Frequency Invariant Transformation of Periodic Signals (FIT-PS) for Classification in NILM

Pirmin Held, Steffen Mauch, Alaa Saleh, Djafar Ould Abdeslam, Dirk Benyoucef

Frequency Invariant Transformation of Periodic Signals (FIT-PS) for Classification in NILM

Pirmin Held, *Student Member, IEEE,* Steffen Mauch, Alaa Saleh, *Student Member, IEEE,* Djaffar Ould Abdeslam, *Senior Member, IEEE,* and Dirk Benyoucef, *Member, IEEE*

***Abstract*—This paper presents a new signal representation called Frequency Invariant Transformation of Periodic Signals (FIT-PS) for the context of Non-Intrusive Load Monitoring (NILM). Compared to former approaches, where a conglomeration of different signal forms has been used, the presented approach is based on a single signal form containing all information. The core idea of this work is to use the original current waveform relatively to the reference voltage as a signature for NILM. In general, the relation of sampling and grid frequency is subject to continuous fluctuations. Therefore, FIT-PS converts uncorrelated sample data to a fixed multiple of the grid frequency. The advantages are that the information of the current signal, as well as the phase shift between voltage and current signal are completely contained in the FIT-PS signal representation. For classification, a neural net was applied to the HELD1 [1] dataset. Features created by FIT-PS are clearly superior to the standard features. With 18 different devices, a detection rate of up to 90 % is achieved. In particular, when several consumers are active at the same time, the new signal representation is much more robust and leads to a better detection rate. However, a Long Short-Term Memory (LSTM) net with FIT-PS signal representation provides the best results.**

***Index Terms*—feature extraction, feed forward neural net, LSTM net, non-intrusive load monitoring, signal form, supervised classification**

I. Introduction

THE idea of measuring the total current and voltage of a house and deducting the individual energy consumption of devices has been published by George W. Hart in 1985. Hart *et al.* clearly described the essential steps of NILM [2], [3], which are shown in Fig. 1 and are still valid today.

The main motivation for NILM is to allocate the power consumption to the individual appliances, which would ideally lead to better awareness and a change in consumption habits in order to save energy [4]–[7]. Other possible applications for NILM are predictive maintenance [8], [9] and analysis of the user behavior in the Ambient Assisted Living (AAL) context [10], [11].

The disaggregation process of NILM can be described as a linear tool chain of six different steps: measurement, signal waveform, event detection, feature extraction, classification, and tracking (see Fig. 1). The first step is required to record the measured current and voltage. Afterwards, the sampled data is converted into a non-periodic signal form like active power, reactive power or other signal forms. The subsequent event detection is used to identify switching on and off processes of individual appliances. After identifying the switching on and off processes of devices, additional information can be obtained by feature extraction. The feature can be differentiated into steady state and transient state feature. In general it is assumed that there are no significant changes during a steady state. Steady State Feature (SSF) are usually generated by comparing the signal characteristics before and after an event, while Transient State Feature (TSF) make use of the signal form during the turning on or off process of a device.

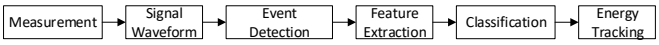

Fig. 1. Conventional approach of the processing chain for NILM applications

The most commonly applied features in NILM are steady state features using active power P and reactive power Q. While calculating P and Q, the high frequency parts of the signal are suppressed because P and Q are defined by evaluating the integral over a period which acts as a low-pass filter. To counteract this loss of information, additionally, numerous features have been combined along P and Q, e.g. current waveform, harmonics, instantaneous admittance waveform, instantaneous power waveform, eigenvalues, and switching transient waveforms [12], [13]. The most widely used complement to P and Q are the harmonics [14]–[16]. Using harmonics as feature for the classification at different datasets leads to most promising results [17].

Until now, no single feature leads to optimal results in the context of NILM because no feature includes all relevant information. It is difficult to choose the right combination of features as the selection depends on numerous factors such as the kind of individual devices and the sampling rate. Furthermore, with changing devices, the optimal combination of the selected waveform is subject to change. Therefore, our main aim was to develop a feature which contains as much information as possible from the original signal. However, the calculation effort for the feature and the subsequent classification should be low in order to allow an implementation on embedded systems.

For periodic signals, we present a signal representation (FIT-PS) which uses the original current waveform relative to the reference voltage. Using FIT-PS for event detection applied on the BLUED dataset [18] led to promising results [19], [20]. In order to compare the proposed signal form with common

P. Held, S. Mauch, A. Saleh and D. Benyoucef are with the Mechanical and Medical Engineering Department, University Furtwangen, 78120 Germany e-mail: {pirmin.held, steffen.mauch, alaa.saleh, dirk.benyoucef}@hs-furtwangen.de

D. O. Abdeslam is with the Department of Electrical Engineering and Computer Science. University Haute-Alsace, 68093 Mulhouse, France e-mail: djaffar.ould-abdeslam@uha.fr

signal forms in the context of the classification in NILM, the Home Equipment Laboratory Dataset 1 (HELD1) dataset [1] is used, and a neural net is applied for classification.

This paper is structured as follows: The preliminaries are given in the subsequent Sec. II. In Sec. III, the proposed signal form is described followed by Sec. IV, which introduces different classification methods being applied on a Feed Forward (FF) and a LSTM net. Afterwards, the results are shown in Sec. V followed by the conclusion in Sec. VI.

II. Preliminaries

The mathematical modeling of the fundamental problem of NILM can be described as follows [21]:

$$y(t) = \sum_{d=1}^{D} y_d(t) \qquad \text{for} \qquad t = 1, ..., T \tag{1}$$

The sum of the power consumption of several devices d at time t is described by $y(t) \in \mathbb{R}$. The number of devices in a setting is given by D whereas the power consumption of the device d at time t by $y_d(t)$. In reality, neither the number of devices nor their specific properties are known in advance. From a mathematical point of view, the inverse solution of (1) is the goal of NILM. More precisely, the goal for energy disaggregation is to find y_d for $d = 1, ..., D$ for given y. Since voltage and current are the measurable quantities, all further signals are derived from these signals. Equation (2) and (3) represent the measured signal [19] (current $i(t)$ and voltage $v(t)$). $I(t)$ and $V(t)$ represent general non-linear signals. The non-linear signals lead to higher harmonics of the mains frequency. The time-dependent frequency is given by $f(t)$ and the phase angle between $v(t)$ and $i(t)$ is given by $\varphi(t)$.

$$v(t) = V(t) \cdot \cos\Big(2\pi f(t) \cdot t\Big) \tag{2}$$

$$i(t) = I(t) \cdot \cos\Big(2\pi f(t) \cdot t + \varphi(t)\Big) \tag{3}$$

For continuous time signals, the root mean square (RMS) can be calculated by

$$i_{rms}(t) = \sqrt{\frac{1}{t_1 - t_0} \cdot \int_{t_0}^{t_1} i(t)^2 dt} \tag{4}$$

$$v_{rms}(t) = \sqrt{\frac{1}{t_1 - t_0} \cdot \int_{t_0}^{t_1} v(t)^2 dt} \tag{5}$$

where $t_1 - t_0$ is the time over one or more periods. By multiplying the current and voltage signal, the active power (P) and reactive power (Q) can be calculated as shown in (6) and (7).

$$P(t) = \frac{1}{2} \cdot v_{rms}(t) \cdot i_{rms}(t) \cdot \cos(\varphi(t)) \tag{6}$$

$$Q(t) = \frac{1}{2} \cdot v_{rms}(t) \cdot i_{rms}(t) \cdot \sin(\varphi(t)) \tag{7}$$

Due to the integration in (4) and (5), the average is determined, and therefore, high-frequency signal components are filtered out.

To take high frequency signals into account, the harmonics have been used in the NILM context. The harmonics $H \in \mathbb{R}^{15}$ are calculated for each device D as follows:

$$H_D(f) = \int_{-\infty}^{+\infty} s(t) e^{-j2\pi ft} \;\Big|\; f = 50, 100, \ldots, 750\,\text{Hz} \tag{8}$$

where $s(t)$ contains 75 periods of the steady state.

The information being provided for the classification is denoted as a feature. Other common literature expressions are finger prints or signatures [22]. The information for SSF is obtained from the steady state, the information for TSF is obtained from the transient state [3]. The use of SSF is more common since it leads to more stable results, especially with a limited number of devices, even at low sampling rates. TSF

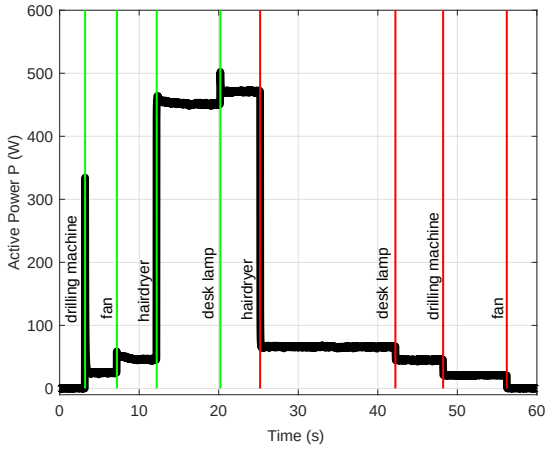

Fig. 2. Typical performance diagram of a household

requires a higher sampling rate of several kHz because the TSF time of most applications is limited to a few seconds. Fig. 2 shows a classic performance diagram with multiple turning on and off events of different appliances of a household. The detailed signal curve in Fig. 3 shows the switching-on behavior (TSF) of a common drilling machine and a fan.

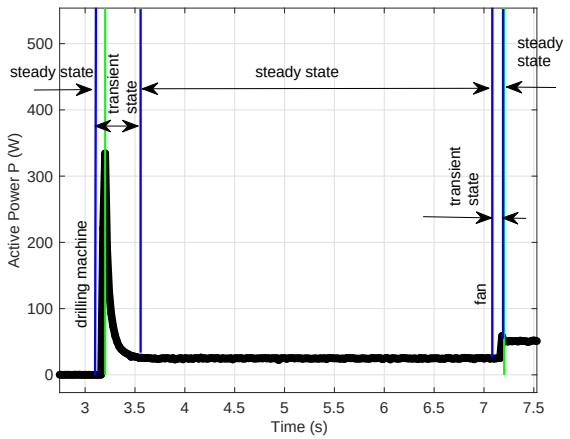

Fig. 3. Details of the switching on process of a drilling machine and a fan

In practice, several devices can be running at the same time. If the absolute values of the signal are used as a

feature, it results in the combinatorial possibilities described in equation (9).

$$N = \sum_{k=1}^{n} \binom{n}{k} \tag{9}$$

Here n depicts the number of applications, and N shows the number of possible combinations of devices which are switched on simultaneously. For instance, if there are only $n = 4$ applications, the number of possible combinations of appliances running at the same time are $N = 15$. In a typical household, if there are more than 20 devices present, then, in order to distinguish each device, more than 1048575 possibilities have to be taken into account. Therefore, it appears that an overlap in the feature space is very likely.

III. Frequency Invariant Transformation (FIT-PS)

The previously described waveforms have not been designed with the goal of classifying different devices. Nevertheless, in the NILM context these waveforms are predominantly used. To improve the results of classification, different signal forms are combined. The most common signal combination are P and Q combined with harmonics. Because of the already mentioned difficulties during the combination of features and because harmonics do not present the information within one period in such a way that the classification could take full advantages out of it, a new approach based on the FIT-PS signal representation is presented here. FIT-PS has the goal of transferring as much information as possible from the periodic current / voltage signal into a non-periodic signal.

The following pseudo code (Algorithm 1 to Algorithm 4) describes the FIT-PS method. The FIT-PS method divides the sampled signal with respect to its fundamental frequency, thus, the line frequency. Therefore, the first step is the detection of the zero crossings: The variable 'IDXZero' has as much

Algorithm 1: Finding zero crossings

$nn = 1$
for $n = 1 : N - 1$ **do**
  if $(v(n) < 0)$ AND $(v(n+1) > 0)$ **then**
    IDXZero$(nn{+}{+}) = n$

entries as the signal has zero crossings. The pseudo code neglects additional checks to eliminate multiple detection of zero-crossings for one period due to simplicity.

Based on the absolute position of the zero crossing for the given data, Algorithm 2 describes the linear approximation of the zero crossing. 'IDXZeroShift' is, for a given zero crossing,

Algorithm 2: Calculation of the exact position of zero crossings

for $l = 1 :$ length(IDXZero) $- 1$ **do**
  $n =$ IDXZero(l)
  IDXZeroShift$(l) = \left(-v(n) \cdot \frac{1}{v(n+1)-v(n)}\right)$

the estimated shift of the crossing itself. IDXZeroShift ranges between zero to one; thus, the maximum deviation is less than one sample. Due to possible variation of the amplitude, the maximum steepness and almost constant derivative of the sinusoidal signal, the zero crossing is chosen to define the length of a period.

After the zero crossings have been found, the indices can be assigned to the individual periods. So far, this has been done by a linear interpolation as described in Algorithm 3.

Algorithm 3: Periodical allocation by interpolation

$X_{l,k} = 0$
for $l = 1 :$ length(IDXZeroShift) $- 1$ **do**
  Len $=$ (IDXZero$(l+1)$ + IDXZeroShift$(l+1)$) $-$ (IDXZero(l) + IDXZeroShift(l))
  Dis $= \frac{\text{Len}}{\#\text{SampPerPeriod}}$
  for $k = 2 :$ #SampPerPeriod **do**
    $k1 =$ IDXZero(l) + IDXZeroShift(l) + Dis $\cdot (k-1)$
    $k2 =$ floor$(k1)$
    $k3 =$ ceil$(k1)$
    $X_{l,k} =$ intp(IDXZeroShift(l), current$(k2)$, current$(k3)$)

$X_{l,k}$ depicts the resulting matrix as having the dimension of $n_l \times n_k$ where n_l is the number of periods and n_k is the number of sampling points of each period. The number of sampling points per period is given in k direction whereas the number of periods is given in l direction. The following parameters have been selected for the application of the HELD1 data set: $n_k = \frac{fs}{fg} = \frac{4\,\text{kHz}}{50\,\text{Hz}} = 80$ and $n_l = 75$ periods. The minimum distance between the events is 3 seconds. While calculating the steady state, 1.5 seconds before and after the event are ignored to exclude the transient state.

The interpolation function ('intp') performs a linear interpolation. The argument 'interval' has a range of zero to one,

Algorithm 4: Function of the linear interpolation

float intp(interval, $f1$, $f2$)
return $f1 + (f2 - f1) \cdot$ interval

whereas $f1$ and $f2$ are two values. 'intp' returns the calculated value at position 'interval'; thus, interpolates the value between $f1$ and $f2$ at position 'interval'. Due to the linear interpolation, a slight deviation from the original signal may occur. To keep this as low as possible, the aliasing filter of the measurement system must be designed correctly.

Fig. 4 describes the method graphically using a voltage and current signal:

(a) The voltage signal is used to detect the length of one period. Therefore, the zero crossings are calculated.
(b) Depending on the length of the period, the new sample positions are defined.
(c) The current signal is interpolated with the new sampling positions. Now, each period has the same number of samples.

(d) The signal is reshaped after each period. The new signal representation can be seen as a k dimensional signal with the length of periods l.

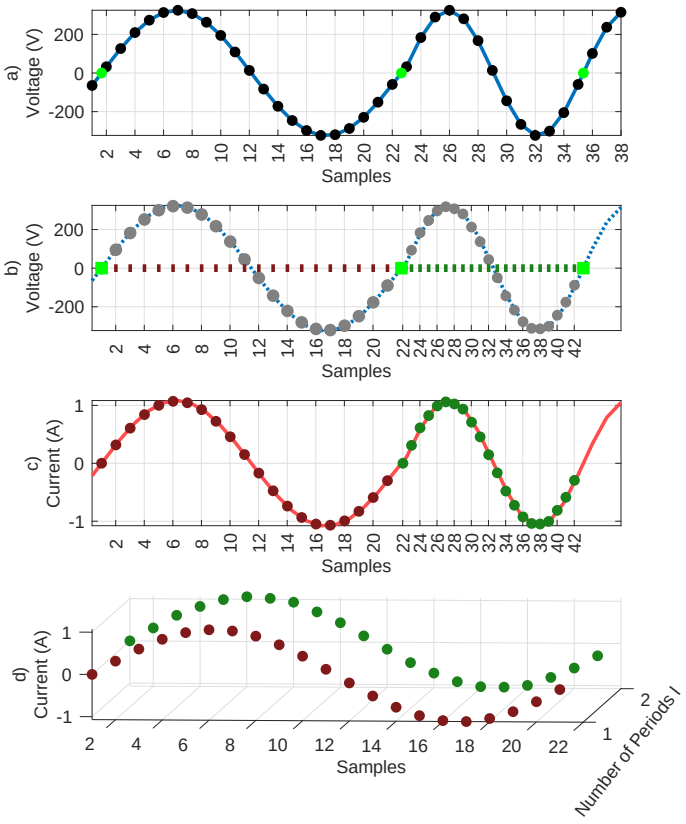

Fig. 4. Graphical representation of FIT-PS. a) finding zero crossings, b) re-sampled sampling positions, c) interpolation of the current signal based on the new sampling positions, d) periodical signal representation. Modified illustration from [20].

Compared to Fig. 4 d) a large number of periods is visualized in Fig. 5. The legend of the current amplitude is additionally colored. Here, a sample frequency of 4 kHz while having a grid frequency of 50 Hz was used. This results in 80 samples per period or a 80-dimensional signal. As one can clearly see, the current signal differs from a perfect sinusoidal signal. Considering each dimension individually, the periodical signal is transformed into a non-periodical signal.

Fig. 6 shows a FIT-PS signal of a switching on event of a refrigerator with l periods and k dimensions. The amplitude of the current is colored. Before switching the refrigerator on, other devices were already running. In Fig. 6, the exact position of switching on can be seen. The ending of the transient state can not be identified clearly. The advantage of this signal form and the corresponding feature is that all the information from the current signal is included. Additionally, the phase shift is also included since the position of the sample points depends on the zero crossing of the voltage signal. This is of particular importance since it allows to distinguish between ohmic, capacitive, or inductive loads which have the same or similar consumption. The calculation effort of FIT-PS

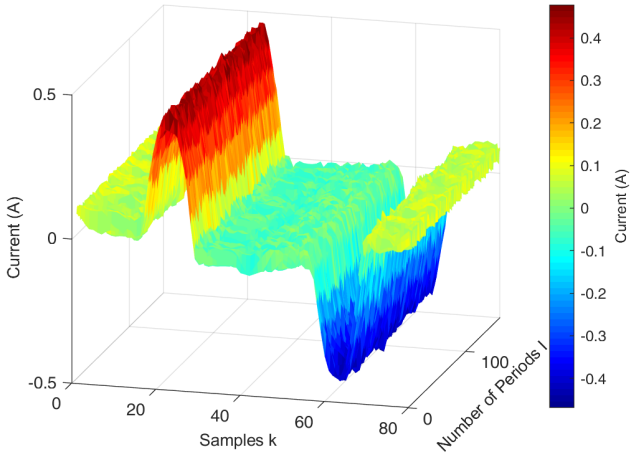

Fig. 5. Illustration of FIT-PS for the steady state of device ID 22 (Multifunction Tool Dremel®). Progress of i_p over several periods over l, k and the amplitude of i_p additionally shown in colors.

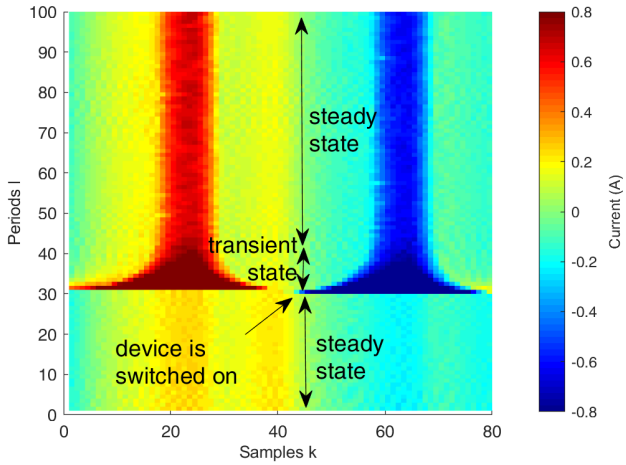

Fig. 6. Signal representation of two steady states and one transient state with FIT-PS . The amplitude of i_p is shown colorized while it is set to a limit of ± 0.8 A for a better presentation. During the transient, a maximum value of ± 3.3 A is reached.

is very low, i.e. we have developed a real-time implementation on one core of a Raspberry Pi2 having a sample rate of 4 kHz (12 bit).

IV. Classification

Feed forward neural networks are investigated with classical features and FIT-PS for classification. For all features scaled conjugate gradient back-propagation is used for training. In order to use the transient part of the signal as well, an additional LSTM network with FIT-PS features is applied. Here, a stochastic gradient descent with momentum (SGDM) is used for training.

For the training and testing, the HELD1 laboratory dataset [1] is used. The dataset includes 18 different appliances and clear reference data. In addition to 100 training measurements of each individual device, there are three different measurement scenarios. These differ in the number (one, up to four,

up to six) of simultaneously active devices. In sum, there are 3,600 events of single measurements and 13,200 events of measurements with multiple appliances recorded with a sample frequency of 4 kHz. The minimum distance between the events of this dataset is 3 seconds. The classifiers were trained with the single measurements based on the HELD1 dataset [1], in which each device is switched on and off 100 times.

To find the optimal number of the harmonics used in combination with P and Q the neural net is trained with different numbers of harmonics. The best result can be achieved with 15 harmonics. A larger number of harmonics causes a deterioration of the classification results. A reason for this is that the amplitude of the harmonics decreases with the higher order of the used harmonic. Depending on the devices and the measurement system, the amplitude of the high order harmonic is not significantly higher than the noise.

Feed forward (FF) neural network

The structure of the feed forward net being applied is shown in Fig. 7. Depending on the waveform, a 17 (P, Q, and 15 harmonics) or 80 (FIT-PS with 80 samples per period) dimensional input vector is prevalent. Since the tests with a higher number of hidden layers could not improve the accuracy of the classifier, only one hidden layer is used. The dimension of the output layer depends on the number of appliances to be distinguished. Since it is easy to distinguish between on or off events, two different classifiers, one for on events and one for off events, are used.

To find the optimal number of neurons, the validation data was applied to the net structures with one neuron up to 500 neurons. Therefore, the data 0003, 0116, and 0201 of the HELD1 dataset have been used as validation data. These measurements of the dataset have been selected for validation because they include all used devices. The accuracy is calculated as follows.

$$\text{accuracy} = \frac{\text{number of correct classified events}}{\text{total number of events}} \tag{10}$$

The results are shown in Fig. 8 and 9 with blue dots. Depending on the initialization values of the neural net, random numbers in general, different results can be achieved. Due to this, 100 iterations with different initialization settings and different numbers of neurons have been simulated. The lower (red line), upper (green line), and average (black line) limits of accuracy were calculated in Fig. 8 and 9. The maximum

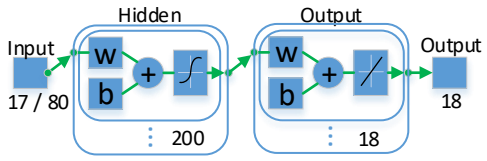

Fig. 7. Structure of the feed forward neural net. The number of input dimensions depends on whether P, Q, and harmonics are used as a feature or FIT-PS.

accuracy of both features can be achieved with approximately 200 neurons. A significant higher number of neurons does

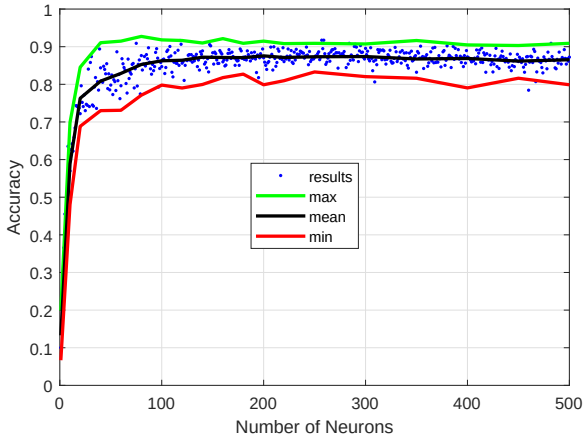

Fig. 8. Results of a feed forward net using P, Q, and harmonics on different numbers of neurons. For training, single measurements have been used, while for testing, the measurements 0003, 0116, and 0201 of HELD1 [1] have been used. The limits of accuracy are calculated by 100 iterations with different initialization settings and different numbers of neurons.

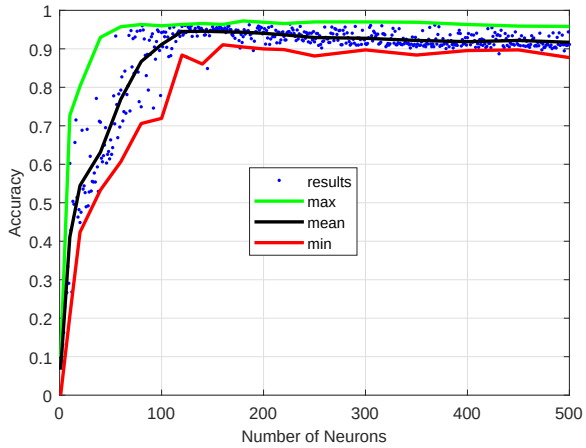

Fig. 9. Results of a feed forward net using FIT-PS on different numbers of neurons. For training the single measurements, for testing the measurement 0003, 0116, and 0201 of HELD1 [1] have been used. The limits of accuracy are calculated by 100 iterations with different initialization settings and different numbers of neurons.

not improve the accuracy significantly. The accuracy achieved with FIT-PS is higher than the accuracy obtained with the standard signal forms P, Q, and harmonics, regardless how many neurons are used. Furthermore, the variance of the results, caused by different initializations is significantly lower with FIT-PS as input signal.

Long short-term memory (LSTM) network

In order to use additional information from the TSF for classification, FIT-PS was extended to a LSTM network, Fig. 10. The used LSTM network consists of $n_l = 100$ units. Thus, a time range of 5 periods before and 95 periods during and after the event can be used for classification. The complete structure of the applied neural net can be seen in Fig. 11. At first, an 80-dimensional sequential input layer is used. The LSTM layer with 80 dimensions and a length of 100 LSTM

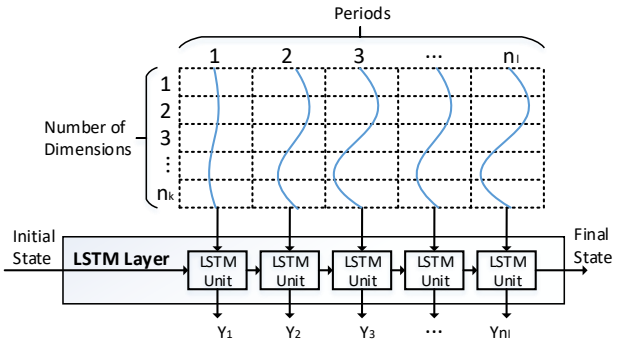

Fig. 10. LSTM layer

units follows. The output of the LSTM net is connected to the fully connected layer which has an output size of 18 (equal to the number of classes). Here the previously learned features from the LSTM net are combined finally for the classification. A softmax activation function calculates a probability between zero and one for each class. The classification or output layer complete the network. The device with the highest probability is classified.

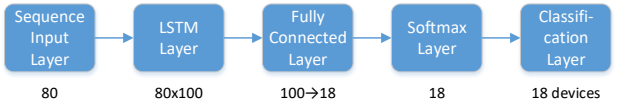

Fig. 11. Complete structure of the LSTM network

Problems with multiple local minimums lead to different results for different initialization seeds. To circumvent the local minimums, different initialization seeds are used on the set of validation data. The results are shown in Fig. 13. For classification of the test data, the initialization setting corresponding to the best results of the validation data is used.

V. Results

The classifiers have been trained with all 18 available appliances. Table I and II show the results of measurement 0194 of the dataset using standard signal forms and the proposed signal form. The ID of the devices in the measurement were listed horizontally and vertically. The devices with the ID 13, 15, 21, 23 did not appear in the measurement, but were falsely detected. The tables show the number of correctly classified devices in the diagonal (marked in green). E.g. in Table I for the device ID 5, this means that it has been recognized correctly 36 times. However, it was incorrectly identified once as device ID 14 and three times as device ID 13.

The results of this measurement are not representative for the whole dataset because the average recognition rate for the dataset is higher. This measurement was selected to explain the following problem: Device ID 13 and ID 14 are refrigerators, which have a very similar power consumption behavior. Therefore, misclassification of both devices occurs frequently for standard and FIT-PS features. Using a FF net with P, Q, and harmonics for this part of the dataset, an accuracy of 66 % could be reached whereas FIT-PS reaches an accuracy of 75 %.

TABLE I
Conversion matrix of the classification results of measurement "0194" using P, Q, and harmonics in a feed forward net [1]

Dev. ID actual / set	3	5	7	9	10	11	14	16	17	19	13	15	21	23
3	20	0	1	0	0	1	0	0	0	0	1	0	3	14
5	0	36	0	0	0	0	1	0	0	0	3	0	0	0
7	0	0	33	0	0	5	0	0	0	0	1	0	1	0
9	2	0	0	29	0	0	0	4	0	0	1	0	4	0
10	1	0	0	0	25	2	0	1	0	0	7	0	0	4
11	1	0	2	0	0	28	0	4	0	0	1	0	1	3
14	0	0	0	0	0	6	6	0	0	0	22	1	4	1
16	0	0	0	0	0	1	0	29	0	1	0	0	0	9
17	0	0	1	0	0	4	0	0	32	0	1	0	2	0
19	3	0	0	0	0	0	0	4	0	26	0	0	6	0

TABLE II
Conversion matrix of the classification results of measurement "0194" using FIT-PS in a feed forward net

Dev. ID actual / set	3	5	7	9	10	11	14	16	17	19	13	20	22	23
3	36	0	0	0	0	2	0	1	0	0	0	0	0	1
5	0	30	0	0	0	3	0	0	0	0	1	0	0	6
7	0	0	40	0	0	0	0	0	0	0	0	0	0	0
9	2	0	0	23	1	3	0	5	0	0	0	0	0	6
10	0	0	0	0	39	0	0	0	0	0	0	0	0	1
11	2	0	0	0	0	34	0	3	0	0	0	0	0	1
14	0	0	0	0	0	3	3	1	0	0	31	0	0	2
16	0	0	0	0	0	0	0	40	0	0	0	0	0	0
17	0	0	0	0	0	3	0	1	32	0	0	2	2	0
19	1	0	0	0	3	4	0	1	0	24	0	0	0	6

Results of all different features are shown in Fig. 12. They are divided into measurements in which one, up to four, and up to six appliances are active. The features and classifiers utilized are visualized in different colors. In order to test whether the initialization parameters selected during the validation also lead to good results for the test data, all 100 initialization parameters were used for the test data. The different initialization parameters lead to different results. These vary according to the feature utilized, classifiers, and the number of simultaneously active devices. The LSTM net has the largest spread. During the validation the classifier with the nearly optimal initialization seed can be selected. The dotted lines in the bars show the results reached with the best initial values from the validation. The solid lines in the bars show the median of all different test initial values of the classifier.

When one appliance is active at one time, the FIT-PS feature with a FF net reaches the highest accuracy with 90 %. FIT-PS outperforms P, Q, and the harmonics clearly in all scenarios. The differences become particularly clear if several devices are switched on at the same time. Here, the results remain most stable when using the LSTM network. Fig. 13 shows the accuracy of the validation data ('$\circ$') and the accuracy of the test results ('$\times$') for various initialization settings. The figure shows the relationship between the accuracy of the validation data and the test data for each initialization. This relationship is most clearly demonstrated for FIT-PS with the LSTM network (orange): the better the initialization parameters were chosen, the higher the accuracy of the test results became, based on the results of the validation data. The results of the test data of P and Q (green) are widely spread. This means that the best initialization setting determined with the validation data does not necessarily represent a good initialization setting for the test data. Consequently, it is difficult to find the optimal initialization setting for P and Q. Although the additional use

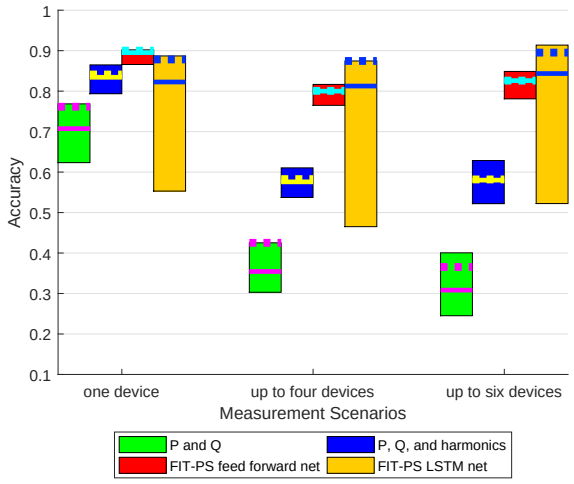

Fig. 12. Accuracy of a feed forward net using P and Q (green); P, Q, and harmonics (blue); FIT-PS (red); LSTM net using FIT-PS (orange). The dotted lines represent the accuracy using the best initialization setting derived from the validation data. The solid lines represent the median of the accuracy of the test data.

of the harmonics (blue) leads to a reduction in the variance of the test results, the increase remains low. This is also the reason why in Fig. 12 the selected initialization (dotted line) is only slightly better than the median (solid line) of all initialization settings. Thus, even with the additional use of harmonics, it remains difficult to determine the optimal initialization setting through the validation data. The same applies to the results with FIT-PS, wherein FIT-PS achieves the lowest variance. However, for all combinations of scenarios and features the use of the best initialization setting based on validation leads to an accuracy similar or most often even superior to the median accuracy. Therefor, the use of the validation data to identify the optimal initialization, especially for the LSTM network, proves to be advantageous.

Because of different dimensions of the features used for the classification, the training time changes. The achieved accuracy and the average training time are presented in Table III. The simulations were done on the following computational environment: Intel(R) Core(TM) i7-6800K CPU@3.40GHz; RAM: 64GB; Win 10 Pro 64-Bit, Matlab 2018a.

TABLE III
TOTAL RESULTS OF CLASSIFICATION USING DIFFERENT FEATURES.

Feature	Accuracy	Training time
P, Q	54.2 %	36.5 s
P, Q, harmonics (15)	67.9 %	97.4 s
FIT-PS feed forward net	84.1 %	259.8 s
FIT-PS LSTM net	87.9 %	670.3 s

Using FIT-PS as a feature increases the accuracy. The higher dimensions of FIT-PS (80 dimensions for the feed forward net) result in a higher computational effort for training of the neural net. FIT-PS in combination with a LSTM net can improve the accuracy again. Here, changes during the transient state are taken into account for classification. However, the complexity of the classifier increases significantly.

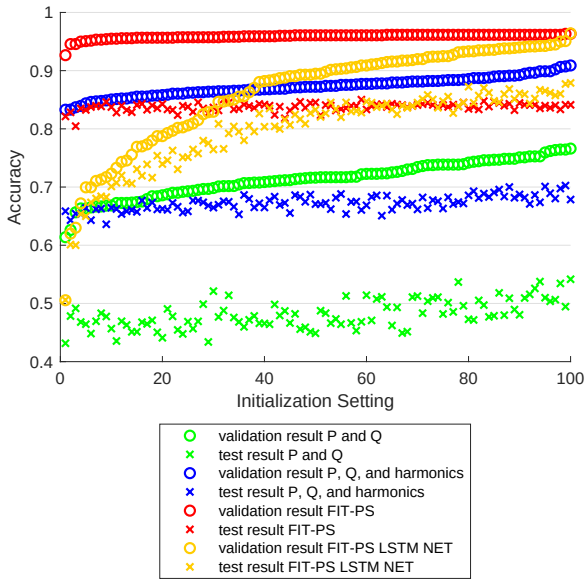

Fig. 13. Comparison of validation results and test results on different initialization settings. The initialization settings are sorted according to increasing accuracy of the validation data. P and Q (green); P, Q, and harmonics (blue); FIT-PS with FF net (red); FIT-PS with LSTM net (orange)

VI. CONCLUSION

A new signal representation for the analysis of periodic signals in the context of NILM has been applied.
The main advantage of FIT-PS is that all information from the current signal remains in the new signal. The information about the phase shift between voltage and current is indirectly represented since the voltage signal is used for calculating the new sampling positions for the current signal. The proposed signal representation and conventional signal forms are applied on a feed forward neural network for comparison. FIT-PS outperforms the compared signal forms clearly, reaching an accuracy of up to 90 %.
Using the information from the transient state by applying a LSTM net, provides the best results for several devices running at the same time. The use of validation data to identify the optimal initialization setting seems to be advantageous, especially for the LSTM net.

ACKNOWLEDGMENT

We thank Andrew McDouall and Dr. Corinna Brichta for proofreading the manuscript.
This work was created as part of the iMon project (funding number 03FH001IX4).

REFERENCES

[1] P. Held, S. Mauch, A. Saleh, D. Abdeslam Ould, and D. Benyoucef, “HELD1: Home Equipment Laboratory Dataset for Non-Intrusive Load Monitoring,” Nice, France, 2018, accepted for publication.

[2] G. Hart, "Residential energy monitoring and computerized surveillance via utility power flows," *IEEE Technology and Society Magazine*, vol. 8, no. 2, pp. 12–16, Jun. 1989.

[3] G. W. Hart, "Nonintrusive appliance load monitoring," *Proceedings of the IEEE*, vol. 80, no. 12, pp. 1870–1891, 1992, 00954.

[4] F. Luo, G. Ranzi, W. Kong, Z. Y. Dong, S. Wang, and J. Zhao, "Non-intrusive energy saving appliance recommender system for smart grid residential users," *IET Generation, Transmission & Distribution*, vol. 11, no. 7, pp. 1786–1793, May 2017.

[5] L. Yu, H. Li, X. Feng, and J. Duan, "Nonintrusive appliance load monitoring for smart homes: Recent advances and future issues," *IEEE Instrumentation & Measurement Magazine*, vol. 19, no. 3, pp. 56–62, 2016.

[6] P. Gao, S. Lin, and W. Xu, "A Novel Current Sensor for Home Energy Use Monitoring," *IEEE Transactions on Smart Grid*, vol. 5, no. 4, pp. 2021–2028, Jul. 2014, 00002.

[7] J.-X. Chin, T. Tinoco De Rubira, and G. Hug, "Privacy-Protecting Energy Management Unit Through Model-Distribution Predictive Control," *IEEE Transactions on Smart Grid*, vol. 8, no. 6, pp. 3084–3093, Nov. 2017.

[8] T. K. Nguyen, I. Azarkh, B. Nicolle, G. Jacquemod, and E. Dekneuvel, "Applying NIALM Technology to Predictive Maintenance for Industrial Machines," vol. IEEE International Conference on Industrial Technology (ICIT), 2018, pp. 341 – 345.

[9] A. Aboulian, D. H. Green, J. F. Switzer, T. J. Kane, G. V. Bredariol, P. Lindahl, J. S. Donnal, and S. B. Leeb, "NILM Dashboard: A Power System Monitor for Electromechanical Equipment Diagnostics," *IEEE Transactions on Industrial Informatics*, pp. 1–1, 2018. [Online]. Available: https://ieeexplore.ieee.org/document/8371632/

[10] P. Klein, "Non-Intrusive Information Sources for Activity Analysis in Ambient Assisted Living Scenarios," Ph.D. dissertation, Universite de Haute Alsace, Mulhouse, France, Nov. 2015. [Online]. Available: https://tel.archives-ouvertes.fr/tel-01526695/

[11] J. M. Alcala, J. Urena, A. Hernandez, and D. Gualda, "Sustainable Homecare Monitoring System by Sensing Electricity Data," *IEEE Sensors Journal*, vol. 17, no. 23, pp. 7741–7749, Dec. 2017.

[12] J. Liang, S. K. K. Ng, G. Kendall, and J. W. M. Cheng, "Load Signature Study-Part I: Basic Concept, Structure, and Methodology," *IEEE Transactions on Power Delivery*, vol. 25, no. 2, pp. 551–560, Apr. 2010, 00195. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5337912

[13] ——, "Load Signature Study-Part II: Disaggregation Framework, Simulation, and Applications," *IEEE Transactions on Power Delivery*, vol. 25, no. 2, pp. 561–569, Apr. 2010, 00083. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5337970

[14] R. Jonetzko, M. Detzler, K.-U. Gollmer, A. Guldner, M. Huber, R. Michels, and S. Naumann, "High frequency non-intrusive electric device detection and diagnosis," in *Smart Cities and Green ICT Systems (SMARTGREENS), 2015 International Conference on*. IEEE, 2015, pp. 1–8, 00001. [Online]. Available: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=7297979

[15] D. Jung, H. H. Nguyen, and D. K. Yau, "Tracking Appliance Usage Information Using Harmonic Signature Sensing," 2015, 00000. [Online]. Available: https://www.cs.purdue.edu/homes/lans/20061129/publications/jung2015sgc.pdf

[16] M. Dong, Meira, W. Xu, and C. Y. Chung, "Non-Intrusive Signature Extraction for Major Residential Loads," *IEEE Transactions on Smart Grid*, vol. 4, no. 3, pp. 1421–1430, Sep. 2013, 00030.

[17] M. M. Kahl, M. T. Kriechbaumer, M. A. Haq, and H.-A. Jacobsen, "Appliance Classification Across Multiple High Frequency Energy Datasets," p. 6, 2017.

[18] K. Anderson, A. Ocneanu, D. Benitez, A. Rowe, and M. Berges, "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)*, Beijing, China, 2012. [Online]. Available: http://inferlab.org/wp-content/uploads/2012/08/2012_anderson_SustKDD.pdf

[19] P. Held, F. Laasch, D. O. Abdeslam, and D. Benyoucef, "Frequency invariant transformation of periodic signals (FIT-PS) for signal representation in NILM," in *Industrial Electronics Society, IECON 2016-42nd Annual Conference of the IEEE*. IEEE, Oct. 2016, pp. 5149–5154, 00001.

[20] P. Held, A. Saleh, D. Abdeslam Ould, and D. Benyoucef, "Frequency Invariant Transformation of Periodic Signals (FIT-PS) for high frequency Signal Representation in NILM," in *3rd Baden-Württemberg Center of Applied Research Symposium on Information and Communication Systems*, Karlsruhe, Dec. 2016, pp. 1–6. [Online]. Available: http://sincom.informatik.hs-furtwangen.de/fileadmin/user_upload/proceedingsSInCom2016.pdf

[21] R. Dong, L. Ratliff, H. Ohlsson, and S. S. Sastry, "Fundamental limits of nonintrusive load monitoring," in *Proceedings of the 3rd international conference on High confidence networked systems*. ACM, 2014, pp. 11–18, 00014.

[22] R. Bonfigli, S. Squartini, M. Fagiani, and F. Piazza, "Unsupervised algorithms for non-intrusive load monitoring: An up-to-date overview," in *Environment and Electrical Engineering (EEEIC), 2015 IEEE 15th International Conference on*. IEEE, 2015, pp. 1175–1180, 00000.

Pirmin Held received his B.Sc. and M.Sc. degrees in Electrical Engineering and Microsystem Engineering at Furtwangen University in 2011 and 2012. He is a Ph.D. student at the department of Electrical Engineering and Computer Science of University Haute-Alsace. He is currently working at the Institute of Intelligent Systems (ISS) of the Faculty of Mechanical and Medical Engineering of the University of Furtwangen. His research interests are Non-Intrusive Load Monitoring, signal processing and classification.

Steffen Mauch received his B.Sc. and M.Sc. degrees in Electrical Engineering and Microsystem Engineering at Furtwangen University in 2010 and 2011. He received his doctoral degree in Control Engineering from Technische Universität Ilmenau in 2016. In his work he concentrates on applied signal processing such as Non-Intrusive Load Monitoring, wavefront sensor evaluation using FPGAs and embedded systems as well as related hardware development.

Alaa Saleh received his B.Sc. in communication engineering from HIAST, Syria in 2007 and M.Sc. in Communication and Information Engineering from TU Darmstadt in 2013. He is currently a research assistant in the Department of Mechanical und Medical Engineering in Furtwangen University, and working simultaneously on his PhD-Thesis in cooperation with University Haute-Alsace. His research interests are Energy Disaggregation, time-frequency analysis, machine learning and signal processing.

Djaffar Ould Abdeslam received the M.Sc. degree in electrical engineering from the University of Franche-Comté, Besançon, France, in 2002. He received the Ph.D. degree in 2005 and the Habilitation degree in electrical engineering in 2014 from the University of Haute-Alsace, Mulhouse, France. Since 2007, he has been an Associate Professor at the University of Haute-Alsace. His research interests include artificial neural networks applied to power systems, signal processing for power quality improvement, smart metering, smart building, and smart grids.

Dirk Benyoucef studied electrical engineering at the HTW of Saarland and at the Saarland University. He has received his doctorate in the Faculty of Engineering in the field of Telecommunication at Saarland University. Since 2004 he is professor at Furtwangen University and director of the Institute of Smart Systems (ISS). His research interests are applied signal processing and embedded systems, especially in the area of non-intrusive and smart sensor signal processing. He is a member of the VDE and IEEE.