Mechanism of Multi-Objective Genetic Algorithm for maintaining the solution diversity using Neural Network

Kenji Kobayashi[1], Tomoyuki Hiroyasu[2], and Mitsunori Miki[2]

[1] Graduate Student, Department of Knowledge Engineering and Computer Sciences,
Doshisha University ,1-3 Tatara Miyakodani,Kyo-tanabe,
Kyoto, 610-0321, JAPAN
kkobayashi@mikilab.doshisha.ac.jp

[2] Department of Knowledge Engineering and Computer Sciences,
Doshisha University ,1-3 Tatara Miyakodani,Kyo-tanabe,
Kyoto, 610-0321, JAPAN
tomo@is.doshisha.ac.jp,mmiki@mail.doshisha.ac.jp

Abstract. When multi-objective genetic algorithm is applied to real world problems for deriving Pareto optimum solutions, high calculation cost becomes a problem. One of solutions of this problem is using small number of population size. With this solution, however, it often happens that the diversity of the solutions is lost. Then the solutions which have the sufficient precisions cannot be derived. For overcoming this difficulty, the solutions should be re-placed when the solutions are converged on a certain point. To perform this re-placement, inverse analyze to derive the design variables from objects since the solutions are located in the objective space. For this purpose, in this paper, the Artificial Neural Network (ANN) is applied. Using ANN, the solutions which are concentrated on certain points are re-placed and the diversity of the solutions is maintained. In this paper, the new mechanism using ANN to keep the diversity of the solutions is proposed. The proposed mechanism is introduced into NSGA-II and applied for the test functions. It is discussed that in some functions the proposed mechanism is useful compared to the conventional method. In other numerical experiments, the results of the proposed algorithm with plentifully population are discussed and the affection of the proposed mechanism is also described.

1 Introduction

In Multi-Objective Optimization Problems, there are several types of objective functions. These objectives often cannot be minimized or maximized at the same time due to a trade-off relationship between them. One of goals of this problem would be to obtain a set of Pareto-optimum solutions. Thus, in Multi-Objective Optimization Problems, Genetic Algorithm(GA) [4] are often used, because they are multi-point search and a set of Pareto-optimum solutions can be obtained by one search. However, the GA search requires a large number of function evaluations until a Pareto-optimum solution is obtained. Therefore, when multi-objective genetic algorithm(MOGA) is applied to real world problems which need a lot of time for each evaluation, high calculation cost becomes a serious problem. To solve this issue, two solutions are mainly considered.

One of them is the application of the response surface methodology [12] which is a technique for approximating objective functions. This method reduces the calculation cost by generating approximations of objective function and treating these approximations as objective functions for each evaluation. There are several response surface methodologies, such as quadratic polynomial model, neural network model, and Kriging model and so on. Quadratic polynomial model is well used, because it is the simplest of these and requires low calculation cost for approximation [12]. Although the cost of the others model for approximation is greater than the quadratic model, the neural network model and kriging model allow approximation of more complicated objective functions [13].

On the other hand, the method discussed here is the search with small number of individuals. For MOGA search, it is critical to search the Pareto-optimal solutions with keeping the diversity of individuals, because it is more likely that solutions with high accuracy and diversity will be obtained. With this approach, it can reduce the calculation cost, however, it often happens that the solutions are converged on a certain point in the search process and the diversity of the solutions may be lost. In this paper, we propose a mechanism that eases the reduction of diversity of solutions during the search process by using an Artificial Neural Network (ANN). It is expected that this mechanism reduces the calculation cost, and obtains a good set of Pareto-optimum solutions which have high diversity and accuracy, even when performing the search with a small number of individuals. With this search, the calculation cost required by the ANN must be considered, but the problems addressed in this paper require large computational cost for evaluation in GA, and thus the relative calculation cost of ANN is negligible. Examples of such problems include airplane design and automobile collision analysis. The calculation cost for ANN will not be discussed here.

In this paper, we will discuss in detail a mechanism to ease the reduction of diversity using ANN. The proposed mechanism was introduced into NSGA-II [6], a typical MOGA, and its effectiveness and influences on search were investigated for mathematical test functions.

1.1 The problem of Multi-Objective Genetic Algorithm

The advantage of multi-objective optimization using a MOGA is that it can derive several Pareto-optimum solutions at once. However, it requires a number of function evaluations until the Pareto-optimum solution is obtained. The calculation cost can be reduced when using a small number of individuals or a small number of generations. However, with this approach, the diversity of individuals is often lost, as shown in Fig. 1(a). This may have a negative influence on the progress of the search, as in MOGA it is important to maintain diversity during the search. In some cases, it may become difficult to obtain the non-dominated

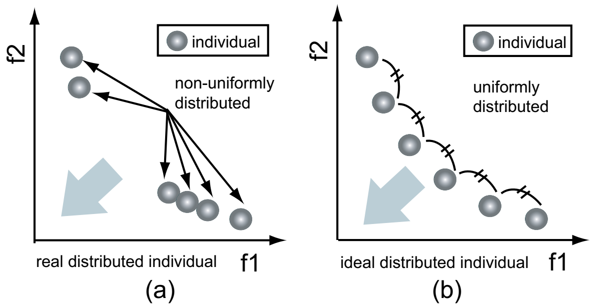

Fig. 1. Diversity of the search individuals.

solutions with the high accuracy and diversity. To overcome this difficulty, whenever the solutions are converged as (see Fig. 1(a)), the solutions are relocated evenly (see Fig. 1(b)). This way, it is expected to maintain the diversity and obtain good Pareto optimal solutions with a smaller number of individuals.

The problem here is how to determine the design variable values of the solutions. For example in Fig. 1, the objective function values of the relocated solutions are known, and the design variable values are not. Thus, the design variable values must be determined through inverse analysis. Here, ANN is used for this inverse analysis.

2 Inverse analysis using ANN

2.1 Artificial Neural Network

An ANN is a sophisticated analytical technique for modeling functions. It is a powerful approach to modeling stochastic and noisy patterns of data in order to produce predicted values of unknown systems [10]. In recent years, there have been many studies on multi-objective optimization using ANN. They are mainly classified into two types: methods that reduce the calculation cost for each evaluation by obtaining an approximation function of the objective function [11], and those to obtain the approximation function that is the inverse of the objective function and apply it for local searches [3, 10]. Multilayer perceptrons, i.e., feed forward neural networks that use back propagation [1, 5, 9] for the learning algorithm, are often used for ANN.

2.2 Diversity maintaining mechanism using ANN

When a set of Pareto-optimum solutions obtained by the MOGA are converged as shown in Fig. 2(a), diversity can be restored if a target solution can be newly derived. However, to obtain the target solution, it is necessary to derive the design variable values corresponding to the objective function values of target individuals. In general, this can be achieved by obtaining the inverse function of the objective function, and we used ANN for this purpose.

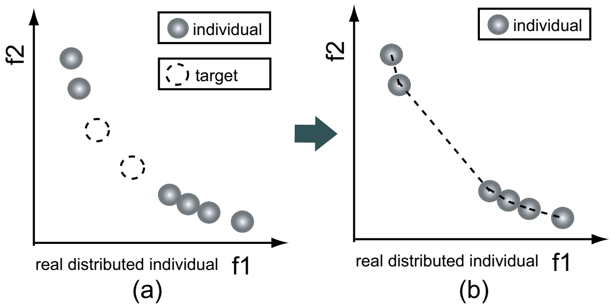

Fig. 2. Diversity maintaining mechanism using ANN.

In this paper, we propose a method to obtain the inverse function in proximity of the non-dominated solutions, such as the broken line in Fig. 2(b). Additional calculation cost is required to train the ANN. However, the problems addressed in this study require a high calculation cost for evaluation, and thus the training cost of ANN is insignificant. As a training data set used to create an inverse function, we need data showing accurate input and output relationships. Therefore a data set actually derived from a MOGA will be used. Using this data, ANN is trained by inputting the objective function values and outputting the design variable values, whereas in MOGA the objective function values are derived from the design variable values.

2.3 Training data filtration

An archive obtained from a MOGA was used as a training data set for ANN. There are two ways of training: one is to use all of the archive, and the other is to use only the non-dominated solutions from the archive as training data. The two methods are compared, using the archive shown in Fig. 3. The results are shown in Fig. 4: the results from the method using the whole archive are shown on the left, while the results using only the non-dominated solutions are shown on the right. In addition, the target Pareto front is indicated with a broken line.

As shown in Fig. 4, when all the solutions in the archive are used, those solutions except for the non-dominated solutions become noise, and thus solutions were derived in the area away from the target line. On the other hand, when only non-dominated solutions were used in the training data set, target solutions were derived appropriately on the Pareto front. Thus, we used only the non-dominated solutions in the archive as a training data set for ANN.

3 Proposed Method

3.1 Outline

The proposed method using both MOGA and ANN aims to reduce the calculation cost and to obtain solutions with high accuracy by maintaining the diversity

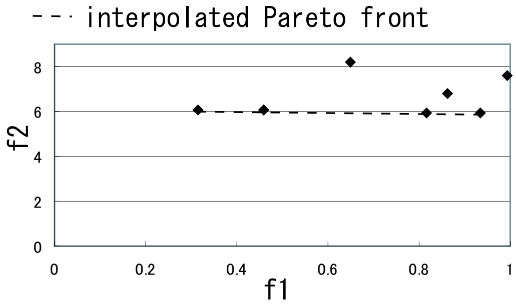

Fig. 3. Archive obtained by MOGA.

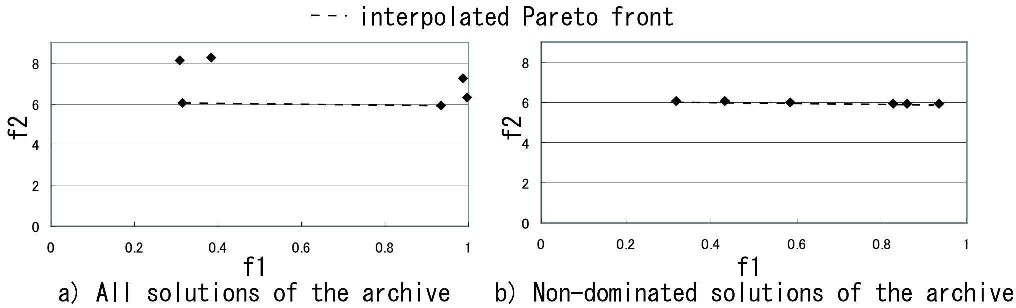

Fig. 4. Change in derived solutions through selection of training data set.

during the search process, even in a search with a small number of individuals. The algorithm of the proposed method is as follows.

N : Number of executions of ANN.
t_{max} : Max number of generations .
t : Number of generations .
k : Number of non-dominated solutions. $i = 1$

Step 1: NSGA-II search is performed up to $i \times t_{max}/N$ generations.
Step 2-1: A set of non-dominated solutions is obtained, and a linear line that passes through the set of non-dominated solutions is obtained through interpolation.
Step 2-2: A set of non-dominated solutions is used as a data set for training the ANN, and an approximation function is created.
(Input: objective function values; Output: design variable values)
Step 2-3: In a set of n non-dominated solutions , all individuals are removed except for those on both end , then $n-2$ target individuals are created so that the distances regarding f1 between adjacent individuals are equal.
Step 2-4: Approximation function created by ANN is used to obtain the design variable values corresponding to the objective function values of the target individuals.

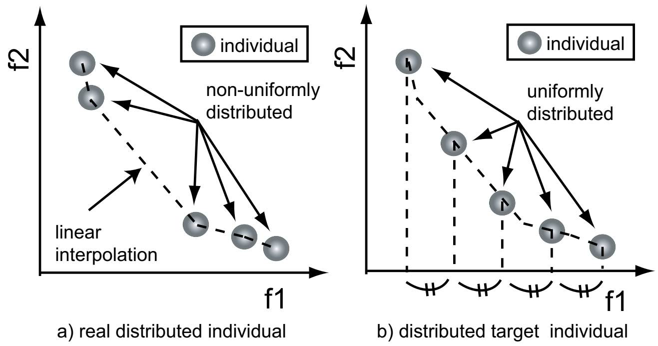

Fig. 5. Linear Interpolation.

Step 2-5: The design variable values obtained by ANN are evaluated using the real objective function, not the approximation function.
Step 2-6: Individuals and archives obtained from ANN are combined and the archive update mechanism of NSGA-II is executed.
Step 2-7: End if all the end conditions are satisfied. If not, return to Step 1($i = i + 1$).

The processes using interpolation mentioned in Steps 2-3 and 2-4 will be discussed in the next section.

4 Relocation of individuals using interpolation

An approximation function that is the inverse of the objective function is created by the ANN which is trained based on individual data derived by the MOGA. Using the approximation function, all non-dominates solutions will be relocated except for those on both end, since it is better to obtain as many individuals as possible in equal distance to the Pareto front. Next, we describe how to determine the objective function values of the target individuals. There are two steps to obtain these values. At the first step, a linear interpolation line is obtained (Step 2-1). A linear interpolation method is adopted, because it showed more positive results in preliminary experiments than two-dimensional interpolation. At the second step, target individuals are relocated so that those satisfy with the following two conditions: 1)individuals are on the interpolation line. 2)individuals are equally distanced regarding f1.

As it is difficult to set the target individuals for interpolation in many objectives (more than 3), this paper focuses on two-objective optimization problems. Many objectives problems (more than 3) will be examined in future studies. The process is shown schematically in Fig. 5. The left diagram in Fig. 5 shows the Pareto optimum solution obtained from the MOGA. Relocation concept is illustrated on the right in Fig. 5.

Table 1. Test Problem(ZDT6)

Problem	Functions
ZDT6	min $f_1 = 1 - \exp(-4x_1)sin^6(6\pi x_1)$ min $f_2 = g \times h$ $g = 1 + 9[(\sum_{i=2}^{n} x_i)/9]^{0.25}$ $h = 1 - (f_1/g)^2$ $x_i \in [0, 1]$, n=2

5 Effectiveness of Diversity Maintaining Mechanism using ANN

The proposed method was tested to examine its effectiveness.

5.1 Test Functions and Evaluation Method

The proposed hybridized method with MOGA and ANN is designed for real-world problems that require large computational cost for each evaluation. As an initial study, we chose a problem where landscape of the function is relatively smooth. ZDT6 [2] unimodal test problems with a non-convex Pareto front is selected as a test function. The equation of test functions is shown in Table 1.

In this experiment, the Ratio of Non-dominated Individuals (RNI) [8] is used as a method to evaluate a set of non-dominated solutions obtained using various methods. RNI measures the accuracy in objective function space. This method compares two sets of non-dominated solutions, and counts the number of solutions that are inferior to those obtained by the other method. This method evaluates items concerning accuracy. The method used by Tan and colleagues [7] was expanded to compare two sets of non-dominated solutions to create this method. The comparison procedure of this method is as follows. The union of the solution sets X and Y obtained by the two methods is set as S^U. Next, solutions not dominated by any solution are selected from S^U , and the selected set of solutions is set as S^P. Then, the ratio of S^P of each method is derived as RNI(X,Y). Examples of RNI(X,Y) are shown in Fig. 6.

The closer this ratio is to the maximum, 100%, the better it is compared to the other method, indicating that a solution that is closer to the true solution is being obtained.

5.2 Examination of diversity improvement using ANN

It is examined whether the reduced diversity of solutions are improved by hybridized NSGA-II. The archive size was set to 10, which showed good results in the preliminary experiment. The parameters used are shown in Table 2.

Fig. 7(a) shows the non-dominated solutions obtained by NSGA-II, and Fig. 7(b) shows the target individuals relocated according to these non-dominated

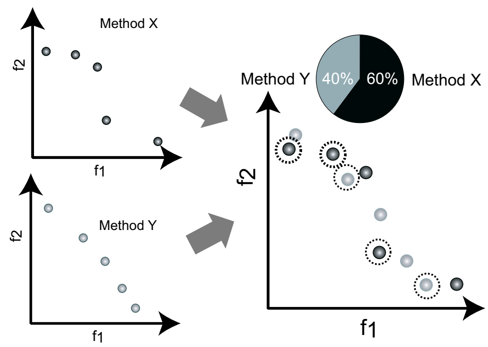

Fig. 6. Ratio of Non-dominated Individuals (RNI).

Table 2. Parameter1

Population size	6
Number of generations	60
Archive size	10
Number of dimensions	2
Crossover rate	1.0
Method of crossover	Two-point crossover
Number of times ANN applied	2
Number of trials	30

Table 3. Parameter2

Search technique	Hybrid	NSGA-II
Number of generations	60	61
Number of evaluations	368	366
Number of times ANN applied	2	None

solutions. Individuals obtained using the proposed mechanism are shown in Fig. 7(c).

Comparison of Fig. 7(a) and (c) indicates that more uniform distribution of the solution is achieved after application of the diversity maintaining mechanism using ANN.

In the next experiment, conventional NSGA-II and hybridized NSGA-II were compared to examine their effectiveness. RNI of conventional NSGA-II and hybridized NSGA-II is shown in Fig. 8. Also, plot diagrams of each method is shown. The parameters used are shown in Table 3. ANN was applied evenly during the search (e.g., 20th and 40th generations)

The results in Fig. 8 indicate that RNI of hybridized NSGA-II is higher than that of conventional NSGA-II. In addition, hybridized NSGA-II could derive the solutions close to the region of the Pareto-optimum solutions in more trials

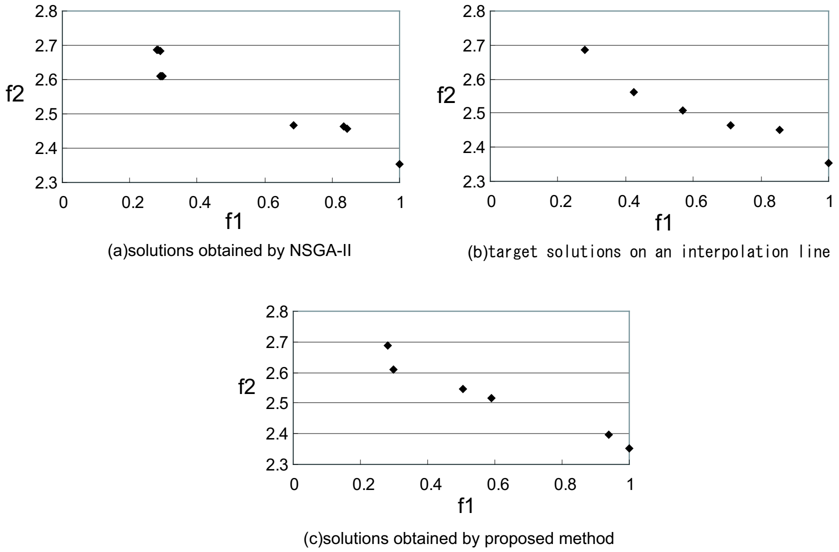

Fig. 7. Results of solutions relocation using ANN.

Table 4. Parameter3

Number of generations	30
Number of times ANN applied	0,3,6,10,15,30

compared to the conventional NSGA-II. From the above, we found that the issue of reduced diversity by the conventional NSGA-II with a small number of individuals can be resolved using the proposed method and it is possible to execute a search with maintaining its diversity.

5.3 Examination of number of times ANN applied

The diversity mechanism using ANN was introduced into NSGA-II, and the influence of the number of times ANN applied was examined. The parameter used is shown in Table 4. The result of this experiment is shown in Fig. 9. RNI was obtained by comparing the various number of times ANN applied with the same number of evaluations.

Fig. 9 shows that RNI was better when ANN was applied often during the search. This indicates that the diversity maintaining mechanism using ANN has a positive effect on the search.

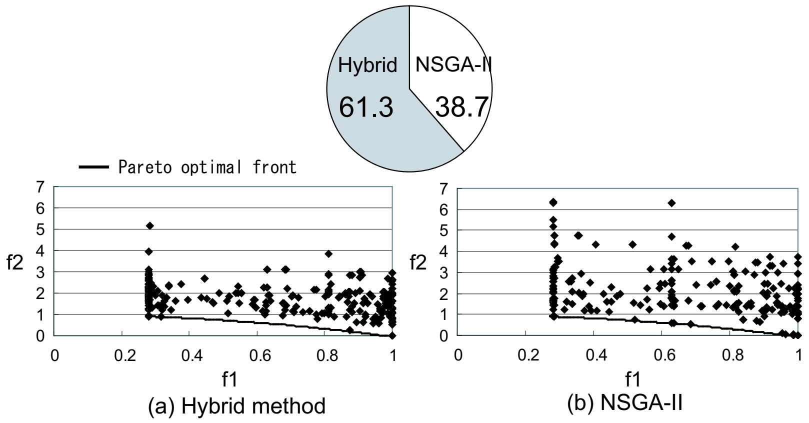

Fig. 8. Comparison of RNI and search result.

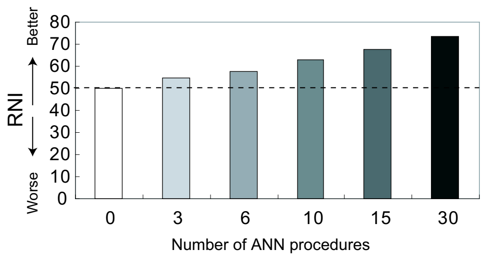

Fig. 9. RNI comparison.

Table 5. Parameter4

Search technique	Hybrid	NSGA-II
Population size	6	18
Number of generations		30
The number of evaluations	300	540
Archive size	10	18
Number of times ANN applied	30	None

5.4 Comparison of a search with small and large number of individuals.

In the previous sections, we have been examining searches with a small number of individuals. Here, we examine the effectiveness of the proposed method related to the reduction of calculation cost. In this experiment, the diversity maintaining mechanism was introduced into NSGA-II and the search was executed with a small number of individuals. The results were compared to a conventional NSGA-II with a large number of individuals. The parameters used are shown in Table 5.

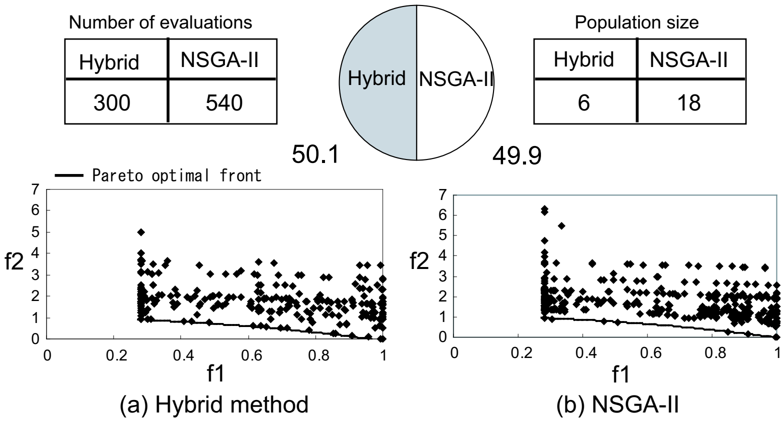

Fig. 10. Comparison of RNI and search result.

RNI and plot diagrams of all non-dominated solutions of each method are shown in Fig. 10.

Fig. 10 shows that the performance of the hybridized NSGA-II with a smaller number of individuals is equivalent to conventional NSGA-II with a larger. These observations indicated the following conclusions. When a conventional MOGA search with a small number of individuals is performed, the searching ability can be improved by using diversity maintaining mechanism. Also, this is comparable to a conventional MOGA search with a larger number of individuals when the number of evaluations is small to some extent. Therefore, this mechanism allows reduction of the calculation cost.

6 Conclusions

In this paper, a mechanism was proposed of a MOGA search that maintains its diversity during the search process, even when performing the search with a small number of individuals. It can restore the reduced diversity of solutions by using ANN, whenever the diversity is lost. The mechanism of maintaining diversity by ANN creates an approximation function which is used to obtain the design variable values corresponding to the objective function values of target individuals. The target objective values are determined by relocating the individuals so that those are equally distanced. It is expected that this proposed method derives individuals with high accuracy and diversity. This proposed hybridized method is designed for real-world problems that require large computational cost for each evaluation, such as airplane design and automobile collision analysis. Additional calculation costs due to ANN training are very small relative to the evaluation costs in GA. Thus, in this paper the cost of training the ANN is not described. Here the proposed mechanism was introduced into NSGA-II, and its effectiveness was examined for mathematical test functions. The results of numerical experiments indicated that the search performance of the proposed hybridized method is comparable to that of the conventional method. Using hybridized method, it is possible to derive solutions which have high accuracy and diversity with a small number of evaluations, even when performing the search with a small number of individuals.

References

1. C.Bishop. Neural networks for pattern recognition. *Oxford University Press*, 1997.
2. K. Deb and T. Meyarivan. Constrained test problems for multi-objective evolutionary optimization. In *KanGAL report 200005, Indian Institute of Technology, Kanpur, India*, 2000.
3. Antonio Gaspar-Cunha and Armondo Vieira. A hybrid multi-objective evolutionary algorithm using an inverse neural network. 2004.
4. D. E. Goldberg. *Genetic Algorithms in search, optimization and machine learning.* Addison-Wesly, 1989.
5. HENSELER J. Back propagation. *Lect Notes Comput Sci*, 931:37–66, 1999.
6. Amrit Pratab Kalyanmoy Deb, Samir Agrawal and T. Meyarivan. A fast elitist non-dominated sorting genetic algorithm for multi-objective optimization: Nsga-ii. In *KanGAL report 200001, Indian Institute of Technology, Kanpur, India*, 2000.
7. K.C.Tan, T.H.Lee, and E.F.Khor. Increamenting multi-objective evolutionary algorithms: Performance studies and comparisons. In *First International Conference on Evolutionary Multi-Criterion Optimization*, pages 111–125, 2001.
8. J. D. Knowles and D. W. Corne. Approximating the nondominated front using the pareto archived evolution strategy. volume 8, pages 149–172, 2000.
9. L.Bull. On modelbased evolutionary computation. 1999.
10. Ian Griffin Salem F. Adra, Ahmed I. Hamody and Peter J. Fleming. A hybrid multi-objective evolutionary algorithm using an inverse neural network for aircraft control system designa. *2005 IEEE Congress on Evolutionary Computation (CEC'2005)*, 1:1–8, 2005.
11. Hirotaka Nakayama Takaharu Sirai, Masao Arakawa. Application of rbf network in approximate optimization in satisfcing method. In *Design and Systems*, pages Vol14th page85–86, 2004.
12. Akira Todoroki. Response surface method, http://florida.mes.titech.ac.jp/responsesurface.pdf.
13. Kikuo Fujita Yusuke Kounoe. Third-order polynominal response surface with optimal selection of interaction terms for global approximation. In *Design and Systems*, volume 15, pages 31–34.