Title	Non-intrusive load monitoring and classification of activities of daily living using residential smart meter data
Author(s)	Devlin, Michael A.; Hayes, Barry P.
Publication date	2019-05-24
Original citation	Devlin, M. A. and Hayes, B. P. (2019) 'Non-Intrusive Load Monitoring and Classification of Activities of Daily Living using Residential Smart Meter Data', IEEE Transactions on Consumer Electronics, In Press, doi: 10.1109/TCE.2019.2918922
Type of publication	Article (peer-reviewed)
Link to publisher's version	https://ieeexplore.ieee.org/document/8721550 http://dx.doi.org/10.1109/TCE.2019.2918922 Access to the full text of the published version may require a subscription.
Rights	**© 2019 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.**
Item downloaded from	http://hdl.handle.net/10468/8119

Non-Intrusive Load Monitoring and Classification of Activities of Daily Living using Residential Smart Meter Data

Michael A. Devlin and Barry P. Hayes, *Senior Member, IEEE*

***Abstract*—This paper develops an approach for household appliance identification and classification of household Activities of Daily Living (ADLs) using residential smart meter data. The process of household appliance identification, i.e. decomposing a mains electricity measurement into each of its constituent individual appliances, is a very challenging classification problem. Recent advances have made deep learning a dominant approach for classification in fields such as image processing and speech recognition. This paper presents a deep learning approach based on multi-layer, feedforward neural networks that can identify common household electrical appliances from a typical household smart meter measurement. The performance of this approach is tested and validated using publicly-available smart meter data sets. The identified appliances are then mapped to household activities, or ADLs. The resulting ADL classifier can provide insights into the behaviour of the household occupants, which has a number of applications in the energy domain and in other fields.**

***Index Terms*—Load identification, non-intrusive load monitoring, energy disaggregation, smart metering, appliance identification, machine learning.**

I. Introduction

ELECTRICITY smart meters are being installed in vast numbers worldwide, with an estimated 72% of European homes to have smart meters installed by the end of 2020 [1]. Research has shown that it is possible to accurately identify household appliance usage by analysing residential smart meter data [2]–[4]. Smart meters typically record the aggregate energy consumption of each customer at the whole building level. Previous studies have shown that it may be possible to "disaggregate" whole building energy profiles into individual appliance-by-appliance energy consumption profiles, since each appliance in the home has a particular power consumption profile and electrical characteristics [5]–[7]. A detailed review of consumer systems and disaggregation methods for residential buildings is provided in [2].

If appliance usage can be accurately identified from smart meter data, this information has a number of useful applications [8], [9]. These include: (i) providing detailed feedback to the consumer on their energy usage and the contributors to that energy usage; (ii) improved short- and long-term forecasting of demand profiles and spatial load forecasting; (iii) design of demand response and demand management schemes; (iv) measurement and validation of building energy efficiency schemes; (v) consumer profiling and classification; and (vi) transactive energy. The behavioural patterns of the householder(s), including occupancy, sleeping patterns, and other Activities of Daily Living (ADLs), can also be inferred from appliance usage data [10]. This has applications both in the energy domain and in other areas, including commercial services (customer profiling, targeted marketing), legal (monitoring of curfews and exclusion orders, detection of illegal activities) and remote healthcare (non-intrusive monitoring of older persons living at home).

Non-Intrusive Load Monitoring (NILM), or Non-Intrusive Appliance Load Monitoring (NIALM), is the process of analysing the aggregate household mains power measurement in the house, and disaggregating this into individual appliances. The NILM concept is was first developed and patented by Hart, Kern and Schweppe in 1986 [11], and published in the seminal paper by Hart in 1992 [12]. With the proliferation of domestic smart meters and the availability of more detailed household power measurements from large numbers of households, there has been a surge in interest in NILM in the last several years [13].

However, energy disaggregation is an extremely challenging problem and research in this area is still at an early stage, with significant technical and practical challenges yet to be overcome. This paper develops and tests a NILM classifier based on deep learning techniques that can detect the activations of selected major household appliances. This NILM classifier is then applied to identify ADLs in a typical European residential household.

The paper is structured as follows. Section II discusses previous literature and the current state of the art in NILM, and its relevance to mass market products and services. Section III outlines the methodology for NILM and the ADL classifier. Section IV provides the results. Section V discusses these results in the context of the current state of the art in NILM. Finally, the conclusions are presented in Section VI.

II. Literature Review and State of the Art

A. Previous Literature on NILM

Early techniques for NILM analysed the electricity mains measurement and applied statistical techniques in order to detect changes in the electrical consumption signal due to appliance on/off events. Similar "steady-state" parts of the

M.A. Devlin is with LM Ericsson Ltd., Athlone, Co. Westmeath, Ireland (e-mail: michael.devlin@ericsson.com).

B.P. Hayes is with the Department of Electrical and Electronic Engineering at University College Cork, 60 College Road, Cork, Ireland (e-mail: barry.hayes@ucc.ie)

power signal are then clustered together and the active and reactive power signatures are matched to the appropriate appliance using a best likelihood algorithm. Such clustering methods have been demonstrated to identify certain two-state (on/off) appliances with a high level of accuracy [14], [15]. However, these approaches have significant issues in identifying more complex appliances with multiple states (e.g. washing machines), and tend to have problems with cases where there is simultaneous operation and switching of multiple appliances [7]. Clustering techniques have also been applied in order to find patterns in electrical usage data and reveal household characteristics [16], [17].

A number of approaches for energy disaggregation use an edge detection algorithm to detect transitions in the power time series and match active and/or reactive power profiles to an existing set of "template" appliance power signatures using dynamic time warping. However, this approach requires a large amount of a priori information from the household in question, including the number, type and power ratings of each appliance and detailed measurements of similar appliances to those installed in the household. This makes it very difficult to generalise this approach to unseen households [18].

Hidden Markov Models (HMMs) are applied to the energy disaggregation problem in [5], [19], [20]. In the Markovian model, the hidden state is the state of a particular appliance, and the observation is the aggregate power demand time series. However, such models may be better suited to applications such as speech recognition, where the time duration of each state is relatively constant. This is a significant disadvantage in energy disaggregation, since appliance run times (and hence state durations) can vary from run to run by several orders of magnitude. It is also necessary include every individual appliance in the household in the HMM which may be undesirable, and/or infeasible.

Other recent approaches to NILM in the literature have included graph signal processing, as presented in [6]. A low-complexity unsupervised NILM algorithm is presented in [21] based on a fuzzy clustering algorithm called entropy index constraints competitive agglomeration. This algorithm showed promising results for practical NILM implementation. A cepstrum-smoothing-based load disaggregation method in order to deal effectively with the simultaneous on/off events of multiple appliances is described in [22]. NILM algorithms based on integer programming [23], and mixed-integer linear programming [24], have also been proposed.

B. Mass Market Products and Services using NILM

NILM techniques have been employed in a number of mass market products and services, where the main application to date has been in smart metering and energy management in residential buildings. Home Energy Management Systems (HEMS) designed to monitor, control and manage building energy use have received significant attention in recent literature [25]–[28]. In [25], each home electrical appliance is interfaced with a data acquisition module that is an IoT object with a unique IP address, resulting in a large mesh wireless network of devices. A HEMS solution based on ZigBee-enabled energy measurement modules, which monitor the energy consumption of home appliances and lights is presented in [26]. Home automation systems based on IEEE802.15.4 and ZigBee commications are also outlined in [27] and [28].

In order to monitor each electrical appliance in a residential building for energy management purposes, some previous work has used a sub-metering approach, where a current sensor is installed on every individual appliance, and this information is wirelessly communicated to a central hub, or HEMS. Sub-metering provides accurate information on all of the selected appliances, which is then processed by the HEMS in order to provide detailed feedback to home users on their energy use.

However, sub-metering is a costly and highly-invasive solution to home energy monitoring, since a large number of sensors need to be installed in the home, and a relatively advanced communications network is needed in order to transmit all of the required data streams to the HEMS. Another significant disadvantage of sub-metering is that new sensors need to be installed each time a new electrical appliance is added or replaced in the home.

NILM has been used in a significant number of mass market home energy management products and services, since it can provide a much less invasive and lower-cost solution than sub-metering. Sense [29] uses NILM to identify patterns in home energy use, with the aim of providing advice to consumers on improving home energy efficiency. Smappee [30] is focused on use of NILM to provide detailed feedback and advice on energy and carbon footprint reduction. SmartB [31] has developed a NILM device for commercial buildings. A number of mass market NILM products are applied to detect potential safety issues in situations where home appliances, such as the oven or iron have being left switched on and/or unattended [32].

Several commercial vendors mention that their products and services use machine learning or artificial intelligence in their algorithms [33], [34]. Bidgely [33] implements machine learning-based NILM algorithms, and holds a number of patents in this area. Verv [34] is a home energy management solution which carries out NILM using high-resolution mains electricity measurements and artificial intelligence approaches, where the output from the NILM classifier is the used to provide advice and recommendations to users.

All of the above examples are mass market products and services which employ NILM techniques as a key part of their technology[1]. However, the details of the NILM techniques used in all of these products are proprietary, and are generally kept as trade secrets, or are patent-protected by the companies involved. It is therefore very difficult to obtain a clear understanding of the NILM methods used in each of these products, and reliable information on the NILM performance and classification accuracy in each case is not currently available.

C. Applying a Deep Learning Approach to NILM

The adaption of deep learning techniques from other fields such as image processing to the NILM problem was proposed

[1]It should be noted that this paper is not intended to provide a complete list of commercial NILM products and services, and there are other existing products and services using NILM not mentioned here.

in [7], where preliminary results showed that deep learning approaches performed well on unseen household smart meter data sets, compared to other approaches in the literature. Deep learning is now a dominant approach in many areas, including image classification, automatic speech recognition, machine language translation [35]. It is anticipated that deep learning techniques could improve NILM performance, since one of the key difficulties in NILM is determining the most discriminative features to extract from a particular household data set. Deep learning techniques have the capability to automatically learn which features to extract from a data set, and to generalise to new and unseen data sets. This allows the development of unsupervised solution to the NILM problem, where the amount of user intervention required to set up and train the system is minimised.

In this paper, a deep learning approach to NILM is developed and validated using demand data taken from existing public smart meter data. The performance of this approach is tested and validated using publicly-available smart meter data sets. The identified appliances are then mapped to household activities, or ADLs.

III. Methodology

A. Identification of Major Appliances from Smart Meter Data

The type of information that can be inferred from smart meter data depends on the time resolution of the data. Figure 1 illustrates different electricity mains sampling frequencies, including the features which can be extracted from the data at each sampling frequency, and the appliances that can be identified. Hourly or half-hourly smart meter data can be used to infer household occupancy, whereas higher meter data resolutions can be used to detect the usage of a range individual appliances [7], [8], [10]. With very high resolution (MHz) data, higher-order harmonics can, in theory, be used to identify a large range of appliances, including consumer electronics and lighting loads. The focus of this paper is on the data sampling frequency to the order of 0.0166 - 1 Hz (corresponding to the area highlighted in red in Fig. 1).

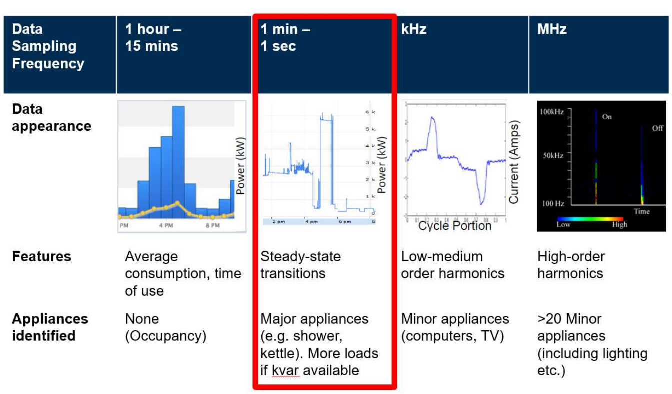

Fig. 1. Data sampling frequency and identification of electrical appliances (adapted from [7], [8], [10]).

Typically, smart meters transmit average kWh consumption data to the utility at regular intervals of 15 or 30 minutes. However, most residential smart meters are capable of recording quantities including kW, kvar, and voltage at a time resolutions in the range 0.1 to 1 Hz, and this data can be accessed in real-time via an appropriate gateway device, such as a Home Area Network (HAN) device based on the wireless personal area network standard IEEE 802.15.4. Figure 2 shows a sample of the recorded active power consumption measured from the household electricity mains and accessed at 10-second intervals.

Previous research has addressed the NILM problem by attempting to separate the aggregate whole building level demand profile such as in Fig. 2 into each of its constituent appliance profiles. However, some appliances are easier to detect in NILM than others. The active power demand profiles of "major" appliances, with large peak demands (greater than 1-2 kW) such as space/water heating and cooking appliances are generally easier to detect, since they have a high signal-to-noise ratio in the aggregated data.

Another important factor in this is the complexity and repeatability of the usage cycle of the appliance. For example, an electric kettle has a very simple cycle, with only two states of operation (on/off). An electric oven has two states in each cycle: an initial heating phase (a steady demand of 2-3 kW while the device reaches the target temperature); and a thermostatic control phase (where the active power demand alternates between 0 W and 2-3 kW as the heating element is switched on/off). Certain multi-state appliances, such as dishwashers and washing machines, are more difficult to detect in NILM since they have multiple stages in each cycle and highly-variable cycle times, which depend on user settings [7].

Other domestic load types, such as consumer electronics and lighting, have small peak demands (less than 100-200 W), and are considered "minor" appliances. These have relatively low signal-to-noise ratios in the aggregated electricity mains data. Their presence can only be identified if specialised power quality recording equipment is installed in the household in order to capture high resolution (kHz-MHz frequency) measurements and analyse the harmonic emission signatures, Fig. 1.

In this paper, efforts are focused on the identification of a number of selected appliances, which have large peak demands (1-2 kW or greater) and less complex power signatures. These appliances make up the majority of household electricity consumption, and can be identified by analysing the active power consumption stream from a regular domestic smart

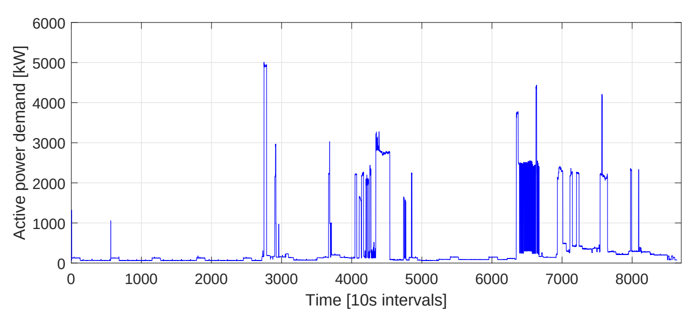

Fig. 2. Recorded active power consumption from electricity smart meter at an individual household over a 24 hour period.

meter, without the need for installation of special equipment.

B. Edge Detection Algorithm

A rule-based algorithm was developed for detecting the "on" and "off" events of the major appliances discussed above. The algorithm needs to be supplied with three parameters for each appliance: average power consumption of the appliance while switched on; and the maximum and minimum cycle time, i.e. time the appliances would typically be active.

The active power measured by the smart meter at time t is:

$$P(t) = \sum_{i=1}^{N_{total}} p_i(t) + e(t) \tag{1}$$

where $p_i(t)$ is the active power demand from appliance i at time instant t, N_{total} is the total set of appliances installed and $e(t)$ is the measurement noise. Appliance on/off events are detected using the adaptive threshold W. If the active power demand from appliance i is large enough, i.e. if:

$$|p_i(t) - p_i(t-1)| \geq W \tag{2}$$

Then appliance i is said to have changed its state. The adaptive threshold W depends on the set of appliances M and needs to be specified so that for all i, if W is exceeded, then appliance i has changed state [5]. W is the minimum state transition that needs to be detected:

$$W = max\{min_{(m \in M)} p_m, max_{(m \in M)} |max(p_m) - min(p_m)|\} \tag{3}$$

where p_m is a vector containing readings from the appliance m. W is adaptively changed every time a new appliance is detected and removed from the data set. A practical initial value for W is 1 kW which would allow for detection of major appliances with power ratings of 1 kW or larger. All appliances with power consumption levels below the threshold W will not be detected.

As shown in Fig. 3, the edge detection algorithm iterates through aggregated smart meter data searching for an "on" event for the target appliance. When searching for an "on" event, the algorithm checks for an increase in power demand in a specific power range by comparing the current power value to the power value of two data points back. Choosing two data points back instead of one data point reduces false negatives where the previous time step records the aggregated power demand when an appliance is midway through turning "on" or "off", and helps the algorithm to ignore transient spikes that can occur during appliance switch on or switch off. The disadvantage of using two time steps is that if another appliance turns "on" or "off" during the two time step window, it will cause the algorithm to miss the target appliance. It was found through experimentation that using two time steps resulted in more reliable appliance detection, and that this outweighed the disadvantage of potentially missing target appliances where another appliance turns "on" or "off" during the two time step window, since the number of instances of this was very small. Such a trade-off is inevitable for low resolution (0.1 to 1 Hz) smart meter data. Improving the detection of multiple, simultaneous switching events requires higher-resolution data than is available from a regular domestic smart meter, and is therefore beyond the scope of this paper.

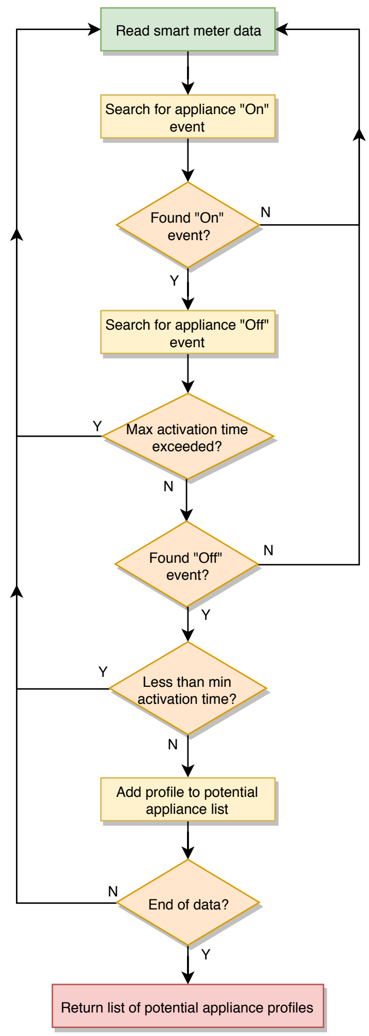

Fig. 3. Rule-based edge detection algorithm used to detect major appliance "on" and "off" events.

After an "on" event is detected, the algorithm will begin searching for a matching "off" event. This search is carried out in a similar fashion as the "on" event detection, except that the algorithm searches for a decrease in power consumption exceeding the adaptive threshold value W. If an "off" event is not detected within a certain time limit the algorithm will discard the previous "on" event and begins searching for a new one. If an appliance is detected but was active for less than the minimum cycle time then it is also discarded. The entire appliance power profile along with its associated times are recorded in an array. Once all the smart meter data has been iterated through, the algorithm returns a list of detected appliance profiles. If no "on" or "off" events are detected within the entire data set, the algorithm returns a blank list of detected appliance profiles, Fig. 3.

C. Design and Training of Deep Neural Network Classifier

A set of feedforward Neural Networks (NNs) were developed to classify potential appliance profiles returned from the edge detection algorithm using the TensorFlow machine learning library [36]. An NN was created and trained to classify each target appliance. As seen in Fig. 4, each neural network contains an input layer, a number of hidden intermediate layers and an output layer. The input layer size is dependent on the maximum number of data points in the potential appliance profiles. The number of intermediate hidden layers and the number of nodes in these hidden layers can be experimented with to find a configuration that yields the best results for the application.

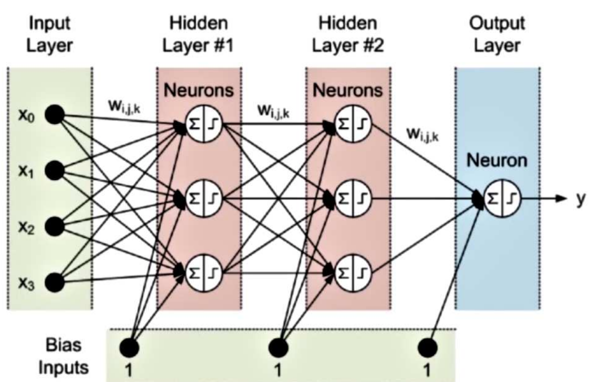

Fig. 4. Feedforward NN structure to classify major appliance profiles [36].

While there are no strict rules for choosing the hidden layers in general the more hidden layers, the more features the NN can extract from the data. Too many hidden layers causes over-fitting of the data and not enough hidden layers will cause the NN to under-fit the data. The larger the hidden layers the more computational power, time, and labelled data is required to train the network. The output layer is dependent on the number of classes that it uses to classify the data. Since an NN was trained for each target appliance, it only needs to use two classes; the input data either is the target appliance, or is not the target appliance (binary classification).

The edge detection algorithm was used to collect and label training and testing data for the NNs. The data was gathered by feeding aggregated smart meter data into the edge detection algorithm. The edge detection algorithm returns all the potential profiles for each target appliance. The disaggregated appliance smart meter data was then fed to the edge detector algorithm, the appliance profiles returned from this data are considered the true appliance activations. Before data is inputted into the NN for training, a base model is created for the network. In the base model, all the layers and nodes are arranged and initialised. All the initial connections between nodes are given random weights and each hidden layer node is given an initial random bias. This is to ensure that nodes in the network do not become "dead" by setting all their input connection weights to zero. Node biases are used to ensure that a node will output a non-zero value even if all inputs are zero.

The times when the potential appliance profiles occur are then compared to the true appliance activation times to determine if a potential appliance profile is true or not. All the potential appliance profiles are labelled in this manner. The labels given are [0, 1] if the profile is an appliance and [1, 0] if it is not an appliance. Since the input layers of the NNs are a fixed size, all potential appliance profiles must be padded out to the length of the input layer using zeroes. The length of the input layer was determined by using the maximum length of the potential appliance profiles encountered in the smart meter data.

The labelled data is then separated into training and testing data. The training data is used to train up the NN and the testing data is held back and used at the end to evaluate the performance of the neural network. Once training begins, the target appliance profiles are fed into the networks in batches of 100. This means that the network attempts to adjust its node connection weights and biases after every 100 appliance profiles using the aggregated error information from all 100 profiles as opposed to doing this for each appliance. The NN is trained in batches until all the training data has been used. This process that uses all the training data is called an epoch. Ten epochs where used to train all the NNs. The number of epochs performed also affects how well the network model can represent the data. Too many epochs lead to over-fitting the data and too few epochs lead to under-fitting the data. The neural network uses a gradient descent approach when adjusting its connection weights and biases.

D. Metrics Used to Calculate NILM Detection Accuracy

The metrics used to evaluate the NILM algorithms are:

- TP, the number of True Positives;
- FP, the number of False Positives; and
- FN, the number of False Negatives.

True positives represent the target appliance activations that are correctly classified. False positives represent non target appliance data that is incorrectly classified as the target appliance. False negatives represent missed target appliance activations.

The precision, p, is defined as:

$$p = \frac{TP}{TP + FP} \tag{4}$$

The recall, r, is defined as:

$$r = \frac{TP}{TP + FN} \tag{5}$$

Finally, the f1 score, $f1$, is defined as:

$$f1 = 2 \times \frac{pr}{p + r} \tag{6}$$

Precision represents the probability that any given appliance activation classified as the target appliance by the algorithm is actually the target appliance. Recall represents the probability that any given appliance activation in the smart meter data is detected by the algorithm. The f1 score is the harmonic average of precision and recall.

IV. Results

A. Description of NILM Test Data

The main dataset used to test the NILM algorithms in this paper is the UK-DALE (United Kingdom-Domestic Appliance Level Dataset) dataset [37], which contains aggregated and disaggregated appliance data for five households in London, England over several years. The aggregated data was recorded using an electricity smart meter. The aggregated data represents the power demand of all appliances in the house (i.e. the mains measurement). The disaggregated power data of each appliance was measured via smart plugs on individual appliances that recorded their individual power demands. The aggregated and disaggregated power data were sampled at a rate of once every 6 seconds (0.1667 Hz).

Seven major appliances are targeted: the hair dryer, vacuum cleaner, iron, microwave, kettle, toaster,and oven. Each of these appliances had labelled measurement data available in the UK-DALE data set. A number of the target appliances had only a few hundred activations in total such as the hair dryer, vacuum and iron whereas the microwave, kettle and toaster had a few thousand activations. To increase the amount of training data available the true appliance profiles where duplicated and augmented. This was done by offsetting the appliance profiles and introducing a small amount of random noise to the data.

B. Selection of NN Configuration

A number of different network configurations were tested, ranging from networks with 1 hidden layer and 500 nodes to networks with 4 hidden layers with 1000 nodes per hidden layer. The networks were subjected to two different tests; a balanced test and an unbalanced test. In the balanced test the NN has to classify equal amounts of true and false appliance profiles. In the unbalanced test, the NN has to classify true and false profiles in the same ratio that the edge detector algorithm returns them from the smart meter data. In the unbalanced test, 30% of the total data is used for testing and the rest for training. In the balanced test majority of the data is used for training and up to 500 instances of each class is used for testing.

The balanced test is a good indicator of the performance of the NNs in isolation and the unbalanced test is a good indicator of how well the NNs perform when run on top of the edge detector algorithm. The results below show the results from the balanced test. The tests were run 10 times on each NN and an average of the metrics was calculated. Note the selection of the testing data was random for every test, this means that data that was chosen for testing in previous tests could be chosen for testing again in future tests.

The two top performing network configurations in the balanced test were the “2 hidden layers, 500 nodes” configuration with an f1 score of 0.773 and the “3 hidden layers, 500 nodes” configuration with an f1-score of 0.765, Fig. 5. The network configuration that was selected was the model with 3 hidden layers and 500 nodes per hidden layer as it had a higher precision score. The trade off in recall (see Fig. 5) was acceptable as the need to reduce false positives in the data was deemed more important than detecting every appliance activation.

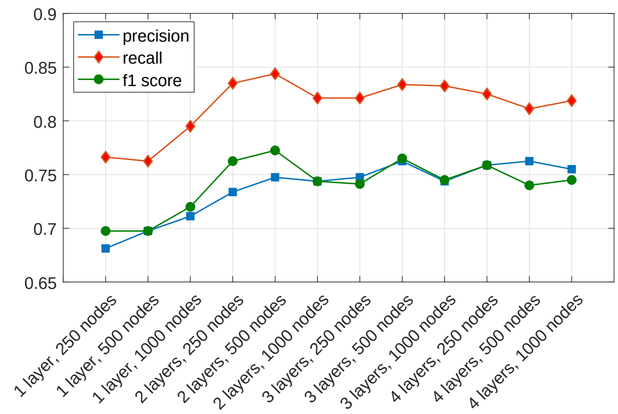

Fig. 5. Selection of number of NN hidden layers and nodes.

C. Results of NILM with NN Classifier

Table I shows the results for seven major household appliances in terms of precision, recall and f1 scores, with the lowest precision being 0.6 for the kettle and the lowest recall being 0.64 for the hair dryer. Appliances associated with cooking achieved the highest recall scores, with average recall scores of 0.92 to 0.99 for the four cooking appliances. The average f1 score is 0.776 or 77.6%.

TABLE I
Balanced test results for the 3 hidden layers, 500 nodes NN configuration using the UK-DALE data set [37].

Appliance name	Precision, p	Recall, r	f1 score
Hair dryer	0.85	0.64	0.70
Vacuum cleaner	0.75	0.87	0.80
Iron	0.85	0.66	0.66
Microwave	0.68	0.96	0.80
Kettle	0.60	0.99	0.74
Toaster	0.74	0.96	0.84
Oven	0.87	0.92	0.89
Average	0.763	0.857	0.776

Table II gives the results from the unbalanced test for seven major household appliances. The ANNs performed well when detecting the appliances associated with cooking (microwave, kettle, toaster and oven) with recall scores over 0.9 and acceptable precision scores. The NNs were also able to achieve good recall scores on other appliances, including the hair dryer, vacuum cleaner and iron, but had poor precision scores. The reason for the differences in results when comparing the unbalanced test to the balanced test is the effect of the edge detector algorithm. The kettle, which had the lowest precision score in the balanced test (0.6) has the high precision score in the unbalanced test. This is a result of the edge detector algorithm detecting significantly less false positives than true positives for the kettle. The precision scores of the hair dryer, vacuum and iron suffer significantly in the unbalanced test due to the large number of false positives from the edge detector algorithm for these appliances.

TABLE II
UNBALANCED TEST RESULTS FOR THE 3 HIDDEN LAYERS, 500 NODES NN CONFIGURATION USING THE UK-DALE DATA SET [37].

Appliance name	Precision, p	Recall, r	f1 score
Hair dryer	0.31	0.79	0.43
Vacuum cleaner	0.32	0.88	0.46
Iron	0.30	0.73	0.36
Microwave	0.61	0.92	0.73
Kettle	0.83	0.99	0.90
Toaster	0.68	0.93	0.78
Oven	0.81	0.99	0.89
Average	0.551	0.890	0.650

D. Classification of Activities of Daily Living

Once the results from NILM are available, the classification of ADLs is carried out by mapping each ADL to a set of criteria based on the appliances used, their consumption in Watt-hours (Wh), and the times that they are switched on. The "Sleeping" ADL is identified by an absence of detections of major appliances during the hours where the householder is most likely to be asleep, Fig. 6. In the period from 23:00 to 8:00, each 1-hour window in the power measurement time series is analysed. If the total household energy demand during this period is less than a pre-defined Wh value (250 Wh), the activity for that time period is classified as "Sleeping". A simple algorithm for the detection of the "Cooking" ADL is illustrated in Fig. 7. The algorithm relies on the detection of one of the two major cooking appliances in the household, the electric oven and the electric hob. If either of these appliances are detected, and if the energy consumption of these appliance is above a pre-defined threshold within the selected time period, then the ADL is classified as "Cooking". Table III illustrates similar requirements for the identification of several other ADLs.

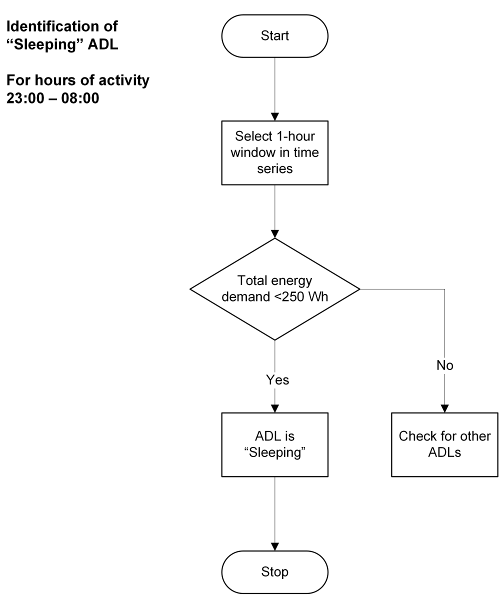

Fig. 6. Algorithm for identification of "Sleeping" ADL.

The criteria for classifying ADLs should be customised to each individual household. A survey of each household is carried out in advance to determine the correct list of major

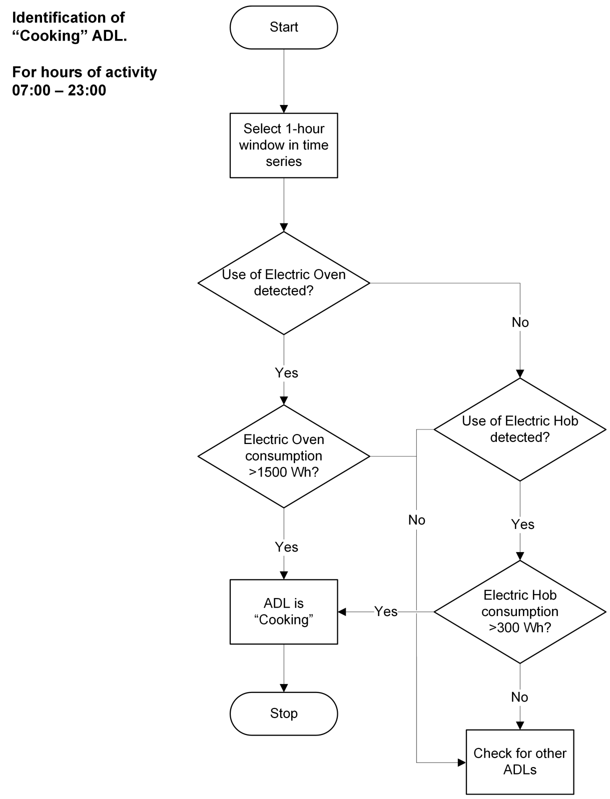

Fig. 7. Algorithm for identification of "Cooking" ADL.

TABLE III
SAMPLE OF REQUIREMENTS FOR IDENTIFICATION OF ADLS.

ADL	Major appliances	Consumption (Wh)	Hours active
Sleeping	None	Total demand < 250 Wh	23:00-08:00
Washing	Electric shower	> 2000 Wh	07:00-23:00
Cooking	Electric oven	> 1500 Wh	07:00-23:00
	Electric hob	> 300 Wh	07:00-23:00
Unoccupied	None	Total demand < 250 Wh	08:00-23:00

appliances in each household and the typical energy usage behaviour of the occupants, allowing the criteria for detection each ADL to be adjusted accordingly. It should be noted that for ADLs that are based on appliance activations (e.g. "Cooking" ADL), it is possible for multiple ADLs to occur at simultaneously, or for ADLs to partially overlap in time; for instance, the household activity might be classified as both "Cooking" and "Washing" during the same hour For ADLs based on the absence of detections of major appliances, such as the "Sleeping" or "Unoccupied" ADL, it is not possible for these to occur at the same time as other ADLs; i.e., the household cannot be classified as "Sleeping" and "Cooking" simultaneously.

E. Description of Experimental Trial Configuration

An experimental household trial designed to test the performance of the entire system, including the ADL classifier was developed. The experimental trial consists of two households, one terraced house located in an urban area of Ireland, and a larger home in a rural location. In Ireland, the electricity smart meter roll-out had not been completed at the time of writing, and electromechanical meters are currently used for residential electricity billing purposes. Another significant issue in collecting experimental trial data is that there are privacy concerns around the use of sensitive, real-time smart meter data (see also the discussion in Section VI of this paper). In order to avoid problems with access to the required data and develop a flexible smart meter configuration, which could be configured to monitor various types of data, an open-source monitoring system was installed in the trial households.

Figure 8 shows the installed monitoring hardware in one of the trial households. Each household is monitored via current and potential transformers connected to the electricity mains supply. The current transformer in in the form of a clamp meter which clips around the household mains cable as labelled in Fig. 8, and the voltage sensor is a potential transformer which is connected to the mains supply via a standard AC household plug. The hardware and software used in the trial is implemented using the Open Energy Monitor system [38], and the data rate is one frame per 10 seconds (0.1 Hz).

The current and voltage measurements are logged remotely via WiFi to a web application, which is used for processing, logging and visualising the recorded energy data. This uses a web Application Programming Interface (API) which exchanges data requests and responses in the JavaScript Object Notation (JSON) format. The resulting energy data feeds can be accessed in real-time, or as historical data sets. The NILM and ADL classifiers described in Section III were developed using Python programming.

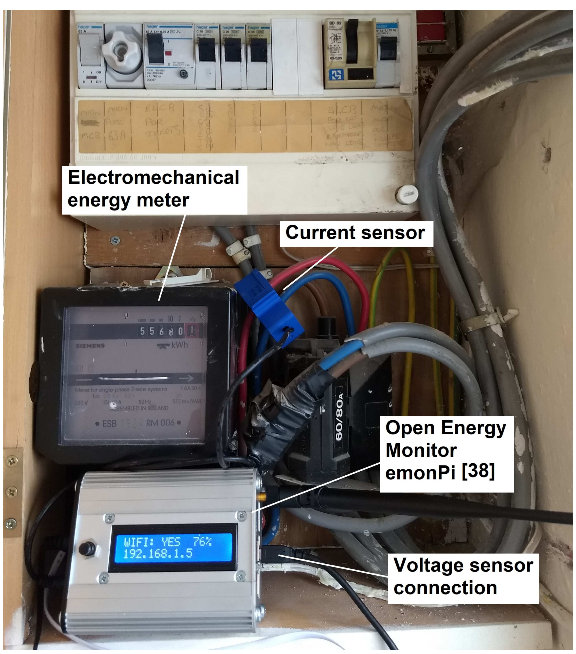

Fig. 8. Monitoring hardware installed in a typical Irish household for ADL classifier experimental trial.

F. Preliminary Results from Experimental Trials

The household ADLs are monitored using a diary kept by the occupants of the home in order to test if the ADLs identified by the classifier match with those recorded by the occupants. Figure 9 shows a sample of the ADLs identified at one of the households using the criteria in Table III. The hours from 00:00-08:00 are classified as the "Sleeping" ADL, due to the low total demand recorded and lack of major appliances detected during this period. The "Washing" ADL occurs at 08:00 (electric shower appliance detected), and cooking appliances are detected around midday and early evening. The household is unoccupied for a number of hours in the early afternoon.

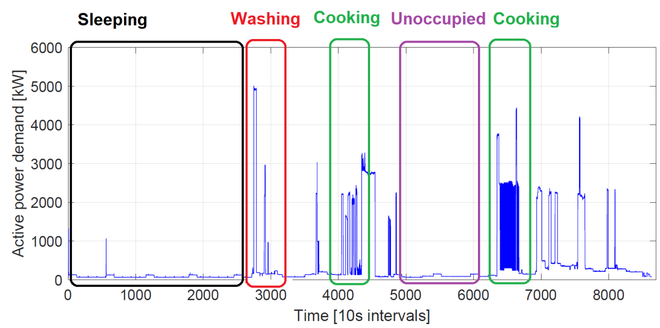

Fig. 9. Sample of ADLs identified from the electricity smart meter data at an individual household in Figure 2.

V. Discussion

The performance of the deep learning NILM classifier developed in this paper is tested using the publicly-available UK-DALE smart meter data set. A NN classifier with 3 hidden layers and 500 nodes per layer was demonstrated to have an average precision of 76.3%, recall of 85.7%, and f1 score of 77.6% over seven household appliances. Direct comparison with NILM classification results from previous research can be difficult, since a number of different data sets have been used in the literature, each with a different set of appliances and appliance characteristics. In addition, there is some variation in the metrics used to assess NILM classifier accuracy. While a large number of previous authors have employed the precision, recall and f1 score metrics described in (4)-(6), these metrics have not been applied universally in NILM.

However, it is possible to take the results of the balanced test in Table I, and consider the f1 score as a measure of the overall NILM classification accuracy in order to make some comparisons to the current state of the art in NILM. The NILM classifier results in [24] show accuracies for individual appliances varying from 0.40 to 0.88, depending on the appliance, with overall accuracies for 6 appliances given as 0.78 [24] and 0.76 [23]. The results in Section IV of this paper compare favourably with individual appliance f1 scores varying from 0.66 to 0.89, and an overall f1 score for 7 appliances of 0.78.

Results from other state of the art NILM classifiers are detailed in [7] and [18]. The accuracy scores for 5 appliances from the UK-DALE data set obtained in [7] ranged from 0.35 to 1.00, with an overall accuracy score of 0.68. This resulted in 92% of the total kWh energy consumed by the 5 appliances being assigned to the correct appliance [7]. A recent review paper [18], compares the NILM performance results from 11 different papers, with the aim of documenting the current state of the art in NILM. These results are based on a range of different data sets and metrics, which makes direct comparison difficult, but the majority of the 11 papers showed NILM accuracies in the 70-80% range, with three papers having an overall f1 score of 0.71 (71%).

In summary, the results in Section IV of this paper compare favourably to the results from other state of the art NILM classifiers detailed in the literature. This suggests that deep learning approaches can be applied successfully to the NILM problem. However, some model-specific information for each appliance being detected is required, and the NN classifier needs to be trained with several labelled appliance activations.

VI. Conclusions

This paper developed an approach for household appliance identification and classification of household Activities of Daily Living (ADLs) using residential smart meter data. Non-Intrusive Load Monitoring (NILM) was applied in order to disaggregate the household mains power measurement into individual appliances. This paper proposed a NILM approach based on multi-layer, feedforward Neural Networks (NNs) that can identify common household electrical appliances from a typical household smart meter measurement. The focus in the paper was on identifying a number of selected appliances, which have large peak demands (1-2 kW or greater) and less complex power signatures. These "major" appliances make up the majority of household electricity consumption, and can be identified by analysing the active power data from a regular domestic smart meter, without the need for installation of special equipment.

Another contribution of this work was the development of a framework for mapping the identified appliances to ADLs. The resulting ADL classifier can provide useful insights into the behaviour of the household occupants. This has a range of applications in the energy domain (e.g. providing feedback on energy usage in the home and the contributors to their energy usage) and also in other areas, including commercial services, legal, and remote healthcare applications. An experimental household trial designed to demonstrate the hardware and software required for implementing the NILM algorithm and test the performance of the ADL classifier was also described in Section IV-E of the paper.

The ADL classifier developed in this paper can provide useful information to the consumer, including detailed feedback on their energy usage and the contributors to that energy usage, allowing the creation of itemised energy bills. This information can then be used to highlight opportunities to save energy and reduce energy costs, and to identify inefficient and/or faulty home appliances. The NILM framework in this paper can also provide insights into the behaviour of the household occupants, which has a number of useful applications outside of the energy domain. A good example of this is the non-intrusive monitoring of elderly persons living alone, where the approach developed in this paper could be used to trigger an alarm if there is a sudden or unexpected change in household behaviours.

The real-time data streams from electricity smart meters used in this paper contain sensitive information about the personal habits and behaviours of the householder. In most countries, the use of smart metering data is highly-regulated and protected by privacy legislation. This is a major inhibitor of real-time smart meter data collection in practice [39]. In order for the proposed NILM and ADL classifier to be realised, the householders involved would need to give consent for their real-time smart meter data to be used. Householders may opt to provide this data in exchange for free access to the online services enabled by their data, such as itemised electricity billing, and remote home monitoring. However, a detailed discussion of smart metering privacy and security issues is outside the scope of this paper.

Future work will investigate the possibilities for improving NILM accuracy by including reactive power profiles, in cases where this information is available from the smart meter. The NILM performance may also be improved by removing the edge detection algorithm, and allow the NN classifier to process the smart meter data time series directly, via a sliding window. Further work will also investigate the benefits of applying the NNs to convert smart meter data directly into classes of user activity, rather than using the ADL mapping to appliances described in Section IV-D of this paper.

Acknowledgement

This work has been funded in part by Enterprise Ireland via the IERC project "EnerPort" (contract number TC-2013-0002B) and Science Foundation Ireland (grant number SFI/12/RC/2289, co-funded by the European Regional Development Fund).

References

[1] European Commission Joint Research Centre. (2018) Smart Metering deployment in the European Union. [Online]. Available: http://ses.jrc.ec.europa.eu/smart-metering-deployment-european-union

[2] M. Zeifman and K. Roth, "Nonintrusive appliance load monitoring: Review and outlook," *IEEE Trans. Consum. Electron.*, vol. 57, no. 1, pp. 76–84, Feb. 2011, doi: 10.1109/TCE.2011.5735484.

[3] M. A. Devlin and B. P. Hayes, "Non-intrusive load monitoring using electricity smart meter data: A deep learning approach," in *IEEE Power and Energy Society General Meeting*, 2019, accepted.

[4] M. A. Devlin and B. P. Hayes, "Load identification and classification of activities of daily living using residential smart meter data," in *IEEE PES Europe PowerTech Conf.*, 2019, accepted.

[5] J. Liao, G. Elafoudi, L. Stankovic, and V. Stankovic, "Non-intrusive appliance load monitoring using low-resolution smart meter data," in *2014 IEEE Int. Conf. on Smart Grid Commun. (SmartGridComm)*, Nov. 2014, pp. 535–540, doi: 10.1109/SmartGridComm.2014.7007702.

[6] K. He, L. Stankovic, J. Liao, and V. Stankovic, "Non-intrusive load disaggregation using graph signal processing," *IEEE Trans. Smart Grid*, vol. 9, no. 3, pp. 1739–1747, May 2018, doi: 10.1109/TSG.2016.2598872.

[7] J. Kelly, "Disaggregation of domestic smart meter energy data," Ph.D. dissertation, 2017.

[8] K. C. Armel, A. Gupta, G. Shrimali, and A. Albert, "Is disaggregation the holy grail of energy efficiency? the case of electricity," *Energy Policy*, vol. 52, pp. 213–234, 2013, doi: 10.1016/j.enpol.2012.08.062.

[9] B. Hayes, I. Hernando-Gil, A. Collin, G. Harrison, and S. Djokic, "Optimal power flow for maximizing network benefits from demand-side management," *IEEE Trans. Power Syst.*, vol. 29, no. 4, pp. 1739–1747, Jul. 2014, doi: 10.1109/TPWRS.2014.2298894.

[10] E. McKenna, I. Richardson, and M. Thomson, "Smart meter data: Balancing consumer privacy concerns with legitimate applications," *Energy Policy*, vol. 41, pp. 807–814, 2012, doi: 10.1016/j.enpol.2011.11.049.

[11] G. Hart, E. Kern, and F. Schweppe, "Non-intrusive appliance monitor apparatus," Patent US4 858 141A, 1989.

[12] G. W. Hart, "Nonintrusive appliance load monitoring," *Proc. of the IEEE*, vol. 80, no. 12, pp. 1870–1891, Dec. 1992, doi: 10.1109/5.192069.

[13] C. Klemenjak and P. Goldsborough, "Non-intrusive load monitoring: A review and outlook," in *Proc. Joint German/Austrian Conf. on Artificial Intell. (KI2016), Klagenfurt, Austria*, 2016, pp. 1–17.

[14] Z. Wang and G. Zheng, "Residential appliances identification and monitoring by a nonintrusive method," *IEEE Trans. Smart Grid*, vol. 3, no. 1, pp. 80–92, Mar. 2012, doi: 10.1109/PerCom.2012.6199866.

[15] M. Weiss, A. Helfenstein, F. Mattern, and T. Staake, "Leveraging smart meter data to recognize home appliances," in *2012 IEEE Int. Conf. on Pervasive Comput. and Commun.*, Mar. 2012, pp. 190–197, doi: 10.1109/PerCom.2012.6199866.

[16] C. Beckel, L. Sadamori, T. Staake, and S. Santini, "Revealing household characteristics from smart meter data," *Energy*, vol. 78, pp. 397 – 410, 2014, doi: 10.1016/j.energy.2014.10.025.

[17] K. Gajowniczek and T. Zabkowski, "Data mining techniques for detecting household characteristics based on smart meter data," *Energies*, vol. 8, no. 7, pp. 7407–7427, 2015, doi: 10.3390/en8077407.

[18] R. Bonfigli, S. Squartini, M. Fagiani, and F. Piazza, "Unsupervised algorithms for non-intrusive load monitoring: An up-to-date overview," in *Proc. IEEE 15th Int. Conf. on Environment and Elect. Eng. (EEEIC)*, Jun. 2015, pp. 1175–1180, doi: 10.1109/EEEIC.2015.7165334.

[19] D. Egarter, V. P. Bhuvana, and W. Elmenreich, "Paldi: Online load disaggregation via particle filtering," *IEEE Trans. Instrum. Meas.*, vol. 64, no. 2, pp. 467–477, Feb. 2015, doi: 10.1109/TIM.2014.2344373.

[20] F. J. Ferrández-Pastor, H. Mora-Mora, J. L. Sánchez-Romero, M. Nieto-Hidalgo, and J. M. García-Chamizo, "Interpreting human activity from electrical consumption data using reconfigurable hardware and hidden markov models," *Journal of Ambient Intell. and Humanized Comput.*, vol. 8, no. 4, pp. 469–483, Aug. 2017, doi: 10.1007/s12652-016-0431-y.

[21] Q. Liu, K. M. Kamoto, X. Liu, M. Sun, and N. Linge, "Low-complexity non-intrusive load monitoring using unsupervised learning and generalized appliance models," *IEEE Trans. Consum. Electron.*, vol. 65, no. 1, pp. 28–37, Feb. 2019, doi: 10.1109/TCE.2019.2891160.

[22] S. Kong, Y. Kim, R. Ko, and S. Joo, "Home appliance load disaggregation using cepstrum-smoothing-based method," *IEEE Trans. Consum. Electron.*, vol. 61, no. 1, pp. 24–30, Feb. 2015, doi: 10.1109/TCE.2015.7064107.

[23] M. Z. A. Bhotto, S. Makonin, and I. V. Bajic, "Load disaggregation based on aided linear integer programming," *IEEE Trans. Circuits Syst. II: Express Briefs*, vol. 64, no. 7, pp. 792–796, Jul. 2017, doi: 10.1109/TCSII.2016.2603479.

[24] F. M. Wittmann, J. C. Lopez, and M. J. Rider, "Nonintrusive load monitoring algorithm using mixed-integer linear programming," *IEEE Trans. Consum. Electron.*, vol. 64, no. 2, pp. 180–187, May 2018, doi: 10.1109/TCE.2011.5735484.

[25] A. R. Al-Ali, I. A. Zualkernan, M. Rashid, R. Gupta, and M. Alikarar, "A smart home energy management system using IoT and big data analytics approach," *IEEE Trans. Consum. Electron.*, vol. 63, no. 4, pp. 426–434, Nov. 2017, doi: 10.1109/TCE.2017.015014.

[26] J. Han, C. Choi, W. Park, I. Lee, and S. Kim, "Smart home energy management system including renewable energy based on zigbee and plc," *IEEE Trans. Consum. Electron.*, vol. 60, no. 2, pp. 198–202, May 2014, doi: 10.1109/TCE.2014.6851994.

[27] D. Han and J. Lim, "Design and implementation of smart home energy management systems based on zigbee," *IEEE Trans. Consum. Electron.*, vol. 56, no. 3, pp. 1417–1425, Aug. 2010, doi: 10.1109/TCE.2010.5606278.

[28] K. Gill, S. Yang, F. Yao, and X. Lu, "A zigbee-based home automation system," *IEEE Trans. Consum. Electron.*, vol. 55, no. 2, pp. 422–430, May 2009, doi: 10.1109/TCE.2009.5174403.

[29] Sense (sense.com). (2019). [Online]. Available: http://www.sense.com

[30] Smappee. (2019). [Online]. Available: https://www.smappee.com/be_en/homepage

[31] SmartB Energy Management. (2019). [Online]. Available: http://www.smartb.de/

[32] Watty. (2019). [Online]. Available: https://watty.io/

[33] Bidgely. (2019). [Online]. Available: http://www.bidgely.com

[34] Verv. (2019). [Online]. Available: https://verv.energy/

[35] J. Kelly and W. Knottenbelt, "Neural NILM: Deep neural networks applied to energy disaggregation," in *Proc. of the 2nd ACM Int. Conf. on Embedded Syst. for Energy-Efficient Built Environments*, ser. BuildSys '15. New York, NY, USA: ACM, 2015, pp. 55–64, doi: 10.1145/2821650.2821672.

[36] TensorFlow: An open source machine learning framework. (2018). [Online]. Available: https://www.tensorflow.org/

[37] J. Kelly and W. Knottenbelt, "The UK-DALE dataset, domestic appliance-level electricity demand and whole-house demand from five UK homes," *Scientific Data*, vol. 2, no. 150007, 2015, doi: 10.1038/sdata.2015.7.

[38] Open Energy Monitor. (2018) Open source monitoring for understanding energy. [Online]. Available: https://openenergymonitor.org/

[39] M. R. Asghar, G. Dan, D. Miorandi, and I. Chlamtac, "Smart meter data privacy: A survey," *IEEE Commun. Surveys Tutorials*, vol. 19, no. 4, pp. 2820–2835, Fourthquarter 2017, doi: 10.1109/COMST.2017.2720195.

Michael A. Devlin received the B. Eng. degree in Electronic and Computer Engineering from National University of Ireland Galway, Galway, Ireland in 2018. He was awarded a University Scholarship by the National University of Ireland Galway in 2015, as a result of his outstanding performance in the 2014-15 academic year. He is currently working as a Trainee Engineer at LM Ericsson Ltd., Athlone, Co. Westmeath, Ireland.

Barry P. Hayes (S'09-M'12-SM'17) received the B. Eng. degree in Electrical Engineering from University College Cork, Cork, Ireland in 2005, the M. Eng. from the National University of Ireland, Maynooth, Ireland in 2008, and the Ph.D degree from the University of Edinburgh, Edinburgh, UK in 2013. He was a Marie Sklodowska-Curie Postdoctoral Research Fellow at IMDEA Energy Institute in Madrid, Spain from 2013-2016, and worked as a Lecturer (Assistant Professor) at National University of Ireland Galway, Galway, Ireland from 2016-2018. He has also held Visiting Researcher positions at National Grid UK and at the University of Tennessee, Knoxville, USA. Currently, he is a Lecturer (Assistant Professor) in Electrical Power Systems at University College Cork, Cork, Ireland. His research interests include the network integration of renewable energy sources, and the operation and planning of future power systems.