Statistical Assessment of Abrupt Change Detectors for Non-Intrusive Load Monitoring

Presentation · February 2018

DOI: 10.13140/RG.2.2.23733.37602

Francois Auger

University of Nantes

Dominique Fourer

Institut de Recherche et Coordination Acoust…

Statistical Assessment of Abrupt Change Detectors for Non-Intrusive Load Monitoring

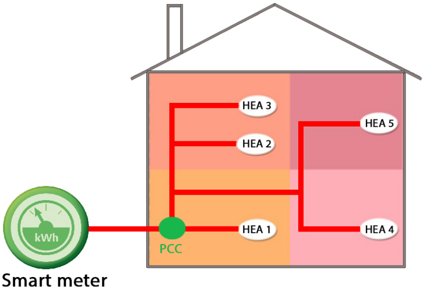
S. Houidi, F. Auger, H. Ben Attia Sethom, L. Miègeville, D. Fourer, X. Jiang

The 19th International Conference on Industrial Technology
ICIT 2018
February 20-22
Lyon- France

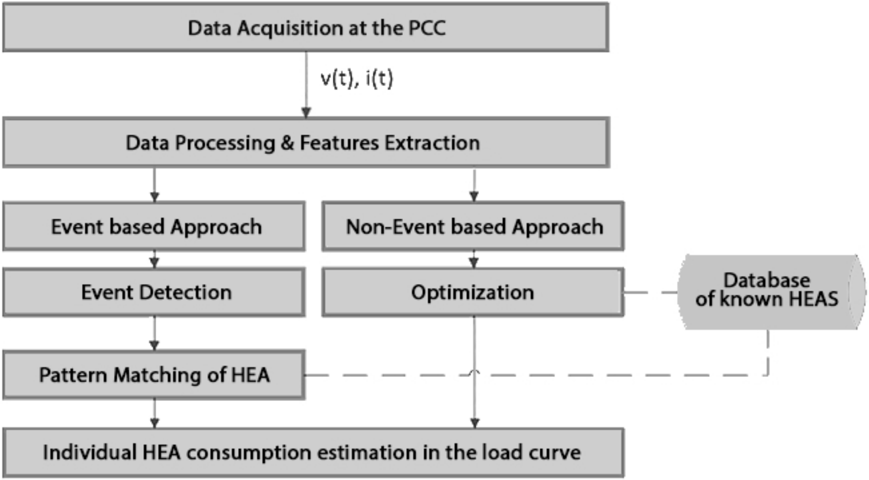
Outline

NILM goals

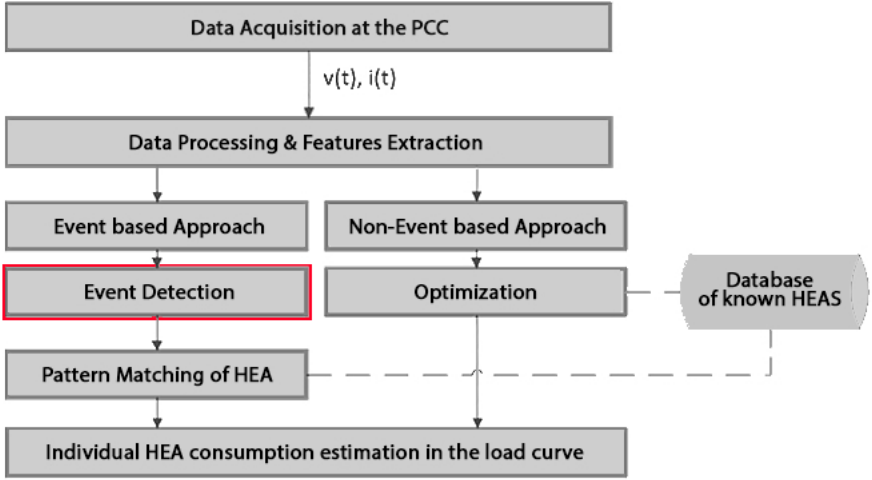
- **NILM:** Process to estimate the energy consumed by individual Home Electrical Appliances (HEAs) with a single meter in a house electrical panel connected at the PCC.
  - $\Rightarrow$ Partition of the load curve into its main components
  - $\Rightarrow$ Assignment of energy expenses per HEA



General framework of supervised NILM methods

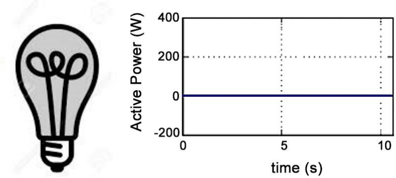


General framework of supervised NILM methods



Abrupt change definition

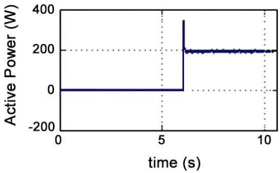
- Fast transition that occurs between stationary states in a signal
  $\Rightarrow$ **NILM:** On/Off and multiple operation modes appliances



Abrupt change definition

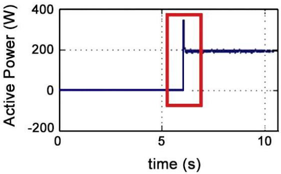
- Fast transition that occurs between stationary states in a signal

  $\Rightarrow$ **NILM:** On/Off and multiple operation modes appliances



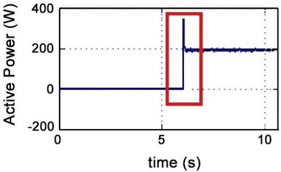
Abrupt change definition

- Fast transition that occurs between stationary states in a signal
  - ⇒ **NILM:** On/Off and multiple operation modes appliances



Abrupt change definition

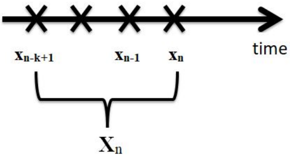
- Fast transition that occurs between stationary states in a signal
  ⇒ **NILM:** On/Off and multiple operation modes appliances



- Need of **tools** to decide whether a change occurs or not in the signal

Mathematical formulation 1/2

- $X_n = \{x_m \in \mathbb{R}, m = n - k + 1, \ldots, n\}$: vector of the last k available samples of a signal at the current time n.
- x_m follows a probability density function (PDF) $p_\theta(x_m)$ depending on a deterministic parameter θ
- **Abrupt change:** modification of θ at a change time n_c.



⇒ Hypothesis Test:

- H_0: **"no change"** versus H_1: **"with a change at time n_c"**

Mathematical formulation 2/2

$$\text{under } \mathrm{H}_0,\ \theta = \theta_0 \quad \text{for} \quad n-k+1 \le m \le n$$

$$\text{under } \mathrm{H}_1,\ \theta = \begin{cases} \theta_{1a} & \text{for} \quad n-k+1 \le m \le n_c - 1 \\ \theta_{1b} & \text{for} \quad n_c \le m \le n \end{cases}$$

$\Rightarrow$ **Decision rule:**
At each time n, comparison of a **decision function** g_n to a **threshold value** h adjusted according to decision probabilities

$$\begin{aligned} &\text{decide } \mathrm{H}_1 \quad \text{if} \quad g_n > h \\ &\text{decide } \mathrm{H}_0 \quad \text{if} \quad g_n \le h \end{aligned}$$

Detectors assessment conditions

⇒ Assessment of all the algorithms in **strictly the same conditions**

⇒ Sliding window of $\mathbf{k} = \mathbf{5}$ **samples** for the three detection algorithms to be studied:

- The Effective Residual algorithm
- The CUmulative SUM (CUSUM) algorithm
- The Bayesian Information Criterion (BIC) algorithm

Effective Residual

- Parity equation-based approach:
  $\Rightarrow$ temporal redundancies of measurements
  $\Rightarrow$ used for sensor fault detection and isolation

Effective Residual decision function

- Absolute variation δ_m between 2 consecutive signal samples:

$$\delta_m = |x_m - x_{m-1}| \quad \text{for} \quad n-k+2 \leq m \leq n$$

- Residual r_m: difference between 2 consecutive variations

$$r_m = |\delta_m - \delta_{m-1}| \quad \text{for} \quad n-k+3 \leq m \leq n$$

- Effective Residual decision function: sum of the last 3 residuals

$$g_n \underset{\mathrm{H}_0}{\overset{\mathrm{H}_1}{\gtrless}} h \quad \text{with} \quad g_n = r_n + r_{n-1} + r_{n-2}$$

$\Rightarrow$ Detection of a mean change at $n_c = n$ from the **last k = 5 samples** of the signal

CUSUM algorithm

- Used for biomedical engineering as well as for NILM applications
- Based on **log-likelihood ratio maximization** over change time n_c
- Most common form: statistical test for the detection of a **mean change** in a Gaussian process $\mathcal{N}(\mu, \sigma)$

CUSUM algorithm principle

- PDFs of X_n under hypotheses H_0 and H_1:

$$p(X_n|\mathrm{H}_0) = \prod_{m=n-k+1}^{n} p_{\theta_0}(x_m) \qquad p(X_n|\mathrm{H}_1) = \prod_{m=n-k+1}^{n_c-1} p_{\theta_{1a}}(x_m) \prod_{m=n_c}^{n} p_{\theta_{1b}}(x_m)$$

- If $\theta_0=\theta_{1a}$, log-likelihood ratio $L(X_n, n_c)$:

$$L(X_n, n_c)=\ln\left(\frac{p(X_n|\mathrm{H}_1)}{p(X_n|\mathrm{H}_0)}\right)=\sum_{m=n_c}^{n} s_m \quad \text{with} \quad s_m=\ln\left(\frac{p_{\theta_{1b}}(x_m)}{p_{\theta_{1a}}(x_m)}\right)$$

- CUSUM decision rule: maximization of the log-likelihood ratio over n_c

$$g_n \underset{\mathrm{H}_0}{\overset{\mathrm{H}_1}{\gtrless}} h, \quad \text{with} \quad g_n = \max_{n-k+1\leq n_c\leq n} \sum_{m=n_c}^{n} s_m$$

CUSUM decision function

- Changing parameter: mean value μ in a **Gaussian process** $\mathcal{N}(\mu, \sigma)$

$$\begin{aligned}\text{under } \mathrm{H}_0,\ \mu &= \mu_0 \quad \text{for} \quad n-k+1 \leq m \leq n\\ \text{under } \mathrm{H}_1,\ \mu &= \begin{cases}\mu_{1a} & \text{for} \quad n-k+1 \leq m \leq n_c-1\\ \mu_{1b} & \text{for} \quad n_c \leq m \leq n\end{cases}\end{aligned}$$

- Instantaneous log-likelihood ratio s_m:

$$s_m = \frac{(x_m-\mu_{1b})^2}{2\sigma^2} + \frac{(x_m-\mu_{1a})^2}{2\sigma^2} = \frac{\Delta\mu}{\sigma^2}\left(x_m - \frac{\mu_{1b}+\mu_{1a}}{2}\right) \quad \text{with} \quad \Delta\mu = \mu_{1b}-\mu_{1a}$$

$\Rightarrow$ **The CUSUM decision rule g_n:**

For an abrupt change occurring at $n_c = n$ in a **sliding window of $k = 5$ samples**

with $\hat{\mu}_{1b} = x_n$, $\quad \hat{\mu}_{1a} = \frac{1}{4}\sum_{m=n-4}^{n-1} x_m$ and $\quad \hat{\sigma}^2 = \frac{1}{4}\sum_{m=n-4}^{n-1}(x_m - \hat{\mu}_{1a})^2$,

$$g_n \underset{\mathrm{H}_0}{\overset{\mathrm{H}_1}{\gtrless}} h, \quad \text{with} \quad g_n = s_n = \frac{(x_n - \hat{\mu}_{1a})^2}{2\,\hat{\sigma}^2}$$

BIC algorithm

- Used for acoustic change detection
- Division of the sequence of observed random samples into **homogeneous segments** by performing a **hypothesis test** at each potential change point
  - ⇒ Hypothesis H_0: on both sides of this point, the signal follows the same probabilistic model
  - ⇒ Hypothesis H_1: a model change occurs

BIC algorithm principle

- The BIC of X_n under hypothesis H_i, $i \in \{0,1\}$: likelihood criterion penalized by the model complexity

$$\mathrm{BIC}(H_i) = \ln(p(X_n|H_i)) - \frac{\lambda}{2} M \ln(k)$$

$p(X_n\|H_i)$	Maximized data likelihood for the given model
λ	Penalty factor (ideally equal to 1)
k	Last available samples
M	Number of parameters in the probabilistic model

- Probabilistic model:

$$\begin{aligned}
\mathrm{H}_0 &: x_{n-k+1}, \ldots, x_n \sim \mathcal{N}(\mu_0, \sigma_0) \\
\mathrm{H}_1 &: x_{n-k+1}, \ldots, x_{n_c-1} \sim \mathcal{N}(\mu_{1a}, \sigma_{1a}); \\
& \quad x_{n_c}, \ldots, x_n \sim \mathcal{N}(\mu_{1b}, \sigma_{1b})
\end{aligned}$$

- Model parameters:
  - ▶ Under H_0: μ_0 and σ_0 ($M = 2$)
  - ▶ Under H_1: μ_{1a}, σ_{1a}, and μ_{1b}, σ_{1b} ($M = 4$)
  - $\Rightarrow$ Maximization of $\mathrm{BIC}(H_i)$ when μ and σ^2 are replaced by their MLEs $\hat{\mu}$ and $\hat{\sigma}^2$

BIC decision function

- The BIC decision function g_n is:

$$g_n \underset{\mathrm{H}_0}{\overset{\mathrm{H}_1}{\gtrless}} h \quad \text{with} \quad g_n = \max_{n-k+1 \leq n_c \leq n} \Delta\mathrm{BIC}(n_c)$$

$$\begin{aligned} \text{where} \quad \Delta\mathrm{BIC}(n_c) &= \mathrm{BIC}(H_1) - \mathrm{BIC}(H_0) \\ &= \frac{k}{2}\ln(\hat{\sigma}_0^2) - \frac{(n_c-n+k-1)}{2}\ln(\hat{\sigma}_{1a}^2) - \frac{(n-n_c+1)}{2}\ln(\hat{\sigma}_{1b}^2) - \lambda\ln(k) \end{aligned}$$

$\Rightarrow$ BIC decision rule:

For an abrupt change occurring at $n_c = n-1$ in a **sliding window** of $k = 5$ samples:

$$g'_n \underset{\mathrm{H}_0}{\overset{\mathrm{H}_1}{\gtrless}} h' \quad \text{with} \quad g'_n = \frac{1}{2}\ln\left(\frac{\hat{\sigma}_0^{10}}{\hat{\sigma}_{1a}^6\hat{\sigma}_{1b}^4}\right), \; h' = h + \lambda\ln(5)$$

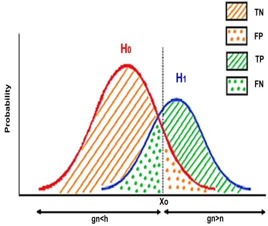
with $\hat{\sigma}_0^2 = \frac{1}{5}\sum_{m=n-4}^{n}(x_m - \hat{\mu}_0)^2$, $\hat{\sigma}_{1a}^2 = \frac{1}{3}\sum_{m=n-4}^{n-2}(x_m - \hat{\mu}_{1a})^2$ and $\hat{\sigma}_{1b}^2 = \frac{1}{2}\sum_{m=n-1}^{n}(x_m - \hat{\mu}_{1b})^2$

Test bench

- Monte Carlo Test repeated 100 000 times
  - $\Rightarrow$ X_n is filled with 5 *i.i.d* samples $x_m \sim \mathcal{N}(0, \sigma)$ with $\sigma = 1$
  - $\Rightarrow$ Under H_1, addition of $\Delta\mu = \mathrm{SNR} \times \sigma$:
    - to the last sample for Effective Residual and CUSUM
    - to the last 2 samples for the BIC
  - $\Rightarrow$ Assessment made for:
    - fixed SNR values
    - varying SNR values ranging from 0.5 to 10
  - $\Rightarrow$ Use of 400 logarithmically spaced values of h

Performance metrics

- **Basic performance metrics**:



True Positive TP $\rightarrow$ detection of a change when there is really one

True Negative TN $\rightarrow$ no detection of a change when there is not

False Positive FP $\rightarrow$ detection of a change when there is not

False Negative FN $\rightarrow$ no detection of a change when there is really one

- Computation of **performance rates**:

True Positive Rate TPR

$$\mathrm{TPR} = \mathrm{TP}/(\mathrm{TP} + \mathrm{FN})$$

False Positive Rate FPR

$$\mathrm{FPR} = \mathrm{FP}/(\mathrm{TN} + \mathrm{FP})$$

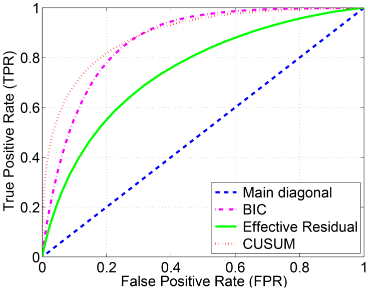
Precision $\mathrm{P_R}$

$$\mathrm{P_R} = \mathrm{TP}/(\mathrm{TP} + \mathrm{FP})$$

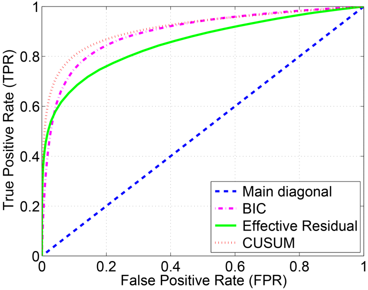
Performance metrics

- **Receiver Operating Characteristics (ROC):**
  Plot of the TPR versus the FPR for varying values of h

$$\lim_{h\to-\infty} TPR = 1, \lim_{h\to-\infty} FPR = 1, \lim_{h\to+\infty} TPR = 0, \lim_{h\to+\infty} FPR = 0$$



$\mathrm{SNR} = 3$



Uniformly distributed SNR between 0.5 and 10

⇒ CUSUM ROC curve is the closest to the “optimal” point (FPR=0;TPR=1)

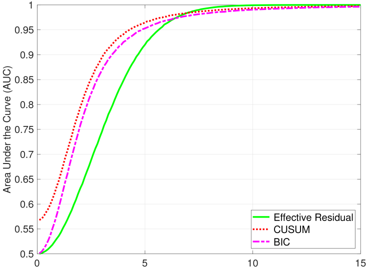
Performance metrics

- **Area Under the Curve (AUC):**
  Numerical integration of the ROC curve

AUC values computed for the three considered detection algorithms

	Effective Residual	CUSUM	BIC
$\mathrm{SNR} = 0.5$	0.51	0.59	0.53
$\mathrm{SNR} = 3$	0.75	0.89	0.87
$\mathrm{SNR} = 6$	0.96	0.98	0.97
variable SNR	0.85	0.91	0.89

$\Rightarrow$ CUSUM: best detection ability



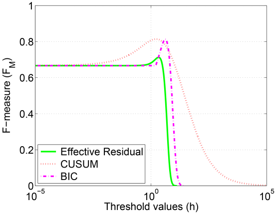
SNR-AUC profile of the Effective Residual, CUSUM and BIC algorithms.

Performance metrics

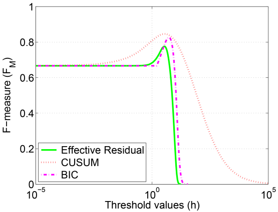
- **F-Measure (F_M)** harmonic mean of $\mathrm{P_R}$ and TPR

$$\mathrm{F_M} = \left(\frac{P_R^{-1} + \mathrm{TPR}^{-1}}{2} \right)^{-1} = 2 \times \frac{P_R \times \mathrm{TPR}}{P_R + \mathrm{TPR}}$$

- F-measure ($\mathrm{F_M}$) depending on threshold values h of each detector:



$\mathrm{SNR} = 3$

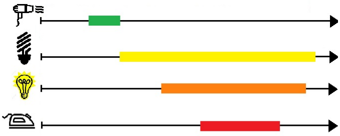


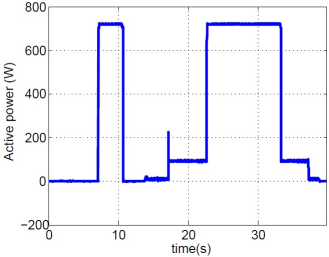
Uniformly distributed SNR between 0.5 and 10

- For small values of h, $P_R = 1/2$ and $\mathrm{TPR} = 1$, so $\mathrm{F_M} = 2/3$.
- For large values of h, $\mathrm{TPR} = 0$, so $\mathrm{F_M} = 0$.
- For the three detectors, the maximum $\mathrm{F_M}$ score is reached for a specific threshold value which is the optimal one.

Practical case study

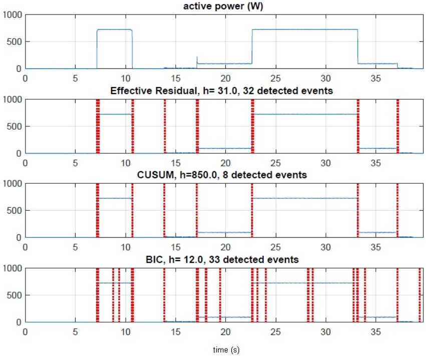
- Realization of a controlled consumption scenario using our own measurement system
- Measured current $i[k]$ and voltage $v[k]$ sampled at $F_s = 1.2$ kHz





Application of the detetctors

- Active power signal $P[n] = \frac{1}{M} \sum_{k=n-M+1}^{n} v[k]i[k]$, with $M = Fs/F$
- For each detector, h is set to allow the detection of the lowest $\Delta P = (P[n] - P[n-1])$ (low-energy lamp switch-on)
- Power time profile and detection results of each algorithm:



Detection results 1/2

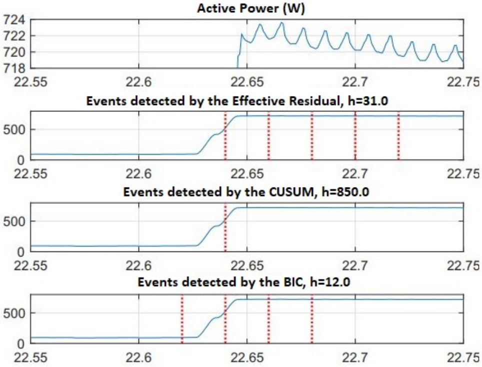
Threshold values h and number of TP, FP and P_R of the three considered detection algorithms applied to the active power signal.

	Effective Residual	**CUSUM**	**BIC**
h	31.0	850.0	12.0
Number of detected events	32	8	33
TP	7	7	7
FP	25	1	26
$\mathbf{P}_R$	21.3%	87.5%	21.2%

$\Rightarrow$ The CUSUM algorithm appears to be the most effective with a high P_R and a small number of false positive.

Detection results 2/2

- Relatively large number of detected events for the Effective Residual and the BIC algorithms
  - $\Rightarrow$ Noisy steady-states leading to false positive F_P and inaccurate detections.
- Zoom overview of the active power signal drawn by the iron switch on:



Conclusion

- **Event-based NILM approach**
- Study of 3 detectors: **Effective Residual**, **CUSUM** and **BIC**
- Test bench for the **statistical assessment** of the detectors under the same conditions
- Definition of **metrics** to judge the detectors performances
- **Practical case study**: application of the detectors to a controlled HEAs consumption scenario.

Prospectives

- Threshold setup from the Probability Density Function of the decision functions
- Extension to multidimensional signals to make a decision from several features

Thank you for your attention.