

Migrant Resettlement by Evolutionary Multi-objective Optimization

Dan-Xuan Liu, Yu-Ran Gu, Chao Qian, *Senior Member, IEEE*, Xin Mu, Ke Tang, *Fellow, IEEE*

Abstract—Migration has been a universal phenomenon, which brings opportunities as well as challenges for global development. As the number of migrants (e.g., refugees) increases rapidly in recent years, a key challenge faced by each country is the problem of migrant resettlement, intending to maximize the employment rate. This problem has attracted scientific research attention. Previous works mainly formulated migrant resettlement as an approximately submodular optimization problem subject to multiple matroid constraints and employed the greedy algorithm, whose performance, however, may be limited due to its greedy nature. In this paper, we propose a new framework MR-EMO based on Evolutionary Multi-objective Optimization, which reformulates Migrant Resettlement as a bi-objective optimization problem that maximizes the expected number of employed migrants and minimizes the number of dispatched migrants simultaneously, and employs a Multi-Objective Evolutionary Algorithm (MOEA) to solve the bi-objective problem. We implement MR-EMO using three MOEAs, the popular NSGA-II, MOEA/D as well as the theoretically grounded GSEMO. To further improve the performance of MR-EMO, we propose a specific MOEA, called GSEMO-SR, using matrix-swap mutation and repair mechanism, which has a better ability to search feasible solutions. We prove that MR-EMO using either GSEMO or GSEMO-SR can achieve better theoretical guarantees than the previous greedy algorithm. Experimental results under the interview and coordination migration models clearly show the superiority of MR-EMO (with either NSGA-II, MOEA/D, GSEMO or GSEMO-SR) over previous algorithms, and that using GSEMO-SR leads to the best performance of MR-EMO.

Index Terms—Migrant resettlement, approximately submodular optimization, matroid constraints, multi-objective optimization, multi-objective evolutionary algorithms.

I. INTRODUCTION

AS an important part of globalization, the phenomenon of migration has become an unstoppable trend. In 2019, the number of international migrants was 272 million, exceeding predictions for the year 2050 [21]. Migration is beneficial to the sustainable development of both host and home countries, as well as migrants themselves. For host countries, massive migrants can fill the manpower shortage [32], and entrepreneurs among them can even create new employment opportunities [9]. What behind the higher employment rates is the prosperity and development of all walks of life [39].

Dan-Xuan Liu, Yu-Ran Gu and Chao Qian are with the State Key Laboratory for Novel Software Technology, Nanjing University, Nanjing 210023, China (e-mails: {liudx, guyr, qianc}@lamda.nju.edu.cn). Chao Qian is the corresponding author.

Xin Mu is with the Peng Cheng Laboratory, Shenzhen 518000, China (e-mail: mux@pcl.ac.cn)

Ke Tang is with the Department of Computer Science and Engineering, Southern University of Science and Technology, Shenzhen 518055, China (e-mail: tangk3@sustech.edu.cn).

Meanwhile, migrants can earn more salaries, and bring back more foreign exchange and trade for home countries [31].

However, migration also brings challenges, one of which is migrant resettlement. For example, there were 82.4 million refugees in the world in 2020, more than double the level in 2010 [40]. Because refugees may lack language skills and have difficulties in adaptation, it is full of challenges and obstacles for policymakers to make them find jobs in the host countries [26], without developed systematic and pellucid resettlement procedures [7], which consequently have attracted scientific research attention.

A paper published in *Science* first tried to select a subset of migrant-locality pairs maximizing the expected number of employed migrants, through a data-driven algorithm combining supervised machine learning with optimal matching [2]. Utilizing historical registry data, supervised learning models are built to predict the employment probability of each migrant at each locality. By assuming that the employment of different migrants at the same locality is independent, the algorithm can lead to a 40%-70% migrants' employment gain relative to the actual outcomes.

However, this additive algorithm ignores the effect of competition in the real-world labor market. That is, whether one migrant gets a job or not is also influenced by other migrants' employment. Thus, Götz and Procaccia [15] introduced two migration models, i.e., interview and coordination models, to simulate the competition among migrants. For these models, the objective function, i.e., the expected number of migrants who find employment, all satisfies the ϵ -approximately submodular property, where $\epsilon \geq 0$. Furthermore, the constraints that each migrant can only go to one locality to apply for a job and each locality can accommodate a limited number of migrants, can be characterized by the intersection of k matroids, where $k \geq 2$. Thus, migrant resettlement is cast as the general problem of ϵ -approximately submodular maximization subject to k matroid constraints. Götz and Procaccia [15] proposed the greedy algorithm, which achieves an approximation guarantee of $1/(k+1 + \frac{4\epsilon r}{1-\epsilon})$, where r is the size of the largest feasible subset of migrant-locality pairs. They empirically showed that the greedy algorithm can bring improvement on both two models, compared with the additive algorithm [2].

Considering that the performance of the greedy algorithm may be limited due to its greedy nature, in this paper, we propose a new framework based on Evolutionary Multi-objective Optimization [13], [22], [33] for Migrant Resettlement, briefly called MR-EMO, which reformulates migrant resettlement as a bi-objective optimization problem that maximizes the ϵ -approximately submodular objective value of a feasible

subset of migrant-locality pairs and minimizes the subset size simultaneously. That is, MR-EMO tries to maximize the expected number of employed migrants while dispatching as few migrants as possible. Note that dispatching unnecessary migrants will lead to more financial expenses. MR-EMO can be equipped with any MOEA to solve this bi-objective problem, and we employ the popular NSGA-II [6], MOEA/D [44] as well as the theoretically grounded GSEMO [23]. Empirical results on the interview and coordination migration models show that using either NSGA-II, MOEA/D or GSEMO, MR-EMO performs better than previous algorithms.

Figure 1 gives the heat map of a migrant resettlement problem instance in the solution space. The instance means that six migrants go to two localities to find jobs under the interview migration model. The two matroid constraints are that each migrant can only go to one locality to apply for a job and each locality can accommodate at most three migrants. And the solution violating these two matroid constraints is infeasible. There are totally 6×2 migrant-locality pairs, and the solution of this instance can be naturally represented by a Boolean vector $\mathbf{x} \in \{0, 1\}^{12}$, where the i -th bit $x_i = 1$ iff the i -th migrant-locality pair is selected. We reduce the dimension of solutions from 12d to 2d using the t-SNE method [41], which can ensure that the solutions with closer positions in the original space also have closer positions in 2D projection space. In Figure 1, the color of a point corresponds to the objective function value, i.e., the expected number of migrants who find employment, of a solution; a point with the darkest color (i.e., the value -1) implies an infeasible solution. Then we can have two observations:

1. The whole space is filled with a large number of infeasible solutions;

2. Good feasible solutions may be surrounded by infeasible ones, as indicated by the circle.

Observation 1 reflects the character of solutions subject to two matroid constraints. Let a migrant-locality pair (v, l) denote that the migrant $v \in V$ is assigned to the locality $l \in L$. Let n denote the total number of migrant-locality pairs, i.e., $n = |V| \cdot |L|$. For any feasible solution $\mathbf{x} \in \{0, 1\}^n$, if the bit corresponding to migrant-locality pair (v_i, l_j) is 1, then any bit corresponding to migrant-locality (v_i, l_k) , where $k \neq j$, must be 0. Otherwise, the first matroid constraint, that each migrant can only go to one locality to apply for a job, will be violated. Note that \mathbf{x} also needs to satisfy the second matroid constraint, that is, for any locality j , the number of migrants assigned to locality j cannot exceed the capacity of the corresponding locality, which limits the number of 1s in those bits corresponding to (\cdot, l_j) . However, the classical bit-wise mutation operator flips each bit of \mathbf{x} independently with equal probability, which will probably generate an infeasible solution rather than a better feasible solution.

To better explore the feasible region, we propose the matrix-swap mutation operator. We first represent a solution $\mathbf{x} \in \{0, 1\}^n$ by a Boolean matrix of size $|V| \times |L|$, where $|V|$ and $|L|$ denote the number of migrants and localities, respectively, and $n = |V| \times |L|$. The matrix representation of a feasible solution satisfies that the number of 1-bits in each row is no more than 1, and the number of 1-bits in each column does

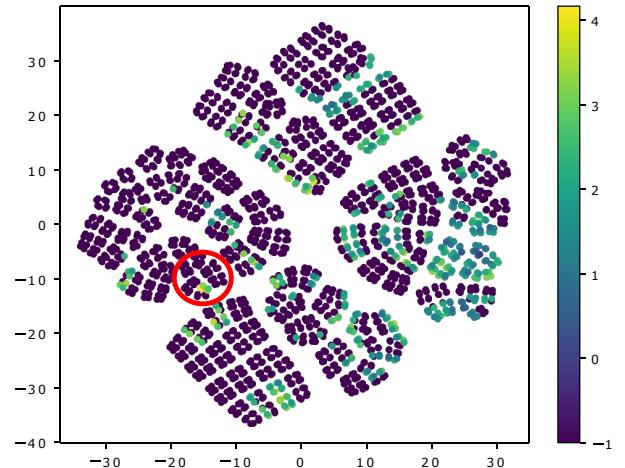


Figure 1. The heat map of a migrant resettlement problem instance (where six migrants go to two localities to find jobs under the interview migration model) in the solution space. The color of a point corresponds to the objective function value, i.e., the expected number of migrants who find employment, of a solution; a point with the darkest color (i.e., the value -1) implies an infeasible solution.

not exceed the capacity of the corresponding locality. Matrix-swap mutation selects two rows (or two columns) of a parent solution uniformly at random and swaps them to generate an offspring, which can better inherit the feasibility of the parent solution to reduce the probability of generating an infeasible offspring. Observation 2 shows that good feasible solutions may be surrounded by infeasible ones, which implies that infeasible solutions may be also valuable. Thus, we introduce a repair mechanism, which randomly flips some migrant-locality bits of an infeasible offspring solution from 1 to 0 to fix it into feasible, instead of discarding it directly.

By incorporating the proposed matrix-Swap mutation operator and the Repair mechanism into GSEMO, we propose a new MOEA, called GSEMO-SR. Empirical results show that MR-EMO applying GSEMO-SR is better than previous algorithms. Moreover, MR-EMO applying GSEMO-SR is significantly better than using NSGA-II, MOEA/D or GSEMO. The ablation experimental results show that GSEMO-SR employing only matrix-swap mutation or repair mechanism can also improve the performance of MR-EMO, and applying both of them performs the best, which verifies the effectiveness of the operators.

In addition to better empirical results, MR-EMO can also achieve a theoretical advantage. We prove that MR-EMO using no matter GSEMO or GSEMO-SR can achieve an approximation guarantee of $1/(k + \frac{1}{p} + \frac{2er}{1-\epsilon})$, which is better than $1/(k+1+\frac{4er}{1-\epsilon})$ of the greedy algorithm [15], where $p \geq 1$. Here, an approximation guarantee of α (where $0 < \alpha < 1$) is the worst-case theoretical guarantee, which is commonly used in the theoretical analysis of optimization algorithms. It implies that for any instance of migrant resettlement, the solution generated by the algorithm has an objective value at least $\alpha \cdot \text{OPT}$, where OPT denotes the optimal function value.

The rest of the paper is organized as follows. Section II first introduces some preliminaries on the studied problem, i.e., migrant resettlement. Section III then presents the general

MR-EMO framework and the specific MOEA (i.e., GSEMO-SR) for migrant resettlement. Sections IV and V give the theoretical analysis and empirical study, respectively. Section VI finally concludes this paper.

II. MIGRANT RESETTLEMENT

Let $V = \{v_1, \dots, v_{|V|}\}$ and $L = \{l_1, \dots, l_{|L|}\}$ denote a set of migrants and localities, respectively. A migrant-locality pair (v, l) means that the migrant $v \in V$ is assigned to the locality $l \in L$. For example, there are eight migrants and four localities in Figure 2, and each (solid or dotted) line denotes a migrant-locality pair. Let $N = V \times L$ denote the set of all migrant-locality pairs. The migrant resettlement problem is to select a subset X of migrant-locality pairs (i.e., $X \subseteq N$), that maximizes the expected number of employed migrants.

Suppose that the number of job opportunities at locality l is job_l . As there are different professions, e.g., drivers, chefs and painters in Figure 2, job_l is further divided, i.e., $\text{job}_l = \sum_{\pi \in \Pi} \text{job}_{l,\pi}$, where Π denotes the set of professions and $\text{job}_{l,\pi}$ denotes the number of job opportunities for profession π at locality l . To simulate the competition among migrants, Götz and Procaccia [15] have introduced two migration models, i.e., interview and coordination models, which will be introduced in the following, respectively. We will also show how to calculate the objective function, i.e., the expected number of migrants who find employment, for each model.

Interview Model. Migrants can only find jobs that match their types of professions, that is, competition exists among migrants with the same profession. Given a set X of selected migrant-locality pairs, let $V_{l,\pi} = \{v \mid (v, l) \in X \text{ and the profession of } v \text{ is } \pi\}$, i.e., the set of migrants with profession π assigned to locality l . To allocate the jobs of profession π at locality l , the interview model assumes that there is a specific order for the migrants in $V_{l,\pi}$, and each one will be interviewed sequentially. For $1 \leq i \leq |V_{l,\pi}|$, let $v_{l,\pi}^i$ denote the i -th migrant in $V_{l,\pi}$. When interviewing the migrant $v_{l,\pi}^i$, we use $\text{job}_{l,\pi}^i$ to denote the remaining number of available jobs of profession π at locality l . Then, $v_{l,\pi}^i$ will try to apply for each of the $\text{job}_{l,\pi}^i$ available jobs, and each attempt will succeed with probability $p_{v_{l,\pi}^i, l}$. If there is at least one success in these $\text{job}_{l,\pi}^i$ attempts, $v_{l,\pi}^i$ will be employed. Thus, the employment of $v_{l,\pi}^i$ can be represented by an indicator function $\mathbb{I}(\text{job}_{l,\pi}^i, p_{v_{l,\pi}^i, l})$, which equals 1 iff the binomially distributed random variable $\mathcal{B}(\text{job}_{l,\pi}^i, p_{v_{l,\pi}^i, l})$ is at least 1. At the beginning, $\text{job}_{l,\pi}^1 = \text{job}_{l,\pi}$, i.e., the initial number of job opportunities for profession π at locality l . As the interview process goes, $\forall i > 1 : \text{job}_{l,\pi}^i = \text{job}_{l,\pi}^{i-1} - \mathbb{I}(\text{job}_{l,\pi}^{i-1}, p_{v_{l,\pi}^{i-1}, l})$. Then, the number of migrants with profession π that find jobs at locality l can be represented as $\sum_{i=1}^{|V_{l,\pi}|} \mathbb{I}(\text{job}_{l,\pi}^i, p_{v_{l,\pi}^i, l})$, and the expected total number of employed migrants for the interview model is

$$f(X) = \sum_{l \in L} \sum_{\pi \in \Pi} \mathbb{E} \left[\sum_{i=1}^{|V_{l,\pi}|} \mathbb{I}(\text{job}_{l,\pi}^i, p_{v_{l,\pi}^i, l}) \right], \quad (1)$$

where the order of migrants in $V_{l,\pi}$ is determined randomly.

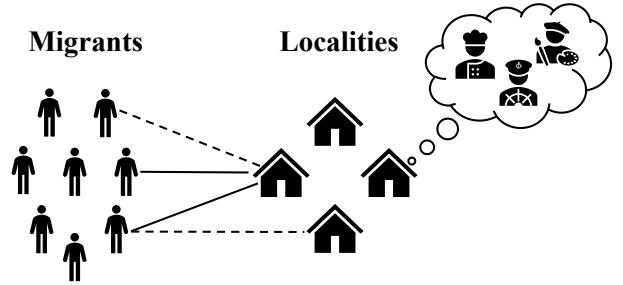


Figure 2. An illustration of migrant resettlement problem, where each edge (migrant-locality pair) denotes that the migrant is assigned to the locality.

Coordination Model. Different from the above interview model where a migrant can only find a job of her profession, the coordination model can coordinate the assignment by compatibility between all migrants and jobs at locality l , demonstrating a more realistic labor market. Each migrant v can compete for a job of profession π with probability $p_{v,\pi}$, which breaks the strict restriction between professions, because migrants also have opportunities to find other professional jobs related to their careers or less skilled. Given a set X of selected migrant-locality pairs, let $V_l = \{v \mid (v, l) \in X\}$ denote the set of migrants assigned to locality l , i.e., $V_l = \cup_{\pi \in \Pi} V_{l,\pi}$. We construct a bipartite graph $B(V_l)$, where the two sets of vertices are V_l and the job_l jobs, respectively. For each $v \in V_l$ and each job j , we add one edge between them with probability p_{v,π_j} , where π_j denotes the profession of job j . Then, the number of migrants that find jobs at locality l is calculated as the size of maximum matching of the bipartite graph $B(V_l)$. Thus, the expected number of employed migrants for the coordination model is

$$f(X) = \sum_{l \in L} \mathbb{E} [\text{size of maximum matching of } B(V_l)]. \quad (2)$$

The objective functions, i.e., Eqs. (1) and (2), of the above two models are all monotone submodular [15]. A set function $f : 2^N \rightarrow \mathbb{R}$ is monotone if $\forall X \subseteq Y \subseteq N$, $f(X) \leq f(Y)$. As more migrant-locality pairs will not worsen the value, the monotonicity is satisfied naturally. Assume w.l.o.g. that monotone functions are normalized, i.e., $f(\emptyset) = 0$. A set function f is submodular [27] if it satisfies the “diminishing returns” property, i.e., $\forall X \subseteq Y \subseteq N = V \times L$, $(v, l) \notin Y$,

$$f(X \cup \{(v, l)\}) - f(X) \geq f(Y \cup \{(v, l)\}) - f(Y), \quad (3)$$

or equivalently $\forall X \subseteq Y \subseteq N$,

$$f(Y) - f(X) \leq \sum_{(v,l) \in Y \setminus X} (f(X \cup \{(v, l)\}) - f(X)). \quad (4)$$

The diminishing returns property Eq. (3) reflects the competition effects naturally, i.e., the chance of migrant v to find a job at locality l decreases as the number of migrants competing for jobs increases.

For the migrant resettlement problem, there are also two constraints to be satisfied. One is that each migrant can only be distributed to at most one locality and get one job, and the other is that each locality l can accept at most cap_l migrants. cap_l is called the capacity of locality l , and can be different

from job_l (i.e., the number of job opportunities at locality l), because some jobs may be allocated to natives. Let us see Figure 2 for an illustration. Assume that the capacity of each locality is two. When we consider the two solid lines only, it is a feasible solution, i.e., satisfies the constraints. But if adding the two dotted lines, the constraints will be violated, since there is a migrant assigned to two localities and also a locality absorbing three migrants more than its capacity two.

Let $\{N_1^{\text{mig}}, \dots, N_{|V|}^{\text{mig}}\}$ and $\{N_1^{\text{loc}}, \dots, N_{|L|}^{\text{loc}}\}$ denote two partitions of N , where $N_i^{\text{mig}} = \{(v_i, l) \mid l \in L\}$ and $N_i^{\text{loc}} = \{(v, l_i) \mid v \in V\}$ are the set of migrant-locality pairs with migrant v_i and locality l_i , respectively. Then, the two constraints can be represented as $\forall 1 \leq i \leq |V| : |X \cap N_i^{\text{mig}}| \leq 1$ and $\forall 1 \leq i \leq |L| : |X \cap N_i^{\text{loc}}| \leq \text{cap}_i$, which are actually two partition matroids. A matroid is a pair $(N, \mathcal{F} \subseteq 2^N)$, satisfying the hereditary (i.e., $\emptyset \in \mathcal{F}$ and $\forall X \subseteq Y \in \mathcal{F} : X \in \mathcal{F}$) and augmentation (i.e., $\forall X, Y \in \mathcal{F}, |X| > |Y| : \exists (v, l) \in X \setminus Y, Y \cup \{(v, l)\} \in \mathcal{F}$) properties [4]. Thus, the constraints of migrant resettlement can be viewed as the intersection of two specific matroids. As observed from Figure 1, these two matroid constraints will lead to a large number of infeasible solutions.

In [15], migrant resettlement has been generally formulated as the problem of maximizing a monotone submodular function subject to multiple matroid constraints, as shown in Definition 1, where the monotone submodular objective function characterizes the expected number of employed migrants under different models.

Definition 1 (Migrant Resettlement). *Given all migrants $V = \{v_1, \dots, v_{|V|}\}$ and localities $L = \{l_1, \dots, l_{|L|}\}$, a monotone submodular function $f : 2^N \rightarrow \mathbb{R}^+$, and k matroids $\mathcal{F}_1, \dots, \mathcal{F}_k \subseteq 2^N$, the goal of migrant resettlement is to find a subset $X \subseteq N$ of migrant-locality pairs such that*

$$\arg \max_{X \subseteq N} f(X) \quad \text{s.t.} \quad X \in \bigcap_{i=1}^k \mathcal{F}_i, \quad (5)$$

where $N = V \times L$ is the set of all migrant-locality pairs.

However, we often cannot obtain the exact value of the objective f in practice, but only a noisy one, denoted as \hat{f} . For example, the objective functions in Eqs. (1) and (2) are all the expectation of a complicated random variable, which can be estimated only by the average of multiple sampling, resulting in the inexactness and even the violation of monotonicity and submodularity. Thus, the objective obtained by algorithms is assumed to be \hat{f} instead of f , which only satisfies the ϵ -approximate submodularity in Definition 2. That is, the value of \hat{f} is bounded between $(1 - \epsilon) \cdot f$ and $(1 + \epsilon) \cdot f$. When $\epsilon = 0$, \hat{f} is exactly the true objective f .

Definition 2 (ϵ -Approximate Submodularity [18], [37]). *Let $\epsilon \geq 0$. A set function $\hat{f} : 2^N \rightarrow \mathbb{R}$ is ϵ -approximately submodular if there exists a monotone submodular function f such that $\forall X \subseteq N$,*

$$(1 - \epsilon) \cdot f(X) \leq \hat{f}(X) \leq (1 + \epsilon) \cdot f(X). \quad (6)$$

For the migrant resettlement problem in Definition 1, Götz and Procaccia [15] proposed the greedy algorithm, which starts from the empty set, and iteratively adds one migrant-locality

Algorithm 1 MR-EMO Framework

Input: a migrant resettlement problem instance

Output: a subset of migrant-locality pairs satisfying multiple matroid constraints

Process:

- 1: Construct the two objective functions $f_1(\mathbf{x})$ and $f_2(\mathbf{x})$ to be maximized, as presented in Eq. (7);
 - 2: Apply an MOEA to solve the bi-objective problem;
 - 3: **return** the best feasible solution in the final population generated by the MOEA
-

pair with the largest marginal gain on \hat{f} while satisfying the constraints. The greedy algorithm achieves a $(1/(k+1+\frac{4\epsilon r}{1-\epsilon}))$ -approximation guarantee, where r denotes the size of the largest feasible solution, i.e., the largest subset of N belonging to $\bigcap_{i=1}^k \mathcal{F}_i$. That is, for any instance of migrant resettlement, the objective function value of the solution X generated by the greedy algorithm satisfies

$$f(X) \geq \left(1 / \left(k + 1 + \frac{4\epsilon r}{1 - \epsilon}\right)\right) \cdot \text{OPT},$$

where OPT denotes the optimal function value.

III. MR-EMO FRAMEWORK

Inspired by the excellent performance of MOEAs for solving general subset selection problems [13], [28], [33], [34], [36]–[38], we propose a new Migrant Resettlement framework based on Evolutionary Multi-objective Optimization, called MR-EMO. Assume that $|N| = |V| \cdot |L| = n$, i.e., there are totally n pairs of migrant-locality. A subset X of N can be naturally represented by a Boolean vector $\mathbf{x} \in \{0, 1\}^n$, where the i -th bit $x_i = 1$ iff the i -th migrant-locality pair in N is contained by X . In the following, we will not distinguish $\mathbf{x} \in \{0, 1\}^n$ and its corresponding subset of migrant-locality pairs for notational convenience.

As presented in Algorithm 1, MR-EMO first reformulates the original migrant resettlement problem in Definition 1 as a bi-objective maximization problem

$$\arg \max_{\mathbf{x} \in \{0, 1\}^n} (f_1(\mathbf{x}), f_2(\mathbf{x})), \quad (7)$$

where

$$f_1(\mathbf{x}) = \begin{cases} \hat{f}(\mathbf{x}), & \mathbf{x} \in \bigcap_{i=1}^k \mathcal{F}_i \\ -1, & \text{otherwise} \end{cases}, \quad f_2(\mathbf{x}) = |\mathbf{x}|_0.$$

That is, the first objective f_1 equals the original objective \hat{f} for feasible solutions (i.e., solutions satisfying the constraints), while -1 for infeasible ones; the second objective f_2 equals the number of 0-bits of a solution. Thus, MR-EMO tries to maximize the expected number of employed migrants (measured by \hat{f}) while dispatching as few migrants as possible. Note that for policymakers, dispatching unnecessary migrants also means more financial expenses. As the two objectives may be conflicting, the domination relationship in Definition 3 is often used for comparing two solutions.

Definition 3 (Domination). *For two solutions \mathbf{x} and \mathbf{x}' ,*

1. \mathbf{x} weakly dominates \mathbf{x}' (i.e., \mathbf{x} is better than \mathbf{x}' , denoted

Algorithm 2 GSEMO Algorithm**Process:**

- 1: Let $P = \{\mathbf{0}\}$;
- 2: **repeat**
- 3: Select \mathbf{x} from P uniformly at random;
- 4: Apply bit-wise mutation on \mathbf{x} to generate \mathbf{x}' ;
- 5: **if** $\nexists z \in P$ such that $z \succ \mathbf{x}'$ **then**
- 6: $P \leftarrow (P \setminus \{z \in P \mid \mathbf{x}' \succeq z\}) \cup \{\mathbf{x}'\}$
- 7: **end if**
- 8: **until** some criterion is met

- by $\mathbf{x} \succeq \mathbf{x}'$) if $\forall i : f_i(\mathbf{x}) \geq f_i(\mathbf{x}')$;
 2. \mathbf{x} dominates \mathbf{x}' (i.e., \mathbf{x} is strictly better than \mathbf{x}' , denoted by $\mathbf{x} \succ \mathbf{x}'$) if $\mathbf{x} \succeq \mathbf{x}' \wedge \exists i : f_i(\mathbf{x}) > f_i(\mathbf{x}')$;
 3. \mathbf{x} and \mathbf{x}' are incomparable if neither $\mathbf{x} \succeq \mathbf{x}'$ nor $\mathbf{x}' \succeq \mathbf{x}$.

After constructing the bi-objective problem in Eq. (7), MR-EMO employs an MOEA to solve it, as shown in line 2 of Algorithm 1. Evolutionary algorithms (EAs), inspired by Darwin's theory of evolution, are general-purpose randomized heuristic optimization algorithms [1], mimicking variational reproduction and natural selection. Starting from an initial population of solutions, EAs iteratively reproduce offspring solutions by crossover and mutation, and select better ones from the parent and offspring solutions to form the next population. The population-based search of EAs matches the requirement of multi-objective optimization, i.e., EAs can generate a set of Pareto optimal solutions by running only once. Thus, EAs have become the most popular tool for multi-objective optimization [5], [16], [17], [43], and the corresponding algorithms are also called MOEAs. After running a number of iterations, the best solution, i.e., $\arg \max_{\mathbf{x} \in P, \mathbf{x} \in \bigcap_{i=1}^k \mathcal{F}_i} \hat{f}(\mathbf{x})$ will be selected from the final population as the output, as shown in line 3 of Algorithm 1.

Note that MR-EMO can be equipped with any MOEA. In this paper, we will apply NSGA-II [6], MOEA/D [44] and GSEMO [14], [23], [29]. NSGA-II is a representative of Pareto dominance based MOEAs, which selects two parents by binary tournament selection, employs crossover and mutation to generate offspring solutions, and updates the population based on non-dominated sorting and crowding distance. MOEA/D is a representative of decomposition based MOEAs, which decomposes a multi-objective optimization problem into a number of scalar optimization subproblems and optimizes them collaboratively. For each current solution, it randomly selects two neighbor solutions as parents, and employs crossover and mutation to generate an offspring solution. By comparing the newly generated offspring solution with neighbor solutions, the better one will be kept in the population.

GSEMO is relatively simple but has shown good theoretical properties in solving many problems [10], [30], [47]. As presented in Algorithm 2, it starts from the all-0 vector $\mathbf{0}$ (i.e., the empty set) in line 1, and iteratively improves the quality of solutions in the population P (lines 2–7). In each iteration, it first selects a parent solution \mathbf{x} from the current population P uniformly at random in line 3. Next, it applies bit-wise mutation on \mathbf{x} (lines 4), which flips each bit of \mathbf{x} with

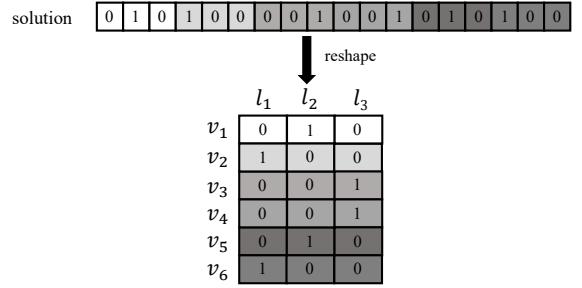


Figure 3. A feasible solution (Boolean vector representation and matrix representation) of a migrant resettlement problem instance with six migrants and three localities, where each locality can accommodate at most two migrants.

probability $1/n$ to generate an offspring solution \mathbf{x}' . Then, \mathbf{x}' is used to update the population P (lines 5–7). If \mathbf{x}' is not dominated by any solution in P (line 5), it will be added into P , and meanwhile, those solutions weakly dominated by \mathbf{x}' will be deleted (line 6). By the updating procedure in lines 5–7 of Algorithm 2, we know that the solutions in the population P must be incomparable. Furthermore, P will always contain the initial solution $\mathbf{0}$, because $\mathbf{0}$ has the largest value (i.e., n) of f_2 , and no other solution can weakly dominate it. Thus, any infeasible solution \mathbf{x} will be excluded from P , because $f_1(\mathbf{x}) = -1$ and $f_2(\mathbf{x}) < n$, implying that \mathbf{x} is dominated by $\mathbf{0}$, whose two objective values are 0 and n , respectively. Note that the original version of GSEMO in [23] starts from a random initial solution, while we modify it to start from the all-0s solution $\mathbf{0}$, as the all-0s solution will be used in theoretical analysis, which will be clear in the next section.

Our experimental results will show that MR-EMO using the existing NSGA-II, MOEA/D or GSEMO has already performed better than previous algorithms. Next, by analyzing the characteristics of the migrant resettlement problem and the limitation of the traditional reproduction operators, we will develop a new MOEA, which is more suitable for MR-EMO and will lead to better performance.

A. GSEMO-SR Algorithm

As observed in Figure 1, there are a large number of infeasible solutions for the migration resettlement problem. In the previous section, we have represented a subset X of N migrant-locality pairs (i.e., $X \subseteq N = V \times L$, where V and L denote the set of migrants and localities, respectively) by a Boolean vector $\mathbf{x} \in \{0, 1\}^n$, where $n = |N| = |V| \times |L|$. If using the traditional one-point crossover and bit-wise mutation operators, the generated offspring solutions are easily infeasible, and thus the feasible regions cannot be explored efficiently. To better illustrate this phenomenon, we also represent a solution $\mathbf{x} \in \{0, 1\}^n$ by a matrix of size $|V| \times |L|$, where each row and column correspond to one migrant v_i and one locality l_j , respectively, and $x_{\{v_i, l_j\}}$ denotes the element at the i -th row and j -th column of the matrix, corresponding to the $((i-1) \cdot |L| + j)$ -th bit of \mathbf{x} . Take a migrant resettlement problem instance with six migrants and three localities as an

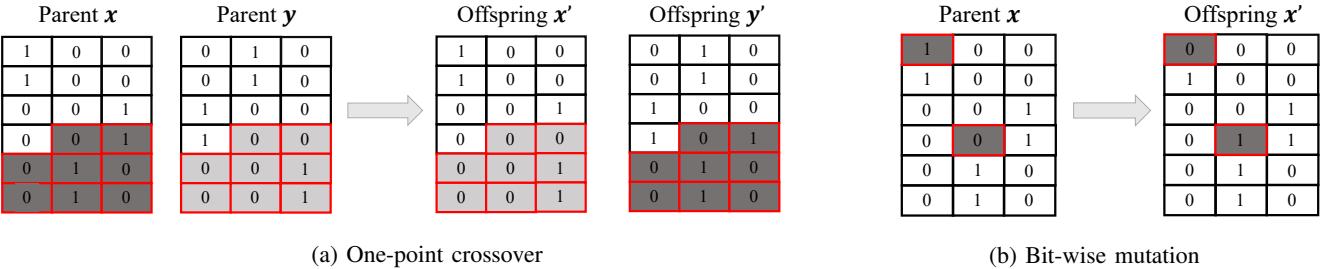


Figure 4. Examples for illustrating one-point crossover (a) and bit-wise mutation (b).

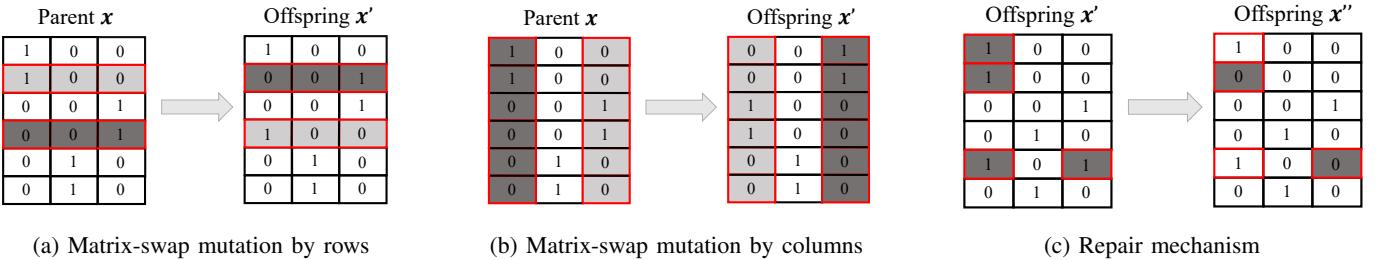


Figure 5. Examples for illustrating matrix-swap mutation by rows (a) and columns (b), and repairing an infeasible solution (c).

example, where each capacity of the locality is two. Figure 3 gives a Boolean vector solution of this instance, as well as its corresponding matrix representation, where migrants $\{v_2, v_6\}$, $\{v_1, v_5\}$ and $\{v_3, v_4\}$ are assigned to locality l_1 , l_2 and l_3 , respectively. This solution is feasible because the number of 1-bits in each row does not exceed 1 and the number of 1-bits in each column does not exceed 2, i.e., the capacity of each locality. That is, each migrant is distributed to at most one locality and each locality accepts at most two migrants; thus, the solution satisfies the two matroid constraints. Note that by the matrix representation, the behaviors of one-point crossover and bit-wise mutation can still be illustrated clearly.

One-point crossover selects $i \in \{1, 2, \dots, n\}$ randomly, and exchanges the bits after the i -th bit of two parent solutions. For Figure 4(a), the one-point crossover operator selects $i = 10$ and then exchanges the bits after the 10-th bit of \mathbf{x} and \mathbf{y} to generate offspring solutions \mathbf{x}' and \mathbf{y}' . Bit-wise mutation flips each bit of a solution independently with probability $1/n$. For Figure 4(b), the bit-wise mutation flips the 1-st and 11-th bits of \mathbf{x} and keeps other bits unchanged to generate an offspring solution \mathbf{x}' . In the above two examples, we can find that the parents are all feasible, but produce infeasible offspring solutions. In fact, this phenomenon will often occur by applying traditional one-point crossover and bit-wise mutation. For parent solutions \mathbf{x} and \mathbf{y} in Figure 4(a), the results of executing one-point crossover on \mathbf{x} and \mathbf{y} all lead to two infeasible offspring solutions, except for exchanging the bits after the 1-st, 17-th or 18-th bit of them. For parent solution \mathbf{x} in Figure 4(b), an extra 1-bit in any row or column of \mathbf{x} will result in an infeasible solution \mathbf{x}' . Thus, the one-point crossover and bit-wise mutation operators may not explore the feasible solution space efficiently.

We then propose a new mutation operator named matrix-swap mutation in Definition 4, which has a better ability to explore the feasible region of solution space. It can be

performed by rows or columns. Figure 5(a) shows that matrix-swap mutation by rows selects the second and the fourth rows uniformly at random, and exchanges the corresponding bits of the second and the fourth rows of \mathbf{x} to generate a new offspring solution \mathbf{x}' . The matrix-swap mutation ensures that \mathbf{x}' is still feasible because it has the same number of 1-bits of each row and column as its feasible parent \mathbf{x} . Also in Figure 5(b), matrix-swap mutation by columns randomly selects the first and the third columns, and then \mathbf{x}' is generated by swapping the corresponding bits of the first and the third columns of \mathbf{x} . Note that \mathbf{x}' is also feasible if the capacity of each locality is the same. Compared with classical operators (i.e., the one-point crossover and bit-wise mutation operators), the matrix-swap mutation operator can better explore the solution space while maintaining solution feasibility.

Definition 4 (Matrix-swap mutation). *Let $M^{|V| \times |L|}$ be a matrix representation of solution $\mathbf{x} \in \{0, 1\}^n$, where $n = |V| \times |L|$. The matrix-swap mutation by rows selects $i, j \in \{1, \dots, |V|\}$ uniformly at random with replacement, and then exchanges the i -th row and the j -th row of M . Similarly, matrix-swap mutation by columns uniformly selects $i, j \in \{1, \dots, |L|\}$ at random with replacement and exchanges the i -th column and the j -th column of M .*

As observed in Figure 1, good feasible solutions may be surrounded by infeasible ones. Thus, we introduce a repair mechanism in Definition 5 to repair an infeasible offspring solution to be feasible, instead of discarding it directly. For an offspring \mathbf{x}' in Figure 5(c), the number of 1-bits of the fifth row of \mathbf{x}' is 2, and then the repair mechanism randomly flips $1 = \max\{2 - 1, 0\}$ 1-bit of the fifth row. Similarly, the total number of 1-bits of the first column is 3, which exceeds the capacity 2 of the first locality, and $1 = \max\{3 - 2, 0\}$ 1-bit of the first column has been flipped. After repairing each row and each column of \mathbf{x}' in this way, the new offspring \mathbf{x}'' must

be feasible. Note that the repair mechanism will not change any bit of a feasible offspring solution.

Definition 5 (Repair mechanism). Let $M^{|V| \times |L|}$ be a matrix representation of solution $\mathbf{x} \in \{0, 1\}^n$, where $n = |V| \times |L|$. The repair mechanism first counts the total number of 1-bits of each row i from M , denoted by $x_{\{v_i, \cdot\}} = \sum_{j=1}^{|L|} x_{\{v_i, l_j\}}$. Then it randomly flips $\max\{x_{\{v_i, \cdot\}} - 1, 0\}$ 1-bits of row i and keeps other 0-bits unchanged. Next, it also counts the total number of 1-bits of each column j , denoted by $x_{\{\cdot, l_j\}} = \sum_{i=1}^{|V|} x_{\{v_i, l_j\}}$, and randomly flips $\max\{x_{\{\cdot, l_j\}} - cap_{l_j}, 0\}$ 1-bits and keeps other 0-bits unchanged.

By incorporating the matrix-Swap mutation operator and the Repair mechanism into GSEMO, we propose a new MOEA, called GSEMO-SR. As presented in Algorithm 3, it starts from the all-0 vector $\mathbf{0}$ (i.e., the empty set) in line 1, and iteratively improves the quality of solutions in the population P (lines 2–13). In each iteration, it first selects a parent solution \mathbf{x} from the current population P uniformly at random in line 3. Next, it applies bit-wise mutation on \mathbf{x} with probability p_m (lines 5–6), which flips each bit of \mathbf{x} independently with probability $1/n$ to generate an offspring solution \mathbf{x}' . Otherwise, GSEMO-SR applies the proposed matrix-swap mutation operator on \mathbf{x} (lines 7–8), which can be performed by rows or columns with equal probability. Then the offspring solution \mathbf{x}' will go through a repair mechanism in line 10, which can repair an infeasible solution to be feasible. Note that \mathbf{x}'' equals to \mathbf{x}' if \mathbf{x}' has already been feasible, and GSEMO-SR always contains feasible solutions. The repaired offspring solution, \mathbf{x}'' is used to update the population P (lines 11–13). If \mathbf{x}'' is not dominated by any solution in P (line 11), it will be added into P , and meanwhile, those solutions weakly dominated by \mathbf{x}'' will be deleted (line 12). Similar to GSEMO, the solutions contained by the population P of GSEMO-SR must be incomparable.

Compared with GSEMO-SR, GSEMO in Algorithm 2 only applies bit-wise mutation to generate a new offspring solution and no repair mechanism. In the experiments, we will show that using GSEMO-SR leads to the best performance of MR-EMO, which is significantly better than using NSGA-II, MOEA/D or GSEMO. We will also perform ablation study to validate the effectiveness of the introduced matrix-swap mutation operator and repair mechanism. Note that the matrix-swap mutation operator and repair mechanism can also be combined with other MOEAs to bring performance improvement, which will be studied in the future.

IV. THEORETICAL ANALYSIS

Before examining the empirical performance of the proposed MR-EMO framework, we show that MR-EMO can achieve better theoretical guarantees than the previous greedy algorithm [15] in this section.

First, we prove in Theorem 1 that MR-EMO using GSEMO (briefly called MR-GSEMO) can achieve an approximation guarantee of $1/(k + \frac{1}{p} + \frac{2\epsilon r}{1-\epsilon} + \frac{(1+\epsilon)\delta r}{1-\epsilon})$ after running at most $O(\frac{r}{\delta} n^{(k+1)p} \log \frac{(1+\epsilon)r}{1-\epsilon})$ expected number of iterations, where k is the number of matroids, ϵ is the degree of approximate

Algorithm 3 GSEMO-SR Algorithm

Parameter: probability $p_m \in \{0, 1\}$

Process:

```

1: Let  $P = \{\mathbf{0}\}$ ;
2: repeat
3:   Select  $\mathbf{x}$  from  $P$  uniformly at random;
4:    $q \leftarrow$  a number selected from  $[0, 1]$  uniformly at random;
5:   if  $q \leq p_m$  then
6:     Apply bit-wise mutation on  $\mathbf{x}$  to generate  $\mathbf{x}'$ 
7:   else
8:     Apply matrix-swap mutation by rows or columns
9:     (each with probability  $1/2$ ) on  $\mathbf{x}$  to generate  $\mathbf{x}'$ 
10:    end if
11:    $\mathbf{x}'' \leftarrow$  Repair( $\mathbf{x}'$ );
12:   if  $\nexists z \in P$  such that  $z \succ \mathbf{x}''$  then
13:      $P \leftarrow (P \setminus \{z \in P \mid \mathbf{x}'' \succeq z\}) \cup \{\mathbf{x}''\}$ 
14:   end if
15: until some criterion is met

```

submodularity of \hat{f} as in Eq. (6) and r is the size of the largest feasible solution. Note that OPT denotes the optimal function value of Eq. (5).

Theorem 1. *MR-GSEMO using $O\left(\frac{r}{\delta} n^{(k+1)p} \log \frac{(1+\epsilon)r}{1-\epsilon}\right)$ expected number of iterations finds a solution $X \in \bigcap_{i=1}^k \mathcal{F}_i$ with*

$$\left(k + \frac{1}{p} + \frac{2\epsilon r}{1-\epsilon} + \frac{(1+\epsilon)\delta r}{1-\epsilon}\right) \cdot f(X) \geq \text{OPT}, \quad (8)$$

where $k \geq 2$, $p \geq 1$, $\epsilon \geq 0$ and $\delta > 0$.

The main proof idea is to show that MR-GSEMO can find a $(1 + \delta)$ -approximate local optimal solution in Definition 6, which guarantees the desired approximation ratio as shown in Lemma 1.

Definition 6 (($1 + \delta$)-approximate local optimum). A $(1 + \delta)$ -approximate local optimal solution X is a feasible solution such that no feasible solution X' with the objective value $\hat{f}(X') > (1 + \delta) \cdot \hat{f}(X)$ can be achieved by deleting at most $k p$ elements from X and adding at most p new elements, where $\delta > 0$ and $p \geq 1$.

Lemma 1. Any $(1 + \delta)$ -approximate local optimal solution X satisfies

$$\left(k + \frac{1}{p} + \frac{2\epsilon r}{1-\epsilon} + \frac{(1+\epsilon)\delta r}{1-\epsilon}\right) \cdot f(X) \geq \text{OPT},$$

i.e., Eq. (8), where $k \geq 2$, $p \geq 1$, $\epsilon \geq 0$ and $\delta > 0$.

The proof of Lemma 1 will use the following two lemmas.

Lemma 2 (Lemma 1.1 in [24]). *Let f be a submodular function. For any $X, C \subseteq N$, let $\{T_j\}_{j=1}^t$ be a collection of subsets of $C \setminus X$ such that each element of $C \setminus X$ appears in exactly k of these subsets. Then, $\sum_{j=1}^t (f(X \cup T_j) - f(X)) \geq k(f(X \cup C) - f(X))$.*

Lemma 3 (Lemma 1.2 in [24]). *Let f be a submodular function. For any $X' \subseteq X \subseteq N$, let $\{T_j\}_{j=1}^t$ be a collection of*

subsets of $X \setminus X'$ such that each element of $X \setminus X'$ appears in exactly k of these subsets. Then, $\sum_{j=1}^t (f(X) - f(X \setminus T_j)) \leq k(f(X) - f(X'))$.

Proof of Lemma 1. Let X denote a $(1+\delta)$ -approximate local optimal solution, and C denote any other feasible solution. By reusing the construction process in Lemma 3.1 of [24], there are two collections of subsets $\{W_j\}_{j=1}^t$ of $C \setminus X$ and $\{\Lambda_j\}_{j=1}^t$ of $X \setminus C$ for $t \leq p2^q|C|$ and $q \geq 0$, such that each element of $C \setminus X$ appears in exactly $p2^q$ of sets $\{W_j\}_{j=1}^t$ and each element $e \in X \setminus C$ appears in exactly n_e of sets $\{\Lambda_j\}_{j=1}^t$, where $n_e \leq (k+1/p-1)p2^q$. Furthermore, $|W_j| \leq p$, $|\Lambda_j| \leq kp$, and the set $(X \setminus \Lambda_j) \cup W_j$ is feasible. As X is $(1+\delta)$ -approximately local optimal, we have

$$\hat{f}((X \setminus \Lambda_j) \cup W_j) \leq (1+\delta) \cdot \hat{f}(X)$$

according to Definition 6. Using the relation between \hat{f} and f in Eq. (6), we can further get

$$f((X \setminus \Lambda_j) \cup W_j) \leq \frac{1+\epsilon}{1-\epsilon} (1+\delta) \cdot f(X). \quad (9)$$

By the submodularity of f , i.e., Eq. (3), we have

$$\begin{aligned} & \sum_{j=1}^t (f(X \cup W_j) - f(X)) \\ & \leq \sum_{j=1}^t (f((X \setminus \Lambda_j) \cup W_j) - f(X \setminus \Lambda_j)) \\ & \leq \sum_{j=1}^t \left(\frac{1+\epsilon}{1-\epsilon} (1+\delta) \cdot f(X) - f(X \setminus \Lambda_j) \right) \\ & \leq \sum_{j=1}^t \left(\frac{1+\epsilon}{1-\epsilon} (1+\delta) \cdot f(X) - f(X \setminus \Lambda_j) \right) \\ & \quad + \sum_{e \in X \setminus C} ((k+1/p-1)p2^q - n_e)(f(X) - f(X \setminus \{e\})) \\ & = t \left(\frac{1+\epsilon}{1-\epsilon} (1+\delta) - 1 \right) f(X) + \sum_{j=1}^t (f(X) - f(X \setminus \Gamma_j)), \end{aligned} \quad (10)$$

where the second inequality holds by Eq. (9), the third inequality holds by $n_e \leq (k+1/p-1)p2^q$ and $f(X) \geq f(X \setminus \{e\})$ due to the monotonicity of f , and the equality holds by letting $\lambda = t + \sum_{e \in X \setminus C} ((k+1/p-1)p2^q - n_e)$, $\forall 1 \leq j \leq t : \Gamma_j = \Lambda_j$, and $\{\Gamma_j\}_{j=t+1}^\lambda$ be a collection of singleton sets $\{e\}$ where for each $e \in X \setminus C$, $\{e\}$ appears $(k+1/p-1)p2^q - n_e$ times.

As $\{W_j\}_{j=1}^t$ is a collection of subsets of $C \setminus X$ such that each element of $C \setminus X$ appears in exactly $p2^q$ of these subsets, applying Lemma 2 to the left side of Eq. (10) leads to

$$\sum_{j=1}^t (f(X \cup W_j) - f(X)) \geq p2^q(f(X \cup C) - f(X)). \quad (11)$$

We know that $\{\Gamma_j\}_{j=1}^\lambda$ is a collection of subsets of $X \setminus C$, and each element $e \in X \setminus C$ appears in exactly $(k+1/p-1)p2^q$ of these subsets (i.e., n_e times in $\{\Gamma_j\}_{j=1}^t = \{\Lambda_j\}_{j=1}^t$ and $(k+1/p-1)p2^q - n_e$ times in $\{\Gamma_j\}_{j=t+1}^\lambda$). Thus, applying Lemma 3 to the last term of Eq. (10) leads to

$$\begin{aligned} & \sum_{j=1}^\lambda (f(X) - f(X \setminus \Gamma_j)) \\ & \leq (k+1/p-1)p2^q \cdot (f(X) - f(X \cap C)). \end{aligned} \quad (12)$$

As the size of a feasible solution is at most r , we have $t \leq p2^q|C| \leq rp2^q$. By applying this inequality, Eqs. (11) and (12) to Eq. (10), we have

$$\begin{aligned} p2^q(f(X \cup C) - f(X)) & \leq rp2^q \left(\frac{1+\epsilon}{1-\epsilon} (1+\delta) - 1 \right) f(X) \\ & \quad + (k+1/p-1)p2^q \cdot (f(X) - f(X \cap C)). \end{aligned}$$

By simple calculation, we have

$$\begin{aligned} & \left(k + \frac{1}{p} + \frac{2\epsilon r}{1-\epsilon} + \frac{(1+\epsilon)\delta r}{1-\epsilon} \right) \cdot f(X) \\ & \geq f(X \cup C) + (k+1/p-1) \cdot f(X \cap C). \end{aligned} \quad (13)$$

Let C be an optimal solution of Eq. (5), i.e., $f(C) = \text{OPT}$. By the monotonicity of f , $f(X \cup C) \geq f(C) = \text{OPT}$, and $f(X \cap C) \geq f(\emptyset) = 0$. Thus, Eq. (13) becomes

$$\left(k + \frac{1}{p} + \frac{2\epsilon r}{1-\epsilon} + \frac{(1+\epsilon)\delta r}{1-\epsilon} \right) \cdot f(X) \geq \text{OPT},$$

implying that the lemma holds. \square

Now we are ready to prove Theorem 1. Note that as presented in Algorithm 2, GSEMO only uses the bit-wise mutation operator to generate a new offspring solution in each iteration, and uses this offspring solution to update the population such that only the non-dominated solutions generated so far are kept in the population.

Proof of Theorem 1. We divide the optimization process of MR-GSEMO into two phases: (1) starts from the initial solution $\mathbf{0}$ (i.e., \emptyset) and finishes after finding a feasible solution \mathbf{x} with the objective value $\hat{f}(\mathbf{x}) \geq ((1-\epsilon)/r) \cdot \text{OPT}$; (2) starts after phase (1) and finishes after finding a feasible solution with the desired approximation ratio in Eq. (8). Note that the population P will always contain feasible solutions, because infeasible solutions are dominated by the initial solution $\mathbf{0}$.

Let \mathbf{x}^* denote an optimal solution, i.e., $f(\mathbf{x}^*) = \text{OPT}$. Let $\{(v, l)^*\} = \arg \max_{\{(v, l)\} \in \bigcap_{i=1}^k \mathcal{F}_i} \hat{f}(\{(v, l)\})$. By the submodularity of f in Eq. (4) and $\hat{f}(\emptyset) = 0$, we have

$$\begin{aligned} \text{OPT} = f(\mathbf{x}^*) - f(\emptyset) & \leq \sum_{(v, l) \in \mathbf{x}^*} (f(\{(v, l)\}) - f(\emptyset)) \\ & \leq \sum_{(v, l) \in \mathbf{x}^*} \frac{\hat{f}(\{(v, l)\})}{1-\epsilon} \leq r \cdot \frac{\hat{f}(\{(v, l)^*\})}{1-\epsilon}, \end{aligned}$$

where the second inequality holds by Eq. (6), and the last inequality holds by the definition of $\{(v, l)^*\}$ and $|\mathbf{x}^*| \leq r$, i.e., the largest size of a feasible solution. Thus,

$$\hat{f}(\{(v, l)^*\}) \geq ((1-\epsilon)/r) \cdot \text{OPT}.$$

Note that the initial solution $\mathbf{0}$ will always exist in the population P , because it has the largest f_2 value and no other solution can dominate it. By selecting $\mathbf{0}$ in line 3 of Algorithm 2 and flipping only one specific 0-bit (corresponding to the migrant-locality pair $(v, l)^*$) by bit-wise mutation in line 4, the solution $\{(v, l)^*\}$ will be generated, which will then always exist in P , because it is the best solution with the f_2 value of $n-1$ and no other solution can dominate it. This implies that the goal of phase (1) has been reached. Because the probability of selecting $\mathbf{0}$ in line 3 is $1/|P|$ due

to uniform selection, and the probability of flipping a specific 0-bit while keeping other bits unchanged in bit-wise mutation is $(1/n) \cdot (1 - 1/n)^{n-1} \geq 1/(en)$, the expected number of iterations of phase (1) is at most $en|P| \leq e(r+1)n$. The population size $|P| \leq r+1$ holds, because MR-GSEMO excludes infeasible solutions from the population P , and P can contain at most one feasible solution for each size in $\{0, 1, \dots, r\}$ due to the incomparability of solutions in P .

Lemma 1 shows that a $(1 + \delta)$ -approximate local optimal solution achieves the desired approximation ratio in Eq. (8). Thus, for phase (2), we only need to analyze the expected number of iterations until generating a $(1 + \delta)$ -approximate local optimal solution. We consider the largest \hat{f} (i.e., f_1) value of the solutions in P , denoted by F_{\max} . That is, $F_{\max} = \max\{\hat{f}(\mathbf{x}) \mid \mathbf{x} \in P\}$. After phase (1), $F_{\max} \geq ((1 - \epsilon)/r) \cdot \text{OPT}$, and let \mathbf{x} be the corresponding solution, i.e., $\hat{f}(\mathbf{x}) = F_{\max}$. Obviously, F_{\max} cannot decrease, because \mathbf{x} cannot be weakly dominated by a solution with a smaller \hat{f} value. As long as \mathbf{x} is not $(1 + \delta)$ -approximately local optimal, we know that a new solution \mathbf{x}' with $\hat{f}(\mathbf{x}') > (1 + \delta) \cdot \hat{f}(\mathbf{x}) = (1 + \delta) \cdot F_{\max}$ can be generated through selecting \mathbf{x} in line 3 of Algorithm 2, and flipping at most $(k+1)p$ bits (i.e., deleting at most kp elements from \mathbf{x} and adding at most p new elements) in line 4, whose probability is at least $(1/|P|) \cdot (1/n)^{(k+1)p} (1 - 1/n)^{n-(k+1)p} \geq 1/(e(r+1)n^{(k+1)p})$. Since \mathbf{x}' now has the largest \hat{f} value and no solution in P can dominate it, it will be included into P . Thus, F_{\max} can increase by at least a factor of $1 + \delta$ with probability at least $1/(e(r+1)n^{(k+1)p})$ in each iteration, implying at most $e(r+1)n^{(k+1)p}$ expected number of iterations. As the maximum value of \hat{f} is at most $(1 + \epsilon) \cdot \text{OPT}$, the required number of such an increase on F_{\max} is at most

$$\log_{1+\delta} \frac{(1 + \epsilon) \cdot \text{OPT}}{((1 - \epsilon)/r) \cdot \text{OPT}} = O\left(\frac{1}{\delta} \log \frac{(1 + \epsilon)r}{1 - \epsilon}\right),$$

until generating a $(1 + \delta)$ -approximate local optimal solution. Thus, the expected number of iterations of phase (2) is at most $e(r+1)n^{(k+1)p} \cdot O\left(\frac{1}{\delta} \log \frac{(1+\epsilon)r}{1-\epsilon}\right) = O\left(\frac{r}{\delta} n^{(k+1)p} \log \frac{(1+\epsilon)r}{1-\epsilon}\right)$, implying that the theorem holds. \square

Thus, we have proven the polynomial-time approximation guarantee, $1/(k + \frac{1}{p} + \frac{2\epsilon r}{1-\epsilon} + \frac{(1+\epsilon)\delta r}{1-\epsilon})$, of MR-GSEMO, which can approximately reach $1/(k + \frac{1}{p} + \frac{2\epsilon r}{1-\epsilon})$ by setting δ sufficiently small. Note that $p \geq 1$. By comparing it with the previously known guarantee, we have

Remark 1. The approximation guarantee, i.e., $1/(k + \frac{1}{p} + \frac{2\epsilon r}{1-\epsilon})$, of MR-GSEMO is better than that, i.e., $1/(k + 1 + \frac{4\epsilon r}{1-\epsilon})$, of the greedy algorithm [15].

Similarly, we can prove that MR-EMO using GSEMO-SR (briefly called MR-GSEMO-SR) can achieve the same approximation guarantee as MR-GSEMO after running $O\left(\frac{r}{\delta \cdot p_m} n^{(k+1)p} \log \frac{(1+\epsilon)r}{1-\epsilon}\right)$ expected number of iterations, as shown in Theorem 2. Different from GSEMO which always performs bit-wise mutation in each iteration, GSEMO-SR performs bit-wise mutation with probability p_m ; otherwise, performs matrix-swap mutation. Furthermore, if the generated

offspring solution is infeasible, GSEMO-SR will repair it to be feasible.

Theorem 2. *MR-GSEMO-SR using $O\left(\frac{r}{\delta \cdot p_m} n^{(k+1)p} \log \frac{(1+\epsilon)r}{1-\epsilon}\right)$ expected number of iterations finds a solution $X \in \bigcap_{i=1}^k \mathcal{F}_i$ with*

$$\left(k + \frac{1}{p} + \frac{2\epsilon r}{1-\epsilon} + \frac{(1+\epsilon)\delta r}{1-\epsilon}\right) \cdot f(X) \geq \text{OPT},$$

where $k \geq 2$, $p \geq 1$, $\epsilon \geq 0$ and $\delta > 0$.

Proof. The proof can be finished by directly following that of Theorem 1. The only difference is that the probability of applying bit-wise mutation here is $p_m < 1$ rather than 1, and thus the expected number of iterations of phases (1) and (2) is at most $e(r+1)n/p_m$ and $O\left(\frac{r}{\delta \cdot p_m} n^{(k+1)p} \log \frac{(1+\epsilon)r}{1-\epsilon}\right)$, respectively. Note that the offspring solution \mathbf{x}' considered in the proof of Theorem 1 is always feasible; thus, the repair mechanism employed by GSEMO-SR will not affect the analysis. \square

If the GSEMO or GSEMO-SR starts from a random solution, we can first derive the expected number of iterations until finding the all-0s solution $\mathbf{0}$, and then follow the current proof procedure of Theorem 1. For the additional analysis about the expected number of iterations until finding the all-0s solution, Lemma 1 in [35] has already shown that for maximizing any bi-objective pseudo-Boolean problem with $|\mathbf{x}|_0$ being one objective, the GSEMO (thus also GSEMO-SR) needs at most $O(rn \log n)$ expected number of iterations. The proof considers the solution with the maximum number of 0-bits in the population P , and then analyzes the expected number of iterations needed to increase the number of 0-bits of this solution, until generating the all-0s solution $\mathbf{0}$.

When using NSGA-II or MOEA/D, the theoretical analysis will be more challenging. For example, the analysis of GSEMO and GSEMO-SR relies on that if a solution is deleted, there must exist a solution weakly dominating it in the population; however, NSGA-II employs non-dominated sorting and crowding distance to maintain a fixed-size population, which may delete a non-dominated solution. As GSEMO uses mutation only, we also need to consider the influence of crossover for NSGA-II and MOEA/D. Note that there is some recent progress on the theoretical analysis of NSGA-II [3], [11], [12], [45], [46] and MOEA/D [19], [20], [25]. Though these works mainly focus on artificial problems, their analysis techniques might be useful for analyzing the approximation guarantee of MR-EMO using NSGA-II or MOEA/D, which will be an interesting future work.

V. EMPIRICAL STUDY

In this section, we empirically examine the performance of MR-EMO on two migration models (i.e., interview and coordination), by comparing its four variants (i.e., MR-NSGA-II, MR-MOEA/D, MR-GSEMO and MR-GSEMO-SR) with two previous algorithms, i.e., the additive [2] and greedy [15] algorithms. Note that MR-NSGA-II, MR-MOEA/D, MR-GSEMO and MR-GSEMO-SR correspond to MR-EMO using NSGA-II, MOEA/D, GSEMO and GSEMO-SR, respectively.

Table I

EXPECTED NUMBER OF EMPLOYED MIGRANTS (MEAN \pm STD) OBTAINED BY THE ALGORITHMS WHEN THE NUMBER OF MIGRANTS $|V| \in \{100, 120, \dots, 200\}$. FOR EACH $|V|$, THE LARGEST NUMBER IS BOLDED, AND ‘•’ DENOTES THAT MR-GSEMO-SR IS SIGNIFICANTLY BETTER THAN THE CORRESPONDING ALGORITHM BY THE WILCOXON SIGNED-RANK TEST WITH CONFIDENCE LEVEL 0.05. AVG.R. DENOTES THE AVERAGE RANK (THE SMALLER, THE BETTER) OF EACH ALGORITHM UNDER EACH SETTING AS IN [8].

$ V $	Interview Model						Avg.R.
	100	120	140	160	180	200	
Additive	66.76 \pm 3.10•	82.70 \pm 3.21•	99.83 \pm 2.65•	114.15 \pm 5.62•	135.09 \pm 3.87•	153.29 \pm 4.59•	6
Greedy	75.15 \pm 1.38•	93.65 \pm 1.25•	112.36 \pm 0.94•	130.84 \pm 1.65•	149.94 \pm 1.48•	168.94 \pm 1.87•	5
MR-NSGA-II	76.13 \pm 1.20•	94.84 \pm 1.27•	113.38 \pm 1.22•	131.96 \pm 1.50•	150.96 \pm 1.69•	169.82 \pm 1.96•	4
MR-MOEA/D	76.83 \pm 1.48•	95.47 \pm 1.36•	114.04 \pm 1.37•	132.27 \pm 1.53•	151.11 \pm 1.76•	170.16 \pm 1.89•	2.5
MR-GSEMO	76.77 \pm 1.40•	95.27 \pm 1.38•	113.95 \pm 1.12•	132.56 \pm 1.60•	151.25 \pm 1.74•	170.18 \pm 1.80•	2.5
MR-GSEMO-SR	77.95\pm1.49	96.99\pm1.40	115.76\pm1.33	134.23\pm1.53	153.29\pm1.71	172.23\pm1.87	1
Coordination Model							
Additive	68.91 \pm 2.41•	85.55 \pm 3.65•	103.68 \pm 5.30•	117.29 \pm 3.42•	136.53 \pm 4.38•	155.81 \pm 5.36•	6
Greedy	80.47 \pm 1.46•	100.40 \pm 1.39•	120.62 \pm 1.16•	140.53 \pm 1.81•	160.92 \pm 1.42•	181.15 \pm 2.05•	5
MR-NSGA-II	81.73 \pm 1.42•	101.82 \pm 1.32•	121.90 \pm 1.21•	141.69 \pm 1.77•	162.19 \pm 1.36•	182.25 \pm 2.14•	4
MR-MOEA/D	83.11 \pm 1.74•	103.78 \pm 1.61•	123.99 \pm 1.30•	143.68 \pm 1.44•	164.15 \pm 1.68•	184.09 \pm 2.06•	2.1
MR-GSEMO	82.87 \pm 1.59•	103.57 \pm 1.70•	123.64 \pm 1.11•	143.60 \pm 1.62•	164.15 \pm 1.55•	183.89 \pm 2.05•	2.9
MR-GSEMO-SR	83.76\pm1.73	104.76\pm1.89	124.79\pm1.31	144.66\pm1.53	164.98\pm1.61	184.92\pm2.09	1

For MR-NSGA-II and MR-MOEA/D, their population size is set to 100 and the initial population consists of the all-0 vector $\mathbf{0}$ and 99 randomly generated solutions. They apply one-point crossover and bit-wise mutation in each iteration with probability 0.9 and 1, respectively. We incorporate an objective normalization into MOEA/D, and perform the Tchebycheff decomposition approach to decompose the bi-objective reformulated migration resettlement problem in Eq. (7) into several scalar optimization subproblems. During the evolutionary process of MR-NSGA-II and MR-MOEA/D, the generated infeasible offspring solutions violating matroid constraints are discarded directly. MR-GSEMO starts from the initial solution $\mathbf{0}$. For MR-GSEMO-SR, the probability p_m of MR-GSEMO-SR applying bit-wise mutation is set to 0.5. These four variants of MR-EMO are all anytime algorithms, whose performance will be gradually improved by increasing the number of iterations. To make a tradeoff between performance and runtime, we set the number of objective evaluations to $100 \cdot |V|^2 \cdot |L|$, while the running time complexity of the greedy algorithm [15] is $|V|^2 \cdot |L|$.

The experimental data is generated as in [15]. For the interview model, the probability $p_{v,l}$ of each migrant $v \in V$ to find employment at each locality $l \in L$ is randomly sampled from $[0, 1]$. For the coordination model, each migrant $v \in V$ has probability $p_{v,\pi}$, randomly sampled from $[0, 1]$, to compete for a job of her profession $\pi \in \Pi$, and is incompatible with the jobs of other professions. We will compare the algorithms by varying the number of migrants, localities, jobs and professions, respectively. For each setting, we generate 10 problem instances randomly, and report the average objective value and the standard deviation, followed by the Wilcoxon signed-rank test [42] with confidence level 0.05. Note that to estimate the objective value, i.e., Eqs. (1) and (2), of a solution, we use the average of 1,000 random simulation (which corresponds to the approximate objective \hat{f}) in the run of each algorithm, while for the final output solution, we report its accurately estimated f value for the assessment of the algorithms by an expensive evaluation, i.e., the average of 10,000 random simulation.

In the following, we will introduce the detailed setting

(e.g., how to assign each migrant’s profession, determine the profession of each job, distribute jobs to each locality, and set the capacity of each locality) of varying the number of migrants $|V|$, localities $|L|$, jobs $|J|$ and professions $|\Pi|$, respectively, followed by the analysis of the experimental results.

Varying the Number $|V|$ of Migrants. We vary $|V| \in \{100, 120, \dots, 200\}$, and fix the number of localities $|L| = 10$, the number of jobs $|J| = |V|$ and the number of professions $|\Pi| = 2$. The two professions are evenly distributed to the migrants as well as the jobs, i.e., half of the migrants (or jobs) have profession π_1 , and the other half have profession π_2 . The jobs are randomly assigned to the localities such that each locality has the same number of jobs, i.e., $|J|/|L|$. The capacity of each locality is set to its number of jobs.

Varying the Number $|L|$ of Localities. We vary $|L| \in \{12, 14, \dots, 30\}$, and fix $|V| = |J| = 100$ and $|\Pi| = 2$. The other setting is the same as that of varying $|V|$, except that the jobs here are randomly assigned to the localities such that each locality has at least one job.

Varying the Number $|J|$ of Jobs. We fix $|V| = 100$, $|L| = 10$ and $|\Pi| = 2$. To simulate different levels of supply and demand, we vary the number of jobs in $\{10, 20, \dots, 90\}$ for one profession and fix 50 jobs for the other profession. The professions of migrants are evenly distributed, the jobs are randomly assigned to the localities, and the capacity of each locality is 10.

Varying the Number $|\Pi|$ of Professions. We fix $|V| = |J| = 100$, $|L| = 10$ and vary $|\Pi| \in \{2, 3, 5, 8, 10, 15, \dots, 30\}$. The professions of migrants are assigned randomly, ensuring that each profession has at least one migrant. And there is a job of the right profession for every migrant, that is, the number of jobs of each profession is exactly the same as the number of migrants of the profession. Then, we randomly distribute 10 jobs to each locality and the capacity of each locality is fixed to 10.

Result Comparison. The detailed results of varying the number of migrants $|V|$, localities $|L|$, jobs $|J|$ and professions $|\Pi|$ are shown in Tables I to IV, respectively. We can observe that the greedy algorithm is always better than the additive

Table II

EXPECTED NUMBER OF EMPLOYED MIGRANTS (MEAN \pm STD) OBTAINED BY THE ALGORITHMS WHEN THE NUMBER OF LOCALITIES $|L| \in \{12, 14, \dots, 30\}$. FOR EACH $|L|$, THE LARGEST NUMBER IS BOLDED, AND ‘•’ DENOTES THAT MR-GSEMO-SR IS SIGNIFICANTLY BETTER THAN THE CORRESPONDING ALGORITHM BY THE WILCOXON SIGNED-RANK TEST WITH CONFIDENCE LEVEL 0.05. AVG.R. DENOTES THE AVERAGE RANK (THE SMALLER, THE BETTER) OF EACH ALGORITHM UNDER EACH SETTING AS IN [8].

Interview Model											
$ L $	12	14	16	18	20	22	24	26	28	30	Avg.R.
Additive	65.02 \pm 2.14•	63.70 \pm 2.03•	63.61 \pm 2.58•	62.80 \pm 2.00•	60.81 \pm 2.62•	61.75 \pm 3.22•	60.39 \pm 2.63•	61.17 \pm 2.57•	58.91 \pm 1.77•	58.87 \pm 1.97•	6
Greedy	74.45 \pm 2.16•	72.10 \pm 1.08•	70.97 \pm 2.31•	69.59 \pm 2.26•	68.28 \pm 2.03•	66.82 \pm 1.97•	65.03 \pm 1.37•	64.70 \pm 1.40•	63.45 \pm 1.49•	61.88 \pm 1.59•	5
MR-NSGA-II	76.28 \pm 2.07•	73.78 \pm 0.86•	73.00 \pm 2.38•	71.25 \pm 2.35•	70.23 \pm 2.23•	68.84 \pm 2.08•	66.84 \pm 1.24•	66.90 \pm 1.62•	65.62 \pm 1.57•	64.01 \pm 1.39•	4
MR-MOEA/D	77.31 \pm 2.00•	74.95 \pm 1.03•	74.51 \pm 2.35•	72.84 \pm 2.38•	71.76 \pm 1.95•	70.26 \pm 2.15•	68.60 \pm 1.44•	68.23 \pm 1.62•	67.61 \pm 1.62•	65.31 \pm 1.30•	2
MR-GSEMO	77.19 \pm 2.03•	74.55 \pm 1.05•	73.87 \pm 2.36•	72.38 \pm 2.33•	71.05 \pm 2.07•	69.70 \pm 2.17•	67.91 \pm 1.60•	67.74 \pm 1.53•	66.47 \pm 1.47•	64.84 \pm 1.32•	3
MR-GSEMO-SR	78.74\pm2.12	76.31\pm0.91	75.83\pm2.45	74.17\pm2.49	72.96\pm2.01	71.40\pm2.12	69.54\pm1.65	69.36\pm1.63	68.07\pm1.64	66.25\pm1.39	1
Coordination Model											
Additive	68.79 \pm 1.86•	65.91 \pm 2.70•	66.96 \pm 2.99•	65.57 \pm 2.48•	64.03 \pm 2.74•	63.97 \pm 2.93•	63.96 \pm 1.93•	63.12 \pm 1.83•	61.86 \pm 2.10•	61.32 \pm 2.24•	6
Greedy	79.36 \pm 2.46•	77.04 \pm 1.24•	75.40 \pm 2.37•	73.94 \pm 2.37•	72.35 \pm 2.16•	70.57 \pm 2.24•	68.75 \pm 1.36•	68.32 \pm 1.51•	66.96 \pm 1.51•	65.21 \pm 1.77•	5
MR-NSGA-II	80.90 \pm 2.32•	78.48 \pm 1.35•	77.16 \pm 2.21•	75.57 \pm 2.28•	74.13 \pm 1.98•	72.52 \pm 2.08•	70.62 \pm 1.51•	70.07 \pm 1.35•	68.69 \pm 1.75•	67.07 \pm 1.45•	4
MR-MOEA/D	83.59 \pm 2.36•	80.61 \pm 1.03•	79.52 \pm 2.46•	77.54 \pm 2.61•	75.98 \pm 2.38•	74.19 \pm 2.24•	72.16 \pm 1.75•	71.69 \pm 1.73•	70.62 \pm 1.81•	68.68 \pm 1.58•	2
MR-GSEMO	82.89 \pm 2.33•	79.89 \pm 1.18•	79.20 \pm 2.33•	77.23 \pm 2.56•	75.75 \pm 2.25•	73.90 \pm 2.38•	71.79 \pm 1.89•	71.09 \pm 1.52•	69.69 \pm 1.87•	67.86 \pm 1.49•	3
MR-GSEMO-SR	84.82\pm2.43	81.75\pm1.10	81.19\pm2.80	79.15\pm2.78	77.60\pm2.38	75.70\pm2.45	73.55\pm1.96	73.10\pm1.76	71.60\pm1.89	69.47\pm1.60	1

Table III

EXPECTED NUMBER OF EMPLOYED MIGRANTS (MEAN \pm STD) OBTAINED BY THE ALGORITHMS WHEN THE NUMBER OF JOBS $|J| \in \{60, 70, \dots, 140\}$. FOR EACH $|J|$, THE LARGEST NUMBER IS BOLDED, AND ‘•’ DENOTES THAT MR-GSEMO-SR IS SIGNIFICANTLY BETTER THAN THE CORRESPONDING ALGORITHM BY THE WILCOXON SIGNED-RANK TEST WITH CONFIDENCE LEVEL 0.05. AVG.R. DENOTES THE AVERAGE RANK (THE SMALLER, THE BETTER) OF EACH ALGORITHM UNDER EACH SETTING AS IN [8].

Interview Model											
$ J $	60	70	80	90	100	110	120	130	140	Avg.R.	
Additive	36.75 \pm 4.24•	41.81 \pm 2.47•	50.25 \pm 2.61•	56.46 \pm 3.64•	61.57 \pm 5.10•	66.15 \pm 2.47•	70.95 \pm 3.05•	73.46 \pm 2.44•	75.28 \pm 3.23•	6	
Greedy	47.98 \pm 1.22•	56.69 \pm 1.74•	64.88 \pm 1.82•	70.25 \pm 1.75•	74.61 \pm 2.45•	76.96 \pm 1.60•	78.36 \pm 2.46•	79.81 \pm 2.16•	80.42 \pm 1.61•	5	
MR-NSGA-II	49.18 \pm 1.21•	57.95 \pm 1.40•	66.12 \pm 1.67•	71.39 \pm 1.58•	75.58 \pm 2.29•	77.94 \pm 1.59•	79.99 \pm 2.21•	81.05 \pm 1.72•	81.70 \pm 1.45•	4	
MR-MOEA/D	49.42 \pm 1.17•	58.34 \pm 1.35•	66.78 \pm 1.46•	72.51 \pm 1.63•	76.68 \pm 2.22•	79.18 \pm 1.51•	81.35 \pm 2.28•	82.59 \pm 1.93•	83.43 \pm 1.47•	2	
MR-GSEMO	49.32 \pm 1.16•	58.25 \pm 1.43•	66.70 \pm 1.55•	72.11 \pm 1.74•	76.35 \pm 2.35•	78.70 \pm 1.58•	80.63 \pm 2.43•	82.00 \pm 1.70•	82.43 \pm 1.58•	3	
MR-GSEMO-SR	50.03\pm1.19	59.40\pm1.43	68.23\pm1.42	73.84\pm1.65	78.27\pm2.20	81.00\pm1.58	83.14\pm2.09	84.33\pm1.80	85.46\pm1.73	1	
Coordination Model											
Additive	37.99 \pm 4.17•	43.41 \pm 2.43•	52.09 \pm 3.91•	58.80 \pm 4.01•	64.68 \pm 4.52•	69.50 \pm 2.80•	74.13 \pm 4.12•	77.10 \pm 2.69•	79.10 \pm 3.64•	6	
Greedy	50.38 \pm 1.76•	59.29 \pm 2.13•	67.99 \pm 1.99•	74.25 \pm 1.65•	78.56 \pm 2.80•	79.85 \pm 1.61•	81.38 \pm 2.70•	81.94 \pm 2.59•	81.84 \pm 2.30•	5	
MR-NSGA-II	51.95 \pm 1.63•	60.86 \pm 1.76•	69.45 \pm 1.96•	75.43 \pm 1.93•	80.04 \pm 2.86•	81.65 \pm 1.66•	83.37 \pm 2.35•	84.31 \pm 1.87•	84.31 \pm 1.79•	4	
MR-MOEA/D	52.71 \pm 1.36•	62.08 \pm 1.73•	71.20 \pm 1.75•	77.87 \pm 2.04•	82.47 \pm 2.55•	84.56 \pm 1.76•	86.38 \pm 2.30•	87.21 \pm 2.25•	88.30 \pm 1.71•	2	
MR-GSEMO	52.27 \pm 1.49•	61.58 \pm 1.59•	70.49 \pm 1.88•	77.31 \pm 1.73•	81.56 \pm 2.57•	83.68 \pm 1.63•	85.05 \pm 2.50•	86.00 \pm 2.09•	86.88 \pm 1.75•	3	
MR-GSEMO-SR	53.29\pm1.48	62.83\pm1.85	72.33\pm1.78	79.33\pm1.89	83.85\pm2.63	86.23\pm1.71	88.11\pm2.19	88.70\pm1.97	89.60\pm1.71	1	

Table IV

EXPECTED NUMBER OF EMPLOYED MIGRANTS (MEAN \pm STD) OBTAINED BY THE ALGORITHMS WHEN THE NUMBER OF PROFESSIONS $|\Pi| \in \{2, 3, 5, 8, \dots, 30\}$. FOR EACH $|\Pi|$, THE LARGEST NUMBER IS BOLDED, AND ‘•’ DENOTES THAT MR-GSEMO-SR IS SIGNIFICANTLY BETTER THAN THE CORRESPONDING ALGORITHM BY THE WILCOXON SIGNED-RANK TEST WITH CONFIDENCE LEVEL 0.05. AVG.R. DENOTES THE AVERAGE RANK (THE SMALLER, THE BETTER) OF EACH ALGORITHM UNDER EACH SETTING AS IN [8].

Interview Model											
$ \Pi $	2	3	5	8	10	15	20	25	30	Avg.R.	
Additive	68.40 \pm 1.92•	57.78 \pm 2.79•	49.99 \pm 3.94•	43.35 \pm 2.27•	42.81 \pm 2.74•	42.82 \pm 2.65•	42.32 \pm 2.10•	43.17 \pm 2.68•	43.75 \pm 1.84•	6	
Greedy	76.35 \pm 1.09•	70.75 \pm 1.94•	65.97 \pm 2.24•	59.92 \pm 2.74•	58.59 \pm 1.74•	55.71 \pm 1.30•	55.39 \pm 2.17•	54.48 \pm 2.10•	52.58 \pm 1.46•	5	
MR-NSGA-II	77.00 \pm 1.07•	72.32 \pm 1.88•	67.55 \pm 2.02•	61.74 \pm 2.98•	60.37 \pm 1.84•	56.95 \pm 1.47•	56.42 \pm 2.16•	55.20 \pm 1.98•	53.14 \pm 1.50•	4	
MR-MOEA/D	77.75 \pm 1.30•	73.40 \pm 2.08•	68.69 \pm 2.14•	62.74 \pm 3.25•	61.14 \pm 1.96•	57.59 \pm 1.66•	56.82 \pm 2.06•	55.67 \pm 2.11•	53.16 \pm 1.47•	2.1	
MR-GSEMO	77.69 \pm 1.22•	73.23 \pm 2.06•	68.50 \pm 2.33•	62.58 \pm 3.01•	60.98 \pm 1.93•	57.55 \pm 1.54•	56.64 \pm 2.05•	55.59 \pm 2.09•	53.31 \pm 1.48•	2.9	
MR-GSEMO-SR	78.97\pm1.19	74.77\pm2.11	69.92\pm2.09	63.73\pm3.40	62.02\pm1.87	58.29\pm1.64	57.34\pm2.01	56.01\pm2.14	53.58\pm1.48	1	
Coordination Model											
Additive	71.43 \pm 3.08•	59.24 \pm 2.73•	50.77 \pm 4.27•	45.33 \pm 2.74•	44.47 \pm 2.28•	43.85 \pm 2.73•	42.93 \pm 2.33•	42.89 \pm 2.89•	44.15 \pm 1.73•	6	
Greedy	81.82 \pm 1.24•	75.28 \pm 2.00•	69.80 \pm 2.50•	62.78 \pm 2.95•	61.28 \pm 1.68•	57.65 \pm 1.41•	56.84 \pm 2.22•	55.84 \pm 2.08•	53.37 \pm 1.44•	5	
MR-NSGA-II	82.80 \pm 1.10•	76.69 \pm 1.86•	71.60 \pm 2.12•	64.11 \pm 2.88•	62.53 \pm 1.74•	58.44 \pm 1.48•	57.43 \pm 2.16•	56.18 \pm 2.16•	53.54 \pm 1.58•	4	
MR-MOEA/D	84.35 \pm 1.44•	78.57 \pm 2.22•	72.38 \pm 2.39•	65.04 \pm 3.43•	63.16 \pm 2.00•	58.89 \pm 1.73•	57.73 \pm 2.15•	56.46 \pm 2.18•	53.77 \pm 1.58•	2.3	
MR-GSEMO	84.15 \pm 1.19•	78.52 \pm 2.42•	72.27 \pm 2.53•	64.71 \pm 3.49•	63.15 \pm 2.04•	59.01 \pm 1.56•	57.73 \pm 2.19•	56.45 \pm 2.24•	53.83 \pm 1.54•	2.7	
MR-GSEMO-SR	85.14\pm1.31	79.89\pm2.46	73.75\pm2.44	66.14\pm3.79	64.13\pm2.02	59.61\pm1.80	58.24\pm2.05	56.76\pm2.22	54.04\pm1.64	1	

algorithm, which is consistent with the observation in [15]; while the four variants of MR-EMO (i.e., MR-NSGA-II, MR-MOEA/D, MR-GSEMO and MR-GSEMO-SR) all surpass the greedy and additive algorithms, exhibiting the superiority of the MR-EMO framework. This may be because MR-EMO naturally maintains a population of diverse solutions due to

the bi-objective transformation, and the employed bit-wise mutation operator has a good global search ability. These characteristics can lead to a better ability to escape from local optima.

Among the four variants of MR-EMO, MR-NSGA-II performs the worst, which may be because the population of

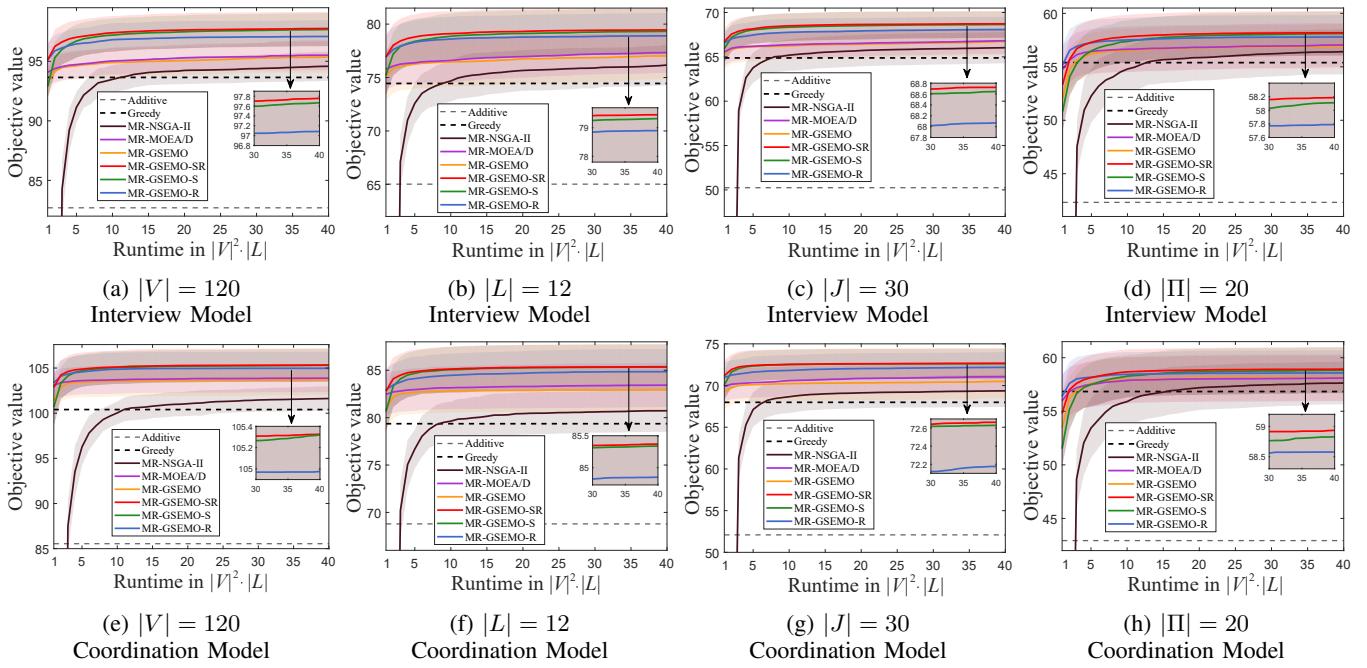


Figure 6. One each migration model, the objective value vs. runtime (i.e., number of objective evaluations) with $|V| = 120$, $|L| = 12$, $|J| = 30$ and $|\Pi| = 20$, respectively.

NSGA-II may contain redundant dominated solutions, leading to the bad performance; MR-MOEA/D performs slightly better than MR-GSEMO; MR-GSEMO-SR always achieves the best average objective value (i.e., the largest expected number of employed migrants), and is significantly better than all the other algorithms by the Wilcoxon signed-rank test [42] with confidence level 0.05. These observations show the effectiveness of the introduced matrix-swap mutation and repair mechanism for the migrant resettlement problem.

Tables I to IV also show that the expected number of employed migrants achieved by each algorithm increases with the number $|V|$ of migrants and the number $|J|$ of jobs, but decreases with the number $|L|$ of localities and the number $|\Pi|$ of professions. More migrants and jobs will bring more employment naturally, while more localities and professions with a fixed number of migrants and the number of jobs will increase the difficulty of resettlement and thus lead to less employment.

Ablation Study. Compared with MR-GSEMO using bit-wise mutation only, MR-GSEMO-SR employs matrix-swap mutation and repair mechanism additionally. To examine the utility of these two introduced components more clearly, we run another two variants MR-GSEMO-S and MR-GSEMO-R, which apply only matrix-swap mutation and repair mechanism, respectively. We plot the curve of objective value over runtime under two migration models with $|V| = 120$, $|L| = 12$, $|J| = 30$ and $|\Pi| = 20$, respectively, as shown in Figure 6. The additive and greedy algorithms are fixed-time algorithms, while the variants of MR-EMO are anytime algorithms, and can get better performance by using more time. MR-NSGA-II exceeds the greedy and additive algorithms by using about $15 \cdot |V|^2 \cdot |L|$ time, and other variants of MR-EMO (i.e., MR-MOEA/D, MR-GSEMO, MR-GSEMO-

SR, MR-GSEMO-S and MR-GSEMO-R) surpass them within $5 \cdot |V|^2 \cdot |L|$ time. We can observe that even applying only one component (i.e., matrix-swap mutation or repair mechanism), both MR-GSEMO-S and MR-GSEMO-R are still better than MR-NSGA-II, MR-MOEA/D and MR-GSEMO, showing the effectiveness of matrix-swap mutation and repair mechanism clearly. Compared with one-point crossover and bit-wise mutation, the matrix-swap mutation operator has a larger probability of generating feasible solutions, and thus can make feasible regions explored efficiently. Instead of discarding infeasible solutions directly, repairing them may lead to good feasible solutions. Meanwhile, MR-GSEMO-S is better than MR-GSEMO-R, disclosing the more important role of the matrix-swap mutation operator. Finally, MR-GSEMO-SR, which employs both matrix-swap mutation and repair mechanism, performs the best, implying that using the two components together can further improve the performance.

VI. CONCLUSION

In this paper, we propose the general framework MR-EMO for the important migrant resettlement problem, which tries to maximize the expected number of employed migrants while dispatching as few migrants as possible. MR-EMO can be equipped with any MOEA to solve the reformulated bi-objective problem, and we apply NSGA-II, MOEA/D and GSEMO, and also propose the specific MOEA, GSEMO-SR, by introducing matrix-swap mutation and repair mechanism. We prove that using either GSEMO or GSEMO-SR, MR-EMO can achieve the approximation guarantee of $1/(k + \frac{1}{p} + \frac{2\epsilon r}{1-\epsilon})$, surpassing the best-known one [15], obtained by the greedy algorithm. It is more challenging to analyze the theoretical guarantees of MR-NSGA-II and MR-MOEA/D, which will be an interesting future work.

By experiments under the two migration models, i.e., interview and coordination models, we show that MR-EMO using any of the four MOEAs can produce better employment than the previous (i.e., additive and greedy) algorithms, and using GSEMO-SR leads to the best performance, which is significantly better than all the other algorithms. The matrix-swap mutation operator and repair mechanism can also be incorporated into other MOEAs besides GSEMO. Thus, it is interesting to examine whether they can still bring performance improvement. In the future, it is also worth trying other advanced MOEAs or designing better ones when applying the proposed framework MR-EMO to solve the migration resettlement problem.

REFERENCES

- [1] T. Bäck, *Evolutionary Algorithms in Theory and Practice: Evolution Strategies, Evolutionary Programming, Genetic Algorithms*. Oxford, UK: Oxford University Press, 1996.
- [2] K. Bansak, J. Ferwerda, J. Hainmueller, A. Dillon, D. Hangartner, D. Lawrence, and J. Weinstein, “Improving refugee integration through data-driven algorithmic assignment,” *Science*, vol. 359, no. 6373, pp. 325–329, 2018.
- [3] C. Bian and C. Qian, “Better running time of the non-dominated sorting genetic algorithm II (NSGA-II) by using stochastic tournament selection,” in *Proceedings of the 17th International Conference on Parallel Problem Solving from Nature (PPSN)*, Dortmund, Germany, 2022, pp. 428–441.
- [4] G. Calinescu, C. Chekuri, M. Pál, and J. Vondrák, “Maximizing a monotone submodular function subject to a matroid constraint,” *SIAM Journal on Computing*, vol. 40, no. 6, pp. 1740–1766, 2011.
- [5] C. A. C. Coello, G. B. Lamont, D. A. Van Veldhuizen *et al.*, *Evolutionary Algorithms for Solving Multi-Objective Problems*. New York, NY: Springer, 2007.
- [6] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, “A fast and elitist multiobjective genetic algorithm: NSGA-II,” *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 2, pp. 182–197, 2002.
- [7] D. Delacrétaz, S. D. Kominers, and A. Teytelboym, “Refugee resettlement,” Department of Economics, University of Oxford, Oxford, UK, Tech. Rep., 2016.
- [8] J. Demsar, “Statistical comparisons of classifiers over multiple data sets,” *Journal of Machine Learning Research*, vol. 7, pp. 1–30, 2006.
- [9] F. Docquier, Ç. Ozden, and G. Peri, “The labour market effects of immigration and emigration in oecd countries,” *The Economic Journal*, vol. 124, no. 579, pp. 1106–1145, 2014.
- [10] B. Doerr and F. Neumann, *Theory of Evolutionary Computation: Recent Developments in Discrete Optimization*. Switzerland: Springer, 2020.
- [11] B. Doerr and Z. Qu, “The first mathematical proof that crossover gives super-constant performance gains for the NSGA-II,” *arXiv preprint arXiv:2208.08759*, 2022.
- [12] B. Doerr and Z. Qu, “A first runtime analysis of the NSGA-II on a multimodal problem,” in *Proceedings of the 17th International Conference on Parallel Problem Solving from Nature (PPSN)*, Dortmund, Germany, 2022, pp. 399–412.
- [13] T. Friedrich and F. Neumann, “Maximizing submodular functions under matroid constraints by evolutionary algorithms,” *Evolutionary Computation*, vol. 23, no. 4, pp. 543–558, 2015.
- [14] O. Giel, “Expected runtimes of a simple multi-objective evolutionary algorithm,” in *Proceedings of the 2003 IEEE Congress on Evolutionary Computation (CEC)*, Newport Beach, CA, 2003, pp. 1918–1925.
- [15] P. Götzl and A. D. Procaccia, “Migration as submodular optimization,” in *Proceedings of the 33rd Conference on Artificial Intelligence (AAAI)*, Honolulu, HI, 2019, pp. 549–556.
- [16] L. He, H. Ishibuchi, A. Trivedi, H. Wang, Y. Nan, and D. Srinivasan, “A survey of normalization methods in multiobjective evolutionary algorithms,” *IEEE Transactions on Evolutionary Computation*, vol. 25, no. 6, pp. 1028–1048, 2021.
- [17] W.-J. Hong, P. Yang, and K. Tang, “Evolutionary computation for large-scale multi-objective optimization: A decade of progresses,” *International Journal of Automation and Computing*, vol. 18, no. 2, pp. 155–169, 2021.
- [18] T. Horel and Y. Singer, “Maximization of approximately submodular functions,” in *Advances in Neural Information Processing Systems 29 (NeurIPS)*, Barcelona, Spain, 2016, pp. 3045–3053.
- [19] Z. Huang, Y. Zhou, Z. Chen, and X. He, “Running time analysis of MOEA/D with crossover on discrete optimization problem,” in *Proceedings of the 33rd AAAI Conference on Artificial Intelligence (AAAI)*, Honolulu, HI, 2019, pp. 2296–2303.
- [20] Z. Huang, Y. Zhou, Z. Chen, X. He, X. Lai, and X. Xia, “Running time analysis of MOEA/D on pseudo-Boolean functions,” *IEEE Transactions on Cybernetics*, vol. 51, no. 10, pp. 5130–5141, 2021.
- [21] IOM, *World Migration Report 2020*. United Nations, 2019.
- [22] J. D. Knowles, R. A. Watson, and D. W. Corne, “Reducing local optima in single-objective problems by multi-objectivization,” in *Proceedings of the 1st International Conference on Evolutionary Multi-Criterion Optimization (EMO)*, Zurich, Switzerland, 2001, pp. 269–283.
- [23] M. Laumanns, L. Thiele, and E. Zitzler, “Running time analysis of multiobjective evolutionary algorithms on pseudo-Boolean functions,” *IEEE Transactions on Evolutionary Computation*, vol. 8, no. 2, pp. 170–182, 2004.
- [24] J. Lee, M. Sviridenko, and J. Vondrák, “Submodular maximization over multiple matroids via generalized exchange properties,” *Mathematics of Operations Research*, vol. 35, no. 4, pp. 795–806, 2010.
- [25] Y.-L. Li, Y. Zhou, Z.-H. Zhan, and J. Zhang, “A primary theoretical study on decomposition-based multiobjective evolutionary algorithms,” *IEEE Transactions on Evolutionary Computation*, vol. 20, no. 4, pp. 563–576, 2016.
- [26] M. Marbach, J. Hainmueller, and D. Hangartner, “The long-term impact of employment bans on the economic integration of refugees,” *Science Advances*, vol. 4, no. 9, p. eaap9519, 2018.
- [27] G. L. Nemhauser, L. A. Wolsey, and M. L. Fisher, “An analysis of approximations for maximizing submodular set functions – I,” *Mathematical Programming*, vol. 14, no. 1, pp. 265–294, 1978.
- [28] A. Neumann and F. Neumann, “Optimising monotone chance-constrained submodular functions using evolutionary multi-objective algorithms,” in *Parallel Problem Solving from Nature XVI (PPSN)*, Leiden, The Netherlands, 2020, pp. 404–417.
- [29] F. Neumann and I. Wegener, “Minimum spanning trees made easier via multi-objective optimization,” *Natural Computing*, vol. 5, no. 3, pp. 305–319, 2006.
- [30] F. Neumann and C. Witt, *Bioinspired Computation in Combinatorial Optimization: Algorithms and Their Computational Complexity*. Berlin, Germany: Springer-Verlag, 2010.
- [31] K. Nurse, “Migration, diasporas, remittances and the sustainable development goals in least developed countries,” *Journal of Globalization and Development*, vol. 9, no. 2, p. 20190006, 2019.
- [32] G. I. P. Ottaviano, G. Peri, and G. C. Wright, “Immigration, offshoring, and american jobs,” *American Economic Review*, vol. 103, no. 5, pp. 1925–59, 2013.
- [33] C. Qian, Y. Yu, and Z.-H. Zhou, “Subset selection by Pareto optimization,” in *Advances in Neural Information Processing Systems 28 (NeurIPS)*, Montreal, Canada, 2015, pp. 1765–1773.
- [34] C. Qian, “Distributed Pareto optimization for large-scale noisy subset selection,” *IEEE Transactions on Evolutionary Computation*, vol. 24, no. 4, pp. 694–707, 2020.
- [35] C. Qian, “Multiobjective evolutionary algorithms are still good: Maximizing monotone approximately submodular minus modular functions,” *Evolutionary Computation*, vol. 29, no. 4, pp. 463–490, 2021.
- [36] C. Qian, J.-C. Shi, Y. Yu, K. Tang, and Z.-H. Zhou, “Subset selection under noise,” in *Advances in Neural Information Processing Systems 30 (NeurIPS)*, Long Beach, CA, 2017, pp. 3560–3570.
- [37] C. Qian, Y. Yu, K. Tang, X. Yao, and Z.-H. Zhou, “Maximizing submodular or monotone approximately submodular functions by multi-objective evolutionary algorithms,” *Artificial Intelligence*, vol. 275, pp. 279–294, 2019.
- [38] V. Roostapour, A. Neumann, F. Neumann, and T. Friedrich, “Pareto optimization for subset selection with dynamic cost constraints,” *Artificial Intelligence*, vol. 302, p. 103597, 2022.
- [39] M. Ruhs and C. Vargas-Silva, “The labour market effects of immigration,” *The Migration Observatory*, 2015.
- [40] UNHCR, *Global Trends: Forced Displacement in 2020*. Geneva, Switzerland: United Nations High Commissioner for Refugees, 2021.
- [41] L. Van der Maaten and G. Hinton, “Visualizing data using t-SNE,” *Journal of Machine Learning Research*, vol. 9, no. 11, 2008.
- [42] F. Wilcoxon, “Individual comparisons by ranking methods,” *Biometrics Bulletin*, vol. 1, no. 6, pp. 80–83, 1945.

- [43] G. Yu, L. Ma, Y. Jin, W. Du, Q. Liu, and H. Zhang, "A survey on knee-oriented multiobjective evolutionary optimization," *IEEE Transactions on Evolutionary Computation*, vol. 26, no. 6, pp. 1452–1472, 2022.
- [44] 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.
- [45] W. Zheng and B. Doerr, "Better approximation guarantees for the NSGA-II by using the current crowding distance," in *Proceedings of the 24th ACM Conference on Genetic and Evolutionary Computation (GECCO)*, Boston, MA, 2022, pp. 611–619.
- [46] W. Zheng, Y. Liu, and B. Doerr, "A first mathematical runtime analysis of the non-dominated sorting genetic algorithm II (NSGA-II)," in *Proceedings of the 36th AAAI Conference on Artificial Intelligence (AAAI)*, Virtual, 2022, pp. 10408–10416.
- [47] Z.-H. Zhou, Y. Yu, and C. Qian, *Evolutionary Learning: Advances in Theories and Algorithms*. Singapore: Springer, 2019.