International Journal of Smart Grid and Clean Energy

# Cost effective disaggregation mechanism for the NIALM system

Shin Jae Kang*, Kangsoo Shin, Rex Yang, Seung Yong Lee, Jinho Sohn

*Data Mining Team, Future IT R&D Lab., LG Electronics, 38, Baumoe-ro, Seocho-gu, Seoul 137-724, Korea*

**Abstract**

Non-intrusive appliance load monitoring (NIALM) technology enables to detect the amount of power usage of each home appliance with single point sensing. In order to improve the performance (i.e., disaggregation/classification accuracy), features such as real power and harmonic components of electrical current of appliances are considered. However, taking more features into an account for better performance also increases the total cost of NIALM-related system. In this paper, a cost effective approach for NIALM is presented. With measured real power data sampled at low speed (about 1~2 Hz) from a single power line, classification of home appliances with high accuracy (about 80%) is achievable based on the proposed NIALM mechanism. The proposed mechanism is based on the real power time window pattern during a transient state and the real power level in a steady state after transition. The performance of the proposed mechanism is evaluated through simulation; classification accuracy is improved by 28% compared to the conventional NIALM approach based on real power delta scheme and 17% compared to the time window pattern.

*Keywords: NIALM, home appliance, classification accuracy, disaggregation, real power time window pattern, real power level, steady state, real power delta value*

## 1. Introduction

In the power and energy system areas, there have been many efforts in establishing efficient energy management in order to resolve problems incurred from the energy shortage issues. The concept of smart grid is one of the solutions to the rising problem. In the concept of smart grid, information of every grid component should be accessible, and all the components comprising the power grid should support two-way communication [1]. Based on available information on the grid, utility companies are able to schedule power transmission and provide information services (e.g., energy usage information in a particular household) to customers in a residential area. Recognizing current energy status or information provides benefits not only to the utility companies but to the customers as well; study [2] showed that residents in the United States are able to save energy from 10 to 15 percent just by knowing how much they are using it.

Methods to monitor the power usage of a customer from a power line can be categorized into three: distributed direct sensing, single point sensing, and intermediate sensing [3]. The distributed direct sensing method requires a sensing device (such as smart plug) for each home appliance. The monitored data from this method is very accurate. However, the cost of installation and maintenance for the distributed direct sensing method is quite expensive (at least USD 150). Apart from the distributed direct sensing method, the single point sensing method requires only one sensing device for monitoring power usage of every home appliance. The accuracy of this method is known to be relatively low compared to the distributed direct sensing method, but the expected cost of installation and maintenance is much lower than the previous method. This encourages the electrical power industry to pay a huge attention to the single point sensing method. The method is well known as non-intrusive appliance load monitoring

* Manuscript received May 9, 2013; revised July 26, 2013.
Corresponding author Tel.: +82-10-7202-3678; *E-mail address*: shinjae.kang@lge.com.

(NIALM). With NIALM system, successful disaggregation of power usage of each appliance from a signal can en achieved with a single point sensing device. There are two major factors, such as accuracy and cost, which are carefully considered and analyzed in NIALM research due to the fact that there is a tradeoff between the accuracy and cost. The powerful benefit of NIALM is cost effective, but the distributed direct sensing method is more accurate than NIALM. Thus, there have been tremendous efforts on improving the accuracy of NIALM [4]-[7]. In the early stage of NIALM research, real power and reactive power values are used for disaggregating power usage of each appliance [4]. In order to improve accuracy, Reference [5] introduced another factor, monitoring of power transient. Reference [6] utilized pattern matching of electrical current harmonic components. Additionally, the frequency analysis of data, sampled at high frequency (~1 MHz), is considered for monitoring voltage noise generated by SMPS (switching mode power supply) of certain types of appliances in [7]. Lastly, the intermediate sensing method, is somewhere between direct and single point sensing.

In NIALM system, the total cost of a system increases as the sampling rate of data increases. This is mainly due to the fact that the hardware cost for the high speed sampling is high, and the overall system complexity for the large-scale data processing with high speed is also high. In a certain case where the data sampling rate is extremely high, the cost for NIALM could be larger than the distributed direct sensing method. Therefore, for the realization and practical use of NIALM system, minimizing the total cost with relatively high and acceptable accuracy is critical.

In order to achieve low cost system with high accuracy, the proposed framework considers both features of real power time window pattern and real power level in the steady state. Those two considerations are formulated as an optimization problem. The main contribution of this paper is summarized as follows. The proposed disaggregation/classification method delivers high accuracy (i) by considering the real power time window pattern if an error in the measured data. In addition, (ii) considering the real power level in the steady state resolves the irregular time window pattern caused by multiple appliances operating simultaneously.

The remainder of this paper is organized as follows. In Section 2, an overview of a conventional NIALM approach and its limits are described. In Section 3, the cost effective algorithm for NIALM is proposed. In Section 4, the performance evaluation for the proposed framework is discussed based on simulations which consist of a power meter and seven different appliances, and the conclusion is drawn in Section 5.

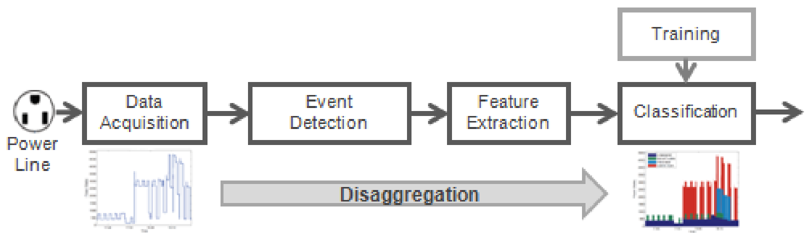


Fig. 1. Block diagram of the NIALM system

## 2. Conventional Real/Reactive Power Approach and Its Limits

### *2.1. Brief overview of NIALM*

Apart from the distributed direct sensing, NIALM requires a single point measurement of an electrical signal. A measured analog signal is then digitized and goes through the event detection. An event detector with certain criteria estimates whether appliances has turned on, off, or changed its operating mode. When the system recognizes an event, it extracts features of appliances from the observed signal. Rather than using raw data, using extracted feature for the remaining process helps to reduce the complexity of computation. Based on the extracted features and trained classifier, the system disaggregates the power usage of each appliance from the monitored signal which contains the power usage of every appliance that are currently operating. There are several methods such as MIP (mixed integer programming) and

FHMM (factorial hidden Markov model) for disaggregating the power usage. Fig. 1 describes the overall structure of the NIALM system.

Pioneering work in NIALM research area was first introduced by George Hart in the 1980s [4]. In the study, the electrical power consumed by each appliance is disaggregated through a single sensor based on power delta value (e.g., the amount of power level change when a certain appliance turns on or off); he especially noted the real power delta and the reactive power delta values at a certain time instance. This approach is based on the assumption that the real and reactive power delta values are different with respect to different types of operating appliances. These features from each appliance can be represented as a P-Q map (real power on the *x*-axis and reactive power on the *y*-axis). Fig. 2 shows one of the examples of P-Q maps based on observed data.

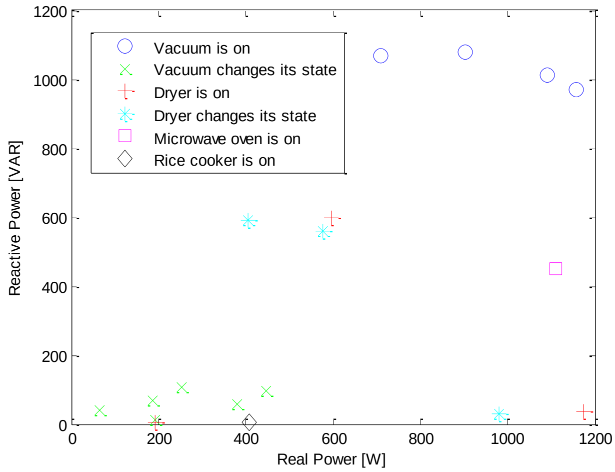


Fig. 2. A P-Q map example; all possible P-Q points of four appliances (vacuum, dryer, microwave oven, rice cooker) are plotted.

The reason for considering both real and reactive power is that these two features increase the diversity of the characteristic. For example, vacuum, microwave oven, and dryer have the similar power level about 1,150 W on the *x*-axis (real power level). On the other hand, those appliances have different reactive power level on the *y*-axis; the reactive power of dryer is about 50 VAR, that of the microwave oven is about 430 VAR, and that of the vacuum is about 1000 VAR. These huge differences on the reactive power can help the NIALM system to disaggregate appliances correctly.

### *2.2. Limits of the conventional method*

Even though the conventional NIALM system considers two aspects of appliances, those two features are not enough to disaggregate each appliance from a single monitored signal without introducing some errors. This problem is well described in the Fig. 2. Three points are crowded around the 200 W on the *x*-axis; especially, the points indicating 'Vacuum status changes' and 'Dryer is on' are close to each other. Thus, if the observed value is slightly changed by dynamics of sensing environment, the system could detect the event as 'Dryer is on' even though it is actually 'Vacuum changes its status' in reality. In addition, measuring the reactive power directly or computing it from the measured real power and power factor causes the cost and complexity of NIALM system to increase.

## 3. The Proposed Mechanism

In order to overcome the limits of the conventional method, several new approaches in NIALM have been introduced; Reference [6] considers the harmonic components of electrical current, and Reference [7] introduced the monitoring of voltage noise from appliances equipped with SMPS (switching mode power supply). Considering certain features, such as harmonics and high frequency noise components of appliances in the frequency domain, definitely overcomes the limits of conventional methods in terms of accuracy, but it increases the total cost of the system. In order to analyze up to 10th harmonic components of electrical current, more than 1.2 kHz sampling rate is required. For analyzing voltage noise from SMPS, about 1 MHz sampling rate should be supported. These high speed sampling rate requirements all increase the system cost and complexity since advanced ADC (analog-to-digital converter) is needed. Additionally, large scale data process and network with high speed should be supported as well.

The increased cost and complexity could not be aligned with the motivation of NIALM, which is the cost reduced system for monitoring power consumption of all individual appliances. For this reason, a new cost effective NIALM mechanism is proposed to overcome the limits of the conventional approaches.

$$\begin{aligned}
&\text{minimize} \quad \sum_{j=1}^{J}\left|\sum_{i=1}^{N}\sum_{k=1}^{K_i} x_i^k P_{\mathrm{tw}_i}^{k_j} - P_{\mathrm{tw}}^{j}\right| + \left|\sum_{i=1}^{N}\sum_{k=1}^{K_i} x_i^k P_{\mathrm{steady}i}^{k} - P_{\mathrm{steady}}\right| \\
&\text{subject to} \quad x_i^k = \{0,1\} \quad \text{for all } i \text{ and } k \\
&\qquad\qquad \sum_{k=1}^{K_i} x_i^k = \{0,1\} \quad \text{for all } i
\end{aligned} \tag{1}$$

where, $P_{\mathrm{tw}_i}^{k_j}$ is the $j$th trained real power component of a time window of the $k$th state of the $i$th appliance

$P_{\mathrm{steady}i}^{k}$ is the trained real power edge value of the $k$th state of the $i$th appliance in the steady state

$P_{\mathrm{tw}}^{j}$ is the $j$th real power component of a monitored time window

$P_{\mathrm{steady}}$ is the monitored real power value in the steady state

$J$ is the total number of time window components

$N$ is the total number of appliances, and $K_i$ is the total number of all states of the $i$th appliance

Equation (1) is the proposed classification algorithm for the cost effective NIALM system. The algorithm tries to minimize the L1-norm difference between observed power level and the linear combination of the trained power level of all appliances. The left term of the cost function indicates the L1-norm difference of real power time window during a transient state in which certain appliances could be turned on/off, or changes its status. The right term of the cost function indicates the L1-norm difference of the real power level during a steady state period in which power fluctuation of all working appliances becomes stable.

The first constraint indicates that the formulated optimization problem is MIP (mixed integer programming), and so, each variable represented by $x_i^k$ indicates power on/off or various mode of each appliance. The second constraint represents that all appliances are restricted to have only one state during the observed time period; for example, two states, ‘Vacuum turns on’ and ‘Vacuum changes its state from a strong mode to a weak mode’ could never happen simultaneously.

$$\begin{aligned}
&\text{minimize} \quad \alpha\sum_{j=1}^{J} t_j + \beta t_{J+1} \\
&\text{subject to} \quad -t_j \le \sum_{i=1}^{N}\sum_{k=1}^{K_i} x_i^k P_{\mathrm{tw}_i}^{k_j} - P_{\mathrm{tw}}^{j} \le t_j \\
&\qquad\qquad -t_{J+1} \le \sum_{i=1}^{N}\sum_{k=1}^{K_i} x_i^k P_{\mathrm{steady}i}^{k} - P_{\mathrm{steady}} \le t_{J+1} \\
&\qquad\qquad 0 \le x_i^k \le 1 \quad \text{for all } i \text{ and } k \\
&\qquad\qquad 0 \le \sum_{k=1}^{K_i} x_i^k \le 1 \quad \text{for all } i
\end{aligned} \tag{2}$$

The formulated optimization problem in Eq. (1) is not linear. In order to reduce the complexity, LP relaxation techniques are introduced for reformulating the problem as a linear problem [8]. For the L-1 norm, a new variable $t_j$ is introduced, and for the discrete variables $x_i^k$, the discrete condition is relaxed as the linear condition. The reformulated problem is described in Eq. (2).

## 4. Simulation Results

### *4.1. Simulation scenario*

In this section, the performance of the proposed algorithm is evaluated with simulation. Fig. 3 represents the simulation environment. The power line from the wall outlet is connected to the input part of the power meter, and the power cable of a power strip is connected to the output part of the power meter. All of the considered appliances such as a vacuum, microwave oven, lamp, TV, dryer, heater, and rice cooker receive the necessary power from the strip. For measuring electrical current, voltage, and real power, the power meter is connected to the system in series. The measured data at the power meter side is used for the processing the classification at the PC side, and the data is transmitted through a serial communication RS-232C. The data is analyzed and utilized in the remaining processing - from the event detection to the classification which is implemented in the PC.

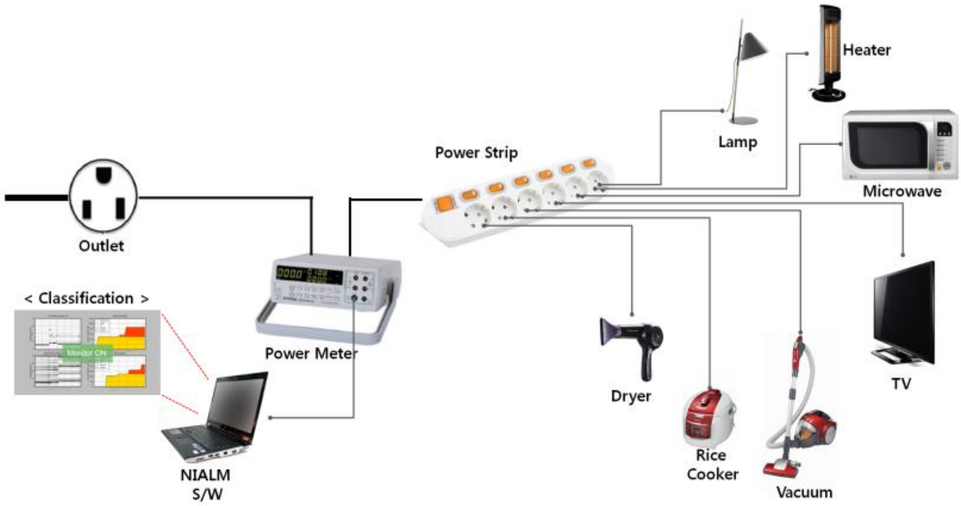


Fig. 3. The simulation environment for NIALM.

For evaluating the performance of proposed algorithm, the classification result is compared with both algorithms based on the real power time window and the real power delta value. In this simulation, overall accuracy is the metric for the classification accuracy of NIALM [9]. The overall accuracy is defined as Equation (3). For the detail terminology, please refer to [9].

$$\eta_{\text{all}} = \frac{N_{\text{dis}}}{N_{\text{true}}} = \frac{N_{\text{dis}}}{N_{\text{det}} - N_{\text{wro}} + N_{\text{miss}}} \tag{3}$$

where $N_{\text{det}} = N_{\text{true}} + N_{\text{wro}} - N_{\text{miss}}$.

There are three simulation cases: single appliance operation, two appliances operation, and three appliances operation. For the reliability of the result, each case is repeated about 100 times. For the single appliance case, ‘on’ and ‘off’ of every seven device is repeated individually. For the two appliances case, ‘on’ and ‘off’ of the following pairs is repeated; a pair of microwave oven and rice cooker, that of a lamp and a TV, that of the heater and TV. For the three appliances case, ‘on’ and ‘off’ of the group, a lamp, heater, and TV, is repeated. The pair for the two or three devices is selected with the consideration that multiple appliances operating simultaneously is highly probable in reality.

### *4.2. Comparison of accuracy*

The accuracy result of three scenarios is summarized in the Table 1. The accuracy of the proposed scheme is greater than other two schemes. For the real power delta scheme, accuracy is relatively lower than other two methods because it performs the classification with observed data sampled at only one time instance. On the other hand, the other two schemes disaggregate the power usage of each appliances with the sequence of monitored data (e.g. observed within the time window length), which adds more feature and ensures more accurate disaggregation. Microwave oven, heater, vacuum, and dryer consume similar level of the real power. This causes the low classification accuracy. In addition, the transient time of the heater is more than one minute while other appliances are about less than 5 seconds. This long transient time makes it hard to disaggregate the heater exactly when heater is working alone or with other appliances.

Table 1. Accuracy performance result

| Case | | Accuracy (%) | | |
|---|---|---|---|---|
| | | Proposed scheme | Time window scheme | Real power delta scheme |
| Case 1 | Microwave oven | 100 | 96 | 20 |
| | Rice cooker | 100 | 100 | 100 |
| | TV | 100 | 100 | 100 |
| | Heater | 77 | 20 | 22 |
| | Vacuum | 21 | 17 | 12 |
| | Dryer | 100 | 91 | 57 |
| | Lamp | 100 | 100 | 100 |
| Case 2 | Microwave oven + rice cooker | 92 | 86 | 33 |
| | Lamp + TV | 100 | 47 | 100 |
| | Heater + TV | 36 | 10 | 19 |
| Case 3 | Lamp + heater + TV | 58 | 31 | 11 |

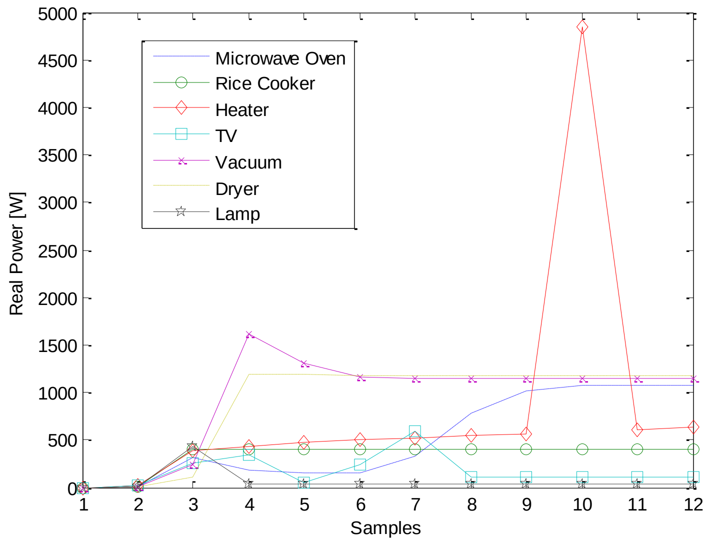


Fig. 4. Trained time window pattern samples (sampling speed: 2 samples/1sec); including the time window pattern of microwave oven, rice cooker, heater, TV, vacuum, dryer, and lamp

Fig. 4 shows the trained time window patterns of several appliances, and it also explains the error of time window scheme thoroughly. Some of the patterns look similar, and there is no large difference in power level; the patterns of the vacuum and dryer show the problem aforementioned. This causes the degradation of classification accuracy for the time window method. In addition, the time window pattern appears to be irregular when multiple appliances are working simultaneously. This problem is not negligible for disaggregating vacuum, so that the accuracy is the lowest among all considered appliances. This could be because appliances mutually have an influence on each other’s power usage, or because the capacity of a single power strip is not enough to provide electricity for all appliances that are currently under operation. The analysis of this problem is beyond the scope of this study, but it is one of the best future work topics in NIALM research area. The proposed method considers the real power level in the steady state which improves the accuracy for the multiple devices cases compared to the time window scheme.

## 5. Conclusion

In this study, we propose a new algorithm for NIALM. The main purpose of the proposed method is to increase the accuracy in a low cost NIALM system, which performs power usage classification with real power data sampled at low speed (about 2 Hz). The proposed method utilizes the time window pattern of the measured real power during the transient state and the real power level in the steady state. The time window pattern ensures the high accuracy performance even if there is an error introduced in the observed data. In addition, the real power level in the steady state ensures high accuracy level when the monitored time window pattern for the multiple devices working case appears to be irregular. The formulated MIP (mixed integer programming) optimization problem considers those two factors. The problem is also reformulated as a linear problem with several LP relaxation schemes. The simulation result verifies that the proposed method improves the overall accuracy about 17% compared to the time window approach and about 28% compared to the real power delta approach.

## References

[1] Farhangi H. The path of the smart grid. *Power and Energy Magazine, IEEE*. 2010; 8(1):18-28.

[2] Neenan B, Robinson J. Residential electricity use feedback: A research synthesis and economic framework. Technical report. Electric Power Research Institute, 2009.

[3] Froehlich J, Larson E, Gupta S, Cohn G, Reynolds MS, Patel SN. Disaggregated end-use energy sensing for the smart grid. *Pervasive Computing*, 2011; 10(1):28-39.

[4] Hart GW. Nonintrusive appliance load monitoring. *Proceedings of the IEEE*. 1992; 80(12):1870-1891.

[5] Cole A, Albicki A. Data extraction for effective non-intrusive identification of residential power loads. In: *Instrumentation and Measurement Technology Conference*, 1998:812-815

[6] Laughman C, Lee K, Cox R, Shaw S, Leeb S, Norford L, Armstrong P. Power signature analysis. *Power and Energy Magazine, IEEE*. 2003; 1(2):56-63.

[7] Patel SN, Supta S, Reynolds M. The design and evaluation of an end-user-deployable whole house, contactless power consumption sensor. In: *Proceedings of CHI2010*, 2010:2471-2480

[8] Boyd S, Vandenberghe L. *Convex optimization*. Cambridge; 2004.

[9] Liang J, Ng SKK, Kendall G, Cheng JWM. Load signature study-part I: basic concept, sitructure, and methodology. *IEEE Transactions on power delivery*, 2010; 25(2):551-560.