International Journal of Fuzzy Logic and Intelligent Systems, vol. 12, no. 3, September 2012, pp. 193-197
http:// dx.doi.org/ 10.5391/IJFIS.2012.12.3.193 PISSN 1598-2645 eISSN 2093-744X

Electric Load Signature Analysis for Home Energy Monitoring System

Lu-Lulu, Sung-Wook Park* and Bo-Hyeun Wang*

Department of Electronic Engineering, Gangneung-Wonju National University, Gangneung, Korea

Abstract

This paper focuses on identifying which appliance is currently operating by analyzing electrical load signature for home energy monitoring system. The identification framework is comprised of three steps. Firstly, specific appliance features, or signatures, were chosen, which are DC (Duty Cycle), SO (Slope of On-state), VO (Variance of On-state), and ZC (Zero Crossing) by reviewing observations of appliances from 13 houses for 3 days. Five appliances of electrical rice cooker, kimchi-refrigerator, PC, refrigerator, and TV were chosen for the identification with high penetration rate and total operation-time in Korea. Secondly, K-NN and Naive Bayesian classifiers, which are commonly used in many applications, are employed to estimate from which appliance the signatures are obtained. Lastly, one of candidates is selected as final identification result by majority voting. The proposed identification frame showed identification success rate of 94.23%.

Keywords : Home energy monitoring system, Load signature, Non-conventional features, Multi-feature multi-algorithm decision method.

1. Introduction

A NILM (Non-intrusive Load Monitoring) is to monitor an electrical circuit that contains a number of appliances that turn on and off independently. By a sophisticated analysis of the current and voltage waveforms of the total load, the NILM estimates the number of the individual loads, their individual energy consumption, and other relevant statistics such as time-of-day variations [1]. Since each individual electrical appliance has the unique load signature described by the common electricity consumption pattern, NILM system should be able to analyze the patterns to identify which appliances that are operating.

MIT (Massachusetts Institute of Technology) pioneered the study of NILM from the early 1980s. They thought they could identify which appliance is currently operating by observing how much real power and reactive power are consumed, assuming that each appliance would consume unique level of real power and reactive power. They measured both reactive power and real power at 0.2Hz interval of 13 appliances at field, and reported identification were successful at the rate of 86 % [2]. CMU (Carnegie Mellon University) tried to show which classification algorithm would yield better success rate if signatures are same [7]. CMU used delta of real power between on-state and off-state, and coefficients obtained by using Fourier regression over real power sampled at 20Hz as features. CMU tested four algorithms over the signatures, i.e., 1-Nearest Neighbor (1-NN), Gaussian Naïve Bayes (GNB), Decision Trees (DT) and Multiclass Adaboost (MultiBoost). The highest success rate was obtained using 1-NN classifier about 79%, and the second was GNB with 67%. CLP (CLP research institute Ltd) proposed novel scheme to improve success rate of load identification. They observed no combination of single-feature and single-algorithm classified all the appliances correctly, i.e. some combination revealed good performance for specific appliances over other combinations, so they got multiple candidate of classification from multiple combinations of single-feature and single-algorithm, and then selected one of the candidates as a final result[3][4]. For the final result, they selected most commonly occurred candidate within the pool, and the scheme showed overall success rate of 92.7% with field data at 12 kHz.

The objective of this paper is developing identification algorithms for the main electrical appliances used in Korean families. And the identification algorithm accuracy can above 92%. This paper is composed of as follows: chapter 2 describes the framework we took in detail, chapter 3 reports experimental results, and chapter 4 closes this paper by summarizing achievements.

2. Proposed Framework of Load Identification

This paper proposes a load signature framework for home electrical power monitoring system as depicted in Fig 1. This approach basically adopts overall structure of [3], and replaced components of it to the better suitable ones for the target devices.

The framework initially extracts features from collected current waveform which were sampled with 1/30 Hz interval during three days at field. The features used are including Duty

Manuscript received Jun. 28, 2012; revised Sep. 18, 2012; accepted Sep. 24. 2012
* Corresponding author: bhw@gwnu.ac.kr

Cycle (DC), Slope of On-state (SO), Variance of On-state (VO), Zero Crossing (ZC), combinations of DC, VO and ZC and combinations of DC, VO, ZC and SO. K-NN and Naive Bayesian classifiers are employed to estimate from which appliance the features are obtained. Each feature is tested by each classification algorithm, so multiple estimates are resulted as candidates of final identification. Since the proposed framework has six features and two algorithms, twelve candidates are resulted and they are marked as follows: DC-KNN, SO-KNN, VO-KNN, ZC-KNN, (DC, VO, ZC)-KNN, (DC, SO, VO, ZC)-KNN, DC-NB, SO-NB, VO-NB, ZC-NB (DC, VO, ZC)-NB and (DC, SO, VO, ZC)-NB. Finally, the most commonly occurred result was selected as the best possible answer, which was named as most common occurrence (MCO) [3][4].

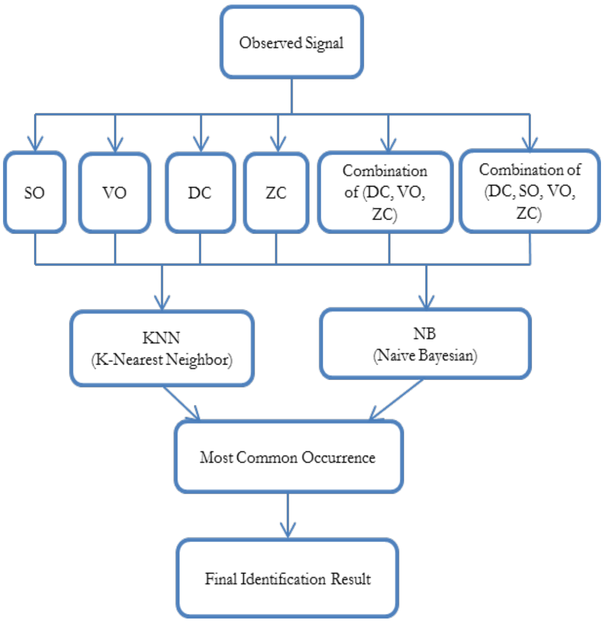

Fig. 1. The proposed load identification

2.1. Data set used for the experiments analysis

This study monitored current waveform of seven kinds of electrical appliances such as TV, PC, refrigerator, kimchi-refrigerator, electric fan, washing machine and electric rice-cooker. They were sampled at 1/30Hz interval and were from 13 homes during 3 days each. The appliances were selected considering high penetration rate and total operation-time in Korea [5][6].

The collected current waveforms were segmented for feature extraction. In order to select proper segments for DC, SO, VO, and ZC, the segments should have 1 hour interval, shall include on-state at least 25 % (15 minutes total), and 18 mA were used as threshold for detection turn-on event. The constraints were selected empirically from the observation of 5 kinds of appliances. Table 1 explains which home volunteered to provide which appliances' current waveform. Table 2 shows how many segments were taken respectively. For experiments we prepared 156 segments (Test Data Set) of 5 appliances.

Table 1 Appliances monitored at each home

Appliance name	refrigerator	kimchi-refrigerator	TV	electric rice-cooker	PC
Home 10	✓	✓	✓	✓	✓
Home 11			✓	✓	✓
Home 12					✓
Home 13		✓	✓	✓	✓
Home 14			✓	✓	
Home 15	✓		✓		✓
Home 16		✓			✓
Home 17	✓	✓	✓	✓	✓
Home 18		✓			
Home 19	✓			✓	✓
Home 20					✓
Home 21	✓	✓			
Home 22	✓		✓	✓	✓
Total	6	6	7	7	10

Table 2 Number of Segments for Classification

Name of Appliances	Number of Test Data Set
Refrigerator	31
Kimchi-refrigerator	53
TV	14
Electric rice-cooker	42
PC	16
Total	156

In order to training classification algorithm, this paper chosen 9 segments per appliances. The number of 9 is around half of number of segments of TV, which has smallest number of segments for experiments.

Test Data Set includes training data set, and at each experiment, 156 segments were used for test data set and 45 segments among the 156 segments were used for training data set.

2.2. Features used for analysis

This paper utilized features of Duty Cycle (DC), Slope of On-State (SO), Variance of On-State (VO), and Zero Crossing (ZC), which are shown in Fig 2.

This paper uses DC (Duty Cycle), which is a ratio between number of samples in on-state and number of samples in off-state. DC is a feature that represents appliances' behavior that turn on and off regularly, so electric rice-cooker (RCKR), refrigerator and kimchi-refregerator are good examples of DC.

SO (Slope of On-State) represents signal's inclination during the on-state. SO can be observed frequently on refrigerator

(FRGR) with which the signal shows small slopes along the time during on states. SO is calculated 1st order regression using regress or robustfit function of MATLAB.

VO (Variance of On-State) represents signal's small change during the on-state. VO is defined as variance/mean during the on-state. VO enables us to distinguish PC and TV from their appliance. SO and VO features are only available during the on-state by their definitions.

ZC (Zero Crossing) represents how frequently an appliance is turn-on or turn-off during an observed segment. ZC can be used to separate refrigerator (FRGR) from kimchi-refrigerator (KFRGR) since KFGR's current waveform shows more frequent crossing than FRGR in a segment. DC and ZC features are utilizing appliance behavior during the on and off-state together.

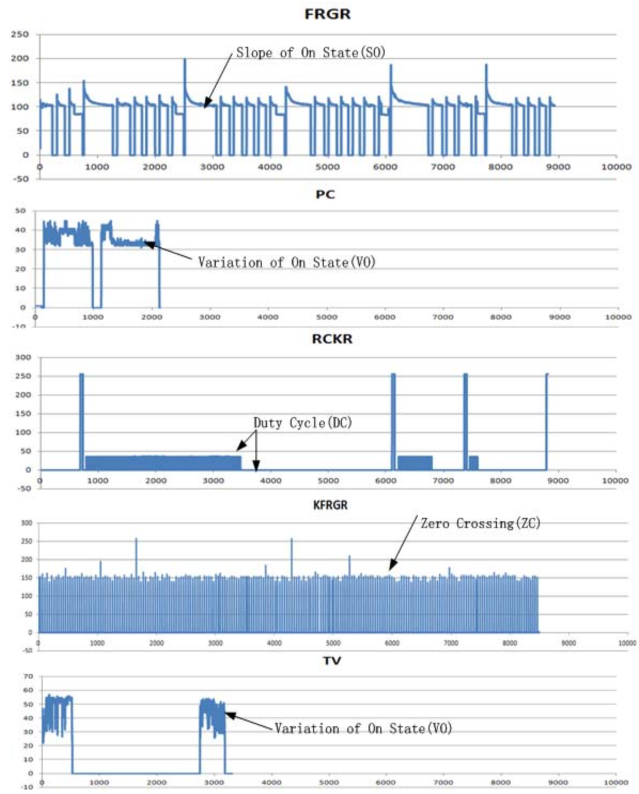

Fig. 2. Four features used

2.3. Algorithms employed for load classification

After careful selection of features the features shall be classified to say which appliance is currently operating. As a classifier this paper employs K-Nearest Neighbor (KNN) classifier and Naive Bayesian (NB) classifier, which were two best classifiers of CMU's study [7].

2.4. Multi-feature/Multi-algorithm Method

A segment of current waveform of an appliance is represented by six features of DC, SO, VO, ZC, combinations of DC, VO and ZC and combinations of DC, SO, VO and ZC, and each feature are classified by K-NN classifier and NB classifier, which lead to twelve candidates called "candidate pool" that are within a circle in the middle of Figure 3. Multi-feature multi-algorithm method is then used to evaluate these potential solutions and render the best final solution among them. This approach is valid since usually single feature can represent some appliances for classification only and different algorithms classifies in different perspectives[3]. In this study multi-feature multi-algorithm method chooses the most commonly occurred candidate within the pool as the final result. This method is selected since it is the least computational demanding mechanism which requires counting the total number of votes within the pool.

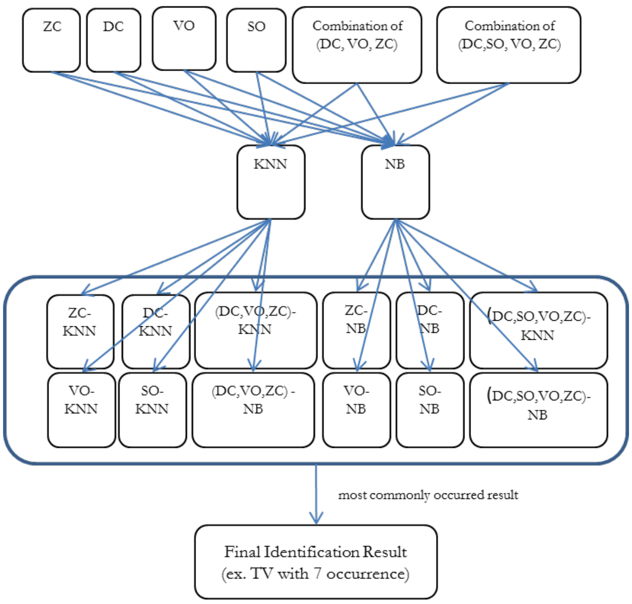

Fig. 3. Flow of Multi-Feature/Multi-Algorithm applied for a test data set

3. Experimental Results & Evaluation

In this experiment, 156 segments were used for testing data set and 45 segments among the 156 segments were used for training data set as described in chapter 2.

3.1 Results with K-NN classifier

Since the performance of a K-NN classifier is primarily determined by the choice of K, several K's were tested by classifying six features in order to select the best classifier for the testing data set. Table 3 lists performance of algorithms.

Table 3 shows SO worked best by 1-NN, DC by 4-NN, VO by 3-NN, ZC by 1-NN and 2-NN, the combinations of DC, VO and ZC by 1-NN and 2-NN algorithms and the combinations of SO, DC, VO and ZC by 1-NN and 2-NN algorithms. From the experiments it was concluded that either 1-NN or 2-NN works better than others.

Table 3 Performance of k-Nearest Neighbor (k=1,2,3,4,5) algorithms

	1NN	2NN	3NN	4NN	5NN
SO	58.33%	58.33%	57.69%	53.85%	48.08%
DC	62.18%	62.18%	69.23%	70.51%	69.87%
VO	78.85%	78.85%	82.69%	80.77%	80.77%
ZC	82.05%	82.05%	81.41%	81.41%	79.49%
DC,VO,ZC	89.74%	89.74%	82.69%	83.97%	83.97%
SO,DC,VO,ZC	92.31%	92.31%	87.82%	88.46%	85.90%

3.2 Results with NB classifier

Table 4 shows the performance for Naive Bayesian (NB) algorithm. NB classifier made the best success rate, when combination of features of DC, VO and ZC was classified, was 89.10%. For the signal feature, Zero Crossing feature and Naive Bayesian (NB) algorithm showed the beat success rate, was 83.33%. For the multi-feature, combination of features of DC, SO, VO and ZC and 1-Nearest Neighbor algorithm showed the beat success rate, was 92.31%.

Table 4 Performance of Naive Bayesian algorithm and 1-Nearest Neighbor algorithm

	1-NN	NB
SO	58.33%	31.41%
DC	62.18%	66.03%
VO	78.85%	66.67%
ZC	82.05%	83.33%
DC,VO,ZC	89.74%	89.10%
SO,DC,VO,ZC	92.31%	88.46%

3.3 Results with proposed framework of load signature analysis

Table 5 showed cases observed during multi-feature multi-algorithm processing when twelve candidates were obtained; then finally obtain the success rate with 94.23%.

Table 5 Multi-feature multi-algorithm method

	FRGR	KFRGR	RCKR	PC	TV
DC-1NN	FRGR	KFRGR	RCKR	TV	TV
VO-1NN	TV	KFRGR	RCKR	PC	RCKR
SO-1NN	PC	KFRGR	PC	PC	RCKR
ZC-1NN	FRGR	KFRGR	RCKR	PC	KFRGR
(DC,VO,ZC)-1NN	TV	KFRGR	RCKR	PC	TV
(DC,VO,SO, ZC)-1NN	FRGR	KFRGR	RCKR	PC	RCKR
DC-NB	FRGR	TV	RCKR	FRGR	TV
VO-NB	RCKR	TV	RCKR	FRGR	TV
SO-NB	PC	KFRGR	KFRGR	PC	TV
ZC-NB	FRGR	TV	RCKR	PC	TV
(DC,VO,ZC)-NB	FRGR	TV	RCKR	PC	TV
(DC,VO,SO, ZC)-NB	FRGR	PC	RCKR	PC	TV
MF/MA	FRGR	KFRGR	RCKR	PC	TV

When reading Table 5, shaded items are observed at almost every single-feature single-algorithm combinations, which mean wrong classification. However, it can be noticed that the results of multi-feature multi-algorithm combinations revealed right classifications. These cased selected indicate that the performance of multi-feature multi-algorithm method is better than the performance of signal-feature signal-algorithm method.

3.4Analysis proposed framework result of load signature

Table 6 Confusion matrix of multi-feature multi-algorithm decision method

Result / Actual	Refrigerator	Kimchi-refrigerator	TV	Electric rice-cooker	PC
Refrigerator	31	0	0	0	0
Kimchi-refrigerator	0	53	0	0	0
TV	3	3	7	0	1
Electric rice-cooker	0	0	0	42	0
PC	2	0	0	0	14

This paper obtained 147 correct decision results and 9 wrong decision results, where 7 wrong decisions from TV and 2 wrong decisions from PC as shown in Table 6. We can say TV is not successfully identified compared with other appliances. With this analysis we made closer look into the case of TV, which is present at Table 7.

Table 7 The proportion of wrong classifier results with TV

Name of Appliances	Number of wrong classifier result	The proportion of wrong classifier results
Refrigerator	30	35.71%
Kimchi-refrigerator	14	16.67%
TV	16	19.05%
Electric rice-cooker	8	9.52%
PC	16	19.05%
Total	84	100.00%

Table 7 presents which candidate results appeared in the middle of MF/MA framework when TV was given. It should be noticed that refrigerator had 35.71% of presence while TV had 19.05% of that. This can be interpreted when TV was identified to Refrigerator, most of features and algorithms thought it was refrigerator. Therefore, in order to improve performance additional features to discriminate TV from Refrigerator are necessary.

When the final result was classified to Kimchi-refrigerator, which has 16.67% of presence, it should be noted that most of the cases TV was rarely a candidate result. This means most features didn't classified as TV, so we need to elaborate to find another feature from the cases that Kimchi-refrigerator was finally answered.

4 Conclusion

This paper focused on identifying which appliance is operating by analyzing electric load signature for home energy monitoring system. Based on different features and different algorithms, a systematic platform of load signature was proposed.

Since smart meter usually measure the power consumption information 1 times for every 15 ~ 60 minutes, this paper analyzed current-waveform collected at every 30 minutes, which were from 13 homes during 3 days and which included 5 kinds of electrical appliance such as electric rice-cooker, PC, refrigerator, kimchi-refrigerator, and TV. To get better identification performance, some non-conventional features were used, which were DC (Duty Cycle), VO (Variance of On-state), SO (Slope of On-state), ZC (Zero Crossing) and the combinations of DC, VO, SO and ZC. 1-NN(1-Nearest Neighbor) and NB (Naïve Bayesian) classifier algorithms are employed in this study to estimate from which appliance the features are obtained. Finally this paper use multi-feature multi-algorithm decision method about twelve candidates that were results from DC-1NN, VO-1NN, SO-1NN, ZC-1NN, (DC, VO, ZC)-1NN, (DC, VO, SO, ZC)-1NN, DC-NB, VO-NB, SO-NB, ZC-NB, (DC, VO, ZC)-NB and (DC, VO, SO, ZC)-NB. And the success rate of multi-feature and multi-algorithm approach was reached 94.23%.

References

[1] http://www.georgehart.com/research/nalm.html

[2] G. W. Hart, "Nonintrusive appliance load monitoring," *Proc. IEEE*,vol.80, no. 12, pp. 1870–1891, Dec. 1992.

[3] Jian Liang, Simon K. K. Ng, Gail Kendall, and John W. M. Cheng, "Load Signature Study - Part I: Basic Concept, Structure, and Methodology," *IEEE Transactions on Power Delivery*, 2010.

[4] Jian Liang, Simon K. K. Ng, Gail Kendall, and John W. M. Cheng, "Load Signature Study - Part II: Disaggregation Framework, Simulation, and Applications," *IEEE Transactions on Power Delivery*, 2010.

[5] Sung-Wook Park, Bo-Hyeun Wang, et al, " A Study on Electric Power Monitoring System per Appliance," *Journal of Korean Institute of Intelligent Systems*, vol. 20, no. 5, pp. 638-644, 2010.

[6] Sung-Wook Park, Jin Soo Seo, Bo-Hyeun Wang, "Development of Home Electrical Power Monitoring System and Device Identification Algorithm," *Journal of Korean Institute of Intelligent Systems*, vol. 21, no. 4, pp. 407-413, 2011.

[7] Mario Berges, Ethan Goldman, H.Scott Matthews, Lucio Soibelman, "Learning System for Electric Consumption of Building," *Carnegic Mellon University*, 2007.

Lu-Lulu
Lu-Lulu received the B.S. degree in Electrical Engineering and Automation from Yancheng Institute of Technology in 2010. He received the M.S. degree in Electronic Engineering from Gangneung-Wonju National University in 2012.
E-mail: lululu19880807@hotmail.com

Sung-Wook Park
Sung-Wook Park received MS and Ph.D degrees from Dept. of Electronics Engineering, Yonsei University, Korea in 1995 and 1998 respectively. After 10 years engineering career at Samsung Electronics Co., LTD, he is currently Assistant Professor at Gangneung-Wonju National University. He's interested in AMI, USN, 3D Audio and Multimedia System.
E-mail:swpark09@gmail.com

Bo-Hyeun Wang
Bo-Hyeun Wang received the B.S. degree in electrical engineering from the Yonsei University, Seoul, Korea, in 1987, and the M.S.E.E. and Ph.D. degrees in electrical engineering from the Georgia Institute of Technology, Atlanta, GA, in 1990 and 1991, respectively. He is currently a professor in the Department of Electrical Engineering of the Gangnung-Wonju National University, Korea. His research interests include computational intelligence, machine learning, and their applications to electrical load forecasting and nonintrusive load monitoring.
E-mail:bhw@gwnu.ac.kr