

Setting the Configuration Parameters of the Algorithm for the Periodic Vehicle Routing Problem by HPC Power

Ekaterina Grakova^{1,a}, Jan Martinovič¹, Kateřina Slaninová¹, Kateřina Janurová¹, Vojtěch Cima¹, Martin Golasowski¹, Roberto Montemanni², Matteo Salani³

¹*IT4Innovations National Supercomputing Center, VŠB - Technical University of Ostrava, 17. listopadu 15/2172, 708 33 Ostrava, Czech Republic*

²*University of Modena and Reggio Emilia, Viale A. Allegri 9, 42121 Reggio Emilia, Italy*

³*Dalle Molle Institute for Artificial Intelligence (IDISA, USI/SUPSI), Galleria 2, 6928 Manno, Switzerland*

Abstract. The quality of an optimal solution of the Vehicle Routing Problem is strongly depended on the setting of the configuration parameters of the algorithm. The paper is focused on the introduction of hyperparameter search for solving the Vehicle Routing Problem using a HyperLoom platform for defining and executing scientific pipelines in a distributed environment. To give a concrete example, we focused on Periodic Vehicle Routing Problem for the waste collection. HyperLoom platform was used to define and execute the hyperparameters sweep pipeline. The heuristic algorithm was tested on a real benchmark of the waste collection in Ostrava, Czech Republic. The aim of our case was to effectively combine the minimization of the total travelled distance and the optimization of the fairness of the routes in terms of the standard deviation of a tour length. The waste collection problem was very extensive and computationally demanding, so it was necessary to use high performance computing architecture for testing a large number of different settings of configuration parameters. The experiments were run on the supercomputer Salomon operated by IT4Innovations National Supercomputing Center in the Czech Republic.

1 Introduction

In 2017, the Czech Republic produced 34.5 million tonnes of all waste, of which 1.5 million tonnes of hazardous waste. The Czech Republic is very intensively engaged in waste recycling. From the 34.5 million tonnes of all waste, 84% was recycled, of which 80.5% was material and 3.6% as energy [1]. Every resident in the Czech Republic produces 3,259 kg of all types of the waste per year, of which about 531 kg is the mixed waste. Ostrava is the third largest city in the Czech Republic. Therefore, the problem of the waste collection is very important.

The waste management process is very complex, involving a range of technological methods and disciplines. Collecting problems in transport planning are usually solved by graph theory. One of the types of combinatorial optimization is the Traveling Salesman Problem (TSP), an extension of the TSP problem is Vehicle Routing Problem (VRP) [2], where the problem aims on creating a route of several vehicles with a limited capacity to determine the number of customers.

In this paper, we used the Periodic Vehicle Routing Problem (PVRP). The typical planning period is one day for the basic formulation of VRP. In this case of PVRP, the classical planning is generalized for M days. The difference is that a vehicle is not returning to a depot on the same day (for example, the day it left the depot). In

this type of the problem, the goal is to minimize the number of vehicles and time needed to serve the customers (for example, the cost for all rounds). Each customer has a given daily demand, which must be satisfied by one vehicle and by one route. If $M = 1$, we are talking about typical traffic delivery problems. In the case of PVRP, the demand of individual customers can be clustered into groups, creating an individual route for each day. Each customer must be served k -times, where $1 \leq k \leq M$. In the literature, PVRP has received considerable attention. The first works on PVRP were presented by the authors [3], and then there are more papers on this type of the problem [4,5]. Periodic deliveries are used in a lot of applications, including courier services, elevator maintenance and repair, vending machine replenishment, the collection of waste or the delivery of interlibrary loan material.

Heuristic and metaheuristic algorithms are used to solve the problem of waste collection. These algorithms need to set the configuration parameters that improve the quality of the provided solution. In this work, the heuristic algorithm that was used for an exhausted oil collection problem in the area of Bali, Indonesia [6,7] will be used. The formulation types of the problems are very similar. In heuristic algorithm, nine configuration parameters for each type of waste need to be set. Datasets were generated on the basis of cooperation with the collecting company in Ostrava. In total, 28,800

^a Corresponding author: ekaterina.grakova@vsb.cz

combinations of configuration parameters will be used for algorithm testing.

The aim of this article is to introduce an implementation of a distributed hyperparameter search for the vehicle routing problem solver based on HyperLoom [8]. Using HyperLoom pipeline we are able to sweep through large parameter spaces and discover the optimal solution for the heuristic algorithm. HyperLoom can discover the optimal solution for the heuristic algorithm.

The waste collection problem is very extensive and computationally demanding, so it is necessary to use high performance computing (HPC) architecture for testing a large number of different settings of configuration parameters. The experiments were run on the supercomputer Salomon operated by IT4Innovations.

2 Periodic VRP for waste collection

Betrami and Bodin devised the PVRP problem of waste collection for the city of New York and Washington DC [3]. The first authors represented the concept of PVRP for the daily planning of collection routes. The aim of this concept of Beltrami and Bodin is to minimize the time to be travelled and to select the feasible number of vehicles to be used instead of minimizing the costs of distribution. The solution of the problem was modified by Clarke-Wright algorithm. A Heuristic algorithm is combined here with randomized search procedures [3].

The PVRP for waste collection represented Russel and Igo, who developed heuristic algorithm assigning points to days of the week in order to minimize weekly travel distance (time). The problem presented herein is expanded to more than 750 points [9].

The four-phase approach to the PVRP for the waste collection was presented by Russel and Gribbin in [10], the number of different solution techniques by Chao, Golden, and Wasil in [11]. This article, the aim is to effectively combine the minimization of the total travelled distance and the optimization of the fairness of the routes in terms of the standard deviation of a tour length.

3 Setting the configuration parameters of PVRP algorithm

The authors in the articles [3,12,13] used heuristics or metaheuristics algorithms for the PVRP. The Clark-Wright algorithm is one of the classic heuristics algorithms that is used to solve the VRP and is based on the notion of savings. The solution of the task in Clarke-Wright algorithm is done through the iterations, incremental steps. First of all, the algorithm finds the least effective solution, which is improving each successive step. The solutions carried out in gradual steps monitor and patrol a defined limiting condition. Local Search (LS) is an algorithm, which iteratively translates from solution to solution in a set of acceptable solutions through local changes until we reach the solution, which we consider to be optimal or does not expire the specified search time. Tabu Search (TS) is a metaheuristic

algorithm, which uses LS with other rules. LS has tended to end up in a local optimum or in the plains, where several solutions are of equal value. TS, therefore, improves LS by changing the basic rules. No step in the environment would improve the solution, we will use a worsening step (this way we can get from the local optimum), and we will introduce bans (hence the notion of tabu). The genetic algorithm (GA) is the metaheuristic algorithm, which attempts to copy the natural selection process. GA belongs to a broader set of so-called evolutionary algorithms, which mimic techniques from nature such as inheritance, mutation, selection, or crossing.

Optimization algorithms for solving VRP problem involve a large number of design choices and algorithm-specific parameters that need to be set to reach their best performance. The maximum performance of optimization algorithms may depend on the setting of tens to hundreds of parameters. There are a lot of methods to solve the setting of the configuration parameters of the VRP algorithm. Authors in [12] used heuristic search techniques for the setting of the configuration parameters for the metaheuristics algorithms. Another group of parameters tuning of optimization algorithms is focused on statistical modelling approaches [13]. Authors in [12] proposed machine learning approaches.

In machine learning, there exists the term hyperparameters. Hyperparameters present configuration variables that are external to the model and whose values cannot be estimated from data. Hyperparameters tuning can be done using Grid Search, Randomized Search or Bayesian optimization.

4 Test case for PVRP waste collection

In this work, we used mixed waste, paper waste, plastic waste, and glass waste. We prepared a set of 3,000 different locations for the mixed waste, a set of 2,000 different locations for the plastic waste, a set of 1,500 different locations for the paper waste and a set of 1,000 different locations for the glass waste. For each location, we know the expected amount of the waste that accumulates within one week. The travel times between any two locations are known. Every waste collection tour starts at the depot with the empty vehicle. Every vehicle starts from the depot, picks-up the waste from the stops until they are full, dumps them at a dump site. The dump site and the depot for the vehicle are at the same GPS location. Each GPS location is one collection point for the vehicle. The travel time for each route is not allowed to exceed a given bound on working hours corresponding to eight hours. The metric distances were created for each dataset in Ostrava. The situation is displayed in Figure 1.

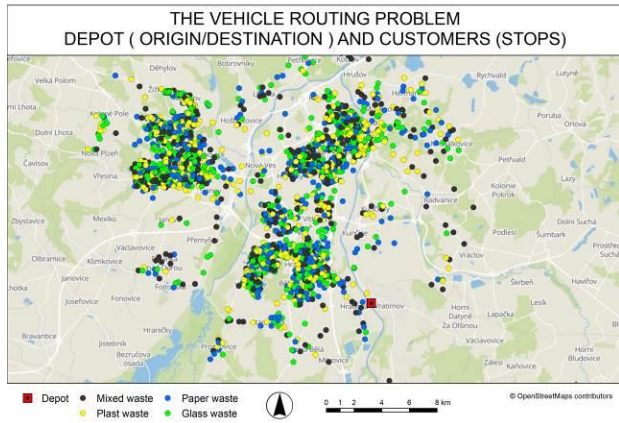


Figure 1. GPS locations in Ostrava.

Planning time for the waste collection was 14 days, including the weekend. The type of collection vehicles depends on the type of the waste.

4.1 HyperLoom for heuristic algorithm

We have explored 7,200 combinations of the configuration parameters for each type of the waste resulting in total of 28,800 unique parameter configurations. We developed a HyperLoom pipeline (see Figure 2) that allows to execute particularly configured instances of the heuristic algorithm in a parallel fashion within a distributed system and dynamically load-balance the workload fairly between the nodes as the instances are being finished.

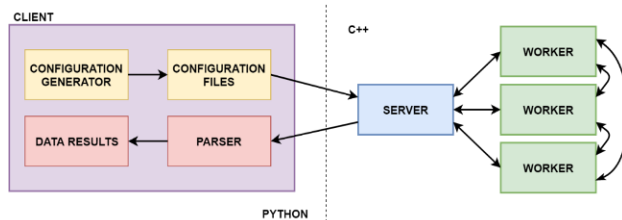


Figure 2. HyperLoom pipeline.

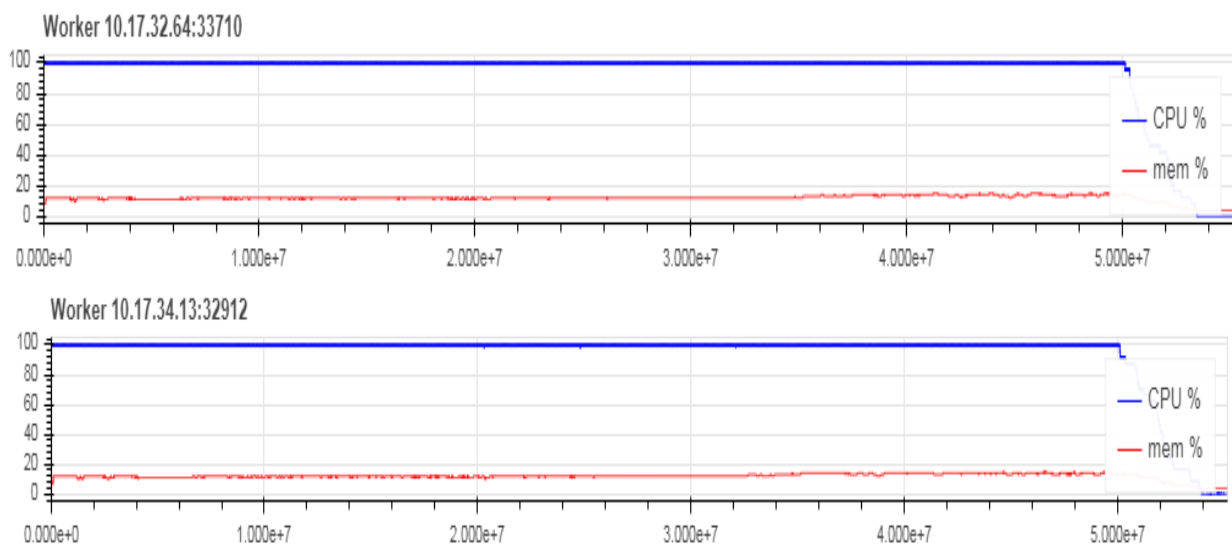


Table 1 contains the settings of configuration parameters for particular types of the waste.

Table 1. Configuration parameters setting for hyperparameter search.

	Mixed waste	Plastic waste	Paper waste	Glass waste
Number of vehicles	50, 51, 52, 53, 54	30, 31, 32, 33, 34	23, 24, 25, 26, 27	23, 24, 25, 26, 27
Depot number	0			
Weight	0.05, 0.50, 0.95			
Random seed	1, 2, 3, 4			
Horizon days	14			
Vehicles capacity	34, 35, 36, 37			
Location per day	100 -109			
Litres per day	200, 250, 300			

The experiments were performed on Salomon supercomputer, where each node has two-twelve-core Intel Xeon processors and 128 GB RAM (2xIntel Xeon E5 – 2680 v3, 2.5 GHz, 12 cores). We have carried out all the experiments on 56 identical physical computational nodes.

Figure 3 illustrates the CPU and memory utilization of a subset of computational nodes during the pipeline execution.

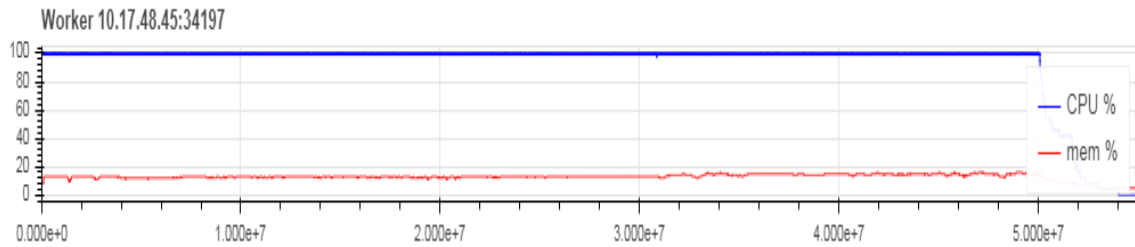


Figure 3. The CPU and memory utilization of computational nodes subset during pipeline execution.

4.2 Test case results

The goal of our work is to optimize the fairness of the routes in terms of the standard deviation of a tour length and to minimize the total of travelled distance. That means that we are looking for a combination of input parameters for which the resulting standard deviation of a tour length value is minimal. If there are more such combinations, we choose one that has a minimum total travelled distance. Figure 4 contains the results of the calculations. The required values of standard deviation vs. total travelled distance are found as far as possible at the bottom left for each type of waste.

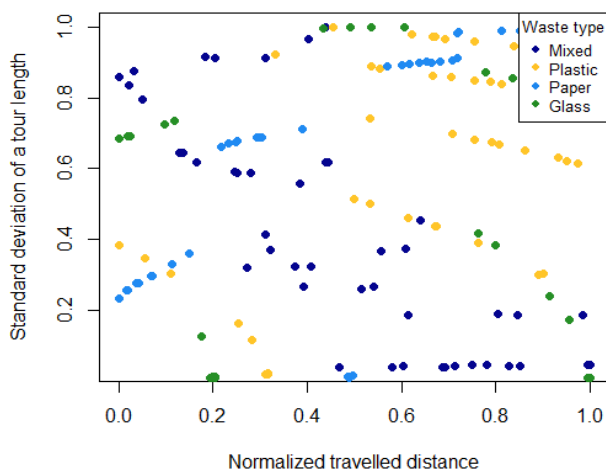


Figure 4. Test case results.

As can be seen from Figure 4, the results for different types of waste show different behaviour. For mixed waste, unlike paper and plastic, there is an increase in the total travelled distance when the standard deviation decreases. This behaviour is also shown by the value of the correlation coefficient $\rho = -0.79$. Plastic and paper waste show the same type of dependency, when the standard deviation increases, the total travelled distance also increases, with the value of correlation coefficients $\rho = 0.43$ and $\rho = 0.59$, respectively. For the glass waste, there is no linear dependence between standard deviation and total travelled distance with the value of correlation parameter $\rho \approx 0$.

The best configuration of input parameters is listed in Table 2. There are listed only those configuration parameters whose values have affected the results

because in our case, we deal with a real dataset, the values in Table 2 can allow waste collection companies to reduce their overall waste collection costs.

Table 2. The best configuration parameters setting.

	Mixed waste	Plastic waste	Paper waste	Glass waste
Number of vehicles	53	30	27	27
Weight	0.95	0.95	0.50	0.50
Litres per day	300	250	300	300

5 Conclusion

The goal of our work was to optimize the fairness of the routes in terms of standard deviation of a tour length and to minimize the total travelled distance. We solved this problem by taking the standard deviation as a priority parameter for minimizing. In practice, companies can put more emphasis on minimizing total travelled distance, or on combining both parameters with different weights. In this case, the problem could be solved by the Euclidean distance approach.

Acknowledgements

This work was supported by The Ministry of Education, Youth and Sports from the National Programme of Sustainability (NPS II) project “IT4Innovations excellence in science - LQ1602” and by the IT4Innovations infrastructure which is supported from the Large Infrastructures for Research, Experimental Development and Innovations project “IT4Innovations National Supercomputing Center – LM2015070”, and partially by the SGC grant No. SP2019/108 “Extension of HPC platforms for executing scientific pipelines”, VŠB - Technical University of Ostrava, Czech Republic.

References

1. Ministry of the Environment of the Czech Republic: <https://www.mzp.cz/cz/news/>. Accessed 17 Oct. 2017
2. G.B. Dantzig, J. H. Ramser. The Truck Dispatching Problem, *Management Science*, **6**, 80–9 (1959)
3. E. J. Beltrami, L.D. Bodin. Networks and vehicle routing for municipal waste collection. *Networks* **4**, chapter 1, 65–94 (1974)

4. S. Coene, A. Amout, C. Frits Spijksma. The periodic vehicle routing problem. Department of decision sciences and information management (2008)
5. R. Cristian Abreu, J. Elias C. Arroyo. An application of ILS heuristic to periodic vehicle routing problem with heterogeneous fleet and fixed costs. IEEE. 1-9 (2015)
6. D. Weyland, M. Salani, R. Montemanni, L.M. Gambardella. Vehicle routing for exhausted oil collection. *Journal of Traffic and Logistics Engineering* **1**, chapter 1, 5-8 (2013)
7. E. Grakova, M. Golasowski, R. Montemanni, K. Slaninová, J. Martinovič, J. Jamal, K. Janurová, M. Salani. Hyperparameter Search in Periodic Vehicle Routing Problem. *Int. Conf. on Intelligent Traffic and Transportation (ICITT 2018)* (2019)
8. V. Cima, S. Böhm, J. Martinovič, J. Dvorský, K. Janurová, T. V. Aa, ... & V. Chupakhin. HyperLoom: a platform for defining and executing scientific pipelines in distributed environments. In *Proceedings of the 9th Workshop and 7th Workshop on Parallel Programming and RunTime Management Techniques for Manycore Architectures and Design Tools and Architectures for Multicore Embedded Computing Platforms*. 1-6 (2018)
9. R. A. Russell, W. Igo. An assignment routing problem, 2 (1979)
10. R. A. Russell, D. Gribbin. A multiphase approach to the period routing problem (1991)
11. I. - M. Chao, B. L. Golden, E. Wasil. An improved heuristics for the period vehicle routing problem. (1995)
12. P. L. N. U. Cooray, T. D. Rupasinghe. Machine Learning-Based Parameter Tuned Genetic Algorithm for Energy Minimizing Vehicle Routing Problem. *Journal of Industrial Engineering*, 13 (2017)
13. M. Mirabi, N. Shokri, A. Sadeghieh. Modeling and Solving the Multi-depot Vehicle Routing Problem with Time Window by Considering the Flexible end Depot in Each Route. **3**, 1373-1390 (2016)