显式 Explicit Evolutionary Multitasking for 组合优化 Combinatorial Optimization: A Case Study on 容量受限 Capacitated Vehicle Routing Problem

Liang Feng, Yuxiao Huang, Lei Zhou, Jinghui Zhong, Abhishek Gupta, Ke Tang, and Kay Chen Tan

***Abstract*—Recently, evolutionary multitasking (EMT) has been proposed in the field of evolutionary computation as a new search paradigm, for solving multiple optimization tasks simultaneously. By sharing useful traits found along the evolutionary search process across different optimization tasks, the optimization performance on each task could be enhanced. The autoencoding-based EMT is a recently proposed EMT algorithm. In contrast to most existing EMT algorithms, which conduct knowledge transfer across tasks *implicitly* via crossover, it intends to perform knowledge transfer *explicitly* among tasks in the form of task solutions, which enables the employment of task-specific search mechanisms for different optimization tasks in EMT. However, the autoencoding-based *explicit* EMT can only work on continuous optimization problems. It will fail on combinatorial optimization problems, which widely exist in real-world applications, such as scheduling problem, routing problem, and assignment problem. To the best of our knowledge, there is no existing effort working on *explicit* EMT for combinatorial optimization problems. Taking this cue, in this article, we thus embark on a study toward *explicit* EMT for combinatorial optimization. In particular, by using vehicle routing as an illustrative combinatorial optimization problem, the proposed *explicit* EMT algorithm (EEMTA) mainly contains a weighted l_1-norm-regularized learning process for capturing the transfer mapping, and a solution-based knowledge transfer process across vehicle routing problems (VRPs). To evaluate the efficacy of the proposed EEMTA, comprehensive empirical studies have been conducted with the commonly used vehicle routing benchmarks in multitasking environment, against both the state-of-the-art EMT algorithm and the traditional single-task evolutionary solvers. Finally, a real-world combinatorial optimization application, that is, the package delivery problem (PDP), is also presented to further confirm the efficacy of the proposed algorithm.**

***Index Terms*—Evolutionary optimization, knowledge transfer, multitask optimization, routing problem.**

Manuscript received July 17, 2019; revised November 14, 2019; accepted December 22, 2019. This work was supported in part by the National Natural Science Foundation of China under Grant 61876025 and Grant 61602181, in part by the Venture and Innovation Support Program for Chongqing Overseas Returnees under Grant cx2018044 and Grant cx2019020, in part by the City University of Hong Kong Research Fund under Grant 9610397, in part by the Research Grants Council of the Hong Kong Special Administrative Region, China under Grant CityU11202418 and Grant CityU11209219, in part by the Shenzhen Scientific Research and Development Funding Program under Grant JCYJ20180307123637294, in part by the Scientific and Technological Innovation Foundation of Chongqing under Grant cstc2019yszx-jscxX0001, and in part by the Fundamental Research Funds for the Central Universities under Grant D2191200. This article was recommended by Associate Editor L. Tang. *(Corresponding author: Liang Feng.)*

Liang Feng, Yuxiao Huang, and Lei Zhou are with the College of Computer Science, Chongqing University, Chongqing 400044, China (e-mail: liangf@cqu.edu.cn; yuxiaohuang@cqu.edu.cn; stone_zhou@cqu.edu.cn).

Jinghui Zhong is with the School of Computer Science and Engineering, South China University of Technology, Guangzhou 510006, China, and also with the Sino-Singapore International Joint Research Institute, Guangzhou, China (e-mail: jinghuizhong@gmail.com).

Abhishek Gupta is with the Singapore Institute of Manufacturing Technology, Agency of Science, Technology and Research, Singapore (e-mail: abhishek_gupta@simtech.a-star.edu.sg).

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

Kay Chen Tan is with the Department of Computer Science, City University of Hong Kong, Hong Kong, and also with the City University of Hong Kong Shenzhen Research Institute, Shenzhen, China (e-mail: kaytan@cityu.edu.hk).

Color versions of one or more of the figures in this article are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TCYB.2019.2962865

I. Introduction

EVOLUTIONARY multitasking (EMT) is a recently emerged search paradigm in the realm of evolutionary computation [1]. In contrast to the traditional evolutionary algorithm (EA), which solves a single task in a single run, the EMT conducts evolutionary search concurrently on multiple search spaces or optimization problems, each possessing a unique function landscape [2]. By leveraging the useful knowledge across tasks while the evolutionary search progresses online, the EMT has demonstrated superior optimization performance in terms of search efficiency and effectiveness on continuous as well as discrete optimization problems, when compared to its single-task counterparts [1]–[5]. Due to its strong search capability and parallelism nature to be applied in today's distributed computing platforms (e.g., cloud computing), EMT has attracted great attention in recent years [6]–[9].

In the literature, existing EMT algorithms can be generally classified into two classes, that is, implicit EMT and explicit EMT. The implicit EMT usually employs a single population with unified solution representation to solve multiple tasks, and the knowledge transfer across tasks is realized implicitly via the chromosomal crossover operator with two solutions possessing different *skill factors*. Particular examples include the single- and multiobjective multifactorial EA proposed by Gupta *et al.* [1], [10], respectively; the generalized EMT algorithm proposed by Ding *et al.* [3]; the linearized EMT algorithm proposed by Bali *et al.* [6]; and the multifactorial genetic programming proposed by Zhong *et al.* [7], etc. On the other hand, the explicit EMT considers multiple populations for multitask optimization, in which each population is for optimizing only one optimization task. In this paradigm, an additional operator has to be designed for conducting knowledge transfer explicitly among the populations of different tasks. In contrast to implicit EMT, the explicit EMT is able to incorporate multiple search mechanisms (e.g., crossover, local search, etc.) to solve different optimization tasks. As different search mechanisms possess various search biases, the employment of problem-specific search operators in explicit EMT could lead to significantly improved multitask optimization performance [11]. However, compared with the accomplishments made in the implicit EMT algorithm design, only few attempts have been conducted for developing the explicit EMT approaches.

The autoencoding-based EMT is a recently proposed algorithm for explicit EMT [11]. By incorporating problem-specific solvers and devising a single-layer denoising autoencoder for explicit knowledge transfer across tasks, the autoencoding-based EMT has demonstrated significantly superior performance in terms of solution quality and search efficiency, over the implicit multifactorial EA, on the commonly used single-objective and multiobjective multitask optimization problems. However, it is worth noting that as the derived denoising autoencoder requires that the input and output must be a continuous real number, the autoencoding-based EMT algorithm proposed in [11] can only deal with the continuous optimization problems. It will fail on combinatorial optimization problems which have solutions represented by discrete integers [12], [13]. Keeping this in mind, here we propose a new explicit EMT algorithm for combinatorial optimization problems. To the best of our knowledge, this is the first attempt in the literature for conducting explicit knowledge transfer across combinatorial optimization problem domains, toward EMT.

In particular, by considering the classical NP-hard VRP [14], [15] as the illustrating combinatorial optimization problem domain, we first propose building a sparse mapping between VRPs via minimizing a weighted l_1-norm-regularized reconstruction error from one VRP to another. In this way, the knowledge transfer could be performed across VRPs possessing diverse problem properties, such as customer topologies, number of customers, and number of vehicles. Next, we propose capturing the useful VRP knowledge by learning a new representation of customers, based on a distance matrix derived from the optimized VRP solution. With the learned sparse mapping between VRPs, the estimated new customer representation can be transferred to generate VRP solutions for the target VRP, via simple clustering and pairwise distance sorting processes. These generated solutions are then served as the useful knowledge transferred across VRPs, to guide the evolutionary search process for solving the target VRP. To confirm the efficacy of the proposed EMT algorithm for combinatorial optimization, comprehensive empirical studies are conducted on both commonly used VRP benchmarks and a real-world routing application, against the single-task evolutionary solver as well as an implicit EMT algorithm for solving VRP.

The remainder of this article is organized as follows. Section II begins with a brief literature reviews on recent research progress in EMT. The definition of the VRP studied in this article is also introduced in this section. Section III presents the details of the proposed explicit EMT algorithm for combinatorial optimization, particularly, VRP. Further, Sections IV and V provide the comprehensive empirical studies of the proposed algorithm against both the single-task counterpart solver and a recently proposed implicit EMT algorithm for solving VRP, on the commonly used VRP benchmarks and a real-world routing application, respectively. Finally, we draw the concluding remarks of this article in Section VI.

II. PRELIMINARY

In this section, we first present a brief literature review of the related works for EMT. Next, the problem definition of the VRP studied in this article is also introduced.

A. Related Work

In the literature, the concept of EMT was first proposed by Gupta *et al.* [1]. In this article, the authors also introduced an implicit EMT algorithm, namely, multifactorial EA (MFEA) and confirmed the efficacy of conducting EMT in contrast to the traditional single-task EA, on a set of continuous, discrete, and the mixtures of continuous and combinatorial tasks. After this, a number of EMT algorithms have been proposed in the literature by extending MFEA for solving the complex optimization problems. In particular, based on the success of MFEA, Gupta *et al.* further developed the multiobjective MFEA for conducting EMT in the multiobjective optimization problem domain. Further, Zhou *et al.* [16] proposed a permutation-based MFEA for conducting EMT on the NP-hard capacitated VRPs, while Bali *et al.* [6] proposed an MFEA variant with linearized domain adaptation strategy for transforming the search space of a simple task to its constitutive complex task which possesses similar search space. Moreover, Liaw and Ting [17] and Wen and Ting [18] extended MFEA for many task optimization by exploring the resource allocation mechanism for reallocating fitness evaluations on offsprings of different tasks. Ding *et al.* proposed a generalized MFEA to solve the expensive optimization problems with the help from cheap optimization problems. More recently, Yang *et al.* [8] presented a multiobjective MFEA for operational indices optimization in beneficiation processes. Li *et al.* [19] applied the MFEA for optimizing multiple sparse reconstruction tasks in hyperspectral image unmixing. In these works, all the tasks are optimized by a single population of individuals, and knowledge transfer across tasks carries out implicitly when two individuals possessing different skill factors are selected for generating the offspring via crossover.

On the other hand, the explicit EMT is a recently proposed new EMT paradigm [11]. In contrast to the existing implicit EMT algorithms, the explicit EMT employs independent population for each task and conducts knowledge transfer across tasks in an explicit manner. In particular, in [11], by deriving a denoising autoencoder, the optimized solutions found for different tasks along the evolutionary search can be explicitly transferred across tasks via a simple matrix multiplication. There are several advantages of explicit EMT. First, since each task has separate population for evolution, task-specific solution encoding schemes can be employed for different tasks. As the encoding scheme defines the search landscape of an optimization problem of interest, an appropriate solution representation is essential to the efficiency and effectiveness of the problem-solving process [20]–[26]. Next, by only designing an explicit knowledge transfer operator, the explicit EMT paradigm can be easily developed by employing different existing evolutionary solvers for each optimization task. In this way, different evolutionary solvers with various search capabilities can learn from each other toward enhanced optimization performance. Further, rather than probabilistically selecting solutions for mating across tasks in the implicit EMT, more flexible solution selection schemes, such as elite selection, can be performed before transfer in the explicit EMT, for reducing negative knowledge transfer effects [27]–[33]. Nevertheless, as the input and output of the developed denoising autoencoder in [11] are required to be real number, the explicit EMT proposed in [11] can only deal with continuous optimization problems, which thus remains the explicit EMT design for combinatorial optimization unexplored.

B. Vehicle Routing Problem

In this article, we consider the capacitated VRP (CVRP) which is one of the most classical VRP variants, as the representative combinatorial optimization problem domain for investigation. CVRP is an NP-hard problem, which was introduced by Dantzig and Ramser [34]. It defines the problem of designing a set of vehicle routes, in which a fleet of delivery vehicles with uniform capacity must serve known customer demands of single commodity, from a common depot at minimum cost [35]–[38]. More mathematically, the CVRP can be defined as follows. Given a connected undirected graph $\mathbf{G} = (\mathbf{V}, \mathbf{E})$ (see Fig. 1), where customer set $\mathbf{V} = \{\mathbf{v}_i\}$, $i = 1, \ldots, n$, n is the number of customers, edge set $\mathbf{E} = \{\mathbf{e}_{ij}\}$, $i, j = 1, \ldots, n$, denoting the edge between customer $\mathbf{v}_i$ and $\mathbf{v}_j$. $\mathbf{v}_0$ denotes the depot at which k homogeneous vehicles are based. Each edge $\mathbf{e}_{ij}$ is associated with a non-negative weight c_{ij}, which represents the distance from customer $\mathbf{v}_i$ to $\mathbf{v}_j$. Considering a demand set $\mathbf{D} = \{d(\mathbf{v}_i)|\mathbf{v}_i \in \mathbf{V}\}$, where $d(\mathbf{v}_i) > 0$ implies that customer $\mathbf{v}_i$ requires to be served, the CVRP aims at designing a set of vehicle routes $\mathbf{S} = \{\mathbf{C}_i\}$, $i = 1, \ldots, k$, with minimal distance cost such that:

1) each route $\mathbf{C}_i$ must start and end at the depot node $\mathbf{v}_0$;
2) the total load of each route must be no more than the capacity Q of each vehicle, $\sum_{\forall \mathbf{v}_i \in \mathbf{C}} d(\mathbf{v}_i) \leq Q$;
3) $\forall \mathbf{v}_i \in \mathbf{V}$, $d(\mathbf{v}_i) > 0$, there exists one and only one route $\mathbf{C}_i \in \mathbf{S}$ such that $\mathbf{v}_i \in \mathbf{C}_i$.

The objective function of CVRP can then be defined as

$$\mathrm{Cost}(\mathbf{S}) = \sum_{i=1}^{k} \mathrm{dis}(\mathbf{C}_i) \tag{1}$$

where $\mathrm{dis}(\mathbf{C}_i)$ denotes the summation of the travel distance (i.e., $\mathbf{e}_{ij}$) contained in route $\mathbf{C}_i$. An illustrative example of CVRP is given in Fig. 1.

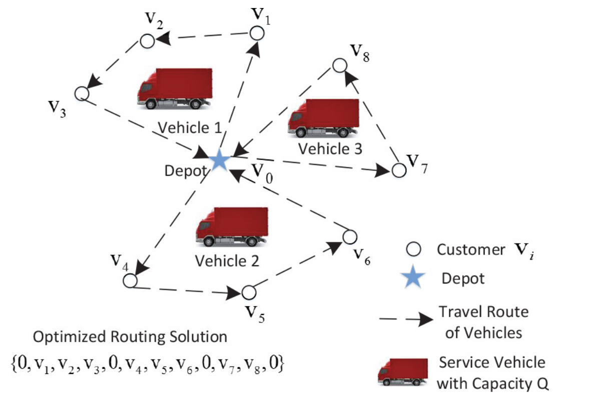


Fig. 1. Example of a CVRP.

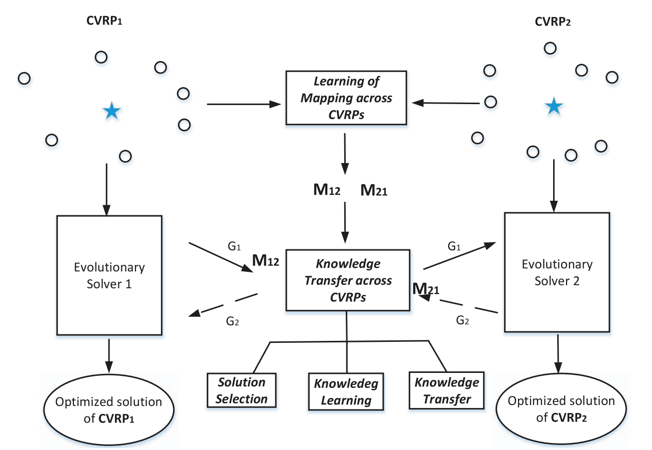


Fig. 2. Outline of the proposed explicit EMT for solving CVRPs.

III. Proposed Explicit EMT for CVRP

In this section, the details of the proposed explicit EMT for the combinatorial optimization problem, particularly, CVRP, are presented. Specifically, as depicted in Fig. 2, given two CVRPs (i.e., CVRP_1 and CVRP_2), the *learning of mapping across CVRPs* first kicks in to build the connections (i.e., $\mathbf{M}_{12}$ from CVRP_1 to CVRP_2 and $\mathbf{M}_{21}$ from CVRP_2 to CVRP_1) between these two problems, based on the problem data, that is, customer distribution. This process guarantees that the useful traits can be explicitly transferred across CVRPs possessing diverse problem properties, such as customer topology, number of vehicles, and customer size. Next, two separate evolutionary solvers are employed to optimize each CVRP, respectively. As the *knowledge transfer across CVRPs* happens while the evolutionary search progresses online, in this article, for simplicity, we define the knowledge transfer happens with a fixed generation interval (see G_1 and G_2 in Fig. 2). However, without loss of generality, other methods for defining the transfer frequency can also be applied. For the knowledge transfer across CVRPs, to reduce negative transfer effect, we first apply *selection* process to identify the high-quality solutions for transfer, which is followed by the *knowledge learning* process to capture the latent useful information that can be transferred across CVRPs. Further, the *knowledge transfer* will be performed with the built mappings to share the learned useful traits across CVRPs.

In what follows, the details of the proposed *learning of mapping* and *knowledge transfer* across CVRPs are presented.

A. Learning of Mapping Across CVRPs

As discussed in Section II-B, the objective of CVRP is to find the optimal assignments of customers to the vehicles and the optimal service orders of customers assigned to a common vehicle. Therefore, the aim of knowledge transfer in EMT across CVRPs is to share the good customer assignments as well as customer service orders found along the evolutionary search across CVRPs. To this end, the key is to learn the proper mapping between the customers across CVRPs. It is straightforward to see that the simplest and best customer mapping is a one-to-one mapping. In another word, for each customer in one CVRP, if there is one and only one corresponding customer in the other CVRP, the customer assignments and service orders can then be transferred accordingly. However, as different CVRPs may have various number of customers, it is unpractical to learn the one-to-one mapping between customers of different CVRPs. In this article, we thus propose to learn a sparse customer mapping to represent one CVRP customer by the most similar customers across CVRPs.

In particular, given two CVRPs, that is, $\mathbf{P}_s$ and $\mathbf{P}_t$, which contain the corresponding customer information, that is, customer locations, that are represented by a $d \times n_s$ and a $d \times n_t$ matrix, respectively, (d denotes the number of features for representing the location of a customer, while n_s and n_t give the number of customers in $\mathbf{P}_s$ and $\mathbf{P}_t$, respectively). In this way, the problem of finding customers from $\mathbf{P}_s$ to represent customers in $\mathbf{P}_t$ can be formulated as the learning of a $n_s \times n_t$ transformation matrix $\mathbf{M}$, so that $\mathbf{P}_s * \mathbf{M} = \mathbf{P}_t$. Furthermore, in order to find the most similar customers in $\mathbf{P}_s$ for customers in $\mathbf{P}_t$, we propose to learn a sparse $\mathbf{M}$ via minimizing a weighted l_1-norm-regularized reconstruction error, which is given by

$$\min_{\mathbf{M}} \|\mathbf{P}_s * \mathbf{M} - \mathbf{P}_t\|_F + ||\mathbf{D} \odot \mathbf{M}||_{l_1} \tag{2}$$

where the first term denotes the reconstruction error of customers in $\mathbf{P}_t$ using the customers in $\mathbf{P}_s$, while the second term is the weighted l_1-norm-based regularization on the mapping $\mathbf{M}$. $\odot$ denotes the elementwise product between two matrices, and F is the Frobenius norm. $\mathbf{D}$ is an $n_s \times n_t$ matrix, which denotes the weight matrix to further reinforce the sparsity of the mapping $\mathbf{M}$ [39]. Each element d_{ij} of $\mathbf{D}$ is calculated by

$$d_{ij} = \exp\left[e_{ij} - e^j_{\min}\right] * e_{ij} \tag{3}$$

where e_{ij} denotes the Euclidean distance between the ith customer in $\mathbf{P}_s$ and the jth customer in $\mathbf{P}_t$, and $e^j_{\min}$ gives the shortest Euclidean distance between all the customers in $\mathbf{P}_s$ to the jth customer in $\mathbf{P}_t$. Further, to solve (2), we propose to learn each column of $\mathbf{M}$ separately, which is given by

$$\min_{\mathbf{M}_{:j}} \left\|\mathbf{P}_s * \mathbf{M}_{:j} - \mathbf{P}^j_t\right\|_F + \left\|\mathbf{D}_j * \mathbf{M}_{:j}\right\|_{l_1} \tag{4}$$

where $\mathbf{M}_{:j}$ denotes the jth column of $\mathbf{M}$. $\mathbf{D}_j$ is an $n_s \times n_s$ diagonal matrix, in which the diagonal element d_{ii} is set as d_{ij} that is calculated via (3). By substituting $\mathbf{D}_j * \mathbf{M}_{:j}$ by $\mathbf{K}$, (4) becomes

$$\min_{\mathbf{K}} \left\|\mathbf{P}_s * \mathbf{D}_j^{-1} * \mathbf{K} - \mathbf{P}^j_t\right\|_F + ||\mathbf{K}||_{l_1} \tag{5}$$

which can be easily solved by the interior-point method [40], Finally, by calculating $\mathbf{M}_{:j} = \mathbf{D}_j^{-1} * \mathbf{K}$, the customer mapping $\mathbf{M}_{st}$ across CVRPs can be obtained by the concat of $\mathbf{M}_{:j}$s.

For two CVRPs, that is, CVRP_1 and CVRP_2, there are two customer mappings. One is from CVRP_1 to CVRP_2, that is, $\mathbf{M}_{12}$, and the other is from CVRP_2 and CVRP_1, that is, $\mathbf{M}_{21}$. Based on the proposed learning approach in (2), $\mathbf{M}_{12}$ is obtained by treating the problem data of CVRP_1 and CVRP_2 as $\mathbf{P}_s$ and $\mathbf{P}_t$, respectively. $\mathbf{M}_{21}$ is then calculated by exchanging the input and output for $\mathbf{M}_{12}$ accordingly.

B. Knowledge Transfer Across CVRPs

Based on the learned customer mapping $\mathbf{M}$s across CVRPs, the knowledge sharing toward enhanced EMT performance happens while the evolutionary search progresses online. In particular, the proposed knowledge sharing across CVRPs consists of three components, which are detailed below.

1) Solution Selection: The selection of solutions for knowledge transfer across CVRP domains is important for enhanced EMT performance, since inappropriate knowledge transfer will bring about negative transfer effect [27], [41]–[43]. In CVRP, as the useful information is embedded in the high-quality CVRP solutions, in this article, we propose to select the best Q number of optimized solutions, in terms of objective value, from the source CVRP domain to be transferred to the target CVRP domain, when the knowledge sharing is triggered.

2) Knowledge Learning: This process is to capture the useful information embedded in each of the selected solutions, which can be transferred across different CVRPs. As aforementioned, the objective of CVRP is to optimize both the assignments of customers to the vehicles and the service orders of customers assigned to a common vehicle. The optimization process of a CVRP thus can be interpreted as two separate phases. The first phase involves the assignment or clustering of the customers that require services to the appropriate vehicles. The second phase then serves to find the optimal service orders of each vehicle for the assigned customers obtained in phase 1. Therefore, if the proper similarity between customers can be learned, the optimal assignments and service order information can be easily obtained via clustering and pairwise distance sorting on the customers, respectively.

Keeping the above in mind, for each of the selected solution for transfer, here, we propose to estimate a new customer representation based on the customer assignments and customer service order information contained in the selected CVRP solution $\mathbf{s}$. This new customer representation seves as the learned knowledge that can be transferred across CVRPs to generate target CVRP solutions with guidance from the optimized assignments and service orders in $\mathbf{s}$. In particular, let $\mathbf{p}_s$ and $\mathbf{p}_t$ denote the source and target CVRP domain, respectively. $\mathbf{s}_s$ is the selected solution of $\mathbf{p}_s$. As given in Fig. 3, using the optimized solution depicted in Fig. 1 as an illustrative example, we first construct an $n_s \times n_s$ distance matrix

Optimized CVRP Solution

$$\{0, v_1, v_2, v_3, 0, v_4, v_5, v_6, 0, v_7, v_8, 0\}$$

$$\begin{pmatrix} 0 & \alpha & 2\alpha & \beta & \beta & \beta & \beta & \beta \\ \alpha & 0 & \alpha & \beta & \beta & \beta & \beta & \beta \\ 2\alpha & \alpha & 0 & \beta & \beta & \beta & \beta & \beta \\ \beta & \beta & \beta & 0 & \alpha & 2\alpha & \beta & \beta \\ \beta & \beta & \beta & \alpha & 0 & \alpha & \beta & \beta \\ \beta & \beta & \beta & 2\alpha & \alpha & 0 & \beta & \beta \\ \beta & \beta & \beta & \beta & \beta & \beta & 0 & \alpha \\ \beta & \beta & \beta & \beta & \beta & \beta & \alpha & 0 \end{pmatrix}$$

Fig. 3. Illustration of the constructed distance matrix using the CVRP example given in Fig. 1.

DM for all the customers in $\mathbf{p}_s$, where n_s is the number of customers. Each element dm_{ij} in **DM** represents the distance between the ith and jth customer. Further, in Fig. 3, α denotes a small real number, while β is a very big real number.[1] The motivation behind this new distance matrix is to make the customers served by a common vehicle to be close to each other, while keep customers served by different vehicles away from one another. Further, the distance between customers severed by the same vehicle increases linearly according to the corresponding service orders. As can be observed, the optimized vehicle assignment and service orders in $\mathbf{s}_s$ can be easily obtained via clustering and pairwise distance sorting using the constructed matrix **DM**, respectively. Next, the new estimated customer representations $\mathbf{P}_s^{\text{new}}$ of $\mathbf{p}_s$ can be obtained via the multidimensional scaling with **DM** [44], which possesses the complexity of $O(n^3)$, where n denotes the number of customers.

3) Knowledge Transfer: With the learned sparse customer mapping $\mathbf{M}$s across CVRP domains (i.e., $\mathbf{M}_{12}$ and $\mathbf{M}_{21}$), and the new CVRP customer representation $\mathbf{P}_s^{\text{new}}$ based on $\mathbf{s}_s$, the knowledge transfer across CVRPs can be performed via the simple operation of matrix multiplication. In particular, as outlined in Algorithm 1, for knowledge transfer from CVRP_1 to CVRP_2, we first set CVRP_1 and CVRP_2 as $\mathbf{p}_s$ and $\mathbf{p}_t$, respectively. Further, for each of the selected solution $\mathbf{s}_s$ for transfer, $\mathbf{P}_s^{\text{new}}$ is estimated as discussed in Section III-B2, which is a $d' \times n_s$ matrix, and d' is the new learned number of customer features. The approximated customers of $\mathbf{p}_t$ is then obtained via $\mathbf{P}_t^{\text{new}} = \mathbf{P}_s^{\text{new}} \times \mathbf{M}_{12}$. Next, to obtain the transferred CVRP solution for $\mathbf{p}_t$, K-means clustering with random initializations is conducted on $\mathbf{P}_t^{\text{new}}$ to derive the customer assignments of vehicles. Moreover, the service orders of each vehicle are subsequently achieved by sorting the pairwise distances among customers using $\mathbf{P}_t^{\text{new}}$ in an ascending order. The two customers with largest distance shall then denote the first and last customer to be served. Taking the first customer as reference, the service order of the remaining customers are defined according to the sorted orders. Finally, the transferred solution across CVRPs, is inserted into the population of $\mathbf{p}_t$ to undergo natural selection, and bias the optimization search process accordingly.

[1]The setting of these two values are based on the rule that the vehicle assignment and service order in $\mathbf{s}_s$ can be accurately obtained when applying clustering and pair-wise distance sorting with **DM**.

Algorithm 1: Proposed Knowledge Transfer in EMT for Solving CVRPs.

Begin
/* Knowledge transfer at G_1 from CVRP_1 to CVRP_2*/
Set $\mathbf{p}_s$ and $\mathbf{p}_t$ as CVRP_1 to CVRP_2, respectively;
$\mathcal{S} = \{\mathbf{s}_i | 1 \le i \le Q\}$, $\mathbf{s}_i$ is selected solution for transfer from $\mathbf{p}_s$, $i = 1$;
for $i \le Q$ **do**
    Set $\mathbf{s}_s = \mathbf{s}_i$, and **Estimate** $\mathbf{P}_s^{\text{new}}$ as discussed in Section III-B2;
    Obtain $\mathbf{P}_t^{\text{new}}$ via $\mathbf{P}_s^{\text{new}} \times \mathbf{M}_{12}$;
    Perform K-means and pair-wise distance sorting with $\mathbf{P}_t^{\text{new}}$ to generate CVRP solution for $\mathbf{p}_t$;
    Insert the generated solution into the population to undergo natural selection;
    $i = i + 1$;
/* Knowledge transfer at G_2 from CVRP_2 to CVRP_1*/
Set $\mathbf{p}_s$ and $\mathbf{p}_t$ as CVRP_2 to CVRP_1, respectively;
$\mathcal{S} = \{\mathbf{s}_i | 1 \le i \le Q\}$, $\mathbf{s}_i$ is selected solution for transfer from $\mathbf{p}_s$, $i = 1$;
for $i \le Q$ **do**
    Set $\mathbf{s}_s = \mathbf{s}_i$, and **Estimate** $\mathbf{P}_s^{\text{new}}$ as discussed in Section III-B2;
    Obtain $\mathbf{P}_t^{\text{new}}$ via $\mathbf{P}_s^{\text{new}} \times \mathbf{M}_{21}$;
    Perform K-means and pair-wise distance sorting with $\mathbf{P}_t^{\text{new}}$ to generate CVRP solution for $\mathbf{p}_t$;
    Insert the generated solution into the population to undergo natural selection;
    $i = i + 1$;
End

For knowledge transfer from CVRP_2 to CVRP_1, $\mathbf{p}_s$ and $\mathbf{p}_t$ are then set as CVRP_2 and CVRP_1 accordingly. The knowledge transfer process is performed exactly the same as discussed above, using the customer mapping $\mathbf{M}_{21}$. Furthermore, as can be observed, the computational cost of the knowledge transfer process contains two parts, one is the K-means clustering with complexity of $O(n * d * k * z)$, and the other is the pairwise distances sorting possessing the complexity of $O(d * n^2)$ in the worst case, where n, d, k, and z denote the number of customers, the feature dimension of each customer, the number of vehicles, and the number of iterations in K-means clustering, respectively.

IV. Empirical Study on VRP Benchmarks

In this section, comprehensive empirical studies using common CVRP benchmarks are conducted to verify the efficacy of the proposed explicit EMT for combinatorial optimization. In particular, we first introduce how the multitasking CVRP benchmarks, which possess different problem similarities, are generated based on the existing CVRP instances. Next, the experimental configurations as well as the obtained results are presented and discussed.

TABLE I
PROPERTY SUMMARY OF THE CVRP INSTANCES

Instance	Customer Number	Vehicle Capacity	Vehicle Number
A-n54-k7	54	100	7
A-n62-k8	62	100	8
A-n80-k10	80	100	10
B-n50-k7	50	100	7
B-n64-k9	64	100	9
B-n78-k10	78	100	10
P-n50-k8	50	120	8
P-n60-k10	60	120	10
P-n76-k5	76	280	5

A. Experimental Setup

In this article, nine CVRP instances of the "AUGERAT" CVRP benchmark set with diverse properties (e.g., number of customers, vehicle number, etc.) are used. The detailed properties of the nine CVRP instances considered are summarized in Table I. To construct the multitasking CVRP instances, according to the recent reports of multitasking benchmarks for continuous optimization [45], [46], here we propose to build the high-, medium-, and low-similarity multitasking CVRP pairs by randomly and independently deleting 10%, 30%, and 50% customers from the CVRP instances, respectively. For instance, to build the high-similarity multitasking pairs of "A-n54-k7," we will perform two separate random deletions of 10% customers from A-n54-k7. The resultant two instances are labeled as "A-n54-k7-h-t1" and "A-n54-k7-h-t2," where "h" stands for high-similarity, while "t1" and "t2" denote task 1 and task 2 of A-n54-k7, respectively. In this way, each CVRP instance will generate three sets of multitasking CVRP benchmarks, and totally, there are 27 sets of multitasking CVRP benchmarks will be built.

Subsequently, according to [47] and [48], we implement a strong memetic algorithm[2] which serves as the single-task evolutionary solver for solving CVRP, that is labeled as EA1. Moreover, as explicit EMT is able to incorporate different solvers for different optimization tasks, to investigate the benefits of employing various search mechanisms for optimization, we further implement another single-task solver, that is, EA2, which shares the same reproduction with EA1, but differs in the local search settings. In the proposed explicit EMT algorithm (EEMTA), EA1 and EA2 are employed for solving task 1 and task 2, respectively. Further, to verify the effectiveness of the transferred solutions across CVRPs, EA1 and EA2 with injection of randomly generated solution along the evolutionary search process, which are labeled as EA1+*R* and EA2+*R*, respectively, are compared. Please note that the frequency and amount of solution for injection in both EA1 and EA2 are kept the same as EEMTA. Moreover, to verify the efficacy of the proposed method, besides the single-task EAs, the recently proposed implicit EMT algorithm, that is, PMFEA [16], for solving combinatorial optimization problems, is also considered as the baseline algorithm. For fair comparison, the search operators of PMFEA are kept the same as EA1 for the comparison on task 1, and changed to be consistent with EA2 for the evaluation on task 2. Finally, all the search operator and parameter settings of the EA1, EA2, EA1+*R*, EA2+*R*, PMFEA, and the proposed EEMTA are referred to [16], [47], and [48] accordingly. Detailed configurations are given as follows.

1) *Parameters for the Proposed EEMTA:*
   a) α and β in DM: $\alpha = 10$ and $\beta = 1000$.
   b) Number of solutions selected for transfer: $Q = 5$.
   c) Generation interval for knowledge transfer: $G_1 = G_2 = 5$.
2) *Population Size:*
   a) EA1, EA2, EA1+*R*, EA2+*R*, and EEMTA: 50.
   b) PMFEA: 100.
3) Maximum generations: 100.
4) Independent runs: 20.
5) *Local Search Settings:*
   a) Local search in EA1 and EA1+*R*: Replace, single-insertion, and two-swap [47], [48].
   b) Local search in EA2 and EA2+*R*: Replace [47], [48].
6) Probability of local search: 0.1.

As can be observed, since PMFEA has only one population for solving two tasks, the population size of PMFEA is doubled when compared to the other algorithms. The reproduction operator settings of the single-task EA1 and EA2 are kept the same and referred to [47]. Finally, the explicit solution transfer across CVRPs in the proposed EEMTA happens in every five generations, and five best solutions in terms of objective value from each task will be selected for transfer from one to the other.

B. Results and Discussion

To investigate the performance of the proposed EEMTA, we present, analyze, discuss, and compare the results obtained against the recently proposed EMT algorithms and the traditional single-task EAs, based on the criteria of solution quality and search efficiency.

1) Solution Quality: To evaluate the solution quality of the proposed approach, Table II tabulates all the results obtained by the respective algorithms over 20 independent runs. In the table, the symbols "h," "m," and "l" in the "Problem" column denote the high-, medium-, and low-similarity multitasking CVRP benchmarks, respectively. The column "B.Cost" and "Ave.Cost" give the best solution and averaged solution obtained across 20 independent runs, respectively. Superior performance is highlighted using bold font in the table. Further, in order to obtain the statistically comparison, Wilcoxon rank-sum test with 95% confidence level has been conducted on the experimental results.

It can be observed from Table II that on task 1, with solutions implicitly transferred across CVRPs, the PMFEA obtained superior or competitive solution qualities on most of the CVRPs, such as "A-n80-k10-h-t1," "B-n78-k10-h-t1," and "P-n60-k10-h-t1," in terms of Ave.Cost when compared against the single-task EA1. As PMFEA and EA1 share the common search operator and parameter settings, these

[2]The implemented solver is a state-of-the-art memetic algorithm, which is able to achieve the best known results of the AUGERAT CVRP benchmark, that are available at: http://neo.lcc.uma.es/vrp/known-best-results/.

TABLE II
SOLUTION QUALITY OF EEMTA, PMFEA, EA1+*R*, EA2+*R*, EA1 AND EA2 ON THE 9 "AUGERAT" MULTITASKING CVRP BENCHMARK SETS. THE SUPERIOR SOLUTION QUALITY OF EACH RESPECTIVE PROBLEM INSTANCE IS HIGHLIGHTED IN BOLD FONT. ("≈," "+" AND "−" DENOTE EEMTA STATISTICALLY SIGNIFICANT SIMILAR, BETTER, AND WORSE THAN PMFEA, RESPECTIVELY)

Problem	Task No.	*EEMTA*		*PMFEA*		*EA1+R*		*EA2+R*		*EA1*		*EA2*	
		B.Cost	*Ave.Cost*	*B.Cost*	*Ave.Cost*	*B.Cost*	*Ave.Cost*	*B.Cost*	*Ave.Cost*	*B.Cost*	*Ave.Cost*	*B.Cost*	*Ave.Cost*
A-n54-k7-h-t1	T1	1072.0	1088.3±4.2≈	1082.0	1088.5±2.5	1087.0	1090.1±2.7	-	-	1072.0	**1088.2±4.0**	-	-
A-n54-k7-h-t2	T2	**1118.0**	**1148.9±13.3+**	1151.83	1166.7±8.6	-	-	1138.0	1160.3±11.4	-	-	1131.0	1154.8±9.6
A-n54-k7-m-t1	T1	899.0	905.8±5.8≈	899.0	904.0±7.7	899.0	**902.8±5.5**	-	-	899.0	904.4±5.6	-	-
A-n54-k7-m-t2	T2	**933.0**	**951.3±9.0+**	938.0	978.5±19.9	-	-	945.0	963.5±16.4	-	-	938.0	960.7±20.6
A-n54-k7-l-t1	T1	690.0	692.2±1.6≈	690.0	691.7±1.1	690.0	691.9±1.4	-	-	690.0	**691.6±1.2**	-	-
A-n54-k7-l-t2	T2	707.0	**724.2±10.3≈**	713.79	727.3±9.1	-	-	714.0	727.5±8.0	-	-	707.0	724.9±12.4
A-n62-k8-h-t1	T1	1211.0	**1222.0±5.1+**	1219.0	1228.0±8.6	**1210.0**	1223.7±6.4	-	-	1217.0	1224.6±6.6	-	-
A-n62-k8-h-t2	T2	1237.0	**1250.7±12.2+**	1257.42	1279.4±10.5	-	-	**1218.0**	1254.9±23.9	-	-	1236.0	1261.8±19.8
A-n62-k8-m-t1	T1	915.0	**916.5±2.3+**	915.5	917.5±2.3	915.0	917.5±3.1	-	-	915.0	917.5±2.9	-	-
A-n62-k8-m-t2	T2	**942.0**	975.9±20.8+	958.55	989.5±18.5	-	-	960.0	972.7±9.4	-	-	952.0	**970.2±9.6**
A-n62-k8-l-t1	T1	765.0	**765.0±0.0≈**	765.0	765.7±1.9	765.0	766.5±3.1	-	-	765.0	767.4±5.2	-	-
A-n62-k8-l-t2	T2	763.0	**789.2±15.1+**	791.7	822.7±15.8	-	-	771.0	814.0±18.7	-	-	763.0	803.4±23.3
A-n80-k10-h-t1	T1	1668.0	1678.8±7.6≈	**1652.7**	1679.7±11.3	1655.0	**1674.7±11.7**	-	-	1658.0	1680.7±10.1	-	-
A-n80-k10-h-t2	T2	**1747.0**	**1775.6±13.1+**	1783.0	1810.7±16.3	-	-	1779.0	1805.6±18.7	-	-	1761.0	1799.1±18.7
A-n80-k10-m-t1	T1	1258.0	**1268.1±8.9≈**	1259.0	1271.0±7.8	1258.0	1270.2±7.4	-	-	1259.0	1270.7±8.1	-	-
A-n80-k10-m-t2	T2	**1346.0**	**1394.1±17.8+**	1382.0	1423.4±21.3	-	-	1367.0	1409.4±16.9	-	-	1379.0	1400.8±15.5
A-n80-k10-l-t1	T1	1022.0	1025.1±2.3≈	1022.0	1025.5±2.1	1022.0	**1024.7±1.6**	-	-	1022.0	1025.6±2.6	-	-
A-n80-k10-l-t2	T2	**1055.0**	**1073.1±8.3+**	1062.0	1083.7±9.2	-	-	1063.0	1074.9±7.7	-	-	1063.0	1079.1±7.3
B-n50-k7-h-t1	T1	655.0	655.0±0.0≈	655.0	655.0±0.1	655.0	655.0±0.0	-	-	655.0	655.2±0.5	-	-
B-n50-k7-h-t2	T2	**696.0**	**697.9±1.2+**	707.53	726.0±14.0	-	-	702.0	715.0±10.8	-	-	703.0	717.6±10.5
B-n50-k7-m-t1	T1	525.0	**525.0±0.0≈**	525.0	525.1±0.2	525.0	525.1±0.2	-	-	525.0	525.2±0.4	-	-
B-n50-k7-m-t2	T2	624.0	**633.3±6.8+**	625.0	643.7±13.0	-	-	623.0	635.6±11.5	-	-	623.0	645.6±9.7
B-n50-k7-l-t1	T1	435.0	435.0±0.0≈	435.0	435.0±0.0	435.0	435.0±0.0	-	-	435.0	435.0±0.0	-	-
B-n50-k7-l-t2	T2	461.0	463.3±1.9≈	461.0	463.4±2.0	-	-	461.0	**463.0±1.8**	-	-	461.0	463.6±1.8
B-n64-k9-h-t1	T1	805.0	805.0±0.0≈	805.0	805.0±0.0	805.0	805.0±0.0	-	-	805.0	805.0±0.0	-	-
B-n64-k9-h-t2	T2	796.0	**800.2±2.3+**	805.85	819.0±6.3	-	-	796.0	811.3±6.3	-	-	802.0	813.4±5.3
B-n64-k9-m-t1	T1	647.0	647.0±0.0≈	647.0	647.0±0.0	647.0	647.0±0.0	-	-	647.0	647.0±0.0	-	-
B-n64-k9-m-t2	T2	**699.0**	**711.6±11.4+**	711.0	733.0±15.7	-	-	700.0	714.9±8.0	-	-	700.0	721.9±13.7
B-n64-k9-l-t1	T1	531.0	531.0±0.0≈	531.0	531.0±0.0	531.0	531.0±0.0	-	-	531.0	531.0±0.0	-	-
B-n64-k9-l-t2	T2	541.0	555.5±9.3≈	545.0	559.8±6.8	-	-	541.0	**552.3±7.1**	-	-	541.0	553.3±8.8
B-n78-k10-h-t1	T1	1056.0	1072.4±8.4≈	**1055.79**	**1069.6±8.5**	1058.0	1070.6±7.9	-	-	1058.0	1072.6±8.8	-	-
B-n78-k10-h-t2	T2	**1175.0**	**1224.9±14.6+**	1221.54	1255.0±14.2	-	-	1200.0	1244.1±14.3	-	-	1217.0	1245.1±13.2
B-n78-k10-m-t1	T1	908.0	908.1±0.4≈	908.0	908.3±0.6	908.0	908.2±0.9	-	-	908.0	908.1±0.4	-	-
B-n78-k10-m-t2	T2	**953.0**	**976.0±12.0+**	971.0	1008.9±20.5	-	-	973.0	1002.9±21.3	-	-	977.0	1003.5±21.7
B-n78-k10-l-t1	T1	781.0	782.4±0.9≈	781.0	**782.1±0.5**	782.0	782.2±0.5	-	-	781.0	782.3±0.9	-	-
B-n78-k10-l-t2	T2	837.0	**847.6±5.7+**	839.43	856.7±10.1	-	-	838.0	851.0±6.3	-	-	**829.0**	851.4±6.5
P-n50-k8-h-t1	T1	563.0	**563.0±0.0≈**	563.0	563.1±0.4	563.0	563.1±0.2	-	-	563.0	563.1±0.3	-	-
P-n50-k8-h-t2	T2	**594.0**	**608.4±5.5+**	602.0	613.6±7.1	-	-	599.0	611.1±6.3	-	-	602.0	613.1±7.7
P-n50-k8-m-t1	T1	460.0	460.0±0.0≈	460.0	460.0±0.1	460.0	460.0±0.0	-	-	460.0	460.0±0.0	-	-
P-n50-k8-m-t2	T2	508.0	519.1±4.6≈	503.0	522.1±7.7	-	-	**499.0**	519.6±6.7	-	-	504.0	**517.6±8.0**
P-n50-k8-l-t1	T1	412.0	412.0±0.0≈	412.0	412.1±0.5	412.0	412.0±0.0	-	-	412.0	412.3±0.7	-	-
P-n50-k8-l-t2	T2	360.0	**367.0±4.2+**	363.72	372.0±5.4	-	-	360.0	368.3±5.0	-	-	363.0	372.2±5.3
P-n60-k10-h-t1	T1	679.0	684.2±4.3≈	679.0	**684.1±4.1**	679.0	684.8±4.9	-	-	679.0	685.2±3.9	-	-
P-n60-k10-h-t2	T2	**713.0**	**727.1±6.1+**	732.36	746.4±6.3	-	-	726.0	745.9±6.5	-	-	734.0	747.4±5.1
P-n60-k10-m-t1	T1	561.0	561.8±1.3≈	561.0	562.7±1.7	561.0	562.3±1.5	-	-	561.0	**561.5±1.1**	-	-
P-n60-k10-m-t2	T2	**554.0**	**571.6±9.7+**	563.0	578.3±8.0	-	-	558.0	574.7±9.5	-	-	557.0	573.8±9.0
P-n60-k10-l-t1	T1	405.0	405.1±0.2≈	405.0	405.2±0.5	405.0	**405.0±0.0**	-	-	405.0	405.1±0.2	-	-
P-n60-k10-l-t2	T2	467.0	482.1±6.4≈	468.0	484.5±8.0	-	-	467.0	**479.1±5.8**	-	-	468.0	480.3±6.6
P-n76-k5-h-t1	T1	583.0	**586.6±2.3≈**	**580.34**	588.3±4.2	585.0	589.1±3.2	-	-	583.0	587.6±3.6	-	-
P-n76-k5-h-t2	T2	**583.0**	**598.8±6.8+**	608.26	627.2±9.6	-	-	609.0	631.4±9.1	-	-	618.0	633.2±9.7
P-n76-k5-m-t1	T1	507.0	507.0±0.0≈	507.0	507.2±0.3	507.0	507.0±0.0	-	-	507.0	507.0±0.0	-	-
P-n76-k5-m-t2	T2	**522.0**	**538.0±8.8+**	531.75	553.4±11.9	-	-	537.0	555.4±10.7	-	-	532.0	548.7±12.2
P-n76-k5-l-t1	T1	422.0	**422.6±0.8≈**	422.0	423.0±0.9	422.0	422.7±0.9	-	-	422.0	423.2±0.9	-	-
P-n76-k5-l-t2	T2	366.0	383.4±9.8≈	374.54	385.6±5.5	-	-	367.0	382.7±7.5	-	-	**362.0**	**381.7±11.8**

obtained results again confirmed the effectiveness of conducting EMT. However, on task 2 of most CVRP instances, although the search operator and parameter settings of PMFEA are now consistent with EA2, PMFEA achieved poor "Ave.Cost" values. This is because EA2 has weaker local search than EA1, the solutions found along the search by EA2 are thus poorer than EA1. As the solution transfer in implicit EMT is via crossover without elite selection, the transferred low-quality solutions caused negative transfer performance in PMFEA.

On the other hand, in Table II, the proposed EEMTA has been observed to outperform the single task EA1 and EA2 on task 1 and task 2 of most multitasking CVRP benchmarks, respectively. In particular, the proposed EEMTA achieved superior solution quality in terms of Ave.Cost, compared to EA1 or EA2 on all the 27 multitasking CVRP benchmarks and obtained improved Ave.Cost values than both EA1 and EA2 on 12 number of multitasking CVRPs. Further, in terms of B.Cost, on task 2 of most CVRP benchmarks, with solution transferred from task 1, EEMTA has found superior best solutions than EA2. On task 1 of the benchmarks, such as "A-n80-k10-m-t1" and "A-n62-k8-h-t1," with solution transferred from task 2, EEMTA has also achieved better best solutions against EA1. As EA1 has more powerful local search capability than EA2, from the obtained results above, we can see that not only strong solver can improve the search process of the weak solver, but the weak solver may also contains information which is useful to enhance the search of the strong solver in EMT. Further, it also can be observed that, EA1+*R* and EA2+*R* achieved competitive solution quality against EA1 and EA2 respectively, on all the CVRP benchmarks. As both EA1+*R* and EA2+*R* share the same frequency and amount of solutions for injection, with the proposed EEMTA, and only differ in the generation of injected solutions, the obtained superior

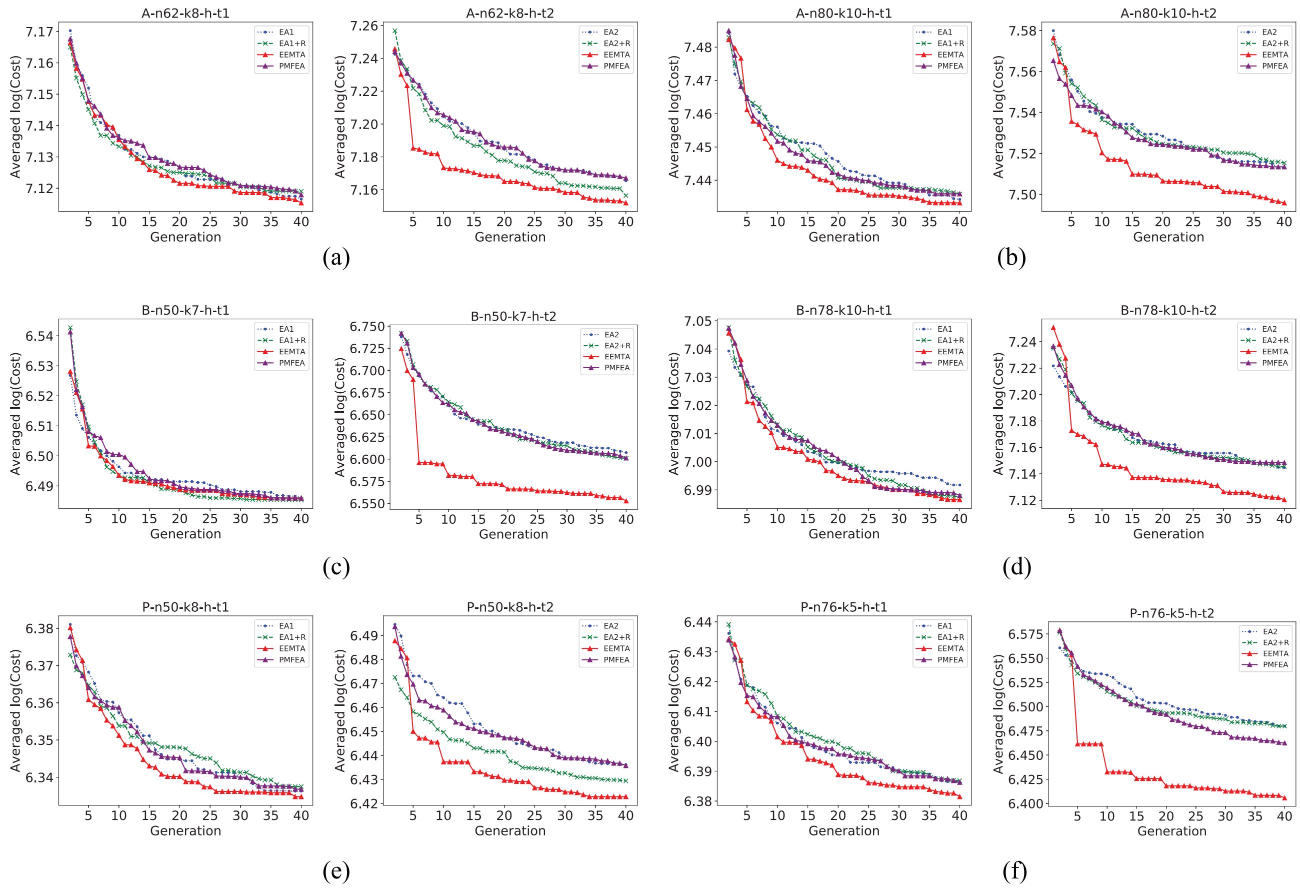


Fig. 4. Convergence traces of EEMTA versus the PMFEA and the single-task EAs on representative high-similarity multitasking CVRPs. y-axis: Log (Averaged travel cost); x-axis: Generation. (a) Convergence traces of multitasking A-n62-k8. (b) Convergence traces of multitasking A-n80-k10-h. (c) Convergence traces of multitasking B-n50-k7-h. (d) Convergence traces of multitasking B-n78-k10-h. (e) Convergence traces of multitasking P-n50-k8-h. (f) Convergence traces of multitasking P-n76-k5-h.

solution quality of EEMTA confirmed the effectiveness of the proposed explicit EMT across CVRPs.

Finally, in contrast to the implicit multitasking algorithm, that is, PMFEA, the proposed EEMTA has obtained superior solutions with respect to Ave.Cost on most of the multitasking CVRP benchmarks. Moreover, due to the advantage of employing different evolutionary solvers in the proposed explicit multitasking paradigm, EEMTA has achieved significantly better Ave.Cost values than PMFEA, on task 2 of totally 22 number of CVRP benchmarks. In summary, since EEMTA shares the common search operators and parameters with PMFEA, EA1+R, EA2+R, EA1, and EA2, for solving the multitasking CVRP benchmarks, the obtained enhanced search performance by EEMTA, in terms of solution quality confirmed the effectiveness of the proposed approach for conducting explicit EMT for combinatorial optimization.

2) Search Efficiency—Convergence Trends: To assess the efficiency of the proposed EEMTA, the representative search convergence traces of EEMTA, PMFEA, EA1, and EA2 on the representative high-, medium-, and low-similarity multitasking CVRP benchmarks are presented in Figs. 4, 5, and 7, respectively. In the figures, the y-axis denotes the averaged travel cost obtained in log scale, while the x-axis gives the respective computational effort incurred in terms of generation made so far.

From these figures, it can be observed that the implicit EMT algorithm, that is, PMFEA converges faster or competitive than the single-task EAs (i.e., EA1 or EA2) on most of the multitasking CVRPs. However, as PMFEA employs a common search mechanism for different tasks, and there is no solution selection in the knowledge transfer process across tasks, the improvements in terms of convergence speed achieved by PMFEA are limited even on the high-similarity multitasking CVRPs (see Fig. 4).

Next, for the proposed explicit EMT algorithm for combinatorial optimization, due to the transfer of high-quality solutions from the strong solver EA1 on task 1 to task 2, EEMTA has obtained significantly faster convergence speed over both the single-task EA2 and the multitasking PMFEA, on task 2 of all the high-, medium-, and low-similarity multitasking CVRPs. For instance, on "A-n62-k8-h-t2," EEMTA uses only about five generations to arrive the solution obtained by EA2 and PMFEA at generation 20. On "B-n50-k7-m-t2," EEMTA takes 20 generations to achieve the solution better than that obtained by EA2 and PMFEA over 30 generations. However, due to the decrease of similarity between tasks from high-similarity to the low-similarity multitasking benchmarks, we can observe that the improvements of convergence speed obtained by EEMTA over EA2 on task 2 of the multitasking CVRPs also decrease from Figs. 4–7. For example, in contrast to the

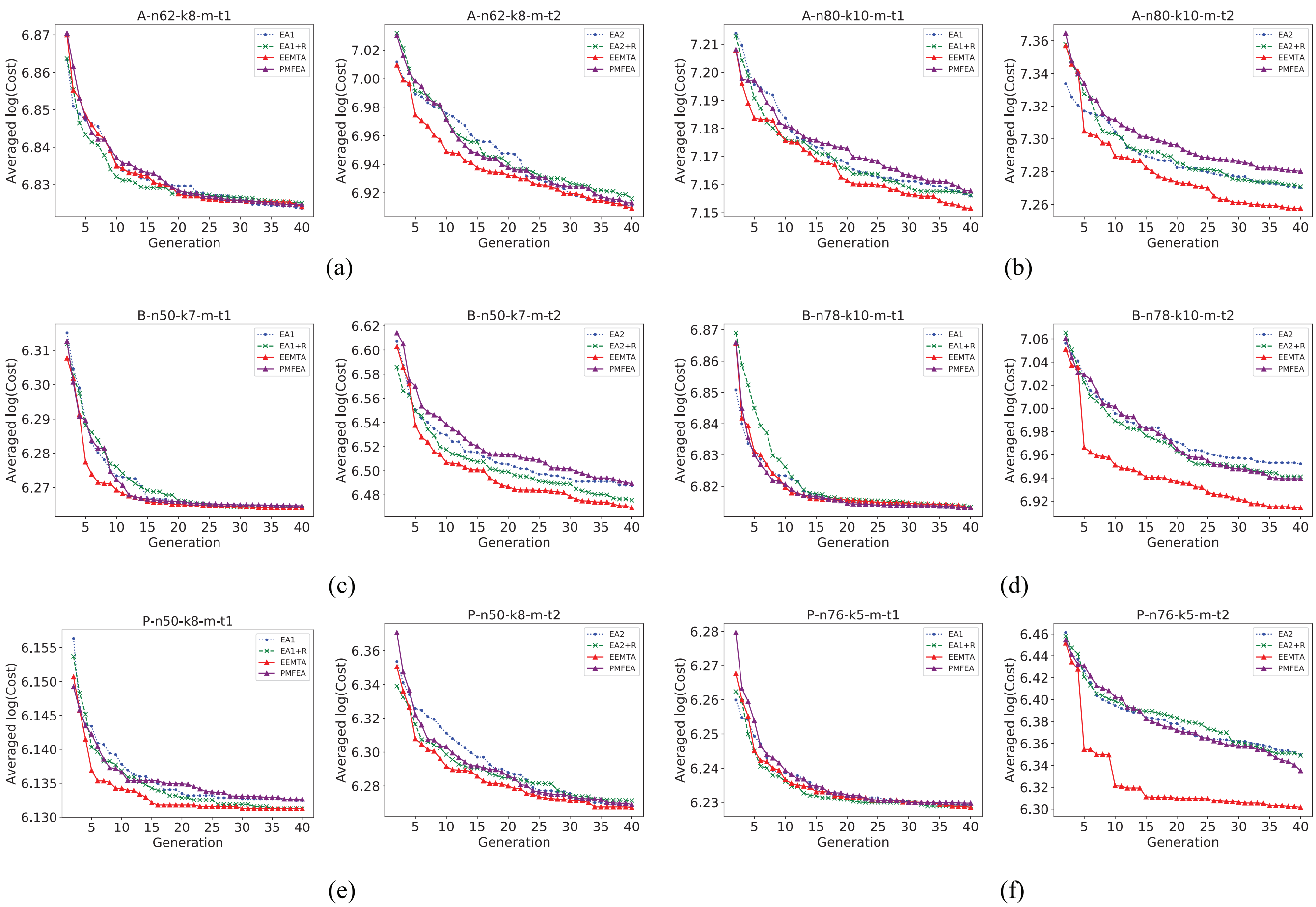
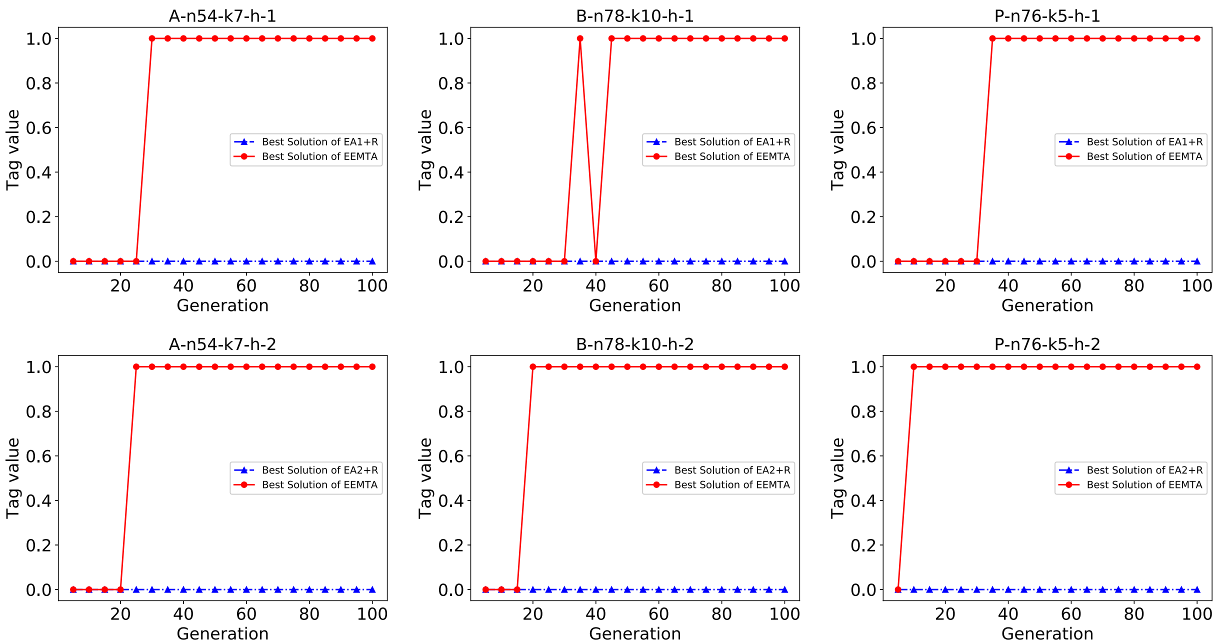


Fig. 5. Convergence traces of EEMTA versus the PMFEA and the single-task EAs on representative medium-similarity multitasking CVRPs. *y*-axis: Log(Averaged travel cost); *x*-axis: Generation. (a) Convergence traces of multitasking A-n62-k8. (b) Convergence traces of multitasking A-n80-k10. (c) Convergence traces of multitasking B-n50-k7. (d) Convergence traces of multitasking B-n78-k10. (e) Convergence traces of multitasking P-n50-k8. (f) Convergence traces of multitasking P-n76-k5.

high-similarity benchmark "A-n62-k8-h-t2" aforementioned [see Fig. 4(a)], on the low-similarity benchmark "A-n62-k8-l-t2" [see Fig. 7(a)], it takes about 15 generations of EEMTA to reach the solution found by EA2 and PMFEA at generation 20. Further, on task 1 of the multitasking CVRPs, although task 2 has a weaker evolutionary solver, that is, EA2, the transferred solutions from task 2 could also be useful, and thus lead to faster or competitive convergence speed of EEMTA over both EA1 and PMFEA on most of the multitasking CVRPs.

Moreover, with the injection of randomly generated solutions in EA1 and EA2, we can observe that EA1+*R* and EA2+*R* obtained faster convergence speed over EA1 and EA2, respectively, on most of the CVRP benchmarks, such as Figs. 4(e), 5(c), and 7(d). This is because of the solution diversity introduced along the evolutionary search of EA1+*R* and EA2+*R*. However, using the same configuration of frequency and amount of solutions for injection, with the proposed knowledge transfer across CVRPs, EEMTA has been observed to achieve faster convergence speed over both EA1+*R* and EA2+*R* on all the high-, medium-, and low-similarity multitasking CVRP benchmarks. Further, to explore the reason behind the superior performance obtained by EEMTA over EA1+*R* and EA2+*R*, in Fig. 6, we plot the tracking of best solutions found along the evolutionary search in EEMTA, EA1+*R*, and EA2+*R*. In particular, if the best solution is generated from the transferred solutions, it will be given tag value 1; otherwise, 0. As can be observed in the figure, in the proposed EEMTA, the transferred solutions across CVRPs successfully lead to the finding of best solutions along the search, while the randomly transferred solutions in both EA1+*R* and EA2+*R* failed to contribute to the localization of best solutions.

Finally, as G (i.e., G_1 and G_2) and Q define the frequency and amount of knowledge sharing between CVRPs, we further study how the configurations of G and Q affect the proposed EEMTA. Generally, a small value of G and a big value of Q will greatly increase the frequency and amount of knowledge sharing across tasks, while a big value of G and a small value of Q will reduce the frequency and amount of solution transfer across tasks significantly. In particular, the averaged travel costs obtained by the proposed EEMTA and the single-task solver (i.e., EA1 and EA2) on all the CVRPs across 20 independent runs with various configuration of G and Q are summarized in Fig. 8. It can be observed from the figures that the superior solution qualities have been obtained



(a) (b) (c)

Fig. 6. Tracking of best solutions along the evolutionary search. (a) A-n54-k7-h. (b) B-n78-k10-h. (c) P-n76-k5-h.

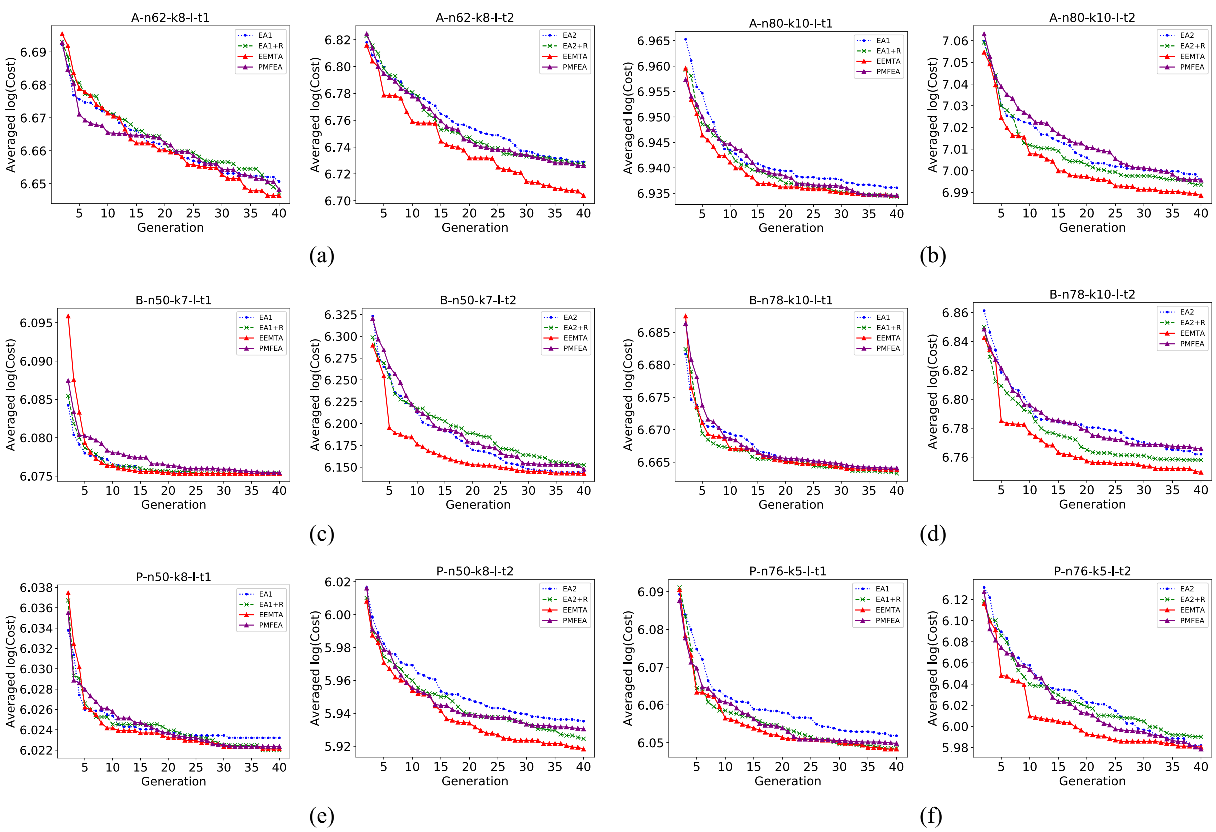


Fig. 7. Convergence traces of EEMTA versus the PMFEA and the single-task EAs on representative low-similarity multitasking CVRPs. y-axis: Log(Averaged travel cost); x-axis: Generation. (a) Convergence traces of multitasking A-n62-k8. (b) Convergence traces of multitasking A-n80-k10. (c) Convergence traces of multitasking B-n50-k7-l. (d) Convergence traces of multitasking B-n78-k10. (e) Convergence traces of multitasking P-n50-k8. (f) Convergence traces of multitasking P-n76-k5.

by the proposed EEMTA when compared to the single-task solvers with all the configurations of G and Q. However, although the optimal confirmations of G and Q are in general problem dependent, fixing $G_1 = G_2 = 5$ and $Q = 5$ is found to provide noteworthy results across a variety of problems encountered.

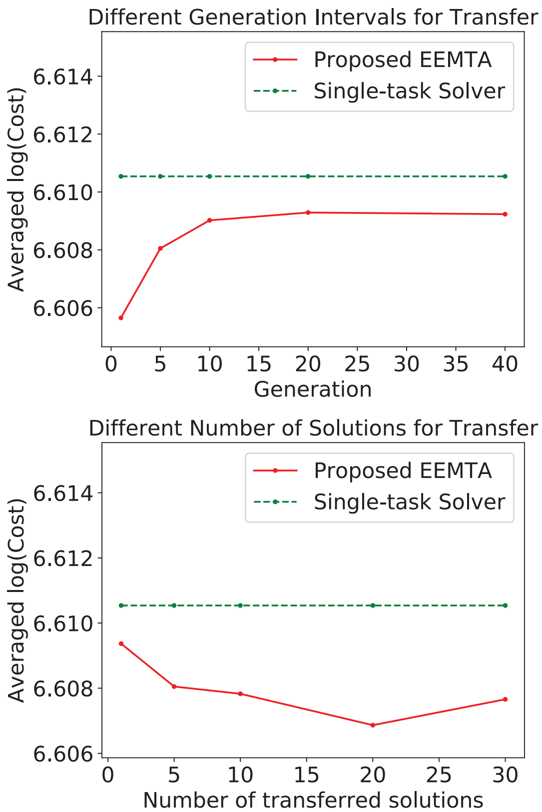


Fig. 8. Averaged travel cost in log scale obtained by the proposed EEMTA and single task solver (i.e., EA1 and EA2) on all the CVRPs.

In summary, it is worth noting here again that EEMTA has the common search mechanisms, that is, evolutionary search operators and parameters, with PMFEA, EA1+*R*, EA2+*R*, EA1, and EA2 for solving task 1 and task 2 of all the multitasking CVRPs. Therefore, the superior convergence speed of EEMTA can clearly be attributed to the efficiency of the proposed explicit multitasking approach for combinatorial optimization.

V. Real-World Routing Application: The Package Delivery Problem

In this section, we further investigate the performance of the proposed algorithm on a real-world vehicle routing application, that is, the package delivery problem (PDP), from the logistics industry. In particular, we aim to carry out the computational demanding optimization of multiple distinct package delivery routing requests from real courier business. It is noted that the requirements of obtaining high-quality routing solution efficiently is becoming as one of the key challenges of logistics in today's global economy era, thereby highlighting the potential real-world implications of our proposition.

The PDP is an NP-hard complex combinatorial optimization task. Due to the fast development of e-commerce, the courier companies confront with a large amount of package delivery tasks everyday. An efficient and effective optimization paradigm for solving PDP cannot only save the operating costs but also improve the service quality of the courier companies. Typically, the PDP can be defined as the task of servicing a set of customers with a fleet of capacity-constrained vehicles located at a single or multiple depot(s), which is a real-world application of the CVRP. In the present context, we have four PDP requests from a courier company in Beijing, China. In particular, the corresponding number of customers need to be served, the vehicle available at the courier company, and the capacity of the vehicles are summarized in Table III.

TABLE III
Property Summary of the PDP Requests

Instance	Customer Number	Vehicle Capacity	Vehicle Number
PDP Request-1	270	862	16
PDP Request-2	255	862	15
PDP Request-3	264	862	16
PDP Request-4	246	862	14

Usually, courier companies always optimize the PDP requests in a sequential manner via the single-task heuristic solvers. In the present study, we employ the proposed multitasking EEMTA to solve the PDP requests concurrently. The implicit PMFEA is also employed for solving the PDP requests to further evaluate the efficacy of the proposed explicit evolutionary multitasking algorithm. In particular, for simplicity, here we keep the configurations of search operators and parameters consistent with those used in Section IV. Moreover, for the four PDP requests, the pyramid match kernel method [49] is used to pair the PDP requests based on the customer distributions. We then have two pairs in this article, which are {PDP request 1, PDP request 2}, and {PDP request 3, PDP request 4}.

The convergence graphs obtained by the proposed EEMTA, the implicit multitasking PMFEA, and the corresponding single-task EAs, that is, EA1 and EA2, on the PDP requests are presented in Fig. 9. As clearly revealed in the figure, both of the EMT algorithms, that is, EEMTA and PMFEA, obtained faster convergence speed against the single-task EAs. Further, it is observed that the knowledge transfer enabled by the proposed explicit multitasking EEMTA provides a strong impetus to the search process, speeding up the discovery of high-quality solutions by a substantial amount. In fact, with knowledge transferred from the paired PDP, the EEMTA is found to receive a significant boost during the initial stages of evolution itself, on PDP request 2 and PDP request 4 [i.e., Fig. 9(a) and (b)], enabling it to quickly achieve high-quality routing solutions while consuming considerably lower computational cost. To demonstrate, it takes EEMTA about five generations on average to attain the routing solution with travel cost significantly lower than that obtained by EA2 over 35 generations. On the other hand, although the knowledge transfer for PDP request 1 and PDP request 2 is from the weak solver EA2, EEMTA is also observed to obtain the solutions achieved by EA1 at generation 35, using around 20 number of generations. It is not hard to imagine that the significant savings of computational cost caused by the proposed EEMTA can play a vital role toward cutting down of optimization time, especially, when faced with real-world complex combinatorial optimization problems, where computational budget available is limited.

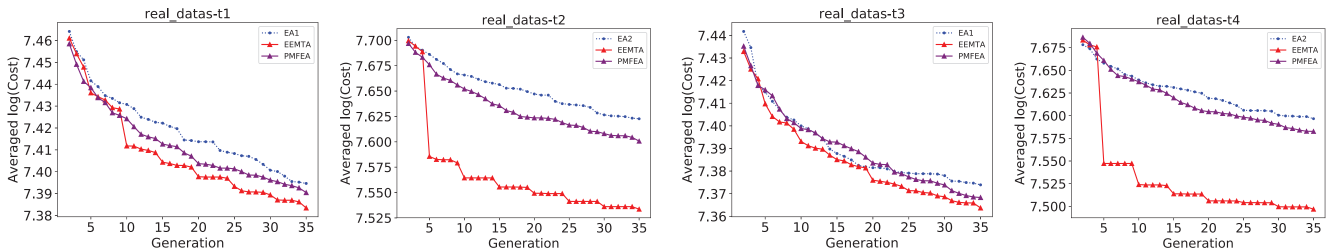


Fig. 9. Convergence traces of EEMTA, PMFEA and the single-task EAs on the real world PDP requests. y-axis: Log(Averaged travel cost); x-axis: Generation. (a) Convergence traces of paired PDP Request-1 and PDP Request-2. (b) Convergence traces of paired PDP Request-3 and PDP Request-4.

VI. Conclusion

In this article, we have proposed an explicit EMT algorithm, that is, EEMTA, for combinatorial optimization. In particular, by employing CVRP as the illustrating combinatorial optimization problem domain, we have first derived a weighted l_1-norm-regularized formulation to learn the sparse mapping between customers across CVRPs. Further, we have proposed to learn the new representation of customers based on a distance matrix derived from the optimized CVRP solutions, so that the useful traits buried in the optimized CVRP solutions can be transferred across CVRPs via the learned customer mapping, using simple clustering and pairwise distance sorting processes. Finally, comprehensive empirical studies on both multitasking CVRP benchmarks and the real-world PDP, against the state-of-the-art EMT algorithm and the traditional single-task evolutionary solvers, confirmed the efficacy of the proposed EEMTA for combinatorial optimization.

In the future, first, we would like to further extend the proposed EEMTA for multitask optimization in cases having more than two tasks, by conducting deeper research studies on the similarity measure between tasks. Further, we would also like to apply the proposed EEMTA for other combinatorial optimization problem domains, such as knapsack problem, scheduling problem, feature selection, etc. The key design issues could be the configuration of proper representation of $\mathbf{P}_s$ and $\mathbf{P}_t$ in (2) and the distance matrix $\mathbf{DM}$, and the design of particular operation to decode the representation as problem solution in knowledge transfer, as discussed in Section III-B3.

References

[1] A. Gupta, Y.-S. Ong, and L. Feng. "Multifactorial evolution: Toward evolutionary multitasking," *IEEE Trans. Evol. Comput.*, vol. 20, no. 3, pp. 343–357, Jun. 2016.

[2] Y.-S. Ong and A. Gupta, "Evolutionary multitasking: A computer science view of cognitive multitasking," *Cogn. Comput.*, vol. 8, no. 2, pp. 125–142, 2016.

[3] J. Ding, C. Yang, Y. Jin, and T. Chai, "Generalized multitasking for evolutionary optimization of expensive problems," *IEEE Trans. Evol. Comput.*, vol. 23, no. 1, pp. 44–58, Feb. 2019.

[4] J. Tang, Y. Chen, Z. Deng, Y. Xiang, and C. Paul Joy, "A group-based approach to improve multifactorial evolutionary algorithm," in *Proc. Int. Joint Conf. Artif. Intell. (IJCAI)*, 2018, pp. 3870–3876.

[5] R.-T. Liaw and C.-K. Ting, "Evolutionary manytasking optimization based on symbiosis in biocoenosis," in *Proc. 33rd AAAI Conf. Artif. Intell. (AAAI)*, 2019, pp. 4295–4303.

[6] K. K. Bali, A. Gupta, L. Feng, Y. S. Ong, and T. P. Siew, "Linearized domain adaptation in evolutionary multitasking," in *Proc. IEEE Congr. Evol. Comput. (CEC)*, 2017, pp. 1295–1302.

[7] J. Zhong, L. Feng, W. Cai, and Y.-S. Ong, "Multifactorial genetic programming for symbolic regression problems," *IEEE Trans. Syst., Man, Cybern., Syst.*, to be published.

[8] C. Yang, J. Ding, Y. Jin, C. Wang, and T. Chai, "Multitasking multiobjective evolutionary operational indices optimization of beneficiation processes," *IEEE Trans. Autom. Sci. Eng.*, vol. 16, no. 3, pp. 1046–1057, Jul. 2019.

[9] J. Luo, A. Gupta, Y.-S. Ong, and Z. Wang, "Evolutionary optimization of expensive multiobjective problems with co-sub-Pareto front Gaussian process surrogates," *IEEE Trans. Cybern.*, vol. 49, no. 5, pp. 1708–1721, May 2019.

[10] A. Gupta, Y.-S. Ong, L. Feng, and K. C. Tan, "Multiobjective multifactorial optimization in evolutionary multitasking," *IEEE Trans. Cybern.*, vol. 47, no. 7, pp. 1652–1665, Jul. 2017.

[11] L. Feng *et al.*, "Evolutionary multitasking via explicit autoencoding," *IEEE Trans. Cybern.*, vol. 49, no. 9, pp. 3457–3470, Sep. 2018.

[12] D. Kobler, *Evolutionary Algorithms in Combinatorial Optimization*. Boston, MA, USA: Springer, 2009, pp. 950–959.

[13] J. Puchinger and G. R. Raidl, "Combining metaheuristics and exact algorithms in combinatorial optimization: A survey and classification," in *Artificial Intelligence and Knowledge Engineering Applications: A Bioinspired Approach*, J. Mira and J. R. Álvarez, Eds. Heidelberg, Germany: Springer, 2005, pp. 41–53.

[14] J. Leung, L. Kelly, and J. H. Anderson, *Handbook of Scheduling: Algorithms, Models, and Performance Analysis*. Boca Raton, FL, USA: CRC, 2004.

[15] G. Laporte, "Fifty years of vehicle routing," *Transp. Sci.*, vol. 43, no. 4, pp. 408–416, 2009.

[16] L. Zhou, L. Feng, J. Zhong, Y.-S. Ong, Z. Zhu, and E. Sha, "Evolutionary multitasking in combinatorial search spaces: A case study in capacitated vehicle routing problem," in *Proc. IEEE Symp. Series Comput. Intell. (SSCI)*, 2016, pp. 1–8.

[17] R.-T. Liaw and C.-K. Ting, "Evolutionary many-tasking based on biocoenosis through symbiosis: A framework and benchmark problems," in *Proc. IEEE Congr. Evol. Comput. (CEC)*, 2017, pp. 2266–2273.

[18] Y.-W. Wen and C.-K. Ting, "Parting ways and reallocating resources in evolutionary multitasking," in *Proc. IEEE Congr. Evol. Comput. (CEC)*, 2017, pp. 2404–2411.

[19] H. Li, Y.-S. Ong, M. Gong, and Z. Wang, "Evolutionary multitasking sparse reconstruction: Framework and case study," *IEEE Trans. Evol. Comput.*, vol. 23, no. 5, pp. 733–747, Oct. 2019.

[20] F. Rothlauf, D. E. Goldberg, and A. Heinzl, "Network random key—A tree network representation scheme for genetic and evolutionary algorithms," *Evol. Comput.*, vol. 10, no. 1, pp. 75–97, 2002.

[21] G. R. Raidl and B. A. Julstrom, "Edge sets: An effective evolutionary coding of spanning trees," *IEEE Trans. Evol. Comput.*, vol. 7, no. 3, pp. 225–239, Jun. 2003.

[22] E. G. Carrano, C. M. Fonseca, R. H. C. Takahashi, L. C. A. Pimenta, and O. M. Neto, "A preliminary comparison of tree encoding schemes for evolutionary algorithms," in *Proc. IEEE Int. Conf. Syst. Man Cybern.*, 2007, pp. 1–8.

[23] J. Wang, Y. Zhou, Y. Wang, J. Zhang, C. L. P. Chen, and Z. Zheng, "Multiobjective vehicle routing problems with simultaneous delivery and pickup and time windows: Formulation, instances, and algorithms," *IEEE Trans. Cybern.*, vol. 46, no. 3, pp. 582–594, Mar. 2016.

[24] L. Zhang, H. Pan, Y. Su, X. Zhang, and Y. Niu, "A mixed representation-based multiobjective evolutionary algorithm for overlapping community detection," *IEEE Trans. Cybern.*, vol. 47, no. 9, pp. 2703–2716, Sep. 2017.

[25] X. Zhang, Y. Zhou, Q. Zhang, V. C. S. Lee, and M. Li, "Problem specific MOEA/D for barrier coverage with wireless sensors," *IEEE Trans. Cybern.*, vol. 47, no. 11, pp. 3854–3865, Nov. 2017.

[26] J. Wang, T. Weng, and Q. Zhang, "A two-stage multiobjective evolutionary algorithm for multiobjective multidepot vehicle routing problem with time windows," *IEEE Trans. Cybern.*, vol. 49, no. 7, pp. 2467–2478, Jul. 2019.

[27] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Trans. Knowl. Data Eng.*, vol. 22, no. 10, pp. 1345–1359, Oct. 2010.

[28] Y. C. Jin, *Knowledge Incorporation in Evolutionary Computation* (Studies in Fuzziness and Soft Computing). Heidelberg, Germany: Springer, 2010.

[29] E. Haslam, B. Xue, and M. Zhang, "Further investigation on genetic programming with transfer learning for symbolic regression," in *Proc. IEEE Congr. Evol. Comput. (CEC)*, 2016, pp. 3598–3605.

[30] M. Iqbal, B. Xue, H. Al-Sahaf, and M. Zhang, "Cross-domain reuse of extracted knowledge in genetic programming for image classification," *IEEE Trans. Evol. Comput.*, vol. 21, no. 4, pp. 569–587, Aug. 2017.

[31] X. Zhang, Y. Zhuang, W. Wang, and W. Pedrycz, "Transfer boosting with synthetic instances for class imbalanced object recognition," *IEEE Trans. Cybern.*, vol. 48, no. 1, pp. 357–370, Jan. 2018.

[32] L. Zhou, P. Yang, C. Chen, and Y. Gao, "Multiagent reinforcement learning with sparse interactions by negotiation and knowledge transfer," *IEEE Trans. Cybern.*, vol. 47, no. 5, pp. 1238–1250, May 2017.

[33] C. Yang, J. Ding, Y. Jin, and T. Chai, "Off-line data-driven multiobjective optimization: Knowledge transfer between surrogates and generation of final solutions," *IEEE Trans. Evol. Comput.*, to be published.

[34] G. B. Dantzig and J. H. Ramser, "The truck dispatching problem," *Manag. Sci.*, vol. 6, no. 1, pp. 80–91, 1959.

[35] A. N. Letchford and J.-J. Salazar-González, "The capacitated vehicle routing problem: Stronger bounds in pseudo-polynomial time," *Eur. J. Oper. Res.*, vol. 272, no. 1, pp. 24–31, 2019.

[36] G. Laporte and F. Semet, "Classical heuristics for the capacitated VRP," in *The Vehicle Routing Problem*. Philadelphia, PA, USA: Soc. Ind. Appl. Math., 2001, pp. 109–128.

[37] T. K. Ralphs, L. Kopman, W. R. Pulleyblank, and L. E. Trotter, "On the capacitated vehicle routing problem," *Math. Program.*, vol. 94, no. 2, pp. 343–359, 2003.

[38] C. Chen, Y. Ding, X. Xie, S. Zhang, Z. Wang, and L. Feng, "TrajCompressor: An online map-matching-based trajectory compression framework leveraging vehicle heading direction and change," *IEEE Trans. Intell. Transp. Syst.*, to be published.

[39] D. Xu, Y. Huang, Z. Zeng, and X. Xu, "Human gait recognition using patch distribution feature and locality-constrained group sparse representation," *IEEE Trans. Image Process.*, vol. 21, no. 1, pp. 316–326, Jan. 2012.

[40] K. Koh, S.-J. Kim, and S. Boyd, "An interior-point method for large-scale l_1-regularized logistic regression," *J. Mach. Learn. Res.*, vol. 8, pp. 1519–1555, Dec. 2007.

[41] Y. Hou, Y.-S. Ong, L. Feng, and J. M. Zurada, "An evolutionary transfer reinforcement learning framework for multiagent systems," *IEEE Trans. Evol. Comput.*, vol. 21, no. 4, pp. 601–615, Aug. 2017.

[42] M. Jiang, Z. Huang, L. Qiu, W. Huang, and G. G. Yen, "Transfer learning-based dynamic multiobjective optimization algorithms," *IEEE Trans. Evol. Comput.*, vol. 22, no. 4, pp. 501–514, Aug. 2018.

[43] B. Tan, Y. Song, E. Zhong, and Q. Yang, "Transitive transfer learning," in *Proc. 21st ACM SIGKDD Int. Conf. Knowl. Disc. Data Min. (KDD)*, 2015, pp. 1155–1164.

[44] J. B. Kruskal, "Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis," *Psychometrika*, vol. 29, no. 1, pp. 1–27, 1964.

[45] B. S. Da *et al.*, "Evolutionary multitasking for single-objective continuous optimization: Benchmark problems, performance metrics and baseline results," Nanyang Technol. Univ., Singapore, Rep., 2016.

[46] Y. Yuan *et al.*, "Evolutionary multitasking for multiobjective continuous optimization: Benchmark problems, performance metrics and baseline results," Nanyang Technol. Univ., Singapore, Rep., 2016.

[47] C. Prins, "A simple and effective evolutionary algorithm for the vehicle routing problem," *Comput. Oper. Res.*, vol. 31, no. 12, pp. 1985–2002, 2004.

[48] T. A. M. Toffolo, T. Vidal, and T. Wauters, "Heuristics for vehicle routing problems: Sequence or set optimization?" *Comput. Oper. Res.*, vol. 105, pp. 118–131, May 2019.

[49] K. Grauman and T. Darrell, "The pyramid match kernel: Efficient learning with sets of features," *J. Mach. Learn. Res.*, vol. 8, pp. 725–760, May 2007.

Liang Feng received the Ph.D. degree from the School of Computer Engineering, Nanyang Technological University, Singapore, in 2014.

He was a Postdoctoral Research Fellow with the Computational Intelligence Graduate Laboratory, Nanyang Technological University. He is currently an Assistant Professor with the College of Computer Science, Chongqing University, Chongqing, China. His research interests include computational and artificial intelligence, memetic computing, big data optimization and learning, as well as transfer learning.

Dr. Feng's research work on evolutionary multitasking won the 2019 IEEE TRANSACTIONS ON EVOLUTIONARY COMPUTATION Outstanding Paper Award. He is an Associate Editor of the *IEEE Computational Intelligence Magazine*, *Memetic Computing*, and *Cognitive Computation*. He is also the Founding Chair of the IEEE CIS Intelligent Systems Applications Technical Committee Task Force on "Transfer Learning & Transfer Optimization."

Yuxiao Huang received the B.E. degree from the Department of Computer Science, Chongqing University, Chongqing, China, in 2018, where he is currently pursuing the Ph.D. degree with the Department of Computer Science.

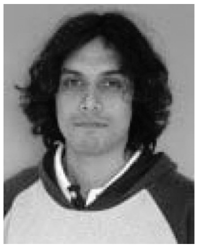
His current research interests include transfer optimization, multitask optimization, transfer learning, and GPU-based computing.

Lei Zhou received the B.E. degree from the School of Computer Science and Technology, Shandong University, Jinan, China, in 2014, and the Ph.D. degree from the College of Computer Science, Chongqing University, Chongqing, China, in 2019.

He is currently a Postdoctoral Research Fellow with the School of Computer Engineering, Nanyang Technological University, Singapore. His research interests include evolutionary computations, memetic computing, as well as transfer learning and optimization.

Jinghui Zhong received the Ph.D. degree from the School of Information Science and Technology, Sun Yat-sen University, Guangzhou, China, in 2012.

He is currently an Associate Professor with the School of Computer Science and Engineering, South China University of Technology, Guangzhou, China. From 2013 to 2016, he was a Postdoctoral Research Fellow with the School of Computer Engineering, Nanyang Technological University, Singapore. His research interests include evolutionary computation (e.g., genetic programming and differential evolution), machine learning, and agent-based modeling.



Abhishek Gupta received the Ph.D. degree in engineering science from the University of Auckland, Auckland, New Zealand, in 2014.

Over the past five years, he has been working in the area of memetic computation, with a particular focus on developing novel theories and algorithms in the topics of evolutionary transfer and multitask optimization. He is currently appointed as a Scientist with the Singapore Institute of Manufacturing Technology, Agency for Science, Technology and Research, Singapore. He also jointly serves as an Adjunct Research Scientist with the Data Science and Artificial Intelligence Research Center, School of Computer Science and Engineering, Nanyang Technological University, Singapore.

Dr. Gupta's pioneering work on evolutionary multitasking, in particular, was bestowed the 2019 IEEE TRANSACTIONS ON EVOLUTIONARY COMPUTATION Outstanding Paper Award by the IEEE Computational Intelligence Society (CIS). He is an Associate Editor of the IEEE TRANSACTIONS ON EMERGING TOPICS IN COMPUTATIONAL INTELLIGENCE, and also the Founding Chair of the IEEE CIS Emergent Technology Technical Committee Task Force on Multitask Learning and Multitask Optimization.

Ke Tang received the B.Eng. degree from the Huazhong University of Science and Technology, Wuhan, China, in 2002, and the Ph.D. degree from Nanyang Technological University, Singapore, in 2007.

From 2007 to 2017, he was with the School of Computer Science and Technology, University of Science and Technology of China, Hefei, China, first as an Associate Professor from 2007 to 2011 and later as a Professor from 2011 to 2017. He is currently a Professor with the Department of Computer Science and Engineering, Southern University of Science and Technology, Shenzhen, China. He has over 7000 Google Scholar citations with an H-index of 41. He has published over 70 journal papers and over 80 conference papers. His current research interests include evolutionary computation, machine learning, and their applications.

Prof. Tang was a recipient of the Royal Society Newton Advanced Fellowship in 2015 and the 2018 IEEE Computational Intelligence Society Outstanding Early Career Award. He is an Associate Editor of the IEEE TRANSACTIONS ON EVOLUTIONARY COMPUTATION and served as a member of editorial boards for a few other journals.

Kay Chen Tan received the B.Eng. degree (First Class Hons.) in electronics and electrical engineering and the Ph.D. degree from the University of Glasgow, Glasgow, U.K., in 1994 and 1997, respectively.

He is a Full Professor with the Department of Computer Science, City University of Hong Kong, Hong Kong. He has published over 200 refereed articles and 6 books.

Prof. Tan is the Editor-in-Chief of the IEEE TRANSACTIONS ON EVOLUTIONARY COMPUTATION, and was the Editor-in-Chief of the *IEEE Computational Intelligence Magazine* from 2010 to 2013. He is currently serves as the Editorial Board Member of over 10 journals. He is an Elected Member of the IEEE CIS AdCom from 2017 to 2019.