A Combined Multi-Objective Memetic Algorithm and ANFIS for Heat Stroke Prediction

Paranya Palwisut[1]

[1] Department of Information Technology,
Faculty of Science and Technology, Nakhon Pathom Rajabhat University,
Nakhon Pathom, Thailand
E-mail: paranya@npru.ac.th

บทคัดย่อ—การทำนายความเสี่ยงในการเกิดโรคลมร้อนเป็นปัญหาหนึ่งที่ต้องการความแม่นยำในการจำแนก เพื่อนำมาใช้เป็นเครื่องมือช่วยในการเฝ้าระวังการเกิดโรคลมร้อน งานวิจัยนี้จึงได้นำเสนอระบบผสมผสานบนพื้นฐานขั้นตอนมีมีติกแบบหลายวัตถุประสงค์ร่วมกับระบบอนุมานนิวโรฟัซซีแบบปรับตัวได้เรียกโดยย่อว่า MOMA-ANFIS เมื่อนำ MOMA-ANFIS ทดสอบแก้ปัญหาในการทำนายความเสี่ยงในการเกิดโรคลมร้อนพบว่าผลการทดลองให้อัตราความถูกต้องของการจำแนกข้อมูลที่สูงถึง 98.51% ซึ่งสูงกว่าการทดลองด้วยแบบจำลอง ANFIS แบบดั้งเดิม และจำนวนกฎของระบบลดลงทำให้โครงสร้างกฎของฟัซซีมีความซับซ้อนลดลง

คำสำคัญ: มีมีติกแบบหลายวัตถุประสงค์, ระบบอนุมานนิวโรฟัซซีแบบปรับตัวได้, โรคลมร้อน

Abstract—Heat stroke risk prediction is a problem that demands high classification accuracy. It is a tool to help prevent heat stroke occurrence. This research introduced a prediction system based on a combination of Multi-Objective Memetic Algorithm (MOMA) and Adaptive Neuro-Fuzzy Inference System (ANFIS) in brief MOMA-ANFIS. When MOMA-ANFIS was applied to solve the problem in the prediction of heat stroke risks, it was found that the accuracy rate of the classification test result was as high as 98.51% which was greater than the rate obtained when the traditional ANFIS. As the number of rules decreased, the fuzzy rule architecture became less complicated.

Keywords— Multi-Objective Memetic Algorithm, Adaptive Neuro-Fuzzy Inference System, Heatstroke

I. INTRODUCTION

Problems in general require accurate classification of big data. Heat stroke risk prediction is a problem that demands high classification accuracy. It is a tool to help prevent heat stroke occurrence. Over the past years, the climate in Thailand has changed dramatically and the temperature tends to increase as a result of global warming or the increasing average global temperature. Human unfortunately suffers from heat exhaustion conditions due to the climate changes. Heat injury often occurs in summer time and causal factors could be: unfamiliarity to a tropical environment; a physical training in a very hot weather; no breeze and high humid climates; thick or cover-up clothes; illness; over exercise or training; long hour outdoor work; insufficient intake of water; overeating; alcohol drinking before exercise; and intake of diuretics and anti-sweating medication. In serious cases, heat stroke can cause death or disability.

Adaptive Neuro-Fuzzy Inference System (ANFIS) is an approach commonly referred to and applied in data classification. It is an adaptive system that integrates the benefits of Artificial Neural Network and Neural Fuzzy System to enhance each other's function and reduce individual limitations. Nonetheless, a problem with big data classification means a considerable number of fuzzy if-then rule nodes and as a result a more complex system [1]. Researches have been conducted in the attempt to reduce nodes in ANFIS rules in order to create rules with higher data classification efficiency. The development of a fuzzy data classification model aims to increase the efficiency, in other words, the most accurate method (or minimum error), and to reduce the complexity of fuzzy rules, that is the least number of rules possible [2].

Researches in the past have revealed that evolutionary optimization algorithms have been widely used to reduce rule nodes in layers and increase data classification efficiency. The consequence is a more accurate prediction. For instance, Stefanie and member

[3] integrated a genetic algorithm with the neural fuzzy system to use for disease diagnosis. Furthermore, multi-objective evolutionary algorithms are employed to reduce fuzzy rule nodes as well [2], [4]-[7].

Memetic Algorithm and Multi-Objective Memetic Algorithm or so-called MOMA is a popular evolutionary optimization algorithm used by researchers for problem optimization as mentioned in some research works, e.g., [8, 9].

Based on the above findings, this research mixes the MOMA and ANFIS in brief MOMA-ANFIS, to boost the data classification efficiency and reduce system rules, and employs it in the classification of heat stroke risks which will become a preventive tool for its occurrence. The approach integrates a combined data classification model with a presentation of an easy-to-comprehend knowledge as a result of the acquisition of more critical rules and higher classification performance. Also, the research could be a guidance for future researchers to improve data classification techniques.

II. LITERATURE REVIEW

Multi-objective problems [10] are multi-objective optimization design problems, consisting of m Objective and Decision Variables as specified in Equation (1)

$$\text{Minimize (or maximize): } \{ f_1(\mathbf{x}), f_2(\mathbf{x}), \ldots, f_m(\mathbf{x}) \} \qquad (1)$$

When $\mathbf{x}$ is vector of the decision variables,
$f_i(\mathbf{x})$ is the objective function i, $i = 1, 2, \ldots, m$

The result of the equation is called Pareto Front or the set of all non-dominated parameterizations as illustrated in Figure 1.

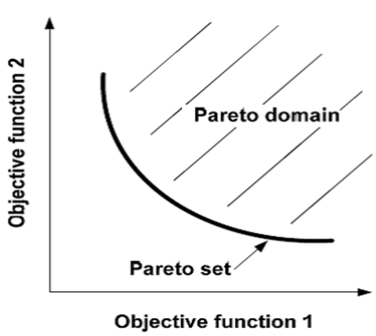


Figure 1. Pareto front of objective function f1 and f2

(Source: Kim, K. *et al*. [11])

ANFIS [1] is a multi-layer artificial neural network that works like the fuzzy inference system but can adapt parameter setting of if-then rule and define the membership function of system input. ANFIS structure comprises a hidden input layer and output layers. The hidden layer substitutes a membership function and fuzzy if-then rule. The architecture of ANFIS is illustrated in Figure 2. Rectangles depict nodes with adjustable parameter and round shapes depict nodes of which parameter is not adjustable.

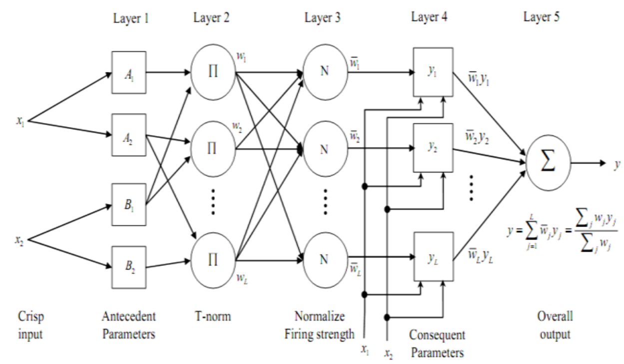


Figure 2. ANFIS Architechture (Source: Meesad, P. [12])

The ANFIS architecture in Figure 2 shows two inputs and each input has two fuzzy sets. Input 1 is composed of A_1 and A_2 while Input 2 consists of B_1 and B_2. The given architecture is based on Takagi Sugeno Kang Model (TSK). ANFIS has 5 layers, i.e.

Layer 1: Antecedent Parameter is the initial parameter of fuzzy rule.

Layer 2: T-norm Operator is connecting fuzzy values between inputs.

Layer 3: Normalized Firing Strength combines all fuzzy values from every rule.

Layer 4: Consequent Parameters are the result value.

Layer 5: Overall Output is the system's output.

Researches have revealed that ANFIS data classification is problematic when the fuzzy rule is very complicated owing to many inputs and nodes of the fuzzy if-then rule. The number of fuzzy if-then rule nodes is p^n where p denotes a division of membership function of input and n represents a number of input [13]. Fuzzy data classification development has double goals: to boost the fuzzy system efficiency, i.e., maximum accuracy (minimum error); and to reduce the complexity of the fuzzy rule, i.e., the lowest number possible. Figure 3 illustrates the goals. The goals can be expressed in two objective functions as per the Equation (2) shown below:

$$\text{Minimize } \{ f_{\text{Error}}(S), f_{\text{Complexity}}(S) \} \qquad (2)$$

Where

$f_{\text{Error}}(S)$ is objective function of classification error in percentage.

$f_{\text{Complexity}}(S)$ is objective function of rule.

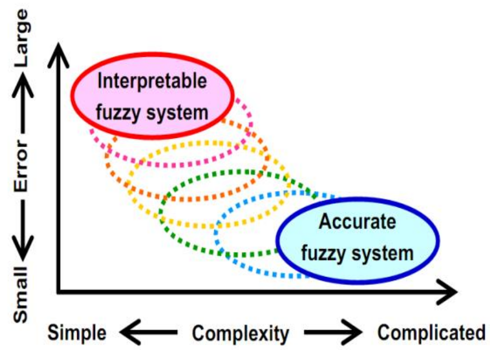


Figure 3. The goals of fuzzy rule system development (Source: Ishibuchi, H. *et al*. [2])

The evolutionary optimization algorithm is a popular method for multi-objective problem optimization. The algorithm [14] imitates the natural evolution, e.g., ants search for the shortest route to reach food source, so-called Ant-Colony Optimization (ACO), or birds try to locate a destination during migration, so-called Particle Swarm Optimization (PSO), etc. MOMA is an evolutionary algorithm, commonly applied in multi-objective problem optimization. Memetic Algorithm, invented by Merz and Freisleben [15], is similar to a genetic algorithm but has a special feature in a way that chromosome can add experience through specific search to enhance efficiency prior to processes.

III. Methodology

The architecture of MOMA-ANFIS applied in this research is depicted in Figure 4 and the details are as follows:

Step 1: Crisp inputs entered the system through nodes. In this layer external signal was transmitted to the next layer.

Step 2: Layer 1-Antecedent Parameter is fuzzification layer. Crisp input was converted to membership function, μ. The calculation was made based on the selected membership function. The Gaussian function was used in this research.

Step 3: Layer 2-T-norm Operator is fuzzy rule base. At this stage, all fuzzy rules and a population were generated. The population, i.e., chromosomes, were encoded. This part of the approach belonged to MOMA, as described below:

1) The initial population was created from ANFIS rule to optimize initial random answers of an equal number to the rules. *N* chromosomes were randomly picked and each had a length equal to the total number of rules.

2) Assessment of population's strength was carried out using Pareto based ranking method. Values of bi-objective functions were calculated with the Equation (2).

3) New generation of population was selected. The selection began (before crossover and mutation procedures) using proportional selection and tournament selection methods.

4) Crossover functioned based on a preset probability of $P_c = 0.6$. One-Point crossover was applied, whereby chromosomes were chosen using tournament selection for crossover procedure.

5) Mutation worked based on a preset probability of $P_m = 0.05$. Mutation was made through an alteration of one value.

6) Population was optimized with local search based on the set probability of $P_{LS} = 0.60$ using insert local search procedure.

7) Population's strength was assessed and chromosomes with the lowest objective function were picked.

8) Population replacement gathered answers from the parent chromosomes and the offspring. Then the optimal values with equal number to that of parent chromosomes were recorded.

Step 4: Layer 3-Normalized Firing Strength is a weighted layer (w). The rule selected in layer 2 was used to calculate the value of Normalized Firing Strength.

Step 5: Layer 4-Consequent Parameters is a layer where the value of each rule was calculated (*y*).

Step 6: Layer 5-Overall Output was calculated by incorporating the outputs from Layer 4.

IV. Efficiency Measurement

The efficiency of the proposed model was measured by accuracy rate, F-measure, and the number of rules acquired. The result of the test demonstrated high percentage of accuracy rate. It suggested the efficiency and suitability of the combined MOMA-ANFIS approach in data classification as expressed in Equation (3) to (6), respectively.

$$\text{Accuracy} = \frac{|A|}{|N|} \quad (3)$$

$$P = \frac{|Ra|}{|A|} \quad (4)$$

$$R = \frac{|Ra|}{|R|} \quad (5)$$

$$\text{F-Measure} = 2 \times \left[\frac{(P \times R)}{(P+R)}\right] \quad (6)$$

Where P is Accuracy rate
R is Recall rate
|N| is total number of sample

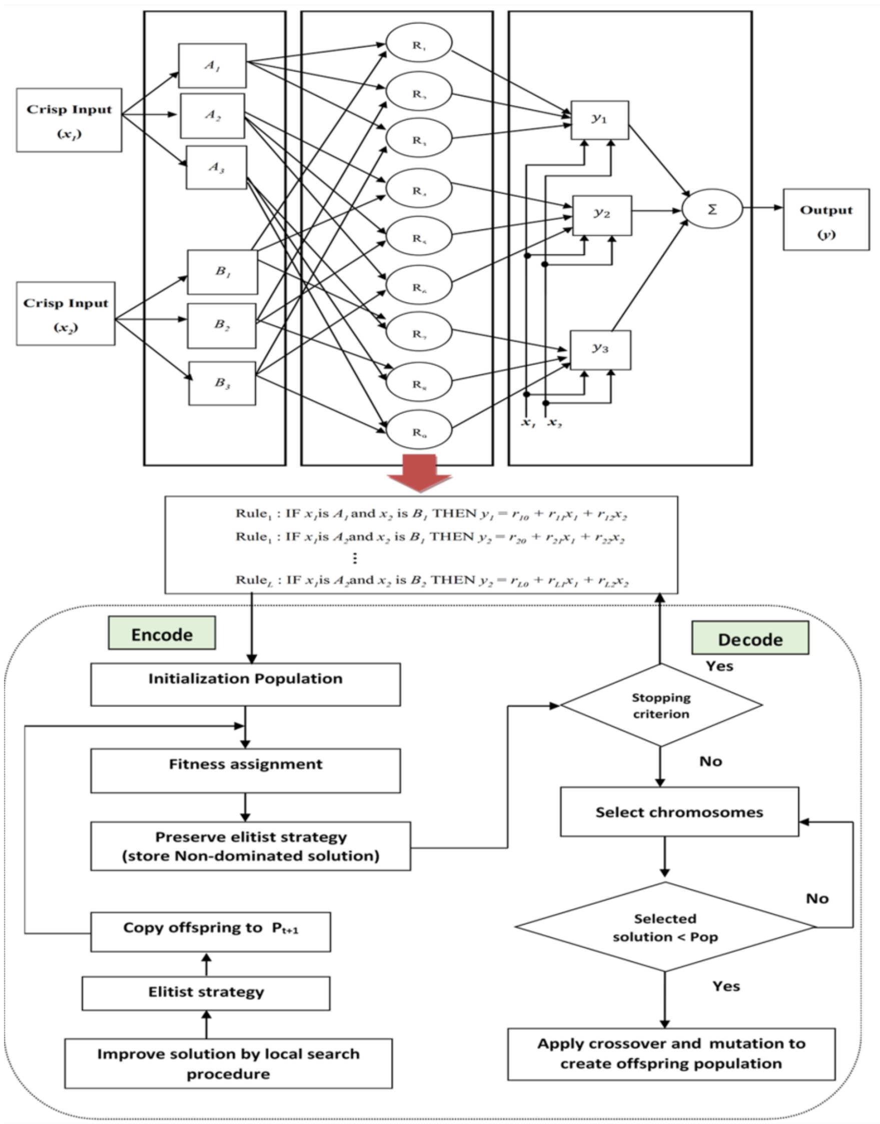
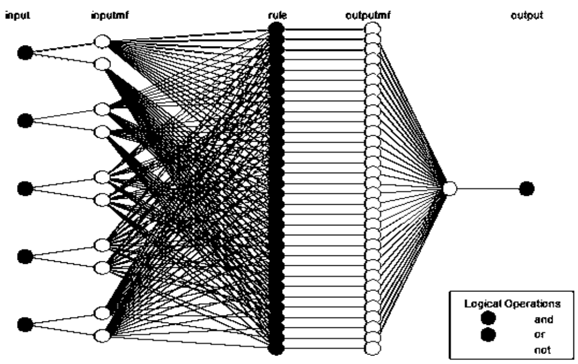


Figure 4. The architecture of the proposed MOMA-ANFIS

|Ra| is number of data of interest, correctly classified.
|A| is all classified data
|R| is all data in specific class of interest



V. RESULT

The problem used in this research is a dataset from the Office of Research Development, Phra Mongkutklao College of Medicine, presented in Table 1.

The dataset is composed of 9987 entities, six attributes, and two classes. Class 1 means risk for heat stroke and Class 0 means normal. Table 1 shows the data.

TABLE I. DATA SPECIFICATIONS

Variable	*Detail*	*key code*
dnk	Water intake	Number of glasses
urn	Urination	Frequency of urination
color	Urine color	1= Dark yellow 2 = Yellow 3 = No color 4 = Others
wg	Weight	Unit of weight in kilogram
Tmp	Body temperature	Unit of temperature
Status	Status	0 = Normal 1 = Risk for heat stroke

Various proportions were suggested for training pattern and testing data, i.e., 40/60 50/5070/30 and 20/80. Each proportion and dataset were tested for 15 times.

TABLE II. TRAINING PATTERN AND DATASET TESTED

Train/Test (%)	*Heat stroke dataset*
50/50	751/751
60/40	901/601
70/30	1001/501
80/20	1202/300

A. Data classification test using the traditional ANFIS

The number of rules in the test using the traditional ANFIS is illustrated in Table 3. The number of fuzzy if-then rule is p^n where p represents a division of number of membership function of each input and n is the number of inputs. This research has set p = 2 for each input, as shown in Figure 5.

TABLE III. NUMBER OF RULES OF THE TRADITIONAL ANFIS

Data set	*Input (n)*	*Rule (p^n)*
Heat stroke	5	32

Figure 5. The architecture of heat stroke rule by the traditional ANFIS

The average result of the 15 tests using the traditional ANFIS at each proposed proportion are demonstrated in Table 4.

TABLE IV. RESULTS OF THE TEST BY THE TRADITIONAL ANFIS

Train/Test (%)	*Accuracy(%)(Avg.)*	*F-Measure (%) (Avg.)*
50/50	98.00	98.03
60/40	98.00	98.64
70/30	98.40	99.13
80/20	98.84	99.25
Average	**98.31**	**98.76**

The 80/20 proportion and heat stroke data had the most efficient data classification. The accuracy rate is 98.84% and F-measure is 99.25%. The average accuracy rate and F-measure of all proportions are 98.31% and 98.76%, respectively.

B. Data classification test using a MOMA -ANFIS

The objective value of ANFIS showed the optimal answer group to the selected data, so-called Pareto Front consisting of two objective functions, i.e., Minimize $\{f_{Error}(S)\}$ and Minimize $\{f_{Complexity}(S)\}$

Samples of Pareto Front graphs from the tests are presented in Figure 6. The number of system rules reduced when the combined MOMA-ANFIS was employed in the tests. Valid result was selected from appropriate group of answer to the dataset. In this research, two objective function values were calculated using Weighted Sum method, and the result was a total value. The lowest total value was picked because minimum error and rules were needed. The number of rules from the comparative testes is shown in Table 5 and Figure 7.

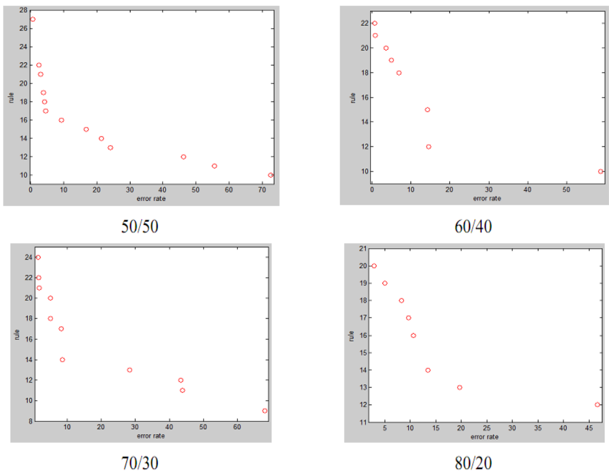


Figure 6. Samples of Pareto Front from MOMA-ANFIS of heat stroke

The answer with the minimum error value was picked by researcher. Since data classification work demanded high accuracy rate, a weighted error value of the objective function was set higher at 0.8. Therefore, the weighted value of objective function complexity was 1 – 0.8 = 0.2. After the calculation, the lowest value of the objective function was chosen as an answer for further process.

The data presented in Table 5 and Figure 7 demonstrate that each proportion after 15 tests similarly generates lower number of rules than 32 rules in the traditional ANFIS test. The decreased number of rules among the proportions are not significantly different. The result of the data classification and the average F-measure throughout the 15 tests are demonstrated in Table 6.

TABLE V. NUMBER OF RULES FROM 15 MOMA-ANFIS TESTS FOR RULE REDUCTION.

Time	**#Rules**			
	50/50	***60/40***	***70/30***	***80/20***
1	18	24	20	22
2	20	25	22	18
3	20	20	18	19
4	24	21	23	21
5	26	22	21	21
6	20	18	18	18
7	21	23	21	21
8	20	24	21	19
9	23	18	18	19
10	21	20	23	18
11	20	20	18	19
12	20	26	22	18
13	20	22	21	19
14	24	21	23	21
15	20	26	22	18

******* ANFIS : 32 Rules**

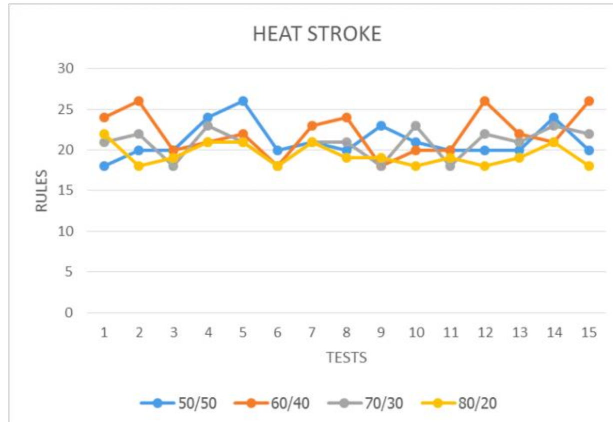


Figure 7. A comparative graph of the number of rules from 15 tests

TABLE VI. RESULTS OF MOMA-ANFIS TESTS

Train/Test (%)	***Accuracy (%) (Avg.)***	***F-Measure (%) (Avg.)***
50/50	98.22	98.42
60/40	98.26	98.45
70/30	98.63	99.12
80/20	98.92	99.14
Average	**98.51**	**98.78**

Table 6 shows that the classification efficiency is best with 80/20 proportion. The accuracy rate is 98.92% and F-measure is 99.14%. The average accuracy rate and F-measure are 98.51% and 98.78%, respectively.

C. Comparative classification efficiency result between MOMA-ANFIS and ANFIS

The results are shown in Table 7.

TABLE VII. COMPARISON OF THE ACCURACY RATE BETWEEN MOMA-ANFIS AND ANFIS

Train/Test (%)	***Accuracy (%)***		***Result (%) (Accuracy)***
	ANFIS	**MOMA- ANFIS**	
50/50	98.00	98.22	Increase
60/40	98.00	98.26	Increase
70/30	98.40	98.63	Increase
80/20	98.84	98.92	Increase
Average	**98.31**	**98.51**	**Increase**

According to Table 7, the test result of MOMA-ANFIS demonstrated a higher accuracy rate in data classification than that from the test using ANFIS. Besides, the number of rules decreased and as a result, the fuzzy architecture became less complicated and data classification rules had higher efficiency.

VI. CONCLUSION

The objective of this research is to propose a model for prediction of heat stroke risks. The study analyzed the data and classified if the population were at risk for heat stroke or normal, through an observation of data pattern fed in the system and an application of MOMA-ANFIS in the classification of heat stroke data. The proposed model focused on the impact of the rule decrease to the system, for instance, whether the system would work properly; the efficiency decreased, remained unchanged, or increased, and what was the alteration value. The result is MOMA-ANFIS increases the classification efficiency, when compared to the ANFIS, even though the fuzzy if-then rules decreased. The enhanced efficiency of the artificial neural fuzzy network can be boosted not only by a decrease of system rules but also by an employment of an evolutionary optimization algorithm in parameter setting for the artificial neural fuzzy network. Further studies in the efficiency enhancement are therefore suggested.

REFERENCES

[1] Jang, J. S.: ANFIS: Adaptive-Network based Fuzzy Inference Systems. IEEE Transactions on Systems, Man and Cybernetics vol.23 (1993)

[2] Ishibuchi, H., Nakashima, T., Murata, T.: Three-objective Genetics-based Machine Learning for Linguistic Rule Extraction. Information Sciences (2001)

[3] Stefania, G., Member, I.: Neuro-Fuzzy Learning and Genetic Algorithm Approach with Chaos Theory Principles Applying for Diagnostic Problem Solving. In: Proceedings of the World Congress on Engineering London, U.K. vol. 1(2009)

[4] Ishibuchi, H., Nakashima, T., Turksen, I.B.: Single-objective and Two-objective Genetic Algorithms for Selecting Linguistic Rules for Pattern Classification Problems. Fuzzy Sets & Systems (1997)

[5] Johansen, T.A., Babuska, R.: Multi-objective Identification of Takagi-Sugeno Fuzzy Models. Fuzzy Systems, IEEE Transactions Vol.11 (2003)

[6] Ishibuchi, H., Yamamoto. T.: Fuzzy Rule Selection by Multi-objective Genetic Local Search Algorithms and Rule Evaluation Measures in Data Mining. Fuzzy Sets & Systems (2004)

[7] Wang, H., Kwong, S., Jin, Y., Wei, W., Man, K.F.: Multi-objective Hierarchical Genetic Algorithm for Interpretable Fuzzy Rule-based Knowledge Extraction. Fuzzy Sets & Systems ((2005)

[8] Alkan, A., Ozcan, E.: Memetic Algorithms for Timetabling. In: CEC '03, The 2003 Congress on Vol.3, pp.1796 - 1802. Evolutionary Computation (2003)

[9] Hung, N. Q. V., Binh, T.V., Anh, D. T.: A Memetic Algorithm for Timetabling. In: Proceedings of 3nd Int ,Conf. RIVF'05 Research Informatics Vietnam-Francophony, pp. 289-294 (2005)

[10] Konak, A., Coit, D.W., Alice E.S.: Multi-objective Optimization using Genetic Algorithms: A Tutorial. Reliability Engineering and System Safety, pp. 992-1007 (2006)

[11] Kim, K., Wang, M., von Spakovsky, M. R., Nelson, D. J.: Stochastic Modeling and Uncertainty Analysis with Multi-objective Optimization Strategies for the Synthesis/Design and Operation / Control of a PEMFC Fuel Processing Subsystem. In: International Mechanical Engineering Congress and Exposition – IMECE'2008, ASME Paper No. IMECE2008-68065, N.Y (2008)

[12] Meesad, P.: Fuzzy Systems and Neural Network. Department of Information Technology, King Mongkut's University of Technology North Bangkok, Bangkok, Thailand (2013)

[13] Branimir, B. J., Irini, S. R.: Modified ANFIS Architecture–Improving Efficiency of ANFIS Technique. In: 7th Seminar on Neural Network Applications in Electrical Engineering (2004)

[14] Emad, E., Tarek, H., Donald, G.: Comparison Among Five Evolutionary based Optimization Algorithm. Advanced Engineering Informatics, pp. 43-53 (2005)

[15] Merz, P., Freisleben, B.: A Genetic Local Search Approach to the Quadratic Assignment problem. In: Back CT, editor. Proceedings of the 7th international conference on genetic algorithms, San Diego, CA: Morgan Kaufmann, pp. 465-72 (1997)