

# Applications of Fuzzy Classification with Fuzzy C-Means Clustering and Optimization Strategies for Load Identification in NILM Systems

Yu-Hsiu Lin

Graduate Institute of Mechanical  
and Electrical Engineering  
National Taipei University of  
Technology  
Taipei, Taiwan  
t6618001@ntut.edu.tw

Men-Shen Tsai

Graduate Institute of Automation  
Technology  
National Taipei University of  
Technology  
Taipei, Taiwan  
[mstsai@ntut.edu.tw](mailto:mstsai@ntut.edu.tw)

Chin-Sheng Chen

Graduate Institute of Automation  
Technology  
National Taipei University of  
Technology  
Taipei, Taiwan  
[saint@ntut.edu.tw](mailto:saint@ntut.edu.tw)

**Abstract**—Due to global warming and climate changes, it is very important to use and conserve the power energy effectively. Monitoring the electrical consumption of consumers is one of the methods that can improve the energy usage efficiency. In this paper, a Non-Intrusive Load Monitoring (NILM) system, which applies a fuzzy classifier with the Fuzzy C-Means (FCM) clustering and optimization algorithms to identify the energizing and de-energizing statuses of each appliance, is proposed. Load energizing and de-energizing transient features are extracted, and the fuzzy classifier performs load identification based on these features. A two-stage fuzzy classifier is used in this paper. For the first stage, the FCM clustering is used to coarsely determine the parameters of the fuzzy classifier. Following this stage, two optimization algorithms, Error Back-Propagation Algorithm (EBPA) and Genetic Algorithm (GA), are employed to fine tune those parameters. As the classification results obtained from different realistic experimental environments, the proposed system is confirmed that it is able to identify the operational status of each appliance.

**Keywords**- Fuzzy C-Means; Genetic Algorithm; Appliance Identification; Non-Intrusive Load Monitoring; Power Signatures

## I. INTRODUCTION

Due to global warming and climate changes, effectively using and conserving the electricity has become a very important issue recently. One of the approaches for restraining the global warming is to improve the energy usage efficiency by monitoring the residential and commercial electrical consumption. If each appliance is efficiently operated and managed with respect to energy policies formulated by the government, the power energy can be used and conserved effectively. Further, as the energy usage efficiency is improved, the production of greenhouse gases, such as CO<sub>2</sub>, CFCs, and so on, can also be reduced. In order to realize these goals, the operational information of each appliance must be understood first. Traditionally, loads are monitored by attaching “smart” sensors. These sensors are able to calculate the energy consumption and transmit the information to an energy management system. Although this approach is easy to implement, it has several disadvantages, e.g., a large amount of

sensors, high installation cost. Therefore, to remedy the disadvantages of the traditional load monitoring approach, several *Non-Intrusive Load Monitoring* (NILM) systems were proposed [1-12]. The NILM approach requires a voltage and/or current sensor that is installed on the main electrical panel in a household. By analyzing the power signals, the NILM system is able to identify the energy consumption information of each appliance.

This paper aims to assess the feasibility of using fuzzy classification technique with the *Fuzzy c-Means* (FCM) clustering technique and two different optimization algorithms, *Error Back-Propagation Algorithm* (EBPA) and *Genetic Algorithm* (GA), for the NILM. In this paper, a two-stage fuzzy classifier is developed. At the first stage, the FCM clustering is used to coarsely determine the parameters of the antecedent and consequent of *fuzzy If-Then rules*. In this stage, the FCM clustering constructs fuzzy rules from a training dataset; each rule corresponds to a partition of feature space that is of *universes of discourse*. Also, this stage can be seen as embedding *heuristic* knowledge into the fuzzy classifier. After this stage, the EBPA and GA are adopted to fine tune the parameters of each fuzzy rule such that the classification accuracy can be improved. The load energizing and de-energizing transient features are extracted from current waveforms that are measured at the power entering point. The fuzzy classifier performs load identification based on the features. As the results, the proposed NILM system is capable of distinguishing the energizing and de-energizing statuses of each appliance. The organization of this paper is as follows: First, the existing NILM methods and power signatures used by the methods are surveyed in Sec. II. The proposed NILM method is discussed in Sec. III. The optimization strategies employed to fine tune the parameters are expressed in Sec. IV. The realistic experiments and their results are shown in Sec. V. The conclusions are summarized in Sec. VI.

## II. EXISTING NON-INTRUSIVE LOAD MONITORING SYSTEMS

Two signal processing approaches, steady-state analysis and transient analysis, for NILM have been studied in the past.

An NILM system developed by the Electric Power Research Institute (EPRI) applied steady-state approach [1]. The system uses the variations of real and reactive power as the identification signatures. A five-stage algorithm was applied to detect the operation ON/OFF status of each load. Although the system with acceptable performance had been applied to some fields, several problems were pointed out by references [2-4]. For example, the system cannot distinguish the loads with similar real and reactive power consumption. In papers [5-6], the authors indicated that the nonlinear loads not only consume real and reactive power, they also introduce harmonics to the power lines. Hence, the harmonics can be used as the features for identifying different loads. In [7], nine time-domain steady-state features were used. Combinations of the features were sequentially inputted to a 1-Nearest-Neighbor Rule recognizer in order to find the best features that result the highest recognition rate. Besides the steady-state analysis, several studies shown that the transient waveforms for different types of loads are unique due to the differences of physical characteristics [2-3, 8-9]. These studies also found that the transient waveforms have completely or partially repeated profiles. Therefore, the transient properties can also be used as the features for load identification. In [10], a load recognition scheme that applied the notion of maximum approaching degree based on fuzzy logic theory was proposed. Since this scheme does not involve any training process, the NILM system is simple in implementation. In articles [4, 11-12], several steady-state features, such as real power, reactive power, and total harmonic distortion, with the turn-on transient energy are used as the features used by different Artificial Neural Network architectures for load identification. These articles also introduced the Coefficients of Variations to validate the repeatability of turn-on transient energy for each load.

### III. PROPOSED NON-INTRUSIVE LOAD MONITORING SYSTEM

In this study, the proposed NILM system that integrates transient load current feature extraction methods with *soft-computing* techniques is used to identify the energizing and de-energizing statuses of each load. The flowchart of the proposed system is shown in Figure 1.

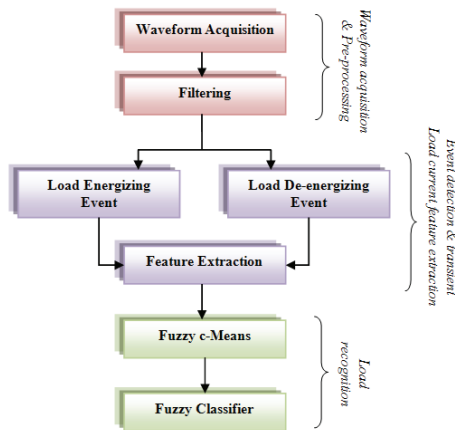


Figure 1. Flowchart of the proposed NILM System. The fuzzy classifier is optimized by optimization strategies.

Detailed descriptions of the system are described in the following sub-sections.

#### A. Waveform Acquisition and Filtering

The aim of waveform acquiring and filtering is to acquire load current waveforms and filter the high-frequency noise, respectively. The cut-off frequency of the low-pass filter of 500 Hz is used in this approach.

#### B. Event Detection and Feature Extraction

In the proposed system, the load energizing or de-energizing event of an appliance can be detected. The feature extraction process is triggered after the event is detected. Following descriptions deal with the processes of event detection and feature extraction.

##### 1) Energizing event detection & Feature extraction:

When the *variation* of current intensity calculated using Eq. (1) is greater than a pre-specified positive threshold,  $\alpha$ , a load energizing event is detected.

$$\Delta[I_{\text{intensity}}] = (I_{\text{intensity}})_{k+1} - (I_{\text{intensity}})_k \quad (1)$$

$$I_{\text{intensity}} = \frac{\sum_{j=1}^N |i(j) - \text{mean}(i)|}{N} \quad (2)$$

In Eq. (1),  $i(j)$  represents the  $j$ -th current sampling point;  $N$  is the total number of sampling points per cycle;  $k$  indicates the  $k$ -th sampling cycle;  $\text{mean}(i)$  denotes the average current value for each sampling cycle.

After the load energizing event is detected, the process for determination of transient period starts. Since the transient period may continue for certain cycles after a load is energized, a differential procedure based on the fundamental property of superposition is performed first. After this procedure, the *change rate* of current intensity is calculated using Eqs. (3) and (4). In Eq. (3),  $i'(t)$  represents a differential waveform that is obtained through the differential procedure. During this process,  $\Delta I'$  is compared with a pre-specified threshold,  $\gamma$ . If  $\Delta I'$  is less than  $\gamma$  for a pre-specified  $\delta$  cycles, the determination of transient period terminates. Finally, the transient period,  $E_t$ , is defined as the total period of transient determination minus  $\delta$  cycle time.

$$I'_{\text{intensity}} = \log\left(\sum_{j=1}^N (i'(j) - \text{mean}(i'))^2\right) \quad (3)$$

$$\Delta[I'_{\text{intensity}}] = |(I'_{\text{intensity}})_{k+1} - (I'_{\text{intensity}})_k| \quad (4)$$

Some load energizing current waveforms for different single and multiple load operation scenarios are shown in Figure 2. The acquired transient current waveform,  $i_{aq}(t)$ , of an unidentified load is shown on the right column of these figures. In multiple load operation scenarios, all loads except the unidentified load were energized in advance.

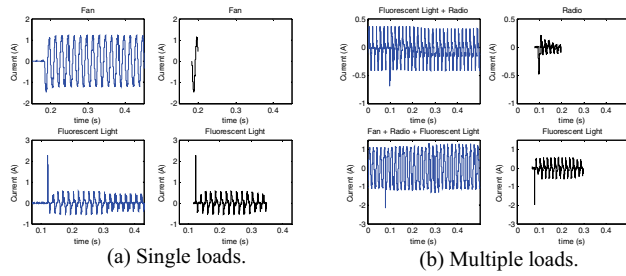


Figure 2. Load current waveforms for different operation scenarios.

Feature extraction is crucial for load identification. The feature extraction process transforms an observation space into a feature space. Generally, the feature space is of much lower dimensions than the observation space. In this paper, *Crest Factor (C.F.)* calculated using Eq. (5) and  $E_t$  are used as the features for load energizing identification. In Eq. (5),  $i_{aq}(t)$  indicates the acquired transient waveform.

$$\text{Crest Factor (C.F.)} = \frac{\max |i_{aq}(t)|}{\sqrt{\int_{E_t} i_{aq}^2(t) dt}} \quad (5)$$

## 2) De-energizing event detection & Feature extraction:

When the variation of current intensity that is obtained from Eq. (1) is less than a pre-specified negative threshold,  $\beta$ , a load de-energizing event is detected. Following the detection of the event, the current waveforms before and after the event are subtracted. By doing so, the load de-energizing transient waveform of an undetermined load can be acquired. For the determination of load de-energizing transient period, the transient period,  $E_t$ , is defined as one cycle.

Figure 3 shows some load de-energizing current waveforms for different single and multiple load operation scenarios. The features, *C.F.* and  $I_{peak}$ , of an acquired de-energizing transient waveform are calculated using Eqs. (5) and (6). These features are used for load de-energizing identification.

$$I_{peak} = \max |i_{aq}(t)| \quad (6)$$

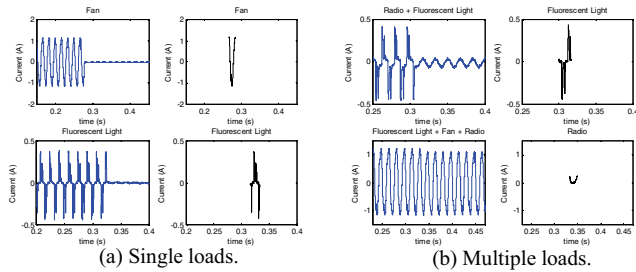


Figure 3. Load current waveforms for different operation scenarios.

## C. Fuzzy Classifier with Fuzzy C-Means Clustering

In this sub-section, the fuzzy classifier and the FCM clustering which is used to roughly construct the fuzzy classifier are discussed. The structure of the fuzzy classifier is shown in Figure 4. In Figure 4, the *fuzzifier* transforms the crisp input values calculated from the acquired current waveform into fuzzy sets. After this fuzzification process, the

fuzzified inputs with fuzzy rules are passed to *inference engine* that combines them using a fuzzy *t*-norm operator to derive the *firing strength* of each fuzzy rule. When the firing strength of each rule is obtained, the crisp output can be calculated using a *de-fuzzifier*. Finally, this crisp output is delivered to the *decision-making unit* to determine the class of an unidentified appliance corresponding to the acquired current waveform.

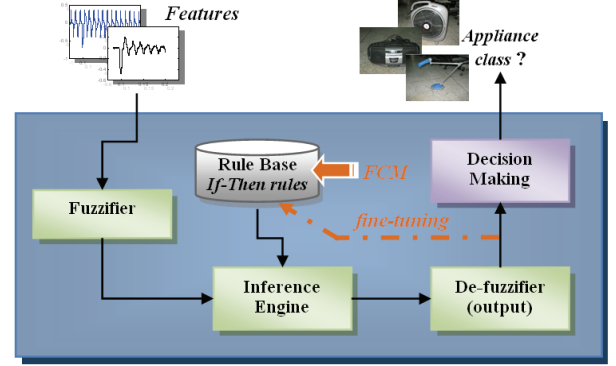


Figure 4. Structure of the fuzzy classifier.

## 1) Fuzzy Classifier:

A *fuzzy Rule Base* that consists of a set of fuzzy IF-THEN rules has the form shown in Eq. (7) [13]:

$$Ru^{(l)} : \text{IF } x_1 \text{ is } A_1^l \text{ and ... and } x_n \text{ is } A_n^l, \text{ THEN } y \text{ is } B^l \quad (7)$$

In Eq. (7),  $A_i^l$  and  $B^l$  are fuzzy sets in  $U_i \subset R$  and  $V \subset R$ , respectively;  $x = (x_1, x_2, \dots, x_i, \dots, x_n)^T \in U \subset R^n$  and  $y \in V \subset R$  are the input and output linguistic variables of the fuzzy classifier, respectively;  $M$  is the total number of rules in the fuzzy rule base (i.e.,  $l=1, 2, \dots, M$ ).

## a) Fuzzifier, Inference Engine, and De-fuzzifier:

Each fuzzy set in Eq. (7) is characterized as a *Gaussian membership function*. The center of the fuzzy set  $B^l$  in Eq. (7) is *singleton* with value of  $\bar{y}^l$ . By using *singleton fuzzifier*, *product inference engine* and *center average de-fuzzifier*, the fuzzy classifier can be formulated as Eq. (8). In Eq. (8),  $f(x) \in V \subset R$  is the de-fuzzified output of the fuzzy classifier;  $a_i^l \in (0, 1]$  are constants;  $\sigma_i^l \in (0, \infty)$ ,  $\bar{x}_i^l \in R$ , and  $\bar{y}^l \in R$  are real-valued parameters.

$$f(x) = \frac{\sum_{l=1}^M \bar{y}^l [\prod_{i=1}^n a_i^l \exp(-\frac{(x_i - \bar{x}_i^l)^2}{\sigma_i^l})]}{\sum_{l=1}^M [\prod_{i=1}^n a_i^l \exp(-\frac{(x_i - \bar{x}_i^l)^2}{\sigma_i^l})]} \quad (8)$$

The Gaussian-type membership function rather than other types of membership functions is chosen in this paper for two reasons. Firstly, the total number of parameters to be tuned or evolved is small since each Gaussian membership function requires two parameters. Secondly, the Gaussian-type membership function ensures that the firing strength of each rule is always non-zero. In addition, the total number of universes of discourse for the input of the fuzzy classifier is

two. Each universe of discourse is further divided into three partitions using three Gaussian membership functions (because three-class energizing or de-energizing classification problem is considered). Thereby, the total number of fuzzy rules is nine. Thus, the feature space is partitioned into  $3 \times 3$  decision regions. In this paper, the partition with the parameters of the fuzzy classifier is initialized using the FCM clustering, which will be discussed in the next sub-section.

#### b) Decision-Making Unit:

After the de-fuzzified output of the fuzzy classifier is calculated using Eq. (8), the output is passed to the decision-making unit to determine the class label of the input using Eq. (9). In Eq. (9),  $Class^k$ , an integer, denotes the pre-assigned class label of the  $k$ -th class;  $w$  is the total number of classes.

$$Class\ label = Class^j; \quad j = \underset{k=1,2,\dots,w}{\operatorname{argmin}} (|f(x) - Class^k|) \quad (9)$$

The recognition rate is one of the criteria for evaluating the performance of a classifier. In this paper, the recognition rate is defined as the number of correctly classifying the test data with the total data in the test dataset.

#### 2) Fuzzy c-Means Clustering:

A 2-stage approach used to design the fuzzy classifier is proposed in this paper. These stages are described as follows:

- **Stage 1.** Apply the FCM clustering with the training dataset to coarsely determine the parameters of the fuzzy If-Then rules. This stage is a “heuristic” stage.
- **Stage 2.** Apply the optimization algorithms with the training dataset to fine tune the parameters determined in Stage 1.

After the fuzzy classifier with the acceptable performance is designed, it can be used to classify new patterns.

The FCM clustering is one of the partition clustering algorithms for data clustering [13, 14]. The FCM clustering algorithm allows each datum belonging to two or more clusters as the sense in *Fuzzy Logic*. The goal of the FCM clustering applied in this research is to find  $c$  *codebook prototypes/centroids/centers* in respect of the training dataset. These  $c$  cluster centers can be used to determine the parameters of each fuzzy If-Then rule. The FCM clustering procedure is summarized as follows:

- **Step 1.** For a dataset  $X = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k, \dots, \mathbf{x}_n\}$ , where  $\{\mathbf{x}_k\}_{k=1}^n \in \mathbb{R}^p$ : fix  $c \in \{2, 3, \dots, (n-1)\}$ , set  $m \in (1, \infty)$ , and initialize  $U^{(0)} \in M_{fc}$ . Here,  $M_{fc} = \{U \in V_{cn} \mid u_{ik} \in [0, 1], 1 \leq i \leq c, 1 \leq k \leq n; \sum_{i=1}^c u_{ik} = 1, \forall k \in \{1, 2, \dots, n\} \text{ is true}\}$ ;  $V_{cn}$  is the set of real  $c \times n$  matrices  $U = [u_{ik}]$ ;  $u_{ik}$  is the membership value of  $\mathbf{x}_k$  that belongs to cluster  $A_i$ .
- **Step 2.** At iteration  $l$ , where  $l=0, 1, 2, \dots$ : compute the  $c$  mean centers using

$$\mathbf{v}_i^{(l)} = \frac{\sum_{k=1}^n (u_{ik}^{(l)})^m \mathbf{x}_k}{\sum_{k=1}^n (u_{ik}^{(l)})^m}, 1 \leq i \leq c$$

- **Step 3.** Update  $U^{(l)}$  to  $U^{(l+1)} = [u_{ik}^{(l+1)}]$  using

$$u_{ik}^{(l)} = \frac{1}{\sum_{j=1}^c \left( \frac{\|\mathbf{x}_k - \mathbf{v}_i^{(l)}\|}{\|\mathbf{x}_k - \mathbf{v}_j^{(l)}\|} \right)^{\frac{2}{m-1}}}, 1 \leq i \leq c, 1 \leq k \leq n$$

- **Step 4.** If  $\|U^{(l+1)} - U^{(l)}\|$  is less than or equal to a pre-determined tolerance, stop; otherwise, set  $l=l+1$  and go to Step 2.

The cluster centers and the distribution of clusters are used in this paper. They are identified using the FCM clustering to roughly allocate the parameters of the fuzzy classifier. The *center* parameter of each Gaussian membership function,  $\bar{x}_{x_i,j}$ , on each universe of discourse is set to the corresponding component of cluster center. Also, the *spread* parameter of each Gaussian membership function on each universe of discourse is determined using Eq. (10).

$$\sigma_{x_i,j} = x_i' - \bar{x}_{x_i,j}; \quad i=1, 2, \dots, n, \text{ and } j=1, 2, \dots, N_{MF} \quad (10)$$

In Eq. (10),  $n$  is the total number of universes of discourse (input variables);  $N_{MF}$  is the total number of membership functions (clusters) on each universe of discourse;  $x_i$  is the component of a datum which belongs to cluster  $j$  with the approximate membership degree of  $e^{-1}$ . For the cases studied in this paper,  $n$  and  $N_{MF}$  are 2 and 3, respectively.

After the center and spread parameters are determined, the singleton parameters,  $\{\bar{y}^l\}_{l=1}^M$ , for the part of consequent of fuzzy rules can also be assigned based on the partitions and the labels of training data. After coarsely determining these parameters, they are further adjusted using the EBPA and GA.

## IV. OPTIMIZING OF FUZZY CLASSIFIER

In this section, two different types of optimization algorithms are applied to fine tune the parameters of fuzzy classifier. The gradient descent algorithm [13, 15-16] used to find the quasi-optimal structure of the fuzzy classifier of (8) according to the training dataset is discussed first.

### A. Gradient Descent Algorithm Approach

Suppose that the following input-output pairs are collected:  $\{(x_0^p, y_0^p)\}_{p=1}^N$ , where  $x_0^p \in U \subset \mathbb{R}^n$  and  $y_0^p \in V \subset \mathbb{R}$ . The goal is to design the fuzzy classifier based on these  $N$  input-output pairs. In Eq. (8),  $\sigma_i^l$ ,  $\bar{x}_i^l$ , and  $\bar{y}^l$  are roughly initialized based on the results of the FCM clustering. Once these parameters are specified, the quasi-optimal structure of the fuzzy classifier can be obtained. The gradient descent algorithm used for tuning the classifier's parameters is summarized as follows:

- **Step 1.** Initialize the parameters. Determine the initial parameters,  $\bar{y}^l(0)$ ,  $\bar{x}_i^l(0)$ , and  $\sigma_i^l(0)$ , according to stage 1.
- **Step 2.** Present an input-output pair  $(x_0^p, y_0^p)$  and calculate the output associated with the inputs of the fuzzy classifier at the  $q$ -th iteration of tuning.
- **Step 3.** Update the parameters. The following updating equations are used to update the parameters:

$$\bar{y}^l(q+1) = \bar{y}^l(q) - \eta \frac{f - y}{a} b^l, \text{ where } \eta \text{ is a tuning constant; } f$$

and  $y$  denote  $f(x_0^p)$  and  $y_0^p$ , respectively;  $a = \sum_{l=1}^M b^l$ ;

$$b^l = \prod_{i=1}^n \exp\left(-\frac{(x_i - \bar{x}_i^l)^2}{\sigma_i^l}\right).$$

$$\bar{x}_i^l(q+1) = \bar{x}_i^l(q) - \eta \frac{f - y}{a} (\bar{y}^l(q) - f) b^l \frac{2(x_{0i}^p - \bar{x}_i^l(q))}{\sigma_i^{l2}(q)}.$$

$$\sigma_i^l(q+1) = \sigma_i^l(q) - \eta \frac{f - y}{a} (\bar{y}^l(q) - f) b^l \frac{2(x_{0i}^p - \bar{x}_i^l(q))^2}{\sigma_i^{l3}(q)}.$$

- **Step 4.** Go to Step 2 with  $q=q+1$ , until  $|f - y_0^p|$  is less than a pre-determined tolerance, or until the value of  $q$  equals to a pre-determined iteration.
- **Step 5.** Go to Step 2 with  $p=p+1$  to update the parameters using next input-output pair,  $(x_0^{p+1}, y_0^{p+1})$ .
- **Step 6.** Set  $p=1$  and re-do steps 2 to 5, until the result of the classifier is satisfactory.

### B. Genetic Algorithm Approach

For the second approach, GA is used to globally search the optimal parameters. [16-18] The GA, which had been applied to different engineering optimization problems, is a global and stochastic search algorithm. Two types of operators for the GA are used: *genetic operators* and *evolution operator*. The genetic operators include *crossover* and *mutation* operators; the evolution operator comprises a *selection* operator. The flowchart of the GA used in this paper is shown in Figure 5.

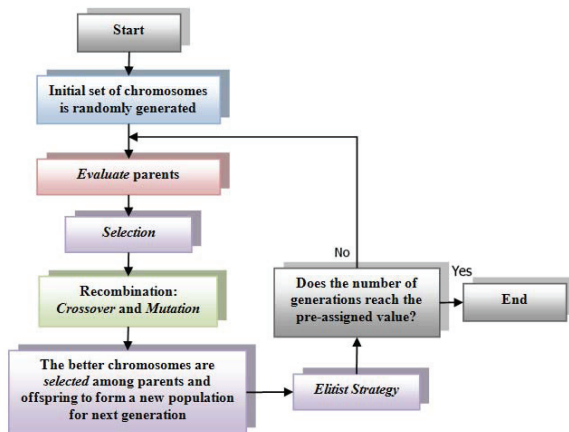


Figure 5. Flowchart of the GA.

With application of GA, the first important step is to *encode* the fuzzy classifier into the *chromosome* (*genotype* representation). In this paper, two universes of discourse are considered. Each universe of discourse is partitioned into three regions using three Gaussian membership functions. Hence, there are nine rules due to the  $3 \times 3$  regions. As a result, a total of 21 parameters ( $3_{\text{membership functions}} \times 2_{\text{parameters per function}} \times 2_{\text{inputs}} + 9_{\text{rules}}$ ) need to be adjusted. Each parameter is encoded into a binary string. The structure illustration of the chromosome is illustrated in Figure 6.

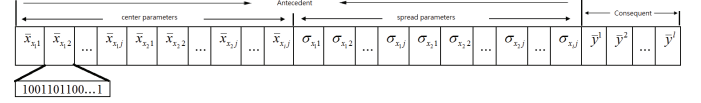


Figure 6. Representation of the chromosome.

In order to evaluate the quality of each chromosome, *fitness function* is needed. A good fitness function can gradually guide the GA to find the optimal solution. In this paper, the fitness function, the reciprocal of *mean-squared error* ( $MSE_{error}$ ), is defined as follows:

$$fitness = \frac{1}{MSE_{error}} = \frac{1}{\frac{1}{N} \sum_{p=1}^N [f(x_0^p) - y_0^p]^2}$$

s.t.

$$\bar{x}_{x_i(j-1)} \leq \bar{x}_{x_i,j} \leq \bar{x}_{x_i(j+1)}$$

$$\bar{x}_{x_i,j} - \rho \cdot \sigma_{x_i,j} \leq \sigma_{x_i,j} \leq \bar{x}_{x_i,j} + \rho \cdot \sigma_{x_i,j}$$

$$\bar{y}^l - 1 \leq \bar{y}^l \leq \bar{y}^l + 1$$

Note that, the boundaries of these constraints are determined based on the results of the FCM clustering; and  $\rho$  is a constant.

During *evaluation*, each chromosome must be *decoded* to the corresponding fuzzy classifier (*phenotype* representation). After each chromosome in current generation is evaluated, the *natural selection* process that mimics the process of Darwinian evolution to create populations from generation to generation is performed. In this paper, the *roulette wheel selection* is adopted; it belongs to the fitness-proportional selection. A new population with respect to the probability distribution based on fitness values is generated. The *enlarged* ( $\mu + \lambda$ ) *sampling space* is considered in this paper. With this sampling space,  $\mu$  parents and  $\lambda$  offspring compete for survival. The  $\mu$  best out of offspring and old parents are selected as the population for the next generation. The genetic operators facilitate an efficient search and guide the search into new regions. The crossover facilitates exploration, while the mutation facilitates exploitation of the search space [18]. In this paper, one-cut-point crossover and bit-wise flipping mutation operators are used. In addition, the *Elitist strategy* is also applied. This strategy ensures that the best chromosome is passed onto next generation if it is not selected via the selection process.

## V. EXPERIMENTAL RESULTS

In this paper, the proposed NILM system is used to identify the operation status of each appliance. The appliances include *Fan*, *Fluorescent Light* (F.L.) and *Radio*. All the experiments in this study are executed at different realistic environments. During identification, no loads were energized or de-energized simultaneously. Experimental 1 describes the results of the appliance energizing identification; Experimental 2 describes the results of the appliance de-energizing identification. A circuit was designed to energize the appliance at different energizing phase angles. As the waveforms shown in Figures 2 and 7, the current waveforms for different types of loads are *unique* and *repeatable*.

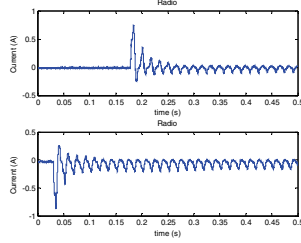


Figure 7. Current waveforms for different load energizing phase angles. The load, Radio, was energized at 73.5 degree (top) and 225 degree (bottom) at different measurement locations.

For all the experiments, 108 single load current measurements are performed. 69 of the measurements are used for training; while the remaining 39 measurements are used for testing. Additionally, 41 multiple load current measurements are used for testing as well.

### A. Experimental 1: Load Energizing Classification

In this experiment, the proposed system is used to identify the energizing status of each load. The results of the fuzzy classifier with the two different optimization strategies for this load energizing classification are summarized below.

#### 1) Fuzzy Classifier with the EBPA

In this experiment, the universes of discourse for the inputs are set to appropriate ranges according to the distribution of the training data. The feature space and the results of the FCM clustering are illustrated in Figure 8.

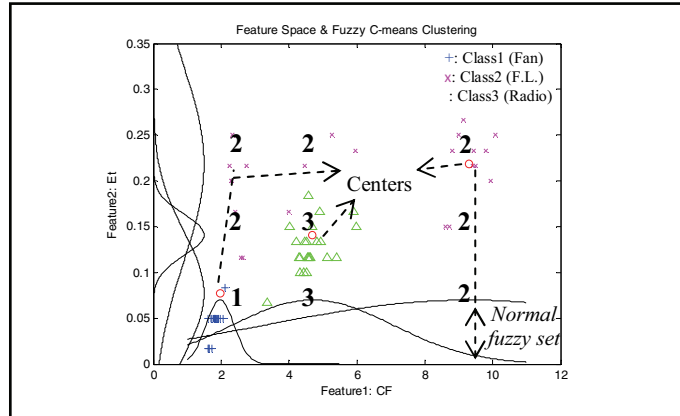


Figure 8. Feature space and fuzzy c-means clustering in Experimental 1.

In Figure 8, the interval of  $CF$  is  $[1, 11]$ , whereas the interval of  $E_t$  is  $[0, 0.3]$ . The label of each partition is assigned according to the clusters and the training data. Analysis of the data shows that  $E_t$  has less ambiguity than  $CF$ . Hence,  $E_t$  is a better feature than  $CF$ .

The values of feature parameters,  $(\alpha, \gamma, \delta)$ , used in this experiment are  $(0.03, 0.065, 10.0)$ , respectively. These values are determined based on the past experience. The fuzziness constant,  $m$ , is 2; the value of  $c$  is 3 because this is a three-class problem. The tuning constant,  $\eta$ , is 0.003; the maximum training epoch is 500. The training and testing results of the fuzzy classifier with the gradient descent algorithm in this experiment are shown in Figure 9. The testing results are tabled in Table I.

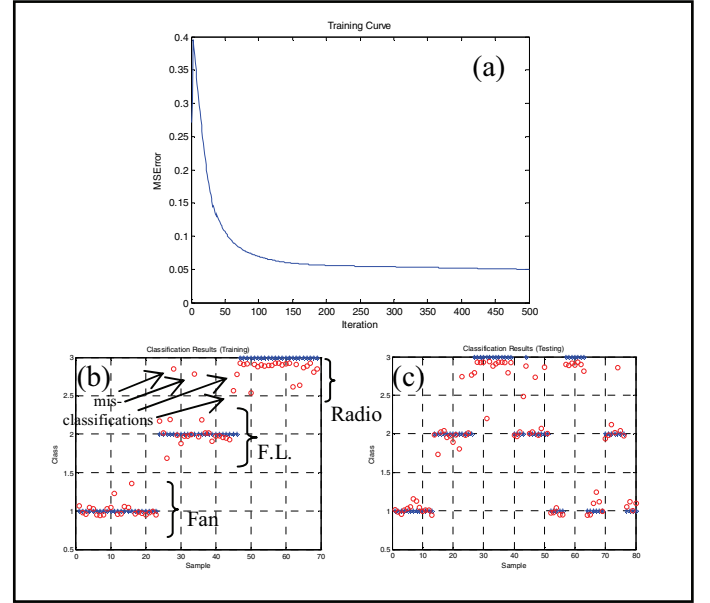


Figure 9. Training and testing results of the fuzzy classifier with the gradient descent algorithm in experimental 1. (a). Training curve. The final  $MSE_{error}=0.04991$ ; (b) Training result. The O.R.R.=94.2%; (c). Testing result. The O.R.R.=92.5%.

TABLE I. IDENTIFICATION RESULTS OF THE FUZZY CLASSIFIER WITH THE GRADIENT DESCENT ALGORITHM IN EXPERIMENTAL 1

| C.M.                    | Single Load and Multiple Load operation |             |             |
|-------------------------|-----------------------------------------|-------------|-------------|
|                         | Fan                                     | F.L.        | Radio       |
| Fan                     | 28 (100%)                               | 0 (0%)      | 0 (0%)      |
| F.L.                    | 0 (0%)                                  | 26 (83.87%) | 5 (16.13%)  |
| Radio                   | 0 (0%)                                  | 1 (4.76%)   | 20 (95.24%) |
| O.R.R. <sup>a</sup> (%) | 92.5                                    |             |             |

O.R.R.<sup>a</sup>: Overall Recognition Rate

The units used in the Confusion Matrix (C.M.) table are the number of successful/unsuccessful recognizing count and percentage, respectively. From Table I, Fluorescent Light has the lowest recognition rate of 83.87%; Radio has the recognition rate of 95.24%; and Fan has the highest recognition rate of 100%. The overall recognition rate is 92.5%.

## 2) Fuzzy Classifier with the GA

The parameters used by the GA are as follows: The total length of each chromosome is 223. The population size and maximum generation are 50 and 200, respectively. For simplicity, the crossover and mutation rates during evolution are held constant; they are set to 0.63 and 0.07, respectively. The training and testing results of the fuzzy classifier with the GA in this experiment are shown in Figure 10. The testing results are shown in Table II.

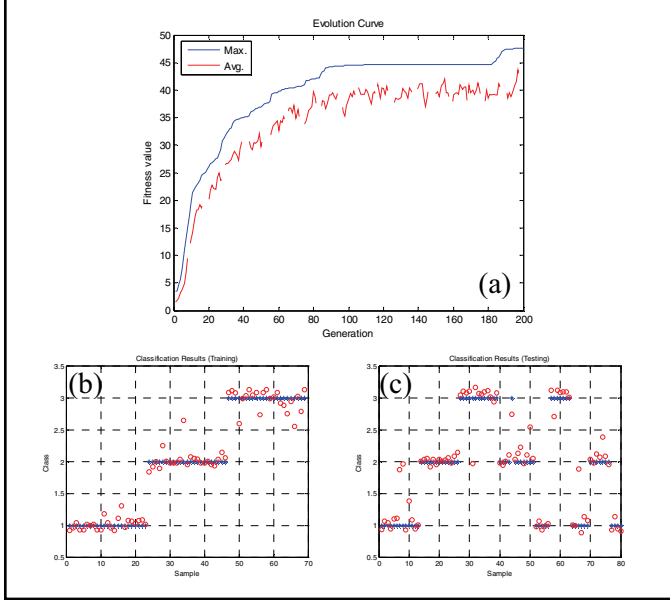


Figure 10. Training and testing results of the fuzzy classifier with the GA in experimental 1. (a). Evolution curves. The final f.v.=47.624; (b) Training result. The O.R.R.=98.55% (+4.35%); (c). Testing result. The O.R.R.=93.75% (+1.25%).

TABLE II. IDENTIFICATION RESULTS OF THE FUZZY CLASSIFIER WITH THE GA IN EXPERIMENTAL 1

| C.M.              | Single Load and Multiple Load operation |             |             |
|-------------------|-----------------------------------------|-------------|-------------|
|                   | Fan                                     | F.L.        | Radio       |
| Fan               | 25 (89.29%)                             | 3 (10.71%)  | 0 (0%)      |
| F.L.              | 0 (0%)                                  | 30 (96.77%) | 1 (3.23%)   |
| Radio             | 0 (0%)                                  | 1 (4.76%)   | 20 (95.24%) |
| <b>O.R.R. (%)</b> | <b>93.75</b>                            |             |             |

From Table II, Fan has the lowest recognition rate of 89.29%; Radio has the recognition rate of 95.24%; and Fluorescent Light has the highest recognition rate of 96.77%. The overall recognition rate is 93.75%.

## B. Experimental 2: Load De-energizing Classification

In this experiment, the proposed system is used to identify the de-energizing status of each load. The feature space and the results of the FCM clustering are illustrated in Figure 11 (a). Analysis of the data shows that  $CF$  as well as  $I_{peak}$  is a good feature for classification. Based on the past experience, the  $\beta$  is set to 0.016. The tuning constant,  $\eta$ , is 0.01; the maximum training epoch is 500. The training and testing results of the fuzzy classifier with the gradient descent algorithm in this experiment are shown in Figure 11 (b) to (d).

The testing results are tabled in Table III. The results show that the overall recognition rate is 100%.

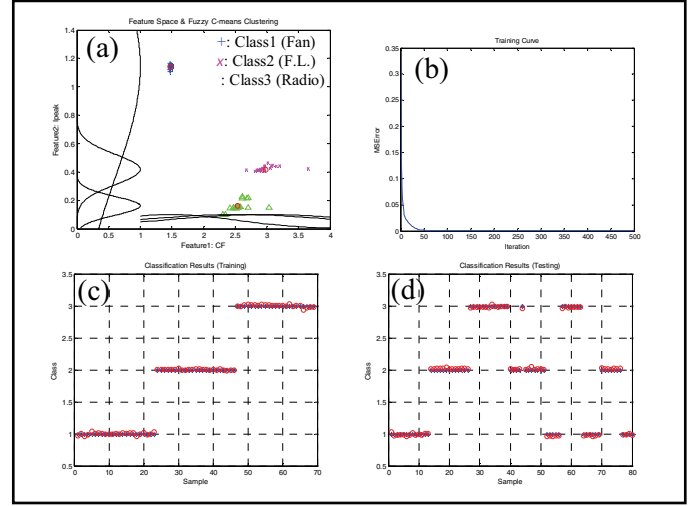


Figure 11. Training and testing results of the fuzzy classifier with the gradient descent algorithm in experimental 2. (a) Feature space and fuzzy c-means clustering in Experimental 2; (b). Training curve. The final  $MSE_{error}$ =0.000268; (c) Training result. The O.R.R.=100%; (d). Testing result. The O.R.R.=100%.

TABLE III. IDENTIFICATION RESULTS OF THE FUZZY CLASSIFIER WITH THE GRADIENT DESCENT ALGORITHM IN EXPERIMENTAL 2

| C.M.              | Single Load and Multiple Load operation |           |           |
|-------------------|-----------------------------------------|-----------|-----------|
|                   | Fan                                     | F.L.      | Radio     |
| Fan               | 28 (100%)                               | 0 (0%)    | 0 (0%)    |
| F.L.              | 0 (0%)                                  | 31 (100%) | 0 (0%)    |
| Radio             | 0 (0%)                                  | 0 (0%)    | 21 (100%) |
| <b>O.R.R. (%)</b> | <b>100</b>                              |           |           |

## VI. CONCLUSIONS

The NILM system acts an important role for effectively using and conserving the electricity. In this paper, an NILM system that integrates transient load current feature extraction methods and soft-computing techniques is proposed. The proposed system, which uses the fuzzy classifier with the FCM clustering and the two optimization algorithms as the load identifier, is capable of identifying energizing and de-energizing statuses of each load under single load and multiple load operation scenarios in different realistic environments. From the experimental results, the classification performance of the fuzzy classifier with the GA is better than the performance of the classifier with the conventional EBPA because the GA is a global search optimization algorithm. The overall recognition rates are above 92.5%. Additionally, the identification is not affected by different experimental locations. These results confirm that the proposed NILM system is able to identify the usage status of each appliance robustly.

## VII. ACKNOWLEDGEMENT

This research is partially supported by National Energy Project of National Science Council, Taiwan, under the grant number NSC 99-3113-P-006-004.

## REFERENCES

- [1] S. Drenker and A. Kader, "Nonintrusive Monitoring of Electric Loads," *IEEE Computer Applications in Power*, vol. 12, no. 3, pp. 47-51, Oct. 1999.
- [2] C. Laughman, K. Lee, R. Cox, S. Shaw, S. B. Leeb, L. Norford, and P. Armstrong, "Power Signature Analysis," *IEEE Power & Energy Magazine*, vol. 1, pp. 56-63, March/April 2003.
- [3] A. Shrestha, E. L. Foulks, R. W. Cox, "Dynamic load shedding for shipboard power systems using the non-intrusive load monitor," *IEEE Electric Ship Technologies Symposium*, pp. 412-419, 20-22 April 2009.
- [4] H. H. Chang, C. L. Lin, H. T. Yang, "Load Recognition for Different Loads with the Same Real Power and Reactive Power in a Non-intrusive load-monitoring System," *12th International Conference on Computer Supported Cooperative Work in Design*, pp. 1122-1127, 16-18 April 2008.
- [5] A. Cole and A. Albicki, "Nonintrusive Identification of Electrical Loads in a Three-Phase Environment Based on Harmonic Content," *IEEE Conf. on Instrumentation and Measurement Technology*, vol. 1, pp. 24-29, 1-4 May, 2000.
- [6] Y. Nakano, H. Murata, K. Yoshimoto, S. Hidaka, M. Tadokoro, and K. Nagasaka, "Non-Intrusive Electric Appliances Load Monitoring System Using Harmonic Pattern Recognition - Performance Test Results at Real Households," *4th International Conference on Energy Efficiency in Domestic Appliances and Lighting*, pp. 477-488, Jun. 2006.
- [7] Takeshi Saitoh, Yuuki Aota, Tomoyuki Osaki, Ryosuke Konishi, Kazunori Sugahara, "Current Sensor based Non-intrusive Appliance Recognition for Intelligent Outlet," *The 23rd International Technical Conference on Circuits/Systems, Computers and Communications*, pp. 349-352, 11:00-12:30 Monday, July 7, 2008.
- [8] S. B. Leeb, S. R. Shaw, and J. L. Kirtley Jr., "Transient Event Detection in Spectral Envelope Estimates For Nonintrusive Load Monitoring," *IEEE Trans. on Power Delivery*, vol. 10, no. 3, pp. 1200-1210, Jul. 1995.
- [9] S. R. Shaw, S. B. Leeb, L. K. Norford and R. W. Cox, "Nonintrusive Load Monitoring and Diagnostics in Power Systems," *IEEE Transactions on Instrumentation and Measurement*, vol. 57, pp. 1445-1454, July 2008.
- [10] S. R. Kamat, "Fuzzy logic based pattern recognition technique for non-intrusive load monitoring," *IEEE Region 10 Conference on TENCON*, Vol. 3, pp. 528-530, 21-24 Nov. 2004.
- [11] H. T. Yang, H. H. Chang, C. L. Lin, "Design a Neural Network for Features Selection in Non-intrusive Monitoring of Industrial Electrical Loads," *11th International Conference on Computer Supported Cooperative Work in Design*, pp. 1022-1027, 26-28 April 2007.
- [12] H. H. Chang, C. L. Lin, L. S. Weng, "Application of artificial intelligence and non-intrusive energy-managing system to economic dispatch strategy for cogeneration system and utility," *13th International Conference on Computer Supported Cooperative Work in Design*, pp. 740-745, 22-24 April 2009.
- [13] L. X. Wang, *A COURSE IN FUZZY SYSTEMS AND CONTROL*. Pearson Education Taiwan Ltd., 2005.
- [14] Jyh-Shing Roger Jang, "Data Clustering and Pattern Recognition", available at the links for on-line courses at the author's homepage at <http://www.cs.nthu.edu.tw/~jjang>.
- [15] H. Nomura, I. Hayashi, and N. Wakami, "A learning method of fuzzy inference rules by descent method," *IEEE International Conference on Fuzzy Systems*, pp. 203-210, 1992.
- [16] Chin-Teng Lin and C. S. George Lee, *Neural fuzzy systems: A neuro-fuzzy synergism to intelligent systems*. Pearson Education Taiwan Ltd., Taiwan, 2003.
- [17] Mitsuo Gen and Runwei Cheng, *Genetic Algorithms & Engineering Design*. John Wiley & Sons, Inc., 1997.
- [18] Chua, T.W., Tan, W.W., Non-singleton genetic fuzzy logic system for arrhythmias classification. *Engineering Applications of Artificial Intelligence* (2010), doi:10.1016/j.engappai.2010.10.003.