ICIC Express Letters
Part B: Applications
Volume **7**, Number **7**, July **2016**
 ISSN 2185-2766
pp. 1511–1517

A COMMUNICATION STRATEGY ENHANCED DIVERSITY HERDS GREY WOLF OPTIMIZER FOR MULTIMODAL OPTIMIZATION

Chin-Shiuh Shieh, Trong-The Nguyen, Mong-Fong Horng and Thi-Kien Dao

Department of Electronic Engineering
National Kaohsiung University of Applied Sciences
No. 415, Jiangong Rd., Sanmin Dist., Kaohsiung City 807, Taiwan
csshieh@cc.kuas.edu.tw; jvnthe@gmail.com

Received December 2015; accepted March 2016

Abstract. *Easy to converge to a local optimum and losing the global optimum nearby could happen in solving the multimodal optimization and complex constrained optimization problems due to interference phenomena among constrained dimensions. In this paper, a communication strategy for diversity herds Grey Wolf Optimizer (GWO) is proposed for solving the multimodal optimization problems. In this proposed method, the whole population is split into several small herds. These herds are regrouped frequently by using various regrouping schedules, and the proposed communications strategy provides the information flow for the search agents to communicate among the herds. A set of the multimodal benchmark functions is used to test the performance quality of the proposed method. According to the experimental result, the proposed method shows the better performance in comparison with the original methods is up to 57%.*
Keywords: Grey wolf optimizer, Diversity grey wolf optimizer, Multimodal optimization

1. **Introduction.** In practical optimization problems, it is often desirable to simultaneously locate multiple global and local optima of a given objective function [1]. For real-world problems due to physical constraints, the best results cannot always be realized. There are different optimal solutions in the search space. Those problems could have been solved by the nature-inspired algorithms [2]. These algorithms have been applied to a wide range of applications; for example, solution for the constrained optimization problems via genetic algorithms [3], training multi-layer perceptron via grey wolf optimizer [4] and unequal clustering formation for wireless sensor networks based on bat algorithm [5].

Moreover, the No Free Lunch (NFL) theorem [6] has logically proved that there is no meta-heuristic best suited for solving all optimization problems. In other words, a particular meta-heuristic may show very promising results on a set of problems, but the same algorithm may show poor performance on a different set of problems. Obviously, NFL makes this field of study highly active which results in enhancing current approaches and proposing new meta-heuristics every year. This also motivates our attempts to consider the strength point of the algorithms to be suitable to the type problems' characteristics. Solving the multimodal optimization and complex constrained optimization problems could be easy to converge to a local optimum and losing the global optimum nearby sometimes due to interference phenomena among constrained dimensions. Enhancing the diversity populations in the optimal algorithms is one of the solutions to this issue. The diversity artificial search agents increase the accuracy and extend the global search capacity versus the original structure [7]. Moreover, with the small sized herds searching using their own best historical information, they are easy to converge to a local optimum because of GWO's convergence property.

In this paper, three main factors of the small size herds, neighborhood topology technique, and its own best historical information are considered to construct an enhanced diversity method for the grey wolf optimizer algorithm. The small size herds could be figured out by dividing the population into subpopulations or sub-regions. The neighborhood topology technique could be implemented by applying some Niching techniques such as the crowding, i.e., the fitness sharing available resources and the speciation. The good information obtained by each herd evolving optimization could be exchanged among the herds. It results to achieve the benefit of cooperation individuals and exploitation through local extreme to the global optimum.

The paper is organized as follows. A brief review of grey wolf optimizer is given in Section 2. Analysis and designs for enhancing diversity GWO are presented in Section 3. A series of experimental results on the multimodal benchmark and the comparison between original GWO and the proposed method are discussed in Section 4. Finally, the conclusion is summarized in Section 5.

2. **Meta-Heuristic Grey Wolf Optimizer.** Grey Wolf Optimizer (GWO) is inspired from observing, imitating, and modeling the leadership hierarchy and hunting mechanism of grey wolf when searching and attacking for the prey [8]. There are four guided types of grey wolves in the leadership hierarchy, alpha (α), beta (β), delta (δ), and omega (ω). The type of α is considered the fittest solution, and then β, and δ are considered the second and the third best solutions respectively. Omega (ω) could be assumed the rest of the candidate solutions. GWO algorithm consists of the constructed mathematical models as follows.

Encircling prey mathematical model: The dominance degree in the social leadership hierarchy is formulated in equations of model as follows:

$$\overrightarrow{D} = \left|\overrightarrow{C} \cdot \overrightarrow{X_p}(t) - \overrightarrow{X}(t)\right| \tag{1}$$

$$\overrightarrow{X}(t+1) = \overrightarrow{X_p}(t) - \overrightarrow{A} \cdot \overrightarrow{D} \tag{2}$$

where $\overrightarrow{D}$ is dominance degree, t indicates the current iteration, $\overrightarrow{A}$ and $\overrightarrow{C}$ are coefficient vectors, $\overrightarrow{X_p}(t)$ is the position vector of the prey, and $\overrightarrow{X}(t)$ indicates the position vector of a grey wolf. Equations (1) and (2) are two-dimensional position vector and some of the possible neighbors. The vectors $\overrightarrow{A}$ and $\overrightarrow{C}$ are calculated as follows:

$$\overrightarrow{A} = 2\overrightarrow{a} \cdot \overrightarrow{r_1} - \overrightarrow{a} \tag{3}$$

$$\overrightarrow{C} = 2 \cdot \overrightarrow{r_2} \tag{4}$$

where components $\overrightarrow{a}$ are linearly decreased from 2 to 0 over the course of iterations and r_1, r_2 are random vectors in $[0, 1]$. A grey wolf in the position of (X, Y) can update its position according to the position of the prey (X^*, Y^*). Different places around the best agent can be reached with respect to the current position by adjusting the value of $\overrightarrow{A}$ and $\overrightarrow{C}$ vectors.

Hunting prey mathematical model: The hunting behavior of grey wolves can be simulated when the alpha (best candidate solution), beta, and delta are supposed to have better knowledge about the potential location of prey. Therefore, the first three best solutions are obtained so far and oblige the other search agents (including the omegas) to update their positions according to the best search agents. This simulating model is formulated as follows:

$$\overrightarrow{D_\alpha} = \left|\overrightarrow{C_1} \cdot \overrightarrow{X_\alpha} - \overrightarrow{X}\right|, \quad \overrightarrow{D_\beta} = \left|\overrightarrow{C_2} \cdot \overrightarrow{X_\beta} - \overrightarrow{X}\right|, \quad \overrightarrow{D_\delta} = \left|\overrightarrow{C_3} \cdot \overrightarrow{X_\delta} - \overrightarrow{X}\right| \tag{5}$$

$$\overrightarrow{X_1} = \overrightarrow{X_\alpha} - \overrightarrow{A_1} \cdot \left(\overrightarrow{D_\alpha}\right), \quad \overrightarrow{X_2} = \overrightarrow{X_\beta} - \overrightarrow{A_2} \cdot \left(\overrightarrow{D_\beta}\right), \quad \overrightarrow{X_3} = \overrightarrow{X_\delta} - \overrightarrow{A_3} \cdot \left(\overrightarrow{D_\delta}\right) \tag{6}$$

$$\overrightarrow{X}(t+1) = \frac{\overrightarrow{X_1} + \overrightarrow{X_2} + \overrightarrow{X_3}}{3} \tag{7}$$

The position of the prey is estimated by alpha, beta, and delta and other wolves update their positions randomly around the prey during the hunt.

Attacking mathematical model: The grey wolves finish the hunt by attacking the prey when it stops moving. The pseudo code of the GWO algorithm is presented in Figure 1.

Initialize the grey wolf population X_i $(i = 1, 2, \ldots, n)$, Initialize a, A, and C
Calculate the fitness of each search agent
X_α = the best search agent, X_β = the second best search agent, X_δ = the third best search agent
while *(t < Max number of iterations)*
for each search agent
Update the position of the current search agent by Equation (7)
end for
Update a, A, and C
Calculate the fitness of all search agents
Update X_α, X_β and X_δ
$t = t + 1$
end while
return X_α

FIGURE 1. Pseudo code of the GWO algorithm

3. **Enhanced Diversity Herds GWO.** The diversity GWO is designed based on original GWO optimization and a neighborhood based on Niching technique is used. There are two considered characters in neighborhood structure, small size and communicating. Not as other evolutionary algorithms that prefer larger population, GWO needs a comparatively smaller population size. Especially for simple problems, a population with three to five wolves can achieve satisfactory results. GWO with small neighborhoods performs better on complex problems. In order to increase diversity, the small sized herds are employed by dividing the wolves in GWO into the groups. Each herd uses its own members to search for better area in the search space. Since the small sized herds are searching using their own best historical information, they are able to converge to a local optimum. So, a randomized regrouping schedule should be set for optimizing by the probability weight setting, and a new configuration of small herds is started the searching for the best global target. The exchangeable information is activated between herds whenever the communication strategy is triggered. The benefit of cooperation and exploitation is achieved through the communicating information. The fitness sharing available resources is one common used in the Niching techniques. The herd GWO has its own wolves as known search agent and the finest agents are evaluated according to the fitness function. These best agents among all the wolves in one group will be assigned to the poorer agents based on the fitness evaluation in the other groups, replace them and update agents for each herd after running the exchanging period.

Let G_j be the group, where j is the index of the group, n is the number of groups, $j = 0, 1, 2, \ldots, n-1$; and m be the number of wolves of a group, called population size of the group. While $t \cap R \neq \theta$, k search agents (where the top k fitness in the group G_j) will be copied to G_{j+1} to replace the same number of search agents with the worst fitness. Every R generations, the population is regrouped randomly and starts searching using a new configuration of small herds. In this way, the good information obtained by each herd is exchanged among the herds. Simultaneously the diversity of the population is increased. It is not surprising that it performs better on complex multimodal problems.

The steps can be described as follows.

1) **Initialization:** Initialize a, A, C, generate $m \times n$ search agents and divide population into n groups randomly, with m individuals in each group G. Assign R the exchanging period for executing X_{ijt} solutions, where $i = 0, 1, \ldots, m-1$; $j = 0, 1, \ldots, n-1$; t is the current iteration and set to 1.
2) **Evaluation:** Evaluate the value of $f(X_{ijt})$ for search agents in the j-th group G_j.
3) **Update:** Update the position of the current search agent by Equations (6) and (7), and a, A, and C by Equations (3), (4) and (5).
4) **Communication Strategy:** Migrate k best agents among G_{tj} to the $(j+1)$-th group G_{tj+1}, and mutate G_{tj+1} by replacing k poorer agents in that group. If $\mathrm{mod}(i, R) == 0$, regroup the herds randomly, and update all of the group in each R iterations.
5) **Termination:** Repeat Step 2 to Step 5 until the predefined value of the function is achieved or the maximum number of iterations has been reached. Record the best value of the function $f(X_{ijt})$ and the best agent solution among all the agent positions X_{ijt}.

4. **Simulation Results.** A set of multimodal benchmark functions [9,10] is used to test the accuracy and the speed of the proposed algorithm dGWO. All the benchmark functions for the experiments are averaged over different random seeds with 25 runs. The goal of the optimization is to minimize the outcome for all multimodal benchmarks. Let $X = \{x_1, x_2, \ldots, x_d\}$ be d-dimensional real-value vector.

The detail of parameter settings of GWO can be found in [8]. The initial range, the dimension and total iteration number for all test functions are listed in Table 1.

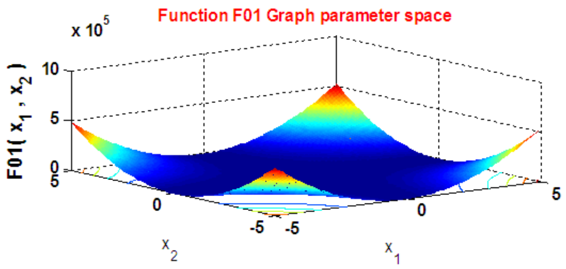
Table 1. The initial range and the total iteration of the multimodal benchmark functions

Multimodal test functions	**Range**	**Dim.**	**Iteration**
$F_1(x) = \left[e^{-\sum_{i=1}^{n}(x_i/\beta)^{2m}} - 2e^{-\sum_{i=1}^{n}x_i^2}\right]\prod_{i=1}^{n}\cos^2 x_i,\ m=5$	± 20	30	2000
$F_2(x) = \sum_{i=1}^{n}\sin(x_i) \times \left(\sin\left(\frac{ix_i^2}{\pi}\right)\right)^{2m},\ m=10$	$0, \pi$	30	2000
$F_3(x) = 0.1\left\{\sin^2(3\pi x_1) + \sum_{i=1}^{n}(x_i-1)^2\left[1+\sin^2(3\pi x_i+1)\right] + (x_n-1)^2\left[1+\sin^2(2\pi x_n)\right]\right\} + \sum_{i=1}^{n}u(x_i, 5, 100.4)$	± 50	30	2000
$F_4(x) = \sum_{i=1}^{n}\left[x_i^2 - 10\cos(2\pi x_i) + 10\right]$	± 5.12	30	2000

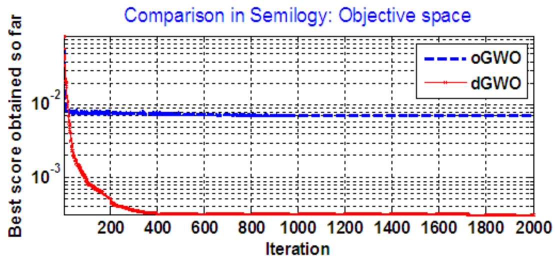
The parameters setting for both dGWO and oGWO are the initial a, A, and C randomly, the total population size N set to 40, and the dimension D set to 30. Further for dGWO, population size is $m \times n$ set to 4×10, number of group n set to 4, and the fixed iteration R set to 20. Each function contains the full iterations of 2000. The final result is obtained by taking the average of the outcomes from all runs. Comparison of the performance quality and running time of the dGWO and oGWO methods for the multimodal optimization problems are shown in Table 2. Clearly, the results of the proposed algorithm on all of these cases of testing multimodal benchmark problems show that dGWO method almost increases higher than those obtained from original method. The maximum case obtained from dGWO method increases higher than those obtained from the oGWO method which is up to 57%. However, the figure for the minimum case

TABLE 2. The quality performance evaluation and speed comparison of oGWO and dGWO for solving the multimodal optimization problems

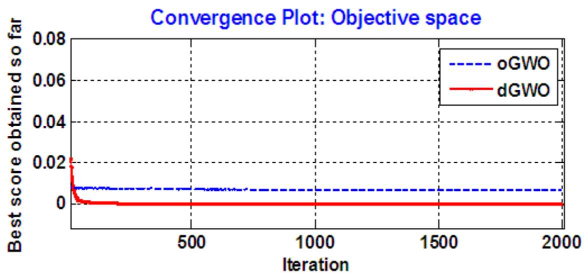
Multimodal test functions	**Time consumption (minutes)**		**Performance evaluation**		**Accuracy %**
	oGWO	**dGWO**	oGWO	**dGWO**	**Comparison**
$F_1(x)$	1.5783	1.5783	9.21E-04	6.49E-04	42%
$F_2(x)$	2.4793	2.4371	1.27E+00	1.09E+00	16%
$F_3(x)$	2.7154	2.7294	1.55E+07	1.15E+07	35%
$F_4(x)$	1.7011	1.7275	1.45E+02	9.21E+01	57%
Average	**2.1185**	**2.1181**	**3.86E+06**	**2.86E+06**	**38%**



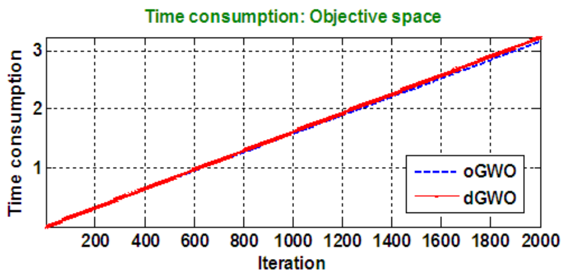
(a) The benchmark function graph



(b) Comparison in semilogy

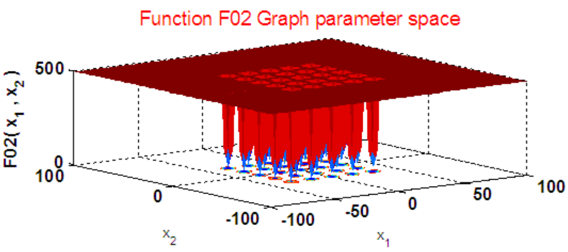


(c) Comparison in performance

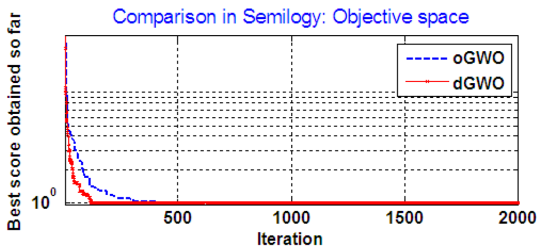


(d) Comparison in time running

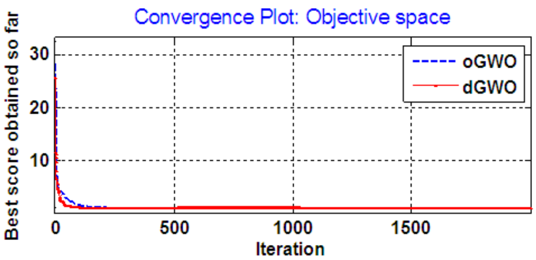
FIGURE 2. The experimental results of function $F_1(x)$



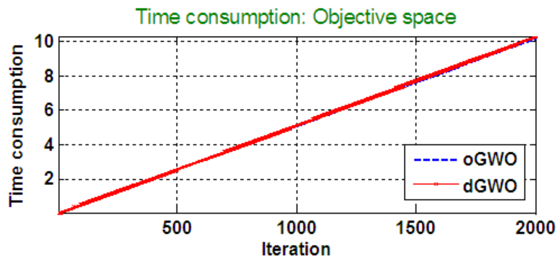
(a) The benchmark function graph



(b) Comparison in semilogy

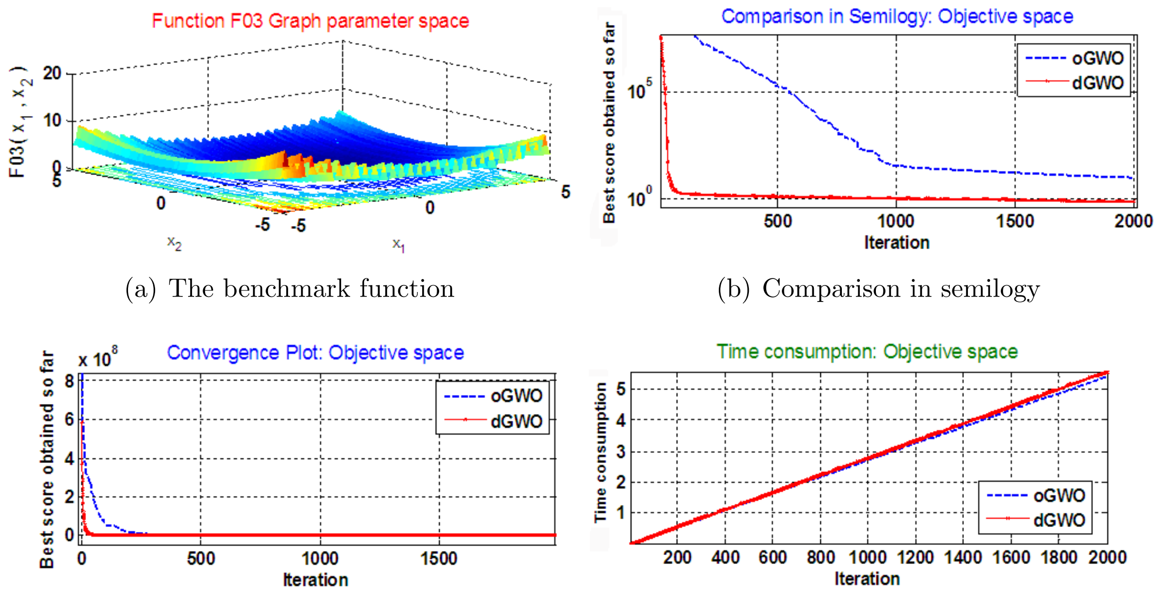


(c) Comparison in performance



(d) Comparison in time running

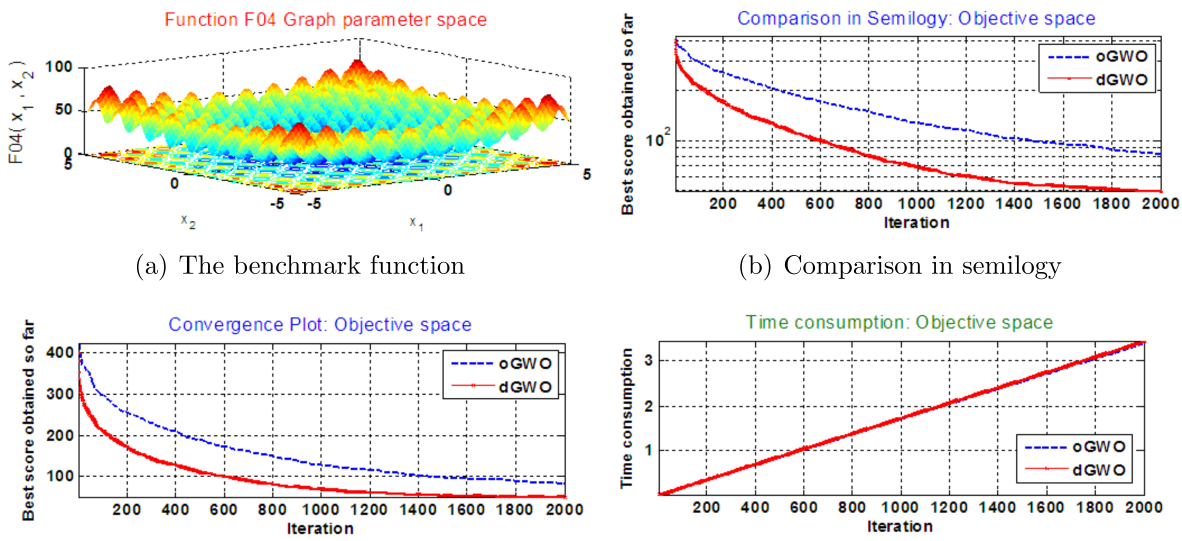
FIGURE 3. The experimental results of function $F_2(x)$



(a) The benchmark function (b) Comparison in semilogy

(c) Comparison in performance (d) Comparison in time running

FIGURE 4. The experimental results of function $F_3(x)$



(a) The benchmark function (b) Comparison in semilogy

(c) Comparison in performance (d) Comparison in time running

FIGURE 5. The experimental results of function $F_4(x)$

is only the increase of 16%. Thus, in general the proposed algorithm obtained the average cases of various tests multimodal optimization problems for the convergence, and accuracy increased more than those obtained from the oGWO method being 38%.

Figures 2-5 show the experimental results of four multimodal benchmark functions in 25 seed runs output obtained from dGWO and oGWO methods with the same iteration of 2000.

In Figures 2-5, semilogy plot measures the performance through using a base 10 logarithmic scale with their index containing real numbers and convergence plot measures the performance through using the convergence. Clearly, all of these cases of testing multimodal benchmark functions for dGWO have performance quality higher than those for oGWO in terms of the accuracy and convergence, even though the time consuming of two methods is equivalent.

5. **Conclusion.** In this paper, a novel proposed method for the multimodal optimization problems was presented with the diversity herds Grey Wolf Optimizer (GWO). The implementation of diversity populations could have important significance for avoiding the easy to converge to a local optimum and losing the global optimum nearby of the optimal algorithms for solving the multimodal optimization and complex constrained optimization problems. In this new proposed algorithm, whole populations of GWO are split into several independent groups based on the original structure of the GWO, and the neighborhood topology technique is used. The communication strategy provides the information flow for the search agents to exchange the signs in different groups. By this way, the poorer solutions in the groups will be replaced with new better solutions from neighbor groups after running the exchanging period. A randomized regrouping schedule is activated by the probability weight setting, and a new configuration of small herds is started the searching for the best global target.

A set of the multimodal benchmark functions is used to test the quality performace, and the speed of the proposed method. According to the experimental result, the proposed method shows the better performance in comparison with the original method. The best case obtained from the proposed method increases higher than those obtained from the original method which is up to 57%. However, the figure for the worst case is only the increase of 16%. Thus, in general the proposed algorithm dGWO obtained the average cases of various tests multimodal optimization problems for the convergence, and accuracy increased more than those obtained from the original method GWO being 38%.

REFERENCES

[1] B. Y. Qu, P. N. Suganthan and S. Das, A distance-based locally informed particle swarm model for multimodal optimization, *IEEE Trans. Evolutionary Computation*, vol.17, pp.387-402, 2013.
[2] X.-S. Yang, *Nature-Inspired Metaheuristic Algorithms*, Luniver Press, 2010.
[3] A. Homaifar, C. X. Qi and S. H. Lai, Constrained optimization via genetic algorithms, *Simulation*, vol.62, pp.242-253, 1994.
[4] S. Mirjalili, How effective is the grey wolf optimizer in training multi-layer perceptrons, *Applied Intelligence*, pp.1-12, 2015.
[5] T.-T. Nguyen, C.-S. Shieh, M.-F. Horng, T.-G. Ngo and T.-K. Dao, Unequal clustering formation based on bat algorithm for wireless sensor networks, in *Knowledge and Systems Engineering*, V.-H. Nguyen, A.-C. Le and V.-N. Huynh (eds.), Springer International Publishing, 2015.
[6] D. H. Wolpert and W. G. Macready, No free lunch theorems for optimization, *IEEE Trans. Evolutionary Computation*, vol.1, pp.67-82, 1997.
[7] H. Wang, H. Sun, C. Li, S. Rahnamayan and J.-S. Pan, Diversity enhanced particle swarm optimization with neighborhood search, *Information Sciences*, vol.223, pp.119-135, 2013.
[8] S. Mirjalili, S. M. Mirjalili and A. Lewis, Grey wolf optimizer, *Advances in Engineering Software*, vol.69, pp.46-61, 2014.
[9] X. Yao, Y. Liu and G. Lin, Evolutionary programming made faster, *IEEE Trans. Evolutionary Computation*, vol.3, pp.82-102, 1999.
[10] M. Jamil and X.-S. Yang, A literature survey of benchmark functions for global optimisation problems, *International Journal of Mathematical Modelling and Numerical Optimisation*, vol.4, pp.150-194, 2013.