Non-intrusive load monitoring based on event detection and unsupervised learning for airport baggage handling systems

Miguel Angel Bermeo-Ayerbe[1,2], David Antonio Cruz-Rangel[2], Javier Diaz-Rozo[2], and Carlos Ocampo-Martinez[1]

[1] *Universitat Politècnica de Catalunya*, Barcelona, Spain,
`miguel.angel.bermeo@upc.edu`,
[2] *Aingura IIoT*, Donostia-San Sebastian, Spain,

Abstract. With a non-intrusive load monitoring paradigm, this paper poses the first steps to monitor the health of airport baggage handling systems. This goal is reached by measuring the energy consumption of the electrical cabinets that power a set of conveyor belt systems. Therefore, using an energy disaggregation approach, each motor in the conveyor system can be monitored. The proposed methodology consists of an algorithm to detect power-on/off events and how those events can be clustered by characterizing transient states, employing unsupervised clustering algorithms. Energy measurements were filtered to remove noise, and the on-off events were detected and characterized. The power-on and -off events were clustered, using k-means and Gaussian mixture model (GMM), showing similar grouping behaviors with discrepancies for labeling samples at the frontier of the clusters. Since the GMM provides more information for samples with ambiguity in overlapping clusters, its results are presented and analyzed. From each cluster important insight are extracted in terms of energy consumption.

Keywords: Non-intrusive load monitoring, Event detection, Gaussian mixture models, k-means clustering, Unsupervised learning

1 Introduction

The contribution of this work is a non-intrusive load monitoring (NILM) methodology of multiple conveyor belt systems, which will allow the detection and disaggregation of each belt conveyors from a single source of electrical measurement. This methodology focuses on detecting on-off events and clustering them, with an unsupervised algorithm, to subsequently monitor them. In this way, this methodology lays the foundation for implementing multiple conveyor condition monitoring at airport baggage handling systems (ABHSs) using electricity consumption with the minimum number of sensors. Furthermore, the energy desegregation of a line of conveyor belt systems can help to understand their consumption and detect energy inefficiencies that can be mitigated, promoting energy efficiency in ABHSs.

2 Problem Statement

ABHSs are made up of many conveyor belt systems for transporting baggage, which are segmented by areas, *e.g.,* check-in area, merging area, security inspection site/screening area, sorting area, loading area [1]. Those areas have many electrical cabinets to supply electricity to a set of conveyor systems, as well as photoelectric sensors, programmable logic controllers (PLCs), and variable frequency drives (VFDs). Therefore, the energy consumption of each electrical cabinet is the superposition of the consumption of motors of each conveyor system and an aggregate constant consumption of sensors and PLCs, *i.e.,*

$$p_t(k) = p_o(k) + p_n(k) + \sum_{i=0}^{n_m} p_{m,i}(k), \tag{1}$$

being $k \in \mathbb{Z}_{\geq 0}$ the discrete-time index, $p_t(k) \in \mathbb{R}$ the total power consumption of electrical cabinet, $p_o \in \mathbb{R}$ the constant consumption (sensors, PLCs, others), n_m the number of motors connected, $p_{m,i} \in \mathbb{R}$ the power consumption of i-th motors and p_n the noise in the energy consumption of the electricity supplier network. The energy consumption of each motor ($p_{m,i}$) has a dynamic behavior with transient and stationary states. The transient state is characterized by having an overshoot, settling time, damped oscillation, and slew rate. Note that the VFDs are in the middle of the electrical cabinet and the motors. Their consumption is divided into the motor consumption and the constant consumption of the VFD electronics. Thereby, the VFD consumption is distributed respectively in equation (1).

Following a NILM scheme, a three-phase smart meter was installed in an electrical cabinet of Bilbao airport ABHS (Spain). The installed smart meter acquires measurements at a sampling frequency of 8 kHz for one month. The electrical cabinet supplies energy to 40 conveyor belt systems, whose motors have different specifications, and their individual energy footprints are unknown. Additionally, the PLC configuration handles multiple motors simultaneously, overlapping their power consumption in the transient state, challenging its energy disaggregation and labeling process.

3 Proposed approach

As a first approach to analyze this power consumption and label motors in a NILM scheme, an event-based methodology is proposed to detect switching on or off of motors. With the on-off events detected, an unsupervised learning process was carried out to cluster each event.

3.1 Power signal pre-processing

The first step in this methodology is to pre-process the power signal in order to reduce signal noise. This preprocessing stage requires inspecting the spectrum of the power signal. This inspection looks for those frequency bands that hold over

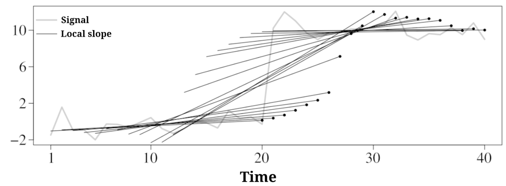


Fig. 1: Moving slope evolution along a transient state of a signal [3].

time and do not depend on the activation/deactivation of the motors. Therefore, those frequency bands must be filtered with a notch filter. Another strategy to eliminate noise is to perform a median filter [2].

3.2 Event detection

The following event detection methodology was developed to detect transient states and estimate when they start and stop. Assume constant power consumption that can be described by a straight line as follows:

$$p(k) = \alpha + \beta k + \eta, \tag{2}$$

being p the power consumption at time instant k, α the constant consumption level, β the slope of the straight line that would be close to zero for constant consumption, and η the additive white Gaussian noise. When any load is switched on/off, those parameters change continuously until reaching a steady state, then both α and β are set to new constant values. Furthermore, when those transient states start or stop, the signal undergoes a sudden change that is perceived with high slope values, which are positive or negative when the signal increases or decreases, respectively (see Figure 1). Thereby, with a sliding window with length n_w, a slope signal can be computed as $\beta_p(k) = \beta(\boldsymbol{t_w}(k), \boldsymbol{P}_w(k))$, where $\boldsymbol{t_w}(k) = [k - n_w, \ldots, k + n_w]$ and $\boldsymbol{P_w}(k) = [p(k + n_w), \ldots, p(k - n_w)]$ are time and power windows with length n_w, respectively. The slope function $\hat{\beta}$ is computed for n values as

$$\beta(\boldsymbol{x}, \boldsymbol{y}) = \frac{\sum_{i=0}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=0}^{n}(x_i - \bar{x})^2}, \tag{3}$$

being $\bar{x} = \text{Mean}(\boldsymbol{x})$ and $\bar{y} = \text{Mean}(\boldsymbol{y})$ the means of vectors $\boldsymbol{x}$ and $\boldsymbol{y}$, respectively. The mean function for a $\boldsymbol{x}$ vector with n values is computed as $\text{Mean}(\boldsymbol{x}) = \frac{1}{n}\sum_{i=0}^{n} x_i$. Performing an inspection of the slope signal $\beta_p(k)$, any time instant k is collected in vector $\boldsymbol{I_{on}}$, whilst $\beta_p(k)$ is a positive peak with a value greater than $\phi_{on} \in \mathbb{R}$. Thus, $\boldsymbol{I_{on}}$ has the time instants with power-on event, according to the threshold ϕ_{on} that defines the level of relevant activation events. Similarly, a vector $\boldsymbol{I_{off}}$ is computed for negative peaks less than $\phi_{off} \in \mathbb{R}$.

With $\boldsymbol{I_{on}}$ and $\boldsymbol{I_{off}}$ vectors, the next step is to determine the time instance k when the transient states start and stop. For load activation events, each time instant of $\boldsymbol{I_{on}}$ is evaluated following this procedure: The instance where the transient state starts is determined by looking back at the first negative peak in the $\beta_p(k)$ signal. To estimate where to stop the transient state, it is required to calculate the angle signal $\theta_p(k) = \theta(\boldsymbol{t_w}(k), \boldsymbol{P}_w(k))$ signal, where θ function is define as $\theta(\boldsymbol{x},\boldsymbol{y}) = \arctan\left(\beta(\boldsymbol{x},\boldsymbol{y})\right)$. Therefore, the transient state stops when a peak in $\theta_p(k)$ is lower than $\theta_{on} \in \mathbb{R}$, being θ_{on} the minimum peak of $\theta_p(k)$ to be considered as a transient state. In this way, a power-on event vector is obtained as

$$\boldsymbol{E_{on}} = \begin{bmatrix} k^{(1)}_{on,1} \; k^{(1)}_{on,2} \ldots k^{(1)}_{on,i_{on}} \ldots k^{(1)}_{on,n_{on}} \\ k^{(2)}_{on,1} \; k^{(2)}_{on,2} \ldots k^{(2)}_{on,i_{on}} \ldots k^{(2)}_{on,n_{on}} \end{bmatrix}^T,$$

for $i_{on} \in \{1,\ldots,n_{on}\}$, being n_{on} the number of power-on events detected, while $k^{(1)}_{on,i_{on}}$ and $k^{(2)}_{on,i_{on}}$ are the start and stop times of each transient state, respectively.

For shutdown events, a new slope signal $\beta_{\mathrm{Var}_p}(k) = \beta(\boldsymbol{t_w}(k), \mathrm{Var}(\boldsymbol{P}_w(k)))$ should to be computed, being $\mathrm{Var}(\boldsymbol{x}) = \frac{1}{n}\sum_{i=0}^{n}\left(x_i - \bar{x}\right)^2$ the variance function of $\boldsymbol{x}$ vector. Thus, the start and stop k instants of the shutdown events $\boldsymbol{I_{off}}$ correspond to the positive and negative closed peak of the $\beta_{\mathrm{Var}_p}(k)$ signal, respectively. Consequently, the power-off event vector is obtained as

$$\boldsymbol{E_{off}} = \begin{bmatrix} k^{(1)}_{off,1} \; k^{(1)}_{off,2} \ldots k^{(1)}_{off,i_{off}} \ldots k^{(1)}_{off,n_{off}} \\ k^{(2)}_{off,1} \; k^{(2)}_{off,2} \ldots k^{(2)}_{off,i_{off}} \ldots k^{(2)}_{off,n_{off}} \end{bmatrix}^T,$$

for $i_{off} \in \{1,\ldots,n_{off}\}$, where n_{off} is the number of power-off events detected, while $k^{(1)}_{off,i_{off}}$ and $k^{(2)}_{off,i_{off}}$ are the start and stop times of i_{off}-th event, respectively.

3.3 Clustering analysis

Throughout the process of characterizing edge events, similar events can be grouped, using clustering techniques. In this way, considering a tolerance error, this process can indicate when a set of motors are turned on/off, using only total energy consumption, which enables to implement a condition monitoring approach. The methodology proposed is to extract characteristics of each event and use an unsupervised clustering algorithm to group the different types of events, either power-on or -off.

Features extraction

Combining the on and off events, a set of indicators is proposed to characterize each edge event. The characteristic extraction stage consists in analyzing the transient state and the changes in the steady states of energy signals due to a given event. For this analysis, four time instants were considered for each event: $k_{i_e,0}$, $k_{i_e,1}$, $k_{i_e,2}$ and $k_{i_e,3}$, for $i_e \in \{1,\ldots,n_e\}$, being n_e the total number of events. The time interval between $k_{i_e,0}$ and $k_{i_e,1}$ corresponds to the steady state prior to the i_e-th event, between times $k_{i_e,1}$ and $k_{i_e,2}$, the transient state occurs of the i_e-th event, and finally, the interval $k_{i_e,2}$ to $k_{i_e,3}$ belongs to the steady state that occurs just after the i_e-th event. Therefore, the following features have been extracted for each event:

(i) *Maximum relative peak.* This feature measures the maximum value of power added during the i_e-th event, *i.e.,*

$$f_{1,i_e} = \max(\boldsymbol{P}_{k_{i_e,1}:k_{i_e,2}}) - \mathrm{Mean}(\boldsymbol{P}_{k_{i_e,0}:k_{i_e,1}}), \tag{4}$$

where the subscripts $\bullet_{k_1:k_2}$ indicate the portion of the signal considered as a matrix in the time interval $[k_1, k_2]$.

(ii) *Minimum relative peak.* This feature measures the minimum value of power during the i_e-th event relative to steady state prior the event,

$$f_{2,i_e} = \min(\boldsymbol{P}_{k_{i_e,1}:k_{i_e,2}}) - \mathrm{Mean}(\boldsymbol{P}_{k_{i_e,0}:k_{i_e,1}}). \tag{5}$$

(iii) *Settling time.* Duration of the transient state,

$$f_{3,i_e} = k_{i_e,2} - k_{i_e,2}. \tag{6}$$

(iv) *Delta mean active power.* The mean added/subtracted to the total active power signal due to the i_e-th event,

$$f_{4,i_e} = \mathrm{Mean}(\boldsymbol{P}_{k_{i_e,2}:k_{i_e,3}}) - \mathrm{Mean}(\boldsymbol{P}_{k_{i_e,0}:k_{i_e,1}}). \tag{7}$$

(v) *Delta variance active power.* The variance added/subtracted to active power signal due to the i_e-th event,

$$f_{5,i_e} = \mathrm{Var}(\boldsymbol{P}_{k_{i_e,2}:k_{i_e,3}}) - \mathrm{Var}(\boldsymbol{P}_{k_{i_e,0}:k_{i_e,1}}). \tag{8}$$

(vi) *Delta RMS filtered current.* The added/subtracted distortion of the i_e-th event was characterized via the delta root mean square (RMS) of filtered current (without fundamental components at 50Hz), *i.e.,*

$$f_{6,i_e} = \mathrm{Mean}\left(\mathrm{RMS}(\boldsymbol{I}_{f,k_{i_e,2}:k_{i_e,3}})\right) - \mathrm{Mean}\left(\mathrm{RMS}(\boldsymbol{I}_{f,k_{i_e,0}:k_{i_e,1}})\right), \tag{9}$$

being $\boldsymbol{I}_f \in \mathbb{R}^3$ the filtered three-phase current and RMS equation defined as $\mathrm{RMS}(\boldsymbol{x}) = \sqrt{\frac{1}{n}\sum_{i=1}^{n} x_i^2}$.

Unsupervized clustering analysis

With the matrix $\boldsymbol{F}_{on-off}$ of features computed by event, $\boldsymbol{F}_{on-off}$ is divided into two matrices $\boldsymbol{F}_{on}$ and $\boldsymbol{F}_{off}$. Thus, using an unsupervised clustering algorithm, both power-on and -off events can be clustered, executing the clustering algorithm for $\boldsymbol{F}_{on}$ and $\boldsymbol{F}_{off}$ individually. In this way, power-on events can be grouped and labeled, allowing monitoring each group of events (similarly for power-off events) and subsequently matching power-on events to power-off events. For the purpose of this paper, two clustering algorithms were considered: the k-means learning and the Gaussian Mixture Model (GMM). The k-means algorithm is based on pairwise Euclidean distances between data points, whereas GMM is a parametric probability density function represented as a weighted sum of multi-dimensional Gaussian probability distributions that models the probability distribution of data [4].

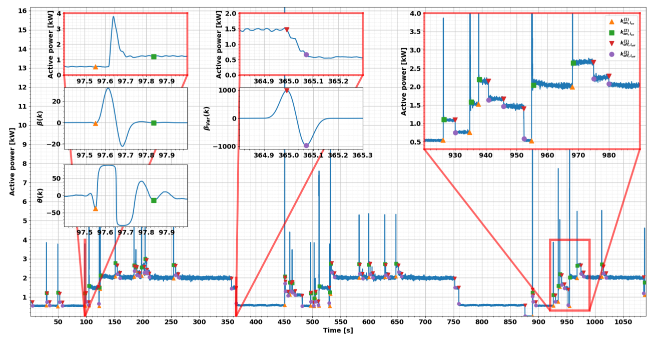
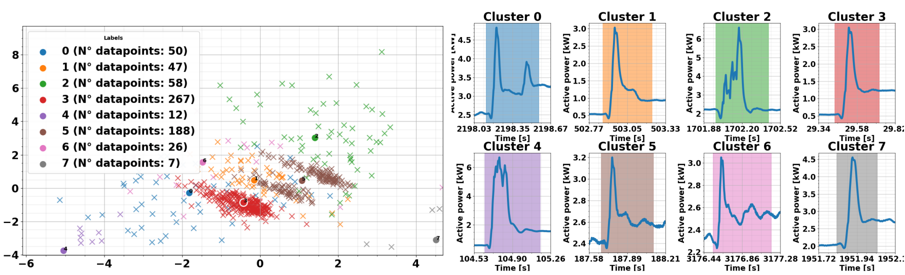


Fig. 2: ON-OFF event detection.



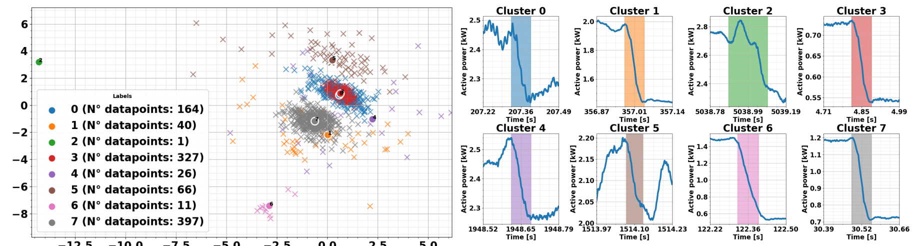
(a) (b)

Fig. 3: Clustering analysis for power-on events: (a) MDS analysis of power-on event clusters. (b) Dynamic behavior for each clustered event.

4 Results

The proposed methodology was validated with one and a half hours of electrical measurement of an ABHS electrical cabinet at Bilbao airport (Spain). These measurements have a sampling frequency of 8 kHz. After an analysis of the frequency spectrum of the current and power signals, the main information is concentrated in the band 0 to 500 Hz. Thereby, the measures were resampled at 1 kHz to reduce the amount of samples. Furthermore, a sanity check was performed on the signal to guarantee data quality. Pre-processing with filtering techniques was carried out to reduce signal noise and increase data quality.

The edge event detection parameters established were: windows length n_w of 0.1s, activation threshold $\phi_{on} = 5$, deactivation threshold $\phi_{off} = 2$, and minimum slope angle $\theta_{on} = 30°$. As Figure 2 shows, the ON-OFF switching events were detected properly, detecting 655 and 1032 power on/off events, respectively. The first column of Figure 2 displays an example of a power-on event with slope signal $\beta(k)$ and angle signal $\theta(k)$; the start and stop of the transient state are marked according to the explained methodology. Similarly, the second column shows a shutdown event with $\beta_{\mathrm{Var}}(k)$ signal. Note that shutdown events have a positive and negative peaks in the $\beta_{\mathrm{Var}}(k)$ signal when the events start and stop, respectively. Whereas with $\beta(k)$ and $\theta(k)$ signals for power-on events, start and stop instances are unclear as power-on events have more fluctuations and more cases need to be considered, *e.g.,* when turned on several engines at the same



(a) (b)

Fig. 4: Clustering analysis for shutdown events: (a) MDS analysis of shutdown event clusters. (b) Dynamic behavior for each clustered event.

time. Furthermore, although the power signals were filtered, the biggest challenge in event detecting in this dataset has been discerning between high noise variations and power-on/off events. This is due to the noise from each conveyor system overlapped, resulting in a total noise comparable to a motor consumption. When the noise variations were comparable to switch on-off events and detected as event, those events were filtered comparing the power signal offset before and after the event, *i.e.,* if the variation of the power signal offset is less than the minimum consumption of the conveyor belt systems (70W), the event is removed.

For the unsupervised analysis, the clustering algorithms, the k-means learning and the Gaussian Mixture Model (GMM), were executed. The clusters resulting from both algorithms were compared to map the closest ones, which resulted in around 78% of the events being labeled in the same group for both on and off events. Since the rest of the 22% of the events belong to boundaries between clusters, the GMM approach is able to provide probabilistic information about each sample related to each cluster, which represents an advantage for labeling samples with ambiguity in overlapping clusters and for measuring the stability of data. Therefore, Figure 3 and 4 show the clustering results using a GMM for power-on/off events, respectively. The number of clusters was determined by computing several GMMs with different number of cluster and measuring their Bayesian information criterion (BIC), *i.e.,* ten GMMs were calculated for two components, for three and so on, up to 20 components. The median of BIC was calculated for models with the same cluster number, for which the number of components that have a minimum median of BIC is selected. This procedure was performed for power-on and -off events, resulting in eight components for both cases.

The Multidimensional scaling (MDS) algorithm was used to visualize the instances of the clustered event in a two-dimensional space [5]. Figure 3a shows the MDS for power-on events, specifying the number of events that belong to each group. Moreover, Figure 3b presents an example of the dynamic behavior of power-on events that were characterized and grouped in each cluster. Note that each type of event has different dynamics and some patterns considering multiple motor activations, which are repeated several times along the dataset, indicating that the programmable logic controller (PLC) surely has the rule to activate those engines when baggage is detected near those conveyor belt systems. Similarly, Figure 4 shows the MDS visualization, dynamic behaviors, and active power labeling for shutdown events, respectively. The MDS visualization of shutdown events has more overlap between clusters than power-on events, this may be because many shutdown events share more similarities but belong to different clusters or due to the view of MDS visualization. An interesting case is the cluster with only one event, giving the same result with the k-means algorithm. This case was the shutdown event that took the longest setting time.

5 Conclusions

ON-OFF event detection is proposed with a nonintrusive load monitoring scheme in airport baggage handling systems, which detects when a set of loads has been power-on or -off and captures when the transient state event started and stopped. Moreover, it is proposed how to characterize these ON-OFF events and group them. In this way, those ON-OFF events can be characterized and clustered to identify which loads are on or off, allowing a condition monitoring approach to be implemented for those loads. Thus, in the case of a failure, the first sign of failure can be detected and alerted, improving the maintenance process and reducing costs related to unplanned shutdowns in the airport system.

This methodology was validated with measurements from an electrical cabinet that feeds multiple conveyor belt systems at Bilbao airport. The energy signals were preprocessed to eliminate noise. With these filtered energy signals, the detection of ON-OFF events was carried out and for each detected event, and a set of characteristics was calculated. Subsequently, two clustering algorithms, k-means and Gaussian mixture model (GMM), were tested. The clustering process of GMM manages to differentiate events that share similar features but belong to different clusters, showing an advantage over the k-means algorithm. On the other hand, it is crucial to highlight the importance of ON-OFF detections and the sensitivity of the clustering procedure when the transient states start and stop are not set correctly, as this can be affected the feature extraction, consequently perform a clustering incorrectly. Finally, with this proposed methodology, future work will focus on matching the on and off events to estimate the activation/deactivation signal for each load.

6 Acknowledgments

This work has been supported by the Doctorats Industrials program from the Catalan Government (2019 DI 4) and TOFMAN project[3]. The authors would like to thank the company Aena SME, S.A. company that manages Spanish airports and heliports of general interest, for the support related to Data acquisition.

[3] https://www.iri.upc.edu/project/show/241

References

1. Wu, X., Xie, L.: Towards an IIoT-Based Architecture for Baggage Handling Systems. Journal of Communications **12**(8), 475–481 (2017)
2. 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 **13**(2), 386–398 (2019)
3. Borowski, M., Fried, R.: Online signal extraction by robust regression in moving windows with data-adaptive width selection. Statistics and Computing **24**(4), 597–613 (2014)
4. Geoffrey McLachlan, D.P.: Finite Mixture Models, 1 edn. Wiley Series in Probability and Statistics. Wiley-Interscience (2000)
5. Bishop, C.M.: Pattern Recognition and Machine Learning. Springer (2006)