This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI10.1109/TNNLS.2023.3295461, IEEE Transactions on Neural Networks and Learning Systems

A Novel Dynamic Multi-objective Optimization Algorithm with Non-inductive Transfer Learning Based on Multi-strategy Adaptive Selection

Han Li, Zidong Wang, *Fellow*, *IEEE*, Chengbo Lan, Peishu Wu, and Nianyin Zeng

***Abstract*—In this paper, a novel multi-strategy adaptive selection-based dynamic multi-objective optimization algorithm (MSAS-DMOA) is proposed, which adopts the non-inductive transfer learning paradigm to solve dynamic multi-objective optimization problems (DMOPs). In particular, based on a scoring system that evaluates environmental changes, the source domain is adaptively constructed with several optional groups to enrich the knowledge. Along with a group of guide solutions, the importance of historical experiences is estimated via the kernel mean matching (KMM) method, which avoids designing strategies to label individuals. The proposed MSAS-DMOA is comprehensively evaluated on 14 DMOPs, and the results show an overwhelming performance improvement in terms of both convergence and diversity as compared with other four popular DMOAs. In addition, ablation studies are also conducted to validate the superiority of the applied strategies in MSAS-DMOA, which can effectively alleviate the negative transfer phenomenon. Without the conventional labeling procedure, the proposed method also yields satisfactory results, which can provide valuable reference for designing other evolutionary transfer optimization (ETO) algorithms.**

Keywords: Dynamic multi-objective optimization algorithm (DMOA); evolutionary transfer optimization (ETO); kernel mean matching (KMM); transfer learning (TL)

I. Introduction

Dynamic multi-objective optimization problems (DMOPs) are frequently encountered in various real-world scenes [22], [30], [34], [44], which consist of multiple conflicting objectives with time-varying characteristics. Particularly, due to the existence of the dynamic behaviors, both constraints and objective functions in DMOPs are changeable with time [6]. As a result, the dynamic multi-objective optimization algorithms (DMOAs) are required to accurately and rapidly converge to the changing Pareto front (PF), and how to cope with the dynamic property (e.g., the frequency, degree, and type of changes) has aroused great research interests.

To deal with the dynamic behaviors in DMOPs, it is feasible and reliable to apply the population-based evolutionary algorithms (EAs) [3], [29], [32], [33], [45], [47], [51], and a great number of EA-based DMOAs have been proposed [8], [16], [26], which can be generally divided into diversity-, memory- and prediction-based methods [1]. It is worth mentioning that, in recent years, a novel paradigm that integrates EA with knowledge learning and transfer across related domains is emerging, which is known as the evolutionary transfer optimization (ETO) [40]. In particular, ETO algorithms for solving DMOPs are dedicated to mining the potential associations between successive changing environment so as to accelerate solution convergence. In this regard, the related methods adopt the transfer learning technique, where the advantages of both memory and prediction mechanisms are integrated, see [15]–[17] for some successful applications.

It should be pointed out that the source domain ($\mathcal{D}_s$) has played an important role in transfer learning [36], which contains rich useful knowledge that is required to assist learning in the target domain ($\mathcal{D}_t$). In the context of applying ETO methods to handle DMOPs, the essence is to learn from previous Pareto solutions and provide an initialized population with high quality in the new environment to accelerate the convergence. A noticeable issue is that some ETO methods have directly employed the Pareto solutions in the previous environment as the $\mathcal{D}_s$ [15], [16], which implicitly assumes that there is a strong correlation between the two successive environments, otherwise selecting $\mathcal{D}_s$ in this way would make little sense. In addition, it is common to adopt a group of individuals with high quality in the new environment to construct the $\mathcal{D}_t$ in many existing studies, and afterwards, some strategies have been designed to label the samples in $\mathcal{D}_t$ [15], [16], [52].

Based on above discussions, two important issues that deserve further attention are listed as follows. 1) It is of vital significance to figure out what the valuable information in previous search is that can be employed to construct $\mathcal{D}_s$, and the previous Pareto solutions are not always helpful in changing environments. 2) In ETO methods that adopt the sample-based inductive transfer learning (TL) paradigm, it is difficult to guarantee the reliability of labels attached on individuals, due to the fact that available data in $\mathcal{D}_t$ may be limited and class-imbalanced, which can lead to great bias in subsequent model training.

To overcome the above mentioned challenges in developing ETO methods for solving DMOPs, a novel multi-strategy adaptive selection-based dynamic multi-objective optimization algorithm (MSAS-DMOA) is proposed in this paper. Particularly, on the basis of Pareto solutions in previous environment,

This work was supported in part by the Natural Science Foundation of China under Grant 62073271, the National Science and Technology Major Project of China under Grant J2019-I-0013-0013, the Independent Innovation Foundation of AECC of China under Grant ZZCX-2018-017, and the Korea Foundation for Advanced Studies. (*Corresponding author: Nianyin Zeng*)

H. Li, C. Lan, P. Wu and N. Zeng are with the Department of Instrumental and Electrical Engineering, Xiamen University, Fujian 361005, China. Email: zny@xmu.edu.cn

Z. Wang is with the Department of Computer Science, Brunel University London, Uxbridge UB8 3PH, U.K. Email: zidong.wang@brunel.ac.uk

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI10.1109/TNNLS.2023.3295461, IEEE Transactions on Neural Networks and Learning Systems

several optional groups are first formed to cover rich information. Then, by a designed scoring system, individuals from each group are selected with different proportions to construct the $\mathcal{D}_s$ in an adaptive manner. It is noticeable that the environmental changes are taken into account in the scoring system and, by doing so, values of the previous searching experiences can be measured in each specific case, which benefits the adaptation to unknown environment. Furthermore, a number of guide solutions are selected from the new environment to construct the target domain $\mathcal{D}_t$. Those guide solutions are deemed to have relative high quality, which can serve as the reference to avoid potential negative transfer phenomenon, and after $\mathcal{D}_t$ is constructed, the kernel mean matching (KMM) method [11] is applied to match the data distribution between the two domains. Instead of selecting a standard to attach the pseudo-labels to samples in $\mathcal{D}_t$ (e.g., the non-domination-based [16] and knee-point-based [15] strategy), in the proposed MSAS-DMOA, $\mathcal{D}_t$ is adopted to assign weights to individuals in $\mathcal{D}_s$ via the KMM method, which can further reflect the importance of each selected historical solution according to the distribution similarity. Consequently, not only the subjective influence in designing label strategy can be avoided, but the trained prediction model can also be robust and adaptive to various unknown environments.

The major contributions of this paper are outlined as follows.

1. A novel non-inductive TL-based ETO algorithm is proposed to solve DMOPs, whose main idea is to design a multi-strategy adaptive selection mechanism to construct the source domain with rich useful knowledge.

2. The KMM method is introduced to match the distribution between $\mathcal{D}_s$ and $\mathcal{D}_t$ so that different weights can be assigned to the historical solutions, which avoids subjectivity when attaching pseudo labels on $\mathcal{D}_t$.

3. The proposed framework is quite universal that can be integrated into any static multi-objective optimization algorithms (SMOAs) to solve the dynamic problems, thereby exhibiting the strong generalization ability.

The remainder of this paper is organized as follows. Preliminaries of this work are provided in Section II. The proposed MSAS-DMOA is comprehensively elaborated in Section III. Experimental results and discussions are presented in Section IV, and finally, conclusions are drawn in Section V.

II. Preliminaries

A. Formulation of DMOPs

Without loss of generality, a minimized DMOP can be formulated as:

$$\begin{cases} \min\ F(\boldsymbol{x},t)=\{f_1(\boldsymbol{x},t),\ f_2(\boldsymbol{x},t),\ ...,\ f_m(\boldsymbol{x},t)\} \\ \qquad s.t.\ g_i(\boldsymbol{x},t)\le 0,\ h_j(\boldsymbol{x},t)=0 \\ \qquad\ i=1,\ 2,\ ...,\ p,\ j=1,\ 2,\ ...,\ q \end{cases} \tag{1}$$

where $\boldsymbol{x}\in\mathbb{R}^n$ is decision vector, t is time variable and $F(\cdot)$ contains m objective functions. $g(\cdot)$ and $h(\cdot)$ are inequality and equality constraints, respectively. Based on above formulation, some definitions are provided as follows.

Definition 1: Dynamic Pareto domination.

At time t, decision vector $\boldsymbol{x}_1$ is regarded to dominate $\boldsymbol{x}_2$ (denoted as $\boldsymbol{x}_1 \succ_t \boldsymbol{x}_2$) only in conditions of:

$$\begin{cases} \forall\ i\in\{1,\ 2,\ ...,\ m\},\ f_i(\boldsymbol{x}_1,t)\le f_i(\boldsymbol{x}_2,t) \\ \exists\ j\in\{1,\ 2,\ ...,\ m\},\ f_j(\boldsymbol{x}_1,t)< f_j(\boldsymbol{x}_2,t) \end{cases} \tag{2}$$

Definition 2: Dynamic Pareto set.

At time t, if none of individuals can dominate $\boldsymbol{x}^*$, then $\boldsymbol{x}^*$ is called a Pareto solution of (1), and all such solutions consist of current dynamic Pareto set (denoted as PS_t):

$$PS_t=\{\boldsymbol{x}^*|\forall\ \boldsymbol{x}\in\mathbb{R}^n,\ \boldsymbol{x}\prec_t \boldsymbol{x}^*\} \tag{3}$$

where $\prec_t$ represents non-domination relationship opposite to $\succ_t$.

Definition 3: Dynamic Pareto front.

At time t, the dynamic Pareto front (denoted as PF_t) is the corresponding mapping results of PS_t in objective space:

$$PF_t=\{F(\boldsymbol{x},t)\mid \boldsymbol{x}\in PS_t\} \tag{4}$$

It should be pointed out that the focus of DMOAs is to handle the dynamic behaviors in DMOPs so that the algorithms can converge to the time-varying PF_t accurately and rapidly. As for obtaining the Pareto solutions in each individual environment, any SMOA (e.g., MOEA/D [53]) can be directly applied as the optimizer.

B. Transfer Learning Paradigm

The core idea of transfer learning is to apply the gained knowledge from previous tasks to assist solving different but related problems. According to [36], a *domain* $\mathcal{D}$ is composed of the sample space $\mathcal{X}$ and marginal probability distribution $P(X)$ of the samples therein, which can be denoted by $\mathcal{D}=\{\mathcal{X},P(X)\}$ where $X\in\mathcal{X}$. Aiming at above domain $\mathcal{D}$, a *task* can be described as $\mathcal{T}=\{\mathcal{Y},f(\cdot)\}$ where $\mathcal{Y}$ is the label space, and the predictor $f(\cdot)$ can be learned from data $\{X\in\mathcal{X},Y\in\mathcal{Y}\}$. Accordingly, $Y=f(X)$ means the label prediction for sample X, and the predictor $f(\cdot)$ can be written as $P(Y|X)$ in the probability form.

With above notations, the aim of TL is to promote the learning of a predictor $f_t(\cdot)$ in target domain $\mathcal{D}_t$ with the obtained knowledge from the source domain $\mathcal{D}_s$ and the task $\mathcal{T}_s$, where $\mathcal{D}_s\neq\mathcal{D}_t$ or $\mathcal{T}_s\neq\mathcal{T}_t$. According to [36], in case of $\mathcal{T}_s\neq\mathcal{T}_t$, it is named inductive TL, where some labeled data in $\mathcal{D}_t$ should be available to induce an objective predictor $f_t(\cdot)$. If it satisfies $\mathcal{T}_s=\mathcal{T}_t$ and $\mathcal{D}_s\neq\mathcal{D}_t$, the term of "transductive" TL has been used in [36]. In addition, when labels in neither $\mathcal{D}_s$ nor $\mathcal{D}_t$ is available, such case is termed as the unsupervised TL.

It is worth mentioning that in this study, only two terms of the *inductive* and *non-inductive* transfer learning are used, and the former/latter refers to the case where labels are available/unavailable in $\mathcal{D}_t$. For a clear view, the adopted taxonomy according to the availability of labeled data is presented in Fig. 1, which benefits a better understanding of this study, and one can refer to [36] for more details of the different TL paradigms.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI10.1109/TNNLS.2023.3295461, IEEE Transactions on Neural Networks and Learning Systems

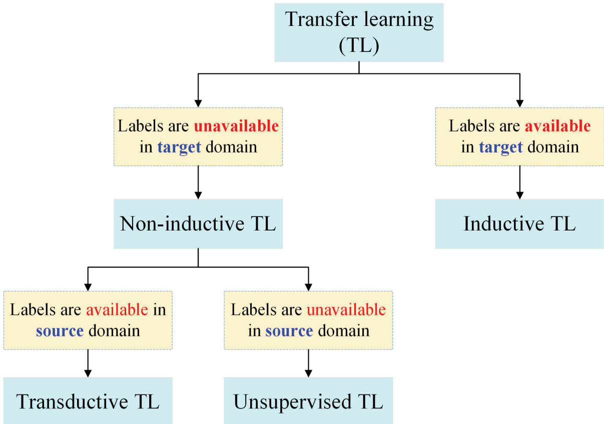


Fig. 1. A simplified taxonomy of different TL paradigms [36].

Remark 1: In this study, both the transductive and unsupervised TL are termed as the non-inductive TL, due to the major focus is whether the samples in target domain $\mathcal{D}_t$ are labeled via specifically designed strategies or not.

C. ETO Methods

Following the idea of TL, the population is encouraged to take full use of the previous searching experiences in ETO methods. In [35], historical solutions obtained from different tasks have been considered to train a generalized model, and an adaptive knowledge transfer framework has been proposed to address the computationally expensive surrogate-assisted evolutionary search problems. A centroid distribution-based framework has been developed in [52], where the probabilistic model of elite candidate solutions is adopted to balance the transfer correlation between multiple domains, which has proven effective in avoiding negative heterogeneous transfer. Additionally, one can refer to [24], [25], [28], [46] for more successful applications of ETO methods in various domains.

In DMOPs, it is noticeable that the Pareto solutions are usually deemed to contain rich valuable information, and as a result, many related ETO algorithms have paid attention to PS in the previous environment (denoted as PS_{t-1}). For example, in [16], PS_{t-1} has been selected as the source domain, where the idea of TrAdaboost [5] is adopted to predict movement tendency of the optimal solutions on the basis of imbalanced knee-points. It should be pointed out that in changing environments, not all solutions in PS_{t-1} are useful, and how to estimate the importance of historical searching experience should be concerned. In [27], aiming at the multi-objective multi-task optimization problems, the authors have investigated how to find valuable solutions to promote the positive knowledge transfer.

III. Methodology

In this section, the proposed MSAS-DMOA is elaborated with details. Briefly, according to the estimation of varying environment, individuals with various kinds of characteristics are adaptively selected to form the source domain $\mathcal{D}_s$, which is conducive to enriching the knowledge so as to promote reliability of the transfer procedure. In addition, a group of guide solutions in the new environment is generated to alleviate the negative transfer, which is mapped into a high-dimensional space along with individuals in $\mathcal{D}_s$. Then, via the KMM method, solutions in $\mathcal{D}_s$ can be assigned with different learning weights to train the prediction model. Finally, above predictor will output an initialized population, which is deemed adaptive to the new environment so as to accelerate the convergence.

A. Environmental Change Evaluation

In a DMOP, it is generally believed that some invisible correlations may exist between the successive environment if the change degree is within a certain range [14]. Following this idea, in the proposed MSAS-DMOA, a population-based evaluation mechanism is designed to characterize the dynamic behaviors in DMOPs, where two changing rates R_d and R_c are defined to reflect fluctuations of the fitness value and diversity, respectively. To be specific, the average distance to the ideal point of solutions in PF_{t-1} is considered in R_d, which can be depicted as:

$$Dis_t = \frac{\sum_{\boldsymbol{x}\in PS_{t-1}} \|F(\boldsymbol{x},t) - P_I\|_2}{|PS_{t-1}|} \tag{5}$$

where t denotes current time, $\|\cdot\|_2$ refers to Euclidean distances, $|\cdot|$ is cardinality of a set and P_I is the ideal point of PF_{t-1}. Referring to the ideal point P_I, Dis_t evaluates the degree that PS_{t-1} has changed. To calculate R_c, following coverage scope is defined and used:

$$Cs_t = \frac{\sum\limits_{\boldsymbol{x}\in PS_{t-1}} \max\limits_{\boldsymbol{z}\in PS_{t-1}} \{\|F(\boldsymbol{x},t) - F(\boldsymbol{z},t)\|_2\}}{|PS_{t-1}|} \tag{6}$$

where Cs_t calculates the average distance between each individual and the farthest one in population, which also evaluates the dispersion extent of PS_{t-1} in the new environment. Based on Eqs. (5)-(6), R_d and R_c are given as:

$$\begin{aligned} R_d &= \frac{|Dis_t - Dis_{t-1}|}{Dis_{t-1}} \\ R_c &= \frac{Cs_t - Cs_{t-1}}{Cs_{t-1}} \end{aligned} \tag{7}$$

By employing R_d and R_c, changes of PS_{t-1} in terms of both location and dispersion can be characterized, which benefits designing different response strategies and details are presented in subsequent subsections. Notice that above two rates are calculated only on the selected sensors (including 50% of individuals in PS_{t-1}) to save the computational resources.

B. Multi-strategy Selection Mechanism

In the Pareto solutions, some inherent information is carried by the special points [21]. For instance, the knee-point represents large gradient that is related to the diversity, center-point can be used to predict the tendency of population movement. Considering that knowledge in $\mathcal{D}_s$ can be enriched by the information of special points, four optional groups are formed to construct $\mathcal{D}_s$, including G_{ps}, G_{re}, G_{kp} and G_{cp}, whose members are introduced in Table I.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI10.1109/TNNLS.2023.3295461, IEEE Transactions on Neural Networks and Learning Systems

TABLE I
FOUR OPTIONAL GROUPS FOR CONSTRUCTING $\mathcal{D}_s$

Groups	Members
G_{ps}	PS_{t-1}
G_{re}	Solutions generated in new environment
G_{kp}	Individuals generated around knee-point
G_{cp}	Individuals generated around center-point

In addition, to realize the adaptive construction of $\mathcal{D}_s$, following scoring system is designed on the basis of Eq. (7) as:

$$\mathscr{F}(S)=\begin{cases}\max\{0,\ 1-R_d\}, & S=G_{ps}\\ \min\{R_d,\ 1\}, & S=G_{re}\\ -\min\{0,\ R_c\}, & S=G_{kp}\\ 0.2, & S=G_{cp}\end{cases} \tag{8}$$

As is shown, scores of G_{ps} and G_{re} are dependent on R_d, which reflect the severity of environmental changes. Notice that $\mathscr{F}(G_{ps})+\mathscr{F}(G_{re})=1$, which implies that the generated individuals in the new environment are complementary to Pareto solutions in the previous one. In an extremely fluctuated situation, knowledge of Pareto solutions in the previous environment may not well adapt to the new one and introducing some new solutions can be helpful. On the contrary, if R_d is small, a fine-tuning on PS_{t-1} is recommended and a few newly generated individuals can supplement diversity to some extent.

Since shrinkage of the PF coverage scope is reflected in R_c, $R_c \geq 0$ implies an unchanged and even better diversity, therefore, only in condition of $R_c < 0$ that individuals in G_{kp} are adopted to supplement the population diversity. It is noticeable that the score of G_{cp} is set to a constant 0.2, which is required to guarantee that even in dramatically changed situations, partial historical information can be retained by reserving a few solutions around PS center. According to the scores calculated by Eq. (8), the weight of each group is allocated as:

$$\omega(G)=\frac{\mathscr{F}(G)}{1+\mathscr{F}(G_{kp})+\mathscr{F}(G_{cp})} \tag{9}$$

where $G \in \{G_{ps}, G_{re}, G_{kp}, G_{cp}\}$. As a result, an adaptive construction of $\mathcal{D}_s$ is realized, and the overall procedure is displayed in Alg. 1 and Fig. 2.

C. Guide Solutions

To alleviate the negative knowledge transfer that may lead the evolution towards wrong direction, a group of guide solutions is generated in the new environment to serve as the reference. Accordingly, it should be addressed that how to screen the guide solutions, and to handle this issue, following ϵ-indicator [2] is employed in the proposed MSAS-DMOA to evaluate the quality of individuals:

$$I_\epsilon(\boldsymbol{x}_1,\boldsymbol{x}_2)=\max\{F(\boldsymbol{x}_1,t)-F(\boldsymbol{x}_2,t)\} \tag{10}$$

where $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ are different solutions, and I_ϵ reflects the largest difference in dimension-wise therein. For instance,

Algorithm 1 Adaptive construction of $\mathcal{D}_s$

Input:
Pareto set in the previous generation PS_{t-1} and a predefined capacity of source domain n_s
Output:
Source domain $\mathcal{D}_s$

```
Initialize D_s ← ∅
Select sensors X from PS_{t-1} such that |X| = ½|PS_{t-1}|
Characterize dynamics based on Eqs. (5)-(6) by x ∈ X
Calculate R_d and R_c by Eq. (7)
G_ps ← PS_{t-1}
Randomly generate solutions in new environment as G_re
Get knee-points in multiple sub-spaces of PS_{t-1} and generate G_kp
Obtain center-point of PS_{t-1} and establish G_cp
Allocate weights for each group according to Eqs. (8)-(9)
For G ∈ {G_ps, G_re, G_kp, G_cp}
    Pick ñ = n_s × ω(G) individuals from G as x̃ = {x_i}_{i=1}^{ñ}
    D_s ← x̃
EndFor
Return D_s
```

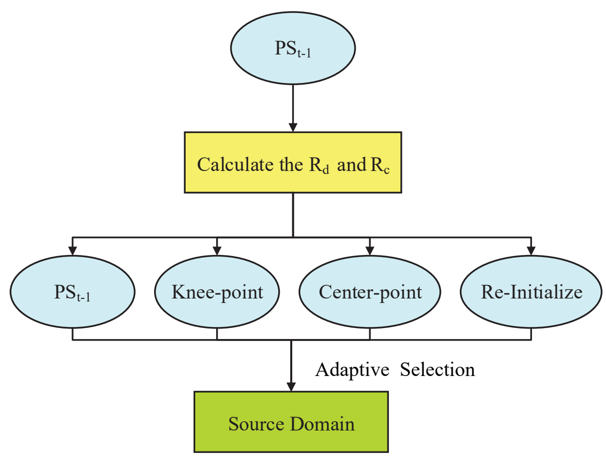


Fig. 2. Adaptive construction of $\mathcal{D}_s$.

given that $F(\boldsymbol{x}_1)=[1,2,3]^T$ and $F(\boldsymbol{x}_2)=[0,4,1]^T$, then $I_\epsilon(\boldsymbol{x}_1,\boldsymbol{x}_2)=\max\{1-0,\ 2-4,\ 3-1\}=2$. Based on Eq. (10), the quality factor of a solution $\boldsymbol{x}$ is defined as:

$$Q(\boldsymbol{x})=\max_{\boldsymbol{z}\neq\boldsymbol{x}}\{I_\epsilon(\boldsymbol{x},\boldsymbol{z})\} \tag{11}$$

where $\boldsymbol{z}$ is another individual different from $\boldsymbol{x}$ in the population, and a larger Q value corresponds to a better solution. By adopting above $Q(\cdot)$, the guide solutions can be screened to ensure that the knowledge is transferred towards a proper direction, and the details are displayed in Alg. 2. It should be pointed out that in the proposed MSAS-DMOA, the target domain $\mathcal{D}_t$ is formed with those eventually selected guide solutions, but different from several existing work, the guide solutions are not attached with pseudo-labels via some specified strategies, and more information is presented in the next subsection.

Remark 2: Mutation operator in Alg. 2 is used to enrich the

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI10.1109/TNNLS.2023.3295461, IEEE Transactions on Neural Networks and Learning Systems

Algorithm 2 Generation of guide solutions

Input:
Mutation operator $\mathcal{M}$ and mutation probability p_M
Output:
Target domain $\mathcal{D}_t$ consisting of guide solutions
Initialize $\mathcal{D}_t = \emptyset$
Randomly generate two populations P_1 and P_2 in the new environment with N individuals
For n from 1 to N
  $\boldsymbol{x}_n^1, \boldsymbol{x}_n^2 \leftarrow n$-th individual in P_1 and P_2
  Generate two random numbers $r_1, r_2 \in (0, 1)$
  For i from 1 to 2
    If $r_i < p_M$
      $\boldsymbol{z}^i = \mathcal{M}(\boldsymbol{x}_n^i)$
      If $Q(\boldsymbol{z}^i) < Q(\boldsymbol{x}_n^i)$
        $\boldsymbol{x}_n^i \leftarrow \boldsymbol{z}^i$
      EndIf
    EndIf
  EndFor
  $\mathcal{D}_t \leftarrow \underset{\boldsymbol{x} \in \{\boldsymbol{x}_n^1, \boldsymbol{x}_n^2\}}{\arg\max}\ Q(\boldsymbol{x})$
EndFor
Return $\mathcal{D}_t$

candidate solutions so as to enhance the diversity.

D. Individual-based Non-inductive Transfer Learning

In a DMOP, it is worth mentioning that labeled data are generally accessible only in $\mathcal{D}_s$, which refer to those already converged solutions in previous search, and such individuals with labels are quite lacked in the target domain as the population enters a new environment. In addition, the non-domination relationship among solutions may also change in the new environment, which leads to the uncertainty of labels in $\mathcal{D}_s$ and impedes the training of an accurate population predictor. To solve above issues, some studies have employed a certain standard to label individuals in $\mathcal{D}_t$. For example, based on the domination relationship, individuals in $\mathcal{D}_t$ are labeled with ± 1 in [16], respectively; in [15], labels are attached according to whether a solution is the knee-point. A noticeable issue is that there may be subjectivity in manual intervention when selecting the criterion, and in the proposed MSAS-DMOA, we aim to avoid labeling individuals in $\mathcal{D}_t$ to alleviate potential subjective effects, which is in the hope of bringing some new ideas to the related studies.

According to Section II-B, the primary purpose of TL is to train a predictor for the target domain $\mathcal{D}_t$, suppose that the model parameter is denoted as θ_t, on the basis of empirical risk minimization, the optimal θ_t can be obtained as follows:

$$\theta_t^* = \underset{\theta_t}{\arg\min}\, \mathbb{E}_{(x,y)\in\mathcal{D}_t}\left[L\big((x,y);\theta_t\big)\right] \tag{12}$$

where $L(\cdot)$ is the loss function regarding to θ_t, $\mathbb{E}$ denotes expectation value and (x, y) is the sample in $\mathcal{D}_t$. As above mentioned, labels in $\mathcal{D}_t$ are always unavailable and there are plenty of labeled data in $\mathcal{D}_s$, therefore, above Eq. (12) is rewritten in following form:

$$\begin{aligned}\theta_t^* &= \underset{\theta_t}{\arg\min}\, \mathbb{E}_{(x,y)\in\mathcal{D}_s}\left[\frac{P_t(x,y)}{P_s(x,y)}L\big((x,y);\theta_t\big)\right] \\ &= \underset{\theta_t}{\arg\min}\, \mathbb{E}_{(x,y)\in\mathcal{D}_s}\left[\frac{P_t(y|x)P_t(x)}{P_s(y|x)P_s(x)}L\big((x,y);\theta_t\big)\right]\end{aligned} \tag{13}$$

where P denotes the data distribution.

It is worth mentioning that in DMOPs, the task in each environment is consistent, which is to minimize the objective functions. Consequently, the conditional probability of $\mathcal{T}_s$ and $\mathcal{T}_t$ is the same, that is, there will be $P_s(y|x) = P_t(y|x)$ in Eq. (13). Furthermore, given a density ratio defined as $\beta_i = P_s(x_i)/P_t(x_i)$ $(x_i \in \mathcal{D}_s)$, and let $\mathcal{D}_s = \{(x_i, y_i)\}_{i=1}^{n_s}$ denote the source domain, then above Eq. (13) can be further converted into following form:

$$\theta_t^* = \underset{\theta_t}{\arg\min} \sum_{i=1}^{n_s}[\beta_i L((x_i, y_i);\theta_t)] \tag{14}$$

where n_s is the cardinality of $\mathcal{D}_s$.

Under the circumstance of few available labels in target domain, $P_t(x_i)$ $(x_i \in \mathcal{D}_s)$ is always inaccessible. Hence, it is tough to directly calculate the defined density ratio β_i. To handle this issue, the KMM method is adopted in the proposed MSAS-DMOA to obtain β_i, which aims at weighting $x_s \in \mathcal{D}_s$ via kernel tricks to match the data distribution between $\mathcal{D}_s$ and $\mathcal{D}_t$, so that using the information of $P_t(x)$ can be avoided. To be specific, both $\mathcal{D}_s$ and $\mathcal{D}_t$ are mapped into a high-dimensional reproducing kernel Hilbert space (RKHS), and according to the maximum mean discrepancy theory [9], the difference between mean values of the mapped $\mathcal{D}_s$ and $\mathcal{D}_t$ is equivalent to the distance of data distribution in the original $\mathcal{D}_s$ and $\mathcal{D}_t$. Consequently, in order to match the data distribution between $\mathcal{D}_s$ and $\mathcal{D}_t$ to promote reliable knowledge transfer, one can re-weight $x_s \in \mathcal{D}_s$ in the RKHS to make the mean value of those mapped x_s as close as that of the mapped $x_t \in \mathcal{D}_t$, which can be depicted as:

$$\beta^* = \arg\min \left\| \frac{1}{n_s}\sum_{x_s\in\mathcal{D}_s}\beta\boldsymbol{\kappa}(x) - \frac{1}{n_t}\sum_{x_t\in\mathcal{D}_t}\boldsymbol{\kappa}(x)\right\|_2 \tag{15}$$

where $n_s = |\mathcal{D}_s|$ and $n_t = |\mathcal{D}_t|$, $\boldsymbol{\kappa}(\cdot)$ is the Gaussian kernel function to generate RKHS.

Remark 3: Obtained β is an n_s-dimension vector, which reflects the distance of n_s samples in $\mathcal{D}_s$ to $\mathcal{D}_t$ in the reproducing kernel Hilbert space.

According to [9], solving above Eq. (15) is equivalent to addressing following quadratic programming problem:

$$\begin{cases} \min\ \mathscr{L}(\beta) = \dfrac{1}{2}\beta^T\mathbf{K}\beta - \boldsymbol{g}^T\beta \\ s.t.\ \left|\dfrac{1}{n_s}\displaystyle\sum_{i=1}^{n_s}\beta_i\ - 1\right| \leq \sqrt{n_s} \end{cases} \tag{16}$$

where $\mathbf{K} = \begin{pmatrix} K_{s,s} & K_{s,t} \\ K_{t,s} & K_{t,t}\end{pmatrix}$ with the element $K_{s,s} = [\boldsymbol{\kappa}(x_i) \cdot \boldsymbol{\kappa}(x_j)]$ $(x_i \in \mathcal{D}_s,\ x_j \in \mathcal{D}_s)$, and so on. $\boldsymbol{g}$ consists of element $g_i = \frac{n_s}{n_t}\sum_{j=1}^{n_t}\boldsymbol{\kappa}(x_i)\cdot\boldsymbol{\kappa}(x_j)$ where $x_i \in \mathcal{D}_s \cup \mathcal{D}_t$ and $x_j \in \mathcal{D}_t$.

As a result, the optimal $\beta = [\beta_i]_{i=1}^{n_s}$ can be obtained via

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI10.1109/TNNLS.2023.3295461, IEEE Transactions on Neural Networks and Learning Systems

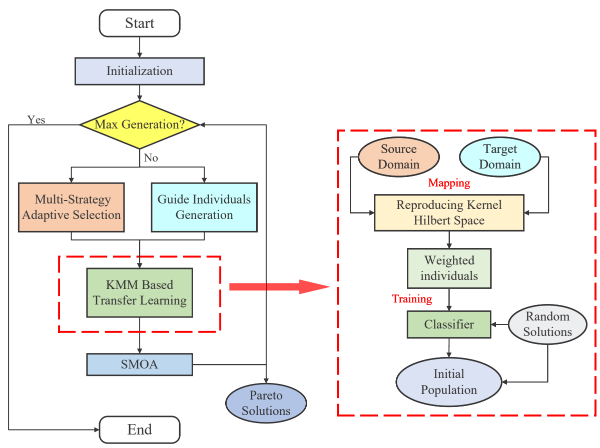


Fig. 3. Overall flowchart of MSAS-DMOA.

solving Eq. (16), where the element stands for the reliability of individuals in $\mathcal{D}_s$. Particularly, if β_i is close to 0, then $x_i \in \mathcal{D}_s$ is deemed useless in the new environment. On the contrary, large $\beta_.$ implies that the historical solutions are valuable [11]. Then, samples in $\mathcal{D}_s$ with β are adopted to train the population predictor according to Eq. (14), where the corresponding normalized β is used as the training weight. By doing so, a great deal of available solutions with labels obtained in previous searching can be fully used, and more importantly, the guide solutions generated in the new environment have not directly participated in the model training, which are only adopted to further estimate the reference value of historical experiences so that reliable knowledge transfer can be realized. At last, via the trained predictor, initial population with high quality in the new environment can be obtained, which is in the hope of accelerating the acquisition of PS_{t+1}.

E. Overall Framework of MSAS-DMOA

The overall framework of the proposed MSAS-DMOA is presented in Alg. 3, and the schematic diagram is illustrated in Fig. 3. Notice that population in the first environment is initialized through the applied SMOA, and it is from the second generation that the initial population is provided by the trained prediction model.

In particular, the support vector machine (SVM) is employed as the prediction model (i.e., the "classifier" in Fig. 3) in the proposed MSAS-DMOA, whose training samples are the previously mentioned $\{\beta_i, x_i, y_i\}_{i=1}^{n_s}$ where $(x_i, y_i) \in \mathcal{D}_s$, and it is noticeable that the labels in source domain $\mathcal{D}_s$ are available based on the dominance relationship. Moreover, the TrAdaboost technique [5] is applied for model training, and by inputting the randomly generated solutions to the trained SVM classifier, individuals with high-quality are outputted as the initial population for searching in the new environment.

IV. Experiments and Results

In this section, the proposed MSAS-DMOA is evaluated on a series of benchmark functions, and other four popular DMOAs are adopted for comparisons.

Algorithm 3 Framework of MSAS-DMOA

Input:
An SMOA optimizer $\mathcal{S}$ and objective functions F
Output:
Pareto set $PS = \{PS_t\}_{t=1}^{T}$ in all T environment

```
PS ← ∅
For t from 1 to T
   Update environmental parameters
   If t = 1
      Generate initial population Pop^t_ini by optimizer S
   else
      Establish source domain D_s based on Alg. 1
      Build target domain D_t according to Alg. 2
      Train a classifier via KMM-based transfer learning
      Output an initial population Pop^t_ini
   EndIf
   Obtain Pareto set PS_t = S(F, Pop^t_ini)
EndFor
Return PS
```

A. Benchmark Functions and Evaluation Metrics

For a comprehensive evaluation, the proposed MSAS-DMOA is tested on 14 benchmark functions DF1-DF14, and more details of these adopted DMOPs can be found in [13]. The inverted generational distance (IGD) [12] and the maximum spread (MS) [16] are employed as the evaluation metrics. In particular, IGD is a comprehensive indicator that can reflect both diversity and convergence, whereas in this study, it is mainly adopted to evaluate the convergence, which calculates the average distance between the obtained PF and corresponding ground-truth one as follows:

$$IGD_t = \frac{1}{|PF_t|} \sum_{\tilde{p} \in \tilde{PF}_t} \min_{x \in PF_t} d_{min}(x, \tilde{p}) \tag{17}$$

where $\tilde{PF}_t$ is the true Pareto front at time t, and d_{min} stands for the minimal Euclidean distance. Apparently, if the true PF has covered the obtained one (i.e., the algorithm directly obtains the ground-truth solutions), then the IGD value equals to zero, which indicates that smaller IGD corresponds to better convergence.

The other indicator MS is adopted to characterize the diversity of algorithm, which measures the overlapping areas between obtained PF and the true one as:

$$MS_t = \sqrt{\frac{1}{m} \sum_{k=1}^{m} \left(\frac{\min\{F_k^{max}, f_k^{max}\} - \max\{F_k^{min}, f_k^{min}\}}{F_k^{max} - F_k^{min}}\right)^2} \tag{18}$$

where m is the number of objective functions, superscripts min and max denote the minimal and maximal value. F and f refer to the true and obtained PF, respectively. Notice that the larger the MS, the better the diversity.

In addition, given that there is T environment in total, the average level of above two evaluation metrics is reported in the experiments. Accordingly, the mean values of IGD and

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI10.1109/TNNLS.2023.3295461, IEEE Transactions on Neural Networks and Learning Systems

MS are calculated as:

$$\begin{aligned} MIGD &= \frac{1}{T}\sum_{t=1}^{T} IGD_t \\ MMS &= \frac{1}{T}\sum_{t=1}^{T} MS_t \end{aligned} \tag{19}$$

B. Comparison Algorithms and Experimental Settings

To further verify the competitiveness of the proposed MSAS-DMOA, other four popular DMOAs are employed for comparisons, which are DNSGA-II-B [6], Tr-DMOEA [14], CR-DMOEA [39] and KT-DMOEA [15]. In both the latter three algorithms and the proposed method, MOEA/D [53] is selected as the SMOA optimizer.

For a DMOP, the dynamic behavior is depicted as $t = \frac{1}{n_t}\lfloor\frac{\tau}{\tau_t}\rfloor$, where τ is the maximum generation and t stands for current environment, n_t and τ_t denote severity and frequency of environmental changes, respectively. Accordingly, on each benchmark function, three groups of above dynamics are adopted, including $(n_t, \tau_t) \in \{(5, 10), (10, 5), (10, 10)\}$. Furthermore, the maximum generation is regulated as $\tau = 30 \times \tau_t$ to ensure that the environment will change 30 times.

In addition, dimension of variable space is set to 10, and population size is fixed at 100 and 150 to solve bi- and tri-objective problems, respectively. To alleviate influences of randomness, experiment on each benchmark function has run 10 times independently and the obtained results in average level are reported.

C. Benchmark Evaluation Results

At first, the benchmark evaluation results in terms of MIGD and MMS are reported in Table II and Table III, respectively, where the best results are shown in boldface. In addition, the Wilcoxon rank sum test [7] is performed at the significance level of 0.05, where "+/−" and "=" denote that the proposed MSAS-DMOA has performed significantly better/worse, and equivalently in comparison to corresponding algorithm. In addition, to provide an intuitive illustration of the algorithm performance, results on two functions (DF9 and DF10) are presented in the box-plot form in Fig. 4.

According to Table II, MSAS-DMOA has presented an overwhelming performance improvement on most benchmark problems in terms of the MIGD indicator. To be specific, in 30 out of 42 test cases, MSAS-DMOA has achieved the best results. Especially on the five problems of DF1, DF2, DF6, DF12 and DF14, the proposed method has ranked first in all cases with different dynamic parameters. It is noticeable that on DF8 and DF11, MSAS-DMOA presents a relatively worse convergence, which may due to the complex property of those problems, where the true PF expands or contracts with diverse scales along the centroid. Moreover, the density of solutions can also dramatically change, which brings difficulties in obtaining the uniformly distributed Pareto solutions. As a result, it is confirmed that while solving DMOPs, it is tough to guarantee that the transferred knowledge obtained from historical searching is always reliable, and the negative transfer can lead to poor convergence.

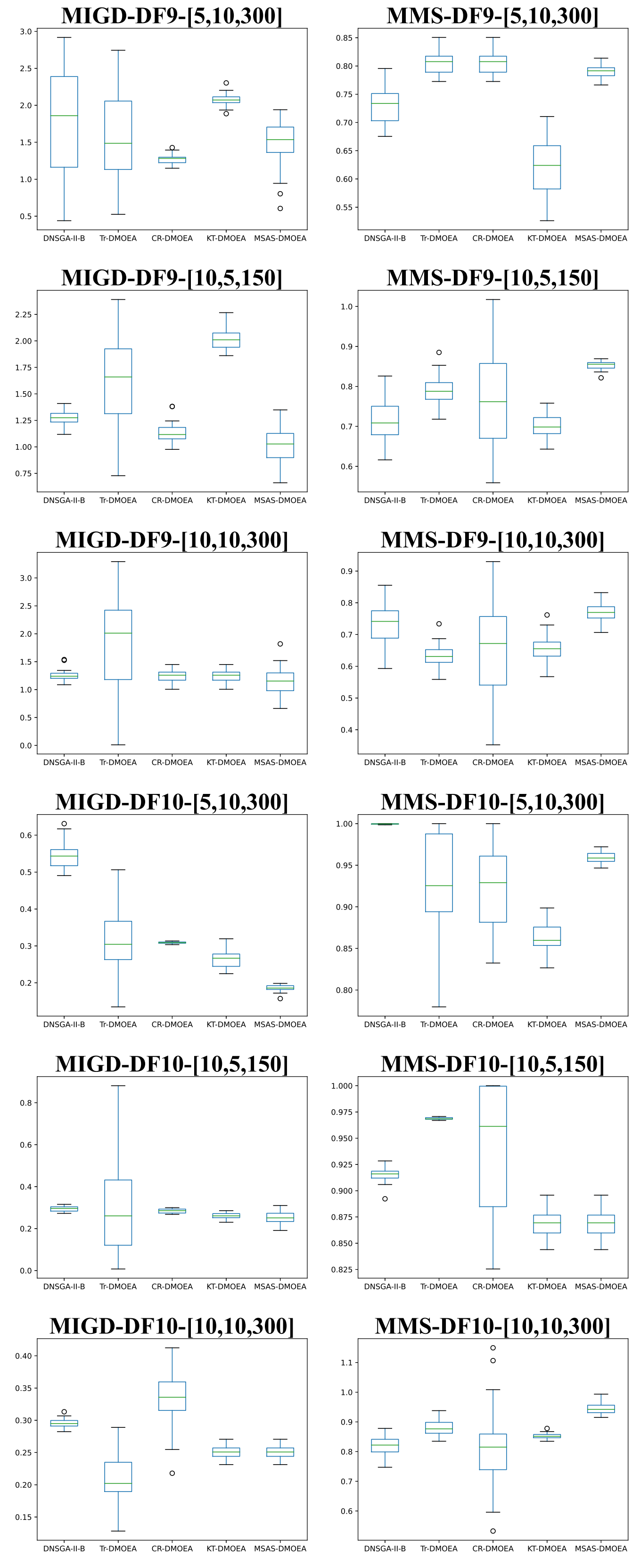


Fig. 4. Box-plots of problems DF9, DF10 on MIGD and MMS.

As reported in Table III, the proposed MSAS-DMOA achieves 27 best results out of 42 comparisons on MMS indi-

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI10.1109/TNNLS.2023.3295461, IEEE Transactions on Neural Networks and Learning Systems

TABLE II
PERFORMANCE COMPARISON OF FIVE ALGORITHMS ON MIGD

Problems	$n_t,\ \tau_t$	Algorithms				
		DNSGA-II-B [6]	Tr-DMOEA [14]	CR-DMOEA [39]	KT-DMOEA [15]	MSAS-DMOA
DF1	5, 10	0.1435±1.67e-01(+)	0.1397±5.17e-02(+)	0.1018±1.61e-02(=)	0.1448±1.35e-02(+)	**0.0983±1.93e-02**
	10, 5	0.2436±2.14e-01(+)	0.2349±3.71e-02(+)	0.2201±5.26e-02(+)	0.2054±5.00e-03(+)	**0.1838±2.47e-02**
	10, 10	0.2215±1.82e-01(+)	0.1487±2.39e-02(+)	0.1407±4.68e-02(+)	0.1324±1.17e-02(=)	**0.1321±1.97e-02**
DF2	5, 10	0.1408±1.77e-02(+)	0.1291±3.43e-02(+)	0.0902±6.30e-03(=)	0.1132±1.84e-02(+)	**0.0876±7.82e-03**
	10, 5	0.1316±1.05e-01(+)	0.2145±3.18e-02(+)	0.1702±2.12e-02(+)	0.1632±2.10e-02(+)	**0.1271±2.09e-02**
	10, 10	0.1684±8.82e-02(+)	0.0835±4.13e-02(=)	0.1045±5.22e-02(+)	0.1090±6.70e-03(+)	**0.0823±6.89e-03**
DF3	5, 10	**0.2803±1.12e-01(+)**	0.2988±9.32e-02(+)	0.3833±4.56e-02(+)	0.4427±3.27e-02(+)	0.3783±3.43e-02
	10, 5	0.5788±1.72e-01(+)	0.5482±1.12e-01(+)	0.2919±2.89e-02(+)	0.4581±2.43e-02(+)	**0.2754±1.30e-02**
	10, 10	0.6022±2.20e-01(+)	0.3516±2.35e-01(+)	0.2692±4.37e-02(+)	0.4101±3.23e-02(+)	**0.2400±3.41e-02**
DF4	5, 10	1.1704±2.40e-01(-)	**1.0988±4.77e-01(-)**	1.3300±2.22e-02(-)	1.1360±7.22e-02(-)	1.3570±4.42e-02
	10, 5	1.8188±3.05e-01(+)	1.2415±4.13e-01(-)	1.1643±1.39e-01(-)	**1.1557±1.04e-01(-)**	1.3492±9.63e-02
	10, 10	1.6794±2.28e-01(+)	1.2154±3.25e-01(+)	1.0269±1.85e-01(+)	1.1736±1.55e-01(+)	**0.9041±1.35e-01**
DF5	5, 10	0.0612±1.72e-02(+)	1.3541±3.66e-01(+)	0.0493±4.90e-03(+)	1.3278±2.68e-02(+)	**0.0356±5.24e-03**
	10, 5	0.1326±8.71e-02(+)	1.8921±6.15e-01(+)	0.1273±3.46e-02(+)	1.4023±4.37e-02(+)	**0.1113±2.40e-02**
	10, 10	**0.1000±8.86e-02(-)**	0.8623±3.51e-01(+)	0.1138±3.78e-02(-)	1.3283±1.56e-02(+)	0.1459±4.23e-02
DF6	5, 10	3.4736±2.61e+00(+)	2.2569±4.63e-01(+)	1.7698±5.45e-01(+)	2.6096±2.46e-01(+)	**1.6742±7.59e-01**
	10, 5	0.7657±3.96e-01(+)	3.4311±8.23e-01(+)	14.5568±7.81e+00(+)	4.5306±4.11e-01(+)	**0.5862±2.04e-01**
	10, 10	0.5435±2.15e-01(+)	1.5813±3.16e-01(+)	7.9292±5.35e+00(+)	3.5304±2.27e-01(+)	**0.3650±3.10e-01**
DF7	5, 10	6.8462±1.17e+00(+)	3.2145±1.12e+00(+)	2.2799±1.17e+00(+)	2.8198±1.91e-01(+)	**2.0305±4.39e-01**
	10, 5	8.3965±1.16e+00(+)	**3.1793±1.75e+00(-)**	13.1652±8.74e+00(+)	4.2143±4.70e-01(+)	3.4137±1.80e-01
	10, 10	7.3865±2.21e+00(+)	2.2264±9.21e-01(+)	6.3117±3.64e+00(+)	3.1263±3.95e-01(+)	**1.9993±8.64e-01**
DF8	5, 10	**0.8208±2.65e-02(-)**	1.3579±2.12e-01(+)	0.8561±7.48e-02(-)	1.0927±2.22e-02(+)	0.9288±4.24e-02
	10, 5	0.7987±1.82e-02(=)	1.2741±8.79e-02(+)	**0.7877±7.41e-02(-)**	1.0493±4.49e-02(+)	0.8575±4.61e-02
	10, 10	0.7902±3.91e-02(+)	1.7253±3.75e-01(+)	0.8088±1.24e-01(+)	1.0572±1.91e-02(+)	**0.6378±4.20e-02**
DF9	5, 10	1.9389±6.00e-01(+)	1.2482±9.48e-01(-)	**1.2479±8.23e-02(-)**	2.0908±8.70e-02(+)	1.4950±3.94e-01
	10, 5	1.2684±6.79e-02(+)	1.6371±3.75e-01(+)	1.1496±1.23e-01(+)	2.0497±9.14e-02(+)	**1.0047±2.04e-01**
	10, 10	1.2510±8.04e-02(+)	1.8649±8.23e-01(+)	1.2620±1.18e-01(+)	2.0631±5.41e-02(+)	**1.1378±2.49e-01**
DF10	5, 10	0.5413±3.27e-02(+)	0.2967±8.43e-02(+)	0.3090±2.50e-03(+)	0.2702±2.19e-02(+)	**0.1867±9.55e-03**
	10, 5	0.2945±1.40e-02(+)	0.2874±1.88e-01(+)	0.2809±1.31e-02(+)	0.2610±1.50e-02(+)	**0.2543±3.00e-02**
	10, 10	0.2933±6.70e-03(+)	**0.2013±3.12e-02(-)**	0.3170±4.02e-02(+)	0.2523±1.01e-02(+)	0.2369±2.53e-02
DF11	5, 10	0.3694±1.84e-02(+)	0.3167±3.29e-02(+)	0.3779±1.27e-02(+)	0.1635±9.10e-03(+)	**0.1587±1.13e-02**
	10, 5	0.5211±3.58e-02(+)	0.3933±4.19e-02(+)	0.3984±4.37e-02(+)	**0.1941±1.62e-02(-)**	0.2241±2.53e-02
	10, 10	0.5309±4.28e-02(+)	0.2819±1.84e-01(+)	0.4658±1.79e-02(+)	**0.1648±1.05e-02(-)**	0.1726±1.33e-02
DF12	5, 10	0.7407±2.96e-02(+)	2.1365±3.79e-01(+)	0.4311±1.64e-02(+)	0.5781±2.50e-02(+)	**0.4073±2.49e-02**
	10, 5	0.6655±8.84e-02(+)	2.9341±7.91e-01(+)	0.4688±3.37e-02(+)	0.5717±1.89e-02(+)	**0.4246±3.05e-02**
	10, 10	0.7072±5.44e-02(+)	1.5938±3.76e-01(+)	0.4353±7.00e-03(+)	0.5703±2.76e-02(+)	**0.4132±3.59e-03**
DF13	5, 10	**0.2533±1.41e-02(-)**	2.1465±8.19e-01(+)	0.2699±2.82e-02(-)	1.3789±1.44e-02(+)	0.3058±1.30e-02
	10, 5	0.4580±1.13e-01(+)	2.8955±1.91e+00(+)	0.4016±5.51e-02(=)	1.4670±3.57e-02(+)	**0.3917±1.11e-02**
	10, 10	0.4158±7.25e-02(+)	2.9658±1.46e+00(+)	0.3321±1.02e-02(+)	1.3892±4.90e-03(+)	**0.2411±2.30e-02**
DF14	5, 10	0.1384±9.20e-03(+)	0.8930±3.21e-01(+)	0.1157±3.60e-03(+)	0.8678±6.60e-03(+)	**0.0946±1.30e-03**
	10, 5	0.4099±6.00e-02(+)	1.2749±6.49e-01(+)	0.1960±2.01e-02(+)	0.9010±9.30e-03(+)	**0.1840±2.59e-02**
	10, 10	0.4387±4.77e-02(+)	0.8524±1.08e-01(+)	0.1801±2.15e-02(+)	0.8723±4.00e-03(+)	**0.1603±3.95e-02**
+ / - / =	\	37 / 4 / 1	36 / 5 / 1	32 / 7 / 3	37 / 4 / 1	\

cator, according to which one can conclude that the diversity performance of MSAS-DMOA has far more exceeded other 4 algorithms. In addition, it is worth mentioning that the TL paradigm has been introduced in both Tr-DMOEA and KT-DMOEA to solve DMOPs, whereas above two methods only obtain 7 and 6 best results on MMS, respectively, which implies that the proposed MSAS-DMOA is a competitive DMOP solver that can realize effective knowledge transfer. Moreover, on 30, 30, 32 and 31 cases, MSSA-DMOA presents significantly better results than DNSGA-II-B, Tr-DMOEA, CR-DMOEA and KT-DMOEA, respectively, which sufficiently exhibits the superiority of the proposed algorithm.

Above analysis has shown that MSAS-DMOA can yield outstanding results on convergence and diversity simultaneously, which can also be reflected in Fig. 4. It is mainly because that the reference values of historical solutions are well considered, which promotes transferring the useful knowledge in complicated varying environment.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI10.1109/TNNLS.2023.3295461, IEEE Transactions on Neural Networks and Learning Systems

TABLE III
PERFORMANCE COMPARISON OF FIVE ALGORITHMS ON MMS

Problems	$n_t,\ \tau_t$	Algorithms				
		DNSGA-II-B [6]	Tr-DMOEA [14]	CR-DMOEA [39]	KT-DMOEA [15]	MSAS-DMOA
DF1	5, 10	0.7887±2.66e-01(+)	0.8291±3.65e-02(+)	0.7194±1.02e-01(+)	0.6979±2.45e-02(+)	**0.8942±1.94e-02**
	10, 5	**0.9230±1.13e-01(-)**	0.8743±5.83e-02(-)	0.8325±1.28e-01(-)	0.7761±5.80e-03(+)	0.8013±2.81e-02
	10, 10	0.8716±3.14e-02(-)	0.8472±6.38e-02(+)	**0.9010±1.14e-01(-)**	0.7984±2.90e-02(+)	0.8565±1.67e-02
DF2	5, 10	0.8716±3.14e-02(-)	0.8472±6.38e-02(+)	**0.9010±1.14e-01(-)**	0.7984±2.90e-02(+)	0.8565±1.67e-02
	10, 5	0.7517±1.24e-01(+)	0.8294±8.88e-02(+)	0.7123±7.60e-02(+)	0.7225±2.12e-02(+)	**0.8439±7.59e-02**
	10, 10	0.8157±1.40e-01(+)	0.8702±5.38e-02(+)	0.8310±8.16e-02(+)	0.7427±1.95e-02(+)	**0.9003±6.64e-02**
DF3	5, 10	**0.4835±1.64e-01(-)**	0.4692±1.09e-02(-)	0.3302±6.41e-02(+)	0.4292±2.21e-02(=)	0.4245±1.60e-02
	10, 5	0.2266±6.08e-02(+)	0.4284±1.34e-02(-)	**0.4694±7.78e-02(-)**	0.4236±3.12e-02(-)	0.4160±1.75e-02
	10, 10	0.2372±1.06e-01(+)	**0.6938±3.41e-02(-)**	0.4984±9.17e-02(-)	0.4200±2.75e-02(=)	0.4230±9.58e-03
DF4	5, 10	0.3671±3.52e-02(-)	**0.4867±1.09e-02(+)**	0.3941±3.74e-02(-)	0.2767±2.22e-02(+)	0.3379±2.18e-02
	10, 5	0.2866±6.57e-02(+)	0.2492±3.55e-02(+)	0.4694±2.12e-02(+)	0.2746±2.02e-02(+)	**0.5932±3.49e-02**
	10, 10	0.2645±3.64e-02(+)	0.3357±3.12e-02(+)	0.4629±2.92e-02(+)	0.3163±2.72e-02(+)	**0.5505±3.00e-02**
DF5	5, 10	**0.9923±3.00e-04(-)**	0.9780±1.46e-02(-)	0.9698±1.93e-02(-)	0.8829±2.99e-02(+)	0.9376±2.63e-02
	10, 5	0.9597±4.00e-04(+)	0.9532±1.11e-02(+)	0.9757±2.73e-02(+)	0.8979±5.87e-02(+)	**0.9855±5.27e-02**
	10, 10	0.9498±2.00e-04(+)	0.9589±2.91e-02(+)	**0.9874±2.53e-02(-)**	0.8569±2.36e-02(+)	0.9755±1.89e-02
DF6	5, 10	**1.0000±0(-)**	0.8362±4.16e-02(+)	0.9408±1.50e-01(+)	**1.0000±0(-)**	0.9874±1.17e-01
	10, 5	**1.0000±0(=)**	0.6297±7.99e-02(+)	**1.0000±0(=)**	**1.0000±0(=)**	**1.0000±0**
	10, 10	**1.0000±0(=)**	0.6787±6.20e-02(+)	**1.0000±0(=)**	**1.0000±0(=)**	**1.0000±0**
DF7	5, 10	0.8192±2.36e-02(+)	**1.0000±0(=)**	0.9484±3.40e-03(+)	**1.0000±0(=)**	**1.0000±0**
	10, 5	0.7976±4.21e-02(+)	0.6620±1.04e-01(+)	0.9981±2.70e-03(+)	**1.0000±0(=)**	**1.0000±0**
	10, 10	0.6072±6.96e-02(+)	0.6268±6.04e-02(+)	0.9867±1.04e-02(+)	**1.0000±0(=)**	**1.0000±0**
DF8	5, 10	0.4146±2.98e-02(+)	0.4975±8.56e-02(+)	0.4574±1.31e-02(+)	0.2317±1.45e-02(+)	**0.5732±6.40e-02**
	10, 5	0.4279±1.93e-02(+)	0.6591±4.20e-02(+)	0.6340±9.82e-02(+)	0.2839±2.67e-02(+)	**0.7492±3.88e-02**
	10, 10	0.4394±3.35e-02(+)	0.6284±3.86e-02(+)	0.5652±1.50e-01(+)	0.2426±2.11e-02(+)	**0.8324±2.03e-02**
DF9	5, 10	0.7297±2.93e-02(+)	**0.8035±1.90e-02(+)**	0.7776±7.90e-02(+)	0.6311±5.30e-02(+)	0.7889±1.28e-02
	10, 5	0.7229±5.34e-02(+)	0.7864±3.69e-02(+)	0.7684±1.21e-01(+)	0.7025±3.25e-02(+)	**0.8541±1.14e-02**
	10, 10	0.7252±6.62e-02(+)	0.6347±5.55e-02(+)	0.6716±1.14e-01(+)	0.6438±3.40e-02(+)	**0.7706±2.22e-02**
DF10	5, 10	**0.9997±6.00e-04(-)**	0.9192±6.90e-02(+)	0.9306±4.75e-02(+)	0.8604±1.96e-02(+)	0.9599±8.03e-03
	10, 5	0.9158±8.90e-03(+)	**0.9683±1.03e-03(=)**	0.9618±7.69e-02(=)	0.8683±1.36e-02(+)	0.9655±1.36e-02
	10, 10	0.8175±2.96e-02(+)	0.8795±1.97e-02(+)	0.8325±1.50e-01(+)	0.8535±1.01e-02(+)	**0.9441±1.64e-02**
DF11	5, 10	0.9514±2.50e-02(+)	0.9734±1.94e-03(=)	0.8898±2.12e-02(+)	0.9629±5.30e-03(+)	**0.9784±1.33e-02**
	10, 5	0.6956±2.78e-02(+)	**0.9725±1.97e-03(-)**	0.9091±3.48e-02(+)	0.9307±1.19e-02(-)	0.9161±1.45e-02
	10, 10	0.6886±2.35e-02(+)	**0.9932±1.35e-03(-)**	0.7853±2.09e-02(+)	0.9550±1.03e-02(-)	0.9431±1.73e-03
DF12	5, 10	0.6428±4.51e-02(+)	0.0152±1.19e-03(+)	0.6695±7.89e-02(+)	0.6990±9.00e-03(+)	**0.8158±2.15e-02**
	10, 5	0.6761±1.27e-02(+)	0.0975±1.11e-03(+)	0.7272±6.67e-02(+)	0.6950±1.50e-02(+)	**0.8443±2.15e-02**
	10, 10	0.6808±2.11e-02(+)	0.0034±3.46e-04(+)	0.6007±5.08e-02(+)	0.6841±2.56e-02(+)	**0.7842±1.62e-02**
DF13	5, 10	0.9090±1.10e-03(+)	0.9295±3.07e-02(=)	0.9038±4.92e-02(+)	0.8358±1.94e-02(+)	**0.9368±1.93e-02**
	10, 5	0.8993±3.00e-04(+)	0.9001±2.12e-02(+)	0.9535±3.36e-02(-)	0.8342±7.30e-03(+)	**0.9728±3.83e-02**
	10, 10	0.8996±2.00e-04(+)	0.9284±1.19e-02(+)	0.9201±4.32e-02(+)	0.8395±9.90e-03(+)	**0.9536±1.17e-02**
DF14	5, 10	**1.0000±0(=)**	0.9737±1.85e-03(+)	0.8085±1.08e-01(+)	0.9371±8.10e-03(+)	**1.0000±0**
	10, 5	**1.0000±0(=)**	0.9208±1.00e-03(+)	0.9183±9.80e-03(+)	0.9365±2.80e-03(+)	**1.0000±0**
	10, 10	**1.0000±0(=)**	0.9497±1.35e-03(+)	0.9062±1.09e-02(+)	0.9203±6.50e-03(+)	**1.0000±0**
+ / - / =	\	30 / 9 / 3	30 / 9 / 3	32 / 8 / 2	31 / 6 / 5	\

D. Comparisons with Other Advanced DMOAs

In this subsection, on five bi-objective benchmark functions DF1-DF5, the proposed MSAS-DMOA is compared with other three advanced DMOAs. In particular, a novel multi-directional prediction approach and an extended autoencoding evolutionary search strategy have been proposed in [38] and [8], respectively, and their implementations on the multi-objective particle swarm optimizer are adopted for comparisons, which are denoted as MDP-MOPSO and AE-MOPSO. In [48], the knowledge guided Bayesian classification has been applied to solve DMOPs, and this recently proposed algorithm is used for comparison as well, which is denoted as KGB-DMOA. According to the experimental settings described in [8], the dynamic parameters are fixed at $n_t = 1$ and $\tau_t = 10$,

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI10.1109/TNNLS.2023.3295461, IEEE Transactions on Neural Networks and Learning Systems

TABLE IV
COMPARISONS WITH OTHER ADVANCED DMOAS ON BI-OBJECTIVE FUNCTIONS DF1-DF5 IN TERMS OF MIGD

Problems	MSAS-DMOA	AE-MOPSO	MDP-MOPSO	KGB-DMOA
DF1	**2.8208e-01±3.15e-02**	1.8842e+01±6.34e-01	2.0982e+01±7.11e-01	4.5977e-01±4.16e-01
DF2	**2.0327e-01±3.21e-02**	2.3671e-01±1.63e-01	2.1988e-01±6.72e-02	5.2668e-01±3.62e-01
DF3	**4.0147e-01±1.71e-02**	4.7168e-01±4.80e-02	5.5480e-01±9.80e-02	9.8680e-01±5.81e-01
DF4	**9.5072e-01±6.20e-02**	1.1727e+00±1.07e-01	1.3074e+00±1.34e-01	2.4960e+00±2.71e+00
DF5	1.7729e+00±4.69e-02	1.7240e-01±2.77e-02	**1.6015e-01±3.28e-02**	5.0357e-01±2.79e-01

TABLE V
COMPARISONS WITH OTHER ADVANCED DMOAS ON TRI-OBJECTIVE FUNCTIONS DF10-DF14 IN TERMS OF MIGD

Problems	$n_t,\ \tau_t$	MSAS-DMOA	SVR-DMOA [4]	MMTL-DMOA [17]	MSTL-DMOA [49]
DF10	(5,10)	**0.1914±1.07e-02**	0.2265±1.90e-02	0.2200±2.02e-02	0.2336±2.90e-02
	(10,10)	**0.1813±1.44e-02**	0.2460±2.31e-02	0.2197±3.00e-02	0.2197±1.46e-02
DF11	(5,10)	0.1262±8.10e-03	**0.1156±3.35e-03**	0.1182±5.14e-03	0.1171±4.16e-03
	(10,10)	0.1290±1.19e-02	0.1167±3.48e-03	**0.1151±3.60e-03**	0.1168±3.42e-03
DF12	(5,10)	0.4741±1.47e-02	0.1883±1.54e-02	0.3236±6.85e-02	**0.1729±3.17e-02**
	(10,10)	0.4166±1.32e-02	0.1765±8.24e-03	0.2556±2.37e-02	**0.1384±8.06e-03**
DF13	(5,10)	**0.2578±7.70e-03**	0.2586±1.59e-02	0.2599±1.01e-02	0.2616±1.15e-02
	(10,10)	**0.2456±9.20e-03**	0.2472±1.02e-02	0.2644±1.34e-02	0.2604±1.51e-02
DF14	(5,10)	**0.0782±3.40e-03**	0.0805±2.10e-03	0.0932±2.94e-02	0.0853±3.80e-03
	(10,10)	**0.0785±3.20e-03**	0.0794±2.36e-03	0.0817±2.81e-03	0.0846±5.06e-03

the number of environmental changes is 20, and the population size is set to 100. The comparison results in terms of the MIGD are displayed in Table IV, where data of the MDP-MOPSO and AE-MOPSO are cited from [8].

As can be seen from Table IV, even in case of large change severity ($n_t = 1$), the proposed MSAS-DMOA outperforms the other three advanced algorithms on four problems, which shows the effectiveness and competitiveness of our method in handling the dynamic behaviors. Notice that on the DF1 problem, the geometric shape of Pareto front changes back and forth from concave to convex, while the result obtained by the proposed MSAS-DMOA reaches 10^{-1} level, which owes to the rich information contained in the source domain, and it can be concluded that the diverse training samples do promote the transfer of useful knowledge. As compared with the KGB-DMOA that also obtains result in 10^{-1} level, the proposed MSAS-DMOA presents more stable performance with smaller standard variance, which demonstrates the reliability of our method as a competent DMOP solver.

In addition, to further validate the superiority of the applied non-inductive TL-based response method, on five tri-objective benchmark functions, the proposed MSAS-DMOA is compared with some other advanced DMOAs, denoted as SVR-DMOA [4], MMTL-DMOA [17] and MSTL-DMOA [49], where in [4], the support vector regression predictor has been adopted to assist evolution; the manifold and multi-source TL paradigms have been applied in [17] and [49], respectively. The comparison results are presented in Table V, where data of the three comparison algorithms are cited from [49], and corresponding experimental settings are maintained the same for a fair comparison.

As is reported in Table V, on 3 out of 5 problems, the proposed MSAS-DMOA obtains both the best results with two dynamic parameter settings, and the other three advanced DMOAs rank the first on 1, 1, and 2 cases, respectively, which demonstrates that the proposed MSAS-DMOA is competent in handling various dynamic behaviors in complex situations. It is noticeable that in the MSTL-DMOA, multiple source domains are selected as candidates where the most similar one to the target domain is finally used for training, which has been proven effective to realize the knowledge transfer. Hence, it also motivates us to store the previously constructed $\mathcal{D}_s$, due to the carefully selected samples in the historical environments may also be valuable in a new situation, which benefits further enriching the potential helpful information in source domain.

It should be pointed out that although the proposed MSAS-DMOA has obtained satisfactory results, there are still spaces for further improvements. Firstly, the dynamic behaviors exist commonly in many situations, thus an in-depth comprehension and analysis of the dynamics can no doubt facilitate designing methods to solve other dynamic optimization problems [23], [37], [43], [54]. Secondly, the quantification of environmental changes in this study is based on the fitness re-evaluation, and it is promising to employ other alternatives such as the surrogate-based methods [41], [42]. Thirdly, in a dramatically changing environment, how to escape from the local optima deserves further attention. Moreover, some DMOPs may even present the multi-modal property, where several $PS_{t,i}$ $(i = 1, 2, \cdots)$ will be mapped to the same PF_t, that is, different Pareto solutions share the same phenotype in objective space, which makes it quite challenging and tough to obtain all of the global optimal solutions. To handle this issue, some advanced heuristic algorithms that pay attention to the balance between global exploration and local exploitation can be employed [19], [31], [50]. Lastly, some applied strategies in the proposed MSAS-DMOA can be integrated into play-and-plug modules,

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI10.1109/TNNLS.2023.3295461, IEEE Transactions on Neural Networks and Learning Systems

which can improve the generalization ability of the proposed methodology to assist other research such as the modeling of complex dynamic systems [10], [20].

E. Analysis of Negative Transfer

While solving DMOPs via TL methods, it is an important issue to effectively alleviate the negative transfer phenomenon, which significantly influences the quality of the predicted initial population in the new environment and may lead to poor convergence. In this subsection, to validate whether the proposed MSAS-DMOA is competent for handling above issue, experiments are performed in terms of $\mathcal{D}_s$ construction and $\mathcal{D}_t$ determination, respectively, of which the obtained results are displayed in Tables VI-VII and Fig. 5.

1) Influences of Source Domain: In the proposed MSAS-DMOA, the source domain is composed of solutions from various optional groups, which is expected to provide rich historical information. Accordingly, two variants are designed for comparisons to validate the effectiveness of the adaptive $\mathcal{D}_s$ construction manner. To be specific, by leaving out the scoring system in Eq. (8), MS-DMOA directly carries out the multiple strategies without in-depth analysis, which picks out individuals from all optional groups with the same proportion to form $\mathcal{D}_s$; another variant SS-DMOA adopts single strategy, which only employs the G_{ps} as $\mathcal{D}_s$. Two groups of environmental dynamics are selected, including $(n_t, \tau_t) = (10, 5)$ and $(n_t, \tau_t) = (10, 10)$, whereas other experimental settings remain unchanged. Comparison results on eight problems (DF7-DF14) of above three algorithms are reported in Table VI.

TABLE VI
PERFORMANCE COMPARISON AMONG ALGORITHMS WITH DIFFERENT SOURCE DOMAINS ON MIGD

Problems	$n_t,\ \tau_t$	MS-DMOA	SS-DMOA	MSAS-DMOA
DF7	10, 5	3.5837±3.23e-01(+)	4.1346±7.49e-01(+)	**3.4137±1.80e-01**
	10, 10	**1.8964±6.39e-01(-)**	2.3954±8.95e-01(+)	1.9993±8.64e-01
DF8	10, 5	1.4580±2.64e-02(+)	0.9875±3.98e-02(+)	**0.8575±4.61e-02**
	10, 10	1.0865±1.59e-02(+)	**0.5836±1.56e-02(-)**	0.6378±4.20e-02
DF9	10, 5	3.2804±2.55e-01(+)	3.2214±1.14e-01(+)	**1.0047±2.04e-01**
	10, 10	2.6093±4.53e-01(+)	3.5229±1.57e-01(+)	**1.1378±2.49e-01**
DF10	10, 5	0.5996±5.00e-02(+)	0.7354±1.04e-02(+)	**0.2543±3.00e-02**
	10, 10	0.4834±9.41e-02(+)	0.3988±2.69e-02(+)	**0.2369±2.53e-02**
DF11	10, 5	**0.2074±2.17e-02(-)**	0.2906±3.25e-02(+)	0.2241±2.53e-02
	10, 10	0.2018±2.89e-02(+)	0.2189±1.98e-02(+)	**0.1726±1.33e-02**
DF12	1, 5	0.4781±5.44e-02(+)	0.4485±6.11e-02(+)	**0.4246±3.05e-02**
	1, 10	0.5725±4.44e-02(+)	0.5854±4.35e-02(+)	**0.4132±3.59e-03**
DF13	10, 5	0.4869±2.71e-02(+)	0.5083±2.75e-02(+)	**0.3917±1.11e-02**
	10, 10	**0.1642±8.36e-02(-)**	0.3111±1.83e-02(+)	0.2411±2.30e-02
DF14	10, 5	0.2404±1.89e-02(+)	0.2451±2.34e-02(+)	**0.1840±2.59e-02**
	10, 10	0.2305±1.03e-02(+)	0.2349±1.31e-02(+)	**0.1603±3.95e-02**
+ / - / =	\	13 / 3 / 0	15 / 1 / 0	\

As can be seen, the original MSAS-DMOA performs significantly better than the variants MS-DMOA and SS-DMOA in 11 and 12 cases, respectively. Without an adaptive selection process to form $\mathcal{D}_s$, it is prone to assign large weight to individuals in a certain source where the data distribution is similar to that of $\mathcal{D}_t$. Hence, an over-fitting phenomenon may occur in model training due to those solutions deemed with high reference values can have a relatively dense distribution, which makes the predicted population easy to be trapped into a local optimum.

In addition, compared with SS-DMOA, solutions from multiple optional sources can provide complementary knowledge in $\mathcal{D}_s$, which enriches the training samples to train a robust predictor. Therefore, based on above discussions, one can conclude that the proposed multi-strategy adaptive selection mechanism for $\mathcal{D}_s$ construction can effectively alleviate the negative transfer phenomenon.

2) Influences of Target Domain: In the proposed MSAS-DMOA, $\mathcal{D}_t$ contains a group of guide solutions determined based on the Q-value (see Section III-C). To verify the effectiveness of those guide solutions, another variant MSAS-DMOA* is designed for comparison, where the fine-tuned Pareto set in the previous environment is adopted as $\mathcal{D}_t$. It should be pointed out that for a fair comparison, samples in $\mathcal{D}_t$ have not directly participated in the model training in both algorithms, which are only adopted to estimate the reference values of historical solutions via KMM method. Above comparison results in terms of MIGD are presented in Table VII.

TABLE VII
PERFORMANCE COMPARISON OF USING DIFFERENT TARGET DOMAINS ON MIGD

Problems	$n_t,\ \tau_t$	MSAS-DMOA*	MSAS-DMOA
DF7	10, 5	**3.2947±9.56e-01(-)**	3.4137±1.80e-01
	10, 10	**1.5624±2.72e-01(-)**	1.9993±8.64e-01
DF8	10, 5	1.1554±4.06e-02(+)	**0.8575±4.61e-02**
	10, 10	1.0418±1.43e-02(+)	**0.6378±4.20e-02**
DF9	10, 5	1.3842±3.43e-01(+)	**1.0047±2.04e-01**
	10, 10	1.2864±3.74e-01(+)	**1.1378±2.49e-01**
DF10	10, 5	0.7815±2.13e-02(+)	**0.2543±3.00e-02**
	10, 10	0.6549±1.28e-02(+)	**0.2369±2.53e-02**
DF11	10, 5	**0.1957±5.71e-02(-)**	0.2241±2.53e-02
	10, 10	**0.1700±4.62e-02(=)**	0.1726±1.33e-02
DF12	10, 5	0.8639±1.58e-02(+)	**0.4246±3.05e-02**
	10, 10	0.8574±3.41e-02(+)	**0.4132±3.59e-03**
DF13	10, 5	0.4826±2.46e-02(+)	**0.3917±1.11e-02**
	10, 10	0.2896±2.19e-02(+)	**0.2411±2.30e-02**
DF14	10, 5	0.2451±1.30e-02(+)	**0.1840±2.59e-02**
	10, 10	0.4963±3.04e-02(+))	**0.1603±3.95e-02**
+ / - / =	\	12 / 3 / 1	\

It is obvious from Table VII that in 12 out of 16 testing cases, the proposed MSAS-DMOA has outperformed its variant MSAS-DMOA*, which demonstrates that the screened guide solutions do play an important role in guiding the knowledge transfer direction. More importantly, it also indicates that re-weighting samples in source domain to match data distribution can lead to comparable results to the conventional methods that label samples in $\mathcal{D}_t$. As a result, the proposed method brings a feasible idea to replace the labeling procedure, which is of vital significance especially in cases where the labeled samples in target domain are quite insufficient.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI10.1109/TNNLS.2023.3295461, IEEE Transactions on Neural Networks and Learning Systems

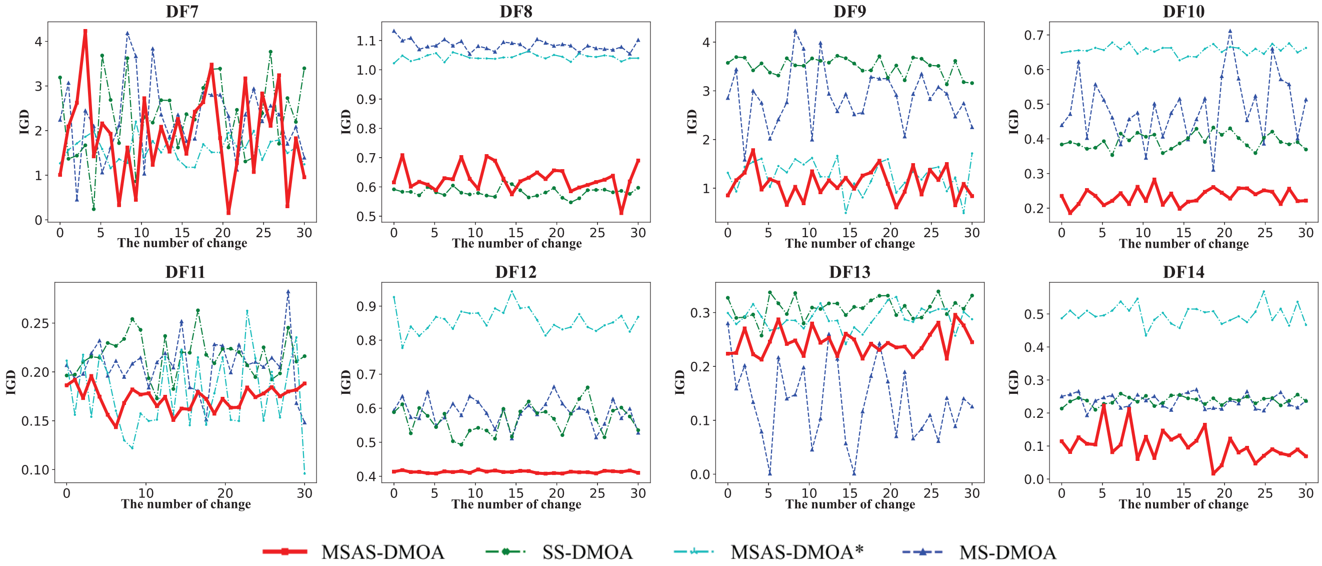


Fig. 5. Ablation study results of four MSAS-DMOA variants with $(n_t, \tau_t) = (10, 10)$ in terms of IGD values.

F. Case Study of Vehicle Speed Control Task

In this subsection, a case study of vehicle speed control task [18] is carried out for performance validation of the proposed method, which is realized by a proportional integral derivative (PID) controller with time-varying system function. Therefore, the proposed MSAS-DMOA is responsible for timely adjusting the key parameters in the PID system (including the K_p, K_i, and K_d) so as to meet the requirement of real-time control, and the schematic diagram of this case study is illustrated in Fig. 6.

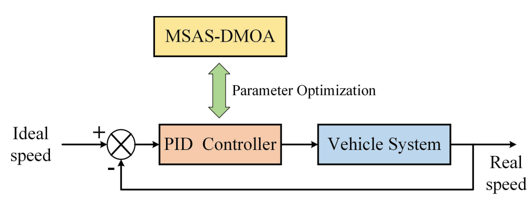


Fig. 6. Determination of the key parameters in a PID system for vehicle speed control.

With the control parameters provided by MSAS-DMOA, the output of PID system in response to a unit step input is adopted as the performance evaluation metric, which contains following two defined objectives:

$$\begin{cases} \min\ J_1 = 0.1 \times \displaystyle\int_{t_1}^{t_2} |e(t)|dt + 10 \times O_m \\ \min\ J_2 = 5 \times T_r \end{cases} \tag{20}$$

where $|e(t)|$ is the absolute error between the real output (in steady state) and the ideal value, O_m stands for the maximum overshoot, and T_r is the rise-time for the response curve to first reach 90% of the defined ideal value. It is noticeable that the objective J_1 reflects whether the PID system can realize accurate speed control, and the J_2 focuses on the speed of reaching a steady state, which requires the proposed MSAS-DMOA to properly adjust the key parameters so as to yield both accurate and fast performance. In addition, as previously mentioned, the system function of the PID controller is time-varying, which makes the speed control task more tough, and the expression of system function is given as:

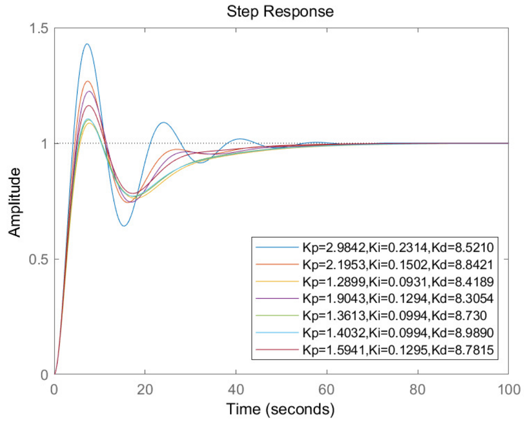
$$G(s,t) = \frac{1.5}{50s^3 + \alpha_2(t)s^2 + \alpha_1(t)s + 1} \tag{21}$$

where $\alpha_1(t) = 3 + 30sin(\frac{\pi t}{5})$ and $\alpha_2(t) = 43 + 30sin(\frac{\pi t}{5})$.

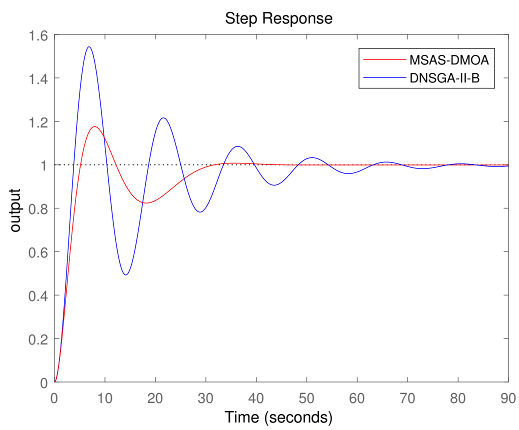
Remark 4: In this case study, the ideal speed is denoted as 1, thus other output values are correspondingly transformed in proportion and presented.

Firstly, at time $t = 0$, the unit step response curves with different parameters (i.e., the Pareto solutions) obtained by the proposed algorithm are displayed in Fig. 7(a), where the interest conflicts between the two objectives can be reflected. When the proportion coefficient K_p is enlarged, the maximum overshoot will significantly increased, meanwhile the rise-time is decreased. Similarly, the collaborative effects of K_i and K_d influence the convergence of the step response curves. According to Fig. 7(a), it is proven that the proposed method is capable of providing several feasible solutions in the vehicle speed control task. A noticeable issue is that in the real-world problems, it is extremely tough to obtain sufficient Pareto solutions due to the complex coupling relationships between objective functions. Hence, above results show the reliability of our MSAS-DMOA in real-world applications.

Then, the comparison between the proposed MSAS-DMOA and classic DNSGA-II-B is presented in Fig. 7(b) and Table VIII, respectively. As is shown, although using the DNSGA-II-B for PID parameters determination realizes faster rise-time, the maximum overshoot is as three times much as that obtained by applying the proposed MSAS-DMOA. Moreover, the reported indicator T_s in Table VIII is the settling



(a) Different PID parameters obtained by MSAS-DMOA in the first environment.



(b) Comparison between MSAS-DMOA and DNSGA-II-B.

Fig. 7. Illustrations of the unit step response curves.

TABLE VIII
COMPARISONS BETWEEN THE PROPOSED MSAS-DMOA AND CLASSIC DNSGA-II-B ALGORITHM

Methods	T_r	O_m	T_s
DNSGA-II-B	**3.45**	54.5%	60.79
MSAS-DMOA	4.48	**17.6%**	**28.30**

time, which refers to the required time for the output to be steady within a pre-defined tolerance band, and according to the results, the proposed MSAS-DMOA takes only half time to reach the steady state in comparison to the DNSGA-II-B, which exhibits the advantages of our algorithm in optimizing parameters of the PID controller.

Finally, according to the presented scatter plots of the obtained PF in Fig. 8, in each environment (markers with the same color), most of the Pareto solutions obtained by DNSGA-II-B can be dominated by those obtained by our method, which indicates that the proposed MSAS-DMOA can well adapt to the changing environments and obtain solutions with higher quality. Hence, it can be concluded that as compared with the re-initialization strategy used in the classic DNSGA-II-B, the

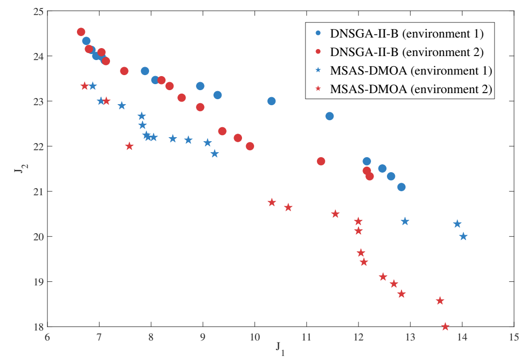


Fig. 8. Pareto fronts obtained by the proposed MSAS-DMOA and the classic DNSGA-II-B (denoted by the stars and circles, respectively) in two environments (shown in different colors).

proposed non-inductive TL-based response method can handle the dynamic behaviors better, thereby exhibiting the competitiveness and superiority in terms of accurately searching for the time-varying Pareto solutions. In our future work, we are prone to apply the proposed MSAS-DMOA to solve more real-world problems so as to validate its engineering practicality in optimizing the complex dynamic systems.

V. CONCLUSION

In this paper, a novel multi-strategy adaptive selection-based DMOA has been proposed, and as an ETO method, the non-inductive transfer learning paradigm has been adopted. Rich knowledge from different optional groups is contained in the formed source domain, and the importance of historical experience has been estimated via the KMM method, which replaces the conventional labeling procedure of target domain, and samples in $\mathcal{D}_t$ are only used to guide the direction of evolution to alleviate negative transfer.

Benchmark evaluations have been carried out on 14 DMOPs, and the results have demonstrated the superiority of the proposed MSAS-DMOA, which outperforms other 4 popular DMOAs in terms of both convergence and diversity in a statistic sense. According to the results of ablation study, the effectiveness of the applied strategies in our method has also been validated. In future, how to make the proposed algorithm adapt to extremely changing situations deserves further investigation, and it is also promising to apply the MSAS-DMOA in more real-world optimization scenes.

REFERENCES

[1] R. Azzouz, S. Bechikh and L. Said, A dynamic multi-objective evolutionary algorithm using a change severity-based adaptive population management strategy, *Soft Computing*, vol. 21, no. 4, pp. 885-906, 2017.

[2] R. Azzouz, S. Bechikh and L. Said, A multiple reference point-based evolutionary algorithm for dynamic multi-objective optimization with undetectable changes, *2014 IEEE Congress on Evolutionary Computation*, pp. 3168-3175, 2014.

[3] X. Cai, H. Hu, M. Li, Y. Xiao, H. Ishibuchi and Z. Fan, A grid-based inverted generational distance for multi/many-objective optimization, *IEEE Transactions on Evolutionary Computation*, vol. 25, no. 1, pp. 21-34, 2021.

[4] L. Cao, L. Xu, E. Goodman, C. Bao and S. Zhu, Evolutionary dynamic multiobjective optimization assisted by a support vector regression predictor, *IEEE Transactions on Evolutionary Computation*, vol. 24, no. 2, pp. 305-319, 2019.

[5] W. Dai, Q. Yang, G. Xue and Y. Yu, Boosting for transfer learning, *International Conference on Machine Learning*, pp. 193-200, 2007.

[6] K. Deb, U. B. Rao and S. Karthik, Dynamic multi-objective optimization and decision-making using modified NSGA-II: a case study on hydro-thermal power scheduling, *International Conference on Evolutionary Multi-criterion Optimization*, pp. 803-817, 2007.

[7] J. Derrac, S. García, D. Molina and F. Herrera, A practical tutorial on the use of nonparametric statistical tests as a methodology for comparing evolutionary and swarm intelligence algorithms, *Swarm and Evolutionary Computation*, vol. 1, no. 1, pp. 3-18, 2011.

[8] L. Feng, W. Zhou, W. Liu, Y-S. Ong and K. Tan, Solving dynamic multiobjective problem via autoencoding evolutionary search, *IEEE Transactions on Cybernetics*, vol. 52, no. 5, pp. 2649-2662, 2022.

[9] A. Gretton, K. Borgwardt, M. Rasch, B. Schoelkopf and A. Smola, A kernel two-sample test, *The Journal of Machine Learning Research*, vol. 13, no. 1, pp. 723-773, 2012.

[10] J. Hu, C. Jia, H. Liu, X. Yi and Y. Liu, A survey on state estimation of complex dynamical networks, *International Journal of Systems Science*, vol. 52, no. 16, pp. 3351-3367, 2021.

[11] J. Huang, A. Smola, A. Gretton, K. Borgwardt and B. Schölkopf, Correcting sample selection bias by unlabeled data, *Advances in Neural Information Processing Systems 19: Proceedings of the 2006 Conference*, vol. 19, pp. 601-608, 2007.

[12] S. Jiang and S. Yang, A steady-state and generational evolutionary algorithm for dynamic multiobjective optimization, *IEEE Transactions on Evolutionary Computation*, vol. 21, no. 1, pp. 65-82, 2016.

[13] S. Jiang, S. Yang, X. Yao, K. Tan, M. Kaiser and N. Krasnogor, Benchmark problems for CEC2018 competition on dynamic multiobjevtive optimisation, *Proceedings of 2018 CEC Competition*, pp. 1-18, 2018.

[14] M. Jiang, Z. Huang, L. Qiu, W. Huang and G. Yen, Transfer learning-based dynamic multiobjective optimization algorithms, *IEEE Transactions on Evolutionary Computation*, vol. 22, no. 4, pp. 501-514, 2017.

[15] M. Jiang, Z. Wang, H. Hong and G. Yen, Knee point-based imbalanced transfer learning for dynamic multiobjective optimization, *IEEE Transactions on Evolutionary Computation*, vol. 25, no. 1, pp. 117-129, 2020.

[16] M. Jiang, Z. Wang, S. Guo, X. Gao and K. Tan, Individual-based transfer learning for dynamic multiobjective optimization, *IEEE Transactions on Cybernetics*, vol. 51, no. 10, pp. 4968-4981, 2021.

[17] M. Jiang, Z. Wang, L. Qiu, S. Guo, X. Gao and K. Tan, A fast dynamic evolutionary multiobjective algorithm via manifold transfer learning, *IEEE Transactions on Cybernetics*, vol. 51, no. 7, pp. 3417-3428, 2021.

[18] H. Li, X. Wang, S. Song and H. Li, Vehicle control strategies analysis based on PID and fuzzy logic control, *Procedia Engineering*, vol. 137, pp. 234-243, 2016.

[19] H. Li, J. Li, P. Wu, Y. You and N. Zeng, A ranking-system-based switching particle swarm optimizer with dynamic learning strategies, *Neurocomputing*, vol. 494, pp. 356-367, 2022.

[20] H. Li, P. Wu, N. Zeng, Y. Liu and F. E. Alsaadi, A survey on parameter identification, state estimation and data analytics for lateral flow immunoassay: from systems science perspective, *International Journal of Systems Science*, 2022. https://doi.org/10.1080/00207721.2022.2083262

[21] Q. Li, J. Zou, S. Yang, J. Zheng and G. Ruan, A predictive strategy based on special points for evolutionary dynamic multi-objective optimization, *Soft Computing*, vol. 23, no. 11, pp. 3723-3739, 2019.

[22] Z. Li, W. Yuan, Y. Chen, F. Ke, X. Chu and C. Chen, Neural-dynamic optimization-based model predictive control for tracking and formation of nonholonomic multirobot systems, *IEEE Transactions on Neural Networks and Learning Systems*, vol. 29, no. 12, pp. 6113-6122, 2018.

[23] Z. Li, J. Hu and J. Li, Distributed filtering for delayed nonlinear system with random sensor saturation: a dynamic event-triggered approach, *Systems Science & Control Engineering*, vol. 9, no. 1, pp. 440-454, 2021.

[24] Z. Liang, W. Liang, Z. Wang, X. Ma, L. Liu and Z. Zhu, Multiobjective evolutionary multitasking with two-stage adaptive knowledge transfer based on population distribution, *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 52, no. 7, pp. 4457-4469, 2022.

[25] Z. Liang, X. Xu, L. Liu, Y. Tu and Z. Zhu, Evolutionary many-task optimization based on multisource knowledge transfer, *IEEE Transactions on Evolutionary Computation*, vol. 26, no. 2, pp. 319-333, 2022.

[26] Z. Liang, T. Wu, X. Ma, Z. Zhu and S. Yang, A dynamic multiobjective evolutionary algorithm based on decision variable classification, *IEEE Transactions on Cybernetics*, vol. 52, no. 3, pp. 1602-1615, 2022.

[27] J. Lin, H. Liu, K. Tan and F. Gu, An effective knowledge transfer approach for multiobjective multitasking optimization, *IEEE Transactions on Cybernetics*, vol. 51, no. 6, pp. 3238-3248, 2021.

[28] X. Liu, Z. Zhan, T. Gu, S. Kwong, Z. Lu, H. Duh and J. Zhang, Neural network-based information transfer for dynamic optimization, *IEEE Transactions on Neural Networks and Learning Systems*, vol. 31, no. 5, pp. 1557-1570, 2020.

[29] Y. Liu, N. Zhu and M. Li, Solving many-objective optimization problems by a Pareto-based evolutionary algorithm with preprocessing and a penalty mechanism, *IEEE Transactions on Cybernetics*, vol. 51, no. 11, pp. 5585-5594, 2021.

[30] R. Liu, J. Li, Y. Jin and L. Jiao, A self-adaptive response strategy for dynamic multi-objective evolutionary optimization based on objective space decomposition, *Evolutionary Computation*, vol. 1, pp. 1-26, 2021.

[31] W. Liu, Z. Wang, X. Liu, N. Zeng and D. Bell, A novel particle swarm optimization approach for patient clustering from emergency departments, *IEEE Transactions on Evolutionary Computation*, vol. 23, no. 4, pp. 632-644, 2019.

[32] W. Liu, Z. Wang, Y. Yuan, N. Zeng, K. Hone and X. Liu, A novel sigmoid-function-based adaptive weighted particle swarm optimizer, *IEEE Transactions on Cybernetics*, vol. 51, no. 2, pp. 1085-1093, 2021.

[33] X. Luo, Y. Yuan, S. Chen, N. Zeng and Z. Wang, Position-transitional particle swarm optimization-incorporated latent factor analysis, *IEEE Transactions on Knowledge and Data Engineering*, vol. 34, no. 8, pp. 3958-3970, 2022.

[34] M. Mazouchi, Y. Yang and H. Modares, Data-driven dynamic multiobjective optimal control: an aspiration-satisfying reinforcement learning approach, *IEEE Transactions on Neural Networks and Learning Systems*, vol. 33, no. 11, pp. 6183-6193, 2022.

[35] A. Min, Y-S. Ong, A. Gupta and C-K. Goh, Multiproblem surrogates: transfer evolutionary multiobjective optimization of computationally expensive problems, *IEEE Transactions on Evolutionary Computation*, vol. 23, no. 1, pp. 15-28, 2017.

[36] S. Pan and Q. Yang, A survey on transfer learning, *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, no. 10, pp. 1345-1359, 2010.

[37] G. Ran, H. Chen, C. Li, G. Ma and B. Jiang, A hybrid design of fault detection for nonlinear systems based on dynamic optimization, *IEEE Transactions on Neural Networks and Learning Systems*, 2022. https://doi.org/10.1109/TNNLS.2022.3174822

[38] M. Rong, D. Gong, Y. Zhang, Y. Jin and P. Witold, Multidirectional prediction approach for synamic multiobjective optimization problems, *IEEE Transactions on Cybernetics*, vol. 49, no. 9, pp. 3362-3374, 2019.

[39] S. Sahmoud and H. Topcuoglu, Exploiting characterization of dynamism for enhancing dynamic multi-objective evolutionary algorithms, *Applied Soft Computing*, vol. 85, Article no. 105783, 2019.

[40] K. Tan, L. Feng and M. Jiang, Evolutionary transfer optimization-a new frontier in evolutionary computation research, *IEEE Computational Intelligence Magazine*, vol. 16, no. 1, pp. 22-33, 2021.

[41] Y. Tian, S. Peng, S. Yang, X. Zhang, K. Tan and Y. Jin, Action command encoding for surrogate-assisted neural architecture search, *IEEE Transactions on Cognitive and Developmental Systems*, vol. 14, no. 3, pp. 1129-1142, 2022.

[42] H. Tong, C. Huang, L. Minku and X. Yao, Surrogate models in evolutionary single-objective optimization: a new taxonomy and experimental study, *Information Sciences*, vol. 562, pp. 414-437, 2021.

[43] H. Tong, L. Minku, S. Menzel, B. Sendhoff and X. Yao, A novel generalized metaheuristic framework for dynamic capacitated arc routing problems, *IEEE Transactions on Evolutionary Computation*, vol. 26, no. 6, pp. 1486-1500, 2022.

[44] D. Wang, F. Liu and Y. Jin, A multi-objective evolutionary algorithm guided by directed search for dynamic scheduling, *Computers and Operations Research*, vol. 79, pp. 279-290, 2017.

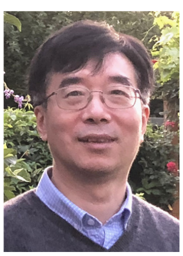
[45] Z. Wang, Z. Zhan, W. Yu, Y. Lin, J. Zhang, T. Gu and J. Zhang, Dynamic group learning distributed particle swarm optimization for large-scale optimization and its application in cloud workflow scheduling, *IEEE Transactions on Cybernetics*, vol. 50, no. 6, pp. 2715-2729, 2020.

[46] X. Wang, Y. Jin, S. Schmitt and M. Olhofer, Transfer learning based co-surrogate assisted evolutionary bi-objective optimization for objectives with non-uniform evaluation times, *Evolutionary Computation*, vol. 30, no. 2, pp. 221-251, 2022.

[47] L. Xu, B. Song and M. Cao, An improved particle swarm optimization algorithm with adaptive weighted delay velocity, *Systems Science & Control Engineering*, vol. 9, no. 1, pp. 188-197, 2021.

[48] Y. Ye, L. Li, Q. Lin, K. Wong, J. Li and Z. Ming, Knowledge guided Bayesian classification for dynamic multi-objective optimization, *Knowledge-Based Systems*, vol. 250, Article no. 109173, 2022.
[49] Y. Ye, Q. Lin, L. Ma, K. Wong, M. Gong and C. Coello, Multiple source transfer learning for dynamic multiobjective optimization, *Information Sciences*, vol. 607, pp. 739-757, 2022.
[50] N. Zeng, Z. Wang, W. Liu, H. Zhang, K. Hone and X. Liu, A dynamic neighborhood-based switching particle swarm optimization algorithm, *IEEE Transactions on Cybernetics*, vol. 52, no. 9, pp. 9290-9301, 2022.
[51] Z. Zhan, Z. Wang, H. Jin and J. Zhang, Adaptive distributed differential evolution, *IEEE Transactions on Cybernetics*, vol. 50, no. 11, pp. 4633-4647, 2020.
[52] J. Zhang, W. Zhou, X. Chen, W. Yao and L. Cao, Multisource selective transfer framework in multiobjective optimization problems, *IEEE Transactions on Evolutionary Computation*, vol. 24, no. 3, pp. 424-438, 2019.
[53] Q. Zhang and H. Li, MOEA/D: a multiobjective evolutionary algorithm based on decomposition, *IEEE Transactions on Evolutionary Computation*, vol. 11, no. 6, pp. 712-731, 2007.
[54] C. Zhang and J. Fu, An efficient dynamic optimization algorithm for path-constrained switched systems, *IEEE Transactions on Neural Networks and Learning Systems*, 2021. https://doi.org/10.1109/TNNLS.2021.3113345

Han Li received the bachelor degree in Measurement and Control Technology and Instrumentation from Xiamen University, Xiamen, China, in 2018. He is currently working towards the Ph.D. degree in Measuring and Testing Technologies and Instruments at Xiamen University, Xiamen, China. His research interests include swarm-based intelligent algorithms and evolutionary transfer optimization techniques.



Zidong Wang (Fellow, IEEE) was born in Jiangsu, China, in 1966. He received the B.Sc. degree in mathematics from Suzhou University, Suzhou, China, in 1986, and the M.Sc. degree in applied mathematics and the Ph.D. degree in electrical engineering from the Nanjing University of Science and Technology, Nanjing, China, in 1990 and 1994, respectively.

From 1990 to 2002, he held teaching and research appointments in universities in China, Germany, and the U.K. He is currently a Professor of dynamical systems and computing with the Department of Computer Science, Brunel University London, Uxbridge, U.K. He has published more than 600 papers in international journals. His research interests include dynamical systems, signal processing, bioinformatics, and control theory and applications.

Prof. Wang is also a member of the Academia Europaea, the European Academy of Sciences and Arts, and the program committee for many international conferences, an Academician of the International Academy for Systems and Cybernetic Sciences, and a fellow of the Royal Statistical Society. He is a holder of the Alexander von Humboldt Research Fellowship of Germany, the JSPS Research Fellowship of Japan, and the William Mong Visiting Research Fellowship of Hong Kong. He also serves (or has served) as the Editor-in-Chief of *International Journal of Systems Science*, *Neurocomputing*, and *Systems Science & Control Engineering* and an Associate Editor for 12 international journals, including IEEE TRANSACTIONS ON AUTOMATIC CONTROL, IEEE TRANSACTIONS ON CONTROL SYSTEMS TECHNOLOGY, IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS, IEEE TRANSACTIONS ON SIGNAL PROCESSING, and IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS—PART C.

Chengbo Lan received the bachelor degree in Machinery Design and Manufacturing and Its Automation from Guangdong University of Technology, GuangZhou, China, in 2018, and the master degree in Electronic Information at Xiamen University, Xiamen, China, in 2023. His research interests include dynamic multi-objective optimization and transfer learning techniques.

Peishu Wu received the bachelor degree in Measurement and Control Technology and Instrumentation from Tianjin University of Science and Technology, Tianjin, China, in 2020. He is currently pursuing the Ph.D. degree in Measuring and Testing Technologies and Instruments at Xiamen University, Xiamen, China. His research interests include computer vision and deep learning techniques.

Nianyin Zeng was born in Fujian Province, China, in 1986. He received the B.Eng. degree in electrical engineering and automation in 2008 and the Ph.D. degree in electrical engineering in 2013, both from Fuzhou University. From October 2012 to March 2013, he was a RA in the Department of Electrical and Electronic Engineering, the University of Hong Kong. From September 2017 to August 2018, he as an ISEF Fellow founded by the Korea Foundation for Advance Studies and also a Visiting Professor at the Korea Advance Institute of Science and Technology.

Currently, he is an Associate Professor with the Department of Instrumental & Electrical Engineering of Xiamen University. His current research interests include intelligent data analysis, computational intelligent, time-series modeling and applications. He is the author or co-author of several technical papers and also a very active reviewer for many international journals and conferences.

Dr. Zeng is currently serving as Associate Editors for Neurocomputing, Evolutionary Intelligence, and Frontiers in Medical Technology, and also Editorial Board members for Computers in Biology and Medicine, Biomedical Engineering Online, and Mathematical Problems in Engineering.