

# Bridging Synthetic and Real Routing Problems via LLM-Guided Instance Generation and Progressive Adaptation

Jianghan Zhu<sup>1</sup>, Yaixin Wu<sup>2</sup>, Zhuoyi Lin<sup>3\*</sup>, Zhengyuan Zhang<sup>4</sup>, Haiyan Yin<sup>5</sup>, Zhiguang Cao<sup>1</sup>, Senthilnath Jayavelu<sup>3, 7</sup>, Xiaoli Li<sup>4, 6</sup>

<sup>1</sup> Singapore Management University, Singapore

<sup>2</sup> Eindhoven University of Technology, Netherlands

<sup>3</sup> Institute for Infocomm Research, Agency for Science, Technology and Research (A\*STAR), Singapore

<sup>4</sup> Nanyang Technological University, Singapore

<sup>5</sup> Centre for Frontier AI Research (CFAR), Agency for Science, Technology and Research (A\*STAR), Singapore

<sup>6</sup> Singapore University of Technology and Design, Singapore

<sup>7</sup> National University of Singapore, Singapore

{zhuj0044, zhengyua002}@e.ntu.edu.sg, y.wu2@tue.nl, {Lin\_Zhuoyi, Yin\_Haiyan, J\_Senthilnath}@a-star.edu.sg, zgcao@smu.edu.sg, xiaoli\_li@sutd.edu.sg

## Abstract

Recent advances in Neural Combinatorial Optimization (NCO) methods have significantly improved the capability of neural solvers to handle synthetic routing instances. Nonetheless, existing neural solvers typically struggle to generalize effectively from synthetic, uniformly-distributed training data to real-world VRP scenarios, including widely recognized benchmark instances from TSPLib and CVRPLib. To bridge this generalization gap, we present Evolutionary Realistic Instance Synthesis (**EvoReal**), which leverages an evolutionary module guided by large language models (LLMs) to generate synthetic instances characterized by diverse and realistic structural patterns. Specifically, the evolutionary module produces synthetic instances whose structural attributes statistically mimics those observed in authentic real-world instances. Subsequently, pre-trained NCO models are progressively refined, firstly aligning them with these structurally enriched synthetic distributions and then further adapting them through direct fine-tuning on actual benchmark instances. Extensive experimental evaluations demonstrate that EvoReal markedly improves the generalization capabilities of state-of-the-art neural solvers, yielding a notable reduced performance gap compared to the optimal solutions on the TSPLib (1.05%) and CVRPLib (2.71%) benchmarks across a broad spectrum of problem scales.

**Code** — <https://github.com/HenryZhu1029/EvoReal>

## 1 Introduction

Being a longstanding NP-hard challenge, Vehicle Routing Problems (VRPs) represent a classic family of combinatorial optimization problems (COPs) (Bengio, Lodi, and Prouvost 2021; Cappart et al. 2023) widely encountered in the field of logistics (Cattaruzza et al. 2017) and public transportation (Konstantakopoulos, Gayialis, and Kechagias 2022). In recent years, Neural Combinatorial Optimization

(NCO) methods have demonstrated remarkable success in tackling various VRPs through deep reinforcement learning and attention-based models, achieving state-of-the-art performance on synthetic VRP instances (Bello et al. 2017; Kool, van Hoof, and Welling 2019; Li, Yan, and Wu 2021; Drakulic et al. 2023; Ma, Cao, and Chee 2023). Despite this progress, existing NCO models frequently exhibit limited generalization when transitioning from synthetic, uniformly distributed training data to real-world problem instances, such as those found in benchmarks like TSPLib (Reinelt 1991) and CVRPLib (Uchoa et al. 2017). This distributional shift significantly restricts their practical applicability (Joshi et al. 2006; Bi et al. 2022; Gao et al. 2024).

On the other hand, LLM solvers exhibit strong generalization capabilities when confronted with new COPs and distributions. Recent works leverage LLMs to dynamically generate heuristics or evolve hyper-heuristics (Duflo et al. 2019; Drake et al. 2020), thus facilitating efficient Evolutionary Search (ES) of heuristic spaces without human bias (Yang et al. 2024). Noteworthy endeavors include FunSearch (Romera-Paredes et al. 2024), EoH (Liu et al. 2024a), and ReEvo (Ye et al. 2024). However, existing LLM-based approaches often encounter challenges when handling tasks that require extensive contextual information, typically demonstrating reduced coherence and accuracy in the processing of extended textual descriptions (Yang et al. 2024; Xu et al. 2025). Consequently, these approaches exhibit significant limitations when addressing medium- or large-size instances (e.g., instances involving more than 50 nodes), thus compromising their ability to effectively generalize to real-world VRP instances.

Motivated by these observations, we propose EvoReal, a novel data-centric framework explicitly designed to enhance the generalization capabilities of NCO models for real-world VRPs. Unlike prior approaches that evolve heuristics or solvers for direct solution construction, we innovatively leverage LLMs to evolve data generators that facilitate model adaptation. At the core of EvoReal is an LLM-driven evolutionary module, which generates synthetic VRP

\*Zhuoyi Lin is the corresponding author.

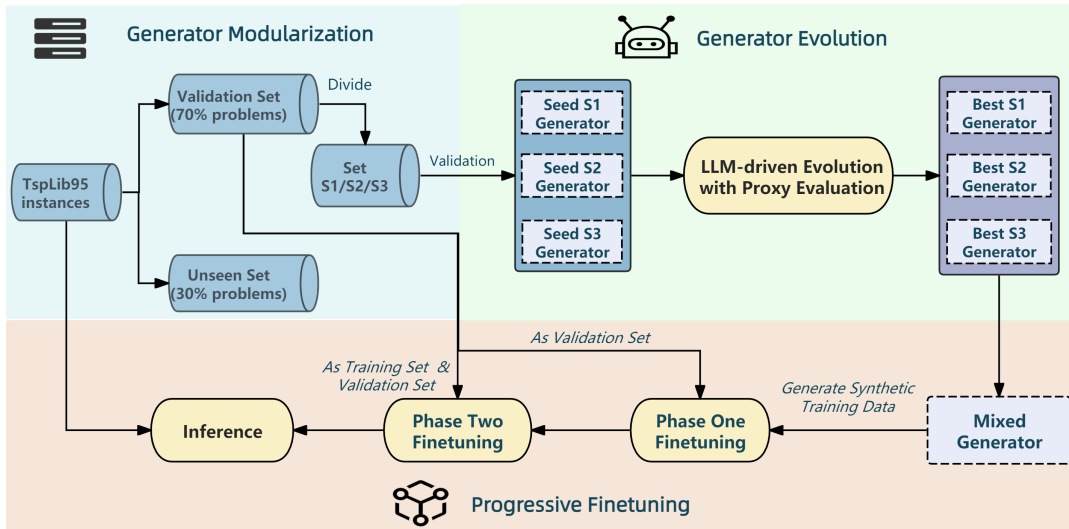


Figure 1: Overall workflow of EvoReal including LLM-guided generator evolution and progressive fine-tuning. **Top Left:** Validation set and unseen test set are split, with validation problems grouped by distribution category for structural-specific generator design. **Top Right:** LLM-driven module evolves generators which are evaluated on specific validation sets. **Bottom:** Pre-trained models are progressively fine-tuned on data from the evolved generators and the validation set’s real instances.

instances structurally aligned with real-world distributions. This module systematically *evolves simplistic synthetic scenarios into diverse and complex distributions*, effectively capturing the intricate characteristics of real-world datasets. Subsequently, we introduce a progressive fine-tuning strategy which incrementally adapts pre-trained neural solvers by transitioning through increasingly complex synthetic instances toward actual real-world VRP scenarios. Consequently, EvoReal *facilitates smoother model adaptation by progressively refining solver parameters and representations*, effectively bridging the generalization gap between synthetic data and real-world benchmark instances. We summarize our contributions as follows. (1) We propose EvoReal, an LLM-guided evolutionary framework to synthesize structurally realistic VRP instances, bridging the distributional gap between synthetic and real-world routing problems. (2) We introduce a progressive fine-tuning strategy, which first adapts neural combinatorial solvers to diverse LLM-evolved distributions and then further refines them on real benchmark data, enabling effective domain adaptation without architectural changes. (3) Extensive experiments on TSPLib and CVRPLib benchmarks demonstrate that EvoReal significantly improves the generalization of SOTA neural solvers across a wide range of problem sizes, achieving new state-of-the-art results and notably reducing the performance gap between small and large instances. Notably, our results demonstrate that this generator-based adaptation consistently outperforms direct fine-tuning on real data alone.

## 2 Related Works

**Constructive Neural VRP Solvers.** Those models learn policies to construct solutions in a step-by-step or one-shot manner: Pointer Network (Ptr-Net) proposed by (Vinyals, Fortunato, and Jaitly 2015) employed a transformer-based

encoder-decoder network that generates solutions to VRPs in a sequential way. Follow-up works utilize reinforcement learning (RL) to improve Ptr-Net’s gradient update to generate better approximate solutions on TSP (Bello et al. 2017) and CVRP (Nazari et al. 2018). More recently, more neural VRP models have evolved with their architectures built upon attention mechanisms, starting with the attention-based model (AM) (Kool, van Hoof, and Welling 2019), which is versatile on a wide range of COPs. Policy optimization with multiple optima (POMO) (Kwon et al. 2020), has further promoted the capability of the attention-based model on TSP and CVRP. Other constructive frameworks developed from AM and POMO illustrate their broader scalability (Kwon et al. 2021; Choo et al. 2022; Kim, Park, and Park 2022; Chen et al. 2023; Chalumeau et al. 2023; Hottung, Mahajan, and Tierney 2024; Lin et al. 2024).

Recent studies have investigated the severe generalization ability decrement in unseen problem sizes or distributions (Joshi et al. 2006; Liu et al. 2023). In the line of size generalization, there are many attempts to generalize solvers from small instances to larger ones (Lisicki, Afkanpour, and Taylor 2020; Kim et al. 2022; Bdeir, Falkner, and Schmidt-Thieme 2022; Hou et al. 2023; Son et al. 2023). Unlike approaches focused on scaling, considerable efforts have been devoted to addressing cross-distribution generalization, including works that augment training instances on multi-distribution and multitasks (Goh et al. 2025), and works which synergetically train a backbone model with multi-distribution instances towards a generalizable solver on out-of-distribution tasks (Zhou et al. 2023, 2024). Recently, training LLM for end-to-end CO has been attempted, moving beyond prompting paradigms (Jiang et al. 2025a).

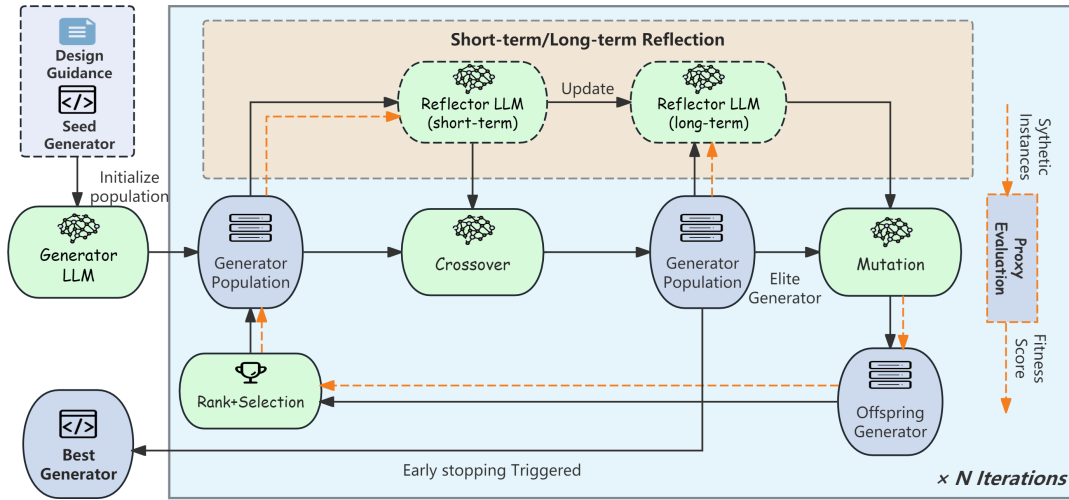


Figure 2: LLM-driven evolution component in Fig.1. The pipeline within the blue rectangle block is repeated for  $N$  iterations. Dotted arrows represent the proxy-evaluation of each generator; black arrows indicate the flow of generators. For each pair of parents, the reflector LLM performs short-term reflection based on their relative performance, and this insight is used to guide crossover for designing new offspring. Accumulated short-term reflections are further distilled into long-term ones, which guide mutation to improve the current best generator. After mutation, populations are ranked and selected to maintain a fixed size.

**Learning to Optimize with LLMs.** Existing approaches to improving generalization in combinatorial optimization (CO) take advantage of the linguistic comprehension and domain-specific competence of LLMs via prompt engineering to tackle COPs. These methods can be mainly divided into two mainstreams: 1) *LLMs are solution generators*. Early explorations in this line of research involve prompting LLMs to solve graph-based COPs (Wang et al. 2023), iteratively refining current heuristics (Iklassov et al. 2024), progressively improving solutions by targeting better objectives (Yang et al. 2024; Liu et al. 2024b; Elhenawy et al. 2024b), and instilling LLMs with visual modalities of problems and solutions (Elhenawy et al. 2024a; Huang et al. 2024). 2) *LLMs are heuristic designers*. Explainable, task-oriented heuristics in the form of codes are iteratively improved throughout automatic heuristic generation and evaluation of Evolution Algorithms (EA), the productivity of which outstrips that of human experts (Romera-Paredes et al. 2024; Liu et al. 2024a; Ye et al. 2024; Tran et al. 2025; Jiang et al. 2025b). Among them, ReEvo (Ye et al. 2024) achieves strong adaptability in many COPs under both white-box and black-box prompt settings. In this paper, we focus mainly on the utilization of LLMs as code designers. Distinct from prior constructive neural solvers and LLM-as-optimizer approaches, our method leverages LLMs to evolve data generators, enabling a novel, generator-based adaptation pipeline that outperforms direct fine-tuning on real data.

### 3 Methodology

In this section, we propose a novel LLM-guided evolution framework (EvoReal) that evolves VRP data generators to generate VRPLib-style distributions and a progressive fine-tuning strategy that gradually shifts the generalization ability of the pre-trained neural model on synthetic uni-

form instances towards VRPLib. The proposed LLM-guided evolution approach is primarily comprised of two components: LLM-driven evolutionary search and proxy evaluation. Specifically, the **LLM-driven evolution** aims at gradually developing structurally specific generators, while in the **proxy evaluation**, performance of each LLM-designed generator is accessed throughout training and monitoring the progressive validation results. In the following part, by taking EvoReal for TSP as an example, we elaborate the generator evolution and progressive fine-tuning framework.

**Generator Modularization.** Fig.1 shows the entire pipeline of our framework. For TSP, the evolution of data generators is driven by various carefully-written prompt strategies crafted for three types of generator, which generate different distributions of TSP. This ensures the coverage of generated TSP distributions that emulate TSPLib-like distributions. We combine the three evolved generators to produce mixed-distribution TSP data for fine-tuning. To ensure a wide range of problem sizes and distributions, we select 70 TSPLib problems of size less than 5000, and then carefully pick 48 problems ( $\approx 70\%$ ) for validation while leaving the remaining 22 problems as a held-out unseen set. Among the 48 validation problems, we further categorize them into three types (i.e., S1, S2 and S3), each serving as a specific validation set for one type of generator. Detailed categorization methods are provided in the extended version. For CVRP, we randomly select 70% problems from CVRPLib (SetX) for validation and direct fine-tuning, and the remaining 30% are left as the unseen set. The specifics of the generation of CVRP are discussed in the extended version.

### Generator Evolution

We prompt the LLMs to transform latent generator configurations into realistic routing distributions. Formally, let

$h_\phi : \mathcal{Z} \rightarrow \mathcal{X}$  represent a data generator parameterized by an LLM model  $\phi$ , mapping a latent variable  $z \sim \mathcal{Z}$  (e.g. seed generator) into a synthetic routing instance  $x = h_\phi(z)$ . Throughout the evolution, we iteratively minimize the divergence between the synthetic distribution of  $h_\phi(\mathcal{Z})$  and the target real-world distribution  $\mathcal{D}_{\text{real}}$ :

$$\min_{\phi} \mathcal{L}_{\text{evolve}}(\phi) = D(\mathcal{D}_{\text{real}}, \mathcal{D}(h_\phi(\mathcal{Z}))) \quad (1)$$

where  $D(\cdot \| \cdot)$  is a divergence measure (e.g., KL divergence, average gap). This dual-level optimization framework explicitly captures how heuristic generation indirectly optimizes the underlying combinatorial optimization problem by aligning evolved heuristics with real-world structures.

**Representation of Generators.** Each generator  $h$  is described by three components: 1) a function description defines the format of input and describes valid output, 2) a code implementing generation, and 3) a fitness score. The fitness score, defined as a function  $f$  of  $h$  ( $f : h \rightarrow \mathbb{R}$ ), serves as the performance indicator of  $h$  on the validation set. An example of an evolved generator that outputs TSP problems is given in Fig.7 in section E of the extended version.

**Proxy Evaluation.** Following recent work (Ramji et al. 2024), we use an approach that leverages external proxy metrics such as gaps or objective values as the divergence measure  $\mathcal{D}$  in Eq.1 to evaluate the quality of LLM-generated outputs, in lieu of direct human or reference-based evaluation. During the LLM-driven evolution, the proxy evaluation is instantiated by measuring the average gap on the validation set, which serves as the proxy metric. The gap is the gap between the objective value found and the best-known optimum. Each generator only fine-tunes the model for a small number of epochs (e.g. 10 epochs for LEHD), and the best average gap on the validation set of each generator during training is taken as the fitness score. Generators with lower fitness scores are retained according to a fixed population size, guiding generator selection without full fine-tuning. This low-fidelity evaluation strategy substantially accelerates the evolution search while presenting reliable indications of the proximity of the generated distributions to the real ones, as is the standard in the neural architecture search and hyperparameter optimization literature.

**Population Initialization.** Our evolutionary search framework begins with population initialization, during which a population of  $N$  generators  $h_1, h_2, \dots, h_N$  is initialized by incorporating initialization prompts. These prompts instill LLMs with the prior knowledge about the distributions that they are encouraged to explore and reproduce. Those prompts are composed of a trivial blueprint of the seed generator and a design guidance, each as a hint for generating possible structures of satisfactory distributions. Ablation studies in Section 5 show that external knowledge plays a significant role in improving the performance of generators.

**Evolution Process.** In order to balance the effectiveness of evolutionary search and the efficiency of fitness evaluation, we incorporate targeted modifications into ReEvo (Ye et al. 2024) framework. Ours introduces a ranking-based selection mechanism that increases the survival rate of elite generators, thereby accelerating the search process and reducing

token costs. In addition, by postponing the selection stage until after mutation, we balance exploration and exploitation, allowing more newly initialized generators to participate in crossover and mutation before selection. Table 5 in Section 5 shows the elevated performance of the evolution framework equipped with these modifications. We perform the workflow in Fig.2 to search for the best generator by repeating the evolution part for  $N$  iterations. The whole evolution procedure can be summarized as the following steps:

- Step 1. Choose a specific type of generator to evolve. Initialize the code population by requesting LLMs to devise  $N$  generators that align with the initialization prompts.
- Step 2. Expand the generator population by four prompt strategies (i.e., crossover, mutation, short- & long-term reflection) as exploration and modification operators.
- Step 3. Each new generator is then evaluated on their structurally-specific validation set. Then it is added to the population if both the code and the fitness score are valid.
- Step 4. Select  $k$  offspring generators based on the rank of each generator individual. Ranks are calculated from fitness scores and transformed into the probability of being selected. The lower the fitness score, the more likely the generator individual will be selected.
- Step 5. Repeat Steps 2 to 4.
- Step 6. Stop the iteration once the number of the generators evaluated reaches pre-set maximum or when no better generator can be found for consecutive  $m$  iterations.

Once the evolutions are completed for all types of generator, we save the best generators for subsequent progressive fine-tuning. Detailed explanations of prompt strategies and evolution mechanisms are listed in the extended version.

## Progressive Fine-tuning

After generator evolution, we progressively fine-tune neural models to real-world distributions. Ablation studies in Section 5 show that directly fine-tuning pre-trained neural solvers with TSPLib leads to limited generalization capability improvement compared to our progressive fine-tuning strategies. Through comprehensive experimental results, we demonstrate that our progressive fine-tuning framework effectively bridges the generalization gap between vanilla synthetic distributions like uniform and heterogeneous real-world instances without modifying model architectures. We divide our progressive strategies into two phases.

**Phase One.** In the first phase of progressive fine-tuning, the best-performing generator of all types evolved collaboratively generate large-scale synthetic data of the same problem size. By training pre-trained models to the disparate, multi-size node patterns in these diverse, TSPLib-style instances, neural solvers are better equipped to handle challenging patterns in real-world scenarios through preliminary exposure to similar structures. Compared to the proxy evaluation conducted during LLM-driven generator evolution, phase one involves a greater number of training epochs. Additionally, we continuously assess model performance throughout phase one using the 48 validation problems.

Method	[0,200)		[200,500)		[500,1000)		[1000,5000)		Overall		Time
	Obj	Gap	Obj	Gap	Obj	Gap	Obj	Gap	Obj	Gap	
#Problems	27		15		6		22		70		
Concorde	30558.67	0.00%	42279.13	0.00%	29658.17	0.00%	158431.73	0.06%	73166.50	0.02%	8h 59min
LKH-3	30558.67	0.00%	42279.13	0.00%	29658.17	0.00%	158384.22	0.03%	73159.18	0.01%	2h 45min
ORTools	31072.05	1.68%	43678.57	3.31%	30752.55	3.69%	165319.37	4.41%	75390.32	3.06%	14h 46min
DIFUSCO (Ts=50)	30830.64	0.89%	43336.11	2.50%	30758.49	3.71%	-	-	-	-	-
SGBS	30803.14	0.80%	45187.94	6.88%	34750.47	17.17%	235035.03	48.44%	86670.34	18.48%	6h 29min
CNF (3)	30867.31	1.01%	46266.05	9.43%	38119.65	28.53%	251882.05	59.08%	90284.04	23.42%	1h 18min
BQ (greedy)	31270.68	2.33%	43623.61	3.18%	32122.76	8.31%	225740.66	42.57%	84614.77	15.67%	13min
BQ (bs16)	30907.04	1.14%	43200.82	2.18%	31295.30	5.52%	216493.80	36.73%	82603.09	12.92%	4h 7min
ELG (no aug)	31255.40	2.28%	44985.00	6.40%	32398.58	9.24%	178508.82	12.74%	78309.08	7.05%	7min
ELG ( $\times 8$ aug)	30910.09	1.15%	43961.84	3.98%	32247.32	8.73%	176323.77	11.36%	77263.00	5.62%	39min
POMO (no aug)	32318.85	5.76%	48481.48	14.67%	38104.81	28.48%	272069.99	71.83%	95375.41	30.38%	-
POMO ( $\times 8$ aug)	31973.53	4.63%	46701.53	10.46%	37461.23	26.31%	261002.25	64.84%	92654.16	26.66%	26min
LEHD (greedy)	31179.01	2.03%	43441.81	2.75%	30859.32	4.05%	176181.27	11.27%	76999.66	5.26%	6min
LEHD (RRC-50)	30659.51	0.33%	42558.18	0.66%	30289.89	2.13%	168581.11	6.47%	74966.04	2.48%	1h 51min
POMO (ours): no aug	30946.76	1.27%	43192.36	2.16%	31366.48	5.76%	213754.57	35.00%	82259.28	12.45%	-
POMO (ours): $\times 8$ aug	30907.04	1.14%	43192.36	2.16%	31366.48	5.76%	209685.32	32.43%	81630.17	11.59%	26min
LEHD (ours): greedy	30986.49	1.40%	43116.26	1.98%	30666.54	3.40%	166839.40	5.37%	75302.53	2.94%	6min
LEHD (ours): RRC-50	<b>30650.34</b>	<b>0.30%</b>	<b>42427.11</b>	<b>0.35%</b>	<b>29874.67</b>	<b>0.73%</b>	<b>162358.47</b>	<b>2.54%</b>	<b>73919.96</b>	<b>1.05%</b>	1h 52min

Table 1: Performance comparison on 70 TSPLIB problems. Each entry shows the range-wise average objective and optimality gap (%), along with the total inference time in the last column. The best results of all learning-based methods are in **bold**.

**Phase Two.** After aligning the neural model with a broader range of structural priors in phase one, we fine-tune the best-performed model saved in phase one with TSPLib instances and validate the model on themselves. This enables the model to shift from easy tasks (synthetic data) to accommodate to hard cases (real-world instances) seamlessly, forming a synthetic-real progression. By combining both phases together, our strategy not only enables models to generalize from synthetic to real distributions, but also goes beyond standard curriculum or domain adaptation by sequentially bridging both distributional and scale gaps.

## 4 Experimental Results

In this section, we empirically validate the effectiveness of EvoReal on two typical VRP benchmarks, TSPLib95 (Reinelt 1991) and CVRPLib setX (Uchoa et al. 2017), across a wide range of problem sizes.

### Experimental Settings

**Baselines.** We employ four classic solvers as non-neural baselines: Concorde (Applegate et al. 2006), ORTools, HGS (Vidal 2022) and LKH-3 (Helsgaun 2017). For deep neural solvers, POMO (Kwon et al. 2020), LEHD (Luo et al. 2023), SGBS (Choo et al. 2022), BQ (Drakulic et al. 2023), ELG (Gao et al. 2024), CNF (Zhou et al. 2024) and DIFUSCO (Sun and Yang 2023). Importantly, POMO is trained using reinforcement learning (RL) with policy gradient methods, whereas LEHD employs a supervised learning (SL) scheme

that leverages optimal objective values. These models sample solely uniform distributions for training.

**Training Setups.** Hyperparameters have two sets: one for proxy evaluation and the other for progressive fine-tuning. We choose two representative attention-based neural models for our study, namely POMO and LEHD. For both sets, we leverage the evolved generators to produce TSP or CVRP of size 100 for fine-tuning. In phase two, instances from VRPLib are expanded by a pre-set batch size as a single batch, and they are expanded only once in one epoch. For LEHD, Concorde solver and HGS are adopted to obtain the optimum of TSP and CVRP, respectively. For generator evolution, we use OpenAI model o3 to construct new generators. All evolution pipelines can be conducted on a single NVIDIA RTX 3090 Ti GPU with 24GB memory, whereas the progressive fine-tuning is performed on one NVIDIA RTX A5000 GPU with 40GB memory. The complete hyperparameters are given in section D of the extended version.

**Metric and Inference.** We compare the results with different metrics with respect to different problem size intervals, including average objective values, average gaps (to the optimal average tour length), and total inference time. In order to monitor the performance of models during the proxy evaluation and progressive fine-tuning, we evaluate the models’ performance at fixed epochs. At each validation point, we calculate the total average gap on the validation set as the performance metric with 8-fold augmentation for POMO or greedy mode for LEHD. The inference time of classic non-neural solvers (CPU) and learning-based

Method	[0,200)		[200,500)		[500,1002)		Overall		Time
	Obj	Gap	Obj	Gap	Obj	Gap	Obj	Gap	
Count	22		46		32		100		
HGS-CVRP	27222.86	0.01%	53334.09	0.06%	102118.28	0.24%	63176.43	0.11%	16h 40min
LKH-3	27315.41	0.35%	53845.79	1.02%	103310.20	1.41%	63738.08	1.00%	27h 29min
ORTools	27802.65	2.14%	55514.15	4.15%	106987.85	5.02%	65637.60	4.01%	8h 4min
CNF (3)	28964.95	6.41%	61244.12	14.90%	143774.47	41.13%	76624.53	21.42%	10min 34s
ELG (no aug)	28921.39	6.25%	57342.41	7.58%	112081.53	10.02%	68199.75	8.07%	2m 15s
ELG ( $\times 8$ aug)	28447.76	4.51%	56244.39	5.52%	109819.94	7.80%	66912.36	6.03%	5m 24s
POMO (no aug)	29822.38	9.56%	63685.36	19.48%	161266.20	58.30%	81862.41	29.72%	-
POMO ( $\times 8$ aug)	29057.50	6.75%	61270.77	14.95%	143794.84	41.15%	76693.95	21.53%	2m 17s
LEHD (greedy)	30309.62	11.35%	58339.16	9.45%	119946.19	17.74%	71008.01	12.52%	2m 18s
LEHD (RRC-50)	28556.65	4.91%	55823.30	4.73%	110298.74	8.27%	66830.32	5.90%	1h 10min
POMO (ours): no aug	28279.00	3.89%	55940.56	4.95%	108505.76	6.51%	66401.20	5.22%	-
POMO (ours): $\times 8$ aug	28178.29	3.52%	55754.01	4.60%	108189.96	6.20%	66180.32	4.87%	2m 6s
LEHD (ours): greedy	28782.57	5.74%	55700.70	4.50%	106172.85	4.22%	66060.42	4.68%	2m 36s
LEHD (ours): RRC50	<b>27916.97</b>	<b>2.56%</b>	<b>54682.63</b>	<b>2.59%</b>	<b>104929.99</b>	<b>3.00%</b>	<b>64817.21</b>	<b>2.71%</b>	1h 9min

Table 2: Performance comparison over 100 SetX problems, with average objective, optimality gap (%), and inference time. The best result of all learning-based methods under each problem size range are marked in **bold**.

POMO (aug $\times 8$ )	[0,200)	[200,1000)	[1000,5000)	Time
w/o Fine-tune	4.63%	14.99%	64.84%	-
w/o Phase 1	1.63%	12.06%	44.85%	7h 41min
w/o Phase 2	<b>1.10%</b>	11.14%	43.06%	5h 42min
<b>Full (ours)</b>	1.11%	<b>6.78%</b>	<b>36.48%</b>	6h 37min

Table 3: Comparison of different POMO fine-tuning setups. We use problem of nodes  $\leq 500$  with batch size 4 in TSPLib for finetuning. All experiments are trained for 600 epochs. The best result for each problem range is in **bold**.

LEHD - greedy	[0,200)	[200,1000)	[1000,5000)	Time
w/o Fine-tune	2.03%	3.12%	11.27%	-
w/o Phase 1	2.55%	3.24%	8.7%	31 min
w/o Phase 2	1.97%	<b>2.35%</b>	6.08%	35 min
<b>Full (ours)</b>	<b>1.40%</b>	2.39%	<b>5.37%</b>	38.5 min

Table 4: Comparison of different LEHD fine-tuning setups. We use problem of nodes  $\leq 500$  with batch size 4 in TSPLib for finetuning. All experiments are trained for 40 epochs. The best result for each problem range are in **bold**.

models (GPU) are not directly comparable. For POMO, LEHD, SGBS, CNF(3), ELG, DIFUSCO ( $T_s=50$ ) we report the performance of their pre-trained models with their default settings. For BQ, we reproduced their model with their original setups. For the classic non-neural solvers for TSP, we set a maximum time limit according to problem size  $n$  ( $T_{max} = 1.8 * n$  seconds for LKH-3 and ORTools, and  $T_{max} = 10$  minutes for Concorde). As for non-neural solvers for CVRP, we cite the results from Table 2 of HGS paper (Vidal 2022). For POMO, we provide the results with and without 8 times augmentation ( $\times 8$  AUG and non aug). With respect to LEHD, the inference results under both greedy mode and RRC (with 50 iterations) are documented.

## Main Results

The main experimental results for TSPLib and CVRPLib are detailed in Table 1 and Table 2. As shown in Table 1, we observed that compared to recent SOTA baselines, our method, namely POMO (ours) and LEHD (ours) both significantly reduce objective values and optimal gaps across all problem

sizes. Markedly, with 50 RRC iterations, LEHD (ours) dominates all other learning-based methods and ORTools across all sizes. Even compared to LEHD (RRC-50) itself, most gap reductions are more than a half, without notably increasing the inference time. POMO (ours) also outstrips solvers like SGBS, BQ and POMO, especially on large instances, both with and without augmentation. Compared to the pre-trained POMO, the gaps also dropped by more than a half. Given that the model architectures for POMO and LEHD remain intact, the inference time is only subject to trivial variance. POMO (non-augmentation) and LEHD (greedy) exhibit the lowest and second-lowest inference costs, respectively, while the costs associated with their augmented variants remain within a manageable range.

Regarding the generalization in CVRPLib, we report the same metrics as before in Table 2. Similarly to their TSP counterparts, POMO (ours) and LEHD (ours) under both non-augmented and augmented versions have witnessed consistent performance improvement, particularly in large-size instances. With 50 RRC iterations, LEHD under pro-



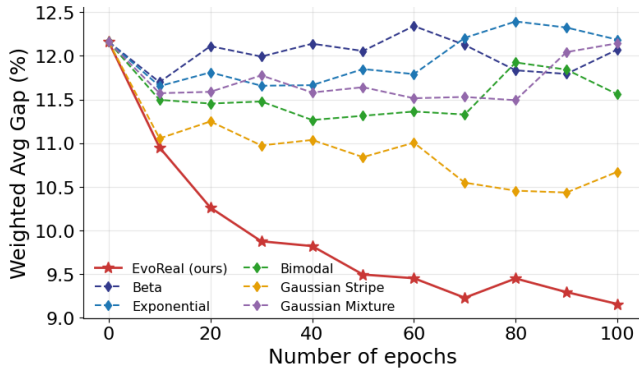


Figure 3: Comparison of the performance of the evolved generator with five the naive-distribution generators.

gressive fine-tuning has outperformed all neural solvers, and the results surpass ORTools in other problem intervals whereas it is still competitive against it with 0-200 nodes. The total average gaps of ours are comparable in all problem ranges (of the largest gap difference being 2.68%), indicating that progressive fine-tuning has significantly narrowed the performance difference between small-size instances and large-size ones. POMO under progressive fine-tuning also demonstrates improved performance over ELG and the original POMO across all problem sizes. It is worth noting that LEHD (ours) show even smaller average gap in large-size instances compared to small-size ones, possibly due to overfitting on small-size problems in the second phase of progressive fine-tuning. Additionally, all inference times vary from those of pre-trained models at a marginal level. Table 6 in the extended version also shows that with our framework, neural solvers can also generalize well on the problems unseen during training.

In summary, our EvoReal framework has substantially enhanced the models’ generalization capability on VRPLib, particularly on large problems. Ablation studies in Table 3 and Table 4 further corroborate the efficacy of our method.

## 5 Ablation Studies

**Phase Level Ablation.** We investigate the impact of each phase in progressive fine-tuning. Under two pre-trained backbone models for TSP, we ablate phase one and phase two separately, and report the range-wise average gaps and the total training time of the full progressive fine-tuning (Full), two ablation setups (w/o Phase 1 and w/o Phase 2) as well as not fine-tuning at all (w/o fine-tune). As shown in Table 3, the gaps in our full framework consistently outperform other setups excluding problem size [0,200], in which the gap marginally degrades 0.01% compared to without phase 2. Furthermore, compared to without phase 1, adding phase 1 markedly contributes to the performance improvement in problem size [200,1000] and [1000,5000], while the performance of without phase 1 is limited by the diversity of training instances. Similarly in Table 4, for LEHD-greedy, our full framework is superior on both small-size and large-size problems, with only minor degradation (at 0.04%) on the

Method	Avg. Aug Gap (%)
W/O Finetuning	22.21 $\pm$ 0.00
W/O LLM	19.79 $\pm$ 0.25 (best: 19.46)
W/O Rank-based Selection	16.88 $\pm$ 0.06 (best: 16.79)
W/O Design Guidance	16.62 $\pm$ 0.05 (best: 16.56)
W/O Late Selection	16.55 $\pm$ 0.07 (best: 16.50)
Full EvoReal	<b>16.52 <math>\pm</math> 0.05 (best: 16.41)</b>

Table 5: Ablation of key mechanisms in our evolution framework. Each entry shows the average fitness score of top-5 generators, with standard deviation and the score of best one.

medium-size one. This further confirms that our progressive design can synergistically narrow the gap while maintaining an acceptable training cost. Fig.6 in the extended version displays the training curves corresponding to Table 3 and 4.

**Comparison with Naive Generators.** We also compare our evolved three TSP generators, combined into a single generator, with five naive distribution generators, including beta, exponential, binomial, Gaussian mixture, and Gaussian stripe. All generators fine-tune the pre-trained POMO TSP model for 100 epochs. The trends of the total average gap in Fig.3 indicate that our evolved generator with mixed distribution exceeds all naive distribution generators from the outset and ultimately converges at a much lower average gap, substantiating that our evolved generator better captures the structural characteristics of TSPLib distributions.

**LLM Mechanism-level Ablation** To verify the usefulness of the modified components in our LLM-driven evolution framework, we perform the generator evolution on S1-type generator under different setups, including the removal of rank-based selection, the removal of the design-guidance prompt, and advancing the selection before evolution. We use the manually-crafted seed generator to fine-tune the pre-trained model for w/o LLM experiment. The performance of the original model is also added for comparison (w/o fine-tuning). The average fitness score of the top five elitist generators, along with the standard deviation and the fitness score of the best generator in each setup, are reported in Table 5.

## 6 Conclusion and Future Works

This work introduces a novel LLM-guided evolution framework for evolving VRP data distributions, combined with a progressive fine-tuning strategy to progressively adapt neural solvers from synthetic to diverse real-world instances. Our approach enables automated discovery and evolution of real-world-aligned VRP data generators, bridging the generalization gap for two neural solvers. Experimental results on two benchmarks demonstrate that the models trained with our evolved structural-aligned generators significantly outperform strong baselines, particularly on large-size instances, without requiring architectural changes or costly re-training. For future work, we will extend the LLM-driven evolution framework to other non-VRP tasks, such as MIS and bin packing, which involve richer structural and combinatorial constraints requiring expressive generator designs.

## 7 Acknowledgments

This research is supported by the National Research Foundation, Singapore, under its AI Singapore Programme (AISG Award No: AISG3-RP-2022-031, and AISG-NMLP-2024-003).

## References

- Applegate, D.; Bixby, R.; Chvatal, V.; and Cook, W. 2006. Concorde TSP solver.
- Bdeir, A.; Falkner, J. K.; and Schmidt-Thieme, L. 2022. Attention, filling in the gaps for generalization in routing problems. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, 505–520. Springer.
- Bello, I.; Pham, H.; Le, Q. V.; Norouzi, M.; and Bengio, S. 2017. Neural Combinatorial Optimization with Reinforcement Learning. arXiv:1611.09940.
- Bengio, Y.; Lodi, A.; and Prouvost, A. 2021. Machine learning for combinatorial optimization: A methodological tour d’horizon. *European Journal of Operational Research*, 290(2): 405–421.
- Bi, J.; Ma, Y.; Wang, J.; Cao, Z.; Chen, J.; Sun, Y.; and Chee, Y. M. 2022. Learning Generalizable Models for Vehicle Routing Problems via Knowledge Distillation. In Oh, A. H.; Agarwal, A.; Belgrave, D.; and Cho, K., eds., *Advances in Neural Information Processing Systems*.
- Briggs, W. L.; and Henson, V. E. 1995. *The DFT: an owner’s manual for the discrete Fourier transform*. SIAM.
- Cappart, Q.; Chételat, D.; Khalil, E. B.; Lodi, A.; Morris, C.; and Veličković, P. 2023. Combinatorial optimization and reasoning with graph neural networks. *Journal of Machine Learning Research*, 24(130): 1–61.
- Cattaruzza, D.; Absi, N.; Feillet, D.; and González-Feliu, J. 2017. Vehicle routing problems for city logistics. *EURO Journal on Transportation and Logistics*, 6(1): 51–79.
- Chalumeau, F.; Surana, S.; Bonnet, C.; Grinsztajn, N.; Pretorius, A.; Laterre, A.; and Barrett, T. 2023. Combinatorial optimization with policy adaptation using latent space search. *Advances in Neural Information Processing Systems*, 36: 7947–7959.
- Chen, J.; Zhang, Z.; Cao, Z.; Wu, Y.; Ma, Y.; Ye, T.; and Wang, J. 2023. Neural multi-objective combinatorial optimization with diversity enhancement. *Advances in Neural Information Processing Systems*, 36: 39176–39188.
- Choo, J.; Kwon, Y.-D.; Kim, J.; Jae, J.; Hottung, A.; Tierney, K.; and Gwon, Y. 2022. Simulation-guided beam search for neural combinatorial optimization. *Advances in Neural Information Processing Systems*, 35: 8760–8772.
- Clark, P. J.; and Evans, F. C. 1954. Distance to nearest neighbor as a measure of spatial relationships in populations. *Ecology*, 35(4): 445–453.
- Drake, J. H.; Kheiri, A.; Özcan, E.; and Burke, E. K. 2020. Recent advances in selection hyper-heuristics. *European Journal of Operational Research*, 285(2): 405–428.
- Drakulic, D.; Michel, S.; Mai, F.; Sors, A.; and Andreoli, J.-M. 2023. Bq-nco: Bisimulation quotienting for efficient neural combinatorial optimization. *Advances in Neural Information Processing Systems*, 36: 77416–77429.
- Duflo, G.; Kieffer, E.; Brust, M. R.; Danoy, G.; and Bouvry, P. 2019. A gp hyper-heuristic approach for generating tsp heuristics. In *2019 IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW)*, 521–529. IEEE.
- Elhenawy, M.; Abdelhay, A.; Alhadidi, T. I.; Ashqar, H. I.; Jaradat, S.; Jaber, A.; Glaser, S.; and Rakotonirainy, A. 2024a. Eyeballing combinatorial problems: A case study of using multimodal large language models to solve traveling salesman problems. In *International Conference on Intelligent Systems, Blockchain, and Communication Technologies*, 341–355. Springer.
- Elhenawy, M.; Abutahoun, A.; Alhadidi, T. I.; Jaber, A.; Ashqar, H. I.; Jaradat, S.; Abdelhay, A.; Glaser, S.; and Rakotonirainy, A. 2024b. Visual Reasoning and Multi-Agent Approach in Multimodal Large Language Models (MLLMs): Solving TSP and mTSP Combinatorial Challenges. *Machine Learning and Knowledge Extraction*, 6(3): 1894–1920.
- Gao, C.; Shang, H.; Xue, K.; Li, D.; and Qian, C. 2024. Towards Generalizable Neural Solvers for Vehicle Routing Problems via Ensemble with Transferrable Local Policy. In *IJCAI*, 6914–6922.
- Goh, Y. L.; Ma, Y.; Zhou, J.; Cao, Z.; Dupty, M. H.; and Lee, W. S. 2025. SHIELD: Multi-task Multi-distribution Vehicle Routing Solver with Sparsity & Hierarchy in Efficiently Layered Decoder.
- Gouveia, L. 1996. Multicommodity flow models for spanning trees with hop constraints. *European Journal of Operational Research*, 95(1): 178–190.
- Helsgaun, K. 2017. An extension of the Lin-Kernighan-Helsgaun TSP solver for constrained traveling salesman and vehicle routing problems. *Roskilde: Roskilde University*, 12: 966–980.
- Herraz, A.; Gutierrez, M.; and Ortega-Mier, M. 2022. Equivalent cyclic polygon of a euclidean travelling salesman problem tour and modified formulation. *Central European Journal of Operations Research*, 30(4): 1427–1450.
- Hottung, A.; Mahajan, M.; and Tierney, K. 2024. PolyNet: Learning Diverse Solution Strategies for Neural Combinatorial Optimization.
- Hou, Q.; Yang, J.; Su, Y.; Wang, X.; and Deng, Y. 2023. Generalize Learned Heuristics to Solve Large-scale Vehicle Routing Problems in Real-time. In *The Eleventh International Conference on Learning Representations*.
- Huang, Y.; Zhang, W.; Feng, L.; Wu, X.; and Tan, K. C. 2024. How Multimodal Integration Boost the Performance of LLM for Optimization: Case Study on Capacitated Vehicle Routing Problems. *CoRR*, abs/2403.01757.
- Iklassov, Z.; Du, Y.; Akimov, F.; and Takáč, M. 2024. Self-Guiding Exploration for Combinatorial Problems. In Globerson, A.; Mackey, L.; Belgrave, D.; Fan, A.; Paquet, U.; Tomczak, J.; and Zhang, C., eds., *Advances in Neural Information Processing Systems*, volume 37, 130569–130601. Curran Associates, Inc.



- Illian, J.; Penttinen, A.; Stoyan, H.; and Stoyan, D. 2008. *Statistical analysis and modelling of spatial point patterns*. John Wiley & Sons.
- Jiang, X.; Wu, Y.; Li, M.; Cao, Z.; and Zhang, Y. 2025a. Large Language Models as End-to-end Combinatorial Optimization Solvers. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems*.
- Jiang, X.; Wu, Y.; Zhang, C.; and Zhang, Y. 2025b. DRoC: Elevating large language models for complex vehicle routing via decomposed retrieval of constraints. In *13th international Conference on Learning Representations, ICLR 2025*.
- Joshi, C.; Cappart, Q.; Rousseau, L.; Laurent, T.; and Bresson, X. 2006. Learning TSP requires rethinking generalization. arXiv 2020. *arXiv preprint arXiv:2006.07054*.
- Kim, M.; Park, J.; and Park, J. 2022. Sym-NCO: Leveraging Symmetricity for Neural Combinatorial Optimization. In Koyejo, S.; Mohamed, S.; Agarwal, A.; Belgrave, D.; Cho, K.; and Oh, A., eds., *Advances in Neural Information Processing Systems*, volume 35, 1936–1949. Curran Associates, Inc.
- Kim, M.; Son, J.; Kim, H.; and Park, J. 2022. Scale-conditioned Adaptation for Large Scale Combinatorial Optimization. In *NeurIPS 2022 Workshop on Distribution Shifts: Connecting Methods and Applications*.
- Kingma, D. P.; and Ba, J. 2014. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.
- Konstantakopoulos, G. D.; Gayialis, S. P.; and Kechagias, E. P. 2022. Vehicle routing problem and related algorithms for logistics distribution: A literature review and classification. *Operational research*, 22(3): 2033–2062.
- Kool, W.; van Hoof, H.; and Welling, M. 2019. Attention, Learn to Solve Routing Problems! In *International Conference on Learning Representations*.
- Kou, S.; Golden, B.; and Poikonen, S. 2022. Optimal TSP tour length estimation using standard deviation as a predictor. *Computers & Operations Research*, 148: 105993.
- Kwon, Y.-D.; Choo, J.; Kim, B.; Yoon, I.; Gwon, Y.; and Min, S. 2020. POMO: Policy Optimization with Multiple Optima for Reinforcement Learning. In Larochelle, H.; Ranzato, M.; Hadsell, R.; Balcan, M.; and Lin, H., eds., *Advances in Neural Information Processing Systems*, volume 33, 21188–21198. Curran Associates, Inc.
- Kwon, Y.-D.; Choo, J.; Yoon, I.; Park, M.; Park, D.; and Gwon, Y. 2021. Matrix encoding networks for neural combinatorial optimization. In Ranzato, M.; Beygelzimer, A.; Dauphin, Y.; Liang, P.; and Vaughan, J. W., eds., *Advances in Neural Information Processing Systems*, volume 34, 5138–5149. Curran Associates, Inc.
- Li, S.; Yan, Z.; and Wu, C. 2021. Learning to delegate for large-scale vehicle routing. *Advances in Neural Information Processing Systems*, 34: 26198–26211.
- Lin, Z.; Wu, Y.; Zhou, B.; Cao, Z.; Song, W.; Zhang, Y.; and Jayavelu, S. 2024. Cross-Problem Learning for Solving Vehicle Routing Problems. In *Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence, IJCAI-24*, 6958–6966.
- Lisicki, M.; Afkanpour, A.; and Taylor, G. W. 2020. Evaluating Curriculum Learning Strategies in Neural Combinatorial Optimization. In *Learning Meets Combinatorial Algorithms at NeurIPS2020*.
- Liu, F.; Xialiang, T.; Yuan, M.; Lin, X.; Luo, F.; Wang, Z.; Lu, Z.; and Zhang, Q. 2024a. Evolution of Heuristics: Towards Efficient Automatic Algorithm Design Using Large Language Model. In *Forty-first International Conference on Machine Learning*.
- Liu, S.; Chen, C.; Qu, X.; Tang, K.; and Ong, Y.-S. 2024b. Large Language Models as Evolutionary Optimizers. In *2024 IEEE Congress on Evolutionary Computation (CEC)*, 1–8.
- Liu, S.; Zhang, Y.; Tang, K.; and Yao, X. 2023. How Good is Neural Combinatorial Optimization? A Systematic Evaluation on the Traveling Salesman Problem. *IEEE Computational Intelligence Magazine*, 18(3): 14–28.
- Luo, F.; Lin, X.; Liu, F.; Zhang, Q.; and Wang, Z. 2023. Neural Combinatorial Optimization with Heavy Decoder: Toward Large Scale Generalization. In *Thirty-seventh Conference on Neural Information Processing Systems*.
- Ma, Y.; Cao, Z.; and Chee, Y. M. 2023. Learning to search feasible and infeasible regions of routing problems with flexible neural k-opt. *Advances in Neural Information Processing Systems*, 36: 49555–49578.
- Nazari, M.; Oroojlooy, A.; Snyder, L.; and Takac, M. 2018. Reinforcement Learning for Solving the Vehicle Routing Problem. In Bengio, S.; Wallach, H.; Larochelle, H.; Grauman, K.; Cesa-Bianchi, N.; and Garnett, R., eds., *Advances in Neural Information Processing Systems*, volume 31. Curran Associates, Inc.
- Rajala, T. A.; Olhede, S. C.; Grainger, J. P.; and Murrell, D. J. 2023. What is the Fourier Transform of a Spatial Point Process? *IEEE Transactions on Information Theory*, 69(8): 5219–5252.
- Ramji, K.; Lee, Y.-S.; Astudillo, R. F.; Sultan, M. A.; Naseem, T.; Munawar, A.; Florian, R.; and Roukos, S. 2024. Self-Refinement of Language Models from External Proxy Metrics Feedback. *CoRR*, abs/2403.00827.
- Reinelt, G. 1991. TSPLIB—A traveling salesman problem library. *ORSA journal on computing*, 3(4): 376–384.
- Romera-Paredes, B.; Barekatin, M.; Novikov, A.; Balog, M.; Kumar, M. P.; Dupont, E.; Ruiz, F. J.; Ellenberg, J. S.; Wang, P.; Fawzi, O.; et al. 2024. Mathematical discoveries from program search with large language models. *Nature*, 625(7995): 468–475.
- Son, J.; Kim, M.; Kim, H.; and Park, J. 2023. Meta-SAGE: Scale Meta-Learning Scheduled Adaptation with Guided Exploration for Mitigating Scale Shift on Combinatorial Optimization. In Krause, A.; Brunskill, E.; Cho, K.; Engelhardt, B.; Sabato, S.; and Scarlett, J., eds., *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, 32194–32210. PMLR.
- Sun, Z.; and Yang, Y. 2023. Difusco: Graph-based diffusion solvers for combinatorial optimization. *Advances in neural information processing systems*, 36: 3706–3731.

Toth, P.; Vigo, D.; Toth, P.; and Vigo, D. 2014. *Vehicle Routing: Problems, Methods, and Applications, Second Edition*. USA: Society for Industrial and Applied Mathematics. ISBN 1611973589.

Tran, C. D.; Nguyen-Tri, Q.; Binh, H. T. T.; and Thanh-Tung, H. 2025. Large Language Models powered Neural Solvers for Generalized Vehicle Routing Problems. In *Towards Agentic AI for Science: Hypothesis Generation, Comprehension, Quantification, and Validation*.

Uchoa, E.; Pecin, D.; Pessoa, A.; Poggi, M.; Vidal, T.; and Subramanian, A. 2017. New benchmark instances for the capacitated vehicle routing problem. *European Journal of Operational Research*, 257(3): 845–858.

Vidal, T. 2022. Hybrid genetic search for the CVRP: Open-source implementation and SWAP\* neighborhood. *Computers & Operations Research*, 140: 105643.

Vinyals, O.; Fortunato, M.; and Jaitly, N. 2015. Pointer Networks. In Cortes, C.; Lawrence, N.; Lee, D.; Sugiyama, M.; and Garnett, R., eds., *Advances in Neural Information Processing Systems*, volume 28. Curran Associates, Inc.

Von Luxburg, U. 2007. A tutorial on spectral clustering. *Statistics and computing*, 17(4): 395–416.

Wang, H.; Feng, S.; He, T.; Tan, Z.; Han, X.; and Tsvetkov, Y. 2023. Can Language Models Solve Graph Problems in Natural Language? In Oh, A.; Naumann, T.; Globerson, A.; Saenko, K.; Hardt, M.; and Levine, S., eds., *Advances in Neural Information Processing Systems*, volume 36, 30840–30861. Curran Associates, Inc.

Xu, X.; Xiao, T.; Chao, Z.; Huang, Z.; Yang, C.; and Wang, Y. 2025. Can LLMs Solve Longer Math Word Problems Better? In *The Thirteenth International Conference on Learning Representations*.

Yang, C.; Wang, X.; Lu, Y.; Liu, H.; Le, Q. V.; Zhou, D.; and Chen, X. 2024. Large Language Models as Optimizers. In *The Twelfth International Conference on Learning Representations*.

Ye, H.; Wang, J.; Cao, Z.; Berto, F.; Hua, C.; Kim, H.; Park, J.; and Song, G. 2024. Reevo: Large language models as hyper-heuristics with reflective evolution. *Advances in neural information processing systems*, 37: 43571–43608.

Zhou, J.; Wu, Y.; Cao, Z.; Song, W.; Zhang, J.; and Shen, Z. 2024. Collaboration! Towards Robust Neural Methods for Routing Problems. In Globerson, A.; Mackey, L.; Belgrave, D.; Fan, A.; Paquet, U.; Tomczak, J.; and Zhang, C., eds., *Advances in Neural Information Processing Systems*, volume 37, 121731–121764. Curran Associates, Inc.

Zhou, J.; Wu, Y.; Song, W.; Cao, Z.; and Zhang, J. 2023. Towards Omni-generalizable Neural Methods for Vehicle Routing Problems. In Krause, A.; Brunskill, E.; Cho, K.; Engelhardt, B.; Sabato, S.; and Scarlett, J., eds., *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, 42769–42789. PMLR.

# Supplementary Materials

## A Pre-processing

### Instance Normalization

We follow pre-trained models such as POMO (Kwon et al. 2020) and LEHD (Luo et al. 2023), which randomly sampled uniformly distributed training instances within a unit square. To maintain consistency in data distribution and scale, we also perform the normalization procedure on the node coordinates of TSPLib (Reinelt 1991) and CVRPLib (Uchoa et al. 2017), rescaling them to the unit square before incorporating these datasets into our training pipeline. The Algorithm 1 is applied to the coordinates of all nodes. We also tried other common techniques like min-max normalization, standardization, and without-normalization for fine-tuning, but they either fail to converge or degrade the performance of the pre-trained model. We empirically find that rescaling TSPLib data to match the scale of synthetically generated uniform data, while preserving the original distributional structure, results in a more favorable convergence landscape during fine-tuning with TSPLib instances. For CVRPLib, we also employ the same approach to normalize customer nodes while re-scale the demands with capacities. For LEHD, we normalize the nodes and other constraints of the TSPLib and CVRPLib to the same scale as their training data.

### Modularization

Several prior studies have used statistics to analyze the structural diversity and heterogeneity of TSP instances (Herraiz, Gutierrez, and Ortega-Mier 2022) or predict optimal tours (Kou, Golden, and Poikonen 2022). Motivated by these findings, our approach adopts a structure-driven modularization strategy: we categorize TSPLib instances according to statistical features and design specialized generators tailored to each identified category. Specifically, we partition TSPLib using two complementary statistical measures: Fast Fourier Transform (FFT) energy and Nearest-Neighbor ratio (NN-ratio, also known as the Coefficient of Variation, CV). The formal definitions of these two statistics are as follows:

- **FFT Energy:** A quantitative statistic that measures the global periodicity or repetitiveness of a spatial point pattern by evaluating the strength of frequency components on its density map. A high FFT energy indicates pronounced global regularity, periodicity, or repeated structures, while a low FFT energy suggests more disordered, chaotic structures. Inspired by classical spectral analysis in signal and spatial statistics (Briggs and Henson 1995; Rajala et al. 2023), FFT energy is widely used to characterize spatial repetition and frequency content in spatial data.
- **NN-Ratio:** The NN-ratio is defined as the coefficient of variation of the nearest neighbor distances:

$$\text{NN-ratio} = \frac{\text{std}(d_{\text{NN}})}{\text{mean}(d_{\text{NN}})}$$

where  $d_{\text{NN}}$  denotes all the nearest neighbor distances in the spatial point set (Clark and Evans 1954; Illian et al.

2008). A low NN-ratio indicates that most points are regularly spaced, as in grid-like or weak-clustered layouts; and a high NN-ratio indicates the presence of clusters or highly varied point spacing. It is a standard, parameter-free measure in spatial statistics to distinguish between regular, random, and clustered point patterns (with self-exclusion, that is,  $k = 2$  for KNN).

Specifically, for FFT energy, we discretize the spatial coordinates into a 2-D histogram (density map), compute the two-dimensional Fast Fourier Transform, and summarize the mean energy of the main frequency region (excluding the DC component).

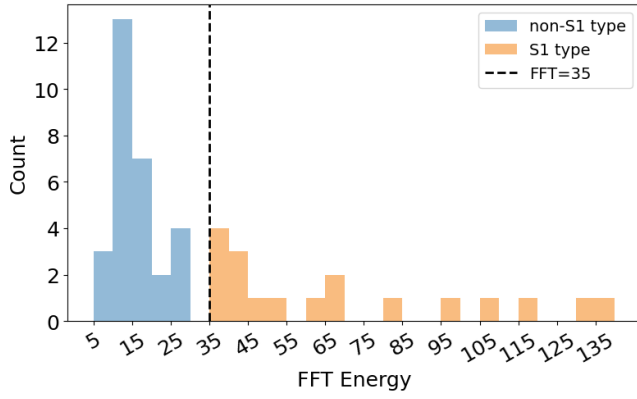
**Instance Segmentation.** We quantify the structural features of the TSPLib validation set using the two aforementioned statistics, and visualize their empirical distributions in Fig.4. We adopt a rule-based segmentation method to classify the TSPLib instances by setting separation thresholds with these statistics. For discriminating repetitive structures, a threshold of FFT=35 is selected based on the observation that the histogram of FFT energy exhibits a clear separation at this value.: almost all non-repetitive instances (non-S1 type) have FFT energy below this threshold, while highly repetitive (S1-type) instances are distinctly partitioned above it. Similarly, to distinguish S2 from S3, a threshold of NN-ratio=0.5 is chosen, as Fig.4b shows a marked jump in the frequency distribution: all S3-type instances cluster tightly above 0.5, whereas S2-type instances are concentrated below it. The box-plots in Fig.5 further suggest that these thresholds form natural boundaries that maximize separability while minimizing misclassifications. The structural descriptions of the three types of distributions are as follows.

- **S1:** Instances formed by the composition of multiple repeated geometric motifs, exhibiting structured spatial regularity with controlled variation.
- **S2:** Instances where points are globally arranged with low inter-nodes proximity variation, presenting weak-clustered or grid-like layouts.
- **S3:** Instances featured by local point aggregations, ranging from dense clusters to uniform distributions.

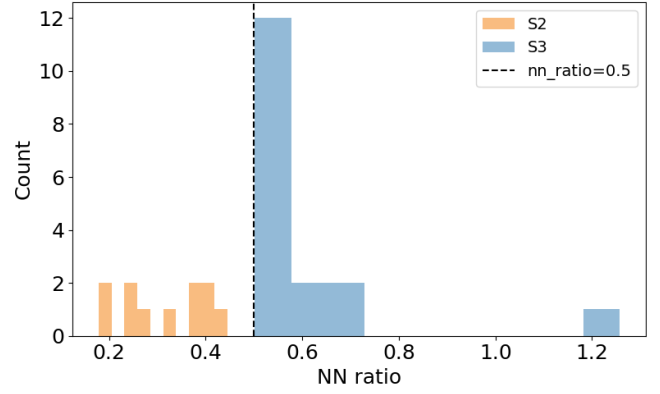
We have also explored other unsupervised machine-learning methods for instance segmentation, such as spectral clustering (Von Luxburg 2007), we found that the classes of most instances (85%) remain the same as those of rule-based segmentation while resulting in trivial performance variation ( $< 5\%$ ) of the best evolved generators found. This consistency further substantiates the interpretability and robustness of our rule-based classification scheme, indicating that the identified structural discrepancies among different TSPLib distributions are intrinsic rather than heuristic artifacts.

## B Evolution Mechanics

In this section, we illustrate more details regarding the LLM-guided generator evolution. Specifically, we firstly provide LLMs with detailed instructions regarding the desired distributional characteristics such as repeated or symmetric node patterns, and subsequently prompt them to design and implement corresponding generators based on these instruc-

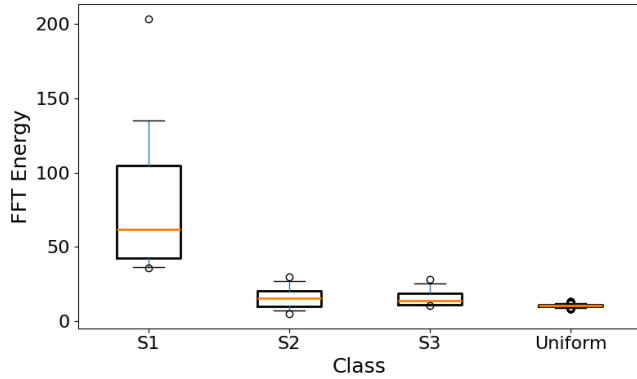


(a) Histogram of the FFT energy of S1 problems and non-S1 problems in the validation set. The black dotted line represents FFT=35.

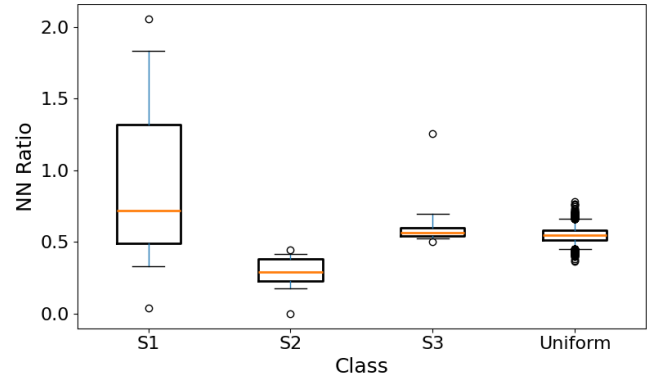


(b) Histogram of the NN-ratio of S2 type instances and S3 type instances. The black dotted line represents NN ratio=0.5.

Figure 4: Left: threshold for FFT energy of S1 type and non-S1 type. Right: threshold for NN-ratio of S2 type and S3 type. The thresholds of the divisions are both marked with black dotted line.



(a) Box-plot for FFT energy of S1, S2 and S3 validation sets.



(b) Box-plot for NN-ratio of S1, S2 and S3 validation problems .

Figure 5: Box-plots of FFT energy and NN-ratio of all validation instances after instance segmentation. The corresponding statistics of 5000 TSP 100 instances sampled from uniform distribution is supplemented for comparison.

---

#### Algorithm 1: Normalize Coordinates

---

**Require:**  $\mathbf{X} \in \mathbb{R}^{n \times 2}$  // Input tensor with  $n$  points in 2-dim  
**Ensure:**  $\mathbf{X}_{\text{norm}}$  // Normalized coordinates

- 1:  $\text{min\_val} \leftarrow \min(\mathbf{X}, \text{axis} = 0)$
- 2:  $\text{max\_val} \leftarrow \max(\mathbf{X}, \text{axis} = 0)$
- 3:  $\text{max\_diff} \leftarrow \max(\text{max\_val} - \text{min\_val})$
- 4:  $\mathbf{X}_{\text{norm}} \leftarrow (\mathbf{X} - \text{min\_val}) / \text{max\_diff}$
- 5: **return**  $\mathbf{X}_{\text{norm}}$

---

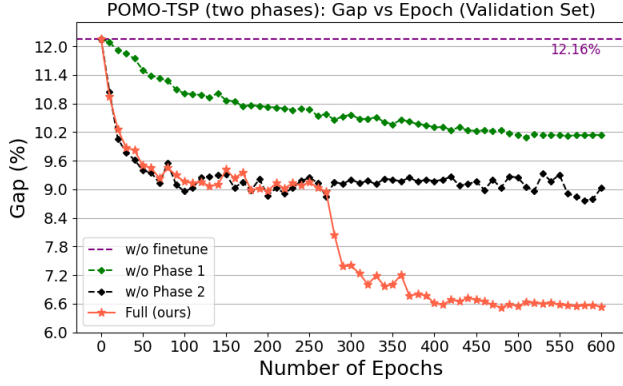
tions. Specifically, we inform LLMs of the encouraged behaviors as well as the discouraged behaviors when designing a PyTorch-based data generator for seamless model fine-tuning. When initializing population, we formulate five categories of prompts to ensure task-aware generation:

- **Problem description:** This prompt informs LLMs about the purpose of devising a data generator.
- **Function format:** For formatting a valid generator, function-formatting prompts consist of function descrip-

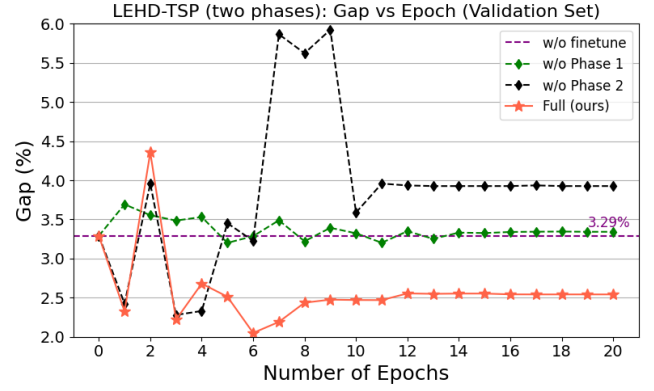
tion and function signature are fabricated to standardize the names, inputs and outputs of the code.

- **System generator:** Further stresses on the function specifications while enumerating all safety instructions, including robust constraints and forbidden operations.
- **Seed generator:** Provides LLMs with a valid exemplification of a manually designed generator for in-context learning with the structural hints and robust designs.
- **Design guidance:** Serves as a external knowledge constraint to encourage LLMs to explore the structures specified in the prompt.

**Evolution Mechanisms for Creating New Generators.** We build our evolution framework upon the structure of ReEvo (Ye et al. 2024), with modifications on several key components for better evolution-search performance. Evolution mechanisms can be classified into two groups: Searching (crossover, mutation) and Reflection (short-term reflection, long-term reflection), and they are arranged into the following sequential order in one evolution iteration:



(a) Training curve for POMO-TSP solver under three different ablation setups: w/o phase 1, w/o phase 2, and w/o finetune.



(b) Training curve for LEHD-TSP solver under three different ablation setups: w/o phase 1, w/o phase 2, and w/o finetune.

Figure 6: Comparison of full progressive fine-tuning of POMO (left) and LEHD (right) TSP solver with different phase-level ablations. The ablation setups are the same as those in Table 3 and Table 4.

**1. Short-term reflection:** Pairs every two generators in the generator population as a parent pair, and requests LLMs to summarize the successes and failures in their design by comparing the parent generators.

**2. Crossover:** Generates a new version of the offspring generator based on the previous parent pair along with short-term reflections.

**3. Long-term reflection:** Synthesizes the newly gained short-term reflections and the previous long-term one.

**4. Mutation:** Modifies the current best generator in the population under the guidance of the previous long-term reflection.

After mutation, the generator population is subject to probability-based selection in order to maintain a fixed population size. Unlike ReEvo, we introduced a new selection mechanism to expedite the convergence of evolution-search given the huge search space of generator structures and limited evaluation costs:

- **Rank-based selection:** the fitness score of each generator is used to calculate the probability of survival. For example, the generators with lower best-average-augmentation-gap have a higher chance to enter the next iteration.

**Postponing early-selection.** To enhance both exploration and robustness of LLM-guided evolutionary search, we delay selection until after crossover and mutation, rather than filtering the initial population early, as in ReEvo. This prevents the premature elimination of structurally diverse or promising candidates, especially those involving complex structures, and results in a richer, more resilient generator pool for TSP data evolution. The previous ablation study has validated the superior performance of our modifications.

## C Routing Problem Formulation

**TSP.** In TSP, each node  $v_i \in V$  is characterized by a 2-D coordinate  $v_i = (x_i, y_i)$ . The weight of each edge between

any two nodes corresponds to their Euclidean distance, computed by  $w_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$ . A solution is considered feasible if each node  $v_i$  is visited exactly once in a tour and the tour is closed. The objective is to find the tour with a minimum total length.

**CVRP.** (Toth et al. 2014) Every node in CVRP has two elements: node coordinate  $v_i$  and demand  $\delta_i$  for  $n \in \{1, 2, 3, \dots, n\}$ . Specifically, the demand  $\delta_0$  for the depot node  $v_0$  is 0. The solution to the problem consists of multiple subtours, each starting and terminating at the depot. A feasible tour should satisfy: 1) every customer node in the problem is visited only once. 2) The total demand in each subtour does not exceed the pre-defined vehicle capacity.

In this paper, we mainly consider the evolution of the distributions of the customer nodes, which is consistent with the design of our framework. To craft the seed generator for CVRP data evolution, we also categorize CVRP instances into three sets according to different depot positioning schemes (Gouveia 1996): 1) S1: depot is located in the center of the unit square. 2) S2: depot is randomly sampled from the unit square  $[0, 1]^2$ . 3) S3: depot is at (0,0). Nonetheless, instead of evolving the generator of each category one at a time, we incorporate these three types of depots into one generator to save evolution cost. Following existing works (Uchoa et al. 2017; Nazari et al. 2018), the coordinates of the customer nodes are sampled uniformly from the unit square while the demand  $d_i$  of each customer node is initialized with  $U[1, 10]$ . The capacity  $Q$  of each vehicle is set according to the following formula:  $Q = \max(\lceil r \cdot \frac{1}{n} \sum_{i=1}^n d_i \rceil, \lceil k \cdot \max_{1 \leq i \leq n} d_i \rceil)$  where  $r \sim \text{Triangular}(3, 6, 25)$  and  $k$  is a constant (e.g.  $k = 2$ ). The demands and capacity are further normalized to  $\delta'_i = \delta_i / Q$  and 1, respectively.

## D Experimental Setup

### Model Setup and Training

Our framework utilizes the identical network topologies of POMO (Kwon et al. 2020) and LEHD (Luo et al. 2023). For

Problem Name	Opt	POMO (×8 aug)		POMO (ours) (×8 aug)		LEHD (RRC50)		LEHD (ours) (RRC50)	
		Obj	Gap	Obj	Gap	Obj	Gap	Obj	Gap
berlin52	7542	9445.60	25.24%	7606.86	0.86%	<b>7544.36</b>	<b>0.03%</b>	7544.66	0.04%
st70	675	699.98	3.70%	682.02	1.04%	<b>677.11</b>	<b>0.31%</b>	<b>677.11</b>	<b>0.31%</b>
pr76	108159	134906.72	24.73%	111295.61	2.90%	<b>108159.44</b>	<b>0.00%</b>	<b>108159.44</b>	<b>0.00%</b>
eil76	538	577.11	7.27%	<b>542.41</b>	<b>0.82%</b>	544.37	1.18%	544.37	1.18%
pr124	59030	60192.89	1.97%	59649.82	1.05%	<b>59030.73</b>	<b>0.00%</b>	59030.74	0.00%
rl1304	252948	464108.99	83.48%	333638.41	31.90%	262932.00	3.95%	<b>256881.20</b>	<b>1.55%</b>
rl1323	270199	493977.81	82.82%	357013.94	32.13%	281240.97	4.09%	<b>275050.94</b>	<b>1.80%</b>
nrv1379	56638	76404.66	34.90%	71375.21	26.02%	63142.54	11.48%	<b>58333.34</b>	<b>2.99%</b>
fl1400	20127	28511.91	41.66%	23258.76	15.56%	21139.27	5.03%	<b>20770.40</b>	<b>3.20%</b>
u1432	152970	205698.76	34.47%	183548.70	19.99%	157660.63	3.07%	<b>154439.91</b>	<b>0.96%</b>
fl1577	22249	35665.15	60.30%	29947.15	34.60%	24010.49	7.92%	<b>22680.66</b>	<b>1.94%</b>
d1655	62128	106487.39	71.40%	83089.99	33.74%	66893.05	7.67%	<b>64340.05</b>	<b>3.56%</b>
vm1748	336556	643293.14	91.14%	439003.65	30.44%	350746.59	4.22%	<b>341082.19</b>	<b>1.34%</b>
u1817	57201	113349.50	98.16%	79366.39	38.75%	60742.67	6.19%	<b>59037.48</b>	<b>3.21%</b>
rl1889	316536	647569.35	104.58%	446410.72	41.03%	331794.13	4.82%	<b>323538.38</b>	<b>2.21%</b>
d2103	80450	139709.47	73.66%	114769.97	42.66%	<b>84679.73</b>	<b>5.26%</b>	85015.16	5.67%
u2152	64253	127041.03	97.72%	90975.82	41.59%	69816.06	8.66%	<b>66345.79</b>	<b>3.26%</b>
u2319	234256	293991.28	25.50%	272791.11	16.45%	240811.14	2.80%	<b>235586.16</b>	<b>0.57%</b>
pr2392	378032	674068.86	78.31%	536087.18	41.81%	415166.63	9.82%	<b>389460.37</b>	<b>3.02%</b>
pcb3038	137694	234782.04	70.51%	193019.45	40.18%	151426.67	9.97%	<b>142091.75</b>	<b>3.19%</b>
fl3795	28772	54635.15	89.89%	43839.90	52.37%	31647.42	9.99%	<b>29807.47</b>	<b>3.60%</b>
fnl4461	182566	330827.85	81.21%	263059.35	44.09%	209771.78	14.90%	<b>191593.55</b>	<b>4.94%</b>
Average	128614.50	203597.92	58.30%	163105.40	26.82%	136344.44	5.52%	<b>131455.05</b>	<b>2.21%</b>

Table 6: Performance on 22 Unseen TSPLib Problems with original POMO and LEHD, and the ones trained with our EvoReal framework. The best gap and the best objective value of each problem are marked in **bold**.

proxy evaluation in LLM-guided evolution and phase one of progressive fine-tuning, we employed the pre-trained TSP and CVRP models trained with problems of size 100, using their released checkpoints. For proxy evaluation and phase one fine-tuning, we also generate TSP100 and CVRP100 for training. Additionally, we employ the same data loader as POMO’s and LEHD’s. Specifically, for fine-tuning LEHD in phase one, training instances are generated only once and shuffled for reuse when training every epoch. For the results in Table 1 and Table 2, we set the hyperparameters in Table 7 for LLM-guided evolution and proxy evaluation.

## Hyperparameter Selection

**Proxy-evaluation & Phase One Fine-tuning Setup.** For proxy-evaluation in LLM-based evolution and phase one fine-tuning, we maintain most of the hyperparameters the same as the original experimental settings in the respective model papers: Adam (Kingma and Ba 2014) optimizer is used with an initial learning rate of  $10^{-4}$  and a decay of 0.9 for both TSP and CVRP, except for the LEHD TSP model, where a lower learning rate of  $10^{-5}$  and a decay of 0.6 is used. We use the same batch sizes as the original papers. However, for phase one fine-tuning, we fine-tuned the pre-trained models with more epochs and instances to see the convergence. For POMO in phase one, the model for TSP

is fine-tuned for 300 epochs with 5000 problems in each episode, while the model for CVRP is fine-tuned for 200 epochs with 500 problems in each episode. For LEHD in phase one, we fine-tuned their pre-trained TSP version and their CVRP version both for 20 epochs. And in each episode for LEHD, we generate 50000 synthetic data once for the repetitive usage in each epoch. For finetuning both models in phase one, we generate mixed distributions of S1, S2, S3 according to the ratio of S1, S2, S3 problems in the validation set in one epoch (e.g. if S1 problems: S2 problems: S3 problems= 17:19:12, then the ratio of the synthetic instances of type S1, S2, S3 is also 17:19:12).

**Phase Two Fine-tuning Setup.** In the case of phase two fine-tuning, the hyperparameter setups are more empirical and pragmatic. The training epochs for both problems of both models are identical to those of phase one. For POMO, we used the same learning rate and decay as those of phase one. Nevertheless, the learning rate and the decay for LEHD-TSP are  $10^{-6}$  and 0.9 while the learning rate and the decay for LEHD-CVRP are  $10^{-5}$  and 0.8. As for batch sizes, since we use the real instances from TSPLib and CVRPLib for fine-tuning, we expand each problem once by batch size as one batch in an epoch. Limited by computational resources, we used an adaptive batch size (*bs*) setting for both VRP problems for POMO (for



Parameter	Value
LLM model (generator and reflector)	OpenAI: o3
Temperature	1
Initial population size	30
Offspring population size	10
Crossover rate	1
Mutation rate	0.5
Maximum number of iterations	10
Maximum number of evaluations	125
POMO (Kwon et al., 2020)	
Epochs	40
Learning rate decay	No decay
Train episodes per epoch	5000
LEHD (Luo et al., 2023)	
Epochs	10
Learning rate decay	0.97
Train episodes per epoch	10000

Table 7: Experiment settings of generator evolution.

POMO-TSP, we set  $bs = 4$  for  $n \in [0, 500)$ ,  $bs = 2$  for  $n \in [500, 750)$ ,  $bs = 1$  for  $n \in [750, 1002)$  where  $n$  is the problem size; for POMO-CVRP,  $bs = 4$  for  $n \in [0, 500)$ ,  $bs = 2$  for  $n \in [500, 650)$ ,  $bs = 1$  for  $n \in [650, 800)$ , which still achieve acceptable generalization performance on both the small- and large-size instances. As for LEHD, we used a fixed batch size and decay of 4 and 0.6 for TSP respectively, and fix a batch size and a decay of 16 and 0.8 for CVRP.

**Training Curves for Phase-level Ablation.** As shown in Fig.6, we compared our full progressive fine-tuning method with ablation setups on different phases with our backbone models, which corresponds to the results in Table 3 and Table 4. These figures further corroborate that our full framework enables the optimization of model parameters toward better generalization optima, guiding the models to achieve dominant performance.

## E Additional Results

### Generalization on Unseen Data

To further evaluate generalization ability of the models trained with our framework, particularly on the data unseen during the generator evolution and progressive fine-tuning, we tested them directly on 22 held-out problems. The 22 unseen problems include 5 small-size instances and 17 instances of sizes  $n \in [1300, 5000]$  proposed in TSPLib95 (Reinelt 1991). Original POMO (Kwon et al. 2020) and LEHD (Luo et al. 2023), along with those trained from progressive fine-tuning, are employed to solve the instances, with both models tested under their augmentation version ( $\times 8$  augmentation or RRC-50). We report the optimum, the objective values, and the gaps for each problem in Table 6. As shown, LEHD trained from progressive finetuning (i.e., LEHD (ours) in the table) outperforms other models in most problems, especially on large-size ones like pcb3038 and

**Function description:** This function returns a batch of synthetic TSP problems as a torch.Tensor of shape (batch\_size, problem\_size, 2). Each individual problem must be well-structured and exactly of size (problem\_size, 2) to ensure safe stacking. The parameter problem\_size defines the number of nodes in each generated TSP instance.

**Code:**

```
def generate_s3_type(batch_size=10, problem_size: int = 100):
    import torch

    def generate_problem():
        # Decide the number of clusters and assign points to clusters
        max_clusters = 6
        num_clusters = torch.randint(2, max_clusters + 1, (1,)).item()
        cluster_sizes = [problem_size // num_clusters] * num_clusters
        for i in range(problem_size % num_clusters):
            cluster_sizes[i] += 1

        points = []
        for size in cluster_sizes:
            # Sample a cluster center and radius
            center = torch.rand(2) * 0.7 + 0.15
            radius = torch.rand(1).item() * 0.15 + 0.05
            # Generate points for this cluster
            pts = center + torch.randn(size, 2) * radius
            points.append(pts)

        # Combine all clusters into a single set of points
        all_points = torch.cat(points, dim=0)
        # Adjust the number of points to match the required problem size
        if all_points.size(0) > problem_size:
            perm = torch.randperm(all_points.size(0))
            all_points = all_points[perm[:problem_size]]
        elif all_points.size(0) < problem_size:
            pad = torch.rand(problem_size - all_points.size(0), 2)
            all_points = torch.cat([all_points, pad], dim=0)
        return all_points

    # Generate a batch of clustered TSP instances
    batch = torch.stack([generate_problem() for _ in range(batch_size)])
    return batch
```

**Fitness value:** 6.17%

Figure 7: An example of an evolved S3-type generator that outputs TSP problems of 100 node, along with function description and its fitness score.

fnl4461, while it is still comparable to the original model on the first five small instances. As for POMO (ours), all problems have seen reductions in gaps by 0.92%  $\sim$  60%+. These results demonstrate that the models with our methods, despite not being exposed to these unseen TSPLib instances during fine-tuning, consistently achieve strong generalization performance, particularly on large-scale and challenging cases. This improvement can be largely attributed to the structural prior alignment introduced in the first phase of progressive fine-tuning, which enables the model to internalize diverse distributional patterns and transfer this knowledge to previously unseen problem instances.

### Effect of different LLMs (for Generator Design)

We evaluate the effectiveness of the generator design using four up-to-date LLMs developed by OpenAI: GPT-4o, GPT-4.1, GPT-4.1 mini, and OpenAI o3 to evolve our S1-type generators for TSP. All experiments were conducted in identical settings to ensure a fair comparison. Our experimental results indicate that integrating the LLM-guided generator for phase one fine-tuning consistently improves the performance on the S1 validation set. Among the language models evaluated, OpenAI o3 achieves the best performance, suggesting that our framework is able to achieve better results with strong LLMs.

## F Prompts

Prompts adopted for EvoReal are gathered in this section. For evolution operators (crossover and mutation) and reflection operators (long-term reflection and short-term reflection), since we only made slight modifications to those of ReEvo (Ye et al. 2024), these prompts are not listed here. Prompts 1–3 are shared by both the TSP and CVRP versions of EvoReal. Problem-specific prompt components are prompt 4–10 as well as Table 8 and Table 9. Table 8 lists all generator descriptions. Table 9 presents the problem descriptions of all COP settings. For demonstration purposes, we only present seed generators and design prompts for CVRP generator (prompt 7 and 9) and TSP S3 generator (prompt 8 and 10). The full ensemble of prompts are in our supplementary materials.

System prompts (prompt 1 and 2) are used to formalize and constrain the task and behavior-including generation procedures, standard formats, encouraged behaviors, and forbidden operations in performing the required task for both the LLM generator and the LLM reflector. The user prompt (prompt 3) is used to construct the task-specific instructions, covering problem description, function name and signature, function description and seed prompt for EvoReal. The function signatures (prompt 4) here specify function names with their types. Seed prompts (prompt 5 and prompt 6) are designed to include the hand-crafted design prompt as well as the seed generator, which together specify the generator styles while narrowing the search space. The design prompts establish verbal boundaries that constrain the generator’s design style.

You are a Python expert in combinatoric optimization. Generate a single Python function that outputs valid 2D coordinates.

### [FUNCTION INTERFACE]

- Define exactly one function.
- All necessary imports (e.g., torch) must appear inside the function.

### [ROBUSTNESS CONSTRAINTS]

- No NaN, Inf, or negative tensor dimensions.
- No use of undefined or invalid Python variables or functions(e.g. `tensor.random` is invalid).
- Use `math.sin/cos/sqrt` for scalars. And `torch.sin/cos/sqrt` only for tensor inputs.
- `torch.stack` must receive a non-empty list.
- Avoid implicit broadcasting (e.g., `(N,) + (N,2)`); always align dimensions explicitly.
- Do not divide by zero or use undefined variables.
- Always validate and correct the number of generated coordinates per problem before stacking.
- Do not assume intermediate steps generate the exact target number of coordinates, explicitly pad or truncate as needed.

### [OUTPUT FORMAT]

- Output only valid Python code (no comments or markdown).
- Wrap code in a single Python code block using triple backticks: ````python ... ````
- Only one top-level function is allowed.

### [FORBIDDEN OPERATIONS]

- Never use `torch.linspace(..., endpoint=...)`. PyTorch does not support the `endpoint` argument and this will raise a `TypeError`.
- Never slice or skip tensor dimensions to fix shape mismatches. You must always fix mismatches explicitly using `reshape`, `pad`, or `truncate`.
- Never use any keyword arguments (e.g., `indexing`, `sparse`) in `torch.meshgrid(...)`. Use `torch.meshgrid(x, y)` without keywords to ensure compatibility.
- Never use negative or zero values as shape arguments for tensor creation functions like `torch.rand`, `torch.randn`, or `torch.empty`. Always clamp or check values (e.g., `max(1, x)`) to ensure strictly positive integer dimensions.

Prompt 1: System prompt for generator LLM.

You are an expert in the domain of optimization. Your task is to give hints to design better generators for combinatoric optimization problems.

Prompt 2: System prompt for reflector LLM.

```

### Task Objective:
{problem_desc}

### Function Name:
{func_name}

### Function Description:
All generated new python code must follow this function signature as the first line of python function:
{func_desc}

### Function Signature:
{func_signature}

### Design Guidance:
- Design generation strategies that reflect spatial complexity, distributional realism, and diverse patterns across samples.
- Encourage variability across samples: avoid excessive reuse of the same generation logic or structure.

### Reference Module:
{seed}

```

Prompt 3: User prompt for generator population initialization.

```

# TSP generator
def generate_{TSP TYPE}_type(batch_size=10, problem_size=100): #. {TSP TYPE} = s1 or s2 or s3

# CVRP generator
def generate_customer_positions(n, cust_type=None):

```

Prompt 4: Function signatures used in EvoReal.

```

Your design for the code should align with the design guidance below:
{external_guide}

seed generator:
{seed_func}

This module (python function) illustrates how to call PyTorch APIs (e.g., `torch.rand`, `torch.cat`,
`torch.randperm`, `torch.linspace`) in a safe and modular way. You may refer to the trivial module above
for inspiration but you should explore more logic, generation strategies, or substructures compositions to
create new `{func_name}`. Output code only and enclose your code with Python code block: ```python ...
...

```

Prompt 5: Seed prompt for TSP generator population initialization.

Your design for the customer position generator should align with the design guidance below:  
{external\_guide}

seed generator:  
{seed\_func}

The generator (python function) above illustrates how to call PyTorch APIs (e.g., `torch.rand`, `torch.cat`, `torch.randperm`, `torch.linspace`) in a safe and modular way. Note that your generator should also have include inner type-selection parameters (e.g. CUST\_TYPES = ['P1', 'P2', 'P3'] for customer position generator). The number of customer position types can be any number. However, be sure to denote each customer position by the capital letter 'P' with index number (i.e., P1,P2,P3...).

You may refer to the trivial generator above but you should explore more logic, generation strategies, or substructures compositions to create a new one in accordance with the design guidance. Output exactly **one** customer position generator only and enclose the generator with a Python code block: ```python ...

Prompt 6: Seed prompt for CVRP generator population initialization.

```
def generate_customer_positions(n, cust_type=None):
    """
    Generate customer positions in [0,1] x [0,1] randomly.
    """
    import random
    import numpy as np

    def ensure_n_unique_points(points, n):
        """
        Ensure the output points are unique and exactly n.
        """
        points = np.unique(points, axis=0)
        while len(points) < n:
            new_pt = np.round(np.random.rand(1, 2), 3)
            while any((new_pt == points).all(axis=1)):
                new_pt = np.round(np.random.rand(1, 2), 3)
            points = np.vstack([points, new_pt])
        if len(points) > n:
            points = points[:n]
        return points

    CUST_TYPES = ['P1']
    if cust_type is None:
        cust_type = random.choice(CUST_TYPES)
    if cust_type == 'P1':
        # Uniformly sample n points, round to 3 decimals
        points = np.round(np.random.rand(n, 2), 3)
        return ensure_n_unique_points(points, n)
```

Prompt 7: Seed CVRP coordinate generator used in EvoReal.

```

def generate_s3_type(batch_size: int = 10, problem_size: int = 100):
    import torch

    def generate_problem():
        max_clusters = 6
        num_clusters = torch.randint(2, max_clusters + 1, (1,)).item()
        cluster_sizes = [problem_size // num_clusters] * num_clusters
        remainder = problem_size % num_clusters

        for i in range(remainder):
            cluster_sizes[i] += 1
            # generate multiple random clusters and combine them
        points = []
        for size in cluster_sizes:
            cluster_center = torch.rand(2) * 0.7 + 0.15
            cluster_radius = torch.rand(1).item() * 0.15 + 0.05
            cluster_points = cluster_center + (torch.randn(size, 2) * cluster_radius)
            points.append(cluster_points)

        all_points = torch.cat(points, dim=0)

        if all_points.size(0) > problem_size:
            perm = torch.randperm(all_points.size(0))
            all_points = all_points[perm[:problem_size]]
        elif all_points.size(0) < problem_size:
            padding_size = problem_size - all_points.size(0)
            padding_points = torch.rand(padding_size, 2)
            all_points = torch.cat([all_points, padding_points], dim=0)

        return all_points
    # create a batch of TSP instances
    batch = torch.stack([generate_problem() for _ in range(batch_size)])
    return batch

```

Prompt 8: Seed TSP coordinate generator used in EvoReal.



You should design a customer position generator that have more diverse generation methods and distributions than seed generators . For designing new `generate\_customer\_positions` function, you are encouraged to more customer position distributions other than pure uniform distribution. For example, including exponential distribution, gaussian distributions, rotation distributions, explosion distributions, a mixtures of those distributions and etc. Generate at least two distributions in the customer position generator. A example design is braketed in ``python code`` below:

```
def generate_customer_positions(n, cust_type=None):
    CUST_TYPES = ['P1', 'P2', 'P3']
    if cust_type is None:
        cust_type = random.choice(CUST_TYPES)
    if cust_type == 'P1':
        points = np.round(np.random.rand(n, 2), 3)
        return ensure_n_unique_points(points, n)
    elif cust_type == 'P2':
        S = np.random.randint(3, 9)
        centers = np.round(np.random.rand(S, 2), 3)
        positions = [tuple(center) for center in centers]
        while len(positions) < n:
            pt = np.round(np.random.rand(2), 3)
            dists = np.linalg.norm(centers - pt, axis=1)
            prob = np.sum(np.exp(-dists / 0.04))
            if np.random.rand() < prob:
                positions.append(tuple(pt))
        positions = np.unique(np.array(positions), axis=0)[:n]
        return ensure_n_unique_points(positions, n)
    elif cust_type == 'P3':
        n_c = n // 2
        n_r = n - n_c
        clustered, _ = generate_customer_positions(n_c, 'P2')
        randoms, _ = generate_customer_positions(n_r, 'P1')
        positions = np.vstack([clustered, randoms])
        if len(positions) > n:
            positions = positions[:n]
        return ensure_n_unique_points(positions, n)
    else:
        raise ValueError(f'Unknown customer type: {cust_type}')
```

Prompt 9: Design prompt (external knowledge) for CVRP coordinate generator.

### Design Guidance for S3-Type Instance Generators:

When generating S3-type (cluster-type) instances, it is essential to capture the full spectrum of structural diversity observed across different cluster-type distributions. Avoid overfitting to prototypes with clear, compact cluster centers. Instead, incorporate instances that range from well-defined dense groupings to those with subtle or diffuse clustering, and even those approaching uniformity. Strive to explore the continuum between localized aggregation and global uniformity. Embrace noise, partial clusters, multiple clusters and varied inter-cluster distances to improve representational breadth. Generators should be flexible enough to mimic both visually distinct clusters and ambiguous arrangements, ensuring broader coverage of naturally occurring patterns. To enhance generalization, prioritize heterogeneity in spatial coherence and avoid repetitive generation of highly clustered-only configurations.

Prompt 10: Design prompt (external knowledge) for TSP coordinate generator.



Problem	Function description
TSP_generator	The function must take two arguments: (batch_size, problem_size). It must return a batch of synthetic TSP problems as a torch.Tensor of shape (batch_size, problem_size, 2), with dtype=torch.float32. Each individual problem must be well-structured and exactly of size (problem_size, 2) to ensure safe stacking. The parameter problem_size defines the number of nodes in each generated TSP instance. All generated problems for stacking in the final return must have exactly problem_size 2D coordinates. Use torch.randperm to truncate or torch.cat + torch.rand() to pad if needed. The output must be stackable via torch.stack without shape errors.
CVRP_generator	<p>The function must take two arguments: (n, cust_type=None). It generates a 2D coordinate array of shape (n, 2) representing customer positions for a CVRP instance. The parameter n is the number of customers in each generated instance, while cust_type=None is a fixed trivial parameter (do not change it). The function must always return a numpy.ndarray of shape (n, 2), where all coordinate values are within the range [0, 1], and all customer positions are unique. To ensure uniqueness, the function must include and call the helper function below (do not modify or remove it):</p> <pre>def ensure_n_unique_points(points, n):     """Ensure the output points are unique and exactly n."""     points = np.unique(points, axis=0)     while len(points) &lt; n:         new_pt = np.round(np.random.rand(1, 2), 3)         while any((new_pt == points).all(axis=1)):             new_pt = np.round(np.random.rand(1, 2), 3)         points = np.vstack([points, new_pt])     if len(points) &gt; n:         points = points[:n]     return points</pre> <p>Both the input and the return of ensure_n_unique_points are (n, 2) numpy.ndarray.</p>

Table 8: Generator descriptions used in prompts.

Problem	Problem description
TSP_generator	The goal is to design a TSP problem data generator — a Python function that outputs synthetic TSP instances for fine-tuning a neural solver. The generated problems should not merely be valid or diverse, but should exhibit structural patterns that are highly representative of a specific class of TSPLib distributions. This includes capturing repeated motifs, geometric arrangements, directional or symmetric alignments, and other spatial regularities. The generator should produce well-formed instances of fixed problem size and exhibit strong generalization on real TSPLib problems of the same type, serving as an abstract yet effective summary of the target distribution’s geometric and distributional characteristics.
CVRP_generator	The goal is to design a CVRP customer position generator — a Python function that outputs synthetic CVRP customer coordinate positions for fine-tuning a neural solver. The generated problems should not merely be valid or diverse, but should exhibit structural patterns that are highly representative of a specific class of CVRPLib distributions. The generator should produce a well-formed set of customer coordinates of fixed problem size n and exhibit strong generalization on real CVRPLib problems of the same type, serving as an abstract yet effective summary of the target distribution’s geometric and distributional characteristics.

Table 9: Problem descriptions used in prompts.