Sensor-aided NILM with Gaussian Mixture Models

1st Nidhal Balti
IRISA, Granit team
Univ Rennes, IRISA, CNRS
Lannion, France
nidhal.balti@irisa.fr

2nd Baptiste Vrigneau
IRISA, Granit team
Univ Rennes, IRISA, CNRS
Lannion, France
baptiste.vrigneau@irisa.fr

3rd Pascal Scalart
IRISA, Granit team
Univ Rennes, IRISA, CNRS
Lannion, France
pascal.scalart@irisa.fr

***Abstract*—Energy disaggregation, also known as Non-Intrusive Load Monitoring (NILM), is the process of analyzing energy consumption in a building and identifying individual appliance-level energy usage. This approach can provide valuable insights into energy consumption patterns and help reduce overall energy usage, costs, and carbon emissions. This paper proposes a new method for tackling the disaggregation problem by using data from low-cost wireless sensor networks. The proposed approach estimates appliance states using a GMM model and uses these states as features to improve energy disaggregation. The performance of the proposed method was evaluated on a real-world dataset called SmartSense deployed in our lab, and the results showed that it significantly improved the accuracy of conventional NILM disaggregation performance.**

***Index Terms*—Energy disaggregation, NILM, GMM, Window-GRU, wireless sensor platform SmartSense**

I. Introduction

The current global energy consumption is estimated to be 29,000 TWh by the International Energy Agency [1], and it is projected to increase to 42,000 TWh in 2040, with an annual growth rate of approximately 2.1%. This substantial rise in energy usage makes the issue of energy conservation increasingly challenging. It has a severe impact on both a country's economy and the environment. Therefore, promoting efficient energy usage is the most practical approach to conserve energy at present. By accurately monitoring energy consumption and communicating this information to consumers, energy waste can be significantly reduced [2], [3]. One solution to aid in this reduction is the disaggregation of energy consumption.

Energy disaggregation can be done using two main methods: Intrusive Load Monitoring (ILM) and Non-Intrusive Load Monitoring (NILM). Intrusive load monitoring involves measuring the power consumption of one or more appliances using a low-end metering device, typically requiring at least one monitoring device per appliance. In contrast, Non-intrusive load monitoring only requires a single measure for the entire home or building being monitored.

The novelty of this paper consists in an intermediate solution combining classical energy consumption data with wireless sensors to form an aided NILM or semi-intrusive load monitoring (SILM). Our lab designed a Smart Building platform, called SmartSense, which collects data from individual and global power meter and from various low-cost wireless sensors like air quality, luminosity and temperature. In this study, we provide a novel dataset that includes not only individual and global power data but also data from numerous sensors that will provide *a-priori* knowledge of the load state.

The paper is organized as follows. In Section II, we remind the state of the art and the work environment. Section III introduces our workflow process associating wireless sensors with NILM and our new platform SmartSense. Finally, we describe and analyze preliminary results.

The authors thanks Lannion Trégor Communauté and Brittany region for their Phd founding. This work is supported by the EEWOK platform.

II. Related works

The concept of Non-Intrusive Load Monitoring (NILM), where energy meter data from the entry point could be used to disaggregate loads, was first introduced by Hart in the 1980s [4]. With the rise of smart meters, there has been a surge in publications on the subject of NILM [5], [6]. The ultimate aim is to analyze each appliance's state and power consumption using only a single reading of the total power usage. This is why a solved NILM problem is frequently referred to as "disaggregated" [7]. Figure 1 explains the basic NILM concept. For instance, from $t = 12$ to $t = 28$ minutes, an oven element is turned on. At $t = 17$, a stove burner is turned on. The first issue is identifying the precise moment when these appliances are turned on and off using their electrical signature, which can take the form of activation height, length, or shapes that are depicted with arrows. Estimating their power loads based on this trace is a second issue.

In literature, several techniques are commonly implemented

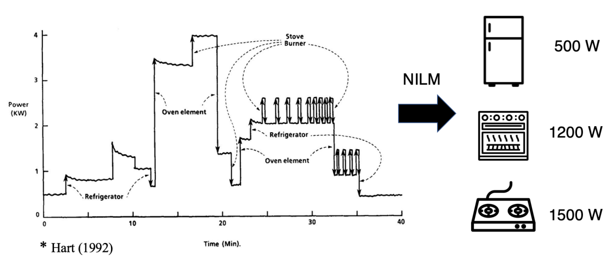


Fig. 1: Concept of typical NILM as presented in [4].

for NILM approach. For instance, signal processing techniques such as Fourier analysis, wavelet transforms, and filtering are used to extract features from the aggregated data [8]. Machine learning algorithms such as Support Vector Machines (SVMs) [9], and Artificial Neural Networks (ANNs) [10] are employed to classify the extracted features and identify the corresponding appliances. Probabilistic models such as Hidden Markov Models (HMMs) [11] [12] and Bayesian Networks [13] are used to model the statistical dependencies between the appliances and improve the accuracy of the disaggregation results. Lastly, some studies have also explored the user behavior patterns [14] to enhance the performance of NILM systems.

In order to design, test and evaluate energy-disaggregation algorithms, free-access energy consumption datasets are crucial for NILM researchers. A NILM dataset is a set of electrical energy measurements captured from real-life scenarios, without perturbing daily routines in the monitored space, hence, keeping the data as close as possible to reality [15]. These datasets include high-resolution measurements of the aggregate electricity consumption of a building, along with detailed information about the individual appliances and their energy consumption patterns. Moreover, they can be categorized according to their high/low sampling frequencies varying from 1Hz to 16kHz. Some of the datasets that have been widely used to evaluate the performance of various NILM algorithms during the last decade are REDD [16], BLUED [17] and UKDALE [18].

Studies considering NILM techniques use metrics to evaluate the performance of their algorithms. Unfortunately, there is no agreement on the measures that should to be used to assess a NILM algorithm's effectiveness. the most used metrics found in the studied literature [18] [19] are disaggregation accuracy, precision, recall, F1-score, the Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE). The two last metrics (MAE and RMSE) correspond to the l2 and l1-norms of thr losses and are defined by:

$$RMSE = \sqrt{\frac{1}{T}\sum_{t=1}^{T}|\hat{y}_t - y_t|^2} \text{ and } MAE = \frac{1}{T}\sum_{t=1}^{T}|\hat{y}_t - y_t| \quad (1)$$

where $\hat{y}_i$ represents the prediction of an appliance's energy usage at time t and y_i represents the corresponding ground truth.

For many years, it was almost impossible to find a way to compare literature findings and experiments in NILM. No previous standard existed in terms of experiment setup, data acquiring and data format. Therefore, an open source toolkit written in python called Non-Intrusive Load Monitoring Toolkit or NILMTK [19] was designed to tackle these issues. It is designed to be a standard tool for NILM tasks, and to be used among researchers. With NILMTK, researchers can have some guidelines on how obtained data and predicted data to be collected, stored, compared, evaluated and even represented in the similar manner or format. Thus, the results from disaggregation can be compared and discussed on the performance of an algorithm used.

III. Semi-Intrusive Load Monitoring

A. *SmartSense dataset*

Using environmental sensing and diverse information can help overcome challenges faced by current NILM techniques [20]. However, this can increase complexity. Hence, a more efficient approach could be to merge supplementary data with the overall power consumption of the building for load monitoring. Furthermore, the increasing number of intelligent sensors integrated into buildings for different purposes has made it more feasible to acquire this information without additional setup, which may explain the recent emphasis on this method. As far as we are aware, no prior research has used environmental sensors in the NILM area.

The dataset used in this study is called SmartSense, which serves as a research platform that employs a sensor network designed for multiple data acquisition purposes. More than 120 nodes are deployed inside our lab and each node within the network comprises fifteen sensors that collect a diverse range of information (see Figure 2), including:

- Video sensors (Video-Graphic-Array (VGA), Infra-Red (IR) cameras) and audio sensors (4 microphones in tetrahedron authorizing localization);
- Radio sensors (2.4 GHz, Sub-GHz, and Ultra-Wide-Band) to sense the radio-frequency band occupancy and to estimate positions;
- Air quality sensors (temperature, humidity, carbon dioxide concentration, air pressure, etc.);
- Light sensors (Ultraviolet and Red-Green-Blue-White);
- Distance sensor (Laser telemeter).

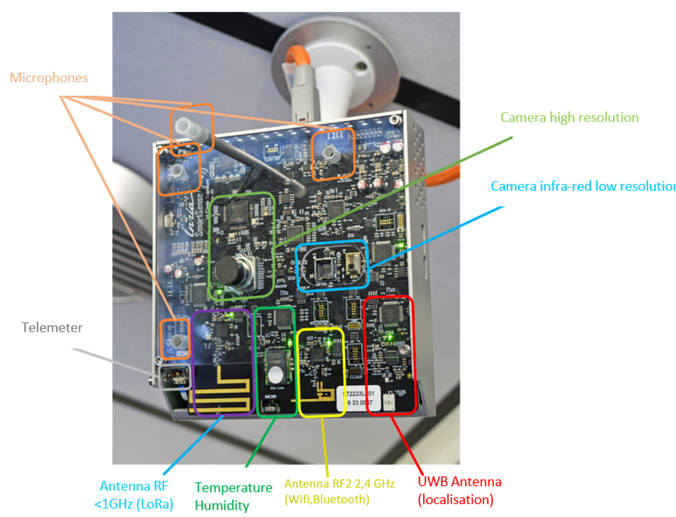


Fig. 2: Different sensors in one node.

Additionally, SmartSense platform contains a record of power consumption in each floor in our lab as a classical NILM (figure 3):

- global power consumption recorded by an EcoCompteur with a frequency of 7Hz;
- individual power consumption of various electrical appliances recorded by a ZigBee meter with a frequency of 100mHz.

A summary of SmartSense features is shown in Table I.

(a) EcoCompteur Legrand module (b) ZigBee Cleode meter

Fig. 3: Hardware devices used to collect power consumption data for SmartSense data

TABLE I: SmartSense dataset features

Measurement device	Data type	Frequency	Appliances
EcoCompteur	global power consumption	7.5 Hz	coffee maker, kettle, fridge, oven, microwave, laptop, printer, light
ZigBee meter	individual power consumption	100 mHz	
Sensors node	sensors data	1 Hz	

B. Proposed workflow

This paper proposes a Semi-Intrusive Load Monitoring architecture (SILM) that uses power consumptions and wireless sensors data to detect individual power consumption of devices from aggregated power data as shown in Figure 4a.

In fact, the SILM architecture is composed mainly of 2 parts:

- The first part called the features extraction phase intends to extract a useful information from sensors data and then feed them to the input of the next step;
- The second part called the disaggregation phase in which the standard NILM workflow is kept. It entails three main phases, including data preprocessing, disaggregation, and performance evaluation.

The aim of our work is to find a good way in exploiting sensors data and extracting a useful information that will help in improving NILM performance. The first naive idea we started with is to add raw sensors data to the aggregated power data as input and let the disaggregator algorithm deal with it. The results show that disaggregation performance compared with the typical NILM improved slightly for the kettle while they stayed almost the same for the coffee maker.
We recognize that raw sensors data could have noise that could limit the disaggreator algorithm's effectiveness. Consequently, it is essential to extract features, such as appliance states, from this data. The disaggregation process involves identifying when an appliance is being used and determining its state. We believe that by focusing on classifying appliances as either `ON` or `OFF`, we can further use this information to enhance the accuracy of the disaggregation process. As a result, we chose to implement for the states estimation task a Gaussian mixture models (GMM) since it showed a good performance in data clustering and classification [21].

C. Gaussian Mixture Models as an estimator

Gaussian mixture models (GMM) are a probabilistic machine learning technique that can provide better approximation when the clusters are overlapping, as compared to k-means clustering [22]. Thus, the GMM model will be used in our case as the estimator/classifier of appliances states as shown in Figure 4b. In Gaussian mixture models, each cluster is modeled as a multivariate Gaussian distribution with the probability density function defined as:

$$f(\boldsymbol{x}) = \frac{1}{\sqrt{(2\pi)^n|\boldsymbol{\Sigma}|}} \exp\left(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right)$$

where $\boldsymbol{x} \in \mathbb{R}^n$ is the random vector, $\boldsymbol{\mu} \in \mathbb{R}^n$ is the mean vector, $\boldsymbol{\Sigma}$ is the $n \times n$ positive-definite covariance matrix, and $|\boldsymbol{\Sigma}|$ is the determinant of $\boldsymbol{\Sigma}$.

If a given appliance can be described by K different states, then for a mixture of Gaussians we have K Gaussian distributions

$$p(X) = \sum_{k=1}^{K} \pi_k \mathcal{N}(X|\mu_k, \Sigma_k)$$

where $\mathcal{N}(X|\mu_k, \Sigma_k)$ represents the multivariate Gaussian distribution with mean vector μ_k and covariance matrix Σ_k, K represents the number of Gaussian mixtures, and π_k is the mixing proportion for the kth component, and $\sum_{k=1}^{K} \pi_k = 1$.

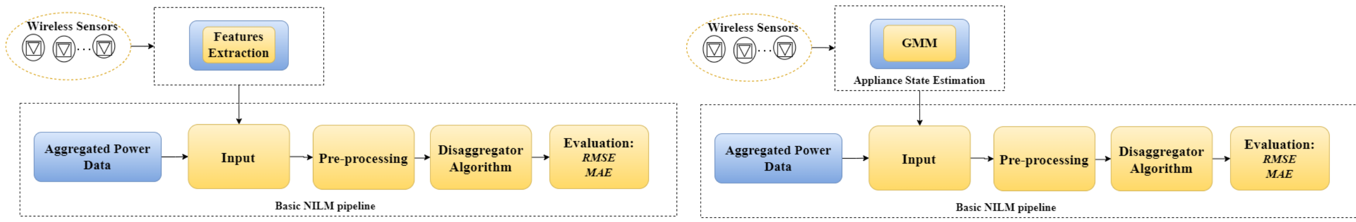
Since there is a mixture of Gaussians, the parameters μ_k, Σ_j, and π_k have to be determined. These parameters are estimated by using the Expectation-Maximization (EM) method [23]. EM consists of two steps (cf Algo 1). In the expectation step (E-Step), the initial estimates of the parameters are assigned using k-mean clustering and the probability of each latent variable is calculated. In the maximization step (M-Step), the parameters are modified to maximize the likelihood of the data with the initial assignments. These steps are repeated until a local optimum is achieved.

Algorithm 1 Expectation-Maximization Algorithm for Gaussian Mixture Model

Require: Data points $x_1, x_2, \ldots, x_N$; number of mixture components K

Initialize parameters: π_j, μ_j, Σ_j for $j = 1, 2, \ldots, K$
while not converged **do**
  E-Step:
  for $i = 1$ to N **do**
    for $j = 1$ to K **do**
      $\gamma_{ij} \leftarrow \frac{\pi_j \mathcal{N}(x_i|\mu_j,\Sigma_j)}{\sum_{k=1}^{K} \pi_k \mathcal{N}(x_i|\mu_k,\Sigma_k)}$
    end for
  end for
  M-Step:
  for $j = 1$ to K **do**
    $\pi_j \leftarrow \frac{\sum_{i=1}^{N} \gamma_{ij}}{N}$
    $\mu_j \leftarrow \frac{\sum_{i=1}^{N} \gamma_{ij} x_i}{\sum_{i=1}^{N} \gamma_{ij}}$
    $\Sigma_j \leftarrow \frac{\sum_{i=1}^{N} \gamma_{ij}(x_i-\mu_j)(x_i-\mu_j)^T}{\sum_{i=1}^{N} \gamma_{ij}}$
  end for
end while

In this code, K is the number of Gaussian components in the GMM, N is the number of data points, π_j is the mixing coefficient for the j-th Gaussian component, μ_j and Σ_j are the mean and covariance of the j-th Gaussian component, γ_{ij}



(a) General Semi-Intrusive Load Monitoring pipeline

(b) Proposed Semi-Intrusive Load Monitoring pipeline

Fig. 4: Semi-Intrusive Load Monitoring synoptic

is the responsibility assigned to the j-th Gaussian component for the i-th data point, $\mathcal{N}(x_i \mid \mu_j, \Sigma_j)$ is the Gaussian density function evaluated at x_i with mean μ_j and covariance Σ_j. The EM algorithm iterates between the E-step (Expectation step), where responsibilities are updated, and the M-step (Maximization step), where parameters of the Gaussian components are updated until convergence (classical threshold).

IV. Results

A. Dataset

The dataset used for the experiments below was exported from our SmartSense platform with a duration of 30 days. Afterwards, it is split into train and test data, each with 15 days. This paper focuses on separating the power consumption data of specific appliances in our lab's kitchen, including the coffee maker, kettle, microwaves, oven, and lights. The data includes both the overall power consumption as well as individual power usage for each appliance. Additionally, the data includes readings from 25 different sensors, measuring factors such as CO2 levels, humidity, and temperature as described in section III.

B. States estimation phase

a) Proof of Concept: Before starting the implementation of the GMM, we made a Proof of Concept (POC) in order to prove that adding an *a-priori* information of appliances states could improve the disaggregation performance. Indeed, the individual power consumption used as reference during the training phase of the disaggregation was transformed to 0 or 1 representing the states `OFF` and `ON` of an appliance, or `No Activity` and `Activity`. Then, these states information was added with the aggregated power data as an additional input of the standard NILM workflow. Finally, with this augmented input a typical disaggregation with NILMTK was performed.

b) GMM: The GMM model was implemented through the Python library Scikit-Learn. For each appliance the training data was used to fit two models for each class, `ON` and `OFF`. The log-likelihoods of each class were then determined, and the test data samples were assigned to the class with the highest log-likelihood ratio. For conciseness, in this paper we present the validation of the proposed semi-intrusive approach (see Fig.4(b)) by using a single measure corresponding to the CO2 sensor. The approach will be extended to higher dimensions. The initial findings focused on determining the states of the coffee maker and kettle and they are presented in Table II and Table III.

TABLE II: Confusion Matrix of coffee maker states estimation from sensors data with K = 9 and CO2 sensor

	Predicted Class		
True Class	Negative	Positive	**Support**
Negative	99726	32464	132190
Positive	141	1229	1370
	Precision	**Recall**	**F1 Score**
Negative	1.00	**0.75**	0.86
Positive	0.04	**0.90**	0.07

TABLE III: Confusion Matrix of kettle states estimation from sensors data with K = 7 and CO2 sensor

	Predicted Class		
True Class	Negative	Positive	**Support**
Negative	91189	41380	132569
Positive	315	676	991
	Precision	**Recall**	**F1 Score**
Negative	1.00	**0.69**	0.81
Positive	0.02	**0.68**	0.03

C. Disaggregation phase

During this stage, the previously estimated states were presensted as an additional input (with the overall power consumption) of the standard NILM pipeline that was executed using NILMTK.
We chose the The Windowed Gated Recurrent Unit (WindowGRU) model [24] as the disaggregation algorithm. In fact, it is a novel Recurrent Neural Network (RNN) architecture that is designed to capture long-term dependencies in sequence data. The algorithm introduces a window-based processing technique that selectively focuses on relevant past information while suppressing irrelevant information. The WindowGRU algorithm showed its effectiveness on various sequence modeling tasks, and showed that it outperforms traditional GRU models and other state-of-the-art approaches. Added to that, it has the potential to improve the performance of RNNs in various sequence modeling tasks that involve long-term dependencies.
The WindowGRU architecture incorporates a single convolutional layer with 16 filters, a 1x8 kernel size, and linear

activation. It also utilizes two Bidirectional GRU layers with 64 units, tanh activation, and a 0.5 dropout rate. Additionally, it employs two fully connected layers with 128 units and a linear activation function.
Table IV presents the outcomes for the coffee maker and kettle and compares it with the disaggregation performance using only the power consumption as input.
The unmarked values indicate the outcomes obtained solely from the total power data, while the values marked with underlines show the results obtained from the suggested input, which includes both the total power data and states information and finally, the values in bold indicate the results of the POC which we want our results to be as similar to them as possible.

TABLE IV: Evaluation of disaggregation performance

Appliances	RMSE	MAE
Kettle	69.82	6.69
	<u>48.66</u>	<u>3.11</u>
	15.37	**1.28**
Coffee maker	26.27	2.65
	<u>21.95</u>	<u>1.64</u>
	7.26	**1.04**

According to the table above, our approach surpassed the conventional NILM disaggregation with only CO2 sensor and demonstrated that incorporating sensors data about the environment can enhance NILM disaggregation.

V. Conclusion

Our study aimed to evaluate the effectiveness of a new NILM approach for disaggregating energy consumption data from smart meters and wireless sensors into individual appliance-level data. This new approach is based on integrating additionnal features to the input that were obtained from specific sensors data that provide additional information from each appliance environment. Our results show that the approach achieved high accuracy in identifying individual appliances compared to the conventional NILM approach.

However, our study also had some limitations that should be considered when interpreting the results. First, the algorithm was trained and tested on a relatively small dataset (15 days), which may limit its generalizability to other settings. Second, the accuracy of the algorithm varied not only across different appliance categories but also the sensors data (in this work we used only CO2 data). The difficulty here is to find the optimal combination of various sensors data that enhance the disaggregation performance for a given appliance but our solution is able to increase the GMM dimension.

Future research could explore ways to further improve the accuracy and robustness of NILM algorithms, particularly for challenging appliance categories such as printer and laptop whose consumption may depend on their activity.

References

[1] I. E. Agency, *World Energy Outlook 2019*. OECD Publishing: Paris, France, 2019.

[2] L. Shorrock, J. Utley *et al.*, *Domestic energy fact file 2003*. Citeseer, 2003.

[3] S. Darby *et al.*, "The effectiveness of feedback on energy consumption," *A Review for DEFRA of the Literature on Metering, Billing and direct Displays*, vol. 486, no. 2006, p. 26, 2006.

[4] G. Hart, "Nonintrusive appliance load monitoring," *Proceedings of the IEEE*, vol. 80, no. 12, pp. 1870–1891, 1992.

[5] M. E. Berges, E. Goldman, H. S. Matthews, and L. Soibelman, "Enhancing electricity audits in residential buildings with nonintrusive load monitoring," *Journal of industrial ecology*, vol. 14, no. 5, pp. 844–858, 2010.

[6] J. Liang, S. K. K. Ng, G. Kendall, and J. W. M. Cheng, "Load signature study—part i: Basic concept, structure, and methodology," *IEEE Transactions on Power Delivery*, vol. 25, no. 2, pp. 551–560, 2010.

[7] A. Zoha, A. Gluhak, M. A. Imran, and S. Rajasegarar, "Non-intrusive load monitoring approaches for disaggregated energy sensing: A survey," *Sensors*, vol. 12, no. 12, pp. 16 838–16 866, 2012.

[8] K. He, L. Stankovic, J. Liao, and V. Stankovic, "Non-intrusive load disaggregation using graph signal processing," *IEEE Transactions on Smart Grid*, vol. 9, no. 3, pp. 1739–1747, 2016.

[9] T. Hassan, F. Javed, and N. Arshad, "An empirical investigation of vi trajectory based load signatures for non-intrusive load monitoring," *IEEE Transactions on Smart Grid*, vol. 5, no. 2, pp. 870–878, 2013.

[10] J. Kelly and W. Knottenbelt, "Neural nilm: Deep neural networks applied to energy disaggregation," in *Proceedings of the 2nd ACM international conference on embedded systems for energy-efficient built environments*, 2015, pp. 55–64.

[11] J. Z. Kolter and T. Jaakkola, "Approximate inference in additive factorial hmms with application to energy disaggregation," in *Artificial intelligence and statistics*. PMLR, 2012, pp. 1472–1482.

[12] M. Zhong, N. Goddard, and C. Sutton, "Signal aggregate constraints in additive factorial hmms, with application to energy disaggregation," *Advances in Neural Information Processing Systems*, vol. 27, 2014.

[13] D. V. Smith and M. S. Shahriar, "A context aware sound classifier applied to prawn feed monitoring and energy disaggregation," *Knowledge-Based Systems*, vol. 52, pp. 21–31, 2013.

[14] J. Kelly and W. Knottenbelt, "Does disaggregated electricity feedback reduce domestic electricity consumption? a systematic review of the literature," *arXiv preprint arXiv:1605.00962*, 2016.

[15] L. Pereira and N. Nunes, "Performance evaluation in non-intrusive load monitoring: Datasets, metrics, and tools—a review," *Wiley Interdisciplinary Reviews: data mining and knowledge discovery*, vol. 8, no. 6, p. e1265, 2018.

[16] J. Z. Kolter and M. J. Johnson, "Redd: A public data set for energy disaggregation research," in *Workshop on data mining applications in sustainability (SIGKDD), San Diego, CA*, vol. 25, no. Citeseer. Citeseer, 2011, pp. 59–62.

[17] A. Filip *et al.*, "Blued: A fully labeled public dataset for event-based nonintrusive load monitoring research," in *2nd workshop on data mining applications in sustainability (SustKDD)*, vol. 2012, 2011.

[18] 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. 1, pp. 1–14, 2015.

[19] N. Batra, J. Kelly, O. Parson, H. Dutta, W. Knottenbelt, A. Rogers, A. Singh, and M. Srivastava, "Nilmtk: An open source toolkit for non-intrusive load monitoring," in *Proceedings of the 5th international conference on Future energy systems*, 2014, pp. 265–276.

[20] A. Zoha, *Load Disaggregation and Monitoring in a Smart Office Space*. University of Surrey (United Kingdom), 2014.

[21] E. Patel and D. S. Kushwaha, "Clustering cloud workloads: K-means vs gaussian mixture model," *Procedia Computer Science*, vol. 171, pp. 158–167, 2020.

[22] D. J. Bora, D. Gupta, and A. Kumar, "A comparative study between fuzzy clustering algorithm and hard clustering algorithm," *arXiv preprint arXiv:1404.6059*, 2014.

[23] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the em algorithm," *Journal of the royal statistical society: series B (methodological)*, vol. 39, no. 1, pp. 1–22, 1977.

[24] O. Krystalakos, C. Nalmpantis, and D. Vrakas, "Sliding window approach for online energy disaggregation using artificial neural networks," in *Proceedings of the 10th Hellenic Conference on Artificial Intelligence*, 2018, pp. 1–6.