Interactive Utility Maximization in Multi-Objective Vehicle Routing Problems: A “Decision Maker in the Loop”-Approach

Martin Josef Geiger, Wolf Wenger, and Walter Habenicht

Abstract—The article presents an interactive multi-criteria approach for the resolution of rich vehicle routing problems. A flexible framework was built to be able to deal with various components of general vehicle routing problems, e.g. the consideration of multiple objectives or different types of specific complex side constraints such as time windows, multiple depots or heterogeneous fleets. In the framework, a local search approach on the basis of variable neighborhood search (VNS) constructs and improves solutions in real time. The decision maker is actively involved into the resolution process as the system allows the interactive articulation of preference information, influencing the global utility function that guides the search. Results of test runs on multiple depot multi-objective vehicle routing problems with time windows are reported, simulating different types of decision maker behaviors.

I. INTRODUCTION

The vehicle routing problem (VRP) is one of the classical optimization problems known from operations research with numerous applications in real world logistics. In brief, a given set of customers has to be served with vehicles from a depot such that a particular criterion is optimized. The most comprehensive model therefore consists of a complete graph \( G = (V, A) \), where \( V = \{v_0, v_1, \ldots, v_n\} \) denotes a set of vertices and \( A = \{(v_i, v_j) \mid v_i, v_j \in V, i \neq j\} \) denotes the connecting arcs. The depot is represented by \( v_0 \), and \( m \) vehicles are stationed at this location to service the customers \( v_1, \ldots, v_n \). Each customer \( v_i \) demands a nonnegative quantity \( q_i \) of goods and service results in a nonnegative service time \( d_i \). Traveling on a connecting arc \((v_i, v_j)\) results in costs \( c_{ij} \) or travel time \( t_{ij} \). The most basic vehicle routing problem aims to identify a solution that serves all customers, not exceeding the maximum capacity of the vehicles \( Q_k \) and their maximum travel time \( T_k \) while minimizing the total distances/costs of the routes.

Various extensions have been proposed to this general problem type. Most of them introduce additional constraints to the problem domain such as time windows, defining for each customer \( v_i \) an interval \([e_i, l_i]\) of service. While arrival before \( e_i \) results in a waiting time, arrival after \( l_i \) is usually considered to be infeasible [21]. In other approaches, the time windows may be violated, leading to a tardy service at some customers. Violations of time windows are either integrated in the overall evaluation of solutions by means of penalty functions [5], or treated as separate objectives in multi-objective approaches [6].

Some problems introduce multiple depots as opposed to only a single depot in the classical case. Along with this sometimes comes the additional decision of open routes, where vehicles do not return to the place they depart from but to some other depot. Also, different types of vehicles may be considered, leading to a heterogeneous fleet in terms of the abilities of the vehicles.

Unfortunately, most problems of this domain are \( NP \)-hard. As a result, heuristics and more recently metaheuristics have been developed with increasing success [8], [17], [18]. In order to improve known results, more and more refined techniques have been proposed that are able to solve, or at least approximate very closely, a large number of established benchmark instances [2]. It has to be mentioned however, that with the increasing specialization of techniques a decrease in generality of the resolution approaches follows.

While the optimality criterion of minimizing the total traveled distance is the most common, more recent approaches recognize the vehicle routing problem as a multi-objective optimization problem [6], [9], [10], [12], [20]. Important objectives besides the minimization of the total traveled distance are in particular the minimization of the number of vehicles in use [13], the minimization of the total tardiness of the orders, and the equal balancing of the routes [11]. Following these objectives, it is desired to obtain solutions that provide a high quality of delivery service while minimizing the resulting costs. As many objectives are however of conflicting nature, not a single solution exists that optimizes all relevant criteria simultaneously. Instead, the overall problem lies in identifying the set of Pareto-optimal solutions \( P \) and selecting a most-preferred solution \( x^* \in P \).

In this context, three different general strategies of solving multi-objective optimization problems can be implemented:

1) \textit{A priori} approaches reduce the multi-objective problem to a single-objective surrogate problem by formulating and maximizing a utility function. The advantage of this approach can be seen in its simplicity given the possibility to specify the precise utility function of the decision maker. The concept may however not be used if the decision maker is not able to state his/her preferences in the required way.

2) \textit{A posteriori} approaches first identify the Pareto set \( P \), and then allow the decision maker to select a most-preferred solution \( x^* \in P \). The main advantage of this resolution principle is, that the computation of the optimal solutions can be done offline without the immediate participation of the decision maker. A large number of elements of the Pareto set are on the...
other hand discarded later during the decision making procedure.

3) Interactive approaches allow the gradual articulation of preferences by the decision maker and compute a sequence of solutions based on his/her individual statements. Several advantages result from this concept. First, the computational effort is smaller in comparison to the identification of the entire Pareto set. Second, the gradual articulation of preferences allows the decision maker to reflect the chosen settings in the light of the obtained results and therefore adapt and react to the optimization procedure. A disadvantage of interactive multi-objective optimization procedures is however the need of the presence of a decision maker and the availability of an interactive software to present the results. Also, comparably little time for computations is allowed as the system should be able to react in (almost) real-time to inputs of the decision maker.

While it has been stressed already quite early, that combining computer programs with interactive planning procedures may be a beneficial way of tackling complex routing problems [14], [15], [22], [23], research in interactively solving multi-objective metaheuristics is a rather newly emerging field of research [16] with so far comparably few applications. Given the increasing computing abilities of modern computers however, approaches can become increasingly interesting as they allow the resolution of complex problems under the consideration of interactive, individual guidance towards interesting solutions.

The article is organized as follows. In the following Section II, a framework for interactive multi-objective vehicle routing is presented that aims to address two critical issues:

1) The necessary generality of resolution approaches when trying to solve a range of problems of different characteristics.
2) The integration of multiple objectives and the consideration of individually articulated preferences of the decision maker during the resolution procedure of the problem.

An implementation of the framework for multi-objective vehicle routing problems is presented in Section III. The system is used to solve instances of multi-objective vehicle routing problems. Conclusions are presented in Section IV.

II. A FRAMEWORK FOR INTERACTIVE MULTI-OBJECTIVE VEHICLE ROUTING

Independent from the precise characteristics of the particular VRP, two types of decisions have to be made when solving the problem.

1) Assignment of customers to vehicles (clustering).
2) Construction of a route for a given set of customers (sequencing).

It is well-known that both types of decisions influence each other to a considerable extent. While the clustering of customers to vehicles is an important input for the sequencing, the sequencing itself is of relevance when adding customers to routes as constraints of maximum distance have to be respected. The two types of decisions can be made either sequential (cluster first-route second vs. route first-cluster second) or in parallel.

Therefore, the framework presented here proposes the use of a set of elements to handle this issue with utmost generality. Figure 1 gives an overview about the elements used.

- The marketplace represents the element where orders are offered for transportation. This element is particularly necessary to allow an exchange of information gathered during the execution of the optimization procedure.
- Vehicle agents place bids for orders on the marketplace. These bids take into consideration the current routes of the vehicles and the potential change when integrating an additional order. Integrating additional orders into existing routes may lead to a deterioration of performance in terms of the underlying objectives, e.g. by increasing traveled routes and/or time window violations. This information is reported back to the marketplace.
- An ontology describes the precise properties of the vehicles such as their capacity, availability, current location, etc. This easily allows the consideration of different types of vehicles. It also helps to model open routes, where vehicles do not necessarily return to the depot where they depart from.
- A decider communicates with the human decision maker via a graphical user interface (GUI) and stores his/her articulated individual preferences. The decider also assigns orders to vehicles, taking into consideration the bids placed for the specific orders.

The described framework is constructing a solution by combining clustering and sequencing decisions in parallel. Therefore orders are placed on the marketplace. To assign

![Fig. 1. Illustration of the framework for interactive multi-objective vehicle routing](image-url)
these orders to vehicles, bids from the vehicle agents for the given orders are needed. Which bid to place is calculated by each vehicle agent while constructing a route by using local search heuristics. The clustering process is done while constantly updating the bids to identify a solution which maximizes the stored individual preferences of the decision maker.

During the construction of the solution the decision maker is kept in the loop. If the presented and visualized compromise alternative is not satisfying, he/she is able to change the articulated preferences. In that case, the decider updates the stored preference information and in consequence, the vehicles resequence their orders such that the updated preference information is met.

III. IMPLEMENTATION AND EXPERIMENTAL INVESTIGATION

A. Configuration of the system

The framework has been implemented in a computer system. In the experiments that have been carried out, two objective functions are considered, the total traveled distance IDIST and the maximum tardiness T\text{max} caused by a vehicle arriving at a costumer v_i after the upper bound t_i of the corresponding time window. It should be noticed however, that neither the concept presented in Section II nor the actual implementation are restricted to two objective functions only. However, a sensible choice had to be made in order to investigate the system in a quantitative way in a controllable experimental setting.

The representation of the decision maker’s individual preferences is implemented using a utility function that aggregates the partial utilities of both objectives. By articulating the relative importance \( w_{DIST} \) of the total traveled distance DIST, the decision maker’s individual relationship between the two objectives can be formulated in a weighted sum approach. Therefore the overall utility \( UTILITY \) of a particular solution can be computed as given in Expression (1).

\[
UTILITY = \frac{w_{DIST}}{BDIST} \frac{\text{DIST}}{BDIST} + (1 - \frac{w_{DIST}}{BDIST}) \frac{T_{\text{max}}}{T_{\text{max}}} (1)
\]

\[
u_{DIST}(DIST) = \frac{UBDIST - \text{DIST}}{UBDIST - LBDIST} (2)
\]

\[
u_{T_{\text{max}}}(T_{\text{max}}) = \frac{UBT_{\text{max}} - T_{\text{max}}}{UBT_{\text{max}} - LB_{T_{\text{max}}}} (3)
\]

After having carried out initial experiments aggregating directly two objective functions without formulating partial utility functions [7], we chose to introduce the partial utility functions given in Expression (2) and (3). These functions compute utility values based on upper and lower bounds of the total distance \((UBDIST, LBDIST)\) and the maximum tardiness \((UBT_{\text{max}}, LB_{T_{\text{max}}})\). The upper bound of the maximum tardiness is derived by computing the latest possible arrival of a vehicle at each customer such that the vehicle may return to the depot within the maximum allowed travel time and subtracting the corresponding \(t_i\). The lower bound of the maximum tardiness is assumed to be 0. For the upper bound of the total distance we computed the direct delivery of each customer from the depot with the assumption of returning directly to the depot. A lower bound of the total distance is derived from the shortest solution of initial experiments [7] and reducing the value by 10%.

The vehicle agents are able to modify the sequence of their orders using four different local search neighborhoods.

- Inverting the sequence of the orders between positions \(p_1, p_2\), \(p_1 \neq p_2\). While this may be beneficial with respect to the distances, it may pose a problem for the time windows as orders are usually served in the sequence of their time windows.
- Exchanging the positions \(p_1, p_2\), \(p_1 \neq p_2\) of two orders.
- Moving an order from position \(p_1\) and reinserting it at position \(p_2\), \(p_1 < p_2\) (forward shift).
- Moving an order from position \(p_1\) and reinserting it at position \(p_2\), \(p_1 > p_2\) (backward shift).

In each step of the local search procedure, a neighborhood is randomly picked from the set of neighborhoods and a move is computed and accepted given an improvement. We select each neighborhood with equal probability of \(\frac{1}{4}\).

Bids for orders on the marketplace are generated by the vehicle agents, taking into consideration all possible insertion points in the current route. The weighted sum of the decrease in the partial utilities of the distance \(u_{DIST}\) and maximum tardiness \(u_{T_{\text{max}}}\) gives the prize for the order. This price reflects the individual preferences articulated by the decision maker using the \(w_{DIST}\) parameter which expresses the tradeoff between the total distance and maximum time window violations.

The decider assigns orders to vehicles such that the maximum regret when not assigning the order to a particular vehicle, and therefore having to assign it to some other vehicle, is minimized. It also analyzes the progress of the improvement procedures. Given no improvement for a certain number of iterations, the decider forces the vehicle agents to place back orders on the market such that they may be reallocated. In the current setting, the vehicle agents are allowed to compute 1000 neighboring solutions without any further improvements before they are contacted by the decider to place back one order on the marketplace. The order to be placed back is the one of the current route that, when removing it from the route, leads to the biggest improvement with respect to the overall evaluation of the route.

B. Experiments

The optimization framework has been tested on ten benchmark instances taken from [3]. The instances range from 48 to 288 customers that have to be served from 4 to 6 depots, each of which possesses 2 to 7 vehicles. The precise description of the instances is given in [3] and therefore not repeated here. Download of the problem files is e.g.
We simulated a decision maker changing the individual relationship between the two considered objectives by adapting the articulated parameter \(w_{DIST}\) during the optimization procedure. First we assumed a decision maker starting with a value of \(w_{DIST} = 1\) and successively decreasing it to 0, second a decision maker starting with a \(w_{DIST} = 0\) and increasing it to 1, and third a decision maker starting with a \(w_{DIST} = 0.5\), increasing it to 1 and decreasing it again to 0. For the instances 1 to 4 the values of \(w_{DIST}\) were adjusted in steps of 0.01, while waiting with a following adjustment until the system has reached (at least) a local optimum, i.e. no improvement has occurred in the last iteration. Instances 5 to 10 were tested for the three simulated decision makers using steps of 0.1 each, but again leaving enough time for computations to allow a convergence of the determined solution.

Every time the decision maker changes the relative importance of the total traveled distance, the system has to follow the updated preference information. Therefore a process of resequencing and reassigning of the customers is taking place using the implemented local search metaheuristics.

Figure 2 to 6 plot the results obtained during the described test runs.

It can be seen, that the results are significantly different depending on the initial chosen value of \(w_{DIST}\). For initial values of \(w_{DIST} = 0.5\), the framework is able to identify solutions with mostly higher values of overall utility compared to other initial parameter settings. Therefore the results obtained by the third simulated decision maker are more closely to the Pareto front than those of decision maker one and two. Moreover, choosing an initial value of \(w_{DIST} = 1\) and decreasing it constantly tends to lead to better solutions than adapting \(w_{DIST}\) from 0 to 1.

To illustrate this behavior more detailed, we are going to discuss the results for instance ‘1a’ more closely and verbally. The first decision maker starts with \(DIST = 1027\), \(T_{max} = 575\). Since the chosen lower bound of the total distance \(LB_{DIST}\) is 852, this leads to \(u_{DIST} = UTILITY = 0.895\). The system is able to improve this solution to \(DIST = 975\), \(T_{max} = 551\), \(UTILITY = 0.926\) while keeping \(w_{DIST} = 1\). Changing \(w_{DIST}\) step by step causes a significant improvement in terms of maximum tardiness and leads to \(T_{max} = 352\) for \(w_{DIST} = 0.89\). Due to the increasing total traveled distance (\(DIST = 1007\)) and a relative importance of that objective, which is still high, the \(UTILITY\) decreases to 0.868. This process of improving solutions concerning the maximum time window violation and a value of \(u_{DIST}\) which tends to decrease can be noticed
while $w_{DIST}$ is adapted. However, when $w_{DIST}$ decreases below 0.61, the negative effect of $w_{DIST}$ on the overall utility is counterbalanced by a positive effect of $w_{T_{max}}$, so that UTILITY increases. Finally when decreasing $w_{DIST}$ from 0.25 to 0.24, a solution without any tardiness can be identified ($DIST = 1359$, $T_{max} = 0$). Since the lower bound of $T_{max}$ is $LB_{T_{max}} = 0$, $w_{DIST} = 0$ leads to an overall utility of 1.

The second decision maker starts with $DIST = 2953$, $T_{max} = 0$ and a total UTILITY = 1, as his initial parameter setting is $w_{DIST} = 0$. For $w_{DIST} = 0.48$ a solution with $DIST = 2225$ was found, where all costumers can still be visited within the given time windows. Obviously the UTILITY decreases with increasing relative importance of the the traveled distance. While there is a short improvement in terms of the overall utility when changing $w_{DIST}$ from 0.48 to 0.49, further steps lead to solutions of at most UTILITY = 0.742. Clearly, the first strategy outperforms the second. While an initial value of $w_{DIST} = 0$ allows the identification of a solution with zero tardiness, it tends to construct routes that, when decreasing the relative importance of the tardiness, turn out to be hard to adapt. In comparison to the strategy starting with a $w_{DIST} = 1$, the clustering of customers appears to be prohibitive for a later improvement.

When comparing the third strategy of starting with a $w_{DIST} = 0.5$, it becomes obvious that this outperforms both other ways of interacting with the system. Here, the solutions start with $DIST = 1272$, $T_{max} = 4.7$ and therefore an UTILITY = 0.871, go to $DIST = 964$, $T_{max} = 432$, UTILITY = 0.933 for $w_{DIST} = 1$, and finally to $DIST = 1255$, $T_{max} = 0$, UTILITY = 1 for $w_{DIST} = 0$. Apparently, starting with a compromise solution is beneficial even for both extreme values of $DIST$ and $T_{max}$.

While the three strategies become slightly more competitive in the following instances, especially in instance 4a, strategy three still remains the best one. For four of the ten instances strategy one actually was unable to find a solution without time window violations, whereas at least one of the other strategies succeeded during the adaptation of $w_{DIST}$. On the other hand, the best obtained value for the total traveled distance $DIST$ by strategy two was outperformed by both other strategies for all of the investigated instances. Again, starting with an extremal value of $w_{DIST}$ seems to be difficult, while an initial solution of $w_{DIST} = 0.5$ leads to higher or at least equal maximum values of overall utility in all cases.

The framework calculates similar results, regardless of whether the decision maker adjusts his/her preferences in rather larger or smaller steps, e.g. by choosing 0.1 or 0.01.
Independent from the initial choice of parameters, feasible solutions have been found by the system for all investigated instances. The quality of the obtained solutions have however differed significantly, depending on the choice of interaction with the system.

As a result of the experiments, it becomes clear that for the investigated case, a compromise value of around $w_{DIST} = 0.5$ should be chosen for the computation of a first solution before starting an interaction with the system. The so constructed alternative can be modified towards the minimization of the traveled distance as well as towards the minimization of the maximum tardiness. Other strategies of initially setting weight parameters such as $w_{DIST} = 0$ and $w_{DIST} = 1$ led to relatively weaker results. Whether the initial setting of $w_{DIST} = 0.5$ however is the only optimal choice cannot be determined as other parameter values have not been investigated yet.

Besides this theoretically gained insight, the contribution of the framework can also be seen in describing a gen-
eral concept for the resolution of complex vehicle routing problems. As practical problems often vary in terms of their characteristics, this may turn out to be beneficial when problems with different side constraints have to be addressed using a single optimization procedure. An additional use can be found for dynamic vehicle routing problems. The market mechanism provides a platform for the matching of offers to vehicles without the immediate need of accepting them, yet still obtaining feasible solutions and gathering a prize for acceptance of offers which may be reported back to the customer.

For the future development of the research carried out, more complex representations of global utility functions could be investigated. This can include nonlinear aggregations as well as the integration of aspiration levels, to mention a few.

Also, the market mechanism has to be improved with respect to the clustering of the customers as the results clearly indicate a certain bias of the solutions towards initial weight settings. More complex reallocations of orders between vehicles will be considered as opposed to the here presented concept.

REFERENCES


