

A MULTIMODAL OPTIMIZATION METHOD FOR SIMULATION SYSTEMS

Zhizhao LIU ^(a), Wei LI ^(b), Ming YANG ^(c)

Control and Simulation Center, Harbin Institute of Technology, Harbin 150080, China

^(a)liuzhizhao2007@163.com, ^(b)fleehit@163.com, ^(c)myang@hit.edu.cn

ABSTRACT

Considering multimodal optimization problem and potentially run times for continuous simulation systems, Particle Swarm Optimization (PSO) method based on Support Vector Machine (SVM) and Cluster Analysis (CA) is proposed. SVM, a global metamodel of simulation system, is built and PSO algorithm is used to search for local optimal points based on SVM. To determine local optimums, the population individuals of PSO are classified by CA. Consequently, six typical multimodal function optimization problems are selected to verify the optimization performance.

Keywords: simulation system, multimodal optimization, SVM metamodel, cluster analysis

1. INTRODUCTION

Simulation optimization is optimization of performance measures based on outputs from simulations (Fu 2001). Optimization problems exist widely in aviation, aerospace, shipbuilding, information and other fields, which is also the main research direction of simulation (Fu 2002, Yang Zhang and Wang 2004, Kleijnen 2007, Min Ma and Yang 2007, Fu and Chen 2008). Therefore, simulation optimization is an important part of simulation theory research.

In this paper, we consider the potentially extensive run times and multimodal optimization problem for continuous simulation system. Multiple extreme values usually exist in real systems, which is the same as simulation systems. The essence of this optimization problem is multimodal function optimization. In the field of simulation, there is little research on multimodal optimization problems. At present, intelligent optimization algorithms are applied for multimodal optimization problem, getting optimization solution by the iteration of population. However, expensive simulation limits simulation run times, leading the result that intelligent optimization algorithms can not be applied directly (Barton and Meckesheimer 2006, Barton 2009, Jin Chen and Simpson 2001). On the other hand, the last generation individuals are located around local optimums. The individuals near the same local optimum cannot be guaranteed to be coincident completely. How to determine the number of local optimal solutions, and obtain the corresponding values is also a valuable problem to study.

Based on the two problems above, PSO optimization method based on SVM metamodel and CA is proposed. The method is applied to multimodal function optimization problems, which is aimed to valid that the method proposed in this paper can get all the effective local optimal solutions in different situations of multimodal optimization.

2. RESEARCH ON THE KEY PROBLEMS

The multimodal optimization problem and determination of the local optimal points will be discussed, Section 2.1 for the first one and Section 2.2 for the second. As the evaluation of multimodal optimization is different from the general optimization, the evaluation indexes are given in Section 2.3.

2.1. Multimodal Optimization Problem

In view of the limitations of optimization method based on gradient, intelligent optimization algorithms are selected for multimodal optimization problem. Simulation optimization based on metamodel has advantages of rapid and practical. In order to avoid too many run times for simulation system, metamodel is introduced to the multimodal optimization method. In this paper, PSO based on SVM metamodel is proposed.

2.1.1. PSO method

In 1995, J. Kennedy and R. Eberhart proposed an intelligent algorithm named PSO based on the foraging behavior of birds, with many advantages such as clear meaning of parameters, small computation of iteration, breaking away from local optimum and so on (Cui and Zeng 2011). The main iterative equation is given as Eq. (1). From the equation, what can be seen is that the population individual is affected by current velocity, the best position of this individual in history and the best position of all the individuals in history. If acceleration C_1 is increased and C_2 is decreased properly, population individual will be stable at/near local optimal points, which can achieve the purpose of local optimum searching.

$$\begin{aligned}
v_{ij}(t+1) &= w \cdot v_{ij}(t) + C_1 r_{1j}(t)(P_{ij}(t) - x_{ij}(t)) \\
&\quad + C_2 r_{2j}(t) \cdot (P_{gj}(t) - x_{ij}(t)) \\
x_{ij}(t+1) &= x_{ij}(t) + v_{ij}(t+1)
\end{aligned} \tag{1}$$

where, $v_{ij}(t)$ denotes the j th dimensional velocity of individual i in the t th generation. w is a inertia constant. C_1 and C_2 are acceleration constants. $r_{1j}(t)$ and $r_{2j}(t)$ are random numbers independently. $P_{ij}(t)$ is the best position in j th dimensional of individual i until the t th generation. $x_{ij}(t)$ is the j th position of individual i in the t th generation. $P_{gj}(t)$ is the best position in j th dimensional of all population individuals until the t th generation.

2.1.2. SVM metamodel

SVM is based on statistical learning theory, which uses structure-risk-minimum criterion instead of empirical risk minimization criterion and transforms low dimensional nonlinear problem into high dimension linear problems. It has the advantages of simple mathematical formula, intuitive geometric meaning, good generalization ability and so on (Pasolli Melgani Tuia et al 2014).

Given a training sample set: $S = \{(x_1, y_1), \dots, (x_l, y_l)\}$, $x_i \in R$, $y_i \in R$, $i = 1, 2, \dots, l$. The Regression function $f(x) = \omega \cdot x + b$ satisfies the following conditions (Nello and John 2004, Mountrakis Im and Ogole 2011):

$$\begin{aligned}
\min & \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^l (\xi_i + \xi_i^*) \\
s.t. & f(x_i) - y_i \leq \xi_i + \varepsilon \\
& y_i - f(x_i) \leq \xi_i + \varepsilon \\
& \xi_i, \xi_i^* \geq 0 \quad i = 1, 2, \dots, l
\end{aligned} \tag{2}$$

where $\xi_i, \xi_i^*, i = 1, 2, \dots, l$ are slack variables. ε - insensitive function is the loss function. By introducing Lagrange function, using kernel function $K(x, y)$ instead of inner product $\langle \psi(x), \psi(y) \rangle$ and combined with the KKT condition, the regression model expression is as shown in Eq. (3).

$$f(x) = \sum_{i=1}^l (\alpha_i - \alpha_i^*) K(x, x_i) + b \tag{3}$$

2.2. Determination of the local optimal points

Since the population individuals of PSO algorithm are stable near the local optimal points of metamodel, each of the local optimal points gathers many individuals. Fig. 1 is an example for single input single

output system where \bullet denotes the PSO individual and $-$ denotes the relationship of input and output variable.

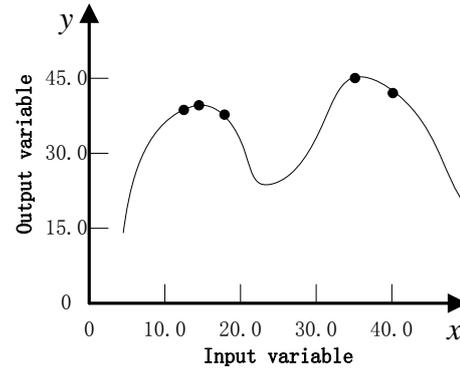


Fig. 1 PSO individuals of last generation

From Fig. 1, it is easy for human to determine how many local optimums the curve has from the curve or the PSO individuals. But for computer, it need specific algorithm. If there are multiple input variables, the input/output relationship could not be shown in a figure and even for people it is hard to find all the local optimal points from dozens or even hundreds of individuals.

In order to solve the problem, CA method is applied to cluster the individuals automatically. CA is a prevail method for data analysis (Yang 2005). It is widely applied in machine learning, data mining and statistics, etc. The basic process is mentioned which consists of 4 basic steps: 1) every sample is a cluster and compute similar measure; 2) the most similar two clusters are incorporated. 3) compute similar measure for new clusters and incorporate the most two clusters again; 4) iteration is done again and again until all the samples are summed up to one cluster.

Due to the advantages of local population gathered, we cluster individuals according to the distance of input variables and get the best output value in each cluster considered as the local optimal points. Suppose the input variable of population individual $A_i = [\alpha_{i1}, \alpha_{i2}, \dots, \alpha_{im}]$, where m is the dimension of input. The distance of input variable is selected to be the similar measure as in Eq. (4).

$$s_{ij} = \sqrt{\sum_{k=1}^m (\alpha_{ik} - \alpha_{jk})^2} \tag{4}$$

Since the population individuals are near the local optimum, similar measure values called linkage distances near the same local optimum are little and the opposite are very large relatively. So the number of clusters can be determined by linkage distances in hierarchical diagram. Suppose that linkage distances between the two clusters are d_1, d_2, \dots, d_k . We use a threshold seen in Eq. (5). If the linkage distance is lower than threshold, the linked two clusters are incorporated to one cluster.

$$C_{sc} \cdot \max_{i=1,2,\dots,k} (d_i) \quad (5)$$

where, $\max_{i=1,2,\dots,k} (d_i)$ is the maximum of d_i and C_{sc} is a constant.

2.3. Evaluation indexes of optimization results

Four indexes are used to evaluate the ability of searching local optimal points (Yang 2004). As bad local optimums have no meaning for the decider, only the local optimums close to the optimization one on the numerical value are required to consider. The detail content of five indexes is as bellow:

Definition 1: Number of Effective Peaks (NEP) for h : suppose that optimization algorithm gets n peak values, denoted by p_i , $i=1,2,\dots,n$. If p_i meets $p_i/p > h$ (in this paper $h = 0.9$), where p is the corresponding peak value of p_i . Then, p_i is an effective peak for h . For p_i , $i=1,2,\dots,n$, the number of effective peaks is NEP for h .

Definition 2: Maximum Peak Ratio (MPR): suppose that the algorithm gets n peaks, denoted by p_i , $i=1,2,\dots,n$ and the corresponding peak values are q_i , $i=1,2,\dots,n$. Then MPR is equal to $\sum_{i=1}^n p_i / \sum_{i=1}^n q_i$.

Definition 3: Degree of Precision (DP): suppose that the algorithm gets n peaks, denoted by p_i , $i=1,2,\dots,n$ and the corresponding testing function's peak values are q_i , $i=1,2,\dots,n$. Then DP is equal to $\sum_{i=1}^n (q_i - p_i)^2$.

Definition 4: Relative Maximum Absolute Error (RMAE): suppose that the algorithm gets n peaks, denoted by p_i , $i=1,2,\dots,n$ and the corresponding peak values are q_i , $i=1,2,\dots,n$. Then RMAE is $\max |(q_i - p_i)/q_i|$.

3. PSO METHOD BASED ON SVM AND CA

In this study, we focus on the ability of searching effective local optimums for the optimization method. We consider the simulation system as a black box and that the number of local optimums is unknown. It is assumed to get the maximum value while the opposite situation is similar. The detailed steps for the method are given as follows:

(i) Sampling and Metamodeling.

Select a proper experiment design method for sampling, and then build a SVM metamodel of simulation system. The SVM metamodel is as shown in Eq. (3).

(ii) Metamodel evaluation.

There are three common performance measures to evaluate metamodel: Multiple Correlation Coefficient (R^2), Relative Average Absolute Error (RAAE) and Relative Maximum Absolute Error (RMAE). They are shown in Eq. (6-8). In engineering application, the main evaluating index, R^2 , should be more than 0.9 and the assistant evaluating indexes are good as small as possible. Under ideal condition, R^2 equals to 1 while RAAE and RMAE equal to 0. If the evaluating indexes are satisfied, Step (iii) is executed. Otherwise, implement (i).

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (6)$$

$$RAAE = \frac{\sum_{i=1}^n |y_i - \hat{y}_i|}{\sum_{i=1}^n |y_i - \bar{y}|} \quad (7)$$

$$RMAE = \frac{\max(|y_i - \hat{y}_1|, \dots, |y_i - \hat{y}_n|)}{\frac{1}{n} \sum_{i=1}^n |y_i - \bar{y}|} \quad (8)$$

where y_i is the i th output value. $p_i, i=1,2,\dots,n$ is the i th predictive value. $\bar{y}, i=1,2,\dots,n$ is the mean of outputs.

(iii) Multimodal optimization.

Suppose that the PSO individual $A_i^k = [\alpha_{i1}^k, \alpha_{i2}^k, \dots, \alpha_{im}^k]$ and the predicted fitness from SVM metamodel $\hat{F}^k = [\hat{\beta}_1^k, \hat{\beta}_2^k, \dots, \hat{\beta}_n^k]$, where $k=1,2,\dots,n$, m is the number of input variables and n is the size of PSO population. The parameters of PSO algorithm is modified properly, making individuals could not skip from local optimum. After repeated iteration as in Eq. (1), all population individuals are stable near several local optimal points.

(iv) Preliminary Screening.

Some individuals at the last generation will be rejected if their fitness is too bad. The screening equation is shown in Eq. (9).

$$\hat{\beta}_i^\xi \geq C_{ps} \cdot \hat{\beta}_{\max}^\xi \quad (9)$$

Where $\hat{\beta}_i^\xi$ is the predicted fitness of individual A_i^ξ at the ξ th generation. ξ is the number of last generation. $\hat{\beta}_{\max}^\xi$ is the maximum value of individual fitness until the ξ th generation. C_{ps} is a constant and its maximum value equals to 1. Go to step (v) directly if there is no need to screen.

(v) Cluster Analysis.

Confirm the number and values of local optimal points as follows:

1) In consideration of magnitude difference, the population individuals of PSO algorithm at the last generation are standardized as shown in Eq. (10).

$$\tilde{\alpha}_{ij}^{\xi} = \frac{\alpha_{ij}^{\xi} - b_j}{a_j - b_j} \quad (10)$$

Where α_{ij}^{ξ} is the j th variable of individual A_i^{ξ} . a_j is the maximum value of the j th variable and b_j is the minimum one.

2) Every individual A_i^k can be seen as a cluster, so there are n clusters, denoted by G_1, G_2, \dots, G_n . The distance s_{ij} can be given in Eq. (4) and the distance matrix $S = \{s_{ij}\}$.

3) Get the minimum value from the matrix S , denoted by s_{ij} . The cluster $G_j = \{A_{j,1}^{\xi}, A_{j,2}^{\xi}, \dots, A_{j,t}^{\xi}\}$ is incorporated into the cluster $G_i = \{A_{i,1}^{\xi}, A_{i,2}^{\xi}, \dots, A_{i,r}^{\xi}\}$. We get the new cluster $G_i = \{A_{i,1}^{\xi}, A_{i,2}^{\xi}, \dots, A_{i,r}^{\xi}, A_{j,1}^{\xi}, A_{j,2}^{\xi}, \dots, A_{j,t}^{\xi}\}$. Go to the step 2) until only one cluster exist.

4) Hierarchical diagram can be drawn and the number of clusters can be determined from the linkage distance. The clusters are denoted by G_1, G_2, \dots, G_v .

5) From the cluster $G_k = \{A_{k,1}^{\xi}, A_{k,2}^{\xi}, \dots, A_{k,t}^{\xi}\}$, we can get the corresponding fitness set $H_k = \{\hat{F}_{k,1}^{\xi}, \hat{F}_{k,2}^{\xi}, \dots, \hat{F}_{k,t}^{\xi}\}$. The maximum value of set H_k is the local optimal value denoted by $\hat{F}_{k,\max}^{\xi}$. The corresponding input variable value is $A_{k,\max}^{\xi}$. Through the v kinds of clusters, we can get v local optimal values denoted by $\hat{F}_{1,\max}^{\xi}, \hat{F}_{2,\max}^{\xi}, \dots, \hat{F}_{v,\max}^{\xi}$ and the corresponding input variable values denoted by $A_{1,\max}^{\xi}, A_{2,\max}^{\xi}, \dots, A_{v,\max}^{\xi}$.

(vi) Verification.

Take local optimal points $A_{1,\max}^{\xi}, A_{2,\max}^{\xi}, \dots, A_{v,\max}^{\xi}$ to the simulation model for achieving the actual output values $F_{1,\max}, F_{2,\max}, \dots, F_{v,\max}$.

4. NUMERICAL EXPERIMENTS AND RESULTS

To analysis the performance of proposed method quantificationally, six well known continuous deterministic multimodal optimization problems are tested. They indicate the input-output relationship. The expressions are as follows:

Example 1 (Lu Liang and Zhang 2008)

$$f_1(x) = \sin^6(5.1\pi x + 0.5), x \in [0, 1]$$

Example 2 (Lu Liang and Zhang 2008)

$$f_2(x) = e^{-4 \ln 2 \times (x-0.0667)^2 / 0.64} \sin^6(5.1\pi x + 0.5), x \in [0, 1]$$

Example 3 (Liu Wang and Wang 2004)

$$f_3(x) = x(x+1) \sin[(2x-0.5)^2 \pi - 1], x \in [-1.5, 1]$$

Example 4 (Zhang and Shao 2008)

$$f_4(x, y) = 200 - (x^2 + y - 11) - (x + y^2 - 7)^2, x, y \in [-6, 6]$$

Example 5 (Bi and Wang 2011)

$$f_5(x, y) = 10e^{(-0.03(x-3)^2 - 0.03(y-3)^2)} + 8e^{(-0.08(x+5)^2 - 0.08(y+5)^2)} + 7e^{(-0.08(x-4)^2 - 0.04(y+7)^2)}, -10 \leq x, y \leq 10$$

Example 6 (Zhang and Shao 2008)

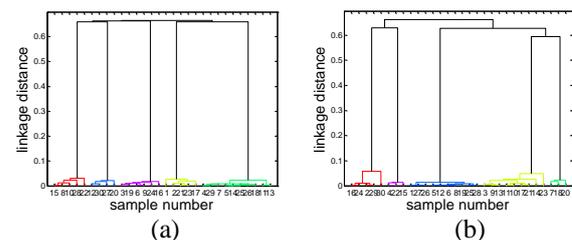
$$f_6(x, y) = \left(\frac{a}{b + (x^2 + y^2)} \right)^2 + (x^2 + y^2)^2, -5.12 \leq x, y \leq 5.12$$

Support Vector Machine (SVM) is chosen to be the metamodel of the six examples. The index values are shown in Table 1. The SVM index R^2 is greater than 0.9.

Table 1: index values of metamodels

NO.	R^2	RAAE	RMAE
1	0.9991751	0.0283195	0.0719181
2	0.9969014	0.0576008	0.2498982
3	0.9996853	0.0270918	0.0481523
4	0.9918178	0.0220666	1.0733524
5	0.9999872	0.0033617	0.0274754
6	0.9238889	0.1059319	2.8144157

PSO algorithm is applied on the SVM metamodel where population size is 50, iteration number is 100, $w=1$, constant $C_1=10$, $C_2=0.1$. So the information of last generation individuals can be got. Cluster analysis is applied according to the distance of inputs. The Hierarchical diagrams are shown as in Fig. 1. The threshold constant $C_{sc}=0.1$, so the numbers of clusters are respectively 5, 5, 8, 4, 3, 5. The local optimal points can be obtained as in Fig. 2.



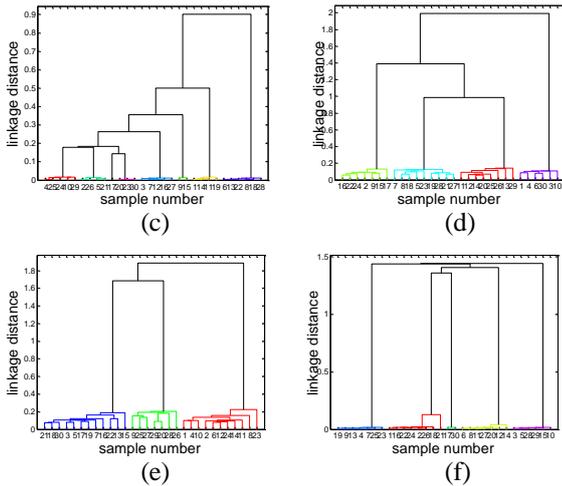


Fig. 1. Hierarchical diagrams (a) function 1 (b) function 2 (c) function 3 (d) function 4 (e) function 5 (f) function 6

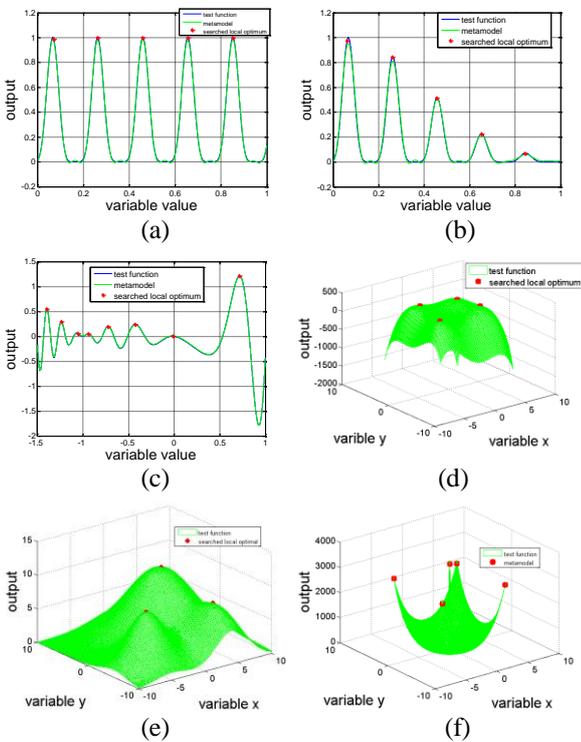


Fig. 2. Multimodal optimization results (a) example 1 (b) example 2 (c) example 3 (d) example 4 (e) example 5 (f) example 6

The ability index values of searching local optimal points are shown in Table 2. The EPN of each example is the same as the true number of examples, the MPR greater than 0.94, and the RMAE is lower than 0.08.

Table 2: Evaluation index values of searching local optimums

NO.	Number of Peaks	NEP	MPR	DP	RMAE
1	5	5	0.994	0.000272	0.013320

2	5	5	0.998	0.000009	0.005434
3	8	8	0.996	0.000038	0.084723
4	4	4	0.988	45.509	0.031813
5	3	3	0.995	0.00691	0.008312
6	5	5	0.949	113112.47	0.066122

We evaluate the efficiency by s , the time ratio of proposed method and the direct optimization, which is good as small as possible. As the SVM modeling time and the optimization time of PSO can be ignored compared with that of simulation system run, the evaluation index s can be seen in Eq. (11).

$$s = \frac{N^* \times T}{N \times T} = \frac{N^*}{N} = \frac{N_{sample}}{N_{pop} \times (N_{gen} + 1) \times N_{times}} \times 100\% \quad (11)$$

where N^* is the run times of simulation system by the proposed method. N is the run times of simulation system by the direct optimization. T is the time of every simulation run. N_{sample} is the times of sampling. N_{pop} is the size of population. N_{gen} is the iteration number. N_{times} is the Monte Carlo times of PSO. In this paper, $N_{pop} = 50$, $N_{gen} = 100$ and $N_{times} = 5$. From the Table 3, the results show that the proposed method can largely shorten the time and increase the efficiency of optimization.

Table 3: Evaluation index values for efficiency increased

NO.	N_{sample}	s
1	101	0.40%
2	101	0.40%
3	251	0.99%
4	169	0.67%
5	441	1.76%
6	529	2.12%

5. CONCLUSION

PSO optimization method based on SVM and CA is proposed for multimodal optimization problem of simulation systems, which consider the multimodal optimization problem and the potentially run times. The number and values of local optimal points can be obtained automatically by computer without too much human labor. Through six typical problems, it can be shown that for the circumstances of equal and unequal peaks this method can find all the effective peaks with a high precision and largely shorten the optimization time. Future work will include the applications of the method to simulation systems and multimodal optimization method to provide robust solutions for discrete-event simulations.

ACKNOWLEDGMENTS

The material is based upon work supported by the "Fundamental Research Funds for the Central Universities"(Grant No. HIT.NSRIF.2015035).

REFERENCES

- Fu M.C., Simulation Optimization. Proceedings of the 2001 Winter Simulation Conference, pp.53-61(2001).
- Yang M., Zhang B., Wang Z. C., The Analysis of Modeling and Simulation Development Direction. Journal of System Simulation 16(9), 1901-1904+1913(2004).
- Fu M.C., Optimization for Simulation, Theory vs. Practice. Inform Journal on Computing 14(3), pp.192-215(2002).
- Fu M.C., Chen C.H., Some Topics for Simulation Optimization. Proceeding of the 2008 Winter Simulation Conference (WSC2008), pp.27-38(2008).
- Min F.Y., Ma P., Yang M., A knowledge-based method for the validation of military simulation. Proceeding of the 2007 Winter Simulation Conference (WSC2007), pp. 1374-1381 (2007).
- Kleijnen J.P.C., Design and Analysis of Simulation Experiments. Springer Science and Business Media (2007).
- Barton R.R., Meckesheimer, M., Metamodel-Based Simulation Optimization. Handbook in OR&MS, pp. 535-574 (2006).
- Barton R.R., Simulation Optimization Using Metamodels. Proceedings of the 2009 Winter Simulation Conference, pp.230-238(2009:).
- Jin R.C., Chen W, Simpson T.W., Comparative studies of metamodeling techniques under multiple modeling criteria. Structural and multidisciplinary optimization, 23(1), pp. 1-13(2001).
- Pasolli E.; Melgan F.; Tuia D.; et al, SVM Active Learning Approach for Image Classification Using Spatial Information, IEEE Transactions on Geoscience and Remote Sensing, 52(4), pp. 2217-2233(2014).
- Nello C., John S.T., An Introduction to Support Vector Machines and Other Kernel-based Learning Methods, Cambridge University Press(2004).
- Mountrakis G.; Im J.; Ogole C., Support vector machines in remote sensing: A review, Isprs Journal of Photogrammetry and Remote Sensing, 66(3), pp. 247-259 (2011).
- Cui Z. H., Zeng J. C., Particle Swarm Optimization. Science Press (2011).
- Yang X. B., Research of Key Techniques in Cluster Analysis. Ph.D. degree of Zhejiang University (2005).
- Lu Q., Liang C.Y., Zhang E. Q., A dynamic sharing scheme based multimodal niche genetic algorithm. Proceeding of the 7th World Congress on Intelligent Control and Automation, pp.5333-5338(2008).
- Liu H. J., Wang X. F., Wang Z. B., Adaptive genetic algorithm for multi-peak searching. China

Academic Journal Electronic Publishing House, pp. 143-146(2004).

- Bi X. J., Wang Y. J., Niche artificial bee colony algorithm for multi-peak function optimization. Systems Engineering and Electronics, 33(11), pp. 2564-2568(2011).
- Zhang M.F., Shao C., Niche artificial fish swarm algorithm for multimodal function optimization. Control Theory & Applications, 25(4), pp. 773-776(2008).
- Yang H.Q., a study on techniques of improving genetic algorithm and its application. Master degree of Zhejiang University of technology (2004).

AUTHORS BIOGRAPHY

ZHIZHAO LIU is a Ph.D. student at HIT, and received the M.E. from HIT in 2012. His research interests include simulation optimization and operational effectiveness evaluation. His email is liuzhizhao2007@163.com.

WEI LI is a Lecture at Harbin Institute of Technology (HIT), and received the B.S., M.E. and Ph.D. from HIT in 2003, 2006 and 2009 respectively. His research interests include simulation evaluation, simulation data analysis, and distributed simulation. His email is fleehit@163.com.

MING YANG is a professor and the director of control and simulation center at HIT. Also he is the vice editor-in-chief for Journal of System Simulation and editor for International Journal of Modeling Simulation and Science Computing. His research interests include system simulation theory and VV&A. His email is myang@hit.edu.cn.