Multivariate Event Detection for Non-Intrusive Load Monitoring

Alexander Gerka and Benjamin Cauchi
OFFIS e.V.
Institute for Information Technology
Oldenburg, Germany
Email: gerka@offis.de
benjamin.cauchi@offis.de

Andreas Hein
Carl von Ossietzky University,
Oldenburg, Germany
Email:andreas.hein@uol.de

***Abstract*—Due to the growing interest for in-home activity monitoring, the tracking of appliances use, usually referred to as Non-Intrusive Load Monitoring (NILM), has to address new challenges. Indeed, as NILM has long been motivated by potential energy savings, most event detectors for NILM have focused on the detection of on- or off-switches of high power devices. On the contrary, in-home monitoring typically relies on the detection of events related to low-power devices from potentially noisy signals. Additionally, approaches that apply expert heuristics to a single-variate input, often favored for their low complexity and real-time applicability, can be overly sensitive to the choice of an arbitrary defined detection threshold. This paper aims at decreasing the sensitivity of a detector based on expert heuristic by applying it to the Hotelling-T^2 statistic of a multivariate input, computed online from the current and voltage inputs. Focusing on realistic scenarios, the approach is evaluated on a dataset recorded in a real apartment using a commercially available smart-meter. The results, expressed in terms of precision, recall and F-score, show that the proposed approach can both yield higher performance and be less sensitive to the choice of the detection threshold.**

Keywords*—*NILM; Event Detection; Hotelling-T^2 statistic

I. INTRODUCTION

Non-intrusive load monitoring (NILM), introduced by George Hart in the 1980s [1], denotes the tracking of appliances use through analysis of their power consumption. The growing presence of smart meters in homes, encouraged, e.g., by a directive of the European Union in 2009 [2], has led to renewed research efforts in NILM. These efforts have been mostly motivated by the potential for energy savings and, therefore, focussed on the monitoring of high power devices. However, due to the ageing population [3], applications such as activity monitoring for disease prevention [4][5] or the surveillance of life-critical devices present in homes, e.g., for respiration support [6], have become of crucial importance. The main advantage of NILM-based monitoring system is their unobtrusiveness, as they do not require an additional installation. Additionally, a NILM-system could prevent other, more obtrusive, installations such as power plugs. Consequently, NILM approaches able to reliably monitor the use of low power devices in realistic settings are urgently needed.

Many approaches extract multiple features from the recorded power consumption signal in order to determine the active appliances during a given time sequence. It has been proposed in [1] to use both reactive and real power. However, due to the high correlation between real and reactive power consumption, these features are not sufficient for the classification of low power devices [7]. As a result, numerous features for NILM have been introduced in recent years such as the current's harmonic [8], the shape of the so-called VI-trajectory [9] or the poles and residues of the power signal's impulse response [10]. The extraction of these complex features require to record the current and voltage of the power supply at a sampling frequency much higher than 50 Hz. Additionally, the applied classifier typically relies on an event based method rather than a state-based method and necessitate to segment the recorded signal into sequences of interest. This segmentation requires an event detector, i.e., the detection of on- and off-switches of appliances. This paper focuses on event detection for NILM in realistic scenarios.

Approaches aiming at event detection for NILM can be broadly split between three categories, namely, based on expert heuristics, matched filtering, or probabilistic methods [11]. Though promising, probabilistic methods based on, e.g., generalized likelihood ratio (GLR) [12], goodness-of-fit (GOF) [13] or cumulative sum control (CuSuM) [14] can be computationaly costly and often rely on long sequences of input signal, making them impractical in many realistic settings. Approaches based on matched filtering, e.g., using Cepstrum smoothing [15] or Hilbert transformation [16] can be sensitive to mismatch between the dataset used to tune the approach and the environment in which it is tested. Though promising, the approach presented in [15] showed largely lower performance when tested on data containing low-power devices [17].

Approaches based on expert heuristics rely on the choice of a somewhat arbitrarily defined threshold or rule-based approach [1][18], on which their performance can be greatly sensitive. However, relying on features whose extraction is typically computationally inexpensive, e.g., standard deviation of the current signal envelope [19]. Approaches based on expert heuristics are easily implemented. Therefore, reducing their sensitivity on an arbitrary defined threshold could be of great interest. Contrary to most expert heuristic approaches that use a single-parameter as input, the approach proposed in

 ISBN: 978-1-61208-739-9

this paper aims at decreasing the sensitivity of a detector based on a expert heuristic by applying the Hotelling-T^2 statistic to a multivariate input.

In order to evaluate the benefit of the proposed approach, a dataset recorded in realistic conditions has to be used. Indeed, most datasets from the literature only considered low frequency systems [1][12][20], specific devices [18] or sets of (mostly) high-power devices [13][14][16][21]. Additionally, those studies were often performed under laboratory conditions [1][14], i.e., in absence of capacitive effects of the long supply line and other environmental noise factors present in a real apartment. Such signal disturbances can have a large impact. Consequently, the evaluation conducted in this paper is done using a dataset recorded in a real apartment using a commercially available smart-meter.

The remainder of this paper is structured as follows. First, the proposed approach and the expert heuristic approach that is used as benchmark is described in Section II. The recorded dataset and experimental framework are described in Section III and the results in Section IV. Section V concludes the paper.

II. Proposed Approach

In this section the proposed approach is described.

A. Threshold based NILM event detection

The signal recorded from a monitoring device, e.g., a smart meter, typically consists of $M = 2$ channels representing the current in Ampere and the voltage in Volts. In the remainder of this paper, we use $x_m(n)$ to denote the signal recorded at a sampling frequency fs, at sample index n, in the m-th channel. We arbitrary set $m = 0$ as the index of the current channel and $m = 1$ as the index of the voltage channel. Event detection methods for NILM based on experts heuristics, such as the one proposed in [19], typically rely on segmenting the input signal into overlapping frames of length L with an hop size of H samples and assigning a label $d(\ell) \in [0, 1]$ equal to 1 if an event is detected in the ℓ-th frame and equal to 0 otherwise.

The methods considered in this paper rely on computing a change quantifying value $v(\ell) \in \mathbb{R}_{\geq 0}$, whose computation is the focus of the next subsection, for each frame, and applying

$$d(\ell) = \begin{cases} 1 \text{ if } v(\ell) \geq \tau(\ell), \\ 0 \text{ otherwise}, \end{cases} \tag{1}$$

where $\tau(\ell)$ denotes a decision threshold. It can be noted that contrary to methods in which the decision threshold is computed from a complete signal utterance, e.g., [13][21], this paper considers event detection for real-time application and uses a frame dependant threshold computed as

$$\tau(\ell) = \alpha \cdot \sigma_v(\ell - \delta), \tag{2}$$

where δ denotes a decision delay and where

$$\sigma_v(\ell) = \sqrt{\frac{1}{\Delta - 1} \sum_{i=0}^{\Delta-1} \left| v(\ell - i) - \frac{1}{\Delta} \sum_{j=0}^{\Delta-1} v(\ell - j) \right|^2} \tag{3}$$

denotes the standard deviation of $v(\ell)$ computed over Δ buffered values. In practical applications, the values of δ and Δ are often hardware dependant. However, the choice of value assigned to the constant α in (2), though of critical importance for the detection performance, is typically arbitrarily defined. The sensitivity of the detection performance to the choice of α can be limited if the value $v(\ell)$ suitably quantifies the potential changes in the signal at a given time frame, i.e. $v(\ell) \approx 0$ when no change is present.

B. Change quantification

As most similar detectors, the approach proposed in [19] that we chose as benchmark (cf. Section IV) due to its low computational complexity and promising performance only uses the recorded current to quantifies the change in the input signal, i.e., computes $v(\ell)$ from $x_0(n)$ only. This computation relies on extracting for each frame, the envelope $e_\ell(n)$ of length $L_e = H \cdot (E - 1) + L$, where E denotes the number of frames used in the envelope extraction. The envelope $e_\ell(n)$ is computed as the interpolation, e.g., linear or cubic, between the maximums of $|x_0(n)|$ in each of the E considered frames and, assuming that $\mathrm{E}\{e_\ell(n)\}$ is constant in absence of event, the change quantifying value $v(\ell)$ can be computed as the Change of Mean Amplitude Envelope (CMAE)

$$v(\ell) = \frac{1}{L_e} \left| \sum_{i=0}^{L_e-1} e_{\ell-1}(i) - e_\ell(i) \right|. \tag{4}$$

Unfortunately computing $v(\ell)$ as in (4) can, as stated in [19], result in an unreliable detector in presence of noisy signals. We aim at improving the reliability of the threshold based detector from (2) by improving the computation of $v(\ell)$.

Preliminary works showed that using multiple features can be extracted from a NILM-signal. In general those features can be separated into different domains:

1) power related features in time domain
2) power related features in frequency domain
3) features related to the phase difference between current an voltage

Therefore, we propose to use a feature from each feature domain as the information about the NILM-signal contained by a feature is different for every domain. More specifically, we have shown in [22] that the variance of the current, taking advantage of both input channels, the phase of the input signal and, the current's frequency ratio were beneficial. In our proposed approach, we extract for each frame a $L_{\mathbf{v}} = 3$ elements feature vector

$$\mathbf{v}(\ell) = \left[\; \sigma_x^2(\ell) \; , \; \phi(\ell) \; , \; \omega(\ell) \; \right]^{\mathrm{T}}, \tag{5}$$

where $\cdot^{\mathrm{T}}$ denotes the transpose operator and where $\sigma_x^2(\ell)$, $\phi(\ell)$ and $\omega(\ell)$ denote the current variance, the phase and the current

 ISBN: 978-1-61208-739-9

frequency ratio computed as

$$\sigma_x^2(\ell) = \frac{1}{L-1}\sum_{i=0}^{L-1}\left|x_0(\ell H+i) - \frac{1}{L}\sum_{j=0}^{L-1}x_0(\ell H+j)\right|^2 \quad (6)$$

$$\phi(\ell) = \cos^{-1}\frac{\sum_{i=0}^{L-1}x_0(\ell H+i)\cdot x_1(\ell H+i)}{\sqrt{\sum_{i=0}^{L-1}x_0(\ell H+i)^2}\cdot\sqrt{\sum_{i=0}^{L-1}x_1(\ell H+i)^2}} \quad (7)$$

$$\omega(\ell) = \frac{\sum_{i=0}^{L-1}\left|\tilde{x}_f(\ell H+i) - \frac{1}{L}\sum_{j=0}^{L-1}\tilde{x}_f(\ell H+j)\right|^2}{\sum_{i=0}^{L-1}\left|\overline{x}_f(\ell H+i) - \frac{1}{L}\sum_{j=0}^{L-1}\overline{x}_f(\ell H+j)\right|^2}, \quad (8)$$

where $\tilde{x}_f(n)$ denotes the output of a bandpass filter applied to $x_0(n)$ and centred around the carrier frequency f of the input signal (cf. Section III-B) and $\overline{x}_f(n) = x_0(n) - \tilde{x}_f(n)$.

We consider each vector $\mathbf{v}(\ell)$ to be a single independent realisation of a random process with an $L_\mathbf{v}$-dimensional F-distribution and define the sequences of vectors $\mathbf{V}_0(\ell)$ and $\mathbf{V}_1(\ell)$ of length $L_\mathbf{0}$ and $L_\mathbf{1}$, respectively, as

$$\mathbf{V}_0(\ell) = \{\mathbf{v}(\ell - L_\mathbf{0})\ ,\cdots, \mathbf{v}(\ell-2)\ ,\ \mathbf{v}(\ell-1)\}, \quad (9)$$

$$\mathbf{V}_1(\ell) = \{\mathbf{v}(\ell)\ , \mathbf{v}(\ell+1)\ ,\cdots,\ \mathbf{v}(\ell+L_\mathbf{1}-1)\}. \quad (10)$$

An event is considered to occur at frame ℓ if the $\mathbf{V}_0()$ and $\mathbf{V}_1()$ are composed of realisations of significantly distinct distributions. This significance of can be expressed by the Hotelling-T^2 statistic [23]. The computation of the Hotelling-T^2 statistic depends on the homogeneity of the partial covariance matrices ($\mathbf{\Gamma}_0(\ell)$ and $\mathbf{\Gamma}_1(\ell)$) computed separately from the vectors $\mathbf{V}_0(\ell)$ and $\mathbf{V}_1(\ell)$. If, using the box test [24], these matrices are determined to be homogeneous, the Hotelling-T^2 statistic is computed as

$$\begin{aligned}T^2(\ell) =&\frac{L_\mathbf{0}\cdot L_\mathbf{1}}{L_\mathbf{0}+L_\mathbf{1}}\cdot\big(\boldsymbol{\mu}_0(\ell)-\boldsymbol{\mu}_1(\ell)\big)^\mathrm{T}\cdot\\ &\mathbf{\Gamma}(\ell)^{-1}\cdot\big(\boldsymbol{\mu}_0(\ell)-\boldsymbol{\mu}_1(\ell)\big),\end{aligned} \quad (11)$$

otherwise,

$$\begin{aligned}T^2(\ell) =&\big(\boldsymbol{\mu}_0(\ell)-\boldsymbol{\mu}_1(\ell)\big)^\mathrm{T}\cdot\left(\frac{\mathbf{\Gamma}_0(\ell)^{-1}}{L_\mathbf{0}} - \frac{\mathbf{\Gamma}_1(\ell)^{-1}}{L_\mathbf{1}}\right)\cdot\\ &\big(\boldsymbol{\mu}_0(\ell)-\boldsymbol{\mu}_1(\ell)\big),\end{aligned} \quad (12)$$

where $\boldsymbol{\mu}_0(\ell)$ and $\boldsymbol{\mu}_1(\ell)$ denote the average of the vectors in $\mathbf{V}_0(\ell)$ and $\mathbf{V}_1(\ell)$, respectively, and $\mathbf{\Gamma}(\ell)$ denotes the $L_\mathbf{v}\times L_\mathbf{v}$ covariance matrix computed using the vectors in both sequences. Finally, the label $d(\ell)$ can be assigned by substituting $v(\ell)$ by $T^2(\ell)$ in (1) and (3).

III. EXPERIMENT

In this section the experiment is presented.

A. Collected Dataset

We evaluated the approach proposed in Section II on a dataset constructed to evaluate the performance of NILM approaches applied to activities monitoring for health applications. The data was recorded in a three-room apartment occupied by two elderly people who agreed to take part on the study. The signals were recorded using a commercially available smart-meter placed on the main power permitting to record a continuous 2-channel stream, i.e., current and voltage, sampled at the sampling frequency fs =4800 Hz. To pass the data to a measurement computer an optical interface on the backside of the smart meter was used. The recorded current and voltage stream was stored with a resolution of 16 bits. Due to the recording conditions, the recorded signals contain the disturbances to be expected in such a realistic setting, e.g., noise generated by the supply line itself. Therefore, the dataset is particularly valuable to assess the performance to be realistically expected from NILM approaches.

A total of 36 individual appliances were present in the apartment and each one was switched on / off at least once during the recording session. Without storing the phase to which the appliances were connected, a total of 142 events to be detected. The distribution of these appliances in terms of both types and location is summarised in Figure 1. It can be seen that, as expected in a real apartment, most appliances were low power and a large number of lamps were present. It can be noted that, e.g., the monitoring of lamp usage can be a good indicator of activity and location of the user in an apartment. Additionally, aiming at health applications, a respiratory support machine (Resp. support) was among the considered appliances and is an example of appliance whose reliable monitoring could potentially be life critical.

All signals were recorded in a session of over an hour during which the timestamp of each on- off-switch event has been manually annotated. Events were setup to occur at large interval from one-another in order to avoid simultaneous events that could hinder the clustering used prior to the computation of the performance of the considered approaches, as described in the next subsection.

B. Evaluation

In order to evaluate the performance of the proposed approach, i.e., the use of Hotelling-T^2 statistic as input of the threshold based detector, and of the considered benchmark, i.e., using CMAE, in realistic framework and to avoid the detrimental effect of repeated approach initialisation, the entire dataset was processed as a single stream. The parameters were extracted using a 40 ms window, $L = 192$, and a 50 % hopsize, $H = 96$. The envelope used in the case of our benchmark CMAE was extracted using $E = 4$ blocks, similarly as in [19]. However, contrary to [19], a linear interpolation was used instead of a cubic one. This choice reduced the number of false positive obtained on the considered dataset. The frequency ratio $\omega(\ell)$ was computed by using a bandpass filter, designed as a second order butterworth filter, centred around the carrier frequency f =50 Hz with lower and higher cutoff set at 35 Hz and 65 Hz, respectively. We fixed $\Delta = 240$ (50 ms) and $\delta = 144$ (30 ms), cf. (2)-(3), and focused on the influence of the arbitrary chosen threshold, considering the

 ISBN: 978-1-61208-739-9

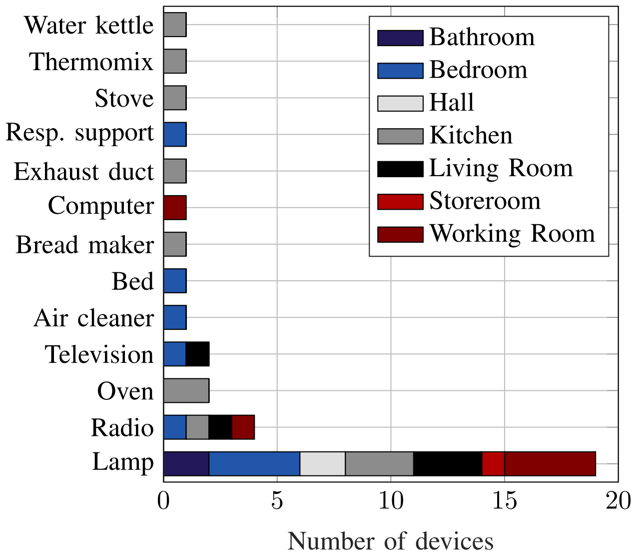

Figure 1. Distribution of the appliances used to record the evaluation dataset.

values $\alpha \in \{1 \cdots 30\}$. It can be noted that the exact values of Δ and δ seemed to have a limited impact on the performance of the considered approaches. The box test and computation of $T^2(\ell)$, cf. (11) and (12), were based on the implementation provided in [25] and according to its recommendation we used $L_0 = L_1 = 100$.

The performance of the considered approaches was determined using the Precision, Recall and F-score defined as [26]

$$\text{Precision} = \frac{\text{True positives}}{\text{True positives} + \text{False positives}}, \tag{13}$$

$$\text{Recall} = \frac{\text{True positives}}{\text{True positives} + \text{False negatives}}, \tag{14}$$

$$\text{F-score} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}, \tag{15}$$

where the number of true positives and false negatives were determined as follows. First, events detected within the same 8 s window were considered as a single-event. The same tolerance was used to determine if a detected event should be assigned to an annotated event, i.e., true positive, or to none, i.e., false positive. Annotated events with no assigned detection were considered false negatives.

IV. Results

An example of the extracted features from the smart meter input signal and the resulting Hotelling value that represents the combination of these features is shown in Figure 2. This figure shows that relatively small changes in the variance (i.e., at the 15th second) are accompanied by relatively big changes in the phase. As a result, the Hotelling value shows a clear peak around the 15th second which should be easier to detect by the event detection algorithm than the pure variance. Therefore, this observation suits the intuition that the usage of multiple non-related features may improve the detection of events in a NILM-signal.

The observed precision and recall as function of α are depicted in Figure 3. It appears that in the case of both CMAE and Hotelling-T^2, the precision increases with the value α, until it reaches a plateau, which in both cases corresponds to a precision of about 0.6. This behaviour is to be expected as increasing the value of α reduces the likelihood of false positives.

On the other hand, high values of α would increase the likelihood of false negative. This behavior can be noticed by observing the recall (Figure 3), for which the proposed approach exhibits a large advantage compared to the use of CMAE. Using CMAE, the recall decrease sharply for values of α ranging from 3 to 7 and ultimately reaches a plateau with a recall of 0.2. On the contrary, using Hotelling-T^2 statistic, recall decreases slowly with increasing values of α with a plateau corresponding to a recall of 0.8. This shows that the proposed approach is less likely to introduce false negative, even for overly large values of α.

The advantage of the proposed method over the use of CMAE is best noticed by observing the F-score depicted in Figure 4. As a logical consequence of the behaviour observed in terms of precision and recall, using CMAE requires an accurate setting of α in order to yield the optimal F-score. Indeed, F-score obtained using CMAE decreases sharply for α values different than 3 or 4. On the other hand, not only does the use of Hotelling-T^2 statistic yields a higher maximum F-score with a value of 0.7, but its performance is much less sensitive to the setting of α value.

Further improvement of the presented approach could be achieved by using the multivariate approach with other event detectors or by improving the multivariate statistic itself, i.e. using multivariate likelihood detectors [27].

V. Conclusion

This paper proposes to use Hotelling-T^2 statistic, computed from a multivariate input, in order to reduce the sensitivity of an event detector for NILM based on expert heuristics to the value of an arbitrary defined detection threshold. The multivariate input is computed online from a recorded signal of current and voltage. The proposed approach is compared to the use of CMAE, which, as many similar approaches, does not take advantage of the available voltage input or frequency-related features.

A dataset recorded in real environment, i.e., containing appliances relevant to activities monitoring and noisy signal, was used to evaluate the proposed approach. The results, expressed in terms of precision, recall and F-score, show that not only does the proposed approach yield better performance than the considered benchmark, this performance is much less sensitive to the setting of the detection threshold.

Even if using multiple features for event detection is computationally more expensive, we think this approach improves the event detection in NILM systems in future implementations, epecially for the detection of low-power devices in high-frequency load signals. In future work, we will evaluate the

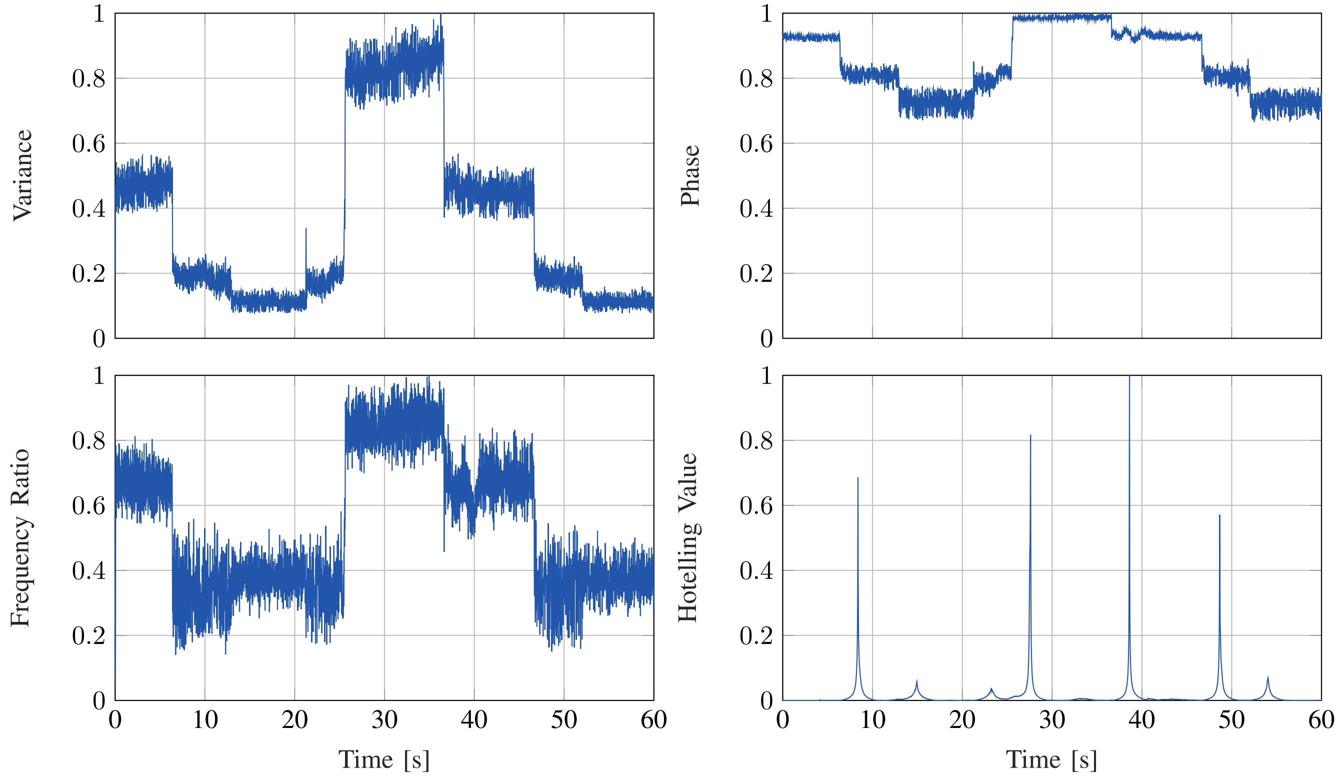

Figure 2. Extracted features and resulting Hotelling value from one phase. Values are normalized between 0 and 1 for readability.

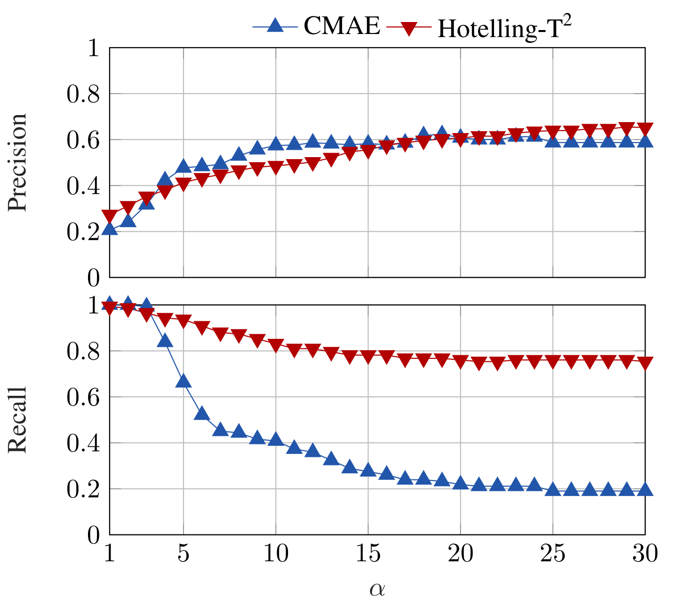

Figure 3. Precision and recall as function of applied threshold.

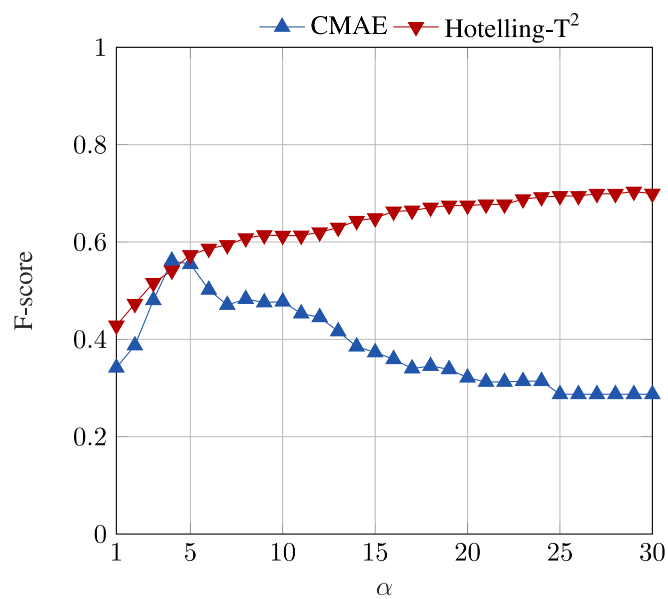

Figure 4. F-score as function of applied threshold.

presented approach with other high-frequency datasets and event detection algorithms.

ACKNOWLEDGMENT

This work is partially funded by the Central Federal Association of the Health Insurance Funds of Germany (GKV-Spitzen-verband) in the context of the QuoVadis research project and by the German Ministry for Education and Research (BMBF) within the joint research projects MeSiB (grant 16SV7723) and Audio-PSS (grant 02K16C201).

 ISBN: 978-1-61208-739-9

References

[1] G. W. Hart, "Nonintrusive appliance load monitoring," *Proceedings of the IEEE*, vol. 80, no. 12, pp. 1870–1891, 1992.

[2] Union, European, "Directive 2009/28/ec of the european parliament and of the council of 23 april 2009 on the promotion of the use of energy from renewable sources and amending and subsequently repealing directives 2001/77/ec and 2003/30/ec," *Official Journal of the European Union*, vol. 5, p. 2009, 2009.

[3] World Health Organization, "World report on ageing and health," 2015.

[4] C. Chalmers, W. Hurst, M. Mackay, and P. Fergus, "Smart health monitoring using the advance metering infrastructure," in *International Conference on Computer and Information Technology; Ubiquitous Computing and Communications; Dependable, Autonomic and Secure Computing; Pervasive Intelligence and Computing (CIT/IUCC/DASC/PICOM)*. IEEE, 2015, pp. 2297–2302.

[5] J. M. Alcalá, J. Ureña, Á. Hernández, and D. Gualda, "Assessing human activity in elderly people using non-intrusive load monitoring," *Sensors*, vol. 17, no. 2, p. 351, 2017.

[6] A. Gerka, C. Lins, C. Lüpkes, and A. Hein, "Zustandserkennung von Beatmungsgeräten durch zentrale Messung des Stromverbrauchs," in *German Medical Science GMS Publishing House*, 09 2017.

[7] A. Zoha, A. Gluhak, M. A. Imran, and S. Rajasegarar, "Non-intrusive load monitoring approaches for disaggregated energy sensing: A survey," *Sensors*, vol. 12, no. 12, pp. 16 838–16 866, 2012.

[8] T. Onoda, Y. Nakano, and K. Yoshimoto, "System and method for estimating power consumption of electric apparatus, and abnormality alarm system utilizing the same," Nov. 9 2004, uS Patent 6,816,078.

[9] H. Y. Lam, G. Fung, and W. Lee, "A novel method to construct taxonomy electrical appliances based on load signatures," *IEEE Transactions on Consumer Electronics*, vol. 53, no. 2, pp. 653–660, 2007.

[10] A. Diop, T. Jouannet, K. E. K. Drissi, and H. Najmeddine, "Method and device for the non-intrusive determination of the electrical power consumed by an installation, by analysing load transients," Oct. 8 2013, uS Patent 8,554,501.

[11] K. D. Anderson, M. E. Bergés, A. Ocneanu, D. Benitez, and J. M. Moura, "Event detection for non intrusive load monitoring," in *IECON 38th Annual Conference on Industrial Electronics Society*. IEEE, 2012, pp. 3312–3317.

[12] L. Norford, S. Leeb, D. Luo, and S. Shaw, "Advanced electrical load monitoring: A wealth of information at low cost," *Diagnostics for Commercial Buildings: from Research to Practice, Pacific Energy Institute, San Francisco, CA*, 1999.

[13] A. R. Rababaah and E. Tebekaemi, "Electric load monitoring of residential buildings using goodness of fit and multi-layer perceptron neural networks," in *International Conference on Computer Science and Automation Engineering (CSAE)*, vol. 2. IEEE, 2012, pp. 733–737.

[14] Z. Zhu, S. Zhang, Z. Wei, B. Yin, and X. Huang, "A novel cusum-based approach for event detection in smart metering," in *IOP Conference Series: Materials Science and Engineering*, vol. 322, no. 7. IOP Publishing, 2018, p. 072014.

[15] L. De Baets, J. Ruyssinck, D. Deschrijver, and T. Dhaene, "Event detection in nilm using cepstrum smoothing," in *3rd International Workshop on Non-Intrusive Load Monitoring*, 2016, pp. 1–4.

[16] J. M. Alcalá, J. Ureña, and Á. Hernández, "Event-based detector for non-intrusive load monitoring based on the hilbert transform," in *Emerging Technology and Factory Automation (ETFA)*. IEEE, 2014, pp. 1–4.

[17] K. Anderson *et al.*, "Blued: A fully labeled public dataset for event-based non-intrusive load monitoring research," in *Proceedings of the 2nd KDD workshop on data mining applications in sustainability (SustKDD)*, 2012, pp. 1–5.

[18] L. Farinaccio and R. Zmeureanu, "Using a pattern recognition approach to disaggregate the total electricity consumption in a house into the major end-uses," *Energy and Buildings*, vol. 30, no. 3, pp. 245–259, 1999.

[19] M. N. Meziane, P. Ravier, G. Lamarque, J.-C. Le Bunetel, and Y. Raingeaud, "High accuracy event detection for non-intrusive load monitoring," in *International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2017, pp. 2452–2456.

[20] A. U. Rehman, T. T. Lie, B. Vallès, and S. R. Tito, "Event-detection algorithms for low sampling nonintrusive load monitoring systems based on low complexity statistical features," *IEEE Transactions on Instrumentation and Measurement*, 2019.

[21] L. Jiang, S. Luo, and J. Li, "Automatic power load event detection and appliance classification based on power harmonic features in nonintrusive appliance load monitoring," in *8th IEEE Conference on Industrial Electronics and Applications (ICIEA)*. IEEE, 2013, pp. 1083–1088.

[22] A. Gerka, F. Lübeck, M. Eichelberg, and A. Hein, "Electricity metering for dementia care," in *VDE-Kongress – Internet der Dinge. Technologien / Anwendungen / Perspektiven.*, V. e. V., Ed. VDE Verlag, 11 2016, inproceedings, p. Paper 54.

[23] H. Hotelling, "The generalization of student's ratio," in *Breakthroughs in statistics*. Springer, 1992, pp. 54–65.

[24] J. P. Stevens, *Applied Multivariate Statistics for the Social Sciences*. Routledge Member of the Taylor and Francis Group, 2009, vol. 5.

[25] A. Trujillo-Ortiz and R. Hernandez-Walls, "Hotellingt2: Hotelling t-squared testing procedures for multivariate tests. a matlab file," 2002.

[26] S. Makonin and F. Popowich, "Nonintrusive load monitoring (nilm) performance evaluation," *Energy Efficiency*, vol. 8, no. 4, pp. 809–814, 2015.

[27] L. I. Kuncheva, "Change detection in streaming multivariate data using likelihood detectors," *IEEE Transactions on Knowledge and Data Engineering*, vol. 25, no. 5, pp. 1175–1180, 2013.