

---

## A research overview of manifold-learning-based multiobjective evolutionary algorithm

---

Wei Zhan\*, Chao Guo and Leiping Xiong

School of Computer Science,  
Yangtze University,  
Jingzhou, Hubei, China  
Email: zhanwei814@gmail.com  
Email: 1442650559@qq.com  
Email: 1074137926@qq.com  
\*Corresponding author

**Abstract:** Manifold learning algorithm can find out the low-dimensional smooth manifold embedded in high-dimensional data. So, in this paper, the manifold learning algorithm is introduced into multiobjective optimisation algorithm for multiobjective optimisation problems (MOPs), and a manifold-learning-based multiobjective evolutionary algorithm (ML-MOEA) is proposed to overcome deficiency of the traditional evolutionary multi-objective optimisation algorithms (EMOAs) and model-based multi-objective optimisation algorithms (MOEAs) for reducing dimension of data and mining manifold in the decision space of MOPs, build accurate model, guide algorithm evolution and accelerate convergence. The steps of ML-MOEA is as follows: 1) randomly initialisation; 2) modelling via manifold learning algorithm; 3) extend and reproduction; 4) elite selection; 5) halt or go to step 2. Based on the framework of ML-MOEA, a ML-MOEA via self-organising maps (ML-MOEA/SOM) and a ML-MOEA via SOM locally linear embedding (ML-MOEA/LLE) is proposed, and comparison experiment of algorithm performance is done.

**Keywords:** manifold learning; multiobjective optimisation; evolutionary algorithm; model-based multiobjective estimation of distribution algorithm; research overview.

**Reference** to this paper should be made as follows: Zhan, W., Guo, C. and Xiong, L. (2015) 'A research overview of manifold-learning-based multiobjective evolutionary algorithm', *Int. J. Computing Science and Mathematics*, Vol. 6, No. 3, pp.287–296.

**Biographical notes:** Wei Zhan received his BEng in Computer Science from Hubei Normal University, Huangshi, China, in 2001, and received his MEng and PhD degrees in Computer Science from China University of Geosciences (CUG), Wuhan, China, in 2006 and 2013, respectively. Currently, he is a faculty member with School of Computer Science, Yangtze University, Jingzhou, Hubei, China. His research interests include computer vision (CV), multiobjective optimisation (MOP), intelligent computation and its applications.

Chao Guo is an undergraduate student in School of Computer Science, Yangtze University, Jingzhou, Hubei, China.

Leiping Xiong is an undergraduate student in School of Computer Science, Yangtze University, Jingzhou, Hubei, China.

This paper is a revised and expanded version of a paper entitled 'A hybrid multiobjective evolutionary algorithm model-based on local linear embedding (LLE)' presented at International Workshop on Swarm Intelligent Systems (IWSIS2014), Nanchang, China, 6–8 June 2014.

---

## 1 Introduction

In many engineering areas, multiobjective optimisation problems (MOPs) arise very often, the objectives in a MOP conflict with each other and no single solution can optimise all the objectives at the same time. The Pareto set/front is the set of all the optimal trade-offs in the decision/objective space.

### 1.1 Evolutionary multi-objective optimisation algorithms

Since 1980s, evolutionary algorithms (EAs) have become one of the most important method to solve MOPs, EAs is a kind of random searching algorithms, which simulate selection and evolutionary of creature and it can work whatever objective functions are continuous and differentiable or not.

In Rosenberg (1967), in his doctoral dissertation, suggested that MOPs can be solved via genetic operation. In Schaffer's (1985), since the publication of seminal work, a number of EAs have been developed for MOPs (Fonseca and Fleming, 1993). On above mentioned traditional EMOAs, particularly NSGA-II, PAES and SPEA2 have become classic algorithms for MOPs, but research (Boyd and Vandenberghe, 2004) shows that when the algorithm closes to convergence, crossover and mutation, in EMOAs, will bring about undesirable impact. So, the current EMOAs research mainly focuses on the following related issues:

- 1 according the different problems' type, design the different crossover and mutation operator
- 2 introduce estimation of distribution algorithms (EDAs) into EMOAs.

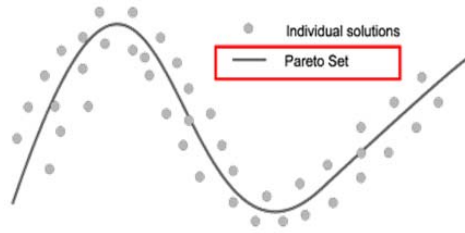
### 1.2 Model-based multi-objective optimisation algorithms

For the shortcomings of EMOAs and EDAs, the model-based multi-objective optimisation algorithms (MOEAs) is proposed, which is based on the following Karush-Kuhn-Tucker condition: under mild smoothness conditions, the Pareto set (in the decision space) of a continuous MOP is a piecewise continuous  $(m - 1) - D$  manifold, where  $m$  is the number of the objectives, as shown in Figure 1. But, this law in the decision space of MOPs does not be utilised in traditional EMOAs and EDAs. Since 2005, Zhou et al. (2005) proposed model-based multiobjective optimisation algorithm, a number of algorithms have been developed for MOPs (Zhou et al., 2009), among which RM-MEDA is typical (Zhou et al., 2009). As shown in Figure 2, via the local principal component analysis algorithm, in RM-MEDA, the local principal component analysis algorithm (local PCA) is used for building clusters (the number is decided by user) of population, expand which via a noise vector, the manifold of the decision space is built

and guide algorithm convergence by this manifold. But the model-based multiobjective optimisation algorithm has its shortcomings also:

- 1 In the early stage, the distribution of population is disorder, so the model built by the algorithm usually far away from the actual Pareto set.
- 2 The model algorithm, such as: Local PCA or PCA are linear modelling algorithm which cannot work if the data sets are nonlinear and failure to modelling.

**Figure 1** The Pareto set of two-objective mops is 1 –  $D$  manifold (see online version for colours)

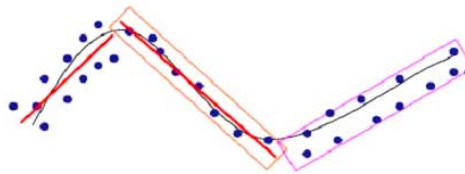


Note: As shown as the black solid line.

### 1.3 Research status of manifold-learning-based multiobjective evolutionary algorithm

So, in this paper, to overcome the shortcoming of the traditional evolutionary multi-objective optimisation algorithms (EMOAs) and MOEAs, the manifold learning algorithm is introduced into multiobjective optimisation algorithm for MOPs, and a manifold-learning-based multiobjective evolutionary algorithm (ML-MOEA) is proposed for reducing dimension of data and mining manifold in the decision space of MOPs, build accurate model, guide algorithm evolution and accelerate convergence.

**Figure 2** Firstly, RM-MEDA build cluster via local PCA (see online version for colours)



Note: Then, uses several segmented eigenvectors (red line) approximately denote manifold (the black solid line).

In Yang et al. (2009), proposed a novel hybrid multi-objective estimation of distribution algorithm by local linear embedding and an immune inspired algorithm. A local linear embedding based manifold algorithm is introduced to build the distribution model of promising solutions, for enhancing local search ability of EDA, an immune inspired sparse individual clone algorithm (SICA) is introduced and combined with the EDA.

From 2008 to 2012, based on EDAs and MOEAs, Guangming and Wang's (2009) of China University of Geosciences, introduced 'manifold learning algorithm' into multiobjective evolutionary algorithm, proposed a kind of multi-objective evolutionary

algorithm via manifold learning for MOPs, the algorithm can find smooth manifolds embedding in decision space in order to guide the algorithm fast convergence to Pareto front.

In Zhang et al. (2011), proposed multiobjective evolutionary algorithm for principal curve model based on multiracial. The algorithm uses principal curve to build nonlinear modelling on the distribution of the solution set and to establish the probability model on the individual distribution of population which can generate the individuals distribute evenly in the objective space and ensure the diversity of optimisation results.

In Li (2011), firstly proposed that fitting the piecewise continuous manifold based on regression analysis in the high-dimensional space. The population is just a lot of discrete data points in the decision space, so it is natural to think of finding a principal curve running through all of them by least square method. The new solutions would be sampled from the round of the principal curve. This paper applied the least squares model into the model-based multi-objective evolutionary algorithm, called MMEA-RA: model-based multi-objective evolutionary algorithm based on regression analysis. Experimental studies have shown that overall, MMEA-RA performs excellent on a set of test instances with variable linkages, especially for the ones with Pareto Set and Pareto Front uniformly distributed.

Based on this regularity property, Zhu (2011) proposed a hybrid multiobjective evolutionary algorithm based on self-organising feature maps (SOM) and genetic operation. At each generation, the algorithm apply self-organising feature map to capture and utilise the regularity by learning the manifold of the solutions. Some of new trial solutions are sampled from the SOM lattice, and the rest of solutions are generated from adaptive genetic operators.

Aim at feature of high-dimensional decision space about satellite constellation optimisation design, via self-organisation feature map (SOM), Zhan (2012) reduced decision space dimensionality of satellite constellation optimisation design. Experiments show that the optimisation results can meet the demand of coverage performance for specific region, but running speed of the algorithm should be improved, the reason is maybe that SOM modelling is time-consuming.

The rest of this paper is organised as follows. Section 2 introduces the continuous MOPs and the regularity property of continuous MOPs, define of manifold and manifold learning. Section 3 describes the details of the proposed algorithm. Section 4 outlines the future work.

## 2 Problem definition

### 2.1 Continuous MOPs

The following continuous multiobjective optimisation problem is considered in this paper:

$$\min_{x \in \Omega} F(x) = (f_1(x), f_2(x), \dots, f_m(x))^T \quad (1)$$

where  $(x_1, x_2, \dots, x_n)^T \in \Omega$  is the decision variable vector,  $\Omega \subseteq R^n$  is the decision space.  $f_i(x)$ ,  $i = 1, 2, \dots, m$  are the continuous objective functions to be minimised,  $R^m$  is the objective space. Let  $u = (u_1, u_2, \dots, u_m)^T$ ,  $v = (v_1, v_2, \dots, v_m)^T$  be two objective vectors,  $u$  is

said to dominate  $v$  (denoted by  $u \prec v$ ) if  $u_i \leq v_i$  for all  $i = 1, 2, \dots, m$ , and  $u \neq v$ . A solution  $x^* \in \Omega$  is called Pareto optimal or non-dominated solution if there is no  $x \in \Omega$  such that  $F(x) \prec F(x^*)$ . All the Pareto optimal solutions in the decision space are made up of the Pareto optimal set, denoted by  $PS$ . The corresponding image of the Pareto optimal set in the objective space is called the Pareto optimal front, denoted by  $PF = \{y \in F(x), x \in PS\}$ .

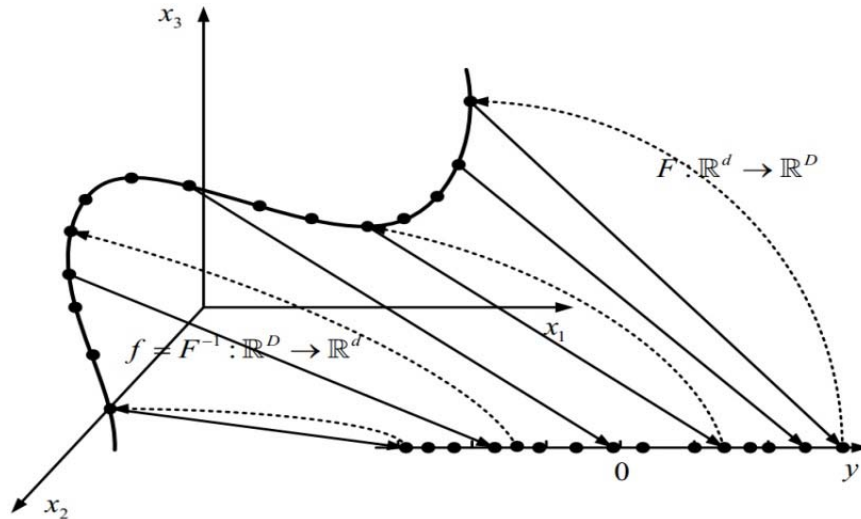
The distribution of the Pareto optimal set of a continuous MOP often shows a high degree of regularity. It can be induced from the Karush-Kuhn-Tucker condition that the PS of a continuous MOP is a piecewise continuous  $(m - 1) - D$  manifold in the decision space, where  $m$  is the number of the objectives (Zhang et al., 2006). Therefore, the PS of a continuous two-objective optimisation problem is a piecewise continuous curve in  $R^n$  (Zhou et al., 2005), while the PS of a continuous three-objective MOP is a piecewise continuous surface, and so on.

## 2.2 Manifolds and manifold learning

Manifold is an important concept in modern mathematics, it combines research area of topology, geometry and algebra and it is the most representative mathematics theory of 20th century. The word of ‘manifold’ has been proposed as early as in Riemannian doctoral dissertation, afterwards, Hilbert gave a rigorous mathematical definition, as follows:

Manifold:  $M$  is a non-empty Hausdorff space, for arbitrary point  $p \in M$ , there is an  $p$ 's open neighbourhood  $U \subseteq M$  of an open subset of  $d$ -dimension Euclidean space, so  $M$  is called  $d$ -dimension topological manifold. Manifold is extension of Euclidean space, in each of the tiny local neighbourhood, the manifold is smoothing and linear as same as in the Euclidean space, and it can be given local coordinates. The nature of manifold is a topological space can be coordinated in the local.

**Figure 3** Mathematics schematic of manifold learning



In Bregler and Omohundro (1995), proposed the term of ‘manifold learning’ in the paper of visual speech recognition and digital image processing. In de Silva and Tenenbaum (2003), the formalised mathematical description of the manifold learning are shown in Figure 3.

Given a set of high-dimensional data  $X = \{x_i \in R^D, i = 1, 2, \dots, N\}$  (among  $D$  is dimension of  $X$  and  $N$  is the number of sample data), and suppose  $X$  comes from a certain  $d$ -dimension ( $d \leq D$ ) smooth manifold. The target of manifold learning is that seeking low-dimensional embedding expression  $y_i \in E^d$  of  $x_i$ , seeking mapping function  $F$  from observation space to low-dimensional space and reconstruction mapping function  $F^{-1}$ , which makes  $y_i = F(x_i)$  and  $x_i = F^{-1}(y_i)$ , at the same time,  $F$  should be satisfy certain constraints in order to keep the global or local geometric structure of the original high-dimensional data in the embedding low-dimensional space.

Manifold learning is one kind of unsupervised learning algorithm, it can find out the low-dimensional smooth manifold embedded in high-dimensional data. In this paper, the manifold learning algorithm is introduced into multiobjective optimisation algorithm for MOPs.

### 3 The ML-MOEA

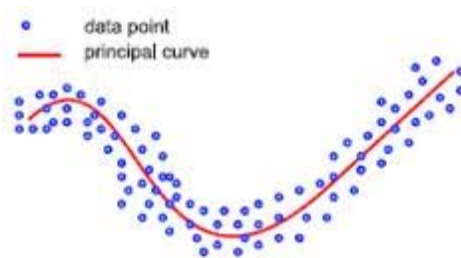
#### 3.1 Basic idea

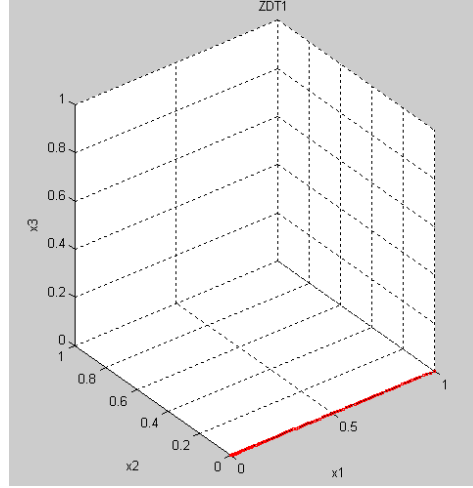
According Karush-Kuhn-Tucker Condition, there are two kind of continuous MOPs,

- 1 The first type: The Pareto front and Pareto set of MOPs are the same dimensional, and the Pareto Set is a piecewise continuous  $(m - 1) - D$  manifold in the decision space, where  $m$  is the number of the objectives
- 2 The second type: The Pareto front is a continuous manifold (where is the number of the objectives), but the Pareto set is a continuous manifold which dimensional is higher than the Pareto front.

The object of this paper’s study object is the above second type MOPs, as shown as Figure 4, the Pareto set of two objectives is a principal curve. As shown as Figure 5, the Pareto set of three objectives is a principal plane.

**Figure 4** A principal curve (see online version for colours)



**Figure 5** Pareto set of ZDT1 is a line (see online version for colours)

### 3.2 Unified algorithm framework of ML-MOEA

Set the population size of ML-MOEA(Zhan and Dai, 2013) is  $N$ , denote by  $Pop(t) = \{x_1, x_2, \dots, x_N\}$ , where  $t$  is generation of algorithm running; The fitness of individual is  $\bar{F}(x) = (f_1(x), f_2(x), \dots, f_m(x))^T$ , where  $m$  is the number of objectives.

#### 3.2.1 Framework of ML-MOEA

- Step 1 *Initialisation*: Set  $t = 0$ , randomly generate initial population  $Pop(0)$  and compute the corresponding objective values of each solution.
- Step 2 *Modelling via manifold learning*: Mining manifold in the current population  $Pop(t)$  via manifold learning algorithm.
- Step 3 *Extend and reproduction*: Set the number of generation is  $N$ , one hand, generate  $N_1$  individual of offspring via manifold learning; on the other hand, generate  $N_2$  individual of offspring via genetic operation, obviously,  $N = N_1 + N_2$  and generation of offspring is denoted by  $Q(|Q| = N_1 + N_2)$ , then, compute fitness ( $\bar{F}$ ) of  $Q$ . Combine  $Q$  and  $Pop(t)$ , which is denote it by  $Q \cup Pop(t)$ .
- Step 4 *Elite selection*: Select out  $N$  solutions from  $Q \cup Pop(t)$  with elite selecting strategy which was proposed in NSGA-II.
- Step 5 *Stopping condition*: If the stopping condition is met, stop and return the non-dominated solutions in  $Pop(t)$ ; otherwise, set  $t = t + 1$  and go to step 2.

### 3.3 Modelling via manifold learning algorithm

Since Seung and Lee (2000), Roweis and Saul (2000) and Tenenbaum et al. (2000) in Science, successively published three classical papers about concept of manifold

learning. A number of manifold learning algorithms have spring up. For different problem type of MOPs, the corresponding manifold learning algorithm can be selected to design ML-MOEA, because different manifold learning algorithms have different feature.

### 3.4 Extend and reproduction

Set the number of generation is  $N$ , one hand, generate  $N_1$  individual of offspring via manifold learning algorithm; on the other hand, generate  $N_2$  individual of offspring via genetic operation, obviously,  $N = N_1 + N_2$  and generation of offspring is denoted by  $Q(|Q| = N_1 + N_2)$ , then, compute fitness ( $\bar{F}$ ) of  $Q$ . Combine  $Q$  and  $Pop(t)$ , which is denote it by  $Q \cup Pop(t)$ .

As described above,

- 1 in the early stage of ML-MOEA, generate most solutions via genetic operators to accelerate convergence
- 2 in the later stage of ML-MOEA, build model of population and then generate most solutions via manifold learning algorithm.

Extend and reproduction work as follows.

- Step 1 *Adaptive strategy*:  $N_2$  new solutions are generated by genetic operators. Where  $N_2$  is defined as  $N_2 = 0.7Ne^{-t/T}$ , where  $N$  is the population size,  $t$  is the current generation of algorithm,  $T$  is the max generation of algorithm. Obviously,  $N_2 \approx 0.7N$  at the beginning of algorithm and  $N_2$  is decreasing as  $t$  is increasing and  $N_2 = 0.7Ne^{-1}$ , finally.
- Step 2 *Crossover and mutation*: Perform crossover and mutation on  $Pop(t)$  to create  $N_2$  new solutions. The crossover and mutation operators used in our algorithm are multiparent crossover (Deb et al., 2000) and polynomial mutation (Pan et al., 2000), respectively.

### 3.5 Elite selection

This selection procedure is the same as that used in NSGA-II (Deb et al., 2000) except that we remove solutions from one by one and we recalculate the crowding distances before deciding which solution should be deleted from , which can increase the diversity of at extra computational cost.

## 4 The future work

The future research topics along this line should include the following:

- 1 Analyse algorithm characteristics of self-organising maps (SOM), design the ML-MOEA via SOM (ML-MOEA/SOM), through neurons of SOM, learn manifold in population, then reproduce individuals via expanding SOM neurons by random noise vector, in order to establish manifold of problem's decision space and guide algorithm evolution.



- 2 Analyse algorithm characteristics of via locally linear embedding (LLE), and design the ML-MOEA via LLE (ML-MOEA/LLE). It is different from ML-MOEA/SOM, in ML-MOEA/LLE, ‘expansion population’ strategy in ML-MOEA/SOM will not be used, but directly reproduce individuals of next population from data points on the ‘manifold’, which will ensure that nonlinear model of decision space can be build.
- 3 Though numerical experiments and contrast algorithm performance between ML-MOEA/SOM and ML-MOEA/LLE.

## References

- Boyd, S. and Vandenberghe, L. (2004) *Convex Optimization*, p.244, Cambridge University Press, Cambridge.
- Bregler, C. and Omohundro, S.M. (1995) ‘Nonlinear image interpolation using manifold learning’, *Advances in Neural Information Processing Systems*, p.7, MIT Press.
- Bregler, C. and Omohundro, S.M. (1995) ‘Nonlinear manifold learning for visual speech recognition’, *The Fifth International Conference on Computer Vision*, California University, Berkeley, CA, pp.494–499.
- Dai, G. and Wang, M. (2009) *Multiobjective Optimization Algorithm and Its Application of Satellite Constellation Design*, China University of Geosciences Press, Wuhan, China.
- de Silva, V. and Tenenbaum, J.B. (2003) ‘Global versus local methods in nonlinear dimensionality reduction’, *Neural Information Processing Systems*, Vol. 15, pp.705–712.
- Deb, K., Agrawal, S., Pratap, A. and Meyarivan, T. (2000) *A Fast and Elitist Multi-Objective Genetic Algorithm: NSGA-II*, Danpur Genetic Algorithms Laboratory (KanGAL), Indian Institute of Technology, Kanpur, Technical Report 200001.
- Fonseca, C.M. and Fleming, P.J. (1993) ‘Genetic algorithms for multi-objective optimization: formulation, discussion and generation’, *Proceedings of the 5th International Conference on Genetic Algorithms*, San Mateo, California.
- Li, M. (2011) *Model-based Multiobjective Evolutionary Algorithm Based on Regression Analyses*, China University of Geosciences, Wuhan, China.
- Pan, Z., Kang, L. and Chen, Y. (2000) *Evolutionary Computation*, Tinghua University Press, Beijing, China.
- Rosenberg, R.S. (1967) *Simulation of Genetic Populations with Biochemical Properties*, University of Michigan, Ann Harbor, Michigan.
- Roweis, S.T. and Saul, L.K. (2000) ‘Nonlinear dimensionality reduction by locally linear embedding’, *Science*, Vol. 290, No. 22, pp.2323–2326.
- Schaffer, J.D. (1985) ‘Multiple objective optimization with vector evaluated genetic algorithms’, *Proceedings of the 1st International Conference on Genetic Algorithms*, Lawrence Erlbaum, Carnegie-Mellon University, pp.93–100.
- Seung, H.S. and Lee, D. (2000) ‘The manifold ways of perception’, *Science*, Vol. 290, No. 5500, pp.2268–2269.
- Tenenbaum, J.B., de Silva, V. and Langford, J.C. (2000) ‘A global geometric framework for nonlinear dimensionality reduction’, *Science*, Vol. 290, No. 5500, pp.2319–2323.
- Yang, D., Jiao, L., Gong, M. and Feng, X. (2009) ‘Hybrid multiobjective estimation of distribution algorithm by local linear embedding and an immune inspired algorithm’, *Proceedings of the 2009 IEEE Congress on Evolutionary Computation*, pp.463–470.
- Zhan, W. (2012) ‘Low earth orbit regional satellite constellation design via self organization feature maps’, *International Journal of Advancements in Computing Technology*, Vol. 4, No. 7, pp.250–260.

- Zhan, W. and Dai, G. (2013) 'Research overview of manifold learning algorithm', *Advances in Information Sciences and Service Sciences*, Vol. 5, No. 1, pp.58–67.
- Zhang, D., Gong, X. and Dai, G. (2011) 'Multi-objective evolutionary algorithm for principal curve model based on multifractal', *Journal of Computer Research and Development*, Vol. 48, No. 9, pp.1729–1739.
- Zhou, A., Zhang, Q. and Jin, Y. (2009) 'Approximating the set of Pareto optimal solutions in both the decision and objective spaces by an estimation of distribution algorithm', *IEEE Transaction on Evolutionary Computation*, Vol. 13, No. 5, pp.1167–1189.
- Zhou, A., Zhang, Q., Jin, Y., Tsang, E. and Okabe, T. (2005) 'A model-based evolutionary algorithm for bi-objective optimization', *Congress on Evolutionary Computation*, pp.2568–2575.
- Zhu, J. (2011) *A Manifold Learning-Based Multi-Objective Estimation of Distribution Algorithm*, China University of Geosciences, Wuhan, China.