MINING ANOMALOUS ELECTRICITY CONSUMPTION USING ENSEMBLE EMPIRICAL MODE DECOMPOSITION

Romain Fontugne[1,2], *Nicolas Tremblay*[3], *Pierre Borgnat*[3], *Patrick Flandrin*[3], *Hiroshi Esaki*[1]

[1] The University of Tokyo [2] JFLI, Tokyo; Japan
[3] CNRS, Ecole Normale Supérieure de Lyon, Physics Laboratory; France

ABSTRACT

Sensor deployments in large buildings allow the administrators to supervise the building infrastructure and identify abnormalities. Nevertheless, the numerous data streams reported by the increasing number of sensors overwhelm the building administrators. We propose a methodology that assists them to identify abnormal devices usages. The proposed method takes advantage of Ensemble Empirical Mode Decomposition (E-EMD) to uncover the patterns of power-draw signals, thereby enabling us to estimate the intrinsic inter-device correlations. By monitoring the devices correlations over time we compute the usual usage of the devices and report the devices that deviate from their normal usage. Our evaluation with 10 weeks of real data shows the efficiency of the proposed method to uncover the devices intrinsic relationships and detect peculiar events that require the administrators attention.

***Index Terms*—** Empirical Mode Decomposition, Energy Consumption, Anomaly Detection

1. INTRODUCTION

Thanks to the mass production of the sensors, modern buildings contain an increasing number of monitoring tools on which administrators rely to remotely identify devices failures or misuses that compromise the buildings operations and are potential energy waste. However, in large buildings, these sensors provide an excessive number of data streams; moreover, accounting for the daily and weekly patterns of human activity, the corresponding signals are usually non-stationary; finally, due to the variety of monitored devices (light, heating, ventilation, elevators,...) and the complexity of human activity, modeling them is hardly possible.

The aim of this work is to propose a method to monitor energy consumption in a building and pinpoint peculiar events that require the administrators attention, without any need for models of the typical sensors' signals. Recent proposals take advantage of frequency analysis to monitor the devices energy consumption and identify abnormalities. A typical approach is to decompose the device's consumption using Fourier transform and detect the outliers using clustering techniques [1, 2, 3]. The problem is that these methods suffer from a high false positive rate, because they assume a constant periodicity in the data (e.g. office hours and weekdays) which is often incorrect (e.g. extra hours at work, holidays). Another take at this problem would be to use fault detection in buildings, for which there has been numerous works [4, 5, 6]. However, it is hard to find in the literature a suitable method that stays usable in a non-stationary context or when there is no model at all.

Our proposal is fundamentally different in that it combines non-stationary analysis of the power-draw signals with a method for anomaly detection that relies on some computed devices' intrinsic relationships. The intuition is that each service provided by the building requires a minimum subset of devices. Within a subset, devices should be used at the same time when the corresponding service is needed, and anomalies are to be found when devices do not share their usual activity pattern. Energy saving opportunities could hence appear; for instance, office comfort is attained through sufficient lighting, ventilation, and air conditioning and they should all be turned off together when the room is emptied. If not, this has to be detected as an anomaly. However, because of the lack of models, one should not expect prior knowledge of the behavior of the devices.

Technically, the first ingredient is to use complete Ensemble Empirical Mode Decomposition (E-EMD) [7, 8], a non-stationary decomposition of signals, so as to exhibit patterns of activities of devices in various domains in frequency. Thanks to that, we are able to estimate correlations between signals in spite of their non-stationarity. By monitoring the inter-device correlations over time we extract the normal behavior of the devices, then we detect anomalous electricity consumption by identifying sudden drops of correlation between devices that are usually highly correlated.

2. ANALYZED DATA

Data collected at the Engineering Building 2 of the University of Tokyo will be used to illustrate and validate the method. The building is a 12-story building where electricity consumption of the lighting and HVAC (Heat Ventilation Air Conditioner) systems of 231 rooms is monitored by 135 sensors. The HVAC systems are of two types: EHP (Electrical Heat Pump) and GHP (Gas Heat Pump). The 5 GHPs of the dataset serve 154 rooms, whereas the EHP and lighting systems serve only pairs of rooms and they are independently controlled by the rooms users.

The dataset contains 10 weeks of data starting from June 27^{th} 2011 and ending on September 5^{th} 2011. This period of time is particularly interesting for two reasons; in Tokyo, the summer is the most energy-demanding season, and, the building manager actively works to curtail energy usage as much as possible due to the Fukushima nuclear accident.

Furthermore, this dataset is a valuable ground truth to evaluate the proposed method. Since the light and HVAC of the rooms are directly controlled by the room's occupants, we expect the method to uncover verifiable devices relationships.

3. METHODOLOGY FOR ANOMALY MINING

The activity of the devices in a building is mainly driven by the building office hours and working days. Therefore, they share similar daily and/or weekly patterns that hide the intrinsic correlations

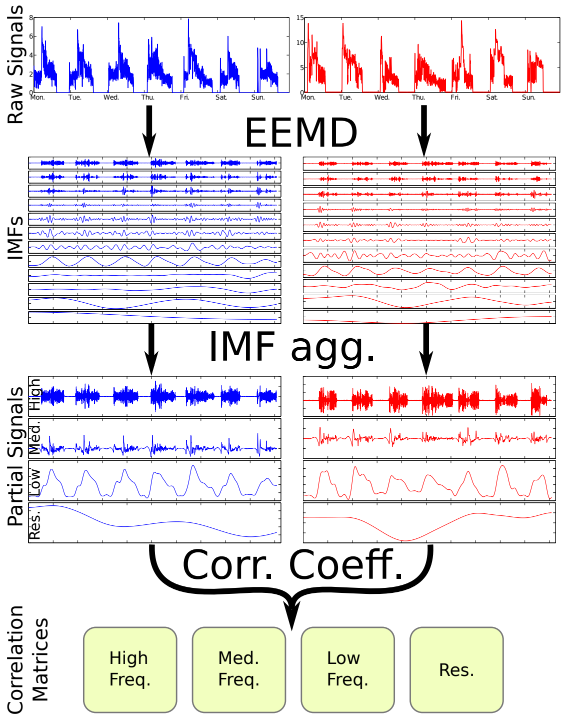


Fig. 1. Overview of the steps 1 to 3 of the proposed method for uncovering devices intrinsic relationships.

of their use and could mislead simple approaches to group devices according to their use, as already observed in [9]. The method proposed here uses non-stationary analysis with E-EMD so as to separate signals in modes for which estimating correlations is relevant. It consists of 4 steps: 1) time series from each device are decomposed in oscillatory functions (called IMFs, see 3.1) using E-EMD; 2) a careful study shows how to aggregate IMFs into several modes of devices' behavior; 3) correlations between devices are estimated on each mode, so as to uncover the intrinsic relationships between devices; 4) a robust anomaly detection method is used to mine anomalous device's behaviors. The first 3 steps are sketched in Fig. 1.

3.1. Decomposition with E-EMD of energy signals

Empirical Mode Decomposition (EMD) [10] is an adaptive and data-driven method designed to decompose non-stationary and/or non-linear signals without requiring to set a priori basis functions or tuning parameters. It finds an additive decomposition of a signal into a set of non-stationary oscillatory components called intrinsic mode functions (IMFs) satisfying two conditions: (1) an IMF contains the same number of extrema and zero crossings (or differ at most by one); (2) the amplitude of its oscillations are symmetric around zero. For that, EMD proposes an iterative algorithm that extracts IMFs one by one, starting from the IMF with locally the highest frequency that can be extracted from the signal. The computed IMF is removed from the signal and the residual is used as input for the next iteration. The process stops when the residual becomes a monotonic function, from which no more IMF can be extracted. The algorithm is recalled in Table 1; an implementation is available in the EMD toolbox: `http://perso.ens-lyon.fr/patrick.flandrin/emd.html`. The result of EMD on a signal X(t) is a set of IMFs

- Input signal $X(t)$; initialize $X_1 = X$ and $k = 1$.
- Iteratively on k, compute $c_k(t)$ by using Sifting on ρ_m with $\rho_1 = X_k$:
  1. Identify all the local maxima and minima of $\rho_m(t)$.
  2. Compute its upper (resp. lower) envelopes by an interpolation of the local maxima (resp. minima) using interpolation.
  3. Compute the local trend $Q_m(t)$ as the average of the upper and lower envelopes.
  4. Extract the local oscillations and update $\rho_m = \rho_m(t) - Q_m(t)$.
  5. If ρ_m is not an IMF (satisfying the two required conditions), iterate Sifting from step 1 for ρ_{m+1}, until the outcome is an IMF. If $\rho_m(t)$ is an IMF, set $c_k(t) = \rho_m(t)$, and extract the residual $X_{k+1}(t) = X_k(t) - c_k(t)$.
- Increase k to $k+1$ and go back at applying Sifting on X_k.
- Stop the iterations on k when residual $X_{k+1}(t)$ has no more oscillations, at index $K+1$ and set the final residual to be $r_K(t) = X_{k+1}(t)$

Table 1. Summary of classical EMD algorithm [10].

c_i and the final residue r_K, such as: $X(t) = \sum_{k=1}^{K} c_k(t) + r_K(t)$.

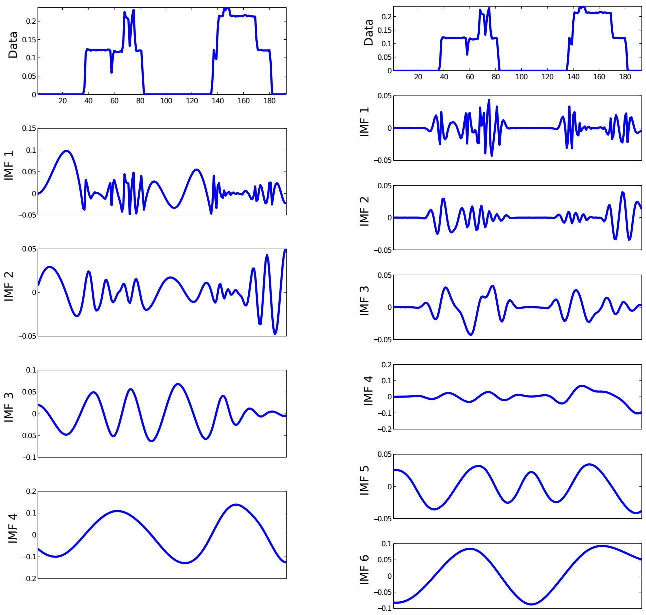
Applying classical EMD to typical power-draw signals, some unexpected difficulties appear because these signals feature long intervals of consecutive zeros. Fig. 2a depicts the results of EMD with the 2-day electricity consumption of the lights in an office. Since the office is unoccupied at night, the lights are turned off, and the corresponding signal is at 0. These flat parts of the signal are problematic for classical EMD: it tries to decompose the flat part in oscillations (as can be seen especially for IMF1 and IMF2 of Fig. 2a) which are irrelevant. Also, this causes a kind of mode mixing because low frequency appears over these intervals in otherwise high frequency IMFs.

In order to overcome this difficulty, we take advantage of a variant of EMD called Ensemble Empirical Mode Decomposition (E-EMD) [7] that was further refined in [8] as a complete E-EMD. This algorithm computes EMD on an ensemble of white noise-added signal and the mean result is the final IMFs. Because of the added noise, the sifting algorithm sees extrema on the intervals where the signal is flat, and extracts here only noise instead of low frequency oscillations. This solves the aforementioned problem and provides relevant IMFs, which are almost flat where the signal is zero. Fig. 2b depicts the result of complete E-EMD for the same 2-day long light energy consumption, and one can see that E-EMD provides more meaningful results than the classical EMD.

3.2. Aggregation of IMFs in modes

Applying complete E-EMD to energy consumption signals, we end up with a set of IMFs that are the consumption patterns of each device. In order to understand the meaning of the consumption patterns, we need to be able to compare the behaviors of different devices at corresponding frequencies. For that, we aggregate IMFs into modes defining relevant frequency domains. These domains, described hereafter, are decided by manual inspections of the IMFs of different power-draw signals, and could not have been set before performing such a decomposition, or doing simple filtering in the frequency domain. We observe four frequency domains that highlight different characteristics of the devices and we sum together the IMFs in each domain. Let us write $X_j^B(t)$ the mode in the frequency domain B (standing for H, M, L or T respectively for the four domains proposed hereafter) of device j (j going from 1 to d the total number of devices):

- *High frequencies* ($B = H$): the sum of the IMFs with a mean pe-



(a) EMD

(b) Ensemble EMD

Fig. 2. Results of EMD and E-EMD for the same signal.

riod lower than 20 minutes capture the noise.
- *Medium frequencies* ($B = M$): the sum of the IMFs with a mean period between 20 minutes and 6 hours stand for the patterns with a period shorter than the usual office hours thus they convey the detailed devices usage.
- *Low frequencies* ($B = L$): the sum of the IMFs with a mean period between 6 hours and 6 days represent the daily pattern common to all the devices.
- *Residual Trends* ($B = T$): the residual data r_K obtained by E-EMD highlights the trend of the devices.

Note that, in the example shown in Fig. 1, the correlation coefficient of the raw signals X_i and X_j is 0.57 and would suggest that they are highly correlated. However, the comparison of the corresponding modes provides new insights; the two devices are poorly correlated at high and medium frequencies (respectively -0.01 and -0.04 for modes X^H and X^M) but highly correlated at low frequencies (correlation for X^L's is 0.79) meaning that these devices share a similar daily pattern, but do not have an intrinsic correlation. It is thus important to separate these modes to analyze the devices behaviors and uncover intrinsic relationships.

3.3. Uncovering Devices Intrinsic Relationships

The domains were chosen so as to uncover similarities between devices used at the same time. Our next step is to quantify these similarities. Let us first define n windows in time that cover the available data; it defines n time bins noted u going from 1 to n. Then, let us compute the pairwise correlations of modes $X_j^B(t)$ on each window; we note $C_{i,j}^{B,u}$ these correlation matrices, with i and j ranging from 1 to d. Each line of $C_{i,j}^{B,u}$ quantifies the behavior of the device i in the chosen frequency domain B at a given time bin u, through its relationship with the other devices. These correlation matrices form the basis for the proposed method to track the behavior of devices and to search for misbehavior.

We define four reference correlation matrices to represent the normal behavior of the devices. They are computed as

$$R_{i,j}^B = \mathrm{median}_u(C_{i,j}^{B,u}) = \mathrm{median}(C_{i,j}^{B,1}, ..., C_{i,j}^{B,n}).$$

3.4. Anomaly Detection

For each device i, a distance $l_i^{B,u}$ is computed to measure the deviation – in terms of correlation coefficients with all the other devices – between the reference $R_{i,*}^B$ previously introduced and the correlations $C_{i,*}^{B,u}$ measured at time bin u:

$$l_i^{B,u} = \left(\sum_{j=1}^{d} w_{ij} |C_{i,j}^{B,u} - R_{i,j}^B|^p \right)^{1/p}$$

where the weights read $w_{ij} = R_{i,j}^B / \sum_{k=1}^d R_{i,k}^B$. We use these weights to give more importance to events characterized by a sudden drop in correlations between devices usually highly correlated than to events characterized by a sudden increase in correlations between devices usually uncorrelated (which are probably coincidences).

We monitor this distance over several time bins, and detect abnormal device behaviors as outlier values. For that, a robust outlier detector based on Median Absolute Deviation (MAD) [11] is implemented. MAD is defined as: $\mathrm{MAD}_i = b\,\mathrm{median}(|l_i - \mathrm{median}(l_i)|)$ where the constant b is usually set to 1.4826 for consistency with the usual parameter σ for Gaussian distributions. The device i at time u is reported as anomalous if it satisfies the following:

$$l_i^{B,u} > \mathrm{median}_t(l_i^B) + \tau\,\mathrm{MAD}(l_i)$$

where τ is a parameter that tunes the detector sensitivity. Empirically, $p = 4$ yields the best results, inhibiting small differences between $C_{i,j}^{B,u}$ and $R_{i,j}^B$ but emphasizing the important ones.

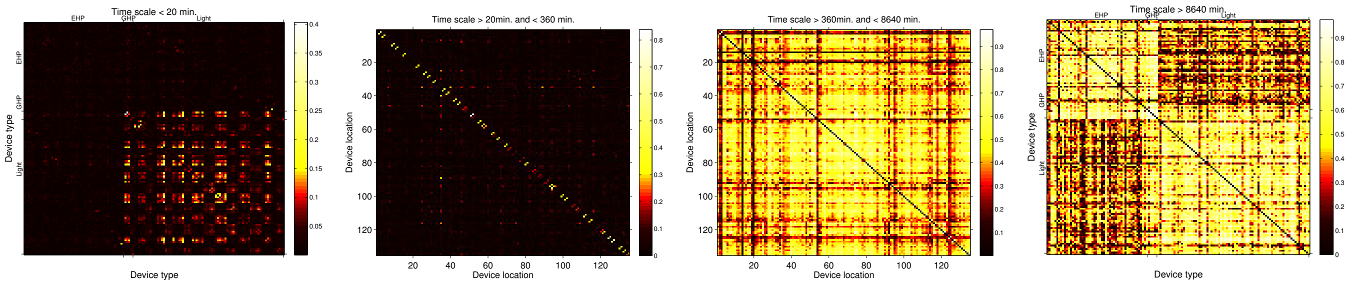
4. RESULTS

4.1. Devices Intrinsic Relationships

This section shows the benefits of comparing energy consumption data at the four proposed frequency bands. We here split the dataset in $n = 10$ bins of 1 week long and compute, for each bin, the 4 correlation matrices corresponding to each frequency band. We thereby obtain the four reference matrices (Fig. 3). To display the matrices, devices are ordered by types (EHP-GHP-Light) for R^H and R^T, and by location proximity for R^M and R^L.

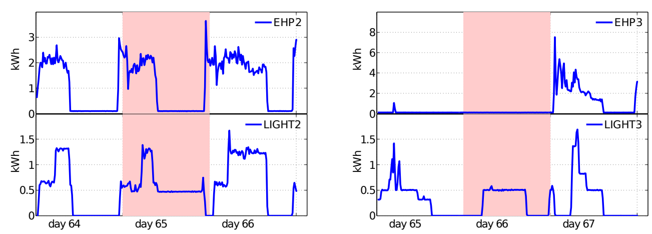
As could be expected, R^H reveals only noise in the data and does not provide any help to determine the inter-device usages (Fig. 3a).

For the medium frequencies (R^M) the high correlation coefficients are clearly concentrated along the diagonal (Fig. 3b). Since devices serving the same or adjacent rooms are nearby in this display of R^M, this means that nearby sensors (typically two sensors in the same room) are usually used in concert. The medium frequency band successfully captures the devices intrinsic relationships. Considering this reference matrix as an adjacency matrix of a graph, in which the nodes are the devices and the links' weights the correlation coefficients, we identify the clusters of strongly correlated devices using a community mining algorithm [12]. Most clusters have only two devices, namely the light and HVAC serving the same pair of room. Still, clusters that are composed of more devices are found. For example a cluster contains 3 HVAC systems serving the three



(a) High freq. (b) Med. freq. (c) Low freq. (d) Residual Trend

Fig. 3. Reference matrices R^B for the 135 power-draw signals. For R^H and R^T, devices are ordered by types for display; for R^M and R^L, devices are ordered by locations for display.



(a) High power usage where the light is left on at night

(b) Low power usage where the HVAC (EHP) is not used during office hours

Fig. 4. Example of alarms (red rectangles) reported by the proposed method. See text for discussion.

server rooms that are managed similarly. Interestingly, we also observe a couple of clusters that consist of independent devices serving adjacent rooms belonging to the same lab. The bigger cluster contains 33 devices that are 2 GHP devices and the corresponding lights. All that confirms that R^M identifies and highlights the hidden inter-device usage relationships.

For low frequencies, R^L shows no particular structure (see Fig. 3c).

The residual data contains the weekly trend, and no information about intrinsic device relationships (see Fig. 3d). Surprisingly though, by ordering R^T in the order of the type of the devices, two major clusters are apparent visually. The first cluster consists of HVAC devices and the second one contains only lights. An in-depth examination of the data reveals that the consumption of both the EHP and GHP devices is driven by the outside temperature. However, the use of light is independent from the outside temperature thus the lighting systems follow a common trend different from the EHP/GHP one.

4.2. Identified Anomalies

Let us now report results for the proposed anomaly detection, applied to the medium frequency domain which is the more relevant for inter-device usage patterns. The dataset in now split in $n = 70$ bins of 1 day long. Parameter τ is set to 5, in order to keep the false alarm rate low. All the alarms were manually inspected to check for ground truth.

The method reported 9 alarms corresponding to high power usage; this, despite the post-Fukushima measures to reduce the building energy consumption. As an example, Fig. 4a depicts the electricity consumption of the light and EHP in the same room. The detector reported these two devices as anomalous because their usually strong relationship is "broken" when the EHP is turned off at night and the light is left on. This anomaly stands for a waste of electricity. The top-priority reported anomaly is caused by the 10 days long constant use of an EHP and this waste of electricity accounts for 165 kWh.

It is also possible to detect anomalies corresponding to abnormally low power use; the method reports 6 of them. Upon further inspection, we notice that it corresponds to device failures or energy saving initiatives from the occupants – likely due to electricity concerns in Japan. One behavior of this type is displayed in Fig. 4b: the room is occupied at the usual office hours (indicated by light usage) but the EHP is not on in order to save electricity.

On the whole, the proposed methodology, going from raw signals to the anomaly detection, is able to find relevant anomalies in energy consumption, without resorting to prior knowledge of expected behavior of the devices.

5. CONCLUSIONS

Administrators of large buildings are overwhelmed with the data streams sent by the numerous sensors monitoring the buildings infrastructures. Our proposed method and our experiments reveal two key benefits of using E-EMD to monitor building data. First, E-EMD allows us to analyze numerous non-stationary signals without prior knowledge of the dataset. Second, E-EMD overcomes the difficulties faced while applying the classical EMD to the signals containing flat parts. Thereby, using E-EMD we uncover four frequency domains that expose distinct features of the monitored devices. Correlations of the medium frequency part of the devices signals are found to accurately measure the inter-devices relationships. It enables us to design an unsupervised anomaly detector that identifies abnormal device usages. Our evaluation with 10 weeks of real data shows that the proposed detector is a valuable help for administrators as the majority of the identified anomalies stand for energy waste.

Acknowledgments

We are grateful to Jorge Ortiz, Kensuke Fukuda, Patrice Abry and Kenjiro Cho for the fruitful discussions.

6. REFERENCES

[1] Gowtham Bellala, Manish Marwah, Martin Arlitt, Geoff Lyon, and Cullen E. Bash, "Towards an understanding of campus-scale power consumption," Seattle, WA, Nov. 1, 2011, Buildsys'11, p. 6.

[2] Michael Wrinch, Tarek H.M. EL-Fouly, and Steven Wong, "Anomaly detection of building systems using energy demand frequency domain anlaysis," in *IEEE Power & Energy Society General Meeting*, San-Diego, CA, USA, 2012.

[3] Chao Chen and Diane J. Cook, "Energy outlier detection in smart environments.," in *Artificial Intelligence and Smarter Living*. 2011, vol. WS-11-07 of *AAAI Workshops*, AAAI.

[4] Srinivas Katipamula and Michael R. Brambley, "Methods for fault detection, diagnostics, and prognostics for building systems, a review, part I," *HVAC&R Research*, vol. 11, no. 1, pp. 3–25, 2005.

[5] Srinivas Katipamula and Michael R. Brambley, "Methods for fault detection, diagnostics, and prognostics for building systems, a review, part II," *HVAC&R Research*, vol. 11, no. 2, pp. 169–187, 2005.

[6] John E. Seem, "Using intelligent data analysis to detect abnormal energy consumption in buildings," *Energy and Buildings*, vol. 39, no. 1, pp. 52 – 58, 2007.

[7] Zhaohua Wu and Norden E. Huang, "Ensemble empirical mode decomposition: A noise-assisted data analysis method," *Advances in Adaptive Data Analysis*, vol. 1, no. 01, pp. 1–41, 2009.

[8] Maria E. Torres, Marcelo A. Colominas, Gaston Schlotthauer, and Patrick Flandrin, "A complete ensemble empirical mode decomposition with adaptive noise," in *Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on*. IEEE, 2011, pp. 4144–4147.

[9] Romain Fontugne, Jorge Ortiz, David Culler, and Hiroshi Esaki, "Empirical mode decomposition for intrinsic-relationship extraction in large sensor deployments," in *IoT-App'12, Workshop on Internet of Things Applications*, Beijing, China, 2012.

[10] Norden E. Huang, Zheng Shen, Steven R. Long, Manli C. Wu, Hsing H. Shih, Quanan Zheng, Nai-Chyuan Yen, Chi Chao Tung, and Henry H. Liu, "The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis," *Proceedings of the Royal Society of London. Series A*, vol. 454, no. 1971, pp. 903–995, 1998.

[11] Peter J. Huber and Elvezio M. Ronchetti, *Robust Statistics*, Wiley Series in Probability and Statistics. Wiley, 2009.

[12] Vincent D. Blondel, Jean-Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre, "Fast unfolding of communities in large networks," *J.STAT.MECH.*, 2008.