A novel CUSUM-based approach for event detection in smart metering

Zhicheng Zhu, Shuai Zhang, Zhiqiang Wei, Bo Yin* and Xianqing Huang

College of Information Science and Engineering, Ocean University of China, Qingdao, China

*Corresponding author e-mail: ybfirst@ouc.edu.cn

Abstract. Non-intrusive load monitoring (NILM) plays such a significant role in raising consumer awareness on household electricity use to reduce overall energy consumption in the society. With regard to monitoring low power load, many researchers have introduced CUSUM into the NILM system, since the traditional event detection method is not as effective as expected. Due to the fact that the original CUSUM faces limitations given the small shift is below threshold, we therefore improve the test statistic which allows permissible deviation to gradually rise as the data size increases. This paper proposes a novel event detection and corresponding criterion that could be used in NILM systems to recognize transient states and to help the labelling task. Its performance has been tested in a real scenario where eight different appliances are connected to main line of electric power.

1. Introduction
Non-Intrusive Load Monitoring (NILM) was proposed by Hart and other scientists 25 years ago when he first put forward an algorithm that was used to decompose household energy consumption [1]. This research area covers three major aims: first, it is to inform residents how much electricity is consumed by each appliance, helping them reduce energy cost. Second, it is to provide personalized feedback by calculating the amount of electricity saved by certain appliance. For example, consumers would be clearly aware of how much energy they have saved by using new energy efficient appliance. Last, if the NILM system could access utility time of each appliance, a system would be built for advising consumers to use certain appliances during off-peak hours [2]. Thus, not only does it allow people to reduce their electricity cost, but it helps avoid problems caused by heavy demand during peak hours.

2. Previous work
The review of event detection can be found in [3] which introduces three different approaches including expert heuristics, probabilistic models, and matched filters. The heuristic method features step changes on the basis of power consumption and those changes are rather fixed and take place between two stable states. Compared to the probabilistic model, this method’s advantage is that a smaller amount of data is required being collected. [4] adopts generalized likelihood ratio (GLR), and matched filters require a higher rate of data acquisition to extract waveform and then connect it with the known signal. [5] puts forward a method combining the average filter and derivative filter to recognize events. In [6], Hsueh-Hsien Chang proposes transient feature analysis, focusing on the transient response time and transient energy among power characteristics. Wavelet transform in time-
frequency domain is therefore used to analyze and detect the transient performance of load, and the experiment results prove this method to be more effective than the one based on steady state.

3. Improved CUSUM algorithm

3.1. CUSUM

Suppose a power’s time series is \( X = \{x_n\}_{n=1}^{\infty} \), we adopt sliding window to conduct continuous observation, with the window size being set as \( m \). First, we calculate average \( K_m \) by using (1).

\[
K_m = \frac{\sum_{n=1}^{m} x_n}{m}
\]

The following recurrence formula (2) defines test statistic \( Z_n \).

\[
Z_n = \max(0, Z_{n-1} + x_n - K_m), \ Z_0 = 0
\]

Take load switch on and off event as an example, when the load of power line is stable without any appliance being added or removed, the CUSUM test statistic \( Z_n \) is a variable that fluctuates randomly around 0, that is, a random variable with its average being 0. When the working state of current load changes making power consumption increase, \( Z_n \) begins its positive shift and accumulates continuously, and vice versa. According to the preset threshold \( H \), we can tell whether switch on and off events happen. Given \( 0 < Z_n < H \) which means the load event might happen but doesn’t reach the threshold level, \( Z_n \) is involved in iterative computation of \( Z_{n+1} \). Until the positive or negative shift meet the criterion, CUSUM identifies the switch on and off events and then we could reversely infer the time when events happened.

However, the recurrence formula shows that the algorithm has accumulative trait, which means the statistic test of the previous point is involved in the calculation of the next point. Consequently, the subsequent statistic tests would all be affected. As the \( n \) value keeps increasing, the algorithm itself becomes more and more sensitive to the fluctuations that make it more likely to cause false detection. To tackle this problem, we attempt to improve the original CUSUM algorithm, making permissible deviation increase gradually as \( n \) goes up. In this way, the detection accuracy and the robustness under complicated circumstance are greatly enhance.

3.2. Improved CUSUM algorithm

To address the false detection problem resulted from a large amount of data that CUSUM needs to process, we improve the test statistic \( Z_n \) by coming up with a new approach named PCUSUM, the “variable sum” CUSUM. First, the average \( K_m \) and variance \( D_m \) of power data in the window need to be calculated by using (3).

\[
D_m = \frac{\sum_{n=1}^{m} (x_n - K_m)^2}{m}
\]

Next, define the new test statistic \( P_n \) and its recurrence formula is shown as (4).

\[
P_n = \max(0, P_{n-1} + x_n - (\sqrt{n - k} - \sqrt{n - k - 1}) \cdot D_m), \ P_0 = 0
\]

Where \( k \) value can be gained from formula (5).

\[
k = \max(i: i < n, P_i = 0)
\]

For the new statistic test \( P_n \), if there were no noticeable deviation in time series of power, \( P_n \) would stay stable. If the switch on or off event occurred, power would change and \( P_n \) would start to accumulate positively or negatively, which is consistent with original CUSUM. However, what is
different is that provided there’s a large amount of data, the increased $n$ value would cause $n-k$ to increase accordingly, leading to less effect produced by $(\sqrt{n-k} - \sqrt{n-k-1}) \ast D_m$. As a result, the statistic test only sees gentle variations so that false detection rate would be reduced. When the power change is finally eliminated, $P_n$ stands stable at the current position. Take the switch on and off events of 280W water fountain for example, the two figures are drawn together to better illustrate how the improved CUSUM respond to power variations. As show in Fig. 1, the blue line represents power waveform while the red line represents PCUSUM chart.

Figure 1. Power waveform vs PCUSUM chart.

3.3. Determination criterion
Due to the fact that improved CUSUM is sensitive to all changes, big or small, the traditional determination criterion is not suitable for switch on and off events detection in NILM system. We therefore introduce a new criteria based on average slope rate difference. Seeing From the waveform of PCUSUM chart in Fig. 1, we can conclude that when a new appliance starts to work, the increased power causes $P_n$ to accumulate positively, whilst the slope rate of waveform begins to rise from 0. By contrast, when the appliance is off, the slope rate returns to around 0. Considering this trait, we make an attempt to use slope rate difference $Rate_n$ as determination criterion, and it is calculated by (6).

$$Rate_n = \frac{P_n - P_{n-1}}{\Delta x}$$  \hspace{1cm} (6)

To speed up the detection of switch on and off events, we divide slope rate into each section with length of $m$, and then calculate the average slope rate difference $Ra$ of every section by using (7).

$$Ra = \frac{\Sigma_{m+1}^{n} Rate_n}{m}$$  \hspace{1cm} (7)

Using the above-mentioned water fountain example again, we record power waveform and average slope rate difference together in the Fig. 2.
It is obvious that the average slope rate difference changes are consistent with power variations in the whole process of switch on and off events. Thus, we realize the accurate event detection of home appliances by means of defining new statistic test and determination criterion.

4. Experiment
To verify the accuracy and robustness of this event detection method, we simulate the real environment of domestic appliance, collecting data from and experimenting on the eight different home appliances, including electric radiator, electric incandescent lamp, kettle, television, fan, microwave oven, oxygen generator, and water fountain, with rated power ranging from 60W to 2400.

4.1. A test case 1
First, we conduct experiments on the switch on and off events of four different appliances respectively. The sample rate is set to 10 kHz for each type of devices, and keep them being switched on and off 100 times repeatedly under the same circumstance.

To collate the two methods, we use the original CUSUM and the newly proposed method based on PCUSUM to process experiment data respectively on the Matlab platform. The experiment results are shown in Table 1.

| Algorithm        | Appliance                | $N_{all}$ | $N_m$ | $N_f$ | $\gamma$  |
|------------------|-------------------------|-----------|-------|-------|-----------|
| Original CUSUM   | Electric incandescent lamp | 200       | 20    | 17    | 82.9%     |
|                  | Electric radiator       | 200       | 1     | 0     | 99.5%     |
|                  | Kettle                  | 200       | 3     | 1     | 98.01%    |
|                  | Television              | 200       | 8     | 11    | 91.00%    |
| Improved CUSUM   | Electric incandescent lamp | 200       | 8     | 9     | 91.87%    |
|                  | Electric radiator       | 200       | 0     | 0     | 100%      |
|                  | Kettle                  | 200       | 2     | 0     | 99.00%    |
|                  | Television              | 200       | 5     | 2     | 96.53%    |

In Table 1, $N_{all}$ represents the total number of switch on and off events of each appliance, and $N_m$ is used to denote the number of missed detection events, and $N_f$ refers to the number of false detection events. The detection accuracy $\gamma$ therefore can be calculated according to formula (8).
\[ \gamma = \frac{N_{all} - N_m}{N_{all} + N_f} \] (8)

The experiment results clearly show that the improved CUSUM achieves overall higher accuracy in identifying various appliances than the original CUSUM. This new method excels particularly in terms of low power appliances with no resistance, with even higher detection accuracy and lower false detection rate. The main reason behind this effective new method is that PCUSUM is very sensitive to tiny deviation and it can successfully avoid exceeding the threshold to cause false warning when the amount of data is increasing.

4.2. A test case 2

In this part, we examine the proposed method’s performance under the circumstance of multiple appliances. Five common appliances are selected to conduct switch on and off operations, and Table 2 elaborates specific steps.

| ID | Operation description                  |
|----|----------------------------------------|
| 1  | Fan went to On                          |
| 2  | Water fountain went to On               |
| 3  | Electric radiator went to On            |
| 4  | Microwave oven went to On              |
| 5  | Oxygen generator went to On             |
| 6  | Oxygen generator went to Off            |
| 7  | Microwave oven went to Off              |
| 8  | Electric radiator went to Off           |
| 9  | Water fountain went to Off              |
| 10 | Fan went to Off                         |

The power waveform and average slope rate difference of five different appliances are illustrated in Fig. 3.

![Power waveform and corresponding average slope rate difference of five appliances.](image)

According to Fig. 3, we can see that average slope rate difference responds to switch on and off events both accurately and promptly. The bigger power variation is, more noticeable changes of
average slope rate difference. To better illustrate experiment results, we repeat the operation by the same order for 10 times, with the total number of switch events reaching 100 times. At the same time, two different algorithms are used to work out detection accuracy rates respectively, and the experiment results are shown in Table 3.

Table 3. The detection probability of two methods under the circumstance of multiple appliances.

| Algorithm | Original CUSUM | Improved CUSUM |
|-----------|----------------|----------------|
| $N_{all}$ | 100            | 100            |
| $N_{m}$   | 7              | 4              |
| $N_f$     | 9              | 0              |
| $\gamma$  | 85.32%         | 96.00%         |

After analyzing experiment results, we find that there truly exists the interference caused by increasing noise of power line if multiple appliances are simultaneously working. Even though that noise gives rise to the declining detection accuracy of both algorithms to some extent, the proposed method still manages to achieve ideal recognition performance. Its overall detection accuracy reaches 96%, much higher than that of the original CUSUM. Moreover, this new method is robust against false detection, as $N_f$ keeps at 0.

5. Conclusion
This paper proposes an approach of non-intrusive appliance load monitoring for event detection, which further improves statistic test of the original CUSUM. Combining with average slope rate difference as the determination criterion, this method is developed to recognize switch on and off events of household appliances of various kinds. This approach sees more than 90% of recognition rate among a range of domestic devices, even under the circumstance where multiple appliances are working simultaneously.

Acknowledgments
This work was supported in part by China International Scientific and Technological Cooperation Special (2015DFR10490), Qingdao national laboratory for marine science and technology Aoshan science and technology innovation project (2016ASKJ07-4) and Qingdao innovation and entrepreneurship leading talent project (13-cx-2).

References
[1] G.W. Hart, “Nonintrusive appliance load monitoring,” Proceeding of the IEEE, vol. 80, no. 12, pp. 1870-1891, 1992.
[2] T. Liang, C. Gao and B. Wang, “Applications of demand side management in smart grid,” Electric Power Automation Equipment, vol. 32, no.5, pp. 81-85, 2012.
[3] K.D. Anderson et al., “Event detection for Non Intrusive load monitoring,” in 38th Annual conf. on IEEE Industrial Electronics Society, 2012.
[4] D. Luo et al., “Monitoring HVAC equipment electrical loads from a centralized location – methods and field test results,” ASHRAE Transaction, vol.108, 2002.
[5] M. Weiss et al., “Leveraging smart meter data to recognize home appliances,” in IEEE Int. Conf. on Pervasive Computing and Communications, 2012.
[6] H.H. Chang et al., “A nonintrusive load identification method for residential applications based on quadratic programming,” Electric Power Systems Research, 133:241-248, 2016.
[7] E.S. Page, “Continuous Inspection Schemes,” Biometrika, 41.1/2:100-115, 1954.