Non-Intrusive Load Monitoring of Single and Aggregated Profiles with a Hidden Markov Model

Nadège Miquey, Etta Grover-Silva

Non-Intrusive Load Monitoring of Single and Aggregated Profiles with a Hidden Markov Model

1st Nadège Miquey
Ecole centrale de Lyon
Lyon, France
nadege.miquey@ecoco2.com

2nd Etta Grover-Silva
Eco CO2
Nanterre, France
etta.grover-silva@ecoco2.com

***Abstract*—Awareness raising programs to encourage energy efficient behaviour is important in the context of the current energy transition. Sensors and connected devices allowing for data collection are easily available providing an opportunity to collect electric consumption data from individual appliances. The effective use of these data sources is necessary to optimize an energy efficiency coaching program. This paper presents a methodology for non-intrusive load monitoring analysis on individual smart plugs to identify an unknown appliance and disaggregated an aggregated load profile such as a power strip. The automatic detection of a connected appliance allows for appliance usage suggestions to be provided quickly without the need of an end-user input. The disaggregation of an aggregated curve such as a power strip allows for the optimization of the number of smart plugs required to represent a significant proportion of the total energy use of the household. The down sampling of high resolution data was also performed to observe the performance of the methodology on lower resolution data. The single appliance identification models all had very high accuracy (between 94 - 100 %). The disaggregation of an aggregated profile in the kitchen use case also had high accuracy for data with a resolution of less than one minute (95 - 99 %). The disaggregation of an aggregated profile for a multi-media use case had a lower performance when more than two appliances were considered (55 - 85 %).**

***Index Terms*—load monitoring, machine learning, disaggregation, energy efficiency**

I. Introduction

The energy transition has become an urgent topic in the context of the accelerating climate change. This transition implies the significant reduction of resource usage in order to accomplish ambitious goals set forth by world leaders including the European Union (EU). The building sector and more precisely the residential building sector is a large contributor to the energy consumption of a country. In the European Union, households represent one-fourth of its total final energy consumption [2]. In recent years, emissions from building energy used has increased reaching an all-time high in 2018 [1].

The energy consumption of a household is dependent on the electric appliances present and the use of these appliances defined by individual appliance usage behaviour. The second factor can have a huge impact on the overall energy use. Inefficient use of building systems and appliances is called the energy efficiency gap. This gap can represent up to 100% of excessive energy use in comparison to the design requirements [6]. The EU has put in place strict regulations about new building construction to improve energy efficient buildings, however few regulations exist addressing the behavioural use of buildings after construction.

Energy efficient behaviour is highly specific to individual households. Therefore, effective awareness raising programs are difficult to implement on a large scale without integrating household specific analysis and advice. Several studies have shown that appliance specific energy use combined with real-time feedback results in the highest energy savings for multiple awareness raising programs [5].

With recent massive deployment of smart meters in multiple countries, individual household electric load data is more easily accessible. Multiple awareness raising programs have been deployed by the company Eco CO2 in the context of a public tender in France put forth by ADEME related to a funding mechanism called *Investissement d'avenir*. The programs deployed include TBH Alliance [4], Picowatty [7] and SEIZE [8]. These programs focus on encouraging energy efficient behaviour by providing useful and user specific information about the environment and the direct impact of their actions.

Recently, individual appliance electric load plugs have been integrated in Eco CO2 data collection platform. These smart plugs allow for the precise measurement of individual appliance load profiles or of a power strip to provide appliance specific advice to individual users. However, the cost of individual smart plugs does not allow for the individual load monitoring of all appliances within a household. In order to maximize the impact of a single smart plug, non-intrusive load monitoring (NILM) techniques [10], [14] can be applied to decompose a combine load curve of up to 6 appliances connected to a power strip.

Multiple algorithms have been implemented in the NILM domain for the classification of appliance profiles including K-Nearest Neighbors (KNN) [9], [11], K-means [3], Decision Tree, Random Forest, Recurrent Neural Network (RNN) [12] and Hidden Markov Model (HMM) [13]. The selection of the appropriate method is based on the data available for model training, the data resolution and the computational requirements.

In [16] individual appliance load data [15] with a one second resolution are modeled with a Random Forest, LogitBoost, Bagging, Decision Tree, Naive Bayes and Support Vector

Machine algorithms. The results were compared for the classification of 33 devices. Accuracy ranged from 90-96% with the Random Forest algorithm performing the best. However, these algorithms were not demonstrated to be capable of analyzing aggregated load profiles as well.

The KNN, K-means and Decision Tree algorithms are well adapted for single appliance identification. In [11] 10 second resolution data of active and reactive power is used for the correct classification of 8 individual appliances. One of these algorithms, the Feed-forward Neural Network, was also applied to aggregated curves to identify individual appliances within an aggregated profile. The aggregated model used three appliances with a 90-98% accuracy depending on the combined states.

A review of event based and non-event based NILM schemes is presented in [13] where two main techniques are discussed: HMM and Convolutional Neural Networks (CNN). The CNN and more generally Deep Neural Networks (DNN) approach was deemed more computationally intensive for the training period and relatively lower accuracy (below 90%) in comparison to HMM techniques. The HMM is classified as a non-event based NILM scheme and is preferential due to the effective application on individual and aggregated curves, the simplicity of the model training and the application on low resolution data.

In [10] advantages of HMM in comparison to other NILM techniques are described as being effective when labeled data is available and on low frequency timeseries data. These models can also have a low computational requirement depending on the model structure. This type of model has been proven to be effective on the identification of single appliances as well as aggregated profiles. It has been noted that this model is highly effective for multi-state appliances with distinct power levels but not ideal for multi-state variable power consumption appliances.

In this paper, a HMM methodology is implemented due to the effectiveness of the model structure to classify profiles for an individual appliance as well as an aggregated load profile. It has been proven to be effective on low resolution data which is important for commercial application of the model and the required data storage necessary for effective classification. The training period is relatively simple and has a low computational requirement. The implementation of this model is innovative to help in providing targeted energy saving advice per appliance. Based on initial appliance classification and individual appliance consumption analysis, awareness raising programs and impact can be more closely studied.

The following sections will detail the simple but robust model developed for the automatic identification of individual appliances that is also applicable to a power strip to identify individual appliances in an aggregated load profile. The methodology is presented in section II followed by a realistic case study in section III. Results are discussed in IV which includes a performance study of the developed model on varying data resolutions.

II. Methodology

In this section the Hidden Markov Model and pre-processing steps of constructing the model features will be explained. The pre-processing and feature extraction of the timeseries data is composed of three main steps: clustering of timeseries values, definition of buckets, feature quantification of buckets. These three steps allow for the effective quantification of the timeseries data to then develop individual appliance Hidden Markov Models for individual appliance recognition. These individual appliance models are then combined to produce an effective state identification model to separate individual appliance timeseries signals out of an aggregated timeseries signal.

A. Data pre-processing

1) Clustering on power consumption values: For each appliance, a Kmeans clustering method is used to identify groups of power consumption values and define the number of states characteristic of each appliance. The clustering method is applied to timeseries data for a duration of 2 to 10 days specified in Table I for each single appliance. The duration of training timeseries is chosen to get a representative number of active periods for each appliance. For appliances functioning more than 6 hours a day as refrigerators, screen or internet router, 3 days allow to identify functioning range of values. For appliances functioning occasionally, up to 10 days of data are necessary to identify range of values of functioning periods.

TABLE I
KMEANS CLUSTERING, SAMPLING RATE = 5 SECONDS

Appliance type	Duration of training timeseries (days)	Number of active periods
Hot-water boiler	10	21
Refrigerator	3	94
Coffee-machine	10	29
Washing-machine	10	10
Screen	3	8
Internet router	3	NA[a]
Laptop charger	6	8
Television	5	12

[a]Not Applicable, internet router is an always on appliance.

A silhouette score is computed to determine the ideal number of clusters for each appliance. The average power value associated with each cluster is also calculated.

2) Interpolation and bucketing of aggregated profile: Load profile data is assigned labels based on assigned clusters from the Kmeans clustering method. This allows for the reconstruction of the timeseries data with labeled time periods associated with each Kmeans cluster. An average value is then assigned to each cluster and used to remove all variation within the period designated as the same cluster label. An interpolation method is used to fill in missing data and reconstruct a continuous timeseries profile. This new constructed load profile is then separated into buckets based on the magnitudes evaluated in the Kmeans clustering method.

3) Test and training data processing: Characteristics of the testing data set can differ from characteristics of the training data set. To avoid this, load profiles are constructed to obtain test load profiles closer to training load profiles and improve models performances. To construct the load profiles, buckets are created around each cluster's value identified on training load profiles. For the appliance i, with n_i clusters of mean values $m_1, m_2, ..., m_n$, created buckets are $m_1 - x\%, m_1 + x\%, m_2 - x\%, ..., m_n + x\%$. Models performances are evaluated and compared on raw test data and on pre-processed load profiles.

B. Hidden Markov Model

The calculated features are then used to construct individual Hidden Markov Models for each individual appliance. A multi-appliance model is then created by combining all possible states of the individual appliance models.

1) Single appliance model: The ideal number of clusters computed with the Kmeans clustering method is used to set the number of hidden states for each single Hidden Markov Model. Parameters of each single appliance HMM are estimated with an Expectation–Maximization (EM) algorithm.

2) Multi-appliance model: Single appliance models are then combined to obtain Hidden Markov Model for multiple appliances. Single appliance models states are combined to form N distinct combinations of states of the combined model.

$$N = \prod_{i=0}^{k} n_i \tag{1}$$

where $k \in \mathbb{N}^*$ is the number of appliances combined in the model, $n_i \in \mathbb{N}^*$ is the number of states of the ith appliance.

Transition matrices of single appliance models are combined using Kronecker product defined by 3:

$$\mathbf{A} \otimes \mathbf{B} = \begin{pmatrix} a_{1,1}\mathbf{B} & \cdots & a_{1,n}\mathbf{B} \\ a_{2,1}\mathbf{B} & \cdots & a_{2,n}\mathbf{B} \\ \vdots & \ddots & \vdots \\ a_{m,1}\mathbf{B} & \cdots & a_{m,n}\mathbf{B} \end{pmatrix} \tag{2}$$

Viterbi algorithm is used to decode sequences and determine the most probable sequence of hidden states (which appliance is in which states) corresponding to the observable sequence of power consumption traces.

3) Splitting data into a training set and a test set: For each appliance class, varying appliance load profiles are used to train the HMM (different refrigerators, washing-machines etc). Feature extraction using bucketing allows to identify specific characteristics for each appliance in the same class (specific load profile patterns, maximum power, average power etc). Testing of the developed model is then performed on load profiles that were not included in the training set.

C. Single model evaluation metrics

Single HMM are evaluated on 15 days of data using the following metrics:

- the accuracy: the proportion of correct On and Off states prediction
- the precision defined in eq. 3: the proportion of correct On states prediction among all On states predicted. TP is the number of True Positive elements, FN is the number of False Negative elements

$$precision = \frac{TP}{TP + FN} \tag{3}$$

- the f1-score defined in eq. 4, where precision is defined in eq. 3 and recall defined in eq. 5

$$F1 = 2 * \frac{precision * recall}{precision + recall} \tag{4}$$

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

D. Combined model evaluation metrics

To evaluate performances of states prediction, two different metrics are used. A micro-averaged $accuracy$ score defined in eq. 6 is computed. This score measures the fraction of correct states prediction among all possible combination of states of the combined model. When power consumption values are below a certain threshold, they are bucketed to zero consumption values. States where all appliances are Off are thus the easiest ones to detect and these states concern a large part of each recording. To get a more representative accuracy score, the micro-averaged $accuracy_{on}$ score defined in eq. 7 does not take into account the state where all appliances are Off.

$$micro\_accuracy = \frac{\sum_{n=0}^{N} TP_n}{\sum_{n=0}^{N} TP_n + \sum_{n=0}^{N} FP_n} \tag{6}$$

$\forall$n $\in [0, N]$ such that $n \neq$ *off state* :

$$micro\_accuracy_{on} = \frac{\sum_{n=0}^{N} TP_n}{\sum_{n=0}^{N} TP_n + \sum_{n=0}^{N} FP_n} \tag{7}$$

Where:

- N is the number of states defined in 1.
- TP_n refer to True Positive elements of states n: predicted states are n and correspond to ground truth.
- FP_n refer to False Positive element of states n: predicted states are n whereas ground truth is not n.

III. Case study

The primary objective of this case study is to show the effectiveness of the defined methodology to automatically classifying individual appliances as well as combined appliances load profiles. A combined appliance load profile could be a smart plug monitoring a power strip where up to six appliances are connected. The data used for testing and validation was collected through an internal experimental study performed with employees of Eco CO2. The recorded power consumption data was collected for 64 smart plugs. More than 3800 days of data have been collected during the experimental study. The appliances used in this case study include refrigerators, coffee-machines, hot-water boilers, microwave-ovens, internet router, screens, computers, televisions, washing-machines. Two realistic power strip configurations have been defined to simulate a realistic application of a power strip found in a typical residential household:

- Aggregated load profile of four kitchen appliances: a hot-water boiler, a refrigerator, a coffee-machine, a washing-machine.
- Aggregated load profile of four multimedia appliances: a screen, an internet router, a laptop charger and a television.

Three models for each category is then computed: a combined HMM model of two single appliance models, a combined HMM model of three single appliance models and a combined HMM model of four single appliance models.

IV. Results

In this section, results of the Kmeans clustering method applied on eight single appliances will be detailed. State prediction performances of the HMM defined on single appliances and aggregated profiles will be evaluated. The aggregated profiles represent two realistic configurations of possible power strips in a kitchen or office scenario presented in section IV. A comparison between results on kitchen appliances and on multimedia appliances is done based on the specificity of each appliance type. The impact of data resolution degradation on combined appliance HMM states predictions will be studied.

A. Model prediction of on and off states evaluation

Table II presents each cluster number and cluster centroide for each appliance computed with a Kmeans method. A majority of appliances analyzed could be coerced into a maximum two states. The one exception being the refrigerator category due to a high peak as a result of the electrical inertia of the compressor therefore creating three distinct states.

TABLE II
KMEANS RESULTS, SAMPLING RATE = 1 SECONDS

Appliance category	Appliance type	Number of clusters	Clusters centroids (W)
Kitchen appliances	Hot-water boiler	2	[0.3839, 2468]
	Refrigerator	3	[0.1490, 117.7, 1269]
	Coffee-machine	2	[1.417, 1576]
	Washing-machine	2	[2.813, 2438]
Multimedia appliances	Screen	2	[1.0, 29.06]
	Internet router	2	[0.0, 8.122]
	Laptop charger	2	[13.04, 0.015]
	Television	2	[0.0, 129.4]

Appliances with no zero state can be interpreted as loads that contribute to the standby consumption. Therefore, the appliance is never completely off and consumes electricity constantly even if the consumption is very low.

B. State prediction appliance models

1) State prediction of single appliance models: Each model is evaluated using the evaluation metrics defined in II-C to determine the accuracy of state prediction. Evaluation metrics are calculated for 15 periods of 24-hours. Each test sample is the addition of 24-hours recordings of single appliances present in the tested model.

TABLE III
SINGLE APPLIANCE HMM RESULTS ON RAW DATA FOR KITCHEN APPLIANCES, SAMPLING RATE = 1 SECOND

Appliance category	Appliance type	Accuracy (%)	Precision	f1-score
Kitchen appliances	Hot-water boiler	99.9	0.82	0.90
	Refrigerator	99.7	NA[a]	NA[a]
	Coffee-machine	99.6	0.99	0.99
	Washing-machine	94.3	0.75	0.84

[a]Not Applicable, refrigerators are three-state appliances.

Table III shows the accuracy results of states prediction for each single appliance HMM when raw data are processed. Low accuracy results are observed for the washing-machine appliance due to the nature of the load profile that can have a higher variation in operational states from one appliance to another. This implies a possible significant difference in shape of a training data set and the test data set of a new appliances not seen by the model in the training set. All other appliances had a high accuracy. The precision of the hot water boiler is lower than the coffee machine. This could be due to a higher variability in the usage duration which is variable with the amount of water being heated.

TABLE IV
SINGLE APPLIANCE HMM RESULTS ON RAW DATA FOR MULTI-MEDIA APPLIANCES, SAMPLING RATE = 1 SECOND

Appliance category	Appliance type	Accuracy (%)	Precision	f1-score
Multimedia appliances	Screen	99.9	0.99	0.99
	Internet router	100	NA[a]	NA[a]
	Laptop charger	99.6	0.93	0.95
	Television	99.9	0.99	0.99

[a]Not Applicable, internet router is an always on appliance.

Similar high accuracy results are found for the multi-media appliances with regardless of the specificity of one appliance in comparison to another.

TABLE V
SINGLE APPLIANCE HMM RESULTS ON PRE-PROCESSED LOAD PROFILES FOR KITCHEN APPLIANCES, SAMPLING RATE = 1 SECOND

Appliance category	Appliance type	Accuracy (%)	Precision	f1-score
Kitchen appliances	Hot-water boiler	99.9	0.97	0.98
	Refrigerator	99.9	NA[a]	NA[a]
	Coffee-machine	99.2	0.72	0.72
	Washing-machine	99.3	0.99	0.98

[a]Not Applicable, refrigerators are 3-states appliances

Table V shows the accuracy results for all state predictions and for on state predictions for each single appliance HMM when load profiles are pre-processed as described in II-A3. Using pre-processed load profiles improve results for washing-machine models significantly. Accuracy is also improved in general for all appliances.

TABLE VI
SINGLE APPLIANCE HMM RESULTS ON PRE-PROCESSED LOAD PROFILES FOR MULTI-MEDIA APPLIANCES, SAMPLING RATE = 1 SECOND

Appliance category	**Appliance type**	**Accuracy (%)**	**Precision**	**f1-score**
Multimedia appliances	Screens	99.9	0.99	0.99
	Internet router	100	NA[a]	NA[a]
	Laptop charger	94.2	0.99	0.91
	Television	99.9	0.99	0.99

[a]Not Applicable, internet router is an always on appliance.

Table VI shows the accuracy results for state predictions for each single multimedia appliance HMM on pre-processed load profiles. In comparison to the analysis done on raw data, pre-processing actually decreased the accuracy of detection for the laptop charger.

2) Combined appliances models: Two, three and four appliance models trained on single kitchen appliances are combined. State predictions of the models are evaluated on combined power load profile of two, three and four appliances respectively. Figure 1 shows states prediction of a two-appliances HMM (hot-water boiler and refrigerator).

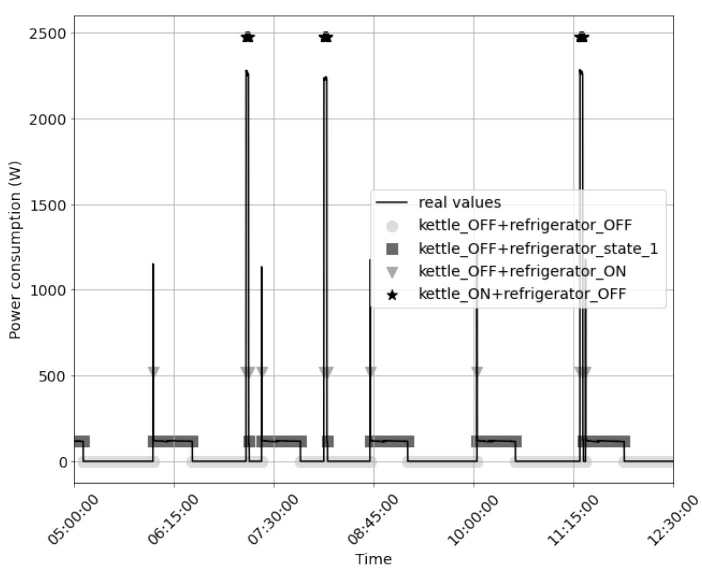

Fig. 1. States predictions HMM of hot-water boiler and refrigerator combined, sampling rate = 1 second

The accuracy of combined 2, 3 and 4 appliance models can be seen in Fig. 2

C. Data resolution degradation

Optimizing the size of the data storage solution and the transmission volume of data is a major issue when it comes to the cost of implementation in a commercialized product. The minimum resolution required while still guaranteeing a high accuracy is tested by purposefully degrading the data resolution of each profile. Power load profiles used for training and for testing are re-sampled to the following data resolutions: [3 sec, 5 sec, 10 sec, 20 sec, 30 sec, 1 min, 2 min, 5 min, 10 min] to study the impact of the down-sampling on the performances of states prediction.

1) Combined HMM for kitchen appliances: For combined kitchen appliance HMM, On and Off state predictions are presented in Fig. 2.

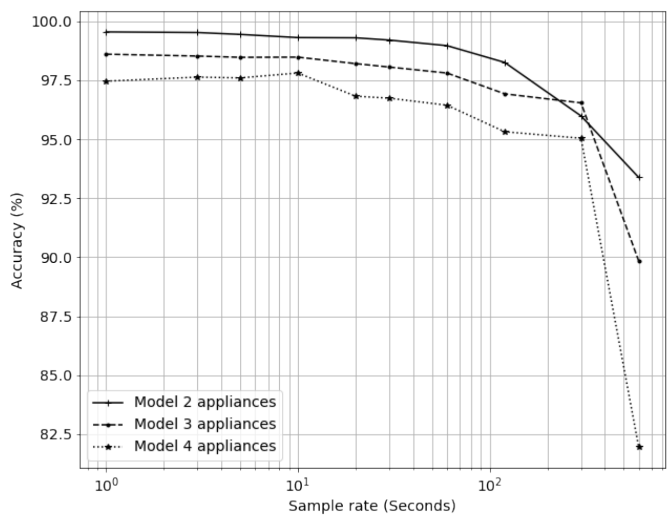

Fig. 2. On and Off state prediction accuracy results (%) versus sample rate (Seconds) for combined kitchen appliances models. Model 2 appliances includes hot water boiler and the refrigerator, model 3 appliances is with the addition of the coffee machine and model 4 appliances includes also the washing machine.

For two HMM combined, overall state predictions and On state predictions accuracy is above 98% for every sample rate value bellow 1 minute. Down-sampling data decreases significantly on-states prediction accuracy for combined HMM for a re-sampling rate above one minute :

- for three HMM combined, on-states prediction accuracy decreased from 96.2% to 87.6%,
- for four HMM combined, on-states prediction accuracy decreased from 94.3% to 89.5%.

For combined multi-media device HMM, On and Off state predictions are presented in Fig. 3.

The combined model for multi-media appliances has low accuracy when three and four appliance curves were combined. However, the accuracy is relatively constant across all resolutions. This is due to the on off state characteristics of these appliances. The internet box and laptop charger have a constant power value when plugged in. They also have relatively low power values and the difference between the load curves are very minimal. This is a limitation of this

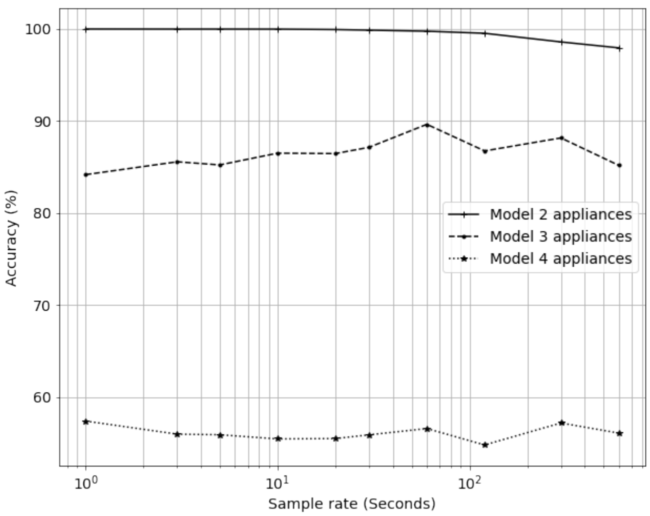

Fig. 3. On and Off state prediction accuracy results (%) versus sample rate (Seconds) for combined multi-media device models. Model 2 appliances includes the screen and internet box, model 3 appliances is with the addition of the television and model 4 appliances includes also the laptop charger.

model. The model is incapable of distinguishing and detecting two low magnitude constant power curve signals.

V. Conclusion

This paper has presented a methodology for a HMM applicable to single appliance identification and multi-appliance load curve disaggregation. High accuracy was achieved for all single appliance models. High accuracy was also achieved for multi-appliance models in the kitchen use case. This is due to the unique characteristics of each appliance considered, creating little difficulty for the model to identify all individual states of each appliance in the aggregated load profile. The combined multi-media accuracy was significantly lower than the kitchen appliance use case when 3 and 4 appliance load curves were combined. This was due to the minimal differences between individual load profiles. A constant power profile with low variation and low magnitude also was difficult to desegregate. These characteristics highlight the limits of this model application. Future work could include a hybrid method to apply the HMM only to variable appliance profiles. For example a linear aggregation model could be used on the constant low consumption appliance profile types which could improve performance for the multi-media disaggregation.

References

[1] Buildings sector CO2 emissions in the sustainable development scenario, 2019. URL: https://www.iea.org/ [accessed: 2019-10-02].

[2] Statistical office of the european union, 2019. URL: https://ec.europa.eu/eurostat [accessed: 2019-10-02].

[3] Vibhatha Abeykoon, Nishadi Kankanamdurage, Anuruddha Senevirathna, Pasika Ranaweera, and Rajitha Udawapola. Real time identification of electrical devices through power consumption pattern detection.

[4] TBH Alliance. Projet tbh, 2015.

[5] K. Carrie Armel, Abhay Gupta, Gireesh Shrimali, and Adrian Albert. Is disaggregation the holy grail of energy efficiency? the case of electricity. *Energy Policy*, 52:213–234, 2013.

[6] Fredrik Berg, Anne-Cathrine Flyen, Åsne Lund Godbolt, and Tor Broström. User-driven energy efficiency in historic buildings: A review. *Journal of Cultural Heritage*, 28:188 – 195, 2017.

[7] Eco CO2. Picowatty.

[8] Eco CO2. Sensibiliser et accompagner les tpes / pmes et les collectivités territoriales dans leur transition énergétique (seize).

[9] J. R. Cuñado and N. B. Linsangan. A supervised learning approach to appliance classification based on power consumption traces analysis. 517:012011. Publisher: IOP Publishing.

[10] Anthony Faustine, Nerey Henry Mvungi, Shubi Kaijage, and Kisangiri Michael. A survey on non-intrusive load monitoring methodies and techniques for energy disaggregation problem.

[11] O. Hamid, M. Barbarosou, P. Papageorgas, K. Prekas, and C-T Salame. Automatic recognition of electric loads analyzing the characteristic parameters of the consumed electric power through a non-intrusive monitoring methodology. 119:742–751.

[12] G.W. Hart. Nonintrusive appliance load monitoring. 80(12):1870–1891. Conference Name: Proceedings of the IEEE.

[13] Yassine Himeur, Abdullah Alsalemi, Faycal Bensaali, and Abbes Amira. Robust event-based non-intrusive appliance recognition using multi-scale wavelet packet tree and ensemble bagging tree. *Applied Energy*, 267:114877, 2020.

[14] Christoph Klemenjak and Peter Goldsborough. Non-intrusive load monitoring: A review and outlook. 10 2016.

[15] Andreas Reinhardt, Paul Baumann, Daniel Burgstahler, Matthias Hollick, Hristo Chonov, Marc Werner, and Ralf Steinmetz. On the accuracy of appliance identification based on distributed load metering data.

[16] Andrea Tundis, Ali Faizan, and Max Mühlhäuser. A feature-based model for the identification of electrical devices in smart environments. 19(11):2611. Number: 11 Publisher: Multidisciplinary Digital Publishing Institute.