Residential load event detection in NILM using robust cepstrum smoothing based method

Nur Iksan, Jaka Sembiring, Nanang Hariyanto, Suhono Harso Supangkat
School of Electrical Engineering and Informatics, Institut Teknologi Bandung, Indonesia

ABSTRACT
Event detection has an important role in detecting the switching of the state of the appliance in the residential environment. This paper proposed a robust smoothing method for cepstrum estimation using double smoothing i.e. the cepstrum smoothing and local linear regression method. The main problem is to reduce the variance of the home appliance peak signal. In the first step, the cepstrum smoothing method removed the unnecessary quefrency by applying a rectangular window to the cepstrum of the current signal. In the next step, the local regression smoothing weighted data points to be smoothed using robust least squares regression. The result of this research shows the variance of the peak signal is decreased and has a good performance with better accuracy. In noise environment, performance prediction quite good with values greater than 0.6 and relatively stable at values above 0.9 on SNR > 25 for single appliances. Furthermore, in multiple appliances, performance prediction quite good at SNR > 20 and begins to decrease in SNR < 20 and SNR > 25.

Keywords:
Event detection
Cepstrum smoothing
Local linear regression
Variance reduction

1. INTRODUCTION
Increased public awareness of energy conservation in recent years has created great interest in monitoring the use of electrical energy. Energy monitoring was originally intended to find out information regarding the amount of energy usage and fees to be paid. Currently, energy monitoring process aims to provide services to the load scheduling and demand response management. Monitoring energy use can be done by measuring each electrical equipment using sensors separately or Intrusive Load Monitoring (ILM) or through the measurement of the overall electrical equipment using a single sensor in the aggregately or non-intrusive load monitoring (NILM).

NILM approach is more efficient because it uses relatively little sensor even using a single sensor. Several studies have been developed regarding the NILM approach [1], [2]. In this approach occurs process of disaggregation to decompose power that measured in aggregate. Disaggregation process be a challenge of this study because of the type of appliances that had non-linear load so the need of development in the selection or feature extraction and appropriate load signature and development on disaggregation algorithms [3], [4]. Through features selection and disaggregation algorithms appropriate, the NILM system can identify a wide range of power that can be used for the calculation of electric power demand, electric power consumption, power scheduling, type of appliances, time of use, overload capacity of load, etc.

NILM analyzes the composite of current and voltage signals measured using current and voltage sensors on the electrical panel so that they can be disaggregated by their energy usage. Research on NILM first performed in 1992 by George Hart [5]. Hart describes the NILM steps, which is measuring aggregate...
electric signal; detecting an event from captured electric data; clustering similar event using feature vector; matching event; recognizing and monitoring each appliance. Event detection is used to accurately detect the switching of appliances state using an event or non-event based methods. Clustering and matching step to determine appliance caused this switch then recognizing to track its electric energy consumption. This paper focuses on the development of event detection methods. The organization of the paper is as follows: First, background and existing methods. Related Works in Section 2. The proposed NILM event detection method is discussed in Section 3 and Experimental results and discussion are shown in Section 4. Some conclusions are discussed in Section 5.

2. RELATED WORKS

Some event-based method using algorithm for real-time event detection such as change point detection [6], Generalized Likelihood Ratio (GLR) [7], the chi-square goodness-of-fit test (X2 GOF) [8], the Cumulative SUM (CUSUM) filtering [9], and cepstrum analysis [10]-[13]. Change point detection to detect abrupt changes in time series data and to identify the specific time instance when the change occurs. In GLR approach, a decision statistic was calculated by the natural log of a ratio of probability distribution before and after the potential change in mean. This approach requires some parameter to be trained, such as the moving windows, the power data and a threshold for detection statistic. The X2 GOF test to decide the detection decision threshold which depends on the data windows size. This method determines decision between two hypotheses. If the null hypothesis is rejected then assumed that an appliance event occurs. The X2 describes decision threshold that depends on windows size and detection confidence level. CUSUM determine changes of transients in sequence data in the average and provides criteria that help make decisions. Its detection both beginning and the end of the transient is based on a simple rule. Since stop rule occurs in the steady state then the beginning of a transient happens and the stop rule occurs in the transient state then the end of a transient occurs. The previous method investigated the power signal in the time domain. Moreover, an analysis can be performed in the frequency domain, e.g., cepstrum analysis. Cepstrum analysis is used to derive the result of taking the Fourier Transform of the log spectrum. The method will produce the cepstrum parameter as one of the characteristics of the signal. To get the value of cepstrum it must pass some sequence block previous diagram that is from the process of FFT which produce spectrum must at first inverse to change electrical signal from frequency domain become time domain, and cepstrum value that results from spectrum inverse value process.

A cepstrum-smoothing-based method is proposed to improve cepstrum performance by considering simultaneous ON/OFF transitions [12]. This method capable of detecting ON/OFF appliances and can be used to handle the similar peak size of characteristic signals from different appliances. The smoothed cepstrum extract the useful features from the current signal generated by a home appliance. This paper focuses on the development of event detection methods. The method is adopted from the cepstrum analysis technique [12] to extract useful feature using smoothing frequency component. However, cepstrum estimator based on the periodogram suffers from large variance and this will cause large estimation errors in the cepstrum coefficients. The contribution of our paper concern on developing a robust smoothing approach based on local linear regression [14], [15] to reduce variance and improve load event identification.

3. LOAD EVENT DETECTION

Cepstrum method used to extract the characteristics of electrical signal on home appliances. Load signals of appliances simulated and measured through the simulator tool. Current signal data from appliances are used to extract its features. Some noise in the Current signal will affect the performance of the event detection result, so it needs to be removed. This chapter discusses the load modeling of home appliances, noise removal processes, detection process of changes in the Current amplitudes and the process of signal extraction using the Cepstrum method.

3.1. Home appliance load modeling

Residential electrical network and some loads of appliances were simulated using MATLAB Simulink model. This model consists of three phases of the power source, three-phase measurement and home appliance. Each phase connected to each home's appliance, ie Phase A to home 1, Phase B to home 2 and Phase C to home 3 as shown in the following Figure 1. Appliances in each house amount four appliances with Breaker Control to turn on/off their operation.

Appliance in each house was simulated non-linearly along with THD analysis. Some non-linear electrical appliances are modeled by refers to the [16]-[18] model as follows Personal Computer (PC),

Residential load event detection in NILM using robust cepstrum smoothing based method (Nur Iksan)
Fluorescents Lamp (FL) and Television (TV). Figure 2 shows examples of waveforms of current and voltage and THD analysis of 47.5% of TV.

3.2. Noise reduction

Electrical signals are obtained from appliances consisting of Voltage and Current and noise that has not been completely erased. This electrical signal is sensitive to white noise due to the small magnitude of the Voltage and Current. This research uses Median Filter method to reduce noise on electrical signal. The filtered signal ($Y_i$) can be obtained using the following formula

$$Y_i = \text{median}(Y(i-m/2)+Y(i-m/2+1)+\cdots+Y(i-m/2-1)+X(i-m/2))$$

(1)
3.3. Change point detection

This event detection process based on changes in the Amplitude of the Current signal. The least square method is used to detect such changes, ie.

$$J_n (\tau, \gamma) = \frac{1}{\alpha} \sum_{(k=1)}^{K} (i-(k-1)^{+1}-(\tau_k)^{Y_i-Y^-k})^2$$ (2)

Figure 3 shows the process of event detection through change point current amplitude in noisy current signal and filtered current signal. In Figure 3(a) the wrong noise is detected as an event. And in Figure 3(b) shows the current trace amplitude after processed by the median filter.

![Figure 3. Point of change detection on current waveform (a) Noised current waveform. (b) Filtered current waveform with m=30](image)

3.4. Robust cepstrum-smoothing

The electrical signal extraction process was performed to obtain the characteristic value from Current signal. The value of this feature to be used as a reference to the new electrical signal to be tested so can produce a decision of the results of the comparison of those values. The signal characteristic extraction process was performed by cepstrum method which will produce the characteristic value of quefrency cepstrum. The application of cepstrum to the signal extraction process can be seen in the block diagram of Figure 4.

![Figure 4. Signal extraction process](image)
function used in this research is window hamming because the hamming function can create data at the first
and the last points in the frame close to zero well. Each frame signal has to be multiplied by the window
function. The value of n in the window process was adjusted by the number of samples used in the framing
process. Hamming window defined by (3), then the signal after Hamming windowing was multiply between
frame signal and the Hamming window, defined by (5).

\[ w(n) = 0.54 - 0.56 \cos \left( \frac{2\pi \phi \cdot n}{n-1} \right), \quad 0 \leq n \leq (n-1) \]  
\[ x(n) = x(n) \times w(n) \quad 0 \leq n \leq (n-1) \]  
\[ x(n) = \begin{bmatrix} x_1 \times 1 & \cdots & x_1 \times n \\ x_2 \times 1 & \cdots & x_2 \times n \\ \vdots & \ddots & \vdots \\ x_n \times 1 & \cdots & x_n \times n \end{bmatrix} \times \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} \]  

Fourier analysis allows for analysis of the spectral properties (spectrogram) of the input signal by
elaborating the signals into the sine wave with different frequency. The purpose of using Fourier analysis on
electrical signal data was to change data from the time domain into spectrum data in frequency domain. The
Current signals were transformed by using the FFT and their spectrum is obtained by (6). The result of the
FFT cannot show the form of the spectrum so that logarithm of the absolute value was used. The coefficient
cepstrum is obtained by transforming this information from the frequency domain into the quefrency or
cepstrum domain. These can be computed by applying the inverse Fourier transform to the logarithm of the
absolute value of the FFT of signal \[ x \].

\[ X[n] = \sum_{n=0}^{N} x(n)e^{-j\frac{2\pi}{N}kn}, \quad 0 \leq k < N \]  
\[ Y(n) = \frac{1}{N} \sum_{k=0}^{N-1} \log |x_m[k]|e^{j\frac{2\pi}{N}kn}, \quad 0 \leq n < N \]  

After obtained cepstrum coefficient, the weighting process was then performed to obtain smoothed
cepstrum value. The value of the cepstrum coefficient resulting from the smoothing process then used as the
feature vector value. This step called as cepstrum-smoothing and furthermore can be defined as (8). This
value can be taken several points that can represent all data and used as a feature that represents each frame.
The next smoothing process is done by using local linear regression method. It is based on locally fitting a
line and calculated by performing a weighted least squares regression. In the weighting process, the
appropriate weight is determined by the distance between the data points and the smoothing points. The data
points that have the greatest weight and the most impact on the fit are used as data points to be smooth.
Weighting is performed by using the following tricube function [14].

\[ W_i = \left( 1 - \frac{x-x_i}{d(x)} \right)^{3} \]  

The regression process performed using 2nd degree polynomial model is known as the 'loess'
method. The robust version of the loess method is used to reduce the 'loess' sensitivity to the outlier. This
robust version is known as the 'rloess' method. This method includes also calculating the weighting using the
bisquare (9) function. The result of robust smoothing is calculated using both the local regression weight and
the robust weight.

\[ W_i = \begin{cases} 
(1 - \left( \frac{r_i}{6MAD} \right)^2, & |r_i| < 6MAD, \\
0, & |r_i| \geq 6MAD,
\end{cases} \]  

Where,
ri : the residuals of the data points generated by the regression smoothing process
MAD : median absolute deviation of the residuals or median(|r|)
4. EXPERIMENT AND RESULT

Electrical signal characteristics are the most important part of the event detection process. The electrical signal characteristics then processed according to the steps shown in Figure 3. Based on the steps, in the framing process, sampling is done every 20ms, and the sampling frequency used is 20 KHz, while the record length is 24 seconds. The number of frames generated is 98 frames per second and 2376 frames for 24 seconds and each has 400 sample data per frame. The window process is used to reduce discontinuous effects at the ends of the frame by using hamming methods. The value of N in the window process is adjusted by the number of samples used in the framing process, which are 400. The FFT process uses 512 points greater than the number of N in the window process. Because the signal FFT result is symmetrical then the data used only half, which is 256 points. The result of the FFT cannot show the shape of the signal spectrum so that logarithm is used. Coefficient cepstrum is obtained by converting the value of the FFT result to frequency domain into time domain through Inverse of FFT process. Simulation dashboard as shown in Figure 5.

The shape of signal transformation at each stage of the cepstrum process can be seen in Figure 6. The amplitude at peak frequency has different values on the spectrum of each appliance. This difference will be the characteristic of each appliance. The peak frequency value of the result of the cepstrum process on the aggregate signal can vary and even the same. The cepstrum method can be used to reduce the peak home signal variance in the FFT spectrum. Furthermore, cepstrum-smoothing can be used to handle peak size signals of the same characteristics of different equipment. Below are the frequency forms of home electrical equipment. For example, TV, PC and TV aggregation and PC.

The cepstrum smoothing method can be used to reduce the variance of the peak signal on the original cepstrum. Furthermore, cepstrum smoothing can be used to overcome the great similarity of the peaks of signal characteristics of different electrical appliances by reducing the characteristics at low frequencies. The next smoothing process is done by using Local Linear Regression method. This process is used to reduce the increase in bias and variance of peak signals resulting from cepstrum smoothing. A robust version of ‘loess’ is used as a method of smoothing. This method assigns the lower weight to outliers in the regression. The method assigns zero weight to data outside mean absolute deviations.

Figure 7 shows the signal smoothing obtained from both of smoothing method of cepstrum and rloess. Signal smoothing of some appliances includes PC equipment (a); TV (b); aggregate PC and TV (c). The cepstrum smoothing signal Y(n) The cepstrum smoothing signal Y(n) is indicated by dot blue and rloess smoothing is indicated by the red line. We determine the span as a percentage of the count points in the data set using 0.2 which is means that a span of 0.2 uses 20% of the count data points in data set. The number of data points cepstrum signal Y(n) will keep the same and used in the local fit when processed by loess method and determined through the span parameter. The number of data point in this simulation are 512 and
span=0.2. Loess will use the 0.2 * 512 nearest data point to signal Y(n) for fitting. Figure 6 shows the result of smoothing using cepstrum and rloes method to determine the feature set then this smoothing signal is extracted through the peak detection process to get the peak of frequency amplitude of the smoothing signal. The peak detection process is used to find the local maxima and minima. The peak threshold of 0.1 is used to determine the minimal difference between the peak and its surroundings and then to declare it as the peak. The peak of frequency amplitude generated by cepstrum smoothing signal has more peak variance than in rloes smoothing signal. More peak variance will increase the complexity of the device features. As it was previously known, the use of this rloes method to reduce the peak of frequency amplitude variance of the cepstrum smoothing results. Therefore, the result of robust smoothing produces peak signals that have relatively less variation and certainly impact on the reduced computational complexity.

![Figure 6](image)

**Figure 6.** Cepstrum signal on single/multiple appliances. The result of peak frequency and amplitude using cepstrum smoothing (right) and simple cepstrum (left) method (a) PC; (b) TV; (c) signal aggregated of PC and TV.
Residential load event detection in NILM using robust cepstrum smoothing based method (Nur Iksan)

Figure 7. Rloess smoothing signal on single/multiple appliances. Signal smoothing using cepstrum smoothing (dot blue) and rloess smoothing (red line) method (a) PC; (b) TV; (c) signal aggregated of PC and TV.

Figure 8 shows the result of smoothing using rloes method on TV equipment. TV operation has several states, which consists of OFF to ON; ON; ON to OFF; and OFF. To detect the occurrence of the state then determined the peak amplitude of the rloes signal. The point of the peak amplitude in each state is different in number and peak point. The number of peak points on the OFF condition is relatively small and does not even exist on certain equipment. This is because the amplitude of the current wave signal is zero. In TV equipment, when the OFF condition has only one peak and will have more peaks when the state transitions ON. In state ON and state ON to OFF, sum and peak points are considered equal.

Figure 9 shows the variance values of the peak amplitude of electrical equipment. The value of this variance indicates a measure of the spread of data in the data distribution and is used as the reference of the
accuracy of the model representing the true value. The higher of variance value then more varied the data and more inaccurate. To calculate the variance, we must first know the mean, and then add the squares of the difference of each data to the mean. Numerically, the variance is the mean of the square of the difference in data to the mean. In the electrical equipment identification process, a low variance value of peak amplitude can help improve the accuracy of identification. The cepstrum smoothing method is good enough at reducing the variance of the peak signal compared when using the FFT spectrum. Furthermore, the use of robust smoothing can improve the ability to reduce the value of variance so it has a lower variance value than using the method of cepstrum smoothing. To test the performance of signal extraction the feature set obtained from the peak amplitude frequency value is used in the classification process. Peak amplitude value of each equipment has the different event time with the number of events is 2. At peak amplitude of aggregate signal there is 3 appliances with time event that varies and based on the signal measurement there are 7 events. This event will then be used to group the vector feature by using k-means method with K is 7. After obtaining the group of vector feature, then the process of classification using Naïve Bayes method with cross-validation is 0.3. The following Figure 10 shows the performance value of the vector feature using the peak amplitude of the Current signal. Figure 10 shows the comparison of the classification performance of the cepstrum smoothing method and the proposed method. The classification of the proposed method shows the better result than the cepstrum smoothing either on single appliance or multiple appliances.

![Figure 9. Peak and min amplitude after smoothed by robust smoothing](image)

![Figure 10. Peak amplitude variance of appliance and precision prediction of cepstrum smoothing and proposed method](image)
Noise added in the simulation is a Gaussian noise that is used for scenarios of sensitivity testing of load event detection against noise. The noise scenario is done by setting the noise ratio of SNR 5; 10, 15, 20; 25; 30; and 50. Figure 11 shows the comparison of the sensitivity of event load detection to noise. At the time of running the simulation, the first training of data on the dataset of single and multiple appliances with SNR of 25. Data trained serve as a model to predict the data generated when running the simulation by using some SNR scenario. Prediction precision in single appliance shows varying results in each scenario. In SNR >20, the precision prediction is quite good with values greater than 0.6 and is relatively stable at values above 0.9 on SNR >25. However, the prediction precision decreases below 0.5 on SNR <20 due to the large noise ratio of the signal. Similarly in multiple appliances, precision prediction is quite good at SNR >20 and begins to decrease in SNR <20 and SNR >25.

![Figure 11. Precision prediction on noise scenario](image)

5. CONCLUSION

In this paper, a robust smoothing method for cepstrum estimation based on local linear regression using rloes method, is presented. This method is capable of producing a better smoothing method than cepstrum smoothing. This method can reduce the variance of peak electrical equipment signals and impact on the reduced computational complexity. The results of this estimation cepstrum are then used in the event identification process of the electrical equipment in operation. In performance testing, the classification results show the proposed method has a better accuracy than the previous method.

ACKNOWLEDGEMENTS

We would like to thank to Blackberry Innovation Center Institute of Technology Bandung (BBIC-ITB); Semarang State University (UNNES); and Ministry of Research, Technology and Higher Education of the Republic of Indonesia (KEMRISTEKDIKTI) for their assistance in facilitating education and research.

REFERENCES


BIographies of Authors

**Nur Iksan** was born in Indonesia in 1983. He received Bachelor’s degree in 2005 from Universitas Negeri Semarang (UNNES), Indonesia, and M.Kom. degrees in Informatic Engineering from Universitas Dian Nuswantoro (UDINUS), Indonesia, in 2010. He is a lecturer at Engineering faculty of UNNES, Indonesia. His main interest is on the application of stochastic system in information technology related field. He is an active member of IEEE as student member.

**Jaka Sembiring** learned the Bachelor’s degree in 1990 from Institut Teknologi Bandung (ITB), Indonesia, Master’s degree in 1997, and Doctoral degree in 2000, from Waseda University, Japan, all in electrical engineering. He is now an associate professor at the School of Electrical Engineering and Informatics (STEI), ITB where he joins the Information Technology Research Division. His main interest is on the application of stochastic system in information technology related field. His research interests include probability, stochastic, information system and its application in enterprise architecture, risk analysis, robotics and agent systems. Dr. Sembiring is an active member of IEEE and ACM.

**Nanang Hariyanto** was born in Jember, Jawa Timur, Indonesia, on July 26, 1962. He graduated from the Electrical Engineering Department, Bandung Institute of Technology. His employment experience included Synatech Industry, and as a researcher at Power System Laboratory, ITB Indonesia. His special fields of interest included power system planning and dynamic stability.

**Suhono Harso Supangkat** a Professor in the School of Electrical Engineering and Informatics Engineering at the Institut Teknologi Bandung, West Java, Indonesia. He received a master’s degree from Meisei University, Japan, and a doctoral degree from the University of Electro-Communications, Tokyo, Japan, in 1998. His research interests include smart cities, smart education, smart health, smart energy, and smart mobility.