Non-intrusive load monitoring and its challenges in a NILM system framework

Qi Liu*

Jiangsu Collaborative Innovation Center of Atmospheric Environment
and Equipment Technology (CICAEET),
Nanjing University of Information Science and Technology,
Nanjing, 210044, China
and
School of Computing,
Edinburgh Napier University,
Edinburgh, Scotland, UK
Email: q.liu@napier.ac.uk
*Corresponding author

Min Lu

School of Computer and Software,
Nanjing University of Information Science and Technology,
Nanjing, China
Email: bmbboy@163.com

Xiaodong Liu

School of Computing,
Edinburgh Napier University,
10 Colinton Road, Edinburgh EH10 5DT, UK
Email: x.liu@napier.ac.uk

Nigel Linge

The University of Salford,
Salford, Greater Manchester, M5 4WT, UK
Email: n.linge@salford.ac.uk

Abstract: With the increasing of energy demand and electricity price, researchers gain more and more interest among the residential load monitoring. In order to feed back the individual appliance's energy consumption instead of the whole-house energy consumption, non-intrusive load monitoring (NILM) is a good choice for residents to respond the time-of-use price and achieve electricity saving. In this paper, we discuss the system framework of NILM and analyse the challenges in every module. Besides, we study and compare the public datasets and accuracy metrics of non-intrusive load monitoring techniques.

Keywords: non-intrusive load monitoring; NILM; data acquisition; event detection; feature extraction; load disaggregation.

Biographical notes: Qi Liu received his BSc degree in Computer Science and Technology from Zhuzhou Institute of Technology, China in 2003, and his MSc and PhD in Data Telecommunications and Networks from the University of Salford, UK in 2006 and 2010. His research interests include context awareness, data communication in MANET and WSN, and smart grid. His recent research work focuses on intelligent agriculture and meteorological observation systems based on WSN.

Min Lu received his Bachelor's degree in Computer Science and Technology from Nanjing University of Information, Science and Technology in 2014. He is currently pursuing his Master's degree in Computer Science and Technology at the Nanjing University of Information Science and Technology. His research interests include the algorithms of non-intrusive load monitoring and disaggregation.

Xiaodong Liu is a Reader and the Director of Centre for Information and Software Systems in School of Computing at Edinburgh Napier University. His research interests include context-aware adaptive services, service evolution, mobile clouds, pervasive computing, software reuse, and green software engineering. He is a member of IEEE Computer Society and British Computer Society.

Nigel Linge received his BSc degree in Electronics from the University of Salford, UK in 1983, and PhD in Computer Networks from the University of Salford, UK in 1987. He was promoted to Professor of Telecommunications at the University of Salford, UK in 1997. His research interests include location based and context aware information systems, protocols, mobile systems and applications of networking technology in areas such as energy and building monitoring.

1   Introduction

Nowadays, with the rapid urbanisation and greenhouse effect, people put more attention on energy saving and environment protection. Related statistics show that residential energy account for almost 30% of the total carbon dioxide in the UK and the figure can achieve 10% reduction by taking some simple energy efficiency measures (Sundramoorthy et al., 2011). Domestic energy consumption makes up over one fifth of the total energy use in the USA and over 40% of this power is wasted (Alahmad et al., 2012). Apart from administrative interventions and energy management (Wang et al., 2016), statistics show direct feedback methods (i.e., real-time energy consumption information of appliance-specific) can achieve maximum energy saving instead of indirect feedback methods (i.e., monthly bills and irregular energy usage suggestions) (Ehrhardt-Martinez et al., 2010). Motivated by this, appliance load monitoring (ALM) has been put forward to reach the goal of energy conservation and emission reduction depends on the use of internet of things (IoT) technologies (Talpur et al., 2015). ALM not only can provide useful feedback to the residents, but also be qualified in fault detection for industry. This can be achieved by two major approaches:

- Intrusive load monitoring – ILM requires individual device and appliance to be installed a sensor with digital communication function to acquire energy usage, then the local area network takes charge of gathering and sending electricity consumption information (Rid et al., 2014).

- Non-intrusive load monitoring – NILM was first proposed by George Hart in the 1980s (Hart, 1992), which only needs to set one sensor to gather aggregated energy information of the total load at the house entry point. Then the raw current and voltage data will be analysed to estimate the appliances that are turned on.

Although ILM have potential high accuracy, the hardware cost and difficulty of implementation will relevantly increase (Froehlich et al., 2011). Due to the existence of multi-sensors, some reliability problems may occur if any sensor does not work, which could lead to a system failure (Laughman et al., 2003). In addition, ILM is not scalable and has poor user acceptance. On the other hand, NILM approaches need no more additional devices and can be easily accepted by consumers because of its convenience and economic efficiency.

The aim of NILM is to disaggregate the whole-house energy consumption data into the information of working appliances that contribute to it. This problem can be stated as follows: The aggregated power signals at the entry point of the residence as $P(t)$ and NILM methods decode the overall data into all kinds of components $P_i(t)$ that are attributed to individual appliances n, which can be mathematically defined as

$$P(t) = p_1(t) + p_2(t) + \cdots + p_n(t) \tag{1}$$

Non-intrusive load monitoring system is analysed in this paper, and we sum up a formal system framework, which can apply to existing methods of NILM through others' works. We also discuss the potential challenges of every module in NILM system framework. Besides, we study different public datasets and tell the difference between them. Some evaluation criterions are proposed to test accuracy and feasibility of diverse load disaggregation algorithms. Finally, we summarise the reasons why NILM can not be put into commercial so far.

In this paper, we explain the resources and algorithms of non-intrusive load monitoring. In Section 2, we introduce some basic concepts of NILM methods. In Section 3, we describe the system framework of NILM approaches. In Section 4, the public datasets are listed and compared. In Section 5, we introduce the accuracy metrics of event detection and load disaggregation algorithms. Challenges and conclusions are finally presented.

2   Concepts and definitions in non-intrusive load monitoring methods

2.1   Load signature

Load signature means reliable and unique load feature, which represents the significant electrical behaviour when

individual appliance is working. Load signature is the amount that can distinguish the operating state and temporal behaviour of appliances. Since every appliance has its own internal structure, working pattern and working environment, load signature is highly contributing to the identification of different appliances, and it is one of the most fundamental elements of the energy disaggregation problems.

2.2 Categories of load signature

In Liang et al. (2010), the authors have divided two forms of load signature. The first one is called snapshot form. In this form, the signature is shown by transient snapshot of appliances' electric behaviour at any fixed time intervals. This form usually contains more than one appliance's operating behaviour simultaneously, which refers to the composite load. The second one is named delta form. This form tells the difference between two sequential snapshot form load signatures. If the time interval is small enough, we can regard the delta form signature as a single appliance's load behaviour more likely than composite load.

2.3 Categories of consumer appliances

The goal of NILM methods is to identify individual working appliance and determine their operating states and corresponding energy consumption. But it is unnecessary to track small devices like phone chargers because such devices consume too less power compared with other appliances in a typical home. Moreover, the load signatures of small devices can be covered by large appliances. The types of appliances which NILM approaches are intended to disaggregate can be classified as follows (Zoha et al., 2012):

- Type-I: On-off appliances. These are appliances that only have two states of operation (ON/OFF), such as table lamp, electric kettle, etc.

- Type-II: Finite state machines (FSM). These refer to the appliances with a limited number of operating states, including washing machine, electric fan, etc. Due to the certainty of the number of operating states and the repeatability of switching pattern, it reduces the difficulty of multi-state appliances' identification.

- Type-III: Continuously variable devices (CAD). This type of appliances has no fixed power draw and no obvious switching signs when it changes the states. Thus it becomes an obstacle to load disaggregation algorithm.

- Type-IV: Permanent consumer devices. In Zeifman and Roth (2011), the authors present the type of appliances which remains active and has approximately constant power draw in a time period. The devices that belong to this type are hardwired smoke detectors and cable TV receivers.

Figure 1 Power curves of three types of loads
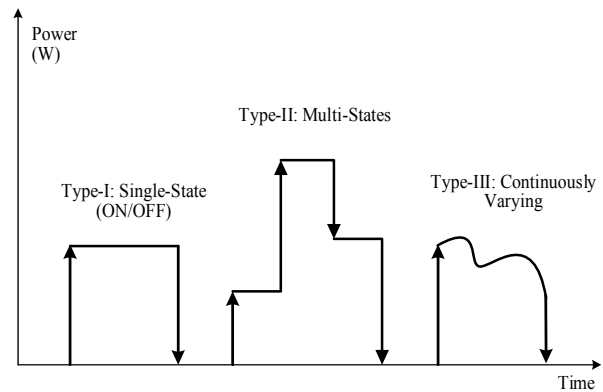
3 System framework of NILM and technical challenges

The general framework of NILM methods can be broadly separated into six modules, including data acquisition module, data processing module, event detection module, feature extraction module, load disaggregation module and application module. Figure 2 shows the logic block diagram of NILM. Above all, event detection module, feature extraction module and load disaggregation module are three key factors that greatly affect the accuracy of the load disaggregation monitoring. The following section introduces the structure and function of every step of the framework in detail.

Figure 2 General logic block diagram of NILM
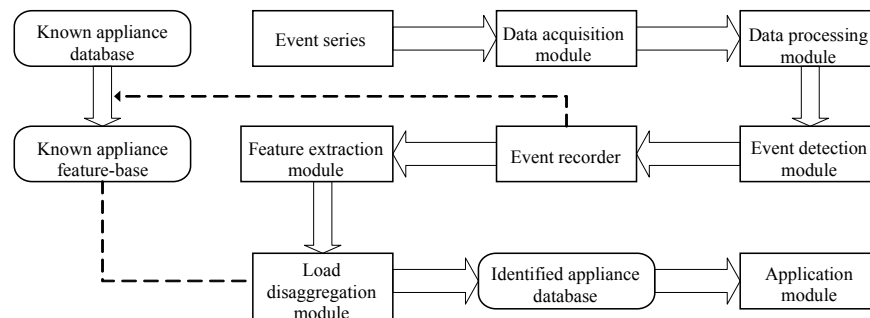
Source: Liang et al. (2010)

3.1   Data acquisition module

Data acquisition module is used to get the measurement of aggregated load for follow-up work. Now, there are various available commercial power meters on the market. In general, the sampling rate of power meters can be classified as low-frequency meters and high-frequency meters, which decides the features that can be extracted from the acquired data. Usually, the power meters measure three values: voltage, current and power factor.

Low-frequency meters: Its aim to capture steady-state features. According to the Nyquist-Shannon sampling theorem, if sampling rate is two times more than the highest frequency of the electrical signal, the digital signal can remain the information of original signal completely. In order to capture the 5th harmonic of the signals, if the fundamental frequency is 60 Hz, the sampling rate should be at least 600 Hz.

High-frequency meters: Its aim to capture transient state and collect fine-grained data to get more unique load features. Since the sampling rate needs to reach a range of 10 to 100 MHz to record waveforms, high resolution power meters usually need to be custom-made and high cost.

Challenges:

1   for the same device, different kinds of power meters could gather different data, which leads to the mismatching of whole-house data and the sum of circuit-level data

2   the colour noise that generated by variable speed devices and white noise that generated by permanent consumer devices will reduce the accuracy of raw data

3   data compression in the power meters will cause the loss of raw data as well.

3.2   Data processing module

Data processing module is used to adjust and process the gathered data, it usually contains three steps:

• Resampling, it requires aligning the current signal with the voltage signal to compute their proper phase relationship. With the phase relationship, we can compute the power factor and other features.

• Quantising, with the initial data standardised, we can calculate useful features, such as reactive power, spectral envelop, higher harmonics and so on.

• Extracting features, we can get specific features using filter bank and down-samples. The task of filtering is to get a good approximate of original waveform and minimise the loss of information. Down-sampling can help us get the appropriate resolution data for further use.

Challenges: Since the information we need is not computed from raw data but from the down-sampled data, the loss of raw data is unavoidable and the amount depends on the resolution that we choose.

3.3   Event detection module

Event detection module is used to decide whether an actual appliance's state switching event occurs by analysing the total power-level changes. The reason for taking this step is to filter out useful information from all captured data because too much indiscriminate data will affect the computational efficiency and data storage. The methods to detect events are major divided into two kinds: using edge detection on the aggregated power curve and probability-based approaches. The detected event can be one of the following types:

1   the changes of one or more devices' operating states

2   the power curve fluctuation caused by noises

3   one part of the working appliance which does not change its operating mode.

Event-based methods usually calculate the difference of continuous values in a sample and compare it to the predefined threshold. In Azzini et al. (2014), the authors present two event detection algorithms to run over the whole house power consumption curve, including the window with margins method and the shifted sample method. The window with margins method uses two different windows and calculates the averages value over the active power consumption curve only with the initial and final samples. Figure 3 illustrates the window with margins method, this method contains four parameters for users to choose: width of the primary window, width of the second window, the number of samples on each margin and threshold of a possible event. The shifted sample method based on the derivative of power also uses two continuous windows to calculate the difference between the averages of the primal and final halves. If the valve is greater than the threshold, the event will be detected and recorded. Compared with the former method, the shifted sample method can reduce computational complexity but with lower detection accuracy. Some methods compare two windows of samples using a probabilistic way to judge whether an event occurs. In Wang and Zheng (2012), the authors record fast switching events as a triangle and steady working events as a rectangle. Figure 4 illustrates the two unit graphics with combined triangles and rectangles in the signal. Rectangle including the data items which are starttime, peaktime, peakvalue, steadytime and steadyvalue. Triangle contains the data items which are starttime, peaktime, endtime and peakvalue. With these data items, the basic unit can be expressed well and since triangles and rectangles cannot cover each other, it will reduce the overlaps of similar features and improve the success rate of event detection. Considering the edge happens as an average distribution, the Poisson probabilistic model is used to calculate the probability of overlaps of edges. In recent year, non-event-based methods have been proposed which do not depend on the edge detection. The examples can refer to hidden Markov model, which take every sample of total power into account for classification and inference.

Figure 3 Diagram of the window with margins method
(see online version for colours)
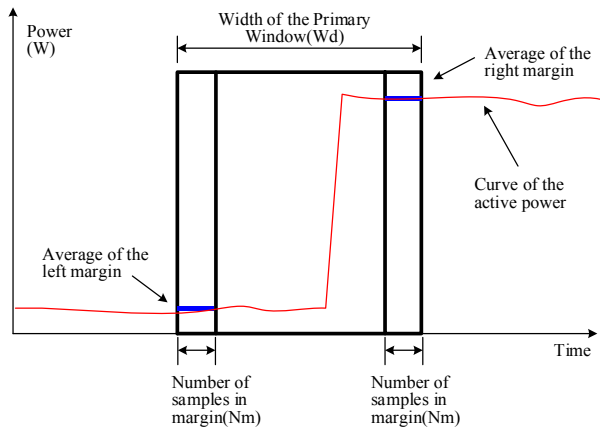
Source: Azzini et al. (2014)

Figure 4 Diagram of the two unit graphics with combined
triangles and rectangles in signal
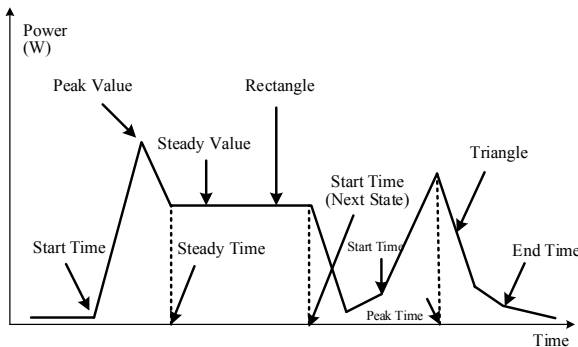
Source: Wang and Zheng (2012)

Challenges: Between the event-based methods and non-event-based methods, the former is more computationally efficient because not the whole data needs to be calculated and estimated. But due to the threshold of edge, if the value is too small, the rate of false detection may increase; if the value is too big, the rate of miss detection may raise up.

3.4 Feature extraction module

Feature extraction module is used to capture features around the event points. The features can be divided into two types according to the sampling frequency: steady-state features and transient-state features. We discuss the advantages and disadvantages of each kind of features below.

3.4.1 Steady-state features

- Power change. Choosing the variation of real power and reactive power as features can easily identify the high-power electrical appliances, but it only works well with on-off devices. Because of the overlapping features in P-Q plane, the performance in identifying type-II, type-III and type-IV appliances is poor.

- Time-frequency analysis of V-I waveform. It needs higher sampling rate to get steady-state harmonics. According to the harmonics, we can easily tell the difference between resistive, inductive and electronic loads. But the accuracy for type-III appliances is low and it is difficult to identify the events that happen at the same time.

- V-I trajectory. It mainly analyses the shape features of V-I trajectory such as looping direction, area enclosed, number of self-intersections and so on (Hassan et al., 2013) It enables the appliances can be distinguished in a detailed way. But it needs complex computational work and the devices with small power consumption have no unique trajectory features.

- Steady-State voltage noise. Due to the EMI features, motor-based appliances and devices with switching mode power supply (SMPS) can be recognised. It is also able to detect the simultaneous activation events. The shortcomings of this feature are not every appliance has SMPS and it is sensitive to signal noise.

3.4.2 Transient-state features

- Transient power. The advantages of using spectral envelopes are feasibility to recognise type-I, type-II and type-III appliances, even including the devices with same power characteristics. But it requires continuous monitoring at a high sampling frequency.

- Start-up current. The current spikes, size and duration can help us distinguish the appliances with multi-states and the accuracy is acceptable. But in the situation of simultaneous activation, the accuracy can be poor. It also does not support type-III and type-IV appliances.

- Transient-state voltage noise. According to the noise fast Fourier transform (FFT), multi-states appliances and devices with SMPS can be easily distinguished. But it needs complexity and large calculating quantity.

3.5 Load disaggregation module

Load disaggregation module is used to do the classification by using the extracted features. Current NILM algorithms can broadly be divided into the supervised learning and unsupervised learning based on whether the approaches using the labelled datasets for training the classifier. The supervised learning can further be divided into optimisation and pattern recognition approaches. Since the training of labelled data is a time-consuming job which increases the cost and human effort, researchers are now looking forward to seeking a solution based on completely unsupervised or semi-supervised methods. We discuss the supervised learning and unsupervised learning approaches below.

3.5.1 Supervised learning

- Optimisation methods. Optimisation methods try to treat the problem of load disaggregation as an

optimisation problem. Once an appliance's event has been detected, this approach extracts feature vector of target event and compare it with the feature vector of known appliance' event stored in the signature database to find the closest possible match. The idea of optimisation is quite simple, but it becomes more and more complicated when taking the combination of appliances into account. Besides, if an unknown appliance without training occurs, the accuracy can be poor. The algorithms belonging to the optimisation methods are integer programming, genetic algorithms and so on.

- Pattern recognition methods. Pattern recognition is one of the most common used approaches by researchers to deal with the problem of load monitoring. This approach classifies the captured features by using machine learning techniques. These algorithms includes artificial neural networks (ANN) (Xuezhi et al., 2015; Gu and Sheng, 2017), naïve Bayes classifier, support vector machines (SVM) (Bin et al., 2015; Li et al., 2016), k-nearest neighbour (kNN) and so on. The shortcomings of these algorithms are the lack of test samples and the overlapping signatures of low power-consuming appliances.

3.5.2   Unsupervised learning

Since the supervised learning approaches need labelled dataset to train classifier and manual labelling is fallible and time-consuming, the unsupervised learning approaches become a hot research content. These approaches base on the mathematical probability models, which including blind source separation, genetic k-means, motif mining, factorial hidden Markov model (FHMM) and its variety. Although these algorithms can reduce human's work, it is computationally expensive because it takes every part of samples into account and the load disaggregated accuracy is not so satisfactory. Besides, it also needs to assume that the number of devices is already known.

3.6   Application module

Since the changes of devices' states have been recognised, application module is used to track the operation pattern and power consumption of individual appliance. Based on this, some personalised and useful suggestions will be provided. Besides, some prediction can also be made to help electric power company know more about the energy demand side.

4   Public dataset

Nowadays, more and more approaches from data mining and machine learning have been used to solve the problem of energy. Although these advanced techniques are presented, it is still hard to put into use and test without the public datasets. The reasons why it is not practical to build their own datasets by every researcher are including:

- Time-consuming, building a dataset needs to take a duration of time to capture and collate data, at least takes a week or much more longer.

- High-costing, not only need to measure the aggregate load at one point, but we also require distributed sensors to get the individual appliance load as ground truth to train and test load disaggregation algorithms. If the features are transient state, high-frequency sampling will greatly increase the cost.

- Hard to compare, just assume that every researcher develops different algorithms on their own datasets, it is actually unable to evaluate the results and judge which method is better.

- Privacy, recording appliance usage and geographical location information may reveal what appliances do occupants have and their usage pattern.

To get a better performance of load disaggregated algorithms, the choice of dataset is also a key factor. Here we list the main reasons for selecting the suitable dataset to evaluate the NILM approaches:

- Feature. Different datasets contain diverse features, such as real power (P) and reactive power (Q) are used to make use of the change of power to judge which appliance is turned on/off, the V-I trajectory is used to make use of the trajectory parameters to uniquely define an appliance's activity, etc. Some external features like weather and climate also benefit the improvement of accuracy.

- Frequency. Some features are in low-frequency and others are in high-frequency, low-frequency features can easily get from the smart meters, but may exist many overlapping in the feature space. High-frequency features can usually distinguish appliances more effective and in a higher precision, but it requires high rate sampling.

- Duration and number of households. Some load disaggregated algorithms need dataset which takes a long period of time to train and get a better prediction of appliance usage, especially periodic consumption behaviours need to be captured. And more number of households is benefit to data statistics for effective consumption pattern research.

- Location. Due to the diversity of electric appliances and usage patterns, electric usage of different countries varies very much. If one algorithm is aimed at load disaggregation for a specific country, it is very necessary to use dataset from that country to get better performance. What's more, different countries may have diverse voltage and frequency, for example 220 V, 50 Hz in China, 110 V, 60 Hz in the USA and 240 V, 50 Hz in the UK

Table 1 The existing public dataset for energy disaggregation research

Dataset	Country	Houses	Sensors (*per house*)	Features	Granularity	Duration
AMPds (Makonin et al., 2013)	Greater Vancouver	1	19	I, V, PF, f, Q, S	1 min	1 year
BLUED (Anderson et al., 2012)	Pittsburg	1	Aggregated	I, V, switch events	12 kHz	8 days
GREEND (Monacchi et al., 2014)	Austria, Italy	8	9	P	1 Hz	1 year
iAWE (Batra et al., 2013)	India	1	33	V, I, f, P, S, E, Φ	1 Hz	73 days
REDD (Kolter and Johnson, 2011)	Boston	6	9–24	Aggregated: V, P; Sub-metered: P	15kHz (aggregated), 3 sec (sub-metered)	3–19 days
Smart* (Barker et al., 2012)	Western Massachusetts	3	25 circuits, 29 appliance monitors	Circuits: P, S; Sub-metered: P	1 Hz	3 months
Tracebase (Reinhardt et al., 2012)	Germany	15	158 appliances in total (43 types)	P	1-10 sec	N/A
UK-DALE (Kelly and Knottenbelt, 2014)	UK	5	house 1:54; house 2-4:5-26	Aggregated: P; Sub-metered: P, switch-states	16 kHz (aggregated), 6 sec (sub-metered)	655 days

Since the reference energy disaggregation dataset (REDD) was publically released by MIT in 2011, which has been widely used in the disaggregation community, the number of datasets from different countries is increasing greatly. We list the public dataset so far in Table 1. It should be noted that the abbreviations of features: current (I), voltage (V), power factor (PF), frequency (f), active power (P), reactive power (Q), apparent power (S), energy (E) and phase angle (Φ).

5 Accuracy metrics

Since researchers use the public datasets to test their load disaggregation algorithms, the accuracy metrics need to be established to evaluate their works. Due to the fluctuation of dynamic loads and the threshold value we set in advance, the type-I error (i.e., no appliance is changed operating state but the event detection module catches an event, which is called false detection) or type-II error (i.e., an appliance is actually turned on or closed but the event detection module misses the event, which is called miss detection) may occur.

In Anderson et al. (2012), the authors present the true positive rate and the true positive percentage to evaluate the performance of the event detection. The true positive rate can be mathematically defined as

$$\text{TPR} = \frac{TP}{TP + FN} \in [0,1] \qquad (2)$$

$$\text{FPR} = \frac{FP}{FP + TN} \in [0,1] \qquad (3)$$

where TP stands for the number of true positives, FP stands for the number of false positives, TN stands for the number of true negatives, and FN stands for the number of false negatives or misses. We can also use receiver operating characteristics (ROC) to make the trade-off between its true positive rate (TPR) and false positive rate (FPR). If the event detector is good enough, the TPR will approach one and the FPR will close to zero.

The true positive percentage compares the percentage of events that the number of true positives and false positives to the ratio of the total number of events E. The true positive percentage can be mathematically defined as

$$\text{TPP} = \frac{TP}{E} \qquad (4)$$

$$\text{FPP} = \frac{FP}{E} \qquad (5)$$

Similar to the rate metric above, the TPP of a good event detector would approach one and the FPP would close to zero.

In research (Liang et al., 2010), the authors propose three different kinds of accuracy metrics. First of all, the total number of detected events is defined as N_{det} in formula (6)

$$N_{\text{det}} = N_{true} + N_{wro} - N_{miss} \qquad (6)$$

where N_{ture} is the true number of events that actually happened, N_{wro} is the number of events that false detected, and N_{miss} is the number of events that the event detector missed. The accuracy measures are proposed as follows.

1   Detection accuracy (η_{det}) refers to the accuracy of disaggregation including the effect of false detected events, which can be mathematically defined as

$$\eta_{\text{det}} = \frac{N_{dis}}{N_{\text{det}}} \qquad (7)$$

2   Disaggregation accuracy (η_{dis}) refers to the accuracy of disaggregation excluding the effect of false detected events, which can be mathematically defined as

$$\eta_{dis} = \frac{N_{dis}}{N_{\text{det}} - N_{wro}} \qquad (8)$$

3   Overall accuracy (η_{all}) refers to the accuracy of disaggregation including the effect of false detected events and missing detected events, which can be mathematically defined as

$$\eta_{all} = \frac{N_{dis}}{N_{tru}} = \frac{N_{dis}}{N_{\text{det}} - N_{wro} + N_{miss}} \qquad (9)$$

Apart from the accuracy metrics of event detection and load disaggregation, the authors in Batra et al. (2014) introduce an open source toolkit for non-intrusive load monitoring which is called NILMTK. This toolkit provides a parser to transform a range of public datasets into the standard data structure which is called NILMTK-DF. The statistical and diagnostic functions can help researchers make a detailed understanding of public dataset. The NILMTK provides two benchmark disaggregation algorithms including combinatorial optimisation (CO) and factorial hidden Markov model (FHMM). Some accuracy metrics are also mentioned in this article including errors in total assigned, EMS errors in assigned power, true positives, false positives, false negatives, true negatives, precision and recall. With this toolkit, researchers can evaluate their load disaggregation algorithms in a more fair and convenient way.

6   Challenges and conclusions

In this paper, we introduce the system framework and its challenges in individual module. Besides, the public datasets are also studied and compared. Although the non-intrusive load monitoring gains a lot of attentions in recent year because of the economic efficiency and convenience, it is still difficult to put the NILM into commercial due to the following reasons:

1   The compatibility of load signatures. Since the factors of different kinds, manufacturers and sizes of appliances will affect the performance of load disaggregation algorithms. Moreover, there are no widely applicable load signatures can model the operations of the four types of appliances well.

2   The comparison of load disaggregation algorithms. As mentioned in Section 4, without the standard and unified public reference dataset, it is quite difficult to fairly compare and test different load disaggregation algorithms. And we look forward to the coming of the testing platform to evaluate the algorithms more easily and equitably.

3   The overlap of load features. Low-power appliances have similar power consumption characteristics and it is difficult to discern them at low-frequency sampling, due to the ambiguous overlapping of steady-state features in the P-Q plane.

4   Manual labelling. For supervised learning approaches, it is quite boring and fallible to turn on/off every appliance in proper order to build a signature database and train the algorithms for classification.

5   The update of signature database. The supervised learning approaches need signature databases to do off-line training. Since any unknown appliance which is not in the appliance signature database appears, the precision of load disaggregation will be poor.

6   Imperfect appliance models. For unsupervised learning approaches, the appliance model generated by the HMM and house power consumption established by the FHMM suffer from non-Gaussian. And since the imperfect manufacturing process and the influence of environment, the precision will decline.

7   Different types of appliances. The precision of load disaggregation for on/off appliances is quite high (more than 90%) so far. But to multi-state appliances, continuously variable devices and rarely used devices, the result is not so satisfactory. And the simultaneous switch events of appliances (like PC and printer etc.) make the disaggregation of load more complex.

8   The security of data transmission. Because most of the data transmission approaches are wireless, thieves may analyse the presence of house owner by counting the number of packets. So it is necessary to reinforce the security of the wireless communications (Jian et al., 2015; Tinghuai et al., 2015).

9   The robustness of algorithms. The NILM approach should be scalable in the sense because the number of used appliance of a typical family can up to 20–30 (Zeifman, 2012).

Although NILM needs to be improved and upgraded before it can be widely spread, the advantages of NILM should be recognised. Compared with ILM, NILM costs less money and time to be implemented and maintained. On the other hand, NILM approaches need no more additional devices and can be easily accepted by consumers because of its convenience and economic efficiency. Since more and more

algorithms are proposed to increase the accuracy of NILM, NILM has a lot of potentials.

References

Alahmad, M.A., Wheeler, P.G., Schwer, A., Eiden, J. and Brumbaugh, A. (2012) 'A comparative study of three feedback devices for residential real-time energy monitoring', *IEEE Transactions on Industrial Electronics*, Vol. 59, No. 4, pp.2002–2013.

Anderson, K., Ocneanu, A., Benitez, D., Carlson, D., Rowe, A. and Berges, M. (2012) 'BLUED: a fully labeled public dataset for event-based non-intrusive load monitoring research', in *Proceedings of the 2nd KDD Workshop on Data Mining Applications in Sustainability (SustKDD)*, pp.1–5.

Anderson, K.D., Berges, M.E., Ocneanu, A., Benitez, D. and Moura, J.M.F. (2012) 'Event detection for non intrusive load monitoring', *Proceedings of the 38th Annual Conference on IEEE Industrial Electronics Society*, Montreal, QC, Canada, pp.3312–3317.

Azzini, H.A.D., Torquato, R. and Silva, L.C.P.D. (2014) 'Event detection methods for nonintrusive load monitoring', *Pes General Meeting | Conference & Exposition*, IEEE.

Barker, S., Mishra, A., Irwin, D., Cecchet, E., Shenoy, P. and Albrecht, J. (2012) 'Smart*: An open data set and tools for enabling research in sustainable homes', in *Proceedings of the 1st KDD Workshop on Data Mining Applications in Sustainability (SustKDD)*, San Diego, California, USA, August.

Batra, N., Gulati, M., Singh, A. and Srivastava, M.B. (2013) 'It's different: insights into home energy consumption in India', in *Proceedings of the 5th ACM Workshop on Embedded Systems For Energy-Efficient Buildings*, pp.1–8.

Batra, N., Kelly, J., Parson, O., Dutta, H., Knottenbelt, W., Rogers, A., Singh, A. and Srivastava, M. (2014) 'NILMTK: an open source toolkit for non-intrusive load monitoring', in *Proceedings of the 5th International Conference on Future Energy Systems*, pp.265–276, Cambridge, UK.

Bin, G., Victor, S.S., Keng, Y.T., Walter, R. and Shuo, L. (2015) 'Incremental support vector learning for ordinal regression', *IEEE Transactions on Neural Networks and Learning Systems*, Vol. 26, No. 7, pp.1403–1416.

Ehrhardt-Martinez, K., Donnelly, K.A. and Laitner, S. (2010) 'Advanced metering initiatives and residential feedback programs: a meta-review for household electricity-saving opportunities', *American Council for an Energy-Efficient Economy*, Washington, DC, USA, June.

Froehlich, J., Larson, E., Gupta, S., Cohn, G., Reynolds, M.S. and Patel, S.N. (2011) 'Disaggregated end-use energy sensing for the smart grid', *IEEE Pervasive Computing*, Vol. 10, No. 1, pp.28–39.

Gu, B. and Sheng, V.S. (2017) 'A robust regularization path algorithm for ν-support vector classification', *IEEE Transactions on Neural Networks and Learning Systems*, Vol. 28, No. 5, pp.1241–1248.

Hart, G.W. (1992) 'Nonintrusive appliance load monitoring', *Proceedings of the IEEE*, Vol. 80, No. 12, pp.1870–1891.

Hassan, T., Javed, F. and Arshad, N. (2013) 'An empirical investigation of v-i trajectory based load signatures for non-intrusive load monitoring', *IEEE Transactions on Smart Grid*, Vol. 5, No. 2, pp.870–878.

Jian, S., Haowen, T., Jin, W., Jinwei, W. and Sungyoung, L. (2015) 'A novel routing protocol providing good transmission reliability in underwater sensor networks', *Journal of Internet Technology*, Vol. 16, No. 1, pp.171–178.

Kelly, J. and Knottenbelt, W. (2014) 'UK-DALE: a dataset recording UK domestic appliance-level electricity demand and whole-house demand', ArXiv, April.

Kolter, J.Z. and Johnson, M.J. (2011) 'REDD: a public data set for energy disaggregation research', in *Proceedings of the SustKDD Workshop on Data Mining Applications in Sustainability (SustKDD)*, San Diego, CA, USA, August.

Laughman, C., Lee, D., Cox, R., Shaw, S., Leeb, S., Norford, L. and Armstrong, P. (2003) 'Power signature analysis', *IEEE Power and Energy Magazine*, Vol. 1, No. 2, pp.56–63.

Li, B., Ding, L. and Li, J. (2016) 'Differential evolution-based parameters optimisation and feature selection for support vector machine', *International Journal of Computational Science and Engineering*, Vol. 13, No. 4, pp.355–363.

Liang, J., Ng, S.K.K., Kendall, G. and Cheng, J.W.M. (2010) 'Load signature study – Part I: basic concept, structure, and methodology', *IEEE Transactions on Power Delivery*, Vol. 25, No. 2, pp.551–560.

Makonin, S., Popowich, F., Bartram, L., Gill, B. and Bajic, I. V. (2013) 'AMPds: a public dataset for load disaggregation and eco-feedback research', in *Proceeding of 2013 IEEE Conference on Electrical Power and Energy (EPEC)*, pp.1–6.

Monacchi, A., Egarter, D., Elmenreich, W., D'Alessandro, S. and Tonello, A.M. (2014) 'GREEND: an energy consumption dataset of households in Italy and Austria', in *Proceedings of 2014 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, pp.511–516.

Reinhardt, A., Baumann, P., Burgstahler, D., Hollick, M., Chonov, H., Werner, M. and Steinmetz, R. (2012) 'On the accuracy of appliance identification based on distributed load metering data', in *Proceedings of the 2nd IFIP Conference on Sustainable Internet and ICT for Sustainability (SustainIT)*, October.

Ridi, A., Gisler, C. and Hennebert, J. (2014) 'A survey on intrusive load monitoring for appliance recognition', *Biochemical & Biophysical Research Communications*, Vol. 94, No. 4, pp.3702–3707.

Sundramoorthy, V., Cooper, G., Linge, N. and Liu, Q. (2011) 'Domesticating energy-monitoring systems: challenges and design concerns', *IEEE Pervasive Computing*, Vol. 10, No. 1, pp.20–27.

Talpur, M.S.H., Bhuiyan, M.Z.A. and Wang, G. (2015) 'Energy-efficient healthcare monitoring with smartphones and IoT technologies', *International Journal of High Performance Computing & Networking*, Vol. 8, No. 2, pp.186–194.

Tinghuai, M., Jinjuan, Z., Meili, T., Yuan, T., Abdullah, A., Mznah, A. and Sungyoung, L. (2015) 'Social network and tag sources based augmenting collaborative recommender system', *IEICE Transactions on Information and Systems*, Vol. 98, No. 4, pp.902–910.

Wang, X., Wu, C. and Nathwani, J. (2016) 'Dynamic response forecasting and optimising paths of energy efficiency', *International Journal of Computational Science & Engineering*, Vol. 12, No. 1, pp.86–93.

Wang, Z. and Zheng, G. (2012) 'Residential appliances identification and monitoring by a nonintrusive method', *IEEE Transactions on Smart Grid*, Vol. 3, No. 1, pp.80–92.

Xuezhi, W., Ling, S., Yu, X. and Wei, F. (2015) 'A rapid learning algorithm for vehicle classification', *Information Sciences*, Vol. 295, No. 1, pp.395–406.

Zeifman, M. (2012) 'Disaggregation of home energy display data using probabilistic approach', *IEEE Transactions on Consumer Electronics*, Vol. 58, No. 1, pp.23–31.

Zeifman, M. and Roth, K. (2011) 'Nonintrusive appliance load monitoring: Review and outlook', *IEEE Transactions on Consumer Electronics*, Vol. 57, No. 1, pp.76–84.

Zoha, A., Gluhak, A., Imran, M.A. and Rajasegarar, S. (2012) 'Nonintrusive load monitoring approaches for disaggregated energy sensing: A survey', *Sensors*, Vol. 12, No. 12, pp.16838–16866.