from two different regions with completely distinct climate conditions; hence, this emphasizes the objectivity of our study. In addition, this helps identifying the advants and limitations of the proposed solution in different scenarios.

Proposed Methodology

In this work, the use of micro-moments as a feature extraction of energy consumption signals is introduced. The latter are acquired from building sensors and utility sub-meters. Precisely, the features are extracted considering occupancy patterns of individuals in order to detect anomalous consumption of the end-users. Moreover, the micro-moment features are employed by Google to model consumers' behavior for marketing applications. To the best of our knowledge, although they have been recently investigated as relevant features that characterize consumer needs and consequently the properties of the marketers required to be a part of, the use of micro-moments for energy applications and particularly for anomalous energy consumption has not yet been explored. Moving to the fieldwork, micro-moments are extracted from various raw signals gathered from multiple sensors. Figure 1 describes a general flowchart of the proposed system. In what follows, the details of the resulting design are presented.

Table 1 An overview of related anomaly detection frameworks

| Work | Feature | Method | Type |
|-------------------------|---|--|-----------------|
| Liu et al. [23] | Multivariate time-series data | One-class DNN | Semi-supervised |
| Chou et al. [27] | Difference between real and predicted consumption | Neural networks-ARIMA | Supervised |
| Ieracitano et al. [34] | Statistical features | Autoencoder base LSTM | Unsupervised |
| Ramchandran et al. [36] | Raw images and detected edges | Convolutional DNN-based autoencoder and LSTM | Unsupervised |
| Janetzko et al. [46] | Power spectrum | Clustering and visual analytics | Unsupervised |
| Ma et al. [47] | Standard deviation of temporal coefficients | FCD-POD-LSE | Unsupervised |
| Cui and Wang [48] | Power consumption time series | Polynomial regression and Gaussian distribution | Semi-supervised |
| Buzau et al. [52] | Historic and non-sequential power data | DNN-based LSTM and MLP | Supervised |
| Lin and Claridge [54] | Power consumption and temperature | deviation between measured and simulated consumption | Unsupervised |
| Araya et al. [55] | Contextual/behavioral features | Detecting incomplete data in sliding window | Unsupervised |
| Liu et al. [56] | Power and temperature | Lambda architecture | Supervised |
| Current paper | Micro-moment consumption and occupancy data | DNN and rule-based algorithm | Supervised |

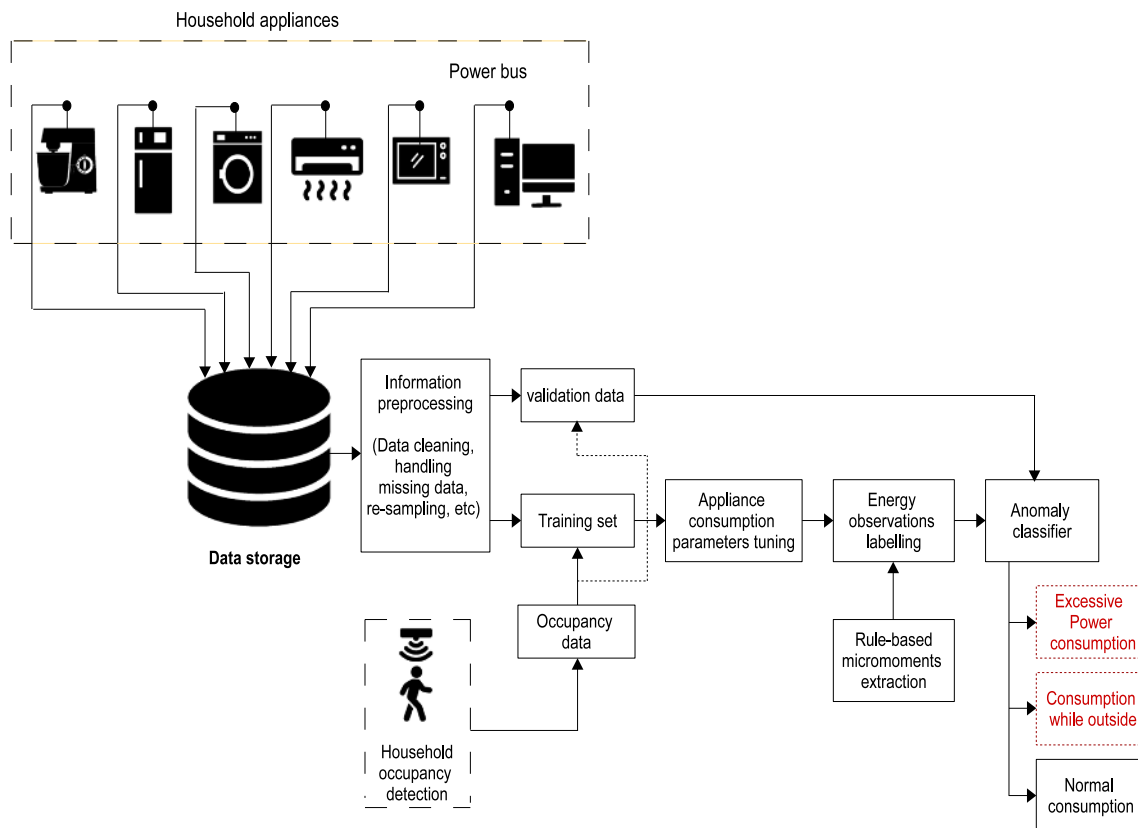


Fig. 1 Block diagram of the proposed system for detecting abnormal energy consumption

Pre-processing

Data collected through different energy and occupancy sensors are firstly cleaned and pre-processed to remove or correct invalid records. Specifically, gathered footprints are raw data or incomplete where missing values are occurring and some interesting attributes are also lost during the collection process. The absence of these values is normally attributed to hardware and/or software failures of measurement devices. Furthermore, other data are noisy, i.e., containing errors or outliers. To this end, making use of a data cleaning process is essential. Consequently, attributed mean values are utilized to fill all missing values in power datasets.

Anomaly Assumption

Defining excessive usage of an electrical device is not a straightforward task, especially with the lack of resources that record appliance power consumption; let us alone set excessive power consumption levels. For that reason, we define “excessive usage” behavior based on two criteria: the appliance operating time and the maximum power the appliance consumes. Fortunately, the study described in [57] provided the maximum operation times of a set of

domestic appliances per day for a period of 4 weeks in an apartment of 3 rooms leased by 2 adults and 2 children. Furthermore, other works have already explored standby power measurements of several home devices in Turkey and Austria [58]. By using the data at hand, Table 2 is introduced, which recapitulates operation times together with maximum power consumption rates at normal and standby modes for a set of domestic appliances.

Furthermore, in order to detect abnormalities related to the presence/absence of the end-users in buildings, the occupancy rules for each appliance need to be set to map the data recordings into their corresponding “consumption while outside the room” micro-moment. Hence, the occupancy flag 0/1 (outside/inside the room) could be set depending on the room of interest and time of the day.

Micro-Moment Feature Extraction

The micro-moment features are extracted from a dataset using a rule-based model. Each database contains multiple information, including dates, time intervals, device ID, power consumption, and occupancy patterns. This work focuses on detecting two main anomalous classes. Firstly, normal consumption contains three subclasses, including

Table 2 Power consumption specifications for different home appliances

| Appliance | Maximum operation time | Maximum power consumption (W) | Maximum standby power consumption (W) |
|-----------------|------------------------|-------------------------------|---------------------------------------|
| Microwave | 1 h | 1200 | 7 |
| Dishwasher | 1 h 45 min | 1800 | 3 |
| Oven | 3 h | 2400 | 6 |
| Washing machine | 1 h | 500 | 6 |
| Tumble dryer | 1 h 45 min | 3000 | 6 |
| Refrigerator | 17 h 30 min | 180 | 0 |
| Light | 8 h | 60 | 0 |
| Television | 12 h 42 min | 65 | 6 |
| Water-heater | 5 h 30 min | 4000 | N/A |
| Air conditioner | 15 h 30 min | 1000 | 4 |
| Laptop | 12 h 42 min | 100 | 20 |
| Desktop | 12 h 42 min | 250 | 12 |

“good usage,” “turn on device,” and “turn off device.” Anomalous consumption encompasses two sub-groups, which are “excessive consumption” and “consumption while outside.” Consequently, a total of five classes is finally considered; they are extracted based on analyzing the occupancy profile (O) and power consumption (P) of each device in reference to its active consumption rate (A), maximum operation time (T_M), and standby consumption level (S). The rule-based model used to extract micro-moment features (MF) over time is summarized in Algorithm 1:

Algorithm 1 The rule-based model used to extract micro-moment features.

Result: MF: the vector of micro-moment features
 Read P, S, O, A, T_M ;
 Initialisation: $MF = \emptyset$
while $t \leq N$ **do**
 if $P(t) \geq \min(A)$ **and** $P(t) \leq 0.95 \times \max(A)$ **then**
 MF(t) = 0 (Good usage) ;
 else if $P(t) \geq \min(A)$ **and** $P(t-1) \leq \max(S)$ **then**
 MF(t) = 1 (Turn on device) ;
 else if $P(t) \leq \max(S)$ **and** $P(t-1) \geq \min(A)$ **then**
 MF(t) = 2 (Turn off device) ;
 else if $P(t) \geq 0.95 \times \max(A)$ **or** $T_M(t) \geq \max(T)$ **then**
 MF(t) = 3 (Excessive consumption) ;
 else
 if $O(t) = 0$ **and** $P(t) \geq 0.95 \times \max(S)$ **then**
 MF(t) = 4 (consumption while outside) ;
 end
end

Table 3 recapitulates the list of micro-moments feature classes extracted using the proposed rule-based algorithm with their corresponding label descriptions.

Handling Imbalanced Data

Most of the energy consumption databases for anomaly detection collected through experimental campaigns express the problem of the enormous imbalance of real anomaly patterns, since anomalous footprints usually pertain to the minority classes. To this end, we have introduced an interesting feature that should be considered to fix this issue, which is the normalized power consumption. It represents the difference of power consumption rates of each current time sample and the previous one. The consumption, in theory, varies through the time intervals; therefore, collecting normalized consumption can provide us information how fast the consumption reacts to the time evolution. If P stands for the power consumption vector of N observations gleaned for a time interval τ , P_N represents the normalized power consumption data. For every power consumption observation at time t , $P_N(t)$ is calculated as follows:

$$P_N(t) = \frac{P(t) - \text{mean}(P)}{\max(P) - \min(P)} \quad (1)$$

In addition, power consumption quantification $P_Q(t)$ is another parameter that could be investigated when managing an imbalanced energy consumption database. This information is of paramount importance to capture the “Turn on” and “Turn off” categories, and thereby it mainly

Table 3 Micro-moments assumption and labeling

| Micro-moment | Label | Description |
|---------------------------|-------|--|
| Good usage | 0 | Non-excessive usage |
| Turn on | 1 | Switching on a device |
| Turn off | 2 | Switching off a device |
| Excessive consumption | 3 | Consumption > 95% of a device's maximum active power consumption level |
| Consumption while outside | 4 | Device consumption without presence of the end-user |

helps improving the efficiency of the deep micro-moment classifier. It can be estimated as follows:

$$P_Q(t) = \begin{cases} 1 & \text{if } P(t) - P(t+1) \geq 0 \\ -1 & \text{if } P(t) - P(t+1) < 0 \end{cases} \quad (2)$$

Classification

Generally speaking, deep learning is a subdiscipline of machine learning built on an idea of learning various grades of representations through developing a hierarchy of characteristics. With this in mind, the DNN arrangement is based on expanding the conventional artificial neural networks (ANN) that encompasses only three layers (i.e., an input layer, a hidden layer, and an output layer) through inserting multiple hidden layers to the network structure between the input and output layers (i.e., DNN includes more than three layers) [59, 60]. This is done for the purpose of making a good ability to deal with complex and nonlinear classification problems. In consequence, DNN attracted scientists' attention during the last years on the ground that it can provide better performance than many other existing solutions, particularly for regression, classification, modeling, forecasting, and prediction objectives.

Deep Learning Classifier

Under this framework, since non-linear separable data are handled, deep learning is highly recommended for such problem. Furthermore, the performance of a deep learning algorithm is usually enhanced by increasing the amount of data used for training. That is why the DNN architecture is selected among other solutions to deal with abnormal power consumption detection.

DNN represents a subgroup of machine learning approaches that has arisen as a flexible classifier and approximator. It usually establishes a trustworthy map between inputs and outputs of different problems with high complexity levels. Usually, a DNN is consisting of various layers l_i , which takes as input x_{l_i} and provides an output

$$y_{l_i} = \sigma(W_{l_i} \cdot x_{l_i} + b_{l_i}) \quad (3)$$

where σ is the nonlinear activation function, W_{l_i} and b_{l_i} are the weights and biases, respectively, to make the connection

between input and output neurons. These connections can be very simple to fit a dense connection between each of the input and output neurons. Nevertheless, dense liaisons may miss preserving regional correlations of input data, generally including valuable information; e.g., this is the context of signal and image classification. Moreover, the time computation can be increased in order to learn dense connectivities between neurons and all connections may not be quite meaningful. In contrast, a convolutional liaison of a dense connection facilitates alleviating these problems. The convolution function (\otimes) can be expressed as:

$$W[m, n] \otimes x[m, n] = \sum_{i=-h}^{i=h} \sum_{j=-l}^{j=l} W[i, j]x[m-i, n-j] \quad (4)$$

The proposed DNN architecture entails 12 layers including 10 hidden layers. The input layer has 10 neurons while the output layer includes 5 neurons. For the hidden layers, each one encompasses 20 neurons. Furthermore, the proposed architecture is implemented in Python, using Keras and TensorFlow deep learning libraries. A typical representation of the proposed architecture based on deep micro-moment classification is outlined in Fig. 2.

Training/Learning Procedure

After applying the rule-based algorithm to annotate the data, we obtained a dataset that contains both normal and anomalous data. Specifically, five feature classes are obtained as discussed in “[Micro-Moment Feature Extraction](#).” Since we are interested in classifying normal vs. anomalous patterns, we use the training data that contain both kinds of observations (i.e., the five micro-moment classes) to train the DNN model. In this direction, a K -fold cross-validation arrangement is deployed to train and test the deep micro-moments solution. This procedure is a statistical analysis method that implies splitting input data and their labels into K subgroups, then applying a supervised training on $K-1$ subgroups and using the remaining subgroup to assess the model's performance in terms of the accuracy and F1 score. We repeat this process K times, in which every subgroup is utilized $K-1$ times to train the DNN and once to test it.

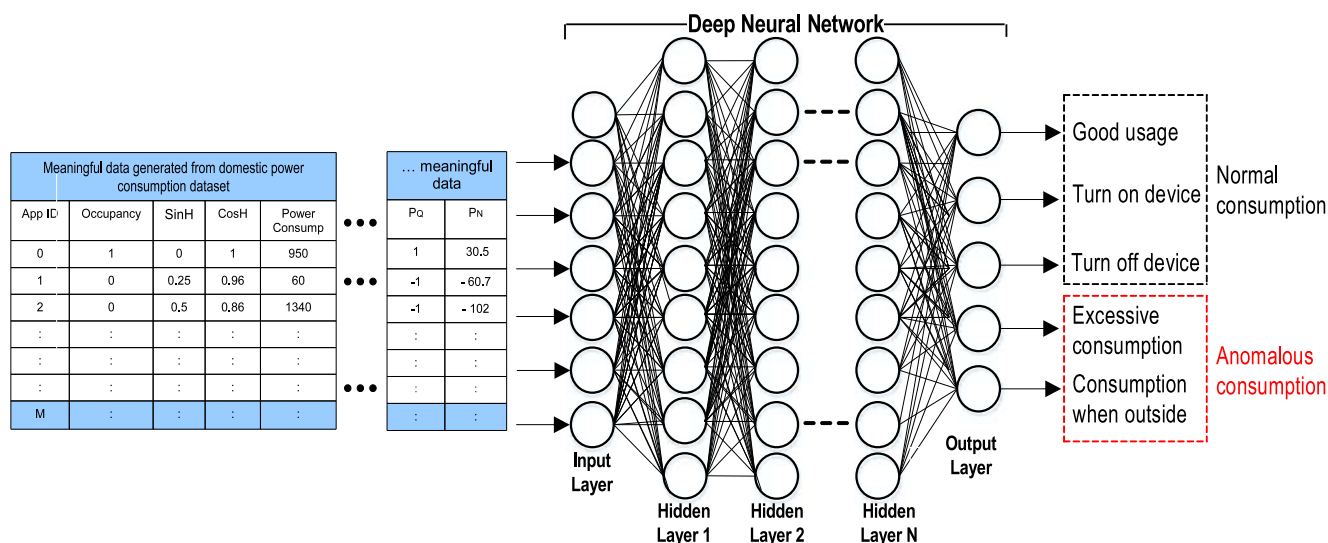


Fig. 2 Typical representation of the DNN architecture used for micro-moment classification

Therefore, for the training phase, input data vectors including the timestamps, appliance ID, P , O , P_N , and P_0 along with their corresponding micro-moments classes MF are fed to the DNN. The latter encompasses multiple hidden layers in order to learn the relation behavior of normal and abnormal observations. At the output end, a rectified linear unit (ReLU) activation function is used.

Other Classifiers

Furthermore, deeply in this framework, several machine learning algorithms are also examined for the classification purpose, including logistic regression (LR), linear discriminant analysis (LDA), naive-Bayes (NB), support vector machine (SVM), random forest (RF), K-nearest neighbors (KNN), decision trees, ensemble classifier, and multi-layer perceptron (MLP). These classification models have been operated using various parameter configurations with the aim of classifying the micro-moment classes.

The main aim here is to find out the best algorithm to identify anomalous consumption. Multiple kinds of data, including timestamps, power consumption records, and occupancy patterns, are utilized to train supervised learning classifiers to forecast the micro-moments labels. Then, relying on these labels, anomalous consumption is identified; and hence, this can help on later notifying the end-users to reduce their wasted energy.

Experimental Results

Aiming to evaluate the performance of our system dedicated for the detection of anomalous consumption based on the deep micro-moment feature extraction, we experiment

with various machine learning classifiers, using various parameter settings for each classifier. This section details the characteristics of the different datasets used in the validation process and emphasizes the performance of the proposed system in terms of various measurement metrics.

QUD Description

To compensate the undersupply of appliance-level datasets, a real-time micro-moment laboratory has been built to collect accurate energy consumption data. In this context, QUD is launched (<http://em3.i-know.org/datasets/>), which is an ensemble of reading from various installed appliances (e.g., light lamp, air conditioner, desktop, heating system) as well as contextual information such as temperature, humidity, ambient light intensity, and room occupancy. To the best of the authors' knowledge, QUD is the first dataset in the Middle East, where ordinarily 240-V voltage is used with varying recording frequency ranging from 3 s to 3 h. As a real data source, the database is employed in the classification and generation of recommendations. Figure 3 illustrates a demonstration setup including all the implemented sensing units.

The consumption profiles of QUD are recorded for a period of 12 weeks until now; however, the collection process is still ongoing in order to cover a period of 1 year. In this direction, time-series data representing power consumption footprints for two appliances are registered along with the corresponding cubicle occupancy, indoor temperature, indoor humidity, and luminosity.

An evaluation of the respective accuracy of each module was conducted with reference to a measurement tool. Test beds for temperature, humidity, electric current, and occupancy are elaborated. Figure 4 illustrates developed

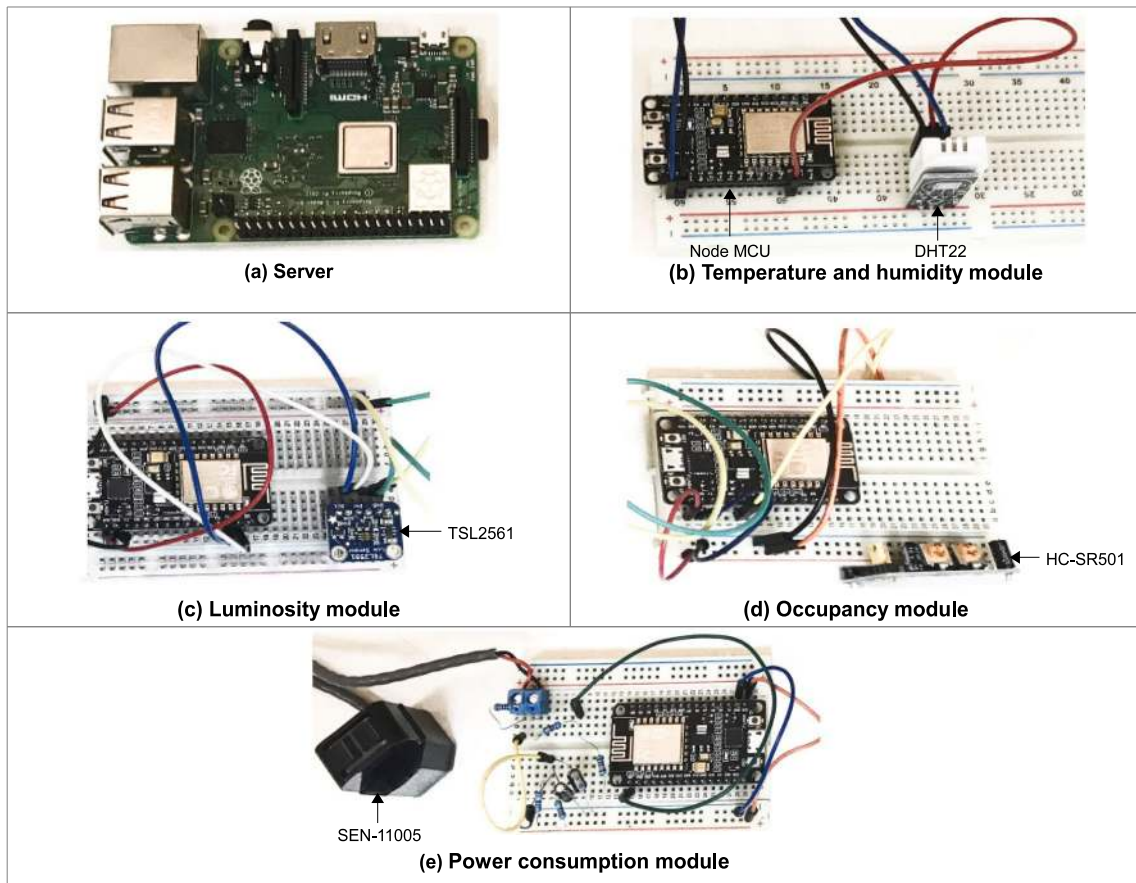


Fig. 3 The sensor configuration implemented to collect different kinds of data on the QUD

test beds to collect the electrical current of the power consumption unit. In general, the temperature sensor achieves the best accuracy with 98.3% compared with 73.3%, 85.9%, and 93.5% for the occupancy, electrical current, and humidity, respectively.

Table 4 describes the essential attributes of QUD and other state-of-the-art datasets described in “[Related Works.](#)” This brief comparison outlines the characteristics of each database in terms of the location and period of data collection, number of observed houses, number of deployed

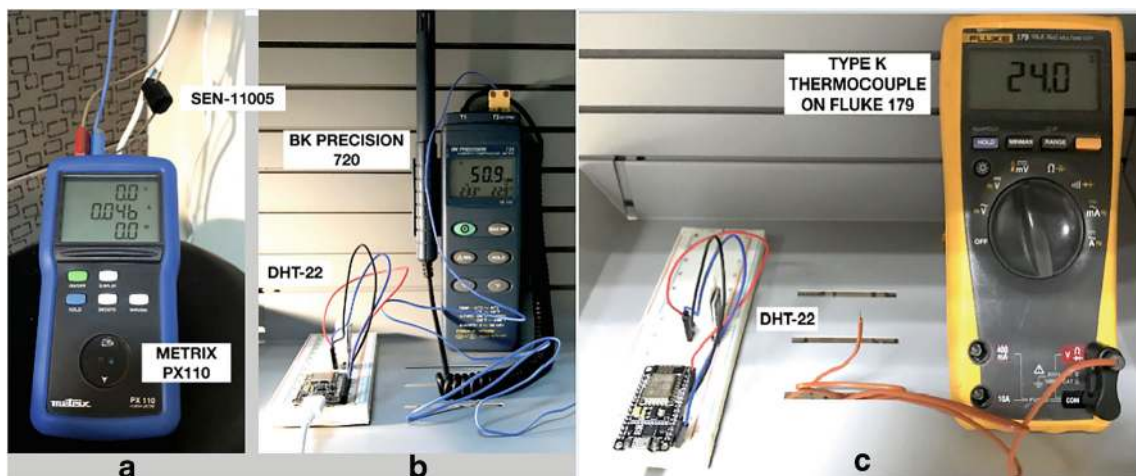


Fig. 4 QUD validation test beds for **a** power consumption, **b** humidity, and **c** temperature

Table 4 Features comparison of the QUD dataset with other of existing real power consumption datasets

| Acronym | Country | Period | NO. of homes | No. of sub-meters | Features | Sampling rate |
|--------------|------------|----------|--------------|-------------------|-------------------|------------------|
| REDD [40] | USA | 119 days | 6 | 24 | I, V, P | 3 s |
| AMPds1 [41] | Canada | 1 year | 1 | 21 | I, V, P, Q, S, pf | 1 min |
| AMPds2 [42] | Canada | 730 days | 1 | 21 | I, V, P, S, F, pf | 1 min |
| UK-DALE [45] | UK | 655 days | 4 | 5 (H4), 53 (H1) | P, agg P | 6 s/6 KHz (agg) |
| DRED [38] | Netherland | 6 months | 1 | 13, agg | P, T, H, Ws, Pr | 1 min/1 Hz (agg) |
| REFIT [44] | UK | 213 days | 20 | 9, agg | P, pf, T, L, \$ | 8 s |
| RAE [43] | Canada | 72 days | 1 | 24 | V, P, Q, S, f, E | 1 Hz |
| QUD | Qatar | 1 year | 3 | 4 | P, H, T, O | 3 s–30 min |

sub-meters, collected features, sampling rates, and release dates. In addition, collected features in each database are also specified in terms of the current (I), voltage (V), active power (P), reactive power (Q), apparent power (S), energy (E), frequency (f), power factor (pf), energy cost (\$), weather (Wt), temperature (T), humidity (H), occupancy (O), and light level (L).

Moreover, it is clearly shown that most of the mentioned databases are collected in totally different regions in terms of climate conditions. QUD is gathered in an arid zone where high temperatures are usually recorded, while DRED belongs to a region with an oceanic climate that is specified with moderate temperatures in the warm months, and low temperatures in the cold periods. The other ones cover humid regions (REDD) and the marine west coast atmosphere (UK-DALE, AMPds1, AMPds2, and REFIT). Additionally, the QUD has the particularity of supplying occupancy patterns for each time sample in each cubicle. This is very helpful when detecting consumption anomalies related to the presence/absence of end-users; therefore, it can be very convenient for applications dealing with power consumption abnormality detection.

Performance Measure

Alongside two kinds of datasets used in the evaluation process are the realistic datasets including QUD and DRED datasets, which are produced through measurement campaigns and the PCSiD repository already described in [39], which is a simulated dataset. In the empirical evaluation, the accuracy of each classifier is investigated using a 10-fold cross-validation arrangement. The latter is a statistical analysis approach used to avoid the overfitting of the classifiers. The accuracy metric represents the percentage of the correctly classified micro-moments in the test bed. Often, however, due to the fact that the accuracy is not considered as a robust metric when evaluating imbalanced data (i.e., typically the case of anomaly detection in energy consumption), the F1 score,

which is a more reliable metric, is assessed as well.

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (5)$$

where TP , TN , FP , and FN depict the number of true positives, true negatives, false positives, and false negatives, respectively.

$$F1\ score = 2 \times \frac{precision \times recall}{precision + recall} \quad (6)$$

where $precision = \frac{TP}{TP+FP}$ and $recall = \frac{TP}{TP+FN}$.

Micro-Moments Visualization

To help consumers in promoting their energy consumption behavior, it is useful to use smart visualization tools, allowing then individuals understanding their power consumption patterns and improving their usage behavior. In this respect, this section describes a new way to visualize electrical consumption profiles through plotting the extracted micro-moments using the proposed algorithm. In this line, the micro-moment samples are drawn using a scatter plot through the time index. Table 5 summarizes the number of micro-moment features in each database.

Figure 5 epitomizes the minutely power consumption of an LCD TV in watts with different visualization plots. Figure 5 (top) outlines the power consumption over the time index. Figure 5 (middle and down) portrays scatter plots of the micro-moment features with sampling intervals of 3 min and 15 min, respectively. As a matter of fact, they summarize in a good way how micro-moments are distributed through the time axis. Correspondingly, we can easily localize abnormal consumption, especially for the micro-moment classes 3 and 4 that represent “excessive consumption” and “consumption while outside”, respectively.

Table 5 Distribution of the micro-moment patterns through the classes in each dataset

| Micro-moment description | Label | No. of micro-moment patterns | | |
|---------------------------|-------|------------------------------|---------|---------|
| | | QUD | DRED | PCSiD |
| Good usage | 0 | 12,114 | 45,482 | 59,425 |
| Turn on device | 1 | 1568 | 3315 | 7780 |
| Turn off device | 2 | 1569 | 3316 | 7779 |
| Excessive consumption | 3 | 3954 | 35,044 | 6343 |
| Consumption while outside | 4 | 27,725 | 79,196 | 23,793 |
| Total | | 46,930 | 166,353 | 105,120 |

Classification Results

For each dataset considered in this investigation, the performance of our anomaly detection system is assessed. To evaluate the classification accuracy and F1 score, collected data along with their related labels are split into 10 independent, complementary subsets, in which 9 subsets are reserved for training (i.e., deployed for training the classifier and defining all parameters) and the remaining subset is reserved for testing (i.e., utilized for validating the results). The process of splitting the data into 10 subset is repeated randomly several times and hence obtained accuracy and F1 score outputs from all partitions are averaged. Table 6 illustrates the comparison of the accuracy results of various machine learning classifiers. Proceeding with the QUD and DRED datasets, DNN outperforms the other classifiers, where an accuracy of 99.58% with an F1 score of 97.85% is obtained under the QUD and 99.29% accuracy with 99.6% F1 score is achieved under DRED. Additionally, considering the PCSiD dataset, DNN offers the best results since it can achieve an accuracy of 93.91% and an F1 score of 89.1. The difference between the results obtained from simulated and real databases is justified by the fact that simulated data are generated randomly, which reduces the realism of the data and thereby they can not match perfectly the electrical consumption performed in real environment. In consequence, this makes the classification process difficult.

Effect of the Number of Hidden Layers

In order to find the optimized configuration of the deep micro-moments solution, the effect of the number of hidden layers used in DNN on the accuracy and F1 score is assessed. Figure 6 a and b illustrate the accuracy and F1 score results achieved when the number of hidden layers is varied from 3 to 50 for the QUD, DRED, and PCSiD datasets. It is clearly seen that the proposed deep classifier attains the best results with 10 hidden layers, and thereafter by increasing the number of hidden layers the results have remained almost the same. In contrast, the computation

time has been increased while the number of hidden layers has risen, as indicated in Fig. 6c. Consequently, selecting a number of hidden layers equals to 10 is the optimal configuration that guarantees the best compromise between the classification performance and computation time.

Permutation Test

In machine learning, it is of utmost importance to assess the proficiency of a classifier using permutation-based p value test. Specifically, the latter estimates to which extent the ascertained accuracy could be achieved by chance. A p value stands for the fraction of random datasets under a certain null hypothesis, in which the classifier acts more effectively than in the original set.

To test the performance of the deep micro-moments classifier, we conducted two permutation tests that rely on 999 permutations plus the original accuracy, as it is discussed in [61]. Test 1 is mainly related to randomly exchange the label and classify the data with a 10 cross-validation procedure, while test 2 is principally dealing with the permutation of data columns per class and hence calculating the accuracy of each new classification [61]. The p value is then measured to be the summation of all the cases when the accuracies of the permutation distributions are equal or higher than the accuracy of the original case divided by the number of permutations.

Table 7 depicts the average errors and p values of permutation test 1 and test 2 when the deep micro-moments classifier is applied on the QUD, DRED, and PCSiD datasets. It is worth noting that the deep micro-moments classifier is efficient when the null assumption is that data and labels are independent (i.e., test 1) for the three datasets. On the other side, for test 2, the deep micro-moments classifier is efficient for both QUD and DRED datasets but not for the PCSiD dataset. Consequently, for both QUD and DRED datasets, the dependence between the columns inside the micro-moments categories is important for achieving excellent classification results, and the deep classifier has been efficient in exploiting that information.

Fig. 5 Time series and scatter plots of power consumption extracted from the DRED: (top) power consumption through time series; (middle) scatter plot of the micro-moments for sampling intervals of 3 min; (down) scatter plot of the micro-moments for sampling intervals of 15 min

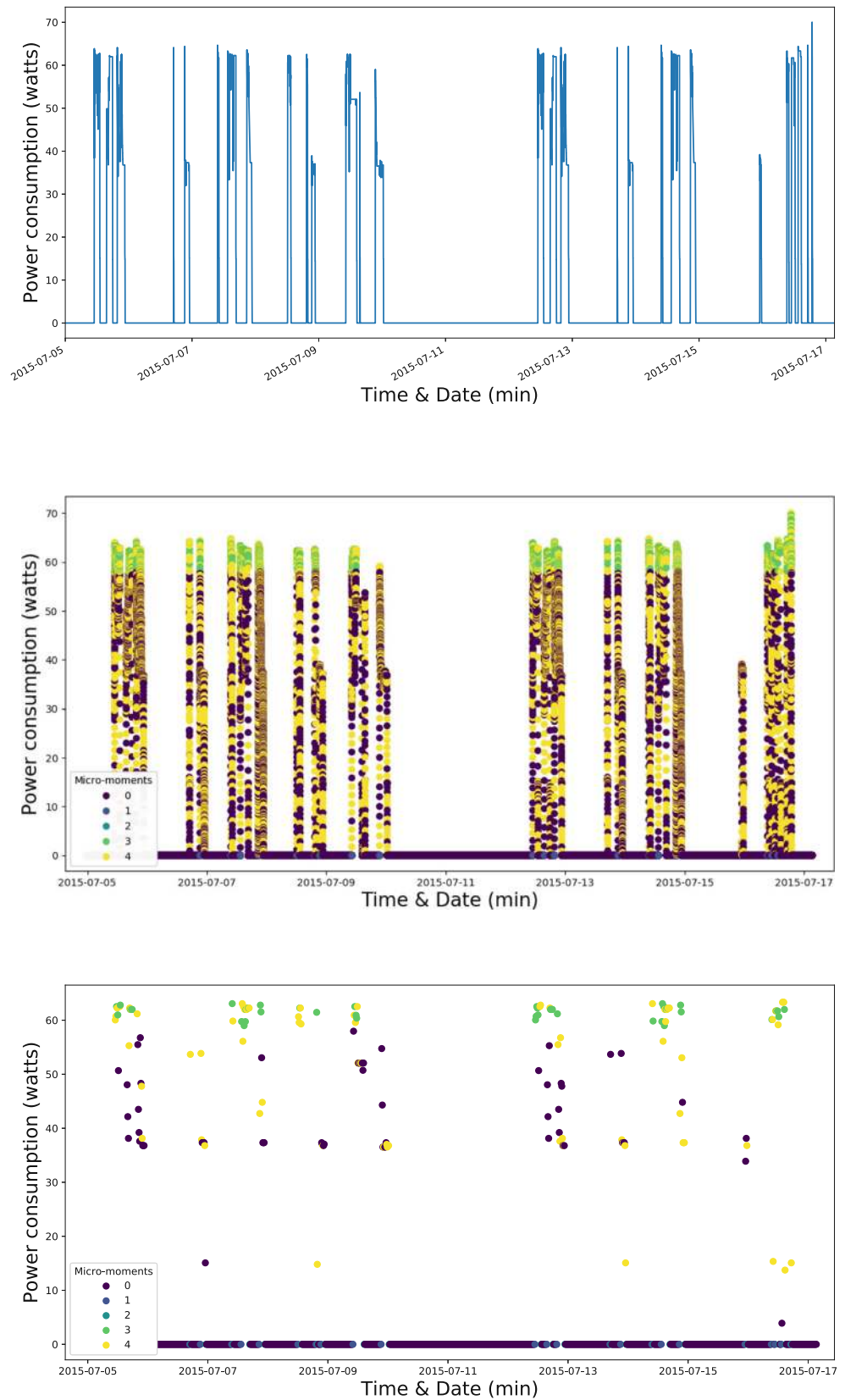


Table 6 Performance comparison of different machine learning algorithms

| ML algorithms | Classifier parameters | QUD | | DRED | | PCSiD | |
|---------------|----------------------------------|----------|----------|----------|----------|----------|----------|
| | | Accuracy | F1 score | Accuracy | F1 score | Accuracy | F1 score |
| LR | One-vs-rest (OVR) | 92.95 | 57.6 | 95.56 | 69.2 | 75.24 | 37.6 |
| LDA | N/A | 84.87 | 36.99 | 89.85 | 53.89 | 75.420 | 4.3 |
| SVM | Linear kernel | 94.98 | 76.23 | 91.10 | 64.66 | 82.29 | 65.63 |
| SVM | Quadratic kernel | 95.08 | 77.08 | 95.92 | 74.07 | 84.08 | 73.85 |
| SVM | Gaussian kernel | 96.18 | 76.75 | 97.09 | 75.03 | 86.65 | 75.16 |
| NB | Gaussian naive Bayes | 95.17 | 70.82 | 96.45 | 74.6 | 50.61 | 39.1 |
| KNN | K=1, Euclidean dist | 95.93 | 72.51 | 93.62 | 74.50 | 83.15 | 71.37 |
| KNN | K=10, Weighted Euclidean dist | 96.28 | 77.90 | 96.91 | 76.99 | 83.69 | 71.38 |
| KNN | K=10, Cosine dist | 96.32 | 78.08 | 96.53 | 70.79 | 82.15 | 65.53 |
| RF | Gaussian classifier | 95.89 | 76.08 | 96.76 | 76.8 | 86.67 | 76.3 |
| MLP | 1 hidden layer | 95.80 | 73.95 | 95.83 | 74.78 | 83.3 | 72.11 |
| Trees | Fine, 100 splits | 96.06 | 74.73 | 96.96 | 75.22 | 86.42 | 71.91 |
| Trees | Medium, 20 splits | 96.17 | 75.69 | 96.97 | 75.03 | 84.32 | 54.43 |
| Trees | Coarse, 4 splits | 95.40 | 70.23 | 96.96 | 74.86 | 82.83 | 51.47 |
| Ensemble | Bagged, 30 learners, 42 k splits | 96.10 | 74.92 | 96.88 | 76.66 | 87.85 | 77.24 |
| DNN | 10 hidden layers | 99.58 | 97.85 | 99.29 | 99.60 | 93.91 | 89.1 |

In contrast, for the PCSiD dataset, it has some issues to effectively utilizing the dependence between the columns inside the micro-moment groups and this has been shown in Table 7, in which a p value = 0.27 has been achieved. Furthermore, this has also been observed in Table 6, where an accuracy and F1 score of 93.91% and 89.1% have been attained, respectively, which are lower than those obtained

for QUD and DRED datasets. This can be justified by the fact that the simulated data in PCSiD does not reflect the real dependency between columns in the micro-moment classes. In addition, PCSiD dataset includes more columns than QUD and DRED datasets, and this may hinder the deep classifier in better exploiting the dependency between those columns.

Fig. 6 Effect of the number of DNN hidden layers on the performance of the deep micro-moments classifier in terms of **a** accuracy, **b** F1 score, and **c** computation time

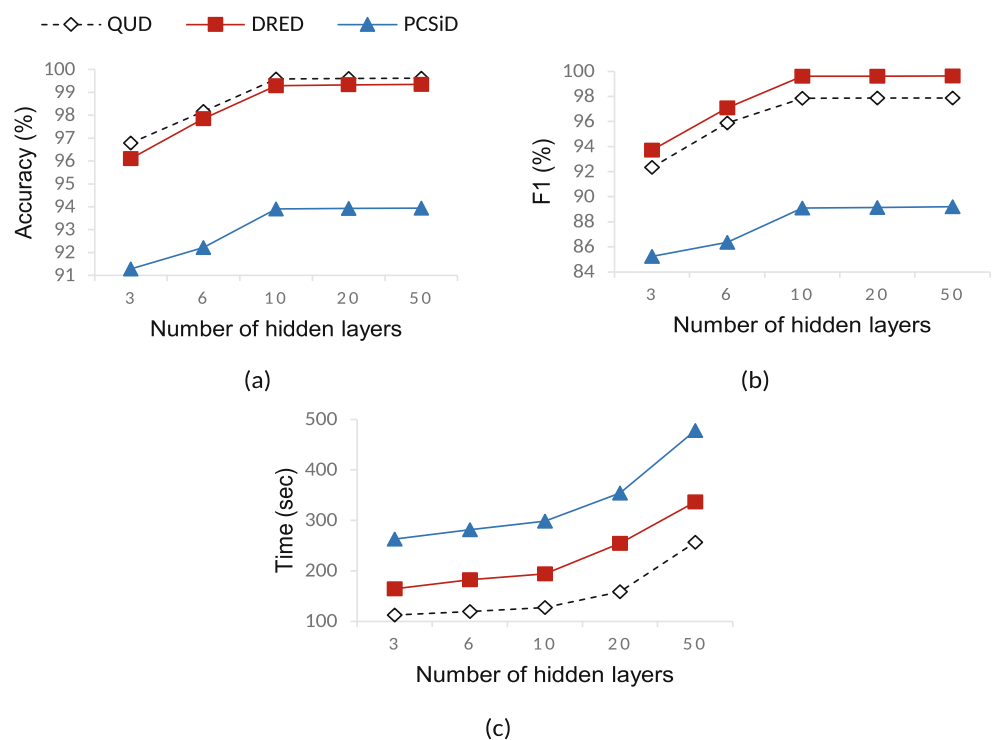


Table 7 Average errors and p values of permutation tests when using the deep micro-moments classifier on QUD, DRED, and PCSiD datasets

| Dataset | Original | Test 1 | | Test 2 | |
|---------|----------|----------------------------|-----------|----------------------------|-----------|
| | Error | Error (standard deviation) | p value | Error (standard deviation) | p value |
| QUD | 0.00 | 0.04 (0.03) | 0.001 | 0.02 (0.02) | 0.001 |
| DRED | 0.00 | 0.03 (0.02) | 0.001 | 0.01 (0.01) | 0.001 |
| PCSiD | 0.00 | 0.07 (0.07) | 0.001 | 0.09 (0.07) | 0.27 |

Effect of Power Normalization/Quantification

In this section, the effect of adding normalized and quantified power consumption data is illustrated in Table 8 using DNN when QUD dataset is considered. Before adding power normalization/quantification patterns our algorithm suffered from a naive behavior, it has some trouble in detecting class 1 and class 2 even if the accuracy was very acceptable. However, generating and inserting new consumption profiles, such as power normalization and power quantification, have led to better accuracy and much higher F1 score results. This confirms how trustable our system is and how robust it is in detecting each micro-moment class. In addition, not only are the anomalous classes well detected, but also the others.

AUROC Curves

Another fascinating benchmark used for checking the performance of multiclass classification architectures is the AUROC curves. It indicates to what extent a classifier can distinguish between the different classes. Figure 7 reports the AUROC curves of the proposed deep micro-moments classification for QUD, DRED, and PCSiD datasets. Figure 7 (top and bottom) clearly shows that the proposed deep model has an AUROC close to 1 for the case of QUD and DRED, denoting that it has a good aptitude of separability. However, for the case of PCSiD database, Fig. 7 (down) accounts that the AUROC has slightly dropped for all the classes involved that the deep micro-moments model encountered some issues. Anyhow,

the proposed deep micro-moments architecture still exhibits satisfactory performance. Furthermore, the latter is slightly dropped due to the fact that the PCSiD repository is generated with a random manner that cannot fit perfectly human consumption behavior. Accordingly, in this case, the proposed solution encounters some difficulties in correctly extracting the micro-moments; and thus, the classification performance is affected as well.

Confusion Matrix

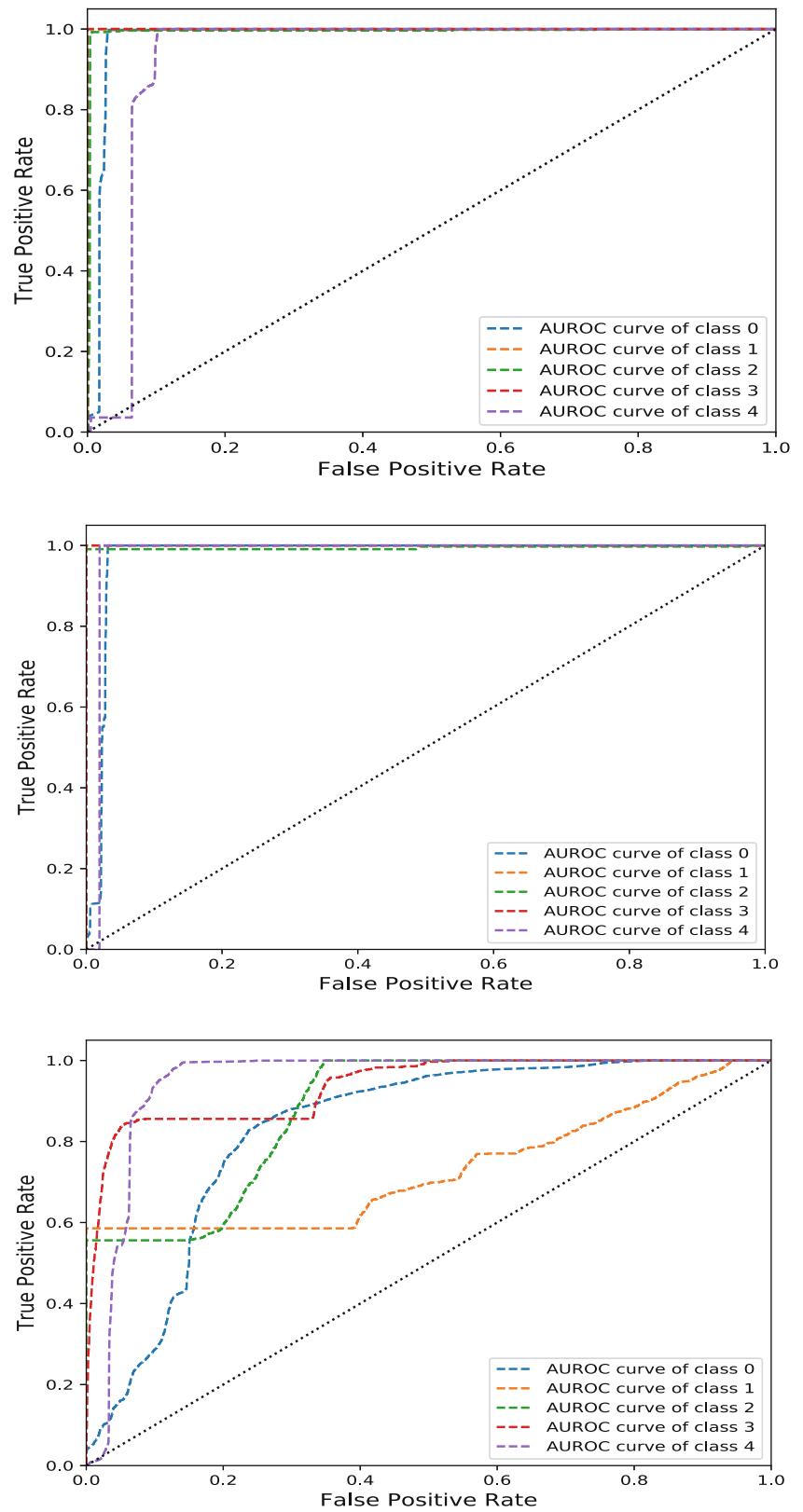
The confusion matrix is judged as an excellent yet simple objective measure that can be useful when managing a classification framework. This measurement provides a valuable summary of how well the classifier is acting. In so doing, it is of utmost importance to trace it when evaluating any classifier. Thus, this section presents a brief performance investigation of the proposed DNN classifier based on the confusion matrix evaluation.

Figure 8 illustrates the confusion matrices extracted using the proposed approach from QUD, DRED, and PCSiD datasets. From Fig. 8a and b, it is clear that the proposed scheme can detect each class with a high accuracy rate and even the anomalous classes are perfectly detected. For example, class 3 is detected with 100% for both QUD and DRED while class 4 is detected with 99.79% and 100% for QUD and DRED, respectively. Unfortunately, Fig. 8c reveals that the performance drops when the PCSiD dataset is considered; this is occurring since the power consumption and occupancy patterns are generated randomly, and hence the classifier has some difficulties to cluster the classes

Table 8 Effect of adding power normalization/quantification on the accuracy and F1 score results

| Micro-moment class | Without power normalization/quantification | | With power normalization/quantification | |
|------------------------------------|--|----------|---|----------|
| | Accuracy | F1 score | Accuracy | F1 score |
| | Class 0: Good usage | 98.45 | 98.86 | 99.45 |
| Class 1: Turn on device | 88 | 0 | 100 | 100 |
| Class 2: Turn off device | 98.45 | 51.88 | 99.65 | 94.34 |
| Class 3: Excessive consumption | 98.38 | 99.38 | 99.35 | 96.34 |
| Class 4: Consumption while outside | 100 | 100 | 99.45 | 99.53 |
| Mean | 95.38 | 70.03 | 99.58 | 97.85 |

Fig. 7 AUROC curves of the different classes obtained using DNN algorithm for **a** QUD, **b** DRED, and **c** PCSiD datasets



correctly. Nevertheless, the performance is quite acceptable given that an accuracy of 93.91% is reached with this database.

In addition, it is worthy to mention that the QUD dataset is very reasonable and the performance obtained using the proposed deep micro-moments anomaly detection under the

Fig. 8 Confusion matrix obtained using the **a** QUD, **b** DRED, and **c** PCSiD datasets

| Micro-moments | # | 0 | 1 | 2 | 3 | 4 |
|-----------------------------|---|---------------|-------------|---------------|-------------|----------------|
| Good Use / No Change | 0 | 1117 97.9% | 1 .08% | 21 1.8% | 2 0.17% | |
| Turn on | 1 | | 171 100% | | | |
| Turn off | 2 | | | 160 98.77% | | 2 1.23% |
| Excessive power consumption | 3 | | | | 416 100% | |
| Consumption when outside | 4 | | | | 6 0.21% | 2797 99.79% |

(a)

| Micro-moments | # | 0 | 1 | 2 | 3 | 4 |
|-----------------------------|---|----------------|-------------|---------------|--------------|--------------|
| Good Use / No Change | 0 | 4538 99.34% | | | 30 0.66% | |
| Turn on | 1 | | 319 100% | | | |
| Turn off | 2 | 3 0.89% | | 335 99.11% | | |
| Excessive power consumption | 3 | | | | 3486 100% | |
| Consumption when outside | 4 | | | | | 7924 100% |

(b)

| Micro-moments | # | 0 | 1 | 2 | 3 | 4 |
|-----------------------------|---|---------------|--------------|--------------|--------------|----------------|
| Good Use / No Change | 0 | 5732 97.1% | 45 0.76% | 65 1.1% | 59 1% | |
| Turn on | 1 | 132 15.9% | 596 71.8% | | 48 5.78% | 54 6.51% |
| Turn off | 2 | 171 22.8% | | 579 77.2% | | |
| Excessive power consumption | 3 | | 10 1.53% | | 636 97.7% | 5 0.77% |
| Consumption when outside | 4 | | 21 0.88% | | 30 1.26% | 2329 97.86% |

(c)

QUD is almost similar to that of the DRED dataset, which is another real repository collected in the Netherlands. Both datasets recorded occupancy patterns in addition to energy consumption footprints. Moreover, QUD has been collected in an academic building (university campus) while DRED has been produced in households. This further proves that the deep micro-moments anomaly detection can be applied in different kinds of buildings that include electrical appliances.

Electricity Saving Rate

The final objective of our micro-moment-based anomaly detection scheme is to reduce wasted electricity and increase energy saving in buildings. In this respect, the amount of saved electricity that can be achieved using our approach depends mainly on the abnormal behavior of end-users, i.e., if they have massive bad consumption habits, a large number of abnormalities will be detected and therefore the

amount of saved energy will be high, and otherwise if they have insignificant bad consumption practices. Moreover, it has been reported in [62] that providing the end-users with indirect feedback on their electricity consumption (i.e., historical consumption data) improves the electricity saving by 12%. On the other hand, the electricity saving level could reach 20% if the end-users are provided with direct appliance-consumption feedback. Consequently, in our case, if we further offer the end-users the anomaly detection feedback, this will increase significantly the saving rate to more than 20%.

To that end, using data from Table 5, the percentage of the anomalous patterns detected in each dataset (excessive consumption and consumption while outside) to the whole number of consumption footprints can be estimated. Explicitly, it can be stated that the anomalous rates represent 67.5%, 68.67%, and 28.66% under QUD, DRED, and PCSiD datasets, respectively. Thus, it is clear that QUD and DRED have massive anomalous behaviors and hence the amount of saved electricity will be very high if the proposed scheme will be adopted. While in PCSiD, the anomalous rate is significantly lower, this will result in a lower energy saving rate. All in all, to measure accurately the amount of saved electricity in real scenarios, a real-time energy saving system based on our anomaly detection system should be implemented. This helps further in transforming end-users' habits via the adoption of a recommender system to trigger end-users with tailored recommendations. Moreover, this will be part of our future work, since the proposed anomaly detection method has shown promising results in identifying abnormal consumption.

Limitation

Finally, it can be noted that the limitation of the proposed work is mainly related to the supervised learning as this latter requires always annotated datasets to train the classifiers. Moreover, this is not unique for our solution, but for all supervised learning models. However, the simplicity and efficiency of the proposed rule-based algorithm deployed to label power consumption observations and design annotated datasets can compensate this disadvantage.

Conclusion

In this paper, a simple yet efficient method for detecting abnormalities in building energy consumption has been introduced. Our technique depended on exploiting micro-moments as features to model energy consumption behavior. In this regard, power consumption signals and occupancy patterns were then analyzed to design the micro-moment classes and identify consumption anomalies. Then,

to carry out the micro-moment classification, we adopted a DNN architecture and conducted a profound comparison process with several machine learning algorithms. Furthermore, a measurement campaign was managed at Qatar University Energy Lab to glean real energy consumption fingerprints and occupancy patterns, which have been used to validate the proposed deep micro-moments solution.

In addition, a new visualization technique was proposed that helps in displaying the energy consumption footprints and easily identifying anomalies. Moreover, the evaluation performed on simulated and real databases revealed the effectiveness of the proposed architecture for detecting abnormal energy consumption. In fact, using the deep micro-moments model, accuracies of 99.58% and 99.29% have been achieved under the QUD and DRED datasets, while 93.91% accuracy has been reached under the PCSiD repository.

Nonetheless, this field of research needs more investigations, especially when considering outdoor environmental and climate conditions, such as temperature, humidity, luminosity, wind speed, and CO₂ emissions and even consumer's preferences including personnel behavior, comfort, and personality. In this case, the system will become more complex and more challenging. Accordingly, our next steps in this work will be twofold. First, the anomaly detection system will be updated with other information sources, such as ambient conditions and outside weather footprints to detect other forms of anomalies, e.g., keeping the doors and windows open while an air conditioner or a heating system is on. Second, it will be also part of our future work to develop a recommender system that will exploit the detected anomalies in order to help end-users in reducing wasted energy and optimizing power usage. In that respect, tailored recommendations will be triggered and sent to end-users to notify them real-time through a mobile application.

Funding Open Access funding provided by the Qatar National Library. This paper was made possible by National Priorities Research Program (NPRP) grant No. 10-0130-170288 from the Qatar National Research Fund (a member of Qatar Foundation).

Compliance with Ethical Standards

Disclaimer The statements made herein are solely the responsibility of the authors.

Conflict of Interest The authors declare that they have no conflict of interest.

Ethical Approval This article does not contain any studies with human participants or animals performed by any of the authors.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Song K, Ahn J, Ahn Y, Park M, Kwon N. Reduction and transformation of energy use data for end-user group categorization in dormitory buildings. *J Build Eng*. 2020;32:101524.
- Chen M, Lin C. Standby power management of a smart home appliance by using energy saving system with active loading feature identification. *IEEE Trans Consum Electron*. 2019;65(1):11–17.
- Wang AL, Chen BX, Wang CG, Hua D. Non-intrusive load monitoring algorithm based on features of v-i trajectory. *Electr Power Syst Res*. 2018;157:134–144.
- Sardianos C, Varlamis I, Dimitrakopoulos G, Anagnostopoulos D, Alsalemi A, Bensaali F, Himeur Y, Amira A. Rehab-c: Recommendations for energy habits change. *Futur Gener Comput Syst*. 2020;112:394–407.
- Himeur Y, Elsalemi A, Bensaali F, Amira A. Improving in-home appliance identification using fuzzy-neighbors-preserving analysis based qr-decomposition. In: *International Congress on Information and Communication Technology (ICICT)*; 2020. p. 1–8.
- Himeur Y, Alsalemi A, Bensaali F, Amira A. Robust event-based non-intrusive appliance recognition using multi-scale wavelet packet tree and ensemble bagging tree. 2020, Vol. 267.
- Kamal A, Al-Ghamdi SG, Koc M. Revaluing the costs and benefits of energy efficiency: a systematic review. *Energy Res Soc Sci*. 2019;54:68–84.
- Cheng L, Yu T. A new generation of ai: a review and perspective on machine learning technologies applied to smart energy and electric power systems. *Int J Energy Res*. 2019;43(6):1928–1973.
- Alsalemi A, Himeur Y, Bensaali F, Amira A, Sardianos C, Varlamis I, Dimitrakopoulos G. Achieving domestic energy efficiency using micro-moments and intelligent recommendations. *IEEE Access*. 2020;8:15047–15055.
- Hu S, Yan D, Azar E, Guo F. A systematic review of occupant behavior in building energy policy. *Build Environ*. 2020;175:106807.
- Delzendeh E, Wu S, Lee A, Zhou Y. The impact of occupants' behaviours on building energy analysis: a research review. *Renew Sust Energy Rev*. 2017;80:1061–1071.
- Sardianos C, Varlamis I, Chronis C, Dimitrakopoulos G, Himeur Y, Alsalemi A, Bensaali F, Amira A. A model for predicting room occupancy based on motion sensor data. In: *2020 IEEE International Conference on Informatics, IoT, and Enabling Technologies (ICIoT)*; 2020. p. 394–399.
- Anjana K, Shaji R. A review on the features and technologies for energy efficiency of smart grid. *Int J Energy Res*. 2018;42(3):936–952.
- Assimakopoulos M, Bamparesos N, Pantazaras A, Karlessi T, Lee S. On the comparison of occupancy in relation to energy consumption and indoor environmental quality: a case study. *Energy Proced*. 2017;134:875–884. *sustainability in Energy and Buildings 2017: Proceedings of the Ninth KES International Conference, Chania*. p. 5–7.
- Himeur Y, Alsalemi A, Bensaali F, Amira A, Sardianos C, Varlamis I, Dimitrakopoulos G. On the applicability of 2d local binary patterns for identifying electrical appliances in non-intrusive load monitoring. *Intelligent Systems and Applications*. In: Arai K., Kapoor S., and Bhatia R., editors. Cham: Springer International Publishing; 2021. p. 188–205.
- Susnea I, Pecheanu E, Cocu A, Hudec G. Improved occupancy-based solutions for energy saving in buildings. In: *2017 5th International Symposium on Electrical and Electronics Engineering (ISEEE)*; 2017. p. 1–5.
- Yousefi F, Gholipour Y, Yan W. A study of the impact of occupant behaviors on energy performance of building envelopes using occupants' data. *Energy Build*. 2017;148:182–198.
- Himeur Y, Alsalemi A, Al-Kababji A, Bensaali F, Amira A. Data fusion strategies for energy efficiency in buildings: overview, challenges and novel orientations. *Information Fusion*. 2020:1–41.
- Bornemann B, Sohre A, Burger P. Future governance of individual energy consumption behavior change—a framework for reflexive designs. *Energy Res Soc Sci*. 2018;35:140–151. *energy and the Future*.
- Al-Marri W, Al-Habaibeh A, Watkins M. An investigation into domestic energy consumption behaviour and public awareness of renewable energy in qatar. *Sustain Cities Soc*. 2018;41:639–646.
- Ali H, Sharif M, Mussarat Y, Rehmani MH, Riaz F. A survey of feature extraction and fusion of deep learning for detection of abnormalities in video endoscopy of gastrointestinal-tract. *Artif Intell Rev*. 2019:1–73.
- Sial A, Singh A, Mahanti A. Detecting anomalous energy consumption using contextual analysis of smart meter data. *Wirel Netw*. 2019:1–18.
- Yan W. One-class extreme learning machines for gas turbine combustor anomaly detection. In: *2016 International Joint Conference on Neural Networks (IJCNN)*; 2016. p. 2909–2914.
- Bialas A, Michalak M, Flisiuk B. Anomaly detection in network traffic security assurance. *Adv Intell Syst Comput*. 2019;987:46–56.
- Hong J, Liu C, Govindarasu M. Integrated anomaly detection for cyber security of the substations. *IEEE Trans Smart Grid*. 2014;5(4):1643–1653.
- Gaur M, Makonin S, Bajic IV, Majumdar A. Performance evaluation of techniques for identifying abnormal energy consumption in buildings. *IEEE Access*. 2019;7:62721–62733.
- Chou J-S, Telaga AS. Real-time detection of anomalous power consumption. *Renew Sust Energy Rev*. 2014;33:400–411.
- Capozzoli A, Piscitelli MS, Brandi S, Grassi D, Chicco G. Automated load pattern learning and anomaly detection for enhancing energy management in smart buildings. *Energy*. 2018;157:336–352.
- Wu C, Zhang Y, Jia J, Zhu W. Mobile contextual recommender system for online social media. *IEEE Trans Mob Comput*. 2017;16(12):3403–3416.
- How mobile became a power tool in idle moments, think with google, <https://www.thinkwithgoogle.com/>, accessed: 2019-08-07.
- Alsalemi A, Sardianos C, Bensaali F, Varlamis I, Amira A, Dimitrakopoulos G. The role of micro-moments: a survey of habitual behavior change and recommender systems for energy saving. *IEEE Syst J*. 2019;13(3):3376–3387.

32. Sardinos C, Varlamis I, Dimitrakopoulos G, Alsalemi A, Bensaali F, Amira A. I want to change” micro-moment based recommendations can change users. In: Proceedings of 8th International Conference on Smart Cities and Green ICT Systems; 2019. p. 1–10.
33. Alsalemi A, Bensaali F, Amira A, Fetais N, Sardianos C, Varlamis I. Smart energy usage and visualization based on micro-moments. *Intelligent Systems and Applications*. In: Bi Y, Bhatia R, and Kapoor S, editors. Cham: Springer International Publishing; 2020. p. 557–566.
34. Ieracitano C, Adeel A, Morabito FC, Hussain A. A novel statistical analysis and autoencoder driven intelligent intrusion detection approach. *Neurocomputing*. 2020;387:51–62.
35. Zhang X, Han Y, Xu W, Wang Q. Hoba. *Information Sciences: A novel feature engineering methodology for credit card fraud detection with a deep learning architecture*; 2019.
36. Ramchandran A, Kumar A. Unsupervised deep learning system for local anomaly event detection in crowded scenes. *Multimed Tools Appl*. 2019: 1–21. <https://doi.org/10.1007/s11042-019-7702-5>.
37. Kim E, Lee J, Shin H, Yang H, Cho S, kwan Nam S, Song Y, a Yoon J, il Kim J. Champion-challenger analysis for credit card fraud detection: Hybrid ensemble and deep learning. *Expert Syst Appl*. 2019;128:214–224.
38. Uttama Nambi AS, Reyes Lua A, Prasad VR. Loced: Location-aware energy disaggregation framework. In: Proceedings of the 2nd ACM International Conference on Embedded Systems for Energy-Efficient Built Environments, BuildSys '15; 2015. p. 45–54.
39. Alsalemi A, Ramadan M, Bensaali F, Amira A, Sardianos C, Varlamis I, Dimitrakopoulos G. Endorsing domestic energy saving behavior using micro-moment classification. *Appl Energy*. 2019;250:1302–1311.
40. Kolter JZ. Redd: a public data set for energy disaggregation research. In: Proceedings of the 1st KDD workshop on data mining applications in sustainability (sustKDD), ACM. San Diego, CA; 2011.
41. Makonin S, Popowich F, Bartram L, Gill B, Bajic IV. Ampds: A public dataset for load disaggregation and eco-feedback research. In: 2013 IEEE Electrical Power Energy Conference; 2013. p. 1–6.
42. Stephen Makonin IVB, Ellert B, Popowich F. Electricity, water, and natural gas consumption of a residential house in Canada from 2012 to 2014. *Sci Data*. 2016;3(180048):1–12.
43. Makonin S, Wang ZJ, Tumpach C. Rae: The rainforest automation energy dataset for smart grid meter data analysis. *Data*. 2018;3(1):1–9.
44. Murray D, Liao J, Stankovic L, Stankovic V, Hauxwell-Baldwin R, Wilson C, Coleman M, Kane T, Firth S. A data management platform for personalised real-time energy feedback. In: Proceedings of the 8th international conference on energy efficiency in domestic appliances and lighting; 2015.
45. Kelly J, Knottenbelti W. The uk-dale dataset, domestic appliance-level electricity demand and whole-house demand from five uk homes. *Sci Data*. 2015;2(150007):1–14.
46. Janetzko H, Stoffel F, Mittelstadt S, Keim DA. Anomaly detection for visual analytics of power consumption data. *Comput Graph*. 2014;38:27–37.
47. Ma Z, Song J, Zhang J. A real-time detection method of abnormal building energy consumption data coupled pod-lse and fcd. *Proced Eng*. 2017;205:1657–1664. 10Th International Symposium on Heating, Ventilation and Air Conditioning, ISHVAC2017, 19-22 October 2017, Jinan, China.
48. Cui W, Wang H. Anomaly detection and visualization of school electricity consumption data. In: 2017 IEEE 2nd International Conference on Big Data Analysis (ICBDA); 2017. p. 606–611.
49. Manimaran A, Chandramohan D, Shrinivas S, Arulkumar N. A comprehensive novel model for network speech anomaly detection system using deep learning approach. *International Journal of Speech Technology*. 2020:1–9. <https://doi.org/10.1007/s10772-020-09693-z>.
50. Aldweesh A, Derhab A, Emam AZ. Deep learning approaches for anomaly-based intrusion detection systems: A survey, taxonomy, and open issues. 2020, Vol. 189.
51. Chalapathy R, Chawla S. Deep learning for anomaly detection: A survey. arXiv:1901.03407. 2019.
52. Buzau M, Tejedor-Aguilera J, Cruz-Romero P, Gomez-Exposito A. Hybrid deep neural networks for detection of non-technical losses in electricity smart meters. *IEEE Trans Power Syst*. 2020;35(2):1254–1263.
53. Shyu M-L, Chen S-C, Sarinnapakorn K, Chang L. A novel anomaly detection scheme based on principal component classifier. In: Proceedings of International Conference on Data Mining; 2003. p. 1–10.
54. Lin G, Claridge DE. A temperature-based approach to detect abnormal building energy consumption. *Energy Build*. 2015;93:110–118.
55. Araya DB, Grolinger K, ElYamany HF, Capretz MAM, Bit-suamlak G. Collective contextual anomaly detection framework for smart buildings. In: 2016 International Joint Conference on Neural Networks (IJCNN); 2016. p. 511–518.
56. Liu X, Iftikhar N, Nielsen PS, Heller A. Online anomaly energy consumption detection using lambda architecture. *Big Data Analytics and Knowledge Discovery*. In: Madria S and Hara T, editors. Cham: Springer International Publishing; 2016. p. 193–209.
57. Rosin A, Auvaart A, Lebedev D. Analysis of operation times and electrical storage dimensioning for energy consumption shifting and balancing in residential areas. *Electron Electr Eng*. 2012;120:1–6.
58. Issi F, Kaplan O. The determination of load profiles and power consumptions of home appliances. *Energies*. 2018;11:(3).
59. Ryu S, Noh J, Kim H. Deep neural network based demand side short term load forecasting. *Energies*. 2017;10:(1).
60. Bourdeau M, qiang Zhai X, Nefzaoui E, Guo X, Chatellier P. Modeling and forecasting building energy consumption: A review of data-driven techniques. *Sustain Cit Soc*. 2019;48:101533.
61. Ojala M, Garriga GC. Permutation tests for studying classifier performance. *J Mach Learn Res*. 2010;11:1833–1863.
62. Maad S. Towards Human, Institutional and Economic Sectors Growth. London: InTechOpen; 2017.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.