UNIVERSIDADE ESTADUAL DE CAMPINAS
Faculdade de Engenharia Elétrica e de Computação

Dimas Augusto Mendes Lemes

Load Disaggregation: Approaches for Application in HEMS

Desagregação de Cargas: Abordagens para Aplicação em Sistemas HEMS

Campinas

2023

Dimas Augusto Mendes Lemes

Load Disaggregation: Approaches for Application in HEMS

Desagregação de Cargas: Abordagens para Aplicação em Sistemas HEMS

Thesis presented to the School of Electrical and Computer Engineering of the University of Campinas in partial fulĄllment of the requirements for the degree of Doctor in Electrical Engineering, in the area of Telecommunications and Telematics.

Tese apresentada à Faculdade de Engenharia Elétrica e de Computação da Universidade Estadual de Campinas como parte dos requisitos exigidos para a obtenção do título de Doutor em Engenharia Elétrica, na Área de Telecomunicações e Telemática.

Supervisor: Prof. Dr. Gustavo Frandenraich

Este exemplar corresponde à versão final da tese defendida pelo aluno Dimas Augusto Mendes Lemes, e orientada pelo Prof. Dr. Gustavo Frandenraich

Campinas

2023

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca da Área de Engenharia e Arquitetura
Rose Meire da Silva - CRB 8/5974

Lemes, Dimas Augusto Mendes, 1986-
L543L Load disaggregation : approaches for application in HEMS / Dimas Augusto Mendes Lemes. – Campinas, SP : [s.n.], 2023.

Orientador: Gustavo Fraidenraich.
Tese (doutorado) – Universidade Estadual de Campinas, Faculdade de Engenharia Elétrica e de Computação.

1. Inteligência artificial. 2. Sistemas de gestão de energia. 3. Detecção de anomalias. 4. Aprendizado de máquina. 5. Transformada Wavelet. 6. Análise de componentes principais. 7. Automação residencial. I. Fraidenraich, Gustavo, 1975-. II. Universidade Estadual de Campinas. Faculdade de Engenharia Elétrica e de Computação. III. Título.

Informações Complementares

Título em outro idioma: Desagregação de cargas : abordagens para aplicação em sistemas HEMS
Palavras-chave em inglês:
Artificial intelligence
Energy management systems
Anomaly detection
Machine learning
Wavelet transform
Principal component analysis
Home automation
Área de concentração: Telecomunicações e Telemática
Titulação: Doutor em Engenharia Elétrica
Banca examinadora:
Gustavo Fraidenraich [Orientador]
Michelle Soares Pereira Facina
Marcos Tomio Kakitani
Fabbryccio Akkazzha Chaves Machado Cardoso
Yusef Rafael Cáceres Zuñiga
Data de defesa: 14-03-2023
Programa de Pós-Graduação: Engenharia Elétrica

Identificação e informações acadêmicas do(a) aluno(a)
- ORCID do autor: https://orcid.org/0000-0001-5189-4129
- Currículo Lattes do autor: http://lattes.cnpq.br/9669584754688405

COMISSÃO JULGADORA - TESE DE DOUTORADO

Candidato: Dimas Augusto Mendes Lemes - **RA**: 163660
Data da Defesa: 14 de março de 2023
Título da Tese: Desagregação de Cargas: Abordagens para Aplicação em Sistema HEMS

Prof. Dr. Gustavo Fraidenraich (Presidente, FEEC/UNICAMP)
Dra. Michelle Soares Pereira Facina (CPQD)
Prof. Dr. Marcos Tomio Kakitani (UFSJ)
Dr. Fabbryccio Akkazzha Chaves Machado Cardoso (CPQD)
Dr. Yusef Rafael Cáceres Zuñiga (EMBRAER)

A ata da defesa, com as respectivas assinaturas dos membros da Comissão Julgadora, encontra-se no SIGA (Sistema de Fluxo de Dissertação/Tese) e na Secretaria de Pós-Graduação da Faculdade de Engenharia Elétrica e Computação.

I dedicate this work to everyone who contributes with their efforts, studies, research, and works towards a better life for humanity and more efficient use of resources.

Acknowledgements

First, I would like to thank God for all the opportunities and achievements and for giving me the strength to continue pursuing my goals.

I am also grateful to my parents, Maria and Dimas, my brother Igor, and my girlfriend Laura, for supporting me and being present in every moment.

I thank my friends, Claudio, Yalena, Nathália, Ruby, Daniela, William, Fábio, and Reginaldo, for the support, hints, and easygoing moments.

I am grateful to professor Gustavo Fraidenraich for all the suggestions, insights, patience, and advice, for sharing his knowledge, and for always being available for support.

I thank my coworkers Carlinho and Thales for all the companionship, tips, ideas, and easygoing moments.

I thank professors Luis Meloni and Paulo Cardieri for all the dedication and orientation during the Projeto Copel.

I thank my colleagues at WissTek for all the companionship.

And I also thank CNPq and Companhia Paranaense de Energia (COPEL) for the investment and for believing in this work.

Resumo

Este trabalho de doutorado investiga métodos de aplicação em sistemas HEMS (*Home Energy Manangement System*) para melhorar a eficiência no consumo de energia e contribuir para que o usuário tenha mais informações quanto ao seu perfil de consumo. Para alcançar este objetivo, uma abordagem utilizando a Análise de Componentes Principais (PCA) para a extração de padrões e um sistema de tomada de decisão baseado em ferramentas de aprendizado de máquina, como *k-Nearest Neighbor* (kNN), *Decision Tree* e *Random Forest* foram empregados para a tarefa de desagregação de cargas. O método proposto atingiu um F1-Score de 95.13% para o *Rainforest Automation Energy Dataset* (RAE), e 95.51% para *Reference Energy Disaggregation Dataset* (REDD). Outra abordagem considerada foi a de detecção de falhas em eletrodomésticos, em que utilizando os dados presentes no *Personalized Retrofit Decision Support Tools for UK Homes using Smart Home Technology* (REFIT) e a transformada Wavelet alcançou-se 87.33% de acurácia na tarefa de detecção de ocorrência de mal funcionamento dos dispositivos presentes no banco de dados.

Palavras-chaves: Inteligência Artificial; Sistemas de Monitoramento de Energia; Detecção de Anomalia; Aprendizado de Máquina; Análise de Componentes Principais; Automação Residencial.

Abstract

This thesis investigates methods of application in HEMS (Home Energy Management System) to improve energy consumption efficiency and contribute to the user understanding of their consumption profile. To achieve this purpose, an approach using Principal Component Analysis (PCA) for pattern and feature extraction and a decision-making system based on machine learning techniques such as k-Nearest Neighbor (kNN), Decision Tree, and Random Forest were employed for the load disaggregation task. The proposed method achieved an F1-Score of 95.13% for Rainforest Automation Energy Dataset (RAE) and 95.51% for Reference Energy Disaggregation Dataset (REDD). Another approach considers the detection of faults in home appliances. By using the data available in the "Personalized Retrofit Decision Support Tools for UK Homes using Smart Home Technology" (REFIT) and the Wavelet transform, it achieves 87.33% of accuracy in the task of detecting the occurrence of malfunctions in the devices present in the database.

Keywords: Artificial Intelligence; Energy Manangement Systems; Anomaly Detection; Machine Learning; Principal Component Analysis; Home Automation.

"If I have seen further than others, it is by standing upon the shoulders of giants."
(Isaac Newton)

List of Figures

List of Tables

List of Acronyms

List of Publications

Journal Articles

- D. A. M. Lemes, T. W. Cabral, G. Fraidenraich, L. G. P. Meloni, E. R. Lima, and F. B. Neto, "Load Disaggregation Based On Time Window for HEMS Application", IEEE Access, vol. 9, pp. 70746 - 70757, May 2021. Impact Factor: 3.367

Co-Authored Journal Articles

- L. C. B. C. Ferreira, A. R. Borchardt, G. S. Cardoso, D. A. M. Lemes, G. R. R. Souza, F. B. Neto, E. R. Lima, G. Fraidenraich, P. Cardieri, and L. G. P. Meloni, "Edge Computing and Microservices Middleware for Home Energy Management Systems", IEEE Access, vol. 10, pp. 2169-3536, October 2022. Impact Factor: 3.367
- T. W. Cabral, D. A. M. Lemes, G. Fraidenraich, F. B. Neto, E. R. Lima, and L. G. P. Meloni, "Load Recognition Based On Hybrid System For HEMS Application", Recently accepted at IEEE Access. Impact Factor: 3.367
- F. A. P. Figueiredo, D. A. M. Lemes, C. F. Dias, and G. Fraidenraich, "Massive MIMO Channel Estimation Considering Pilot Contamination and Spatially Correlated Channels", Electronic Letters, vol. 56, issue 8, pp. 410-413, April 2020. Impact Factor: 1.202

Articles Under Review

- D. A. M. Lemes, T. W. Cabral, L. L. Motta, G. Fraidenraich, E. R. Lima, F. B. Neto, and L. G. P. Meloni, "Low Runtime Approach for Fault Detection for Refrigeration Systems in Smart Homes Using Wavelet Transform", Under Review at the IEEE Transactions On Cosumer Electronics.

Co-Authored Articles Under Review

- L. L. Motta, L. C. B. C. Ferreira, T. W. Cabral, D. A. M. Lemes, G. S. Cardoso, A. R. Borchardt, G. Fraidenraich, E. R. Lima, F. B. Neto, and L. G. P. Meloni, "General Overview and Proof of Concept of a Smart Home Energy Management System Architecture", Under Review at Elsevier.

Contents

1 Introduction

Nowadays, an issue attracting attention in academic and industrial environments relates to residential energy consumption. According to data, affected by climate change, consumption will increase by 84% by 2030 (VORSATZ *et al.*, 2015). Therefore, the technological evolution endeavors to provide tools allowing the end user to control their energy expenditures and to contribute to a more efficient distribution of resources by the companies, enhancing aspects of sustainability (AL-KABABJI *et al.*, 2022).

In this context, the analysis of the consumption profile is particularly relevant. Therefore, as stated in (YUE *et al.*, 2013), energy expenditures are closely influenced by the habits of the residents, age, gender, income level, educational level, and structure of the residence. Hence, it is essential to have information regarding energy consumption and to understand its pattern.

The recent developments of technologies capable of collecting and storing data provide studies and acquisition of patterns applying machine learning and artificial intelligence methods, either supervised or unsupervised. The higher capacity of data transmission offered by 5G and the lower latency to be achieved with the advent of 6G will provide the user the control of their energy costs. Among these processes, a few examples are load disaggregation (LAOUALI *et al.*, 2022) and anomaly detection in home appliances (HIMEUR *et al.*, 2021b). Such solutions fit inside the context of the Home Energy Management System.

The Home Energy Management System (HEMS) architecture consists of the integration of hardware and software to monitor and improve the efficiency of the energy consumption of a residence, allowing the user to influence the decisions (MAHAPATRA; NAYYAR, 2019). The system collects features and transmits the data to the HEMS via wireless communication, employing smart outlets and smart meters. Hence, the system stores the information of interest and provides it for functionalities available in the architecture, such as the previously mentioned load disaggregation and fault detection.

Load disaggregation consists of identifying the set of devices that compose the consumption of a household at a specific period. Through the analysis of features such as current and voltage (YANG *et al.*, 2020), or power measurements, either active or reactive (ALCALA *et al.*, 2017), the user is informed regarding the appliances included in the group that consumes the total energy during a given instant of time. Based on this information, the user can identify which devices are consuming more energy according to the time, resulting in better financial control for the user.

Another relevant factor regarding the idea of sustainability relates to the detection of faults in household appliances. Identifying equipment consuming more power than usual, obsolete, or presenting anomalous behavior results in the reduction of energy consumption and the improvement of the quality of life of the customer. However, due to the stochastic and dynamic nature of the faults, it is difficult to identify and characterize the failures. Consequently, the task is highly complex and usually strategically restricted to a specific group of devices (FERNANDES *et al.*, 2020).

This thesis introduces approaches with potential application in home management systems, presenting them in the form of adapted reproduction of articles produced during the research project in collaboration with Companhia Paranaense de Energia and the Instituto de Pesquisas Eldorado. The thesis is structured as follows: Chapter 2 provides a brief introduction to the techniques and methods applied in this work. Chapter 3 presents the load disaggregation method, published in the Institute of Electrical and Electronics Engineers (IEEE) Access magazine in 2021, Chapter 4 presents the anomaly detection approach, under review in the IEEE Transactions on Consumers Electronics magazine.

2 Brief Introduction of the Technical Concepts

In this Chapter, the concepts used in this work are discussed. The objective is to clarify the use of the methods in the tasks in which they were employed. In Chapter 3, the proposed disaggregation approach uses machine learning techniques as the decision-making system following the acquisition of patterns via Principal Component Analysis as a clustering process. In Chapter 4, the proposed anomaly detection system uses the Wavelet transform as the primary tool. Based on the level 1 coefficient of detail, it is possible to verify the duty cycle of refrigeration appliances and, using this information, analyze if a malfunction occurs, and estimate the beginning and end instants of the problem.

2.1 Load Monitoring

To apply any technique to improve the efficiency of energy usage in a household, it is essential to monitor the consumption of the residence. In this context, there are different methods to accomplish this task, and the two most common are Non-Intrusive Load Monitoring (NILM) and Intrusive Load Monitoring (ILM). The use of load monitoring techniques is restricted to the method employed in each of the databases used in the thesis, Reference Energy Disaggregation Dataset (REDD), Rainforest Automation Energy Dataset (RAE), and Personalized Retrofit Decision Support Tools for UK Homes using Smart Home Technology (REFIT).

The NILM method consists of one smart meter positioned at the residence. In other words, there is no requirement for access inside the house to collect information. All quantities gathered indicate the values consumed by all appliances, i.e., the set of feature measurements for the aggregated values. Therefore, if additional information regarding energy consumption is desirable solely via monitoring methods, it is necessary to access the residence.

For this purpose, the ILM method exists. Here, several meters (named smart outlets) are installed inside the residence in addition to the smart meter. Hence, it is possible to collect information from each appliance individually and the aggregate, enabling the user to identify which devices are consuming the most energy or requiring maintenance. However, such a method is expensive and directly impacts privacy, motivating studies regarding alternative approaches to obtain information regarding energy consumption without accessing the household (ZOHA *et al.*, 2012).

An alternative is directly related to load disaggregation. Based on the acquisition of patterns collected in monitored residences where monitoring is intrusive, the method can be applied in a household lacking intrusive measurement, providing the information of interest. Therefore, despite the non-intrusive approach, it is possible to acquire the knowledge to determine how to improve consumption efficiency (ZHU; LU, 2014).

However, regarding the detection of anomalies, the monitoring is supposed to be intrusive due to the need to measure the required magnitudes directly from the appliance. As the fault resides in the waveform of the collected feature, adding to the complexity of the task when intending to detect the malfunction from the aggregated magnitudes (HIMEUR *et al.*, 2021b).

2.2 Machine Learning Techniques

One of the most famous definitions of Machine Learning was introduced by Tom Mitchell, who states that *"A computer program is said to learn from experience E with respect to some task T and some performance measure P, if its performance on T, as measured by P, improves with experience E"* (MITCHELL, 1997). In other words, Machine learning is an evolving field of computer algorithms designed to emulate human intelligence by learning and understanding from the environment (DIETTERICH, 1990).

With the increase in computational power, the volume of data available for analysis has increased significantly, establishing a challenge for humans to deal successfully with all of this data. In this perspective, machine learning gained room since, in addition to its capability to handle large volumes of information, a large amount of data allows the learning to achieve the desired success in the stages of interest.

The learning of a system consists of three fundamental choices: structure, adaptation criterion, and optimization algorithm. The structure of a method consists of several parameters, which require adjustment according to a specific objective. All aspects related to the problem, such as causality, linearity, and the presence of memory, must be considered. The adaptation criterion relates to the amount of *a priori* information regarding the desired behavior for the system. It is the mathematical expression of the target to be achieved during the learning process, reflecting the level of success achieved by the model. Finally, the optimization algorithm is responsible for determining the optimal parameters of the system. Essentially, it establishes the mathematical operations of minimization or maximization of the functions of interest (BISHOP, 2006).

The application of machine learning techniques intends to identify patterns, feature extraction, and a decision-making system. The identification of patterns is relevant

to recognize the behavior of each appliance to determine which devices are connected to the smart outlets. Feature extraction is a compact representation of the observed data preserving most of the information or structure in the original space (DING *et al.*, 2012). The decision-making system corresponds to the classification process. Here, the system endeavors to assign each data to a label corresponding to an existing class in the case (DUDA *et al.*, 2001).

Pattern identification is employed in Chapter 3, considering Principal Component Analysis (PCA) as a tool to perform the task. Feature extraction is used in Chapters 3 and 4, where the PCA and Wavelet transform methods are applied, respectively. Finally, in Chapter 3, the decision-making systems are consolidated solutions in the machine learning field, such as k-Nearest Neighbor (k-NN), Decision Tree, and Random Forest.

2.2.1 Principal Component Analysis

Principal Component Analysis is a widespread technique for dimensionality reduction, feature extraction, data compression, and data visualization (Jolliffe, 2002). It was first introduced by Hotelling and Karhunen-Loève (HOTELLING, 1933), initially applied in the field of statistics, and subsequently extended to other areas.

Considering the data

$$\mathbf{X} = \begin{bmatrix} \mathbf{x}_1^T \\ \vdots \\ \mathbf{x}_N^T \end{bmatrix} \in \mathbb{R}^{N \times K}$$

where $\mathbf{X}$ is a matrix of available data with $\mathrm{E}\{\mathbf{x_i}\} = 0$, where $1 \leq i \leq N$.

The objective involves finding the linear projection of the data into a new space, with reduced dimension, containing maximum variance, i.e. maximum energy. These projections are named principal components, denoted by $\mathbf{y}$, and the directions are defined through $\mathbf{W} = [\mathbf{w}_1 \dots \mathbf{w_M}] \in \mathbb{R}^{K \times M}$, where $\mathbf{w}_i \in \mathbb{R}^{K \times 1}$ and $\mathbf{W}^T\mathbf{W} = \mathbf{I}$. From this property, the columns of $\mathbf{W}$ represent an orthonormal basis in the subspace of dimension M. $\mathbf{y}$ is defined by

$$\mathbf{y} = \mathbf{W}^T\mathbf{x}. \tag{2.1}$$

The $\mathbf{W}$ matrix is determined by PCA, seeking to achieve the lowest mean square error in the reconstruction of the original data. This error is expressed as

$$J = \mathrm{E}\{||\mathbf{x} - \hat{\mathbf{x}}||^2\} \tag{2.2}$$

$$= \mathrm{E}\{\mathrm{tr}\left((\mathbf{x} - \hat{\mathbf{x}})(\mathbf{x} - \hat{\mathbf{x}})^T\right)\} \tag{2.3}$$

where one can define $(\mathbf{x} - \hat{\mathbf{x}}) = \mathbf{e}$, where $\hat{\mathbf{x}} = \mathbf{W}\mathbf{W}^T\mathbf{x}$. Therefore, one can determine

$$\mathbf{e}\mathbf{e}^T = \begin{bmatrix} e_1e_1 & \dots & e_1e_K \\ \vdots & \ddots & \vdots \\ e_Ke_1 & \dots & e_Ke_K \end{bmatrix} \tag{2.4}$$

where the sum of the elements of the main diagonal is

$$\mathrm{tr}(\mathbf{e}\mathbf{e}^T) = \sum_{i=1}^{K} e^2 \tag{2.5}$$

$$= \mathbf{e}^T\mathbf{e} \tag{2.6}$$

$$= \mathbf{e}^2. \tag{2.7}$$

Therefore, eq. (2.3) is rewritable as

$$J = \mathrm{E}\{\mathrm{tr}(\mathbf{x}\mathbf{x}^T - \mathbf{x}\hat{\mathbf{x}}^T - \hat{\mathbf{x}}^T\mathbf{x} - \hat{\mathbf{x}}^T\hat{\mathbf{x}}^T)\}. \tag{2.8}$$

Based on the properties of two matrices $\mathbf{A} \in \mathbb{R}^{K\times M}$ and $\mathbf{B} \in \mathbb{R}^{M\times K}$:

- $\mathrm{tr}(\mathbf{A}) = \mathrm{tr}(\mathbf{A}^T)$;
- $\mathrm{tr}(\mathbf{A}\mathbf{B}) = \mathrm{tr}(\mathbf{B}\mathbf{A})$,

and the autocorrelation of the data, $\mathbf{R}_\mathbf{x} = \mathrm{E}\{\mathbf{x}\mathbf{x}^T\}$, eq. (2.8) can be rewritten as

$$J = \mathrm{tr}(\mathbf{R}_\mathbf{x}) - \mathrm{tr}(\mathbf{W}^T\mathbf{R}_\mathbf{x}\mathbf{W}). \tag{2.9}$$

PCA seeks to minimize the function J in the terms depending on the matrix W. Therefore

$$\min J = \max\ \mathrm{tr}(\mathbf{W}^T\mathbf{R}_\mathbf{x}\mathbf{W}). \tag{2.10}$$

In terms of the number of principal components, the $k-th$ principal component of $\mathbf{x}$ is determined by y_k, and $\mathbf{w}$ is the eigenvector of the autocorrelation matrix $\mathbf{R}_\mathbf{x}$ associated with the $k-th$ largest eigenvalue, where the matrix $\mathbf{R}_\mathbf{x}$ is defined by

$$\mathbf{R}_{\mathbf{x}} \approx \frac{1}{N} \sum_{i=1}^{N} \mathbf{x}_n \mathbf{x}_n^T \tag{2.11}$$

Based on the percentage of the total variance of data, it is possible to establish the trade-off between the approximation error and the compression degree achieved. Therefore, the number of selected principal components reflects the balance between these variables once the approximation error decreases by selecting more components and the compression degree increases using fewer components. Hence, the variance percentage minimizes the approximation error and maximizes the compression degree, evaluating the number of principal components required to achieve the balance (Jolliffe, 2002).

2.2.2 k-Nearest Neighbor

The k-Nearest Neighbor (kNN) is a supervised learning algorithm used for classification and regression. It is a non-parametric method. In other words, no model exists to be adjusted, and no assumptions performed concerning the input data that makes a decision for a new input pattern based on the information of its nearest neighbors in the input feature space.

Distance metrics and the decision of the hyperparameter value, k, turn the algorithm into a method called competitive learning because it is the data in question (the k nearest neighbors) that compete to influence the classification result whenever distance calculated for each new input data. The most common distance metric used in kNN refer to Minkowski distance of p order and is given by

$$d(x; y) = \left(\sum_{i=1}^{K} |x_i - y_i|^p \right)^{\frac{1}{p}}, \tag{2.12}$$

where p is parameter adjustable. Being $p = 1$ the Manhattan distance, and $p = 2$ the Euclidean distance.

The kNN is considered an algorithm named lazy learning, which means the algorithm does not use the training data to create a model until the time it is necessary to perform the prediction. It reduces the training process complexity, however, the testing phase increases computational effort since all data requires to be stored and consulted to identify the nearest neighbors. Therefore, the algorithm performs better when considering fewer features (ALTMAN; Hart, 1992).

Figure 2.1 depicts an example of kNN classification, considering the hyperparameter value $k = 3$. The new pattern, in yellow, is classified based on the votes of the

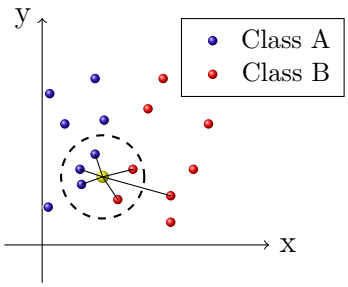


Figure 2.1 – kNN classification for $k = 3$.

nearest neighbors in the attribute space. As highlighted in the Fig. 2.1, the new pattern is assigned as part of the class A.

2.2.3 Decision Tree

The Decision Trees algorithm is another supervised learning strategy commonly used for classification and regression problems. The tree is depicted as a non-directional graph where two vertices connected by a single path.

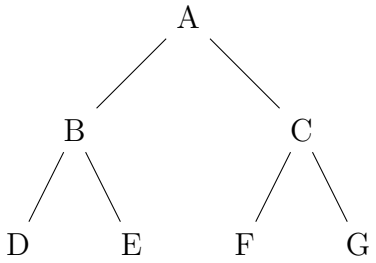


Figure 2.2 – Decision Tree.

Figure 2.2 depicts a tree has a root node, node A, from where the complete decision process starts. In this way, different values of the data attributes generate branches that, when they reach a node, named leaf (B, C, ..., G), they assign to a class. Each attribute yields a binary response leading to each node leaf a decision about death or survival.

Some advantages of using trees include the fact that their structure is capable of reflecting relationships between data, represents hierarchies, provides efficient insertion and research, and is flexible (KAMIńSKI *et al.*, 2028).

2.2.4 Random Forest

Random Forest is a supervised learning method introduced by (BREIMAN, 2001), inspired by earlier work produced by Amit and Geman (AMIT; GEMAN, 1997).

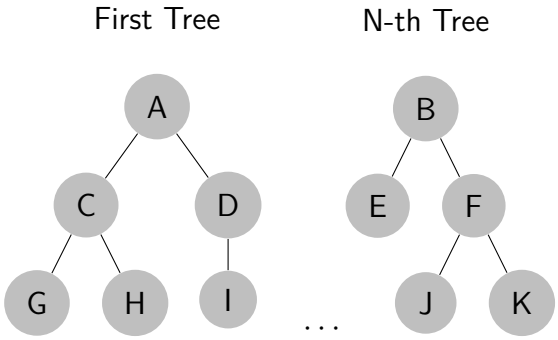


Figure 2.3 – Random Forest.

In his work, Breiman defines the classifier as a combination of predictor trees, as illustrated in Figure 2.3, where each tree depends on the value of a random vector sampled independently and with the same distribution as all other trees in the forest.

As described in (CUTLER *et al.*, 2011), the process consists of finding the prediction function $f(\mathbf{x})$, where $\mathbf{x}$ is an m-dimensional random vector defined as $\mathbf{x} = (x_1, \ldots, x_m)$, for the prediction of $\mathbf{y}$, determined via the loss function, $L(\mathbf{y}, f(\mathbf{x}))$. Formally, it is described as

$$L(\mathbf{y}, f(\mathbf{x})) = \begin{cases} 0, & \text{if } \mathbf{y} = f(\mathbf{x}) \\ 1, & \text{otherwise.} \end{cases}$$

Therefore, the process of minimizing the expected value of $L(\mathbf{y}, f(\mathbf{x}))$ is described via Rule of Bayes.

$$f(x) = \arg\max_{\mathbf{y} \in \mathcal{Y}} P(\mathbf{y} = y | \mathbf{x} = x), \tag{2.13}$$

where $\mathcal{Y}$ denotes the set of all possible values of $\mathbf{y}$ and P is the probability function.

2.3 Wavelet Transform

The well-known Fourier Transform is the most popular method to analyze signals in the frequency domain (OPPENHEIM, 1999). Nevertheless, the Fourier Transform is a suitable method for stationary signals since it considers the frequency components for the entire period over time. In the case of transient (nonstationary) signals, such as frequency variations over different periods, it requires a decomposition that includes a time-frequency perspective.

The time-frequency perspective is suitable for analyzing data from different areas of application. For example, the analysis of vital signals in the biomedical area,

the study of stocks in the financial market, the analysis of power signals in the energy sector, and others. A typical example of such decomposition is the Short-Time Fourier Transform (STFT), computed via the transform of segments of the data in different time intervals, named windows, according to (2.14) (MALLAT, 1999).

$$S_w(\tau, \omega) = \int_{-\infty}^{\infty} f(t) g^*_{\tau,\omega}(t) dt \qquad \tau \in \mathbb{R}, \tag{2.14}$$

where $g^*_{\tau,\omega}(t)$ is the complex conjugate of $g_{\tau,\omega}(t) = g(t-\tau)e^{j\omega t}$ and $g(\cdot)$ is a real, non-zero, symmetric window function. The length of the window used in STFT relates to the time-frequency resolution of the signal. I.e., it determines if it provides better frequency resolution or better time resolution. A wide window gives better resolution in frequency and low resolution in time. On the other hand, a narrow window provides better time resolution and lower frequency resolution. The STFT is constrained to a fixed resolution for the entire time-frequency domain of the signal. This limitation is a reason for the introduction of the Continuous Wavelet Transform (CWT) and Multiresolution Analysis (MRA) methods that provide good resolution of time for high-frequency events and good resolution of frequency for low-frequency events. Therefore, the most suitable choice for many practical signals (MALLAT, 1999). The CWT is derived by replacing the STFT kernel $g(\tau, \omega)$ by $\psi(a, b)$ from the Wavelet

The time-frequency perspective is suitable for analyzing data from different areas of application. A typical example of such perspective is the CWT that provide good resolution of time for high-frequency events and good resolution of frequency for low-frequency events. Therefore, the most suitable choice for many practical signals, e.g. the analysis of vital signals in the biomedical area, the study of stocks in the financial market, the analysis of power signals in the energy sector, and others (MALLAT, 1999). The CWT is defined by

$$W_\psi(a, b) = \int_{-\infty}^{\infty} f(t) \psi^*_{a,b}(t) dt \qquad (a, b) \in \mathbb{R}_+ \times \mathbb{R}, \tag{2.15}$$

where $f(t)$ is the signal of interest, $\psi_{a,b} = \frac{1}{\sqrt{a}}$ is a source of the branch wavelet and, $\psi(\cdot)$ is a function named mother wavelet which is zero mean, centered at $t = 0$, and usually time-limited.

When analyzing practical signals, i.e., discrete signals of finite length, discrete versions of the transforms are employed. Similar to the Discrete Fourier Transform (DFT), there is also the Discrete Wavelet Transform (DWT) for the Wavelet transform. The concept of DWT defines itself in the mother wavelet function $\psi(\cdot)$. The CWT retains all the information regarding the transformed signal when sampled at a given discrete subset of the time-scale field. For this purpose, one considers (2.15) instead of the pair

of variables (a, b). The pair $(a^{-j}, a^{-j}bk)$ are generally set $a = 2$ and $b = 1$, respectively, where the support is given by $(a^{-j}, a^{-j}bk)$, according to (2.16)

$$\psi_{2^{-j},2^{-j}k}(t) = \frac{1}{\sqrt{2^{-j}}}\psi(\frac{t - 2^{-j}k}{2^{-j}}) = 2^{j/2}\psi(2^j t - k). \tag{2.16}$$

Hence, the coefficients of detais of the DWT are given by

$$d_{m,n} = \int_{-\infty}^{+\infty} f(t)\psi_{m,n}(t)dt, \tag{2.17}$$

where d represents the coefficients of details in scale and position described by the indices m and n (ADDISON, 2005). Figure 2.4 depicts the structure of the DWT decomposition tree.

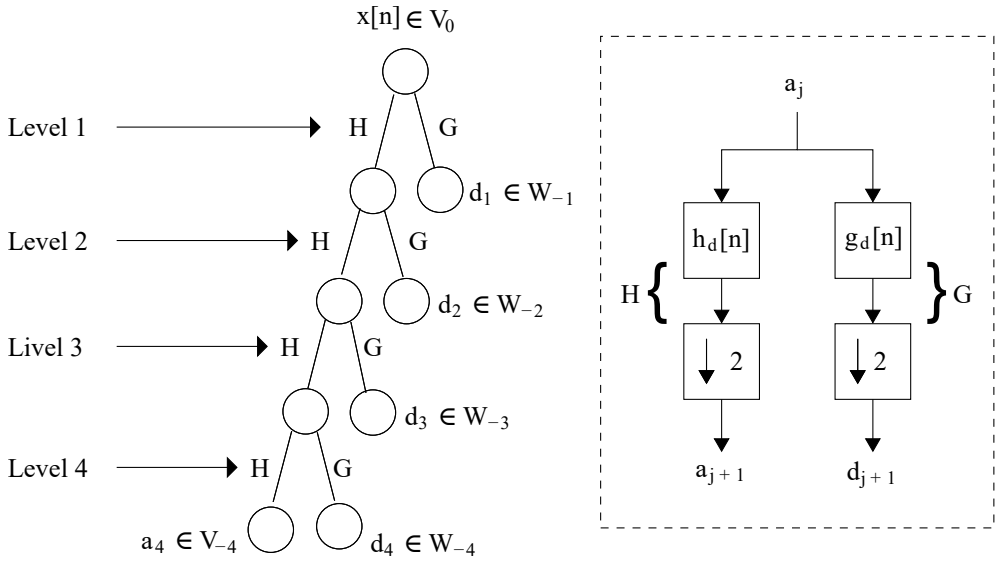


Figure 2.4 – Structure of the DWT decomposition tree.

In this novel form of the transform, one denotes $w_{j,k}$ as the Wavelet coefficients representing the DWT values corresponding to the correlation between the analyzed signal $f(t)$ and the Wavelet function $\psi_{a,b}$ at the specific points $(a, b) = (2^{-j}, 2^{-j}k)$ as follows

$$w_{j,k} = \langle f(t), \psi_{j,k} \rangle. \tag{2.18}$$

Therefore, to reconstruct the signal, the synthesis equation is

$$f(t) = \sum_{j,k} w_{j,k}\psi_{j,k} \qquad j, k \in \mathbb{Z}. \tag{2.19}$$

One sufficient condition for the perfect reconstruction of the signal $f(t)$ via (2.19) lies in the requirement that functions $\psi_{j,k}$ form orthonormal bases in the space

$L^2(\mathbb{R})$. The feasibility of constructing a function $\psi(\cdot)$ such that $(\psi_{j,k})_{j,k}$ forms an orthonormal basis directly ties in a few Wavelet mother conditions (MALLAT, 1999). Several bases were constructed over the years by researchers. Among them is the minimal support basis known as Daubechies (DAUBECHIES, 1992).

The orthogonality and MRA also implies the existence of the sequences h_k and g_k satisfying the following identities

$$h_k = \langle \varphi_{0,0},\, \varphi_{-1,k} \rangle \;\text{ so that }\; \varphi(t) = 2\sum_{k\in\mathbb{Z}} h_k \varphi(2t-k), \tag{2.20a}$$

$$g_k = \langle \psi_{0,0},\, \varphi_{-1,k} \rangle \;\text{ so that }\; \psi(t) = 2\sum_{k\in\mathbb{Z}} g_k \varphi(2t-k). \tag{2.20b}$$

where $\varphi(t)$ and $\psi(t)$ are referred scaling equation and Wavelet equation, respectively. The identity (2.20a) is the refinement equation for the father wavelet φ.

The two equations (2.20a) and (2.20b) form the basis for the Fast Wavelet Transform (FWT).

Similar to the Fast Fourier Transform (FFT), the Fast FWT requires less computational effort for execution, and for an input signal of length k, the complexity of $O(k)$ is lower than that of the FFT, $O(k; ln\ k)$.

The FWT represents a theorem developed by Mallat (MALLAT, 1999) that proves the existence of the expressions for the decomposition process

$$a_{j+1}[p] = \sum_{k=-\infty}^{\infty} a_j[k]\; h[k-2p] = a_j * \bar{h}[2p], \tag{2.21a}$$

$$d_{j+1}[p] = \sum_{n=-\infty}^{\infty} a_j[k]\; g[k-2p] = a_j * \bar{g}[2p] \tag{2.21b}$$

where $*$ is the convolution operation and,

$$\begin{aligned} a_j[p] &= \sum_{k=-\infty}^{\infty} a_{j+1}[k]h[p-2k] + \sum_{k=-\infty}^{\infty} d_{j+1}[k]g[p-2k] \\ &= \check{a}_{j+1} * h[p] + \check{d}_{j+1} * g[p]. \end{aligned} \tag{2.22}$$

The equations (2.21a) and (2.21b) indicate that the decomposition process of a signal $f[n]$ to a transform level ($j_0 = 1$) can be performed via filtering and decimation.

2.4 Evaluation Metrics

Throughout data classification and prediction, it is possible to evaluate the efficiency of the proposed approach. In Machine Learning and statistical classification

problems, the Confusion Matrix is widespread. The Confusion Matrix, also named Error Matrix, consists of a table where each row represents the instances corresponding to the actual classes and the columns represent the instances corresponding to the predicted classes, or vice-versa (FAWCETT, 2006).

Figure 2.5 depicts the confusion matrix for two classes, Positive and Negative. Nevertheless, it is scalable to a multiclass as the problem requires. Some parameters are available from the confusion matrix, where:

- True Positive (TP): Corresponds to predictions of positive classes that actually are members of the positive class;
- True Negative (TN): Corresponds to predictions of the negative class that are actually members of the negative class;
- False Positive (FP) - Type Error 1: Corresponding to predictions of positive values, although pertaining to the negative class;
- False Negative (FN) - Type Error 2: Corresponding to predictions of negative values, although pertaining to the positive class.

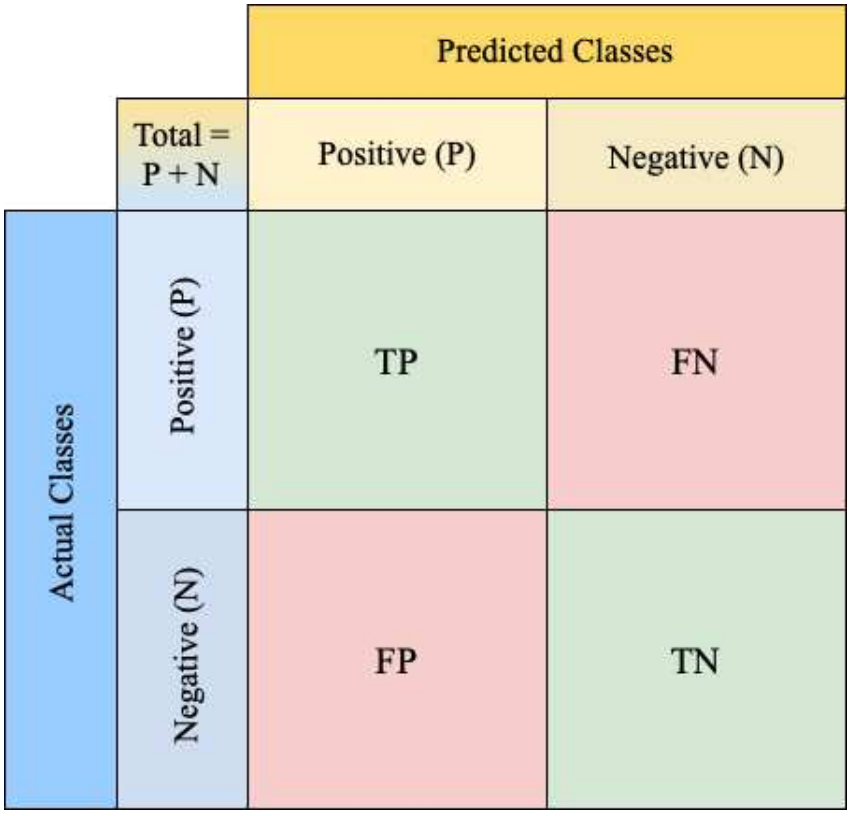


Figure 2.5 – Example of Confusion Matrix.

These parameters allow calculating metrics that provide validation of the approach involved via qualitative terms. These metrics are Accuracy, Precision, Recall, and F-Measures.

2.4.1 Accuracy

Accuracy is an overall measure of performance, which corresponds to how close a predicted sequence is to its actual values. The equation corresponding to the calculation of accuracy based on the values extracted from the confusion matrix is given by

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}}. \tag{2.23}$$

2.4.2 Precision

Precision is defined as the measure of correct predictions considering especially the data referring to the positive class. The precision equation is given by

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}. \tag{2.24}$$

2.4.3 Recall

Recall corresponds to the number of predictions attributed to the positive class compared to the total number of elements of the positive class in the dataset. The recall equation is given by

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}. \tag{2.25}$$

2.4.4 F-Measure

Strictly speaking, neither precision nor recall provides complete information concerning the performance of the classifier on the task. F-Measure is a metric combining precision and recall to offer an overview concerning the performance of the classifier. Also named F-Score or F1-Score, a such metric is the most widely used metric in classification problems involving imbalanced data. The equation for F1-Score is given by

$$\text{F1-Score} = 2\left(\frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}}\right). \tag{2.26}$$

Based on the nature of the problem being addressed, each of the metrics may reveal more meaningful information to report the results (GRANDINI *et al.*, 2020a).

3 Load Disaggregation Based On Time Window for HEMS Application

This Chapter is an adapted reproduction of the article:

D. A. M. Lemes, T. W. Cabral, G. Fraidenraich, L. G. P. Meloni, E. R. Lima, and F. B. Neto, "Load Disaggregation Based On Time Window for HEMS Application", IEEE Access, vol. 9, pp. 70746 - 70757, May 2021. Impact Factor: 3.367

This Chapter presents an approach for the load disaggregation task. This procedure consists of time windows, seeking to identify the appliances in operation during this specific window, using pattern identification via Principal Component Analysis. The decision-making system consists of Machine Learning methods, such as k-Nearest Neighbor, Decision Tree, and Random Forest. The results are described in terms of comparative performance and reveal the efficiency of the approach to perform the task, outperforming other solutions present in the literature, according to Table 3.4.

Abstract

This work investigates the efficiency of the process of load disaggregation, considering only the values of active power. To perform the task, we use data collected from the NILM measurement method, presented in the RAE and REDD database. A strategy of assigning labels using combinations of equipment in use, by status ON/OFF, and also by choosing an appropriate temporal data window is discussed. Also, the performance of very well-known machine learning algorithms such as kNN, Decision Tree, and Random Forest are evaluated. The results show a very efficient and low computer complexity strategy presenting values of F1-Score above 95%, for RAE and REDD database. As presented in Table 3.4, the proposed approach presents the highest F1-Score, compared to other methods in the literature, considering all appliances in the REDD database. The greatest benefit of the approach consists in the possibility of applying the disaggregation process in a household without smart outlets, under the restriction that the training and test houses hold identical or similar appliances.

3.1 Introduction

Electricity is one of the primary sources of energy used globally, especially in densely populated regions. According to UN (United Nations) investigations, data published in the World Urbanization Prospects, inform an increase in the number of people residing in urban areas around the year 2050 (NATIONS, 2014). Hence, the supply of electricity may be an issue if it does not develop to the point of accompanying the growing tendency in demand of the sector. For this reason, the subject received more attention in academic and industrial environments seeking to optimize the processes of generation and distribution.

One of the issues tackled in this context relates to the seasonality of residential electricity use. In other words, to endeavor to know which periods of the day have the highest consumption and which are those where consumption is the lowest. A reliable tool to assist the data treatment, allowing the acquisition of relevant information is the load disaggregation, through the device named HEMS.

The HEMS architecture defines a system that consists of hardware and software linked and integrated to monitor energy consumption, providing feedback, and improving control over household devices. It intends to control the expenditure of electrical energy usage, performing the role of a modern load meter. That is, HEMS intelligently monitors and adjusts electricity consumption by smart meters, devices, and outlets, resulting in efficient use management, through a wireless communication network (MAHAPATRA; NAYYAR, 2019). Knowing that it is a system that attempts to optimize energy management, the HEMS architecture is a foundation for the load disaggregation in houses.

The load disaggregation method consists of an individualized measurement of the equipment allowing the mathematical modeling of the energy consumption of each installed system, according to the variables that influence the devices. Therefore, information as the consumption by equipment, defective or obsolete appliances, time of use, and the individual consumption of each electrical load will be available. Hence, it is possible to understand electric power consumption, as this factor is essential for energy efficiency management and expenditure management.

The most common types of measurement are ILM and NILM methods as described in Section 2.1. Among all types of measurement, the NILM is regarded as the most promising in literature (Azzini *et al.*, 2014) and widely studied for application in HEMS and Ambient Assisted Living (AAL) (RUANO *et al.*, 2019). In the non-intrusive measurement mode, the data is collected from the electrical measurement panel outside the house or building, avoiding the installation of any equipment inside the property (ZOHA *et al.*, 2012). Diverse NILM methodologies are studied to allow the production

of accurate and affordable tools, which are employed for monitoring (ZEIFMAN; ROTH, 2012), supporting the process of disaggregation, and estimation of energy consumption (GIRI; BERGÉS, 2015) including data prediction (Kampouropoulos *et al.*, 2013).

From the aggregate measurement of the loads, as in NILM, it is possible to perform the disaggregation process estimating the power consumed by equipment. The most commonly used methods for this purpose are based on event detection and load identification. Event detection is typically defined as the ON/OFF switching of electrical devices (MASSIDDA *et al.*, 2020). In (MORADZADEH *et al.*, 2020), the authors apply PCA to perform the task based on a dimensional reduction to identify the patterns of power consumption of the devices. In (HUANG *et al.*, 2021), the authors employ PCA to reduce the algorithm complexity using current as initial features. In (ZHU; LU, 2014), PCA is employed as dimensionality reduction technique for substation load profile disaggregation. Yang et al. proposed an approach based in V-I trajectory and the PCA is applied for feature extraction (YANG *et al.*, 2019). In (ALCALA *et al.*, 2017), the authors consider trajectories of active, reactive, and distortion power (Power Quality Division (PQD)) to achieve general models of the device classes using PCA. Methods based on Markov chains are also used. In (MENGISTU *et al.*, 2018), the events are extracted from the aggregate active and reactive powers, and then, the authors used Factorial Hidden Markov Model (FHMM) for disaggregation. In (KONG *et al.*, 2016), the appliances are modeled as Hidden Markov Model (HMM) and the disaggregation process is implemented via Segmented Integer Quadratic Constraint Programming (SIQCP).

Other models applied recently for event detection are Decision Tree and Long Short-Time Memory (LSTM) yielding detection accuracy over 92% (LE; KIM, 2018). The methods that use LSTM, Dense Neural Network (DNN), and Convolutional Neural Network (CNN) are characterized as Deep Learning (DL) methods. In (YANG *et al.*, 2020), the authors used patterns of the curve of the current, and the identification process is based on CNN. In this case, the features need to be converted into 2D images to use CNN. In (KASELIMI *et al.*, 2019), the authors used a hybrid proposal, where a recurrence of CNNs is employed. In this case, active power, reactive power, current and apparent power must be used. Although there are many modern methods, such as DL, the problem of disaggregation is still a challenge, composed of many issues and difficulties.

One of the main difficulties in load disaggregation is the development of an efficient method to extract information about operating appliances based exclusively on aggregate active power. The need for prior knowledge of specific features of each device for the pattern formation and the lack of this information when considering the signature of the aggregate active power turn the problem extremely complex.

To avoid the problem, we consider an approach based on the creation of time

windows where every element on the window consists of a combination of operating devices represented by the respective active power measured by the smart meter. The previously required information is exclusively applied for the label assignment process. Therefore in the stages related to the disaggregation and classification tasks, only the aggregated measurements are considered. Results achieved via simulations reveal the robustness of the approach regarding accuracy, precision, recall, and F1-Score for all classifiers employed.

The main contributions of this work are the following:

1. very low complexity approach with high efficiency on the identification of the appliances. The proposed method achieves the highest F1-Score, $95,51\%$, compared to other algorithms in the literature (see Table 3.4), when applied to the REDD dataset;
2. no need for pre-processing of the data;
3. no need for re-training process at the disaggregation stage;
4. the method is applicable in a household without smart outlets regarding the restriction that the houses need to contain the same set of devices.

The Chapter is organized as follows: in Section 3.2 describes the NILM technique. In Section 3.3 briefly describes the proposed approach. In sections 3.4 and 3.5 present the Load ID and time window disaggregation, respectively. In Section 3.6 the evaluation metrics are explained, section 3.7 describes the databases used to test the algorithm. Section 3.8 presents a discussion regarding the results achieved by the simulations. Finally, Section 3.9 presents the conclusions concerning the efficiency of the approach.

3.2 Non-Intrusive Load Monitoring

The measurement method based on NILM architecture, developed in the decade of 1980, aims to collect electrical data for information processing. Non-intrusive implies no physical access required to the interior of the electrical installation. The measurement is received from a smart meter, installed in the main branch of the residence (KOLTER; JOHNSON, 2011). The aim is to detect variations in the ON\OFF status of devices through differences in total consumed power. From the load, the aggregate power is disaggregated through machine learning and pattern recognition algorithms (JIANG *et al.*, 2011).

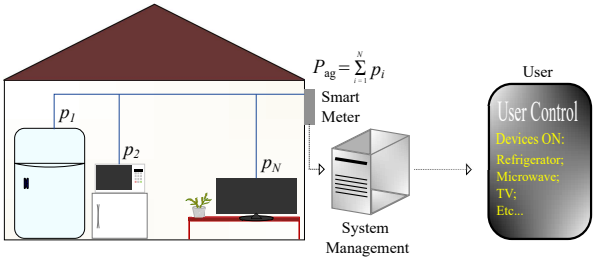


Figure 3.1 – Example of Non-Intrusive Load Monitoring scheme.

As shown in Figure 3.1, we consider a set of N devices, whose individual power per equipment is p_i, where $1 < i < N$. The total aggregate power, P_{ag}, can be defined as (SCHIRMER; MPORAS, 2019)

$$P_{\text{ag}} = \sum_{i=1}^{N} p_i. \tag{3.1}$$

All monitored loads can be classified considering regime operation according to power demand. The power demand can be a flat, variant, and flat or variant, where the type of load can be categorized as a single state (ON = OFF), continuous variation (ON $\neq$ OFF), and multiple states (multiple events), respectively.

This type of information is available through the signature of the load, achieved by monitoring, which is the key method for identifying the loads. Further, the load signatures yield the active and reactive power level, current waveforms and harmonic components, admittance, start transient, and others (WU *et al.*, 2020).

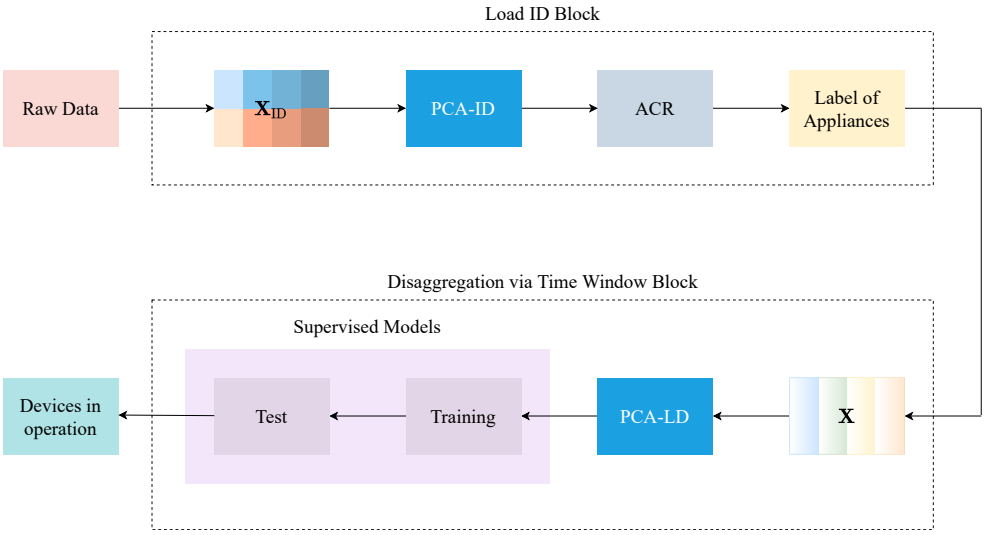


Figure 3.2 – Flowchart of the proposed approach.

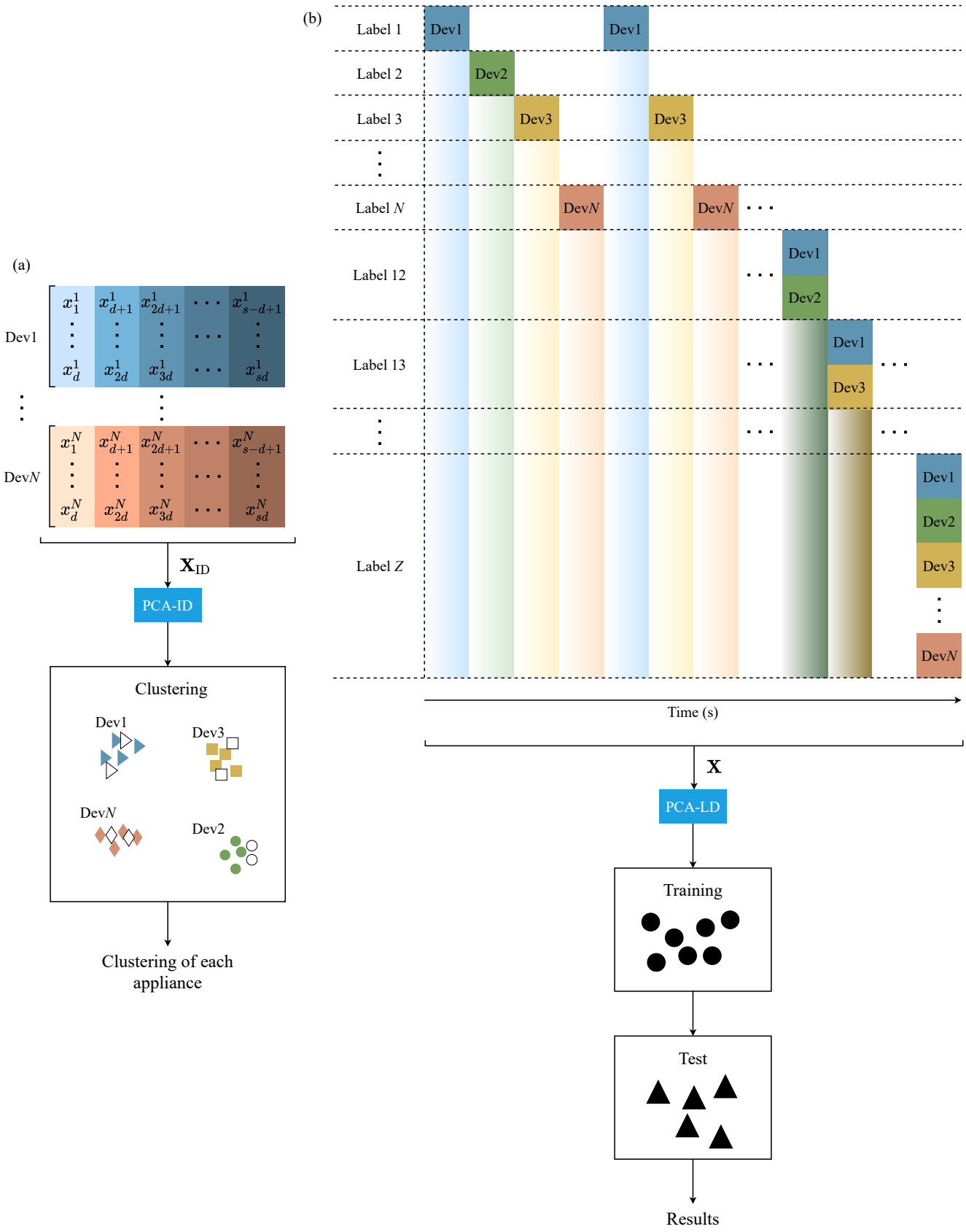


Figure 3.3 – (a) Load ID stage and (b) Time Window scheme.

3.3 Overview Of The Proposed Method

To summarize, the proposed method consists of determining the appliances in operation in a household. The devices are identified via the Load ID process. This procedure initiates by formatting the $\mathbf{X}_{\text{ID}}$ matrix, reading the data from the smart outlets. In this case, the labels of each device are obtained via Principal Component Analysis for Identification (PCA-ID) decomposition. Once the equipment is identified, the disaggregation is performed via time window. The data from the main circuit are formatted according to matrix $\mathbf{X}$, and it flows to Principal Component Analysis for Load Disaggregation (PCA-LD) decomposition for feature extraction. Next, the supervised models are used to classify the devices in operation during the time window. Finally, the proposed approach presents low computational complexity, although considering many devices from different databases. The complexity reduces since the method uses combinations of appliances with no requirement of previously knowing all of them in the training stage. Also, it is the method that provides the one of the highest F1-Score for all devices and a significant amount of samples, as described in Table 3.4. Here, the proposed method was compared with approaches in the literature that present low computational complexity. Figures 3.2 and 3.3 depict the flowchart and the scheme of the blocks in the proposed system, respectively. Further, in Section 3.5, Algorithm 3.1 and Algorithm 3.2 fully describe the stages of Load ID and disaggregation, respectively.

3.4 Principal Component Analysis

The first stage to perform the disaggregation process consists of identifying the load connected to the smart outlet. This stage, named Load ID (depicted in Figure 3.3 (a)) consists of the pattern recognition from the appliance of the active power and works as a previous process to the disaggregation task.

For load identification, the method is based on the approach suggested by (MORADZADEH *et al.*, 2020) which employed PCA as a technique to extract the necessary attributes for equipment clustering. The proposed method differs slightly from (MORADZADEH *et al.*, 2020) due to the formatting of the input data matrix. To format the matrix, the data considered regards to the active power measurements collected during a specific number of days. Each device is assigned to several rows, referring to the intervals described in equation (3.2) rather than a single column containing all the data equipment measurements.

$$\mathbf{X}_{\text{ID}} = \begin{bmatrix} x_1^1 & x_{d+1}^1 & x_{2d+1}^1 & \dots & x_{sd-d+1}^1 \\ \vdots & \vdots & \vdots & \dots & \vdots \\ x_d^1 & x_{2d}^1 & x_{3d}^1 & \dots & x_{sd}^1 \\ & & \vdots & & \\ x_1^N & x_{d+1}^N & x_{2d+1}^N & \dots & x_{sd-d+1}^N \\ \vdots & \vdots & \vdots & \dots & \vdots \\ x_d^N & x_{2d}^N & x_{3d}^N & \dots & x_{sd}^N \end{bmatrix}, \tag{3.2}$$

where $(\cdot)^N$ is the N-th appliance, d is the number of days for the identification process, and s is the number of samples. The dimensions of matrix $\mathbf{X}_{\text{ID}}$ are $Nd \times Nsd$.

Subsequently, we split the data into training and testing for principal component analysis for feature extraction. In this Chapter, the PCA will be used for dimensional reduction (HYVÄRINEN *et al.*, 2001) and, consequently, pattern recognition (THEODORIDIS *et al.*, 2010). In this case, the input matrix for the PCA can be organized as $\mathbf{X}_{\text{ID}}^T$, where $(\cdot)^T$ is the transpose of the matrix.

Initially, the input matrix is centered, i.e., its mean is subtracted, and then the covariance matrix is calculated as in (3.3) (HONEINE, 2011). In the sequel, the eigen-decomposition process takes place and the eigenvalues, λ_i, and eigenvectors, $\mathbf{a}_i$, for $1 \leq i \leq Nd$, are obtained.

$$\mathbf{C} = \frac{1}{Nd} \sum_{i=1}^{Nd} \mathbf{x}_{\text{ID}i} \cdot \mathbf{x}_{\text{ID}_i}^T, \tag{3.3}$$

where $\mathbf{x}_{\text{ID}_i}$ is the i-th row vector of matrix (3.2). The feature extraction is performed according to the number of principal components established as necessary to form the clusters. To set the number of principal components, m, the process employed uses the Accumulative Contributory Ratio (ACR) metric. The ACR indicates the cumulative contribution of the number of eigenvalues examined, sorted in descending order, as expressed in (3.4). Based on a previously defined threshold, the number m of eigenvalues is defined as the number of eigenvalues required for the ACR to be greater or equal than the threshold ρ.

$$\rho_m = \frac{\sum_{i=1}^{m} \lambda_m}{\sum_{i=1}^{Nd} \lambda_i} \tag{3.4}$$

By defining m, where $1 \leq m \leq Nd$, the feature extraction process is complete after selecting the columns of the eigenvectors matrix $\mathbf{A}$, associated with the eigenvalues identified as the most contributive, as described in (3.5). Equation (3.6) establishes the identified patterns of each appliance. Consequently, it is possible to identify the devices

connected to the outlets, and the disaggregation process can start. It is worth highlighting that, although load identification is a prior process to disaggregation, the approaches discussed in this paper are independent of the other.

$$\mathbf{A} = \left[\underbrace{\mathbf{a}_1 \quad \mathbf{a}_2 \quad \dots \quad \mathbf{a}_m}_{\mathbf{A_m}} \quad \dots \quad \mathbf{a}_{Nd} \right] \tag{3.5}$$

Note that the original dimension of the data is reduced. The resultant matrix $\mathbf{P}$, described in (3.6), maintains information concerning energy consumption.

$$\mathbf{P} = \mathbf{A}_m^T \cdot \mathbf{C}. \tag{3.6}$$

For the disaggregation task, another PCA decomposition is applied. In our proposed method, PCA is present in two stages: load identification and disaggregation. The entire process involved in deriving the patterns of loads appears in detail in Algorithm 3.1, where the input parameters and equations required describe the operation of the Load ID Block. In the identification stage, PCA plays the essential role of finding the individual patterns of each device, named PCA-ID. In the load disaggregation stage, as described in (3.7), the length of the window depends on two variables, the measurement frequency and the time set to form the window. In this case, the time windows might contain a large amount of data, resulting in the classification of operating appliances unfeasible. Therefore, PCA,

Algorithm 3.1 Load ID Block

Input: Raw data from the intelligent outlets, N, s, d, ρ
Output: $\mathbf{P}$ with (3.6)

Initialization:
1: first statement
Formatting $\mathbf{X}_{\text{ID}}$ *as in* (3.2)
2: *second statement*:
Apply PCA-ID
Extract the mean of $\mathbf{X}_{\text{ID}}$
Estimate the covariance matrix $\mathbf{C}$ *via* (3.3)
Extract λ *and* $\mathbf{A}$
Sort λ *in descending order*
Sort columns of $\mathbf{A}$ *according to* λ
3: *third statement*:
Apply the ACR
Calculate ρ_m *via* (3.4)
if $\rho_m \geq \rho$
Define m
end if
Select the m *columns of* $\mathbf{A}$ *via* (3.5)
return P

named PCA-LD, reduces the dimensionality of the windows and approximates patterns that present similar combinations of functioning devices.

3.5 Time Window Disaggregation

Once the loads are identified, the method aims to classify operation devices from the measurements of the main circuit of the house, as depicted in Figure 3.1. To achieve this, the algorithm establishes combinations of devices based on the ON/OFF status at a specific time. In other words, the classes are determined following the ON equipment at the measuring time. Thereby, labels are assigned to the classes, as shown the Table 3.1, where N is the number of devices, yielding $2^N - 1$ possible classes of combinations. The disaggregation stage is depicted in Figure 3.3 (b).

Table 3.1 – Labels assignment for the devices classes.

Devices ON	**Labels**
Dev1	1
Dev2	2
Dev3	3
⋮	⋮
Dev1 + Dev2	12
Dev1 + Dev3	13
⋮	⋮
Dev1+Dev2+...+Dev$N$	Z

Table 3.1 refers to the labels of the combinations of the appliances, where Z corresponds to the label $12 \dots N$. The labels are associated with the aggregated power computed as in (3.1), measured by the main circuit at the time of the measurement.

The time window is adjusted by setting a period of measurements which forms a vector with n samples. For the disaggregation task, we use the active power readings from the main circuit. The data is assigned into the vector according to the measured time, the timestamp. For example, the aggregated power value read at time t_1 is placed in the first position of the vector, the value read at time t_2 is placed in the second position of the vector, and so on until time t_l, where l is the number of samples present in the window. Hence, the first row of the $\mathbf{X}$ is set and, consequently, the first time window. The following time windows are the other vectors that compose the $\mathbf{X}$ matrix. In other words, the time window is a time interval that contains the aggregate power samples, where the operating devices are verified and can be defined as in (3.7) with dimension $\frac{n}{l} \times l$.

$$\mathbf{X} = \begin{bmatrix} P_{\mathrm{ag}_1} & P_{\mathrm{ag}_2} & \dots & P_{\mathrm{ag}_l} \\ P_{\mathrm{ag}_{l+1}} & P_{\mathrm{ag}_{l+2}} & \dots & P_{\mathrm{ag}_{2l}} \\ \vdots & \vdots & \dots & \vdots \\ P_{\mathrm{ag}_{(n-1)l+1}} & P_{\mathrm{ag}_{(n-1)l+2}} & \dots & P_{\mathrm{ag}_{nl}} \end{bmatrix} \tag{3.7}$$

where P_{ag} is the aggregate power defined in Equation (3.1).

The prior information for the disaggregation process consists of reading the measurements of the aggregate active power and smart outlets. It is important to highlight that the data of the smart outlets is used solely at the training phase and load identification. Subsequently, the assignment of labels is performed according to Table 3.1. The adjustment of the time windows is performed according to (3.7), and the extraction of features is computed using another PCA-LD (PCA for Load Disaggregation), which is used during the training and testing phases.

Algorithm 3.2 Disaggregation via Time Window Block

Input: Read raw data from the main circuit, n, l
Output: Appliances in operation
Initialization:
1: first statement
Label assignment according to table 3.1
Formatting $\mathbf{X}$ *as in* (3.7)
Dividing $\mathbf{X}$ *in training and test*
2: second statement
Apply PCA-LD
Extract the mean of $\mathbf{X}$
Estimate the covariance matrix $\mathbf{C}$ *via* (3.3)
Extract λ *and* $\mathbf{A}$
Sort λ *in descending order*
Sort columns of $\mathbf{A}$ *according to* λ
Select the m *columns of* $\mathbf{A}$ *via* (3.5)
Extract coordinates via (3.6)
3: third statement
Training of the supervised models
4: fourth statement
Label assignment of the test data according to labels attributed in train data
5: fifth statement
Extract coordinates via (3.6)
6: sixth statement
Test of the supervised methods
7: seventh statement
Extract the combinations from the Time Windows
return Devices in operation

It is noteworthy that after applying the PCA-LD, each row of the transformed matrix is assigned to a new label. This label represents the entire time window and allows the decision-making process for the classification task. In the sequel, the data are scattered, and the labels of the test data are verified by the minimum Euclidean distance via

$$\mathbf{d}(\mathbf{v}, \mathbf{w}) = \underset{w \in \mathcal{R}_+}{\arg\min} \sqrt{\sum_{i=1}^{m} (v_i - w_i)^2}, \tag{3.8}$$

where $\mathbf{v}$ and $\mathbf{w}$ are vectors m-dimensional and m is the number of main components. For the classification task, well-known supervised learning algorithms such as kNN, Decision Tree, and Random Forest algorithms are used. Algorithm 3.2 details the required steps for disaggregation according to the equations and parameters requested as presented in this Section.

3.6 Evaluation Metrics

To evaluate the proposed approach, we used the evaluation metrics provided by the confusion matrix (FAWCETT, 2006). Such metrics are accuracy, precision, recall, and F1-Score (GRANDINI *et al.*, 2020a). These scores rely on rates derived from the classification process to provide a percentage of success of the technique employed to perform the task. Those rates concern the Positive (ON devices) and Negative (OFF devices) classes in our problem, and are described in Section 2.4.

Once these rates have been computed, one can calculate the evaluation metrics. Therefore the accuracy which concerns the overall performance of the model is calculated via equation (2.23). The precision measures the correct predictions of the model concerning the positive class and calculated via equation (2.24). The recall evaluates the model by matching the specific predictions concerning the expected data as belonging to the positive class and derived via equation (2.25), and finally the F1-Score corresponds to the harmonic mean between precision and recall, via equation (2.26).

3.7 Databases

Our investigation consists of using data from RAE Dataset (MAKONIN *et al.*, 2018) and REDD Dataset (KOLTER; JOHNSON, 2011) for the load disaggregation process using active power readings. The reason for choosing only the active power is to attempt to disaggregate households with no smart outlets.

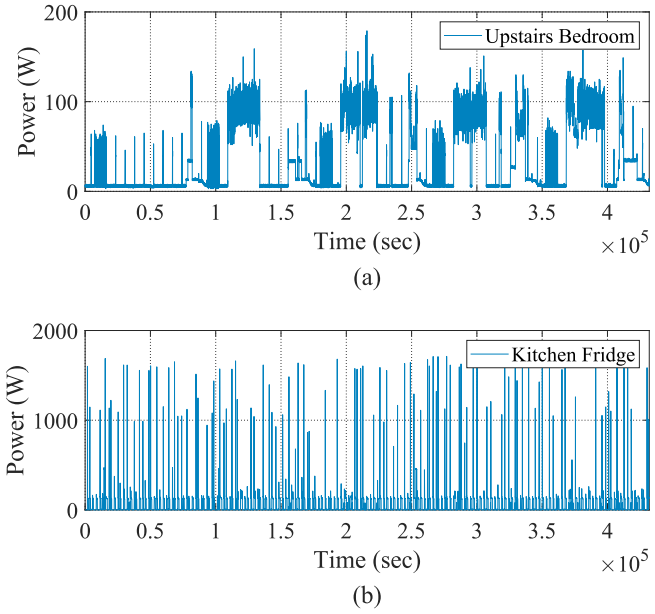


Figure 3.4 – Example of signatures of the RAE dataset (MAKONIN *et al.*, 2018).

RAE contains 11.3 million measurements of electrical power values by the NILM method sampled at 1 Hz, monitoring loads of electronic devices in two residential houses. There are readings of 24 loads captured for 72 days, in House 1, and 21 loads for 59 days, in House 2.

REDD is the dataset built by Massachusetts Institute of Technology (MIT) researchers containing low-frequency measurements in 6 homes in Massachusetts during the year of 2011. Different numbers of appliances per residence were monitored, sampled at 1/3 Hz each, and the main circuit sampled at 1 Hz using the NILM method. According to the authors, the measurements resulted in 119 days of data when all houses are combined.

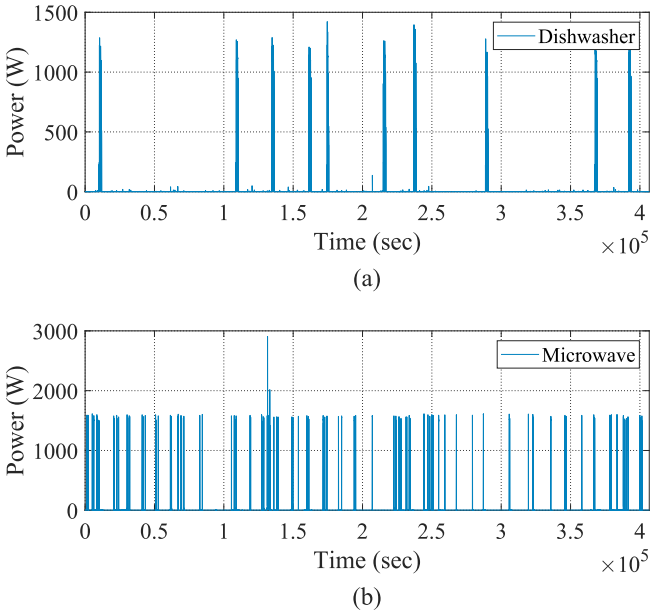


Figure 3.5 – Example of signatures of the REDD dataset (KOLTER; JOHNSON, 2011).

Figures 3.4 and 3.5 depict examples of signatures in RAE and REDD, respectively.

3.8 Simulation and Results

This work consists of the investigation of the efficiency of the strategy of load disaggregation of a specific time window, obtained from the reading of the aggregate active power and smart outlets. For this reason, the proposed method is evaluated in two different scenarios. It is possible to verify the generalization capacity of the approach through the evaluation metrics, i.e., situations including measurement frequencies and increasing numbers of devices at the residence.

Households typically have different amounts and types of appliances. Therefore, as a preliminary process to the disaggregation task, one needs to identify the loads connected to the intelligent outlets described in Section 3.4. Nevertheless, for disaggregation, it is necessary to find out which appliances contribute to the summation given in equation (3.1). Following the same data formatting model used for the identification stage and applying dimensionality reduction would yield one coordinate per column of the input matrix. Such a fact would require an assumption that, during the corresponding period, a single device is in operation. To avoid this problem, we have formatted the data using the Time Windows approach, as described in Section 3.5.

To verify the behavior of the proposed approach, the RAE and REDD datasets were employed. For the RAE dataset, it is considered the data in the table *house1_power_blk1.tab*, which contains data referring to 9 days of measurements performed in House 1 with 17 different devices and the main circuit. For the REDD dataset, it is considered the low-frequency measurements from house 1. Here, the database presents a different sample rate for the appliances and the main circuit. Therefore, via the timestamp, it is possible to determine the instants of time that the aggregated and disaggregated data match. In this case, the frequency of measurement corresponds to 0.33 Hz. Accordingly, the REDD dataset provides data of 8 days of readings for 11 appliances and the central circuit. The difference between aggregate power and the sum of the values of all equipment is attributed to an unknown device.

For the classification task, the kNN algorithm employs the exhaustive search method for the nearest neighbor based on Minkowski distance. The hyperparameter k was set by k-fold cross-validation, yielding $k = 2$. The Decision Tree algorithm establishes the search by entropy criteria and the *random state* value used is 10. Random Forest algorithm uses 100 trees in the forest, in addition to the same criteria and *random state* value used by the Decision Tree algorithm.

All employed methods are implemented in Python 3, available in the *sklearn* library. For the computation of the metric ACR, as described in (BINFENG *et al.*, 2007), the parameter m is determined as the value of ACR that surpasses $\rho = 0.85$. The results are evaluated accordingly to the metrics described in Section 2.4, which are Accuracy, Precision, Recall, and F1-Score.

3.8.1 Scenario 1: RAE Disaggregation

In the RAE dataset, the measurements are sampled at a frequency of 1 Hz. For the Load ID process, the matrix in (3.2) received data referring to 6 days ($d = 6$), yielding six active power columns for each device in the input matrix. For each appliance, four columns are used for clustering and two columns for validation. The first step of the process consists of setting the number of principal components required for the problem. Through ACR metric, as it is depicted in Figure 3.6, three principal components are necessary to achieve the threshold $\rho = 0,85$, therefore $m = 3$.

Figure 3.7 depicts the result of the Load ID process. The model is capable of clustering according to the equipment patterns extracted via dimensionality reduction. The colored markers indicate the data presented in the input matrix, whereas the black markers correspond to the data used to verify the standardization of the devices. Figure 3.7(b) is a zoom, emphasizing the dashed region of Figure 3.7(a), pointing out that even at points where the density of different equipment is higher, it is possible to identify well-defined clusters. Hence, it is possible to recognize the devices connected to the smart outlets.

It is possible to initiate the disaggregation process via the time window. For this stage, the window is structured assuming the number of samples regarding 1 minute of measurements. i.e., every row of the matrix given in (3.7) contains readings referring to 1 minute of house consumption, and the disaggregation algorithm is responsible for

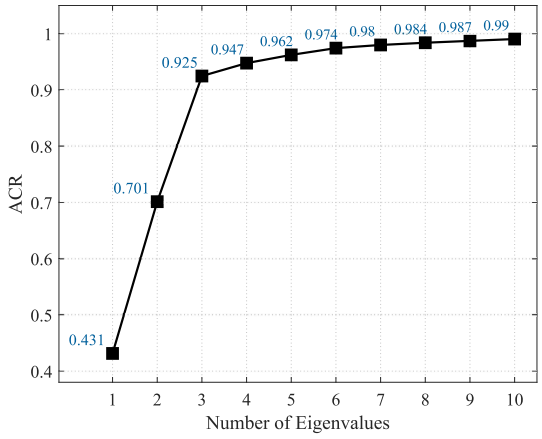


Figure 3.6 – ACR metric for RAE dataset.

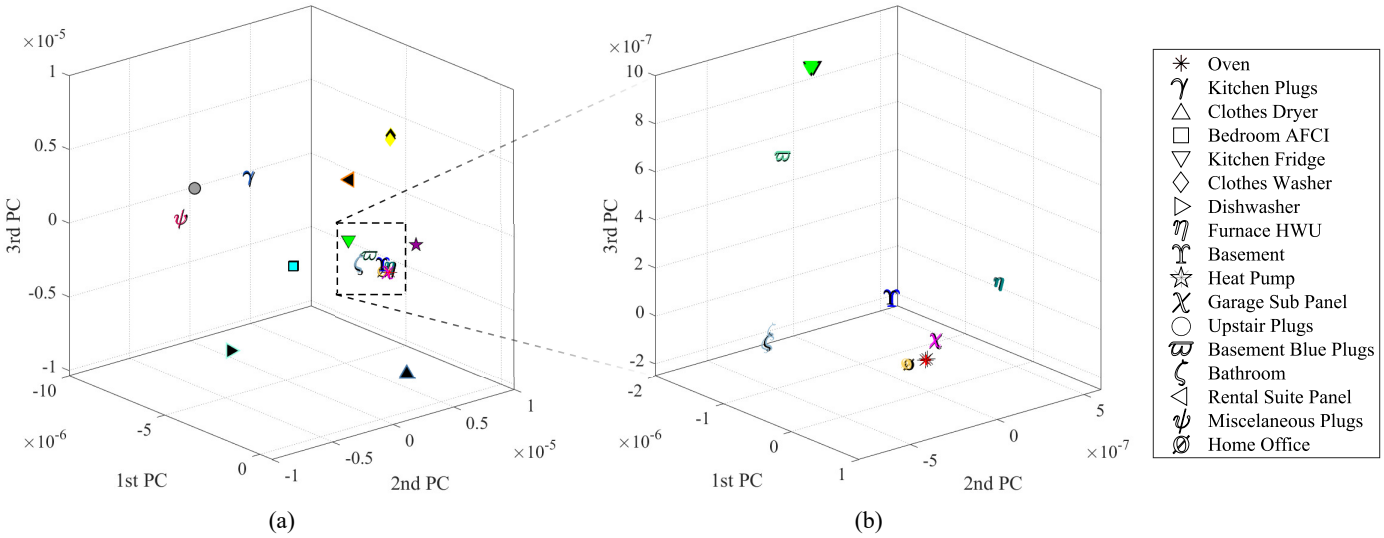


Figure 3.7 – Patterns of devices in RAE dataset.

determining which appliances were in operation during the period. Twenty-four hours of measurements are used and divided equally, 50% for training purposes and 50% for test purposes. Figure 3.8 depicts the arrangement of the data after the feature extraction.

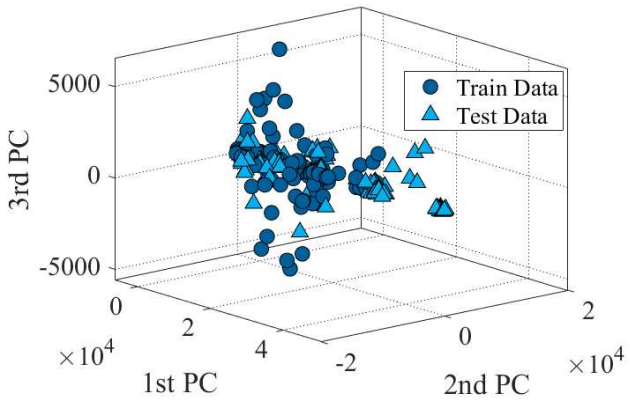


Figure 3.8 – Train and test data disposal for RAE disaggregation.

Table 3.2 shows the percentage of success achieved by the classifiers in Scenario 1 after 125 minutes of simulation. The results are very similar for all evaluated metrics.

Table 3.2 – Scenario 1 results.

Classifier	Accuracy	Precision	Recall	F1-Score
kNN	92.68%	95.43%	95.32%	95.15%
Decision Tree	92.75%	95.25%	95.56%	95.21%
Random Forest	92.50%	95.12%	95.35%	95.04%

Note that for problems involving combinations of appliances, as described in Table 3.1, at every instant of measurement, a new label can show up. This event would

require a re-training phase of the classifiers. However, considering the time window approach, this step is not necessary since, even if it belongs to an untrained class, a device can be classified as ON by checking other classes that were presented in the training stage. That is, a new class appears during the test phase representing a new sequence of connected devices. However, the re-training stage is not necessary since most of the appliances that compose it can already be noticed in classes that were presented in the training stage. Hence, by checking the time window, the non-existence of this new class has no significant impact on the classification error.

In this scenario, in comparison to the other classifiers, the Decision Tree algorithm presented better results. Despite the slight difference to the others, it achieved advantages in terms of accuracy and F1-Score, reaching 92.75% and 95.21%, respectively.

3.8.2 Scenario 2: REDD Disaggregation

In this scenario, we have used the REDD dataset. Similar to Scenario 1, matrix $\mathbf{X}$ receives the data of six measurement days, i.e., $d = 6$. Figure 3.9 shows the ACR metric versus the number of eigenvalues. Note that when the number of eigenvalues is greater than $m = 2$, the threshold $\rho = 0.85$ is attained, therefore we consider $m = 2$.

Figure 3.10 depicts that the Load ID worked for the case where the frequency of measurement is lower. Again, the colored markers correspond to data from the input matrix, and the black markers represent data used for validation.

For disaggregation, 24-hour data were considered, the number of samples presented in the time window corresponds to 1 minute, and the proportion between training and testing is preserved. Figure 3.11 depicts the disposal of training and testing data following the feature extraction process.

The results are presented in Table 3.3, and the runtime, the spent time to

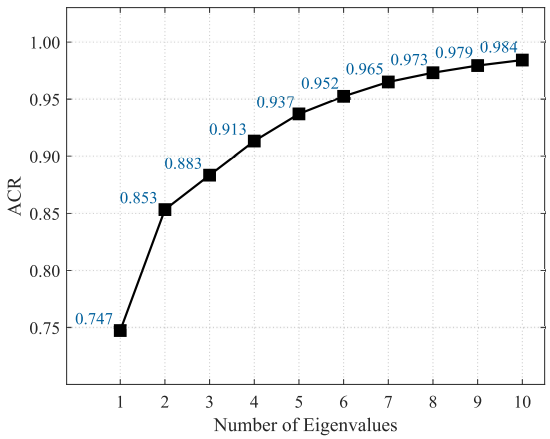


Figure 3.9 – ACR Metric for REDD dataset.

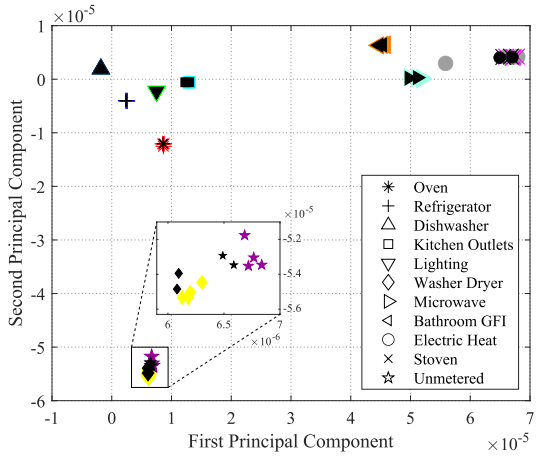


Figure 3.10 – Patterns of devices in the REDD dataset.

run all the steps of the algorithm, summed up 26 minutes. As in the previous scenario, the metrics indicate similarity in the performance of the classifiers employed in the task. However, the kNN algorithm presented a very narrow advantage compared to the others, reaching 94.54% of Accuracy and 95.70% of F1-Score.

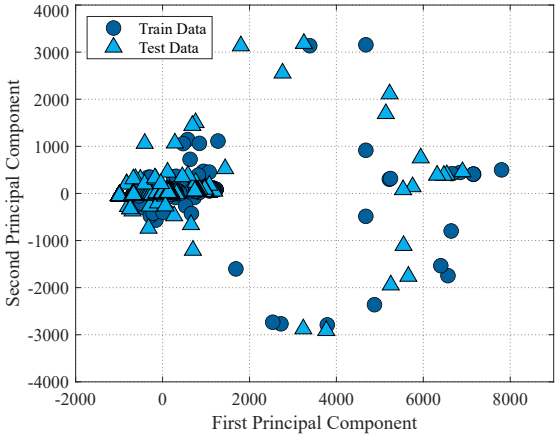


Figure 3.11 – Train and test data disposal for REDD disaggregation.

Table 3.3 – Scenario 2 results.

Classifier	Accuracy	Precision	Recall	F1-Score
kNN	94.54%	95.90%	96.29%	95.70%
Decision Tree	94.36%	95.54%	96.41%	95.58%
Random Forest	93.91%	95.48%	95.84%	95.25%

3.8.3 Brief Discussion of The Results

Table 3.4 presents a comparison of the F1-Score for many disaggregation methods available in the literature. A significant number of alternatives are available to be applied for this purpose. Some of these alternatives provide a relevant percentage of success by increasing the number of equipment in the process. However, due to the system

Table 3.4 – Comparative performance of disaggregation methods in REDD and in another dataset.

Disaggreg. systems	**Num. of apps. in REDD**	**F1-Score REDD**	**F1-Score (2nd Dataset)**
Proposed method	All	95.51 %	95.13 % (RAE)
PCA Clustering (MORADZADEH *et al.*, 2020)	All	94.68 %	-
PBN (WELIKALA *et al.*, 2017)	All	88.50 %	91.60 % (Tracebase)
Viterbi (ZEIFMAN; ROTH, 2012)	9	88.10 %	-
Bayesian classifier (ZEIFMAN; ROTH, 2012)	9	83.00 %	-
CNN and clustering (ZHANG *et al.*, 2020)	10	81.30 %	-
Basic NILM (DINESH *et al.*, 2015)	All	79.70 %	82.20 % (Tracebase)
LDwA (PICCIALLI; SUDOSO, 2021)	3	76.73 %	79.06 % (UK-DALE)
Supervised DT (LIAO *et al.*, 2014)	9	76.40 %	87.64 % (Test House)
Unsupervised GSP (ZHAO *et al.*, 2016)	5	72.20 %	53.42 % (UK-REFIT)
Additive FHMM (KOLTER; JAAKKOLA, 2012)	7	71.30 %	-
P-SGSP (ZHAO *et al.*, 2018)	All	71.00 %	61.50 % (REFIT)
ML-Sparse Repr. (SINGH; MAJUMDAR, 2019)	4	69.97 %	70.75 % (PECAN)
SCANet (CHEN *et al.*, 2019)	3	69.92 %	74.65 % (UK-DALE)
P-UGSP (ZHAO *et al.*, 2018)	All	69.50 %	56.00 % (REFIT)
P-DT (ZHAO *et al.*, 2018)	All	69.00 %	59.50 % (REFIT)
Unsupervised DTW (LIAO *et al.*, 2014)	9	68.60 %	84.28 % (Test House)
Supervised GSP (STANKOVIC *et al.*, 2014)	5	64.00 %	-
Unsupervised HMM (PARSON *et al.*, 2012)	7	62.20 %	-
SGN (SHIN *et al.*, 2019)	4	61.28 %	72.09 % (UK-DALE)
Multi-Label kNN (TABATABAEI *et al.*, 2016)	9	59.10 %	-

complexity, some approaches are restricted to a limited amount of samples. Verifying the F1-Score for REDD, the proposed method presents the highest performance with no restriction on the number of devices and a considerable amount of data. Additionally, considering RAE, the time window yields high performance even for a larger number of devices. These results evidence the generalization capability of the proposed method.

3.9 Conclusion

This work investigated a new strategy for load disaggregation concerning measurements of active power and time window formatting. The analysis revealed similar results of the algorithms applied for the classification task achieving values of 95.13% for F1-Score in the RAE dataset, and 95.51% in REDD dataset. In a case with less data, the kNN algorithm proved to be more effective in identifying equipment in operation. On the other hand, the Decision Tree algorithm was superior in a scenario with a larger volume of measurements.

Among the key advantages of the proposed approach is the possibility of disaggregating a household, without smart outlets, restricted to the need for the same or similar appliances present at the residence used for the training stage. Another highlight concerns the fact that it is not necessary to train all the possible combinations of ON devices to perform the test stage. Once the device appears in the window, it can be identified regardless of the set of appliances never appeared at the training stage.

4 Low Runtime Approach for Fault Detection for Refrigeration Systems in Smart Homes Using Wavelet Transform

This Chapter is an adapted reproduction of the article:

D. A. M. Lemes, T. W. Cabral, L. L. Motta, G. Fraidenraich, E. R. Lima, F. B. Neto, and L. G. P. Meloni, "Low Runtime Approach for Fault Detection for Refrigeration Systems in Smart Homes Using Wavelet Transform", Under Review at the IEEE Transactions On Cosumer Electronics.

This paper presents an approach proposed for detecting malfunctions in refrigeration systems. Using the Wavelet Transform and the REFIT database, a method is used to verify the duty cycles for comparison with those labeled as malfunction events. The decision-making system is focused solely on the occurrence of the anomaly due to the dynamic nature of the failures and the requirement of prior knowledge of the fault patterns. The results indicate that the proposed method is effective in the fault detection task and in the instants of faults begin and cease to occur.

Abstract

This work investigates the efficiency of a fault detection system for refrigeration equipment based on Wavelet Transform. To perform the task, the well–known database REFIT (Personalised Retrofit Decision Support Tools for UK Homes using Smart Home Technology) is used considering data from six devices presented in three households of the database considering readings of active power as features. The results revealed an approach with high efficiency in detecting abnormal behavior of the tested appliances and presented low runtime despite large volumes of input data. Another positive point to highlight, in the proposed method, is the capacity to estimate the instants at which faults start and cease to occur. Moreover, by setting parameters according to the characteristics of the appliances, the proposed method is suitable for different equipment, thereby establishing feasible fault detection for diverse groups of devices.

4.1 Introduction

The advances in technology developments enable diverse analyses concerning large volumes of data. Consequently, efforts focused on the energy demand gathered more attention in the industrial and academic areas, aiming to improve production, distribution, and converging to efficiency and sustainable consumption (CARY; BENTON, 2012; FLETCHER; MALALASEKERA, 2016; LI *et al.*, 2018; IWAFUNE *et al.*, 2017; LEE *et al.*, 2019). In this context, minimizing wasteful spending on energy consumption becomes essential. Therefore, the fault detection process in household appliances increases in relevance to avoid unnecessary consumption by the provision of repairs (ARCOS *et al.*, 2020).

Globally, approximately 80 percent of energy generation originates from fossil fuels, motivating the research for sustainable methods of production and consumption aiming at green energy sources and decreasing the total percentage consumption (HIMEUR *et al.*, 2021b). According to (CARY; BENTON, 2012), among the commercial, industrial and domestic segments, the household sector holds the potential for energy savings amounting to 66.5 TWh. In the United States, domestic electricity consumption reached 38.5% of total annual consumption in 2018 (EIA, 2021), and during the previous year, approximately 7 billion dollars was spent unnecessarily (MATTERA *et al.*, 2019). According to (RASHID; SINGH, 2018) and (BANG *et al.*, 2019), approximately 30% of energy waste is due to consumer habits or equipment malfunctions, also known as anomalies.

According to (HOSSEINI *et al.*, 2020), the anomaly possesses a dynamic and stochastic nature, contributing to power and operation time deviations in household appliances. The irregular consumption of electric heaters, freezers, and refrigerators is related to the non-regular operation time. In addition, one of the main motivations in anomaly detection stems from the structural characteristics of residential networks. For most domestic networks, the infrastructure is not enough to detect all sorts of anomalies, especially operation time abnormalities with typical power demand (CHANDOLA *et al.*, 2009; HOSSEINI *et al.*, 2020). Therefore, many strategies developed in the literature intend to minimize the energy waste occasioned by anomalies.

Among the proposed methods, some authors argue for the employment of unsupervised learning. They consider that anomalous samples belong to the minority class. In this instance, the abnormal patterns represent a minor proportion concerning the total data volume. Therefore, detection techniques based on dimensionality reduction (SIAL *et al.*, 2021; HIMEUR *et al.*, 2021a; KAMARAJ *et al.*, 2019), one-class learning (JAKKULA; COOK, 2011; CHALAPATHY *et al.*, 2018; OZA; PATEL, 2018; ZHANG *et al.*, 2017;

DÉSIR *et al.*, 2013), and clustering algorithms (PASTORE *et al.*, 2020; ARJUNAN *et al.*, 2015; ROSSI *et al.*, 2016; HENRIQUES *et al.*, 2020; IZAKIAN; PEDRYCZ, 2013) can be useful. On the other hand, other authors prefer to use supervised methods due to the tendency to achieve high accuracy (HIMEUR *et al.*, 2021b). The most known supervised methods for anomaly detection comprise convolutional neural networks (CNN) (MUNIR *et al.*, 2019; LI *et al.*, 2019; ZHENG *et al.*, 2017), Recurrent Neural Network (RNN) (SILVA *et al.*, 2019; HOLLINGSWORTH *et al.*, 2018), Long Short-Term Memory (LSTM) (WANG *et al.*, 2019b), Autoencoders (WENG *et al.*, 2018; PEREIRA; SILVEIRA, 2018; YUAN; JIA, 2015), regression models (KROMANIS; KRIPAKARAN, 2013; ZHANG *et al.*, 2011), and classifier models (SIAL *et al.*, 2021; JAKKULA; COOK, 2010; DEPURU *et al.*, 2011; KORBA; KARABADJI, 2019; CODY *et al.*, 2015). Other researchers employ feature extraction methods for the task (GU *et al.*, 2019; ZHANG; WANG, 2018; MAO *et al.*, 2018; LEE; KIM, 2018; PETLADWALA *et al.*, 2019). Nevertheless, due to the complexity of the problem, high accuracy is not often achieved in low time processing approaches.

In the literature, some works achieved success with accuracy considered high. In (WANG *et al.*, 2019a), the authors proposed Sample Efficient Home Power Anomaly Detection (SEPAD), capable of detecting anomalies considering active power as features for two classifiers. The first is responsible for classifying appliance patterns, and the second classifies energy consumption habits. The tests involved 52 households with 150 occupants, and the maximum performance of SEPAD achieved 87.30% in terms of accuracy. In the reference (FERNANDES *et al.*, 2020), the authors describe a solution for failure prediction Heating, Ventilation, and Air-Conditioning (HVAC) systems. For this purpose, the method used temperature data, duration of each cycle, number of heating requests, and number of boiler starts. The work presented a pre-processing stage, aggregating the data, labeling, and scaling it. Afterward, the data was ready for the model training procedure. To perform the classification, models such as random stratified classifier, random tree classifier, Neural Network (NN), the sample weighted neuronal network (weighted NN), Decision Tree, and different LSTM varieties participated in the analysis. The results indicated that $LSTM_0$-1 presented the best accuracy performance with 78.67%. However, as the results reported, the higher the performance, the higher the complexity of the model.

On the other side, in (HOSSEINI *et al.*, 2020), the emphasis is on the low-time processing. In this study, the central issue is the inspection of anomalies in refrigerators. For this task, the authors propose to detect the anomaly via the time-window method and define a framework for identifying anomalous behavior. The results show that it requires some cycles to build an efficient classifier, a minimum of 76 cycles.

In this context, this research intends to introduce a novel approach for fault

detection in refrigeration systems seeking to provide information regarding the occurrence of anomalies for the user. Detecting malfunctioning of a household appliance requires methods of considerable computational complexity to achieve high accuracy, as previously mentioned. To avoid such a problem, we consider an approach based on the Wavelet Transform to minimize the runtime and achieve high accuracy in the anomaly detection task. The Wavelet Transform is a widely known tool in the literature and is applied to electrical power grid-related problems, most notably in arc fault detection (YU *et al.*, 2020; TANG; ZHANG, 2022; QI *et al.*, 2017). Further applications concerning home appliances include air conditioning (KIM, 2016) and specific groups of equipment such as incandescent, lamp rice, cooker hair dryer, induction cooker, and electric kettle (LI *et al.*, 2018). To the best of our knowledge, this is the first study involving the application of the wavelet transform to refrigeration systems, such as freezers and fridges. In addition, active power is one of the most relevant parameters to be measured in a household for anomaly detection in appliances. Since a signal refers to any manifestation that enables the knowledge, recognition, or prediction of some event, active power can be described as a signal. Such a fact justifies the use of the wavelet as a tool for analyzing the behavior of appliances. By adjusting the proposed method according to the characteristics of the appliances, it is feasible to employ it for analysis regarding several groups of devices besides the ones verified in this work. The main contributions of this work are the following:

- High accuracy in anomaly detection;
- The approach requires only active power as a feature for the problem, and it is capable of verifying abnormalities related to operation time;
- The system does not require data pre-processing and training stage;
- Low runtime. The method handles a large volume of data in small time of execution;
- Provide the estimated beginning and end instants of appliance malfunctions to the customer.

The paper is organized as follows: Section 4.2 provides an overview of the proposed method presenting the average accuracy achieved in the simulations, Section 4.3 describes the REFIT database, Section 4.4 gives an overview of the Wavelet Transform, Section 4.5 explains the detection process, Section 4.6 contains the metrics employed for the performance evaluation of the proposed method, Section 4.7 presents the simulation results, and Section 4.8 concludes on the described results.

4.2 Overview of the Proposed Method

The proposed method consists in collecting raw data of active power of appliances, such as refrigeration equipment, and applying the Wavelet Transform to identify instants of abnormal functioning of these devices. After receiving the data collected by the smart outlets, the fault detection system employs the Wavelet Transform to the power readings, extracting the level 1 coefficients of detail. The reason for adopting the wavelet coefficients lies in the possibility of detecting the occurrence of appliance activity by analyzing the level 1 coefficient of detail. More precisely, sufficient information for detection resides at level 1. I. e. no further complex analysis is required to expand the decomposition to higher levels. This process can be seen as a feature extraction stage since the information of interest lies in the frequency domain. The anomalies of the investigated devices distinguish by the longer duration of the operating cycles as described in the REFIT database, and depicted in Figure 4.4. Therefore, the analysis of the coefficients provided by Wavelet enables the estimation of the instants at which faults start and cease to occur within the cycles, allowing the identification of occurrences of periods of longer cycles than usual and hence characterizing these events as abnormal operating instants. As a result, the system notifies the customer of the estimated time of occurrence of the malfunction. Figures 4.1 and 4.2 depict the flowcharts of the proposed method. As shown in Figure 4.2(b), applying the Discrete Wavelet Transform to the collected data, it is feasible to verify the existence of operating cycles with longer duration by analyzing the coefficients of detail. As aforementioned, such cycles represent anomalous behavior. Therefore, by checking the length of these cycles via the coefficients of detail and the parameters specified in Section 4.5, the proposed method identifies these cycles as faults, providing an estimate of the beginning and end of the malfunction of the appliance in question. Further details such as the initial conditions and information regarding the data and the simulation parameters will be discussed in sections 4.3, 4.5, and 4.7.

The proposed system achieved 87.33% accuracy on the task, the highest result reported in the literature. It is worth highlighting that other approaches utilized in the same problem do not consider the same appliances, referring to boilers and other measured quantities (FERNANDES *et al.*, 2020). Another point to be emphasized relates to the absence of databases to provide a comparison of approaches. Therefore, methods using active power as input data performed proper measurements to test their system (WANG *et al.*, 2019a; HORI *et al.*, 2017; CHOU; TELAGA, 2014; NAGI *et al.*, 2009).

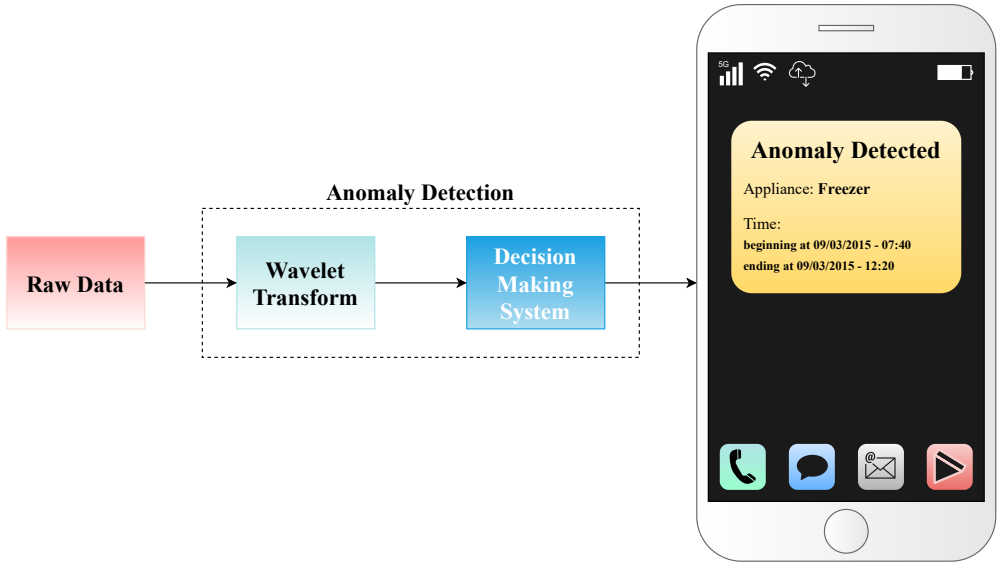


Figure 4.1 – Flowchart of the proposed method.

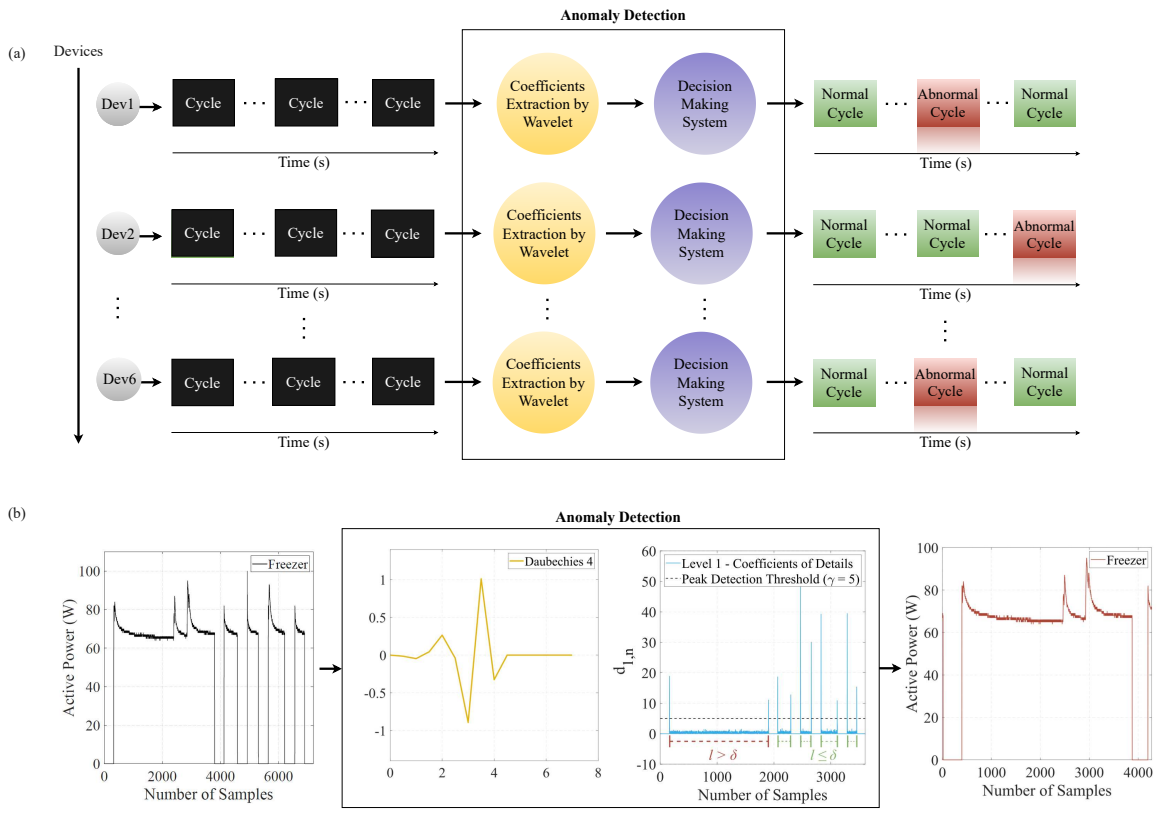


Figure 4.2 – Flowchart of the anomaly detection of the active power via Wavelet Transform considering the family Daubechies.

4.3 Database

In this work, the data employed refers to the Personalised Retrofit Decision Support Tools For UK Homes Using Smart Home Technology (REFIT Dataset) (MURRAY *et al.*, 2017) seeking to identify malfunction instants of appliances regarding active power readings as features.

REFIT contains 1,194,958,790 active power readings of appliances present in 20 homes, at a frequency of 1 measurement every 8 seconds, in the UK, gathered during two years, from 2013 to 2015.

To verify the effectiveness of this approach, we selected data referring to the freezers and fridges of Houses 1, 2, and 4, resulting in 6 devices for the analysis. Fig. 4.3 displays the example of the freezer signature found in House 4 of the database.

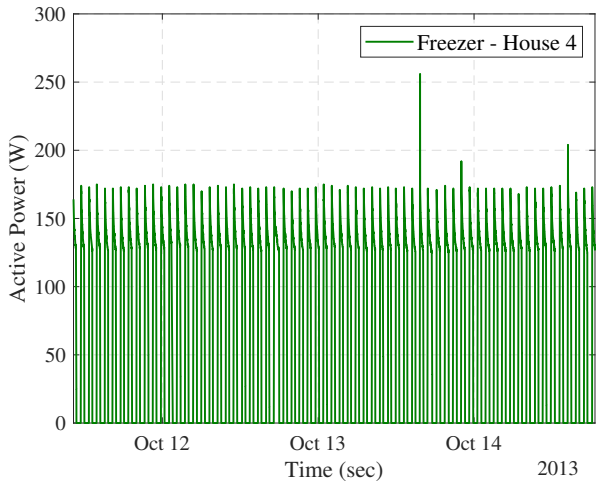


Figure 4.3 – Example of a Freezer Signature in REFIT Dataset.

4.4 Time-Frequency Analysis and Wavelet Transform

As previously mentioned in Section 2.3, the (CWT) provide good resolution of time for high-frequency events and good resolution of frequency for low-frequency events, and it is is defined in equation (2.15).

When analyzing practical signals, i.e., discrete signals of finite length, discrete versions of the transforms are employed. Similar to the Discrete Fourier Transform (DFT), there is also the Discrete Wavelet Transform (DWT) for the Wavelet transform. The concept of DWT defines itself in the mother wavelet function $\psi(\cdot)$. The CWT retains all the information regarding the transformed signal when sampled at a given discrete subset of the time-scale field. For this purpose, one considers (2.15) instead of the pair of variables (a, b). The pair $(a^{-j}, a^{-j}bk)$ are generally set $a = 2$ and $b = 1$, respectively, where the support is given by $(a^{-j}, a^{-j}bk)$, according to (2.16), yielding the coefficients of details given by (2.17).

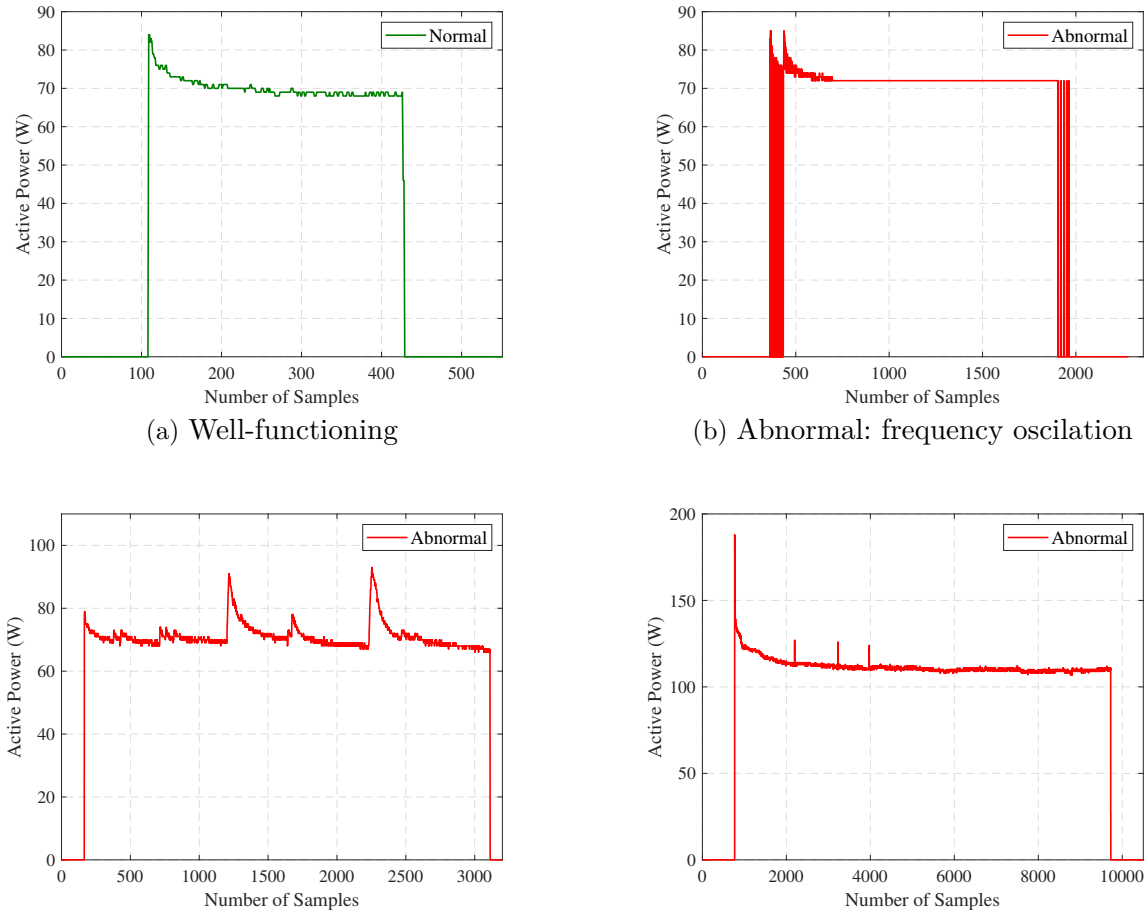
Every Wavelet family has unique characteristics and specific parameters. One refers to the number of null moments (*vanish moments*) of the functions $\psi(\cdot)$ and $\phi(\cdot)$. This parameter relates directly to the selectivity of the filters composed by such functions,

and larger values imply higher computational cost since the number of coefficients of the decomposition and reconstruction filters increase.

The impulse responses of such filters are assumed to satisfy the conditions of perfect reconstruction and orthogonality imposed in the transform. Examples of filters fulfilling the criteria are the Daubechies, Haar, Morlet, Coiflets, and Biorthogonal (MALLAT, 1999).

4.5 Fault Detection Method

To detect anomalies in the active power readings of an appliance requires prior knowledge of the difference between periods of malfunction and normal instants. The table *REFIT_anomalies.csv* contains information regarding the moments of occurrence of malfunctions of the devices present in the database. Therefore, it is possible to assign labels to the power samples according to the characteristics of the highlighted instants. Figure 4.4 depicts examples of the moments of occurrence of devices malfunctions.



(a) Well-functioning

(b) Abnormal: frequency oscilation

(c) Abnormal: distortions in the curve

(d) Abnormal: longer cycles

Figure 4.4 – Examples of normal and abnormal cycles.

As illustrated in Figure 4.4 and described in (HOSSEINI *et al.*, 2020), the faults in refrigeration equipment differ according to frequency oscillations pre and post-functioning cycles (Figure 4.4(b)), distortions in the curve (Figures 4.4(c) and 4.4(d)), and especially the duration time of the functioning cycles (Figures 4.4(b), 4.4(c), and 4.4(d)). In other words, the malfunction instants present longer cycles compared to the moments when the device operates normally (Figure 4.4(a)).

Due to this characteristic, the Wavelet Transform is an alternative to identify the periods when anomalies appear. The mother waveform Daubechies 4 is suitable for the case according to the characteristics of the transitions at the beginning of the activity of the equipment. In other words, the selected waveform matches the active power oscillation characteristics of refrigeration systems, meaning the Daubechies family, especially the previously mentioned Daubechies 4, seeks a signal shaped according to the studied waveform. Therefore, the coefficients of details are efficiently derived to detect the occurrence of activity, providing support for the anomaly detection process. Applying the DWT via (2.16), through the coefficients of detail derived from (2.17), one can estimate the duration of the functioning cycles. Hence, it is possible to identify the malfunctions of the appliances and verify the beginning and end instants of the faults.

Figure 4.2(b) depicts an example of the coefficients of detail feature analysis. As plotted in Figure 4.2(b), peaks of higher magnitude indicate the beginning and the end of the operating cycles. t is possible to set the number of samples, δ, to establish the optimal length of the well-functioning cycles. Meaning,

$$\text{if } \begin{cases} l \leq \delta, & \text{Normal} \\ l > \delta, & \text{Abnormal,} \end{cases}$$

where l is the number of samples in the observed cycle.

In other words, as described in Algorithm 4.1 and depicted in Figures 4.1 and 4.2, the system consists of reading the raw data gathered from the appliances. Subsequently, the data passes through the wavelet transform to detect the instants of activity, and afterward, a scanning function verifies the size of the operation cycles. Hence, the decision-making system can determine if the cycle size is greater than the threshold and assign the label referring to the anomaly, besides storing the start and ending timestamps of the fault occurrence.

Algorithm 4.1 Fault Detection in Refrigeration Systems

Input: Raw data from the appliances, δ, γ, η_{OFF}, l, *Mother Wavelet*
Output: Fault Instants
Initialization:
1: first statement
Calculate $\mathbf{d}_{m,n}$ *as in* (2.17)
2: *second statement*:
Extract the activity instants of $\mathbf{d}_{m,n}$ via peak detection threshold γ
for $i \leftarrow 1$:*number of cycles*
for $j \leftarrow 1$: *length of i-th cycle*
Update l
end for
if $l > \delta$
save j, j+length of *i-th* cycle
end if
end for
return Fault Instants

4.6 Evaluation Metrics

To evaluate the efficiency of the proposed approach, the metrics employed are based on the binary class classification in machine learning (FAWCETT, 2006; GRANDINI *et al.*, 2020b). The labels are assigned following the positive class, the faults, and the negative class, the well-functioning of the appliances. Therefore, the detection accuracy is computed via

$$Acc = \frac{TC}{TA}, \tag{4.1}$$

where TC is the number of correct predictions performed by the Wavelet Transform, and TA is the total of anomalies.

Generally speaking, the metric adopted is the average prediction accuracy, i.e., the average accuracy value considering all the evaluated appliances via,

$$\overline{Acc} = \frac{\sum_{i=1}^{N} Acc_i}{N}. \tag{4.2}$$

Where Acc_i is the value of the accuracy per appliance given by (4.1), and N is the number of investigated devices.

The detection errors, i.e., instants detected by the Wavelet Transform as anomalous although not corresponding to segments displaying the characteristics depicted in Figures 4.4b, 4.4c and 4.4d are denoted as false positives, calculated via,

$$FP = \frac{\epsilon}{TD}. \tag{4.3}$$

Where ϵ is the number of errors, and TD is the total number of detected segments.

4.7 Simulation and Results

This work aims the detection of the fault occurrences in household appliances based on active power readings. Typically, the devices display different power consumption patterns, meaning that abnormalities differ according to the equipment. Therefore, establishing a feasible and efficient general solution is extremely difficult and requires high computational complexity. Hence, as described in Section 4.5, refrigeration equipment, such as freezers and fridges are devices that present similar behavior concerning power signature and anomalies, enabling the development of a method concerning both appliances.

Based on the previous knowledge of the fault characteristics, it is possible to employ the Wavelet Transform as a tool for detecting the occurrence of malfunctioning instants. Therefore, from (2.17), the level 1 coefficients of details are derived, as depicted in Fig. 4.2(b) (sub-figure in blue line), allowing this information to be used as the features required throughout the process. After obtaining the coefficients of details, the information regarding the moments of activity of the devices is available, enabling the analysis of the time the devices remained ON. Based on this information, one can estimate the duration of the operation cycle, verifying the existence of anomalies at the instant in question.

Table 4.1 describes the households selected in the REFIT database, the investigated appliances in those households, and the number of days corresponding to the data considered in the simulations.

Table 4.1 – Device selection for simulations.

House No.	**Appliances**	**Number of Days**
1	Fridge Freezer 1 Freezer 2	639
2	Freezer	617
4	FridgeFreezer Freezer	634

The simulation was performed in a computer Dell Inspiron 7560, CPU i7-7500U, and 8 GHz of RAM. The software employed in this study is MATLAB R2018a, considering level 1 Wavelet decomposition. The δ and the threshold of peak detection magnitude were set to 630 and $\gamma = 5$ established via exhaustive search, respectively. The waveform of the Wavelet family employed was Daubechies 4. The frequency considered is the frequency of measurement of the database, $f = \frac{1}{8}$ Hz. An OFF State Filter is

considered to check if a sequence of samples, η_{OFF}, contains values of $P \leq 5$ W. In this case, the device is assumed to be OFF. The value of η_{OFF} is 500 and set according to the spacing between cycles of the household appliances. The metrics used to evaluate the proposed method are described in Section 4.6, which are Accuracy and False Positive Rate.

Table 4.2 shows the results of the fault detection process for every investigated device, where t indicates the execution time of the algorithm.

Table 4.2 – Simulation results per appliances.

House No.	**Appliances**	**Acc**	**False Positive**	**t** (sec.)
1	Freezer 1	60.0%	0.0%	2.43
1	Freezer 2	100%	12.0%	2.42
1	Fridge	83.0%	0.0%	2.64
2	Freezer	94.0%	19.0%	2.00
4	FridgeFreezer	92.0%	20.0%	3.62
4	Freezer	95.0%	21.0%	2.99

The results presented in Table 4.2 reveal a high ability to identify instants of anomalies with low computational complexity. Excluding Freezer 1 of House 1, the remaining accuracy results indicate that the proposed method succeeded in detecting over 80% of the existing faults in the functioning of the devices, reaching 100% of efficiency for Freezer 2. In the case of Freezer 1, the reduced number of instants of abnormal functioning contributes to the lower efficiency of the performance regarding this device.

Regarding the false positive rate, in two cases, the system did not recognize periods of good functioning as anomalous events. These were the cases of Freezer 1 and the Fridge of House 1. For the others, the false positive rate never exceeded 21%, and the average rate for this metric reached 12.5%, indicating a reliable system for the execution of the task.

Another interesting issue concerns the execution time of the proposed algorithm. Despite working with large volumes of data in all cases, the average execution time reached 2.68 seconds, indicating very low computational complexity. Hence, it is possible to emphasize that the proposed system performed successfully in the anomaly detection process, with no constraints on the amount of data to analyze.

As previously mentioned, the proposed detection system provides an estimation of the beginning and end times of faults, achieved by the analysis of the coefficients of details. Hence, the system delivers the instants of malfunctioning based on the timestamp of the smart outlets connected to the devices. Figure 4.5 exemplifies the estimated information extracted by the Wavelet in comparison to data present in the database, con-

sidering an anomaly selected from Freezer 2 of House 1. Based on the result illustrated, the estimated instants of fault events approximate the ones highlighted in the database. Therefore, it is possible to employ the proposed approach to alert the consumer to indicate the occurrence of an anomaly. Hence, the user can verify the alert and perform maintenance on the device, avoiding unnecessary expenses related to additional power consumption and potential future failures.

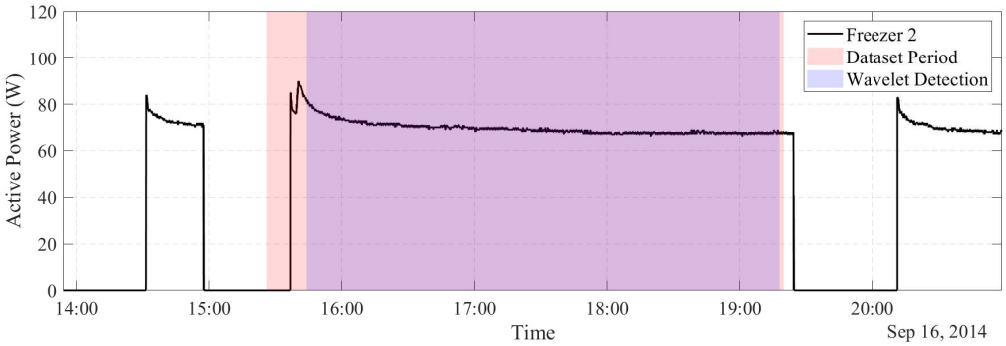


Figure 4.5 – Comparison of a Malfunction Period Reported in the Database and Estimated Via Proposed Method.

An additional interesting point to highlight concerns the issue of generalizing the proposed method to other equipment presenting power consumption characteristics different from the ones studied in this work. By setting the attributes of the Wavelet Transform according to the features of the power signatures of each group of devices, it is possible to verify the existence of anomalies via decomposition analysis of the coefficients of the Wavelet Transform. Therefore, it is feasible to employ the approach to generalize the detection task to other household appliances.

4.8 Conclusion

This Chapter investigated the process of anomaly detection in household appliances performing the Wavelet Transform on the active power readings gathered. The results revealed that the proposed method achieved an efficiency of 87.33% in tests considering the REFIT Dataset for refrigeration equipment from three houses in the database. Therefore, the algorithm provided the most effective result compared to the methods in the literature.

Other relevant advantages of the proposed method include the return to the user of the estimated beginning and end instants of the anomaly occurrence, enabling the user to prevent unnecessary expenses and future malfunctions of the devices. Moreover, the proposed method presented low computational complexity despite the large volume of data, and does not require pre-processing and training steps, simplifying its operation.

Conclusion

This thesis is the result of the study of techniques applicable in HEMS to provide detailed information regarding household energy consumption. In the proposed method for load disaggregation, the use of PCA to extract patterns and features allows the Machine Learning methods to identify the connected appliances during the time window period, achieving an F1-Score of 95.13% in the RAE Dataset and 95.51% in the REDD Dataset. These results revealed that the approach outperformed other methods present in the literature in low execution time. Regarding fault detection, the proposed method using the Wavelet transform achieved 87.33% of accuracy in identifying fault events in refrigeration systems. These results indicate that the system is efficient in this task.

The HEMS system aims to provide additional information regarding energy consumption by establishing methods to improve consumption efficiency, avoid unnecessary waste, and improve sustainability. In this context, the proposed algorithms of disaggregation and fault detection fit in approaches that might be applicable in HEMS to approximate the information to the user. Therefore, the user is capable of controlling their expenses and avoiding waste. Lastly, the electric concessionaire can address regions requiring improved policies for responsible consumption, and offer customers specific policies according to their consumption profile.

Bibliography

ADDISON, P. Wavelet transforms and the ecg: A review. *Physiological measurement*, v. 26, p. R155–99, 11 2005. Cited in page 27.

AL-KABABJI, A.; ALSALEMI, A.; HIMEUR, Y.; FERNANDEZ, R.; BENSAALI, F.; AMIRA, A.; FETAIS, N. Interactive visual study for residential energy consumption data. *Journal of Cleaner Production*, v. 366, p. 132841, 2022. ISSN 0959-6526. Cited in page 17.

ALCALA, J.; UREñA, J.; HERNáNDEZ, ; GUALDA, D. Event-based energy disaggregation algorithm for activity monitoring from a single-point sensor. *IEEE Transactions on Instrumentation and Measurement*, PP, p. 1–12, 07 2017. Cited in pages 17 and 33.

ALTMAN, N.; Hart, P. An introduction to kernel and nearest-neighbor nonparametric regression. *The American Statistician, 46*, v. 46, n. 3, p. 175–185, 1992. Cited in page 23.

AMIT, Y.; GEMAN, D. Shape quantization and recognition with randomized trees. *NEURAL COMPUTATION*, v. 9, p. 1545–1588, 1997. Cited in page 24.

ARCOS, B. P.; BAKKER, C.; FLIPSEN, B.; BALKENENDE, R. Practices of fault diagnosis in household appliances: Insights for design. *Journal of Cleaner Production*, v. 265, p. 121812, 04 2020. Cited in page 52.

ARJUNAN, P.; KHADILKAR, H. D.; GANU, T.; CHARBIWALA, Z. M.; SINGH, A.; SINGH, P. Multi-user energy consumption monitoring and anomaly detection with partial context information. In: *Proceedings of the 2nd ACM international conference on embedded systems for energy-efficient built environments*. [S.l.: s.n.], 2015. p. 35–44. Cited in page 53.

Azzini, H. A. D.; Torquato, R.; da Silva, L. C. P. Event detection methods for nonintrusive load monitoring. In: *2014 IEEE PES General Meeting | Conference Exposition*. [S.l.: s.n.], 2014. p. 1–5. Cited in page 32.

BANG, M.; ENGELSGAARD, S. S.; ALEXANDERSEN, E. K.; SKYDT, M. R.; SHAKER, H. R.; JRADI, M. Novel real-time model-based fault detection method for automatic identification of abnormal energy performance in building ventilation units. *Energy and Buildings*, Elsevier, v. 183, p. 238–251, 2019. Cited in page 52.

BINFENG, Y.; FEILU, L.; DAN, H. Research on edge identification of a defect using pulsed eddy current based on principal component analysis. *NDT E International*, v. 40, n. 4, p. 294–299, 2007. ISSN 0963-8695. Cited in page 45.

BISHOP, C. M. *Pattern Recognition and Machine Learning*. [S.l.]: Springer, 2006. Cited in page 20.

BREIMAN, L. Random forests. *Machine Learning*, v. 45, p. 5–32, 10 2001. Cited in page 24.

CARY, R.; BENTON, D. *Creating a market for electricity savings: Paying for energy efficiency through the energy bill*. [S.l.]: Green Alliance, 2012. Cited in page 52.

CHALAPATHY, R.; MENON, A. K.; CHAWLA, S. Anomaly detection using one-class neural networks. *arXiv preprint arXiv:1802.06360*, 2018. Cited in pages 52 and 53.

CHANDOLA, V.; BANERJEE, A.; KUMAR, V. Anomaly detection: A survey. *ACM computing surveys (CSUR)*, ACM New York, NY, USA, v. 41, n. 3, p. 1–58, 2009. Cited in page 52.

CHEN, K.; ZHANG, Y.; WANG, Q.; HU, J.; FAN, H.; HE, J. Scale-and context-aware convolutional non-intrusive load monitoring. *IEEE Transactions on Power Systems*, IEEE, v. 35, n. 3, p. 2362–2373, 2019. Cited in page 49.

CHOU, J.-S.; TELAGA, A. S. Real-time detection of anomalous power consumption. *Renewable and Sustainable Energy Reviews*, Elsevier, v. 33, p. 400–411, 2014. Cited in page 55.

CODY, C.; FORD, V.; SIRAJ, A. Decision tree learning for fraud detection in consumer energy consumption. In: IEEE. *2015 IEEE 14th international conference on machine learning and applications (ICMLA)*. [S.l.], 2015. p. 1175–1179. Cited in page 53.

CUTLER, A.; CUTLER, D.; STEVENS, J. Random forests. In: _____. [S.l.: s.n.], 2011. v. 45, p. 157–176. ISBN 978-1-4419-9325-0. Cited in page 25.

DAUBECHIES, I. *Ten Lectures on Wavelets*. [S.l.]: Siam, 1992. v. 61. Cited in page 28.

DEPURU, S. S. S. R.; WANG, L.; DEVABHAKTUNI, V. Support vector machine based data classification for detection of electricity theft. In: IEEE. *2011 IEEE/PES Power Systems Conference and Exposition*. [S.l.], 2011. p. 1–8. Cited in page 53.

DÉSIR, C.; BERNARD, S.; PETITJEAN, C.; HEUTTE, L. One class random forests. *Pattern Recognition*, Elsevier, v. 46, n. 12, p. 3490–3506, 2013. Cited in pages 52 and 53.

DIETTERICH, T. G. Machine learning. *Annual Review of Computer Science*, v. 4, n. 1, p. 255–306, 1990. Cited in page 20.

DINESH, C.; NETTASINGHE, B. W.; GODALIYADDA, R. I.; EKANAYAKE, M. P. B.; EKANAYAKE, J.; WIJAYAKULASOORIYA, J. V. Residential appliance identification based on spectral information of low frequency smart meter measurements. *IEEE Transactions on smart grid*, IEEE, v. 7, n. 6, p. 2781–2792, 2015. Cited in page 49.

DING, S.; ZHU, H.; JIA, W.; SU, C. A survey on feature extraction for pattern recognition. *Artif. Intell. Rev.*, v. 37, p. 169–180, 03 2012. Cited in page 21.

DUDA, R. O.; HART, P. E.; STORK, D. G. *Pattern Classification*. 2. ed. New York: Wiley, 2001. ISBN 978-0-471-05669-0. Cited in page 21.

EIA, T. U. E. I. A. *Electricity Consumption in the United States was About 3.95 Trillion Kilowatthours (KWH)*. 2021. <https://www.eia.gov/energyexplained/>. Accessed: 2022-03-02. Cited in page 52.

FAWCETT, T. An introduction to roc analysis. *Pattern Recognition Letters*, v. 27, n. 8, p. 861 – 874, 2006. ISSN 0167-8655. ROC Analysis in Pattern Recognition. Cited in pages 29, 42, and 60.

FERNANDES, S.; ANTUNES, M.; SANTIAGO, A. R.; BARRACA, J. P.; GOMES, D.; AGUIAR, R. L. Forecasting appliances failures: A machine-learning approach to predictive maintenance. *Information*, Multidisciplinary Digital Publishing Institute, v. 11, n. 4, p. 208, 2020. Cited in pages 18, 53, and 55.

FLETCHER, J.; MALALASEKERA, W. Development of a user-friendly, low-cost home energy monitoring and recording system. *Energy*, Elsevier, v. 111, p. 32–46, 2016. Cited in page 52.

GIRI, S.; BERGÉS, M. An energy estimation framework for event-based methods in non-intrusive load monitoring. *Energy Conversion and Management*, Elsevier, v. 90, p. 488–498, 2015. Cited in page 33.

GRANDINI, M.; BAGLI, E.; VISANI, G. Metrics for multi-class classification: an overview. *arXiv preprint arXiv:2008.05756*, 2020. Cited in pages 30 and 42.

GRANDINI, M.; BAGLI, E.; VISANI, G. Metrics for multi-class classification: an overview. *ArXiv*, abs/2008.05756, 2020. Cited in page 60.

GU, X.; AKOGLU, L.; RINALDO, A. Statistical analysis of nearest neighbor methods for anomaly detection. *Advances in Neural Information Processing Systems*, v. 32, 2019. Cited in page 53.

HENRIQUES, J.; CALDEIRA, F.; CRUZ, T.; SIMÕES, P. Combining k-means and xgboost models for anomaly detection using log datasets. *Electronics*, Multidisciplinary Digital Publishing Institute, v. 9, n. 7, p. 1164, 2020. Cited in page 53.

HIMEUR, Y.; ALSALEMI, A.; BENSAALI, F.; AMIRA, A. Smart power consumption abnormality detection in buildings using micromoments and improved k-nearest neighbors. *International Journal of Intelligent Systems*, Wiley Online Library, v. 36, n. 6, p. 2865–2894, 2021. Cited in page 52.

HIMEUR, Y.; GHANEM, k.; ALSALEMI, A.; BENSAALI, F.; AMIRA, A. Artificial intelligence based anomaly detection of energy consumption in buildings: A review, current trends and new perspectives. *Applied Energy*, v. 287, p. 116601, 04 2021. Cited in pages 17, 20, 52, and 53.

HOLLINGSWORTH, K.; ROUSE, K.; CHO, J.; HARRIS, A.; SARTIPI, M.; SOZER, S.; ENEVOLDSON, B. Energy anomaly detection with forecasting and deep learning. In: IEEE. *2018 IEEE international conference on big data (Big Data)*. [S.l.], 2018. p. 4921–4925. Cited in page 53.

HONEINE, P. Online kernel principal component analysis: A reduced-order model. *IEEE transactions on pattern analysis and machine intelligence*, IEEE, v. 34, n. 9, p. 1814–1826, 2011. Cited in page 38.

HORI, M.; HARADA, T.; TANIGUCHI, R.-i. Anomaly detection for an elderly person watching system using multiple power consumption models. In: *ICPRAM*. [S.l.: s.n.], 2017. p. 669–675. Cited in page 55.

HOSSEINI, S. S.; AGBOSSOU, K.; KELOUWANI, S.; CARDENAS, A.; HENAO, N. A practical approach to residential appliances on-line anomaly detection: A case study of standard and smart refrigerators. *IEEE Access*, IEEE, v. 8, p. 57905–57922, 2020. Cited in pages 52, 53, and 59.

HOTELLING, H. Analysis of a complex of statistical variables into principal components. *Journal of educational psychology*, Warwick & York, v. 24, n. 6, p. 417, 1933. Cited in page 21.

HUANG, L.; CHEN, S.; LING, Z.; CUI, Y.; WANG, Q. Non-invasive load identification based on lstm-bp neural network. *Energy Reports*, Elsevier, v. 7, p. 485–492, 2021. Cited in page 33.

HYVÄRINEN, A.; KARHUNEN, J.; OJA, E. *Independent component analysis*. New York: John Wiley & Sons, 2001. v. 46. Cited in page 38.

IWAFUNE, Y.; MORI, Y.; KAWAI, T.; YAGITA, Y. Energy-saving effect of automatic home energy report utilizing home energy management system data in japan. *Energy*, Elsevier, v. 125, p. 382–392, 2017. Cited in page 52.

IZAKIAN, H.; PEDRYCZ, W. Anomaly detection in time series data using a fuzzy c-means clustering. In: IEEE. *2013 Joint IFSA world congress and NAFIPS annual meeting (IFSA/NAFIPS)*. [S.l.], 2013. p. 1513–1518. Cited in page 53.

JAKKULA, V.; COOK, D. Outlier detection in smart environment structured power datasets. In: IEEE. *2010 sixth international conference on intelligent environments*. [S.l.], 2010. p. 29–33. Cited in page 53.

JAKKULA, V.; COOK, D. Detecting anomalous sensor events in smart home data for enhancing the living experience. In: *Workshops at the twenty-fifth AAAI conference on artificial intelligence*. [S.l.: s.n.], 2011. Cited in pages 52 and 53.

JIANG, L.; LI, J.; LUO, S.; JIN, J.; WEST, S. Literature review of power disaggregation. *Proceedings of 2011 International Conference on Modelling, Identification and Control, ICMIC 2011*, 06 2011. Cited in page 34.

Jolliffe, I. *Principal component analysis*. New York: Springer Verlag, 2002. Cited in pages 21 and 23.

KAMARAJ, K.; DEZFOULI, B.; LIU, Y. Edge mining on iot devices using anomaly detection. In: IEEE. *2019 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*. [S.l.], 2019. p. 33–40. Cited in page 52.

KAMIńSKI, B.; JAKUBCZYK, M.; SZUFEL, P. A framework for sensitivity analysis of decision trees. *Central European Journal of Operations Research*, v. 26, p. 135 – 159, 03 2028. Cited in page 24.

Kampouropoulos, K.; Cárdenas, J. J.; Giacometto, F.; Romeral, L. An energy prediction method using adaptive neuro-fuzzy inference system and genetic algorithms. In: *2013 IEEE International Symposium on Industrial Electronics*. [S.l.: s.n.], 2013. p. 1–6. Cited in page 33.

KASELIMI, M.; PROTOPAPADAKIS, E.; VOULODIMOS, A.; DOULAMIS, N.; DOULAMIS, A. Multi-channel recurrent convolutional neural networks for energy disaggregation. *IEEE Access*, IEEE, v. 7, p. 81047–81056, 2019. Cited in page 33.

KIM, J. Fault detection for manufacturing home air conditioners using wavelet transform. *Int. J. Precis. Eng. Manuf.*, v. 17, p. 1299–1303, 2016. Cited in page 54.

KOLTER, J.; JOHNSON, M. Redd: A public data set for energy disaggregation research. *Artif. Intell.*, v. 25, 01 2011. Cited in pages 10, 34, 42, and 43.

KOLTER, J. Z.; JAAKKOLA, T. Approximate inference in additive factorial hmms with application to energy disaggregation. In: PMLR. *Artificial intelligence and statistics*. [S.l.], 2012. p. 1472–1482. Cited in page 49.

KONG, W.; DONG, Z. Y.; MA, J.; HILL, D. J.; ZHAO, J.; LUO, F. An extensible approach for non-intrusive load disaggregation with smart meter data. *IEEE Transactions on Smart Grid*, IEEE, v. 9, n. 4, p. 3362–3372, 2016. Cited in page 33.

KORBA, A. A.; KARABADJI, N. E. I. Smart grid energy fraud detection using svm. In: *Proceedings of the 2019 International Conference on Networking and Advanced Systems (ICNAS), Annaba, Algeria*. [S.l.: s.n.], 2019. p. 26–27. Cited in page 53.

KROMANIS, R.; KRIPAKARAN, P. Support vector regression for anomaly detection from measurement histories. *Advanced Engineering Informatics*, Elsevier, v. 27, n. 4, p. 486–495, 2013. Cited in page 53.

LAOUALI, I.; GOMES, I.; RUANO, M. d. G.; BENNANI, S. D.; FADILI, H. E.; RUANO, A. Energy disaggregation using multi-objective genetic algorithm designed neural networks. *Energies*, v. 15, n. 23, 2022. ISSN 1996-1073. Disponível em: <https://www.mdpi.com/1996-1073/15/23/9073>. Cited in page 17.

LE, T.-T.-H.; KIM, H. Non-intrusive load monitoring based on novel transient signal in household appliances with low sampling rate. *Energies*, v. 11, p. 3409, 12 2018. Cited in page 33.

LEE, H.-T.; SONG, J.-H.; MIN, S.-H.; LEE, H.-S.; SONG, K. Y.; CHU, C. N.; AHN, S.-H. Research trends in sustainable manufacturing: a review and future perspective based on research databases. *International Journal of Precision Engineering and Manufacturing-Green Technology*, Springer, v. 6, n. 4, p. 809–819, 2019. Cited in page 52.

LEE, S.; KIM, H. K. Adsas: Comprehensive real-time anomaly detection system. In: SPRINGER. *International Workshop on Information Security Applications*. [S.l.], 2018. p. 29–41. Cited in page 53.

LI, C.; ZHOU, D.; ZHENG, Y. Techno-economic comparative study of grid-connected pv power systems in five climate zones, china. *Energy*, Elsevier, v. 165, p. 1352–1369, 2018. Cited in page 52.

LI, S.; HAN, Y.; YAO, X.; YINGCHEN, S.; WANG, J.; ZHAO, Q. Electricity theft detection in power grids with deep learning and random forests. *Journal of Electrical and Computer Engineering*, Hindawi, v. 2019, 2019. Cited in page 53.

LI, S.; WANG, X.; SUN, X. Household appliance fault detection based on wavelet denoising and hht. *DEStech Transactions on Computer Science and Engineering*, 2018. Cited in page 54.

LIAO, J.; ELAFOUDI, G.; STANKOVIC, L.; STANKOVIC, V. Power disaggregation for low-sampling rate data. In: *2nd International Non-intrusive Appliance Load Monitoring Workshop, Austin, TX*. [S.l.: s.n.], 2014. v. 1, p. F1. Cited in page 49.

MAHAPATRA, B.; NAYYAR, A. Home energy management system (hems): concept, architecture, infrastructure, challenges and energy management schemes. *Energy Systems*, 11 2019. Cited in pages 17 and 32.

MAKONIN, S.; WANG, Z.; TUMPACH, C. Rae: The rainforest automation energy dataset for smart grid meter data analysis. *MDPI Data*, v. 3, p. 1–9, 02 2018. Cited in pages 10, 42, and 43.

MALLAT, S. *A Wavelet Tour of Signal Processing*. [S.l.]: Elsevier, 1999. Cited in pages 26, 28, and 58.

MAO, W.; CAO, X.; YAN, T.; ZHANG, Y. *et al.* Anomaly detection for power consumption data based on isolated forest. In: IEEE. *2018 international conference on power system technology (POWERCON)*. [S.l.], 2018. p. 4169–4174. Cited in page 53.

MASSIDDA, L.; MARROCU, M.; MANCA, S. Non-intrusive load disaggregation by convolutional neural network and multilabel classification. *Applied Sciences*, v. 10, n. 4, 2020. ISSN 2076-3417. Cited in page 33.

MATTERA, C. G.; SHAKER, H. R.; JRADI, M. Consensus-based method for anomaly detection in vav units. *Energies*, Multidisciplinary Digital Publishing Institute, v. 12, n. 3, p. 468, 2019. Cited in page 52.

MENGISTU, M. A.; GIRMAY, A. A.; CAMARDA, C.; ACQUAVIVA, A.; PATTI, E. A cloud-based on-line disaggregation algorithm for home appliance loads. *IEEE Transactions on Smart Grid*, IEEE, v. 10, n. 3, p. 3430–3439, 2018. Cited in page 33.

MITCHELL, T. M. Does machine learning really work? *AI Magazine*, v. 18, n. 3, p. 11, Sep. 1997. Disponível em: <https://ojs.aaai.org/index.php/aimagazine/article/view/1303>. Cited in page 20.

MORADZADEH, A.; SADEGHIAN, O.; POURHOSSEIN, K.; MOHAMMADI-IVATLOO, B.; ANVARI-MOGHADDAM, A. Improving residential load disaggregation for sustainable development of energy via principal component analysis. *Sustainability*, MDPI AG, v. 12, n. 8, p. 3158, Apr 2020. ISSN 2071-1050. Disponível em: <http://dx.doi.org/10.3390/su12083158>. Cited in pages 33, 37, and 49.

MUNIR, M.; SIDDIQUI, S. A.; CHATTHA, M. A.; DENGEL, A.; AHMED, S. Fusead: unsupervised anomaly detection in streaming sensors data by fusing statistical and deep learning models. *Sensors*, Multidisciplinary Digital Publishing Institute, v. 19, n. 11, p. 2451, 2019. Cited in page 53.

MURRAY, D.; STANKOVIĆ, L.; STANKOVIĆ, V. An electrical load measurements dataset of united kingdom households from a two-year longitudinal study. *Scientific Data*, v. 4, 2017. Cited in page 56.

NAGI, J.; YAP, K. S.; TIONG, S. K.; AHMED, S. K.; MOHAMAD, M. Nontechnical loss detection for metered customers in power utility using support vector machines. *IEEE transactions on Power Delivery*, IEEE, v. 25, n. 2, p. 1162–1171, 2009. Cited in page 55.

NATIONS, U. World urbanization process: The 2014 revision. In: _____. [S.l.: s.n.], 2014. p. 1–32. Cited in page 32.

OPPENHEIM, A. V. *Discrete-time signal processing*. [S.l.]: Pearson Education India, 1999. Cited in page 25.

OZA, P.; PATEL, V. M. One-class convolutional neural network. *IEEE Signal Processing Letters*, IEEE, v. 26, n. 2, p. 277–281, 2018. Cited in pages 52 and 53.

PARSON, O.; GHOSH, S.; WEAL, M.; ROGERS, A. Non-intrusive load monitoring using prior models of general appliance types. In: *Proceedings of the AAAI Conference on Artificial Intelligence*. [S.l.: s.n.], 2012. v. 26, n. 1. Cited in page 49.

PASTORE, V. P.; ZIMMERMAN, T. G.; BISWAS, S. K.; BIANCO, S. Annotation-free learning of plankton for classification and anomaly detection. *Scientific reports*, Nature Publishing Group, v. 10, n. 1, p. 1–15, 2020. Cited in page 53.

PEREIRA, J.; SILVEIRA, M. Unsupervised anomaly detection in energy time series data using variational recurrent autoencoders with attention. In: IEEE. *2018 17th IEEE international conference on machine learning and applications (ICMLA)*. [S.l.], 2018. p. 1275–1282. Cited in page 53.

PETLADWALA, M.; ISHII, Y.; SENDODA, M.; KONDO, R. Canonical correlation based feature extraction with application to anomaly detection in electric appliances. In: IEEE. *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. [S.l.], 2019. p. 2737–2741. Cited in page 53.

PICCIALLI, V.; SUDOSO, A. M. Improving non-intrusive load disaggregation through an attention-based deep neural network. *Energies*, Multidisciplinary Digital Publishing Institute, v. 14, n. 4, p. 847, 2021. Cited in page 49.

QI, P.; JOVANOVIC, S.; LEZAMA, J.; SCHWEITZER, P. Discrete wavelet transform optimal parameters estimation for arc fault detection in low-voltage residential power networks. *Electric Power Systems Research*, v. 143, p. 130–139, 2017. ISSN 0378-7796. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0378779616304059>. Cited in page 54.

RASHID, H.; SINGH, P. Monitor: An abnormality detection approach in buildings energy consumption. In: IEEE. *2018 IEEE 4th international conference on collaboration and internet computing (CIC)*. [S.l.], 2018. p. 16–25. Cited in page 52.

ROSSI, B.; CHREN, S.; BUHNOVA, B.; PITNER, T. Anomaly detection in smart grid data: An experience report. In: IEEE. *2016 ieee international conference on systems, man, and cybernetics (smc)*. [S.l.], 2016. p. 002313–002318. Cited in page 53.

RUANO, A.; HERNANDEZ, A.; UREñA, J.; RUANO, M.; GARCíA, J. Nilm techniques for intelligent home energy management and ambient assisted living: A review. *Energies*, v. 12, p. 2203, 06 2019. Cited in page 32.

SCHIRMER, P.; MPORAS, I. Statistical and electrical features evaluation for electrical appliances energy disaggregation. *Sustainability*, v. 11, p. 3222, 06 2019. Cited in page 35.

SHIN, C.; JOO, S.; YIM, J.; LEE, H.; MOON, T.; RHEE, W. Subtask gated networks for non-intrusive load monitoring. In: *Proceedings of the AAAI Conference on Artificial Intelligence*. [S.l.: s.n.], 2019. v. 33, n. 01, p. 1150–1157. Cited in page 49.

SIAL, A.; SINGH, A.; MAHANTI, A. Detecting anomalous energy consumption using contextual analysis of smart meter data. *Wireless Networks*, Springer, v. 27, n. 6, p. 4275–4292, 2021. Cited in pages 52 and 53.

SILVA, A. da; GUARANY, I.; ARRUDA, B.; GURJÃO, E. C.; FREIRE, R. A method for anomaly prediction in power consumption using long short-term memory and negative selection. In: IEEE. *2019 IEEE international symposium on circuits and systems (ISCAS)*. [S.l.], 2019. p. 1–5. Cited in page 53.

SINGH, S.; MAJUMDAR, A. Non-intrusive load monitoring via multi-label sparse representation-based classification. *IEEE Transactions on Smart Grid*, IEEE, v. 11, n. 2, p. 1799–1801, 2019. Cited in page 49.

STANKOVIC, V.; LIAO, J.; STANKOVIC, L. A graph-based signal processing approach for low-rate energy disaggregation. In: IEEE. *2014 IEEE symposium on computational intelligence for engineering solutions (CIES)*. [S.l.], 2014. p. 81–87. Cited in page 49.

TABATABAEI, S. M.; DICK, S.; XU, W. Toward non-intrusive load monitoring via multi-label classification. *IEEE Transactions on Smart Grid*, IEEE, v. 8, n. 1, p. 26–40, 2016. Cited in page 49.

TANG, J.; ZHANG, L. An adaptive feature recognition algorithm and its hardware accelerator for arc fault recognition: The hardware accelerator of the algorithm is improved from the discrete wavelet transform algorithm and is used for power grid arc fault detection. In: *Proceedings of the 8th International Conference on Computing and Artificial Intelligence*. New York, NY, USA: Association for Computing Machinery, 2022. (ICCAI '22), p. 375–379. ISBN 9781450396110. Disponível em: <https://doi.org/10.1145/3532213.3532269>. Cited in page 54.

THEODORIDIS, S.; PIKRAKIS, A.; KOUTROUMBAS, K.; CAVOURAS, D. *Introduction to pattern recognition: a matlab approach*. [S.l.]: Academic Press, 2010. Cited in page 38.

VORSATZ, D. Ürge; CABEZA, L. F.; SERRANO, S.; BARRENECHE, C.; PETRICHENKO, K. Heating and cooling energy trends and drivers in buildings. *Renewable and Sustainable Energy Reviews*, v. 41, p. 85–98, 2015. ISSN 1364-0321. Cited in page 17.

WANG, X.; YANG, I.; AHN, S.-H. Sample efficient home power anomaly detection in real time using semi-supervised learning. *IEEE Access*, IEEE, v. 7, p. 139712–139725, 2019. Cited in pages 53 and 55.

WANG, X.; ZHAO, T.; LIU, H.; HE, R. Power consumption predicting and anomaly detection based on long short-term memory neural network. In: IEEE. *2019 IEEE 4th international conference on cloud computing and big data analysis (ICCCBDA)*. [S.l.], 2019. p. 487–491. Cited in page 53.

WELIKALA, S.; DINESH, C.; EKANAYAKE, M. P. B.; GODALIYADDA, R. I.; EKANAYAKE, J. Incorporating appliance usage patterns for non-intrusive load monitoring and load forecasting. *IEEE Transactions on Smart Grid*, IEEE, v. 10, n. 1, p. 448–461, 2017. Cited in page 49.

WENG, Y.; ZHANG, N.; XIA, C. Multi-agent-based unsupervised detection of energy consumption anomalies on smart campus. *IEEE Access*, IEEE, v. 7, p. 2169–2178, 2018. Cited in page 53.

WU, X.; JIAO, D.; YOU, L. Nonintrusive on-site load-monitoring method with self-adaption. *International Journal of Electrical Power & Energy Systems*, Elsevier, v. 119, p. 105934, 2020. Cited in page 35.

YANG, C. C.; SOH, C. S.; YAP, V. V. A systematic approach in load disaggregation utilizing a multi-stage classification algorithm for consumer electrical appliances classification. *Frontiers in Energy*, Springer, v. 13, n. 2, p. 386–398, 2019. Cited in page 33.

YANG, D.; GAO, X.; KONG, L.; PANG, Y.; ZHOU, B. An event-driven convolutional neural architecture for non-intrusive load monitoring of residential appliance. *IEEE Transactions on Consumer Electronics*, IEEE, v. 66, n. 2, p. 173–182, 2020. Cited in pages 17 and 33.

YU, Q.; HU, Y.; YANG, Y. Identification method for series arc faults based on wavelet transform and deep neural network. *Energies*, v. 13, n. 1, 2020. ISSN 1996-1073. Disponível em: <https://www.mdpi.com/1996-1073/13/1/142>. Cited in page 54.

YUAN, Y.; JIA, K. A distributed anomaly detection method of operation energy consumption using smart meter data. In: IEEE. *2015 international conference on intelligent information hiding and multimedia signal processing (IIH-MSP)*. [S.l.], 2015. p. 310–313. Cited in page 53.

YUE, T.; LONG, R.; CHEN, H. Factors influencing energy-saving behavior of urban households in jiangsu province. *Energy Policy*, v. 62, p. 665–675, 2013. ISSN 0301-4215. Cited in page 17.

ZEIFMAN, M.; ROTH, K. Disaggregation of home energy display data using probabilistic approach. In: IEEE. *2012 IEEE International Conference on Consumer Electronics (ICCE)*. [S.l.], 2012. p. 630–631. Cited in pages 33 and 49.

ZHANG, C.; WANG, F. Multi-feature fusion based anomaly electro-data detection in smart grid. In: IEEE. *2018 15th international symposium on pervasive systems, algorithms and networks (I-SPAN)*. [S.l.], 2018. p. 54–59. Cited in page 53.

ZHANG, M.; WU, J.; LIN, H.; YUAN, P.; SONG, Y. The application of one-class classifier based on cnn in image defect detection. *Procedia computer science*, Elsevier, v. 114, p. 341–348, 2017. Cited in pages 52 and 53.

ZHANG, Y.; CHEN, W.; BLACK, J. Anomaly detection in premise energy consumption data. In: IEEE. *2011 IEEE power and energy society general meeting*. [S.l.], 2011. p. 1–8. Cited in page 53.

ZHANG, Y.; YIN, B.; CONG, Y.; DU, Z. Multi-state household appliance identification based on convolutional neural networks and clustering. *Energies*, v. 13, n. 4, 2020. ISSN 1996-1073. Cited in page 49.

ZHAO, B.; HE, K.; STANKOVIC, L.; STANKOVIC, V. Improving event-based non-intrusive load monitoring using graph signal processing. *IEEE Access*, PP, p. 1–1, 09 2018. Cited in page 49.

ZHAO, B.; STANKOVIC, L.; STANKOVIC, V. On a training-less solution for non-intrusive appliance load monitoring using graph signal processing. *IEEE Access*, IEEE, v. 4, p. 1784–1799, 2016. Cited in page 49.

ZHENG, Z.; YANG, Y.; NIU, X.; DAI, H.-N.; ZHOU, Y. Wide and deep convolutional neural networks for electricity-theft detection to secure smart grids. *IEEE Transactions on Industrial Informatics*, IEEE, v. 14, n. 4, p. 1606–1615, 2017. Cited in page 53.

ZHU, Y.; LU, S. Load profile disaggregation by blind source separation: A wavelets-assisted independent component analysis approach. In: IEEE. *2014 IEEE PES General Meeting| Conference & Exposition*. [S.l.], 2014. p. 1–5. Cited in pages 20 and 33.

ZOHA, A.; GLUHAK, A.; IMRAN, M.; RAJASEGARAR, S. Non-intrusive load monitoring approaches for disaggregated energy sensing: A survey. *Sensors (Basel, Switzerland)*, v. 12, p. 16838–16866, 12 2012. Cited in pages 19 and 32.