



**HAL**  
open science

# Machine Learning for Developing Guidance to Improve Metaheuristic Algorithm

Bachtiar Herdianto, Romain Billot, Flavien Lucas, Marc Sevaux

► **To cite this version:**

Bachtiar Herdianto, Romain Billot, Flavien Lucas, Marc Sevaux. Machine Learning for Developing Guidance to Improve Metaheuristic Algorithm. The 12th International Conference on Complex Networks and their Applications, Nov 2023, Menton Riviera, France. hal-04586400

**HAL Id: hal-04586400**

**<https://hal.science/hal-04586400v1>**

Submitted on 24 May 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Machine Learning for Developing Guidance to Improve Metaheuristic Algorithm

Bachtiar Herdianto<sup>1</sup>, Romain Billot<sup>1</sup>, Flavien Lucas<sup>2</sup>, and Marc Sevaux<sup>3</sup>

<sup>1</sup> IMT Atlantique, Lab-STICC, UMR 6285, CNRS, Brest, France  
{bachtiar.herdianto, romain.billot}@imt-atlantique.fr,

<sup>2</sup> IMT Nord Europe, CERI Systèmes Numériques, Douai, France  
flavien.lucas@imt-nord-europe.fr,

<sup>3</sup> Université Bretagne Sud, Lab-STICC, UMR 6285, CNRS, Lorient, France  
marc.sevaux@univ-ubs.fr

## 1 Introduction

In this study, we develop guidance based on supervised machine learning to improve the metaheuristic algorithm for solving capacitated vehicle routing problem (CVRP). The CVRP is a combinatorial optimization problem that determines a least-cost set of routes from a depot node for a fleet of capacitated vehicles to meet the demands of a set of customer nodes [1]. Mostly, algorithms for tackling this problem still solve from scratch, even for the same problem type. Meanwhile, leveraging historical data could prove invaluable for achieving efficient and effective solutions [2]. Furthermore, the incorporation of machine learning (ML) offers the promise of real-time problem learning and guiding the algorithm towards efficient problem-solving [3]. Therefore, this research aims to achieve two primary objectives: (1) to understand the connection between the quality of the solutions, their features, and the associated problem instances, and (2) to construct an efficient learning process consolidated with a robust optimization algorithm to solve the problems effectively.

## 2 State-of-the-art

The CVRP is characterized as an undirected graph with a specific number of nodes, where the primary objective is to explore permutation sequences of nodes to identify the most cost-effective routes for each vehicle. These routes collectively should minimize the total cost while adhering to vehicle capacity constraints and fulfilling all customer requirements [1]. Here, the weight of each edge represents the Euclidean distances of the corresponding pair of nodes. In general, the complexity of the problem in terms of the number of variables (and computational time) increases with the number of edges in the graph. The CVRP is often applied in the fields of logistics transportation and good distributions [4]. Even with this, existing optimization algorithms typically solve the problem from scratch, even for a similar type of problem, and nothing useful is extracted from past solutions [3]. On the other hand, machine learning (ML) aims to use historical data for handling new circumstances [5]. Consequently, by integrating optimization algorithms with ML, the algorithm can learn knowledge from past experiences, effectively steering towards optimal solutions [6].

The concept of integrating machine learning (ML) and optimization algorithms can be categorized into three approaches: (1) end-to-end learning, (2) learning based on problem properties, and (3) learning repeated decisions [2]. The idea of end-to-end learning means the ML has a role as the optimization algorithm to solve the optimization problem. Yet, we may need a huge amount of resources when facing large-scale problems [7]. Another idea is to utilize the ML to develop offline guidance for the optimization algorithm. However, using this idea, the full potential of ML may still need to be fully unlocked during optimization processes [4]. And lastly, learning repeated decisions by constructing an in-loop ML-assisted optimization algorithm. This approach allows the algorithm to learn from its own decisions and adapt its behaviour for improved performance. However, its implementation necessitates the development of an efficient learning process combined with a simple yet powerful optimization algorithm to achieve enhanced solutions and reduced computation time.

### 3 Proposed Approach

In this study, we develop guidance based on supervised machine learning to improve the metaheuristic algorithm for solving the CVRP. The first step in this study is to identify a comprehensive set of features that influence the resulting solutions by the knowledge extraction process, using XML100 instances by [8]. Subsequently, we construct a guided metaheuristic algorithm, an improvement of the benchmark *Feature-Guided Multiple Neighborhood Search* (FG-MNS) algorithm proposed by [6], based on the most important feature derived from the knowledge extraction procedures. Moreover, to develop the metaheuristic algorithm, we will also incorporate the path relinking processes [9] to enhance the search capability of the optimization algorithm.

## 4 Results

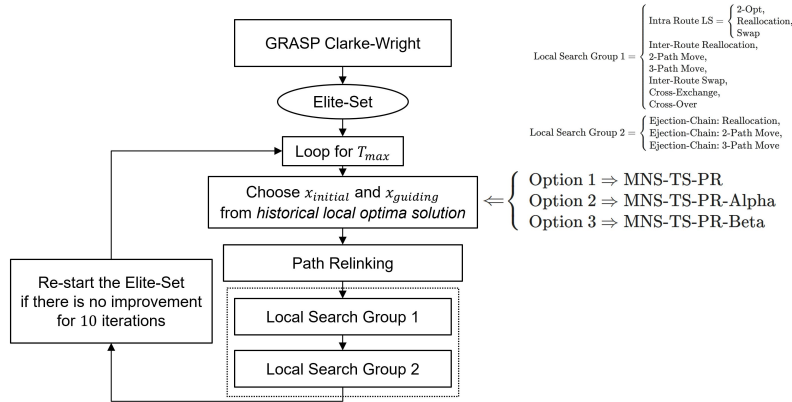
### 4.1 Knowledge Extraction Processes

The beginning step involves performing knowledge extraction processes using XML100 instances. In this context, we employ binary classification technique, segregating the data into optimal and near-optimal solutions. Optimal solution data is sourced from the optimal solution files for each instance. Meanwhile, near-optimal solution data is obtained through the use of the FG-MNS algorithm, as outlined by [6]. For the knowledge extraction processes, we perform the prediction model, using Gradient Boosting Classifier algorithm. As a result, we got 66,12% of prediction accuracy, with 0.6625 of F1 score. Moreover, we learn that the average capacity utilization and standard deviation of capacity utilization are more important than other features. Later, We also perform SHAP explainer to learn a bit more about the resulted model. Here, we learned that the larger capacity utilization for every route will result in a better quality solution of VRP. Therefore, we only consider those features to develop guidance for the proposed metaheuristic algorithm.

## 4.2 Integrating the Knowledge to the Metaheuristic Algorithm

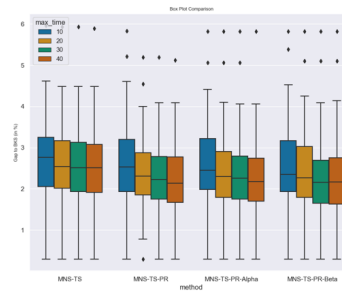
In this study, we introduce three variants of an enhanced FG-MNS [6]. These variants incorporate guidance in the selection of initial and guiding solutions, particularly when the path relinking process is initiated. The guidance mechanisms introduced are denoted as  $\alpha$  and  $\beta$ .

The  $\alpha$  value represents guidance derived from offline learning, while the  $\beta$  value is based on online learning, employing simple linear regression during the optimization process. The three proposed metaheuristic variants are as follows: MNS-TS-PR (an improved metaheuristic without guidance), MNS-TS-PR-Alpha (incorporating  $\alpha$  guidance), and MNS-TS-PR-Beta (utilizing  $\beta$  guidance). Here, during the path relinking, the the initial and guiding solutions are chosen as the most distant a particular value in the historical local optima. Where for; option 1 (MNS-TS-PR): the initial and guiding solutions are from the most distant value of solution  $s$  in the historical local optima; option 2 (MNS-TS-PR-Alpha): the initial and guiding solutions are from the most distant value of  $\alpha$  in the historical local optima; and option 3 (MN-TS-PR-Beta): the initial and guiding solutions are from the most distant value of  $\beta$  in the historical local optima.

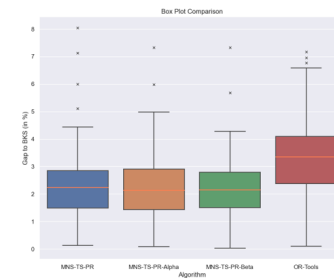


**Fig. 1.** The detailed mechanism about the proposed algorithm

The computational results are presented below. In Fig. 2, by solving XML100 instances within varying time constraints, our findings indicate that all the proposed metaheuristics improve the benchmark algorithm. Moreover, by using statistical t-test, we found that all three proposed method is statistically better than the benchmark algorithm. However, also by using t-test, we found that among three proposed algorithm, that although there are some slightly better results, there still no statistically significant improvement between with and without guidance among the proposed algorithms. Furthermore, in Fig. 3, by solving X instances, we know that the proposed algorithm achieves better results than the OR tools derived from [10]. Also, by performing t-test between all proposed algorithm and the OR-Tools, we found that all three proposed method is statistically better than the result from OR-Tools.



**Fig. 2.** Computational results using XML100 instances [8]



**Fig. 3.** Computational results on X instances [11]. Results for the OR-Tools is derived from [10]

## References

1. Labadie, N., Prins, C., Prodhon, C.: Metaheuristics for vehicle routing problems. John Wiley & Sons. (2016)
2. Bengio, Y., Lodi, A., Prouvost, A.: Machine learning for combinatorial optimization: a methodological tour d’horizon. *EJOR*. 290.2: 405-421 (2021)
3. Lucas, F., Billot, R., Sevaux, M., Sörensen, K.: Some insights about the use of machine learning for solving VRP. In *EURO 31*. (2021)
4. Herdianto, B., Komarudin.: Guided Clarke and Wright Algorithm to Solve Large Scale of Capacitated Vehicle Routing Problem. In *2021 IEEE 8th ICIEA*. pp. 449-453. (2021)
5. Russell, S.J.: Artificial intelligence a modern approach. Pearson Education, Inc. (2010)
6. Lucas, F., Billot, R., Sevaux, M., Sörensen, K.: Reducing space search in combinatorial optimization using machine learning tools. In *LION 14*. pp. 143-150. (2020)
7. Bogrybayeva, A., Meraliyev, M., Mustakhov, T., Dauletbayev, B.: Learning to solve vehicle routing problems: A survey. *arXiv:2205.02453*. (2022)
8. Queiroga, E., Sadykov, R., Uchoa, E., Vidal, T.: 10,000 optimal CVRP solutions for testing machine learning based heuristics. In *AAAI-22 ML4OR*. (2022)
9. Resende, M.G., Ribeiro, C.C., Glover, F., Martí, R.: Scatter search and path-relinking: Fundamentals, advances, and applications. *Handbook of metaheuristics*. pp. 87-107. (2010)
10. Vidal, T.: Hybrid genetic search for the CVRP: Open-source implementation and SWAP\* neighborhood. *CAOR*. 140. (2022)
11. Uchoa, E., Pecin, D., Pessoa, A., Poggi, M., Vidal, T., Subramanian, A.: New benchmark instances for the capacitated vehicle routing problem. *EJOR*. 257. (2022)