*International Journal of Innovative Computing, Information and Control*
Volume **8**, Number **1(B)**, January **2012**
ICIC International ©2012 ISSN 1349-4198
pp. 617–631

# EVOLVING OPTIMAL FUZZY-CONNECTIVE-BASED HIERARCHICAL AGGREGATION NETWORKS USING GENETIC ALGORITHMS

Chao-Ton Su[1], Fang-Fang Wang[1] and Taho Yang[2]

[1]Department of Industrial Engineering and Engineering Management
National Tsing Hua University
No. 101, Section 2, Kuang-Fu Road, Hsinchu 30013, Taiwan
ctsu@mx.nthu.edu.tw; d9634816@oz.nthu.edu.tw

[2]Institute of Manufacturing Information and Systems
National Cheng Kung University
No. 1, University Road, Tainan 701, Taiwan
tyang@mail.ncku.edu.tw

Received September 2010; revised February 2011

Abstract. *Multilayer fuzzy connective-based hierarchical aggregation networks provide a flexible and intuitive approach to decision analysis. This approach simulates the decision-making processes performed by humans, and the results can be interpreted as a set of rules with which to fashion an abstract model of the problem. Identifying the relative importance of the criteria helps to identify redundancies that do not contribute to the decision-making process. However, a gradient-based learning approach tends to generate local solutions, and requires the aggregation function to be continuous and differentiable. This study proposes a GA-based learning approach to identify the connective parameters, exploiting the global exploration ability of GAs to improve the quality of solutions. This approach does not require gradient information, making it applicable to both differentiable and nondifferentiable aggregation functions. The benefits of this method were demonstrated using eight datasets with different characteristics. Statistical analysis of the experimental results confirms that the proposed approach outperforms the gradient-based learning approach, generating more accurate estimates for both generalized mean and gamma operators. The proposed approach is well suited to a broad range of fuzzy aggregation connectives, which further expands its applicability.*
**Keywords:** Fuzzy connectives, Multilayer hierarchical aggregation, Genetic algorithms (GAs), Decision analysis

## 1. Introduction.

Decision analysis is an important area of research, involving the aggregation of information according to multiple criteria at several levels simultaneously. Researchers have proposed a variety of fuzzy set connectives for the purpose of aggregating information, according to the specific applications involved. Many decision-making situations require a degree of compensation [1], and generalized mean operator and gamma operator can be applied specifically to such situations.

Krishnapuram and Lee [2,3] proposed a fuzzy connective-based aggregation network capable of aggregating and propagating information hierarchically, according to the degree of satisfaction. Their gradient-based learning approach systematically identifies the parameters associated with generalized mean and gamma operators at each node. It is also capable of determining the nature of the connectives and interpreting the resulting network as a set of rules with which to fashion an abstract model of the problem. Identifying the relative importance of criteria can help to identify redundancies that do not contribute to the decision-making process. However, a gradient-based learning approach requires that the aggregation function be continuous and differentiable, and this can be problematic when dealing with discrete variables. When gradient search techniques are applied to complex nonlinear optimization problems, they are prone to becoming stuck in local minima and often produce inconsistent and unpredictable results during training procedures [4].

Genetic algorithms (GAs) are stochastic search techniques, based on biological evolution in which global solutions evolve through the processes of competition and variation. This study proposes a GA-based learning approach to identify the parameters associated with aggregation connectives in multilayer hierarchical networks. In addition to retaining the advantages of fuzzy connective-based networks, our proposed approach exploits the global exploration ability of GAs to increase the likelihood of arriving at a global (or near global) optimum. This approach does not require gradient information, making it applicable to both differentiable and nondifferentiable aggregation functions. This study demonstrates the promising results that can be obtained using a GA-based learning approach.

## 2. **Fuzzy Aggregation Network.**

In a multilayer hierarchical decision process, support for a decision often depends on the degree of satisfaction according to several criteria, which may in turn depend on the degree of satisfaction based on other subcriteria [2]. Figure 1 graphically illustrates this concept, where each node in the network represents a criterion. In a fuzzy aggregation network, the input for each node is information related to each subcriteria and the output is the aggregated evidence evaluated by the fuzzy connectives. The output of one node is the input of the node above.

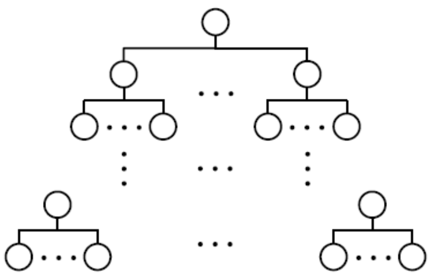

FIGURE 1. Multilayer decision-making architecture

Krishnapuram and Lee [2,3] proposed a gradient-based learning approach to train multilayer fuzzy aggregation networks. This approach is an adaptation of the backpropagation algorithm [5] used to determine the parameters of generalized mean and gamma operators. Forward processing propagates the degree of satisfaction for each input node through the network to the output layer. Beginning with the nodes in the bottom level, the process of aggregation progresses upward through each level until reaching the top of the hierarchy. Backward processing is an optimization process minimizing the degree of output discrepancy in the network. The generalized delta rule uses gradient information to modify weights and parameters. The following section provides the aggregation function and partial derivatives of the generalized mean and gamma operators as they pertain to these weights and parameters.

### 2.1. **Generalized mean operator.**

The generalized mean operator was introduced by [6]. This operator is defined by the following equation

$$f(x_1, x_2, \cdots, x_n; p, w_1, w_2, \cdots, w_n) = \left( \sum_{i=1}^{n} w_i x_i^p \right)^{1/p} \tag{1}$$

where $x_i \in [0,1]$ is the $n$ input to be aggregated, $p$ is the parameter controlling the degree of compensation between intersection and union operators with $p \in \mathrm{R}$ and $p \neq 0$, and $w_i$ is the weight representing the relative importance of various criteria with $w_i \geq 0$ and $\sum_{i=1}^{n} w_i = 1$. To eliminate the constraints on $w_i$, the aggregation function can be written as

$$f = \left( \frac{{w_1}^2}{\sum_{i=1}^{n} w_i^2} x_1^p + \cdots + \frac{w_n^2}{\sum_{i=1}^{n} w_i^2} x_n^p \right)^{1/p} \tag{2}$$

Thus, the partial derivatives of the generalized mean operator are

$$\frac{\partial f}{\partial w_i} = \frac{2w_i}{p\sum_{i=1}^{n} w_i^2} f^{1-p}\left(x_i^p - f^p\right) \tag{3}$$

$$\frac{\partial f}{\partial p} = \frac{f^{1-p}}{p^2}\left( \sum_{i=1}^{n} \frac{w_i^2}{\sum_{i=1}^{n} w_i^2} x_i^p \ln x_i^p - f^p \ln f^p \right) \tag{4}$$

2.2. **Gamma operator.** The gamma operator was introduced by [7,8], and is provided in the following equation

$$y = \left( \prod_{i=1}^{n} x_i^{\delta_i} \right)^{1-\gamma} \cdot \left( 1 - \prod_{i=1}^{n} (1-x_i)^{\delta_i} \right)^{\gamma} \tag{5}$$

where $x_i \in [0,1]$ is the $n$ input to be aggregated, $\delta_i$ represents the weight associated with $x_i$, and $\gamma$ is a parameter controlling the degree of compensation. This parameter indicates where the actual operator is located between the logical 'and' and 'or' with $\sum\limits_{i=1}^{n} \delta_i = n$ and $0 \leq \gamma \leq 1$ [6,9]. The intersection part and union part of the gamma operator can be denoted as

$$y_1 = \prod_{i=1}^{n} x_i^{\delta_i} \tag{6}$$

$$y_2 = 1 - \prod_{i=1}^{n} (1-x_i)^{\delta_i} \tag{7}$$

To eliminate the constraints on $\gamma$ and $\delta_i$, the definition of $\gamma$ and $\delta_i$ can be modified as follows [2,3]:

$$\gamma = \frac{a^2}{a^2+b^2} \quad \text{and} \tag{8}$$

$$\delta_i = \frac{nd_i^2}{\sum_{i=1}^{n} d_i^2} \tag{9}$$

Thus, the partial derivatives of the gamma operator are

$$\frac{\partial y}{\partial a} = \frac{2ab^2}{(a^2+b^2)^2} y \ln\left(\frac{y_2}{y_1}\right) \tag{10}$$

$$\frac{\partial y}{\partial b} = \frac{2ba^2}{(a^2+b^2)^2} y \ln\left(\frac{y_1}{y_2}\right) \tag{11}$$

$$\frac{\partial y}{\partial d_j} = y\frac{2nd_j}{(\sum_{i=1}^{n} d_i^2)^2} \left\{ (1-\gamma)\left[\sum_{i=1}^{n} d_i^2 \ln\left(\frac{x_j}{x_i}\right)\right] + \gamma\left[\frac{y_2-1}{y_2}\sum_{i=1}^{n} d_i^2 \ln\frac{1-x_j}{1-x_i}\right] \right\} \tag{12}$$

Many researchers have discussed fuzzy aggregation networks, applying them in a variety of fields, including safety assessment [10], software quality evaluation [11], structural damage detection [12], fault diagnosis [13] and satellite data classification [14]. Chiang [15] and Chiang and Kuo [16] also proposed gradient-based training algorithms to determine the parameters of Choquet integral and gamma operator.

3. **Genetic Algorithms.** Genetic algorithms are stochastic search and optimization techniques based on the processes associated with biological evolution. Given a specified optimization problem, GAs begin with an initial population of candidate solutions, referred to as chromosomes. These chromosomes evolve over time through the processes of competition and variation. The members of the population that are most likely to produce offspring are those possessing traits favorable to solving the optimization problem. A fitness function assesses the performance of each chromosome indicating how good the solution is. The fitness values of each chromosome are then used to identify which chromosomes reproduce in the competition process, referred to as selection. The higher the fitness value that a chromosome has, the greater its chances of contributing to the next generation are. Genetic operators such as crossover and mutation create new offspring, inheriting a blend of traits from the previous generation. This process is repeated through many generations, with the traits best able to reach a solution being passed on, and undesired traits dying out [17,18].

Genetic algorithms possess a number of advantages over traditional methods, such as their ability to search from population to population, without the limitations imposed by point-to-point search methods. In this manner, searches sweep through the parameter space in many directions simultaneously, greatly enhancing the probability of finding the global optimum [4]. This makes GAs particularly useful for optimizing variables with extremely complex solution surfaces. Genetic algorithms enable the optimization of continuous or discrete variables, without the need for derivative information, making them applicable to a wide range of differentiable, nondifferentiable, continuous, and discontinuous optimization problems encountered in a variety of disciplines [18,19]. Genetic algorithms have enabled researchers to solve many problems, such as the matching and scheduling of extraction, transformation and loading (ETL) tasks [20], overcoming transportation problems [21] and developing medications [22].

4. **Proposed Approach.** This study uses generalized mean and gamma operators to aggregate information at each node in multilayer fuzzy aggregation networks, employing a GA-based learning approach to determine connective parameters. The learning procedures were implemented in the following, and a flowchart is presented in Figure 2.

Step 0. Construct the topology of a fuzzy aggregation network according to the information related to the specific problem.

Step 1. Encode the weights and parameters in the network as chromosomes, represented by a vector of floating point numbers.

Step 2. Select the size of the population and probability of crossover and mutation.

Step 3. Randomly create an initial population of $n$ chromosomes.

Step 4. Aggregate the information upward layer by layer using the aggregation function and evaluate the fitness value of each chromosome in the current population.

Step 5. Create an intermediate population by extracting members from the current population using a selection operator.

Step 6. Generate new offspring by applying the crossover operator to the intermediate population according to the probability of crossover.

Step 7. Mutate new offspring at each position in the chromosome according to the probability of mutation.

Step 8. Adopt an elitist strategy to ensure that the best-performing chromosome always survives.

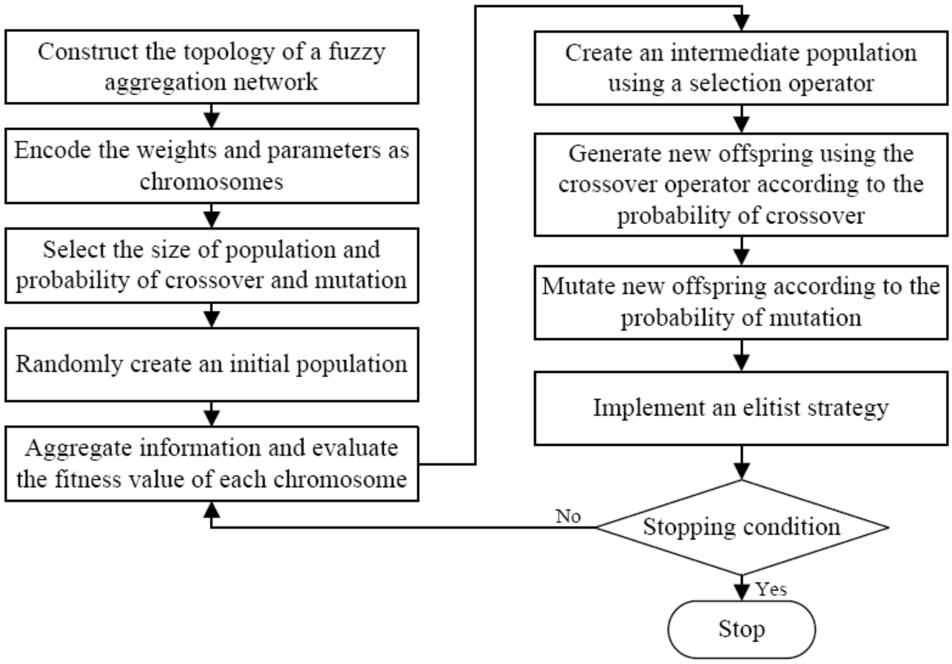


Figure 2. Flowchart of the GA-based learning approach

Step 9. If the conditions required for stopping are met, then stop; otherwise, return to Step 4.

The following section provides detailed information about the processes involved in implementation.

4.1. **Representation of chromosomes.** Genes can be represented directly as real numbers for variables within a continuous domain, to avoid the problems associated with a lack of precision, storage requirements, and computational load incurred through binary representation [23]. In this case, chromosomes are vectors of floating point numbers. For a multilayer fuzzy aggregation network with $L$ layers and $n_l$ nodes in layer $l$, a chromosome comprising the weight of the $i$th node in layer $l-1$ to the $j$th node in layer $l$, $w_{l-1ij}$, and parameter of the $i$th node in layer $l$, $p_{li}$, of the generalized mean operator can be represented as

$$chromosome = [w_{111}, \cdots, w_{L-1n_l n_{l-1}}; p_{11}, \cdots, p_{Ln_l}] \tag{13}$$

A chromosome comprising the weight of the $i$th node in layer $l-1$ to the $j$th node in layer $l$, $d_{l-1ij}$, and the parameters of the $i$th node in layer $l$, $a_{li}$ and $b_{li}$, of the gamma operator can be represented as

$$chromosome = [d_{111}, \cdots, d_{L-1n_l n_{l-1}}; a_{11}, \cdots, a_{Ln_l}; b_{11}, \cdots, b_{Ln_l}] \tag{14}$$

4.2. **Fitness function.** The fitness function assesses the performance of each chromosome to evaluate the quality of the solution. This study selected $1/m_e$ as a fitness function, where $m_e$ is the root mean squared error (RMSE) defined as

$$RMSE = \sqrt{\frac{\sum_{k=1}^{n}\sum_{p=1}^{m}(Y_{pk} - y_{pk})^2}{n \times m}} \tag{15}$$

where $n$ denotes the number of training patterns, $m$ is the number of outputs, and $Y_{pk}$ and $y_{pk}$ are the target and actual value of output $p$ and input pattern $k$, respectively.

The best individual in the population is that with the maximum fitness value, i.e., the minimum error.

4.3. **Initial population.** A group of chromosomes represents the population. The initial generation of the population and the size are two important aspects of populations used in GAs [24]. Consider an initial population of $N_{pop}$ chromosomes. For an $N_{var}$ dimensional optimization problem, the population is represented by an $N_{pop} \times N_{var}$ matrix filled with uniform random numbers [19]. For each problem, the size of the population depends on the complexity of the problem. The larger the population is, the easier it is to explore the search space. However, this requires much higher computational cost, memory, and time. The size of the population can be altered according the available time, computer memory, and desired quality of the results [24].

4.4. **Selection.** Selection is a method in which chromosomes are randomly selected from the population with an emphasis on the fitness of individuals [24]. Consider $P$ a population with chromosomes $C_1, \ldots, C_N$. The selection mechanism creates an intermediate population, $P'$, to produce copies of chromosomes in $P$. The higher the fitness value, the greater the chance that a chromosome will contribute copies to $P'$. The selection mechanism comprises two steps: calculating the probability of selection and implementing the sampling algorithm [17].

- **Calculating the Probability of Selection.** Proportional selection is the best known and most commonly used selection mechanism [25,26]. For each chromosome $C_i$ in $P$ where $i = 1, \ldots, N$, the probability of including a copy of $C_i$ in $P'$ can be calculated using Equation (16) [17].

$$p_s(C_i) = \frac{f(C_i)}{\sum_{j=1}^{N} f(C_j)} \tag{16}$$

  where $f(C_i)$ is the fitness value of chromosome $C_i$.
- **Implementing the Sampling Algorithm.** Copies of chromosomes are reproduced to form $P'$ according to the probability of being selected. Stochastic sampling with replacement, referred to as roulette wheel selection, is a traditional selection technique [25,26]. The principle of roulette wheel selection involves a linear search through each chromosome $C_i$ in the population mapped onto a roulette wheel with a space proportionally to $p_s(C_i)$. With each spin, an individual chromosome is selected for inclusion in the intermediate population of the next generation. The roulette wheel is spun repeatedly until all available positions in $P'$ are filled [17].

4.5. **Crossover.** Crossover is a method used to share information between chromosomes according to the probability of crossover, $p_c$. The traditional crossover operator in GAs is the simple crossover [27,28]. First, chromosomes in the intermediate population $P'$ are randomly grouped into pairs. For each pair, assume two chromosomes have been selected for the application of the crossover operator:

$$C_1 = \left(c_1{}^1 \ldots c_n^1\right) \tag{17}$$

$$C_2 = \left(c_1{}^2 \ldots c_n^2\right) \tag{18}$$

A crossover site $i \in \{1, 2, \ldots, n-1\}$ is then selected at random along the length of the string. Finally, position values are swapped between the two strings according to the crossover site to construct two new chromosomes based on Equations (19) and (20).

$$H_1 = \left(c_1^1, c_2^1, \ldots, c_i^1, c_{i+1}^2, \ldots, c_n^2\right) \tag{19}$$

$$H_2 = \left(c_1^2, c_2^2, \ldots, c_i^2, c_{i+1}^1, \ldots, c_n^1\right) \tag{20}$$

4.6. **Mutation.** Mutation is a method used to recover lost or unexplored genetic material in the population according to the probability of mutation, $p_m$. Random mutation is the most commonly used mutation mechanism for the representation of real-values. This method involves the replacement of selected chromosome with a value uniformly selected from the parameter space [4,23]. Assume that $C = (c_1, \ldots, c_i, \ldots, c_n)$ is a chromosome and $c_i \in [a_i, b_i]$ is the gene to be mutated. The gene $c_i'$ resulting from the application of mutation operators is a uniform random number from the domain $[a_i, b_i]$ [28].

4.7. **Elitist strategy.** An elitist strategy may be adopted following the implementation of crossover and mutation [29]. This strategy involves copying either the best chromosome or a few of the best chromosomes from each generation to the succeeding generation. This method ensures that the best-performing chromosome always survives from one generation to the next. This is necessary to avoid the loss of the best chromosome due to crossover or mutation [17].

4.8. **Stopping conditions.** Searches are terminated at the end of each generation if stopping conditions are satisfied. Examples of stopping conditions include reaching a specified number of generations, a lack of observable change in the best fitness value over a specified number of generations, or a situation in which all chromosomes and associated fitness values would converge, if it were not for mutation.

5. **Experimental Study.** This study conducted experiments on eight datasets to evaluate the effectiveness of the GA-based learning approach in a fuzzy connective-based hierarchical aggregation network. The following section presents details of the datasets, implementation of the experiment, and a comparison of the results.

5.1. **Datasets.** The eight datasets used in this paper varied in terms of the numbers of criteria, subcriteria and volume of data. Table 1 summarizes the characteristics of the datasets with detailed information provided below.

TABLE 1. Summarized characteristics of dataset

| Dataset | No. of Criteria | No. of Subcriteria | Size |
|---|---|---|---|
| Abalone | 3 | 10 | 4177 |
| Bank | 4 | 8 | 8192 |
| Investment | 11 | 35 | 29 |
| Society | 9 | 79 | 51 |
| Economy | 6 | 27 | 51 |
| Politics | 5 | 35 | 51 |
| Environment | 7 | 22 | 51 |
| Health and Education | 4 | 29 | 51 |

**Example 5.1.** *The growing awareness of the profound influence that humans have on marine systems has been the impetus for increased advocacy on the part of marine conservation. Faced with strict size limits, the age and growth of fish populations must be accurately estimated to protect the reproductive capacity of stocks and preserve biodiversity. Many researchers have attempted to develop techniques for estimating the age of abalone [30]. The "Abalone" dataset used in this study was obtained from the UCI machine learning repository, including 4177 instances. The age of abalone was estimated according to 3 criteria comprising 10 subcriteria. Figure 3 illustrates the network architecture of this dataset.*

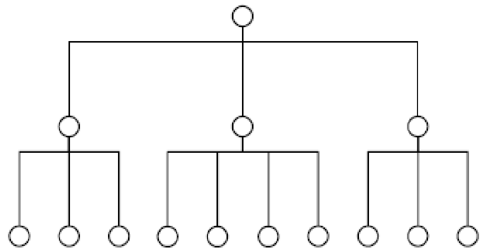

Figure 3. Network architecture of the "Abalone" dataset

**Example 5.2.** *Research into service management and marketing has provided a strong theoretical base on which to pursue an empirical exploration of the relationships among customer satisfaction, customer loyalty and profitability [31]. The "Bank" dataset used in this study was obtained from the Delve datasets, including 8192 instances. Data were generated from a simulation of queues in a series of banks. In this scenario, customers came from various residential areas, selecting their preferred bank according to distance from their home. They had to perform tasks of varying complexity and were equipped with various levels of patience. Customers were liable to change queues when they lost patience. The rejection rate in this problem was determined by 4 criteria comprising 8 subcriteria. Figure 4 illustrates the network architecture of this dataset.*

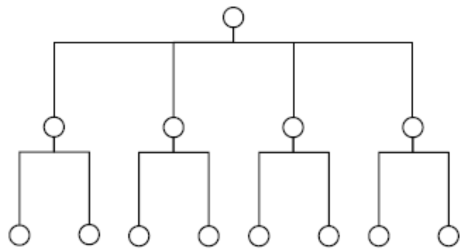

Figure 4. Network architecture of the "Bank" dataset

**Example 5.3.** *With increases in globalization, enterprises are continually seeking suitable overseas investment opportunities to maintain long-term competitiveness. This example, obtained from [32], investigated the experiences of investors in the electrical and electronic appliance industry in Taiwan, concerning their evaluation of the investment environment of Mainland China and Southeast Asia. The degree of satisfaction (subjective) investors felt for the investment environment was aggregated through an hierarchical decision-making structure to produce an overall satisfaction value. Data related to the investment environment were collected from 29 questionnaires, and evaluated according to 11 criteria comprising 35 subcriteria. Figure 5 depicts the network architecture of this dataset.*

The following section introduces 5 datasets related to the evaluation of quality of life (QOL). Quality of life generally refers to the level of satisfaction or happiness experienced by an individual. Quality of life measures are used to assess social progress and social accounting, making them useful for national goal setting, the evaluation of programs, and ranking of priorities. Quality of life is an output of an aggregate function linking a group of quantifiable factors, used to statistically assess the health status of the nation. Quality of life can be evaluated from various aspects related to national welfare, each of which represents a major national objective, in areas pertaining to society, economics, the

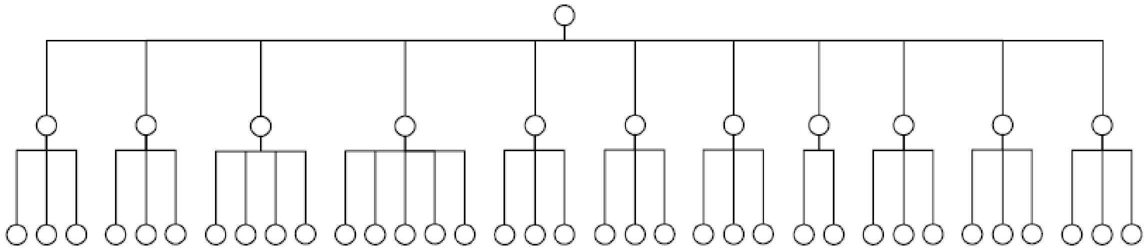

FIGURE 5. Network architecture of the "Investment" dataset

environment, politics, and health and education [33]. This study analyzed data obtained from [33], covering 50 states and the District of Columbia-Washington metropolitan area in the United States. Examples 5.4 through 5.8 provide detailed information, evaluating various aspects of the QOL.

**Example 5.4.** *Due to the wide range of social concerns, evaluation of the social aspect of QOL encompasses many factors reflecting social issues, including equality, personal development and community living conditions. This study evaluated the social aspects pertaining to QOL by measuring 9 criteria comprising 79 subcriteria. Figure 6 illustrates the network architecture of this dataset.*

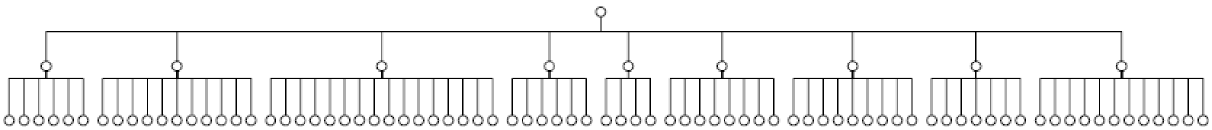

FIGURE 6. Network architecture of the "Society" dataset

**Example 5.5.** *Inputs for modeling QOL from an economic perspective include variables related to the economic well-being of individuals and economic health of the community. All selected variables in this study were designed to measure either the command over goods and services or the capability of satisfying basic needs to provide a better quality of life for the entire population within each state. This study evaluates QOL data from an economic perspective by measuring 6 criteria comprising 27 subcriteria. Figure 7 shows the network architecture of this dataset.*

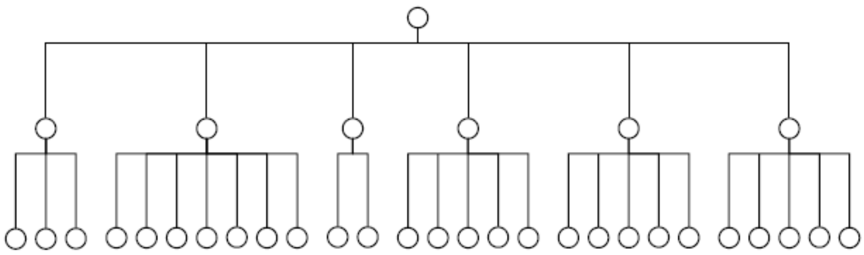

FIGURE 7. Network architecture of the "Economy" dataset

**Example 5.6.** *Evaluating QOL from a political perspective deals with institutional factors and functional operations of a democratic system, in which all individuals are organized within a community to achieve a common goal or public objectives. This study measured QOL from this perspective by measuring 5 criteria comprising 35 subcriteria. Figure 8 shows the network architecture of this dataset.*

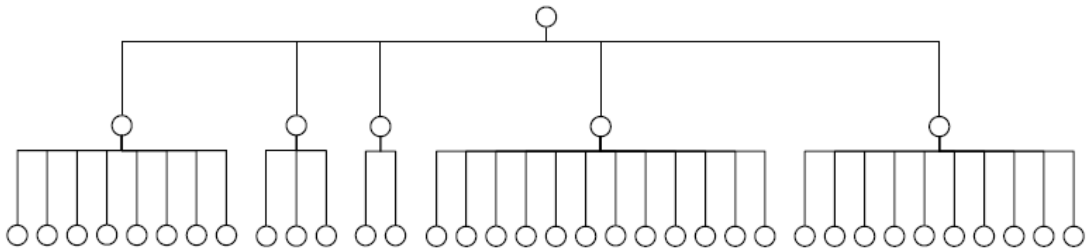

Figure 8. Network architecture of the "Politics" dataset

**Example 5.7.** *Humans use natural resources to satisfy various needs and desires and achieve the goals of economic growth. However, energy related crises and environmental problems are indications that the degree to which nature can be exploited without repercussions is limited. Thus, the protection of the environment and conservation of natural resources have become focal points of public interest and national concern. Evaluating QOL from an environmental perspective can provide a better understanding of environmental issues. This study evaluated this perspective by measuring 7 criteria comprising 22 subcriteria. Figure 9 depicts the network architecture of this dataset.*

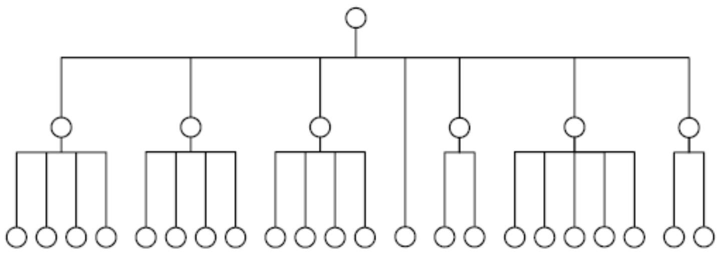

Figure 9. Network architecture of the "Environment" dataset

**Example 5.8.** *The quality of health and education is another principal concern in the quality of life accounting systems. Three primary health concerns are: longevity, freedom from disability, and the availability and accessibility of medical care. Achieving a basic level of education and the opportunity to pursue goals of higher learning are the primary concerns of intellectual health. This study evaluated the conditions of health and education among individuals and the community by measuring 4 criteria comprising 29 subcriteria. Figure 10 shows the network architecture of this dataset.*

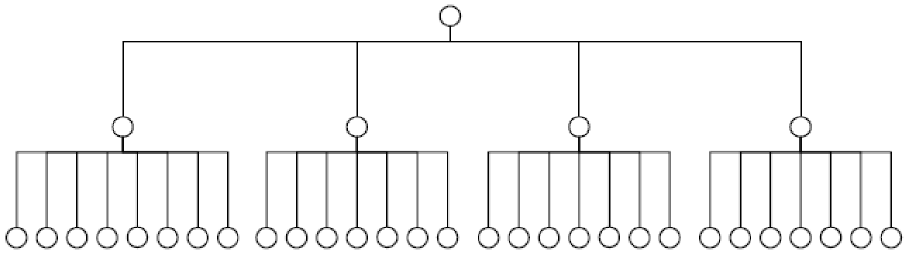

Figure 10. Network architecture of the "Health and Education" dataset

5.2. **Implementation.** This section compares the performance of the gradient-based learning approach in multilayer aggregation networks with the proposed learning approach. Experiments were conducted using a programming environment developed in C# on a Windows XP operating system, a 1.6 GHz dual CPU and 3.0 GB of RAM. Data were transformed within the range [0, 1]. The network topology was modeled according to the structure of the criteria and subcriteria pertaining to each problem. For each model, the initial weights and parameters were established randomly within the range [–1, 1]. In addition, we adopted a 10 five-fold cross-validation strategy to increase the accuracy and reliability of the results.

The learning rate of the gradient-based approach was optimized by trial and error to determine the setting with the minimum RMSE, and the test range gradually increased from 0.05 to 1.0. The training in the gradient-based learning approach was terminated when a specified number of generations was reached or changes in accuracy, weights, or parameters were less than $10^{-8}$. The combination of parameters in the GA-based learning approach was tested using the range recommended in the literature [18,24,34]. We set the size of the population at 100, probability of crossover at 100%, probability of mutation below 20%, and kept the best 10% of the chromosomes from each generation for the elitist strategy. The training of the GA-based learning approach was terminated when a specified number of generations had been reached.

Figure 11 shows the training process for the gamma operator in Example 5.3 with the gradient-based and GA-based learning approaches. It can be observed that the GA-based learning approach provided superior results in the beginning and at the end of training. This can be attributed to the fact that GAs search for solutions in many directions simultaneously and the evolution mechanism helps to focus the search in better regions of the solution space. Table 2 summarizes the experimental results, including the average, best, worst and standard deviation (SD) for RMSE. The GA-based learning approach had a smaller RMSE in the average, best and worst cases, indicating that the proposed approach can provide more accurate estimates. The GA-based learning approach also had smaller variations in RMSE, demonstrating that the proposed approach is capable of generating more stable and reliable solutions for both operators.

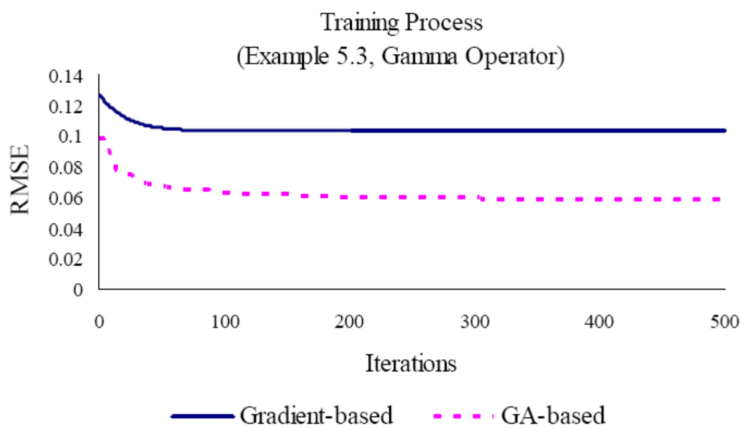


FIGURE 11. Training process for the gamma operator in Example 5.3 using gradient-based and GA-based learning approaches

5.3. **A comparison.** This study was concerned whether there is a statistically significant difference in the accuracy of evaluation between the gradient-based and proposed learning approaches. Therefore, this study adopted two-way ANOVA for the block design using

Table 2. Summarized experiment results

| Dataset | Operator | Learning Approach | RMSE | | | |
|---|---|---|---|---|---|---|
| | | | Mean | Best | Worst | SD |
| Abalone | Generalized Mean | Gradient-based | 0.1873 | 0.1378 | 0.3257 | 0.0556 |
| | | GA-based | 0.1004 | 0.0970 | 0.1038 | 0.0022 |
| | Gamma | Gradient-based | 0.0944 | 0.0877 | 0.1396 | 0.0159 |
| | | GA-based | 0.0877 | 0.0867 | 0.0888 | 0.0006 |
| Bank | Generalized Mean | Gradient-based | 0.1337 | 0.1224 | 0.1493 | 0.0089 |
| | | GA-based | 0.0986 | 0.0923 | 0.1005 | 0.0025 |
| | Gamma | Gradient-based | 0.0685 | 0.0683 | 0.0687 | 0.0002 |
| | | GA-based | 0.0684 | 0.0682 | 0.0690 | 0.0002 |
| Investment | Generalized Mean | Gradient-based | 0.0998 | 0.0860 | 0.1300 | 0.0140 |
| | | GA-based | 0.0966 | 0.0802 | 0.1116 | 0.0103 |
| | Gamma | Gradient-based | 0.1447 | 0.0906 | 0.2861 | 0.0691 |
| | | GA-based | 0.1054 | 0.0949 | 0.1392 | 0.0127 |
| Society | Generalized Mean | Gradient-based | 0.1519 | 0.1245 | 0.1802 | 0.0178 |
| | | GA-based | 0.1495 | 0.1409 | 0.1681 | 0.0090 |
| | Gamma | Gradient-based | 0.4239 | 0.3073 | 0.4918 | 0.0571 |
| | | GA-based | 0.1130 | 0.1053 | 0.1267 | 0.0075 |
| Economy | Generalized Mean | Gradient-based | 0.1417 | 0.1134 | 0.1566 | 0.0134 |
| | | GA-based | 0.1392 | 0.1271 | 0.1650 | 0.0123 |
| | Gamma | Gradient-based | 0.2739 | 0.1198 | 0.3877 | 0.0804 |
| | | GA-based | 0.1122 | 0.1014 | 0.1289 | 0.0082 |
| Environment | Generalized Mean | Gradient-based | 0.1475 | 0.1081 | 0.1930 | 0.0312 |
| | | GA-based | 0.1427 | 0.1254 | 0.1634 | 0.0141 |
| | Gamma | Gradient-based | 0.2277 | 0.1756 | 0.2712 | 0.0325 |
| | | GA-based | 0.1257 | 0.1062 | 0.1427 | 0.0127 |
| Politics | Generalized Mean | Gradient-based | 0.1710 | 0.1514 | 0.2140 | 0.0189 |
| | | GA-based | 0.1459 | 0.1344 | 0.1756 | 0.0126 |
| | Gamma | Gradient-based | 0.3423 | 0.2554 | 0.4993 | 0.0787 |
| | | GA-based | 0.1364 | 0.1293 | 0.1447 | 0.0051 |
| Health & Education | Generalized Mean | Gradient-based | 0.1607 | 0.1313 | 0.2295 | 0.0268 |
| | | GA-based | 0.1576 | 0.1357 | 0.1954 | 0.0173 |
| | Gamma | Gradient-based | 0.2443 | 0.1324 | 0.3721 | 0.0773 |
| | | GA-based | 0.1264 | 0.1165 | 0.1445 | 0.0085 |

ten replications. We adopted the learning approach as the "factor" and the dataset as the "block", running ten replications for each combination of factor and block. Tables 3 and 4 show the results of ANOVA for the generalized mean and gamma operators, respectively. The $p$-values of learning approaches in the ANOVA tables are smaller than 0.05, indicating a statistically significant difference between the accuracy of gradient-based and GA-based learning approaches for both aggregation operators at a 95% confidence level. Figure 12 shows the 95% confidence interval in the accuracy of evaluation using the gradient-based and GA-based learning approaches for generalized mean and gamma operators. Both intervals of the GA-based learning approach had smaller mean estimates of RMSE and widths of confidence intervals, particularly for the gamma operator. The above statistical analysis provides sufficient evidence to conclude that the proposed approach is capable of generating more accurate estimates with a smaller degree of variation.

6. **Conclusions.** Multilayer fuzzy connective-based hierarchical aggregation networks provide a flexible and intuitive method for decision analysis. The approach is similar

TABLE 3. ANOVA for the generalized mean operator

| Source | Df. | SS | MS | F | $p$ |
|---|---|---|---|---|---|
| Approach (Factor) | 1 | 0.0167 | 0.0167 | 38.60 | 0.00 |
| Dataset (Block) | 7 | 0.0633 | 0.0090 | 20.93 | 0.00 |
| Interaction | 7 | 0.0307 | 0.0044 | 10.15 | 0.00 |
| Error | 144 | 0.0623 | 0.0004 | | |
| Total | 159 | 0.1730 | | | |

TABLE 4. ANOVA for the gamma operator

| Source | Df. | SS | MS | F | $p$ |
|---|---|---|---|---|---|
| Approach (Factor) | 1 | 0.5574 | 0.5574 | 312.17 | 0.00 |
| Dataset (Block) | 7 | 0.6772 | 0.0967 | 54.18 | 0.00 |
| Interaction | 7 | 0.3977 | 0.0568 | 31.82 | 0.00 |
| Error | 144 | 0.2571 | 0.0018 | | |
| Total | 159 | 1.8894 | | | |

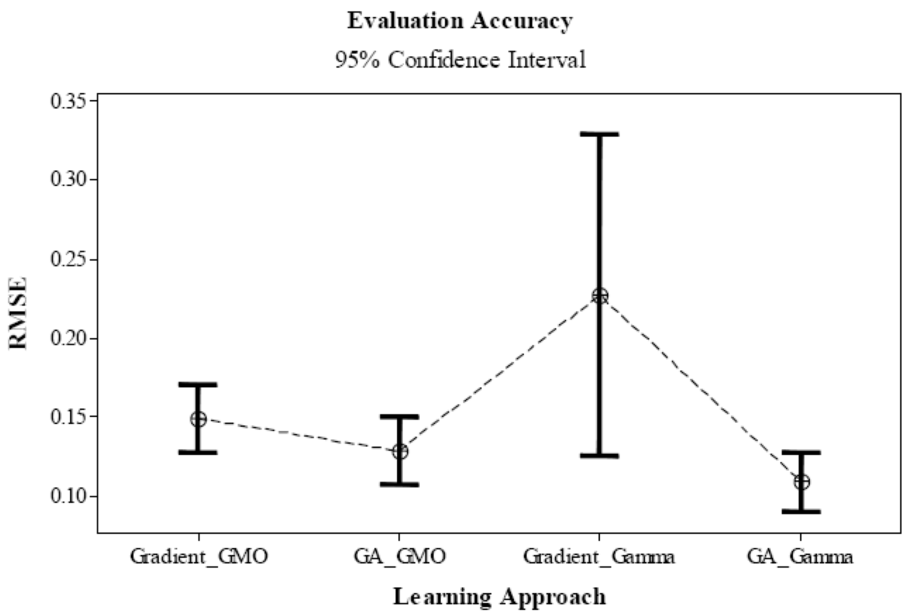


FIGURE 12. 95% confidence interval of the accuracy mean

to the decision-making processes performed by humans, providing a number of attractive features. However, the gradient-based learning approach requires that the aggregation function be continuous and differentiable, requiring considerable computation to adjust the weights and parameters, and learning speeds are slow when discrepancy surfaces are flatter [3]. Problems involving a high degree of complexity in the network architecture tend to obtain local solutions. To enhance effectiveness and applicability, this study proposed a GA-based learning approach to determine weights and parameters, taking advantage of GAs global exploration to improve the quality of solutions. The proposed approach retains the advantages of fuzzy connective-based networks and does not require gradient information, making it applicable to a wider range of aggregation functions. The effectiveness of the proposed learning approach was demonstrated using eight datasets with different characteristics. Statistical analyses of the experimental results in Sections 5.2 and

5.3 indicate statistically significant differences between the GA-based and gradient-based learning approaches for both generalized mean and gamma operators. The GA-based learning approach outperforms the gradient-based learning approach, and is capable of generating estimates of higher accuracy with less variation.

This study makes three main contributions to the literature. First, we successfully implemented a heuristic learning approach using GAs to determine the connective parameters in multilayer hierarchical aggregation networks. This approach preserves the advantages of fuzzy connective-based networks, while improving the quality of the solutions by increasing the likelihood of arriving at the global optimum. The GA-based learning approach is capable of providing a higher degree of accuracy and more reliable solutions than the gradient-based learning approach. Second, the proposed approach is widely applicable, and not limited to continuous or differential aggregation functions. Additional fuzzy aggregation connectives can be applied, thereby expanding the applicability of this approach. Third, the GA-based learning approach requires less mathematical computation than the gradient-based approach does. The mechanism integrating evolution with search solutions from many directions simultaneously helps focus quickly on better regions of the solution space thereby achieving superior results.

The GA-based learning approach has good global search capability, enabling it to locate the region of a better solution quickly. However, it is difficult to find the exact location of the global optimum using the processes of crossover or mutation. Hybridization with a local optimization step may help to obtain more ideal solutions. In addition, the size of the population and probability of crossover and mutation influence the results. Optimizing the settings for the GA-based learning approach and the effect of combination with other aggregation connectives are issues for future research.

## REFERENCES


[1] D. Dubois and H. Prade, A review of fuzzy set aggregation connectives, *Information Sciences*, vol.36, pp.85-121, 1985.

[2] R. Krishnapuram and J. Lee, Fuzzy-connective-based hierarchical aggregation networks for decision making, *Fuzzy Sets and Systems*, vol.46, pp.11-27, 1992.

[3] R. Krishnapuram and J. Lee, Fuzzy-set-based hierarchical networks for information fusion in computer vision, *Neural Networks*, vol.5, pp.335-350, 1992.

[4] R. S. Sexton, R. E. Dorsey and J. D. Johnson, Toward global optimization of neural networks: A comparison of the genetic algorithm and backpropagation, *Decision Support Systems*, vol.22, pp.171-185, 1998.

[5] D. E. Rumelhart, J. M. McClelland and the PDP Research Group, *Parallel Distributed Processing*, The MIT Press, Cambridge, 1986.

[6] H. Dyckhoff and W. Pedrycz, Generalized means as a model of compensation connectives, *Fuzzy Sets and Systems*, vol.14, pp.143-154, 1984.

[7] H. J. Zimmermann and P. Zysno, Latent connectives in human decision making, *Fuzzy Sets and Systems*, vol.4, pp.37-51, 1980.

[8] H. J. Zimmermann and P. Zysno, Decision and evaluations by hierarchical aggregation of information, *Fuzzy Sets and Systems*, vol.10, pp.243-260, 1983.

[9] H. J. Zimmermann, *Fuzzy Set Theory and Its Applications*, 2nd Edition, Maw Chang Book Company, Taipei, 1991.

[10] Z. Yang and Z. Zhang, Hierarchical network-based safety assessment decision support system for thermal power plants, *International Journal of Innovative Computing, Information and Control*, vol.6, no.7, pp.2955-2964, 2010.

[11] L. Lin and H. M. Lee, A new algorithm of software quality evaluation for user satisfaction, *International Journal of Innovative Computing, Information and Control*, vol.4, no.10, pp.2639-2647, 2008.

[12] H. Y. Guo, Structural damage detection using information fusion techniques, *Mechanical Systems and Signal Processing*, vol.20, pp.1173-1188, 2006.

[13] R. Q. Li, J. Chen and X. Wu, Fault diagnosis for rotating machinery using knowledge-based fuzzy neural network, *Applied Mathematics and Mechanics*, vol.27, no.1, pp.99-108, 2006.

[14] P. S. Sangle and S. M. George, Alternate neural network models as supervised classifiers for satellite data, *Journal of Environmental Informatics*, vol.6, no.2, pp.80-92, 2005.

[15] J. H. Chiang, Choquet fuzzy integral-based hierarchical networks for decision analysis, *IEEE Transactions on Fuzzy Systems*, vol.7, no.1, pp.63-71, 1999.

[16] J. H. Chiang and Y. H. Kuo, Learning decision functions in the fuzzy $\gamma$-models, *International Journal on Artificial Intelligence Tools*, vol.9, no.4, pp.459-471, 2000.

[17] F. Herrera, M. Lozano and J. L. Verdegay, Tackling real-coded genetic algorithms: Operators and tools for behavioral analysis, *Artificial Intelligence Review*, vol.12, pp.265-319, 1998.

[18] R. E. Dorsey and W. J. Mayer, Genetic algorithms for estimation problems with multiple optima, nondifferentiability, and other irregular features, *Journal of Business & Economic Statistics*, vol.13, no.1, pp.53-66, 1995.

[19] R. L. Haupt and S. E. Haupt, *Practical Genetic Algorithms*, 2nd Edition, John Wiley & Sons, New Jersey, 2004.

[20] X. Song and X. Liu, ETL tasks matching and scheduling based on genetic algorithm, *ICIC Express Letters*, vol.4, no.4, pp.1245-1250, 2010.

[21] F.-T. Lin, Performance comparison of differential evolution and genetic algorithms for the fuzzy transportation problem, *ICIC Express Letters*, vol.4, no.6(B), pp.2469-2474, 2010.

[22] H.-Y. Chung, S.-C. Ou and C.-Y. Chung, Drug design via novel directivity genetic algorithms and Lyapunov principle, *International Journal of Innovative Computing, Information and Control*, vol.5, no.7, pp.2061-2070, 2009.

[23] M. C. Su and H. T. Chang, Application of neural networks incorporated with real-valued genetic algorithms in knowledge acquisition, *Fuzzy Sets and Systems*, vol.112, pp.85-97, 2000.

[24] S. N. Sivanandam and S. N. Deepa, *Introduction to Genetic Algorithms*, Sprnger-Verlag, New York, 2008.

[25] J. H. Holland, *Adaption in Natural and Artificial Systems*, The University of Michigan Press, Ann Arbor, 1975.

[26] D. E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison-Wesley, New York, 1989.

[27] A. H. Wright, Genetic algorithms for real parameter optimization, in *Foundations of Genetic Algorithms*, G. J. E. Rawlin (ed.), San Mateo, Morgan Kaufmann Publishers, 1991.

[28] Z. Michalewicz, *Genetic Algorithms + Data Structures = Evolution Programs*, Sprnger-Verlag, New York, USA, 1992.

[29] K. A. De Jong, *An Analysis of the Behavior of a Class Genetic Adaptive Systems*, Ph. D. Thesis, University of Michigan, 1975.

[30] L. J. Gurney, C. Mundy and M. C. Porteus, Determining age and growth of abalone using stable oxygen isotopes: A tool for fisheries management, *Fisheries Research*, vol.72, pp.353-360, 2005.

[31] R. Hallowell, The relationships of customer satisfaction, customer loyalty, and profitability: An empirical study, *International Journal of Service Industry Management*, vol.7, no.4, pp.27-42, 1996.

[32] L. H. Chen and D. C. Heh, Evaluation for overseas investment environments: The case of evaluating the Mainland China by electrical and electronic appliance industry, *Chiao Ta Management Review*, vol.15, vol.2, pp.107-127, 1995.

[33] B. C. Liu, *Energy, Income, and Quality of Life Management in U.S.A.: An Information Systems Approach to Decision Analysis*, Tamkang University, Taipei, 1988.

[34] Z. Liu, A. Liu, C. Wang and Z. Niu, Evolving neural network using real coded genetic algorithm (GA) for multispectral image classification, *Future Generation Computer Systems*, vol.20, pp.1119-1129, 2004.