TQrouting: A Hybrid, Adaptive Framework for General Vehicle Routing Problems

Giorgi Tadumadze,* Mikhail Somov, Katsiaryna Tsarova, and Michael Perelshtein

Terra Quantum AG, Kornhausstrasse 25, 9000 St. Gallen, Switzerland

E-mail: gt@terraquantum.swiss

Abstract

TQrouting is an in-house route optimization engine developed at Terra Quantum, designed to efficiently and robustly solve a wide spectrum of Vehicle Routing Problem (VRP) variants. The solver combines mathematical optimization components from state-of-the-art routing heuristics with selected machine-learning concepts.

To handle large-scale instances, TQrouting employs an in-house machine-learning-based decomposition approach that recursively divides a large VRP instance into subproblems by identifying spatial and temporal clusters of customers. The number and shape of clusters are chosen so that each cluster forms a coherent subproblem, i.e., customers assigned to different clusters would be unlikely to be routed together in an optimal solution. Clusters that remain too large are split further, forming a hierarchical decomposition tree in which the root represents the original instance (full customer set and available fleet), internal nodes represent subproblems (subsets of customers with a corresponding number of vehicles), and leaf nodes are the smallest subproblems solved without further decomposition. Once all subproblems of the same parent node are solved, their solutions are merged and further improved at higher tree levels while preserving high-quality substructures found at lower levels. This divide-and-conquer strategy improves scalability and enables parallelization because independent subproblems can be solved concurrently. Conceptually, this contrasts with decomposition strategies that start from an incumbent global solution and then decompose and refine it (e.g., (Santini et al., 2023)), whereas our decomposition proceeds bottom-up by first identifying strong structures on small subproblems, then merging them and using the merged solutions as warm starts for higher-level subproblems. It is worth mentioning that this decomposition approach is largely solver-agnostic and can be paired with a wide range of VRP solution methods as subproblem solvers, ranging from (meta-)heuristics to exact approaches, provided they can exploit an incumbent after merging, which serves as a warm start.

The optimization engine of TQrouting combines fast construction heuristics designed to handle very large-scale instances following the scalability principles highlighted by Arnold et al. (2019), enabling high-quality initial solutions using problem-aware constructive strategies. In addition, it uses randomized initial-solution generators to increase diversity and explore a broader region of the solution space, which is particularly important for population-based improvement methods. The initial solution construction is followed by an iterative improvement phase that employs a range of intensification and diversification mechanisms. The balance between intensification and diversification is adjusted adaptively based on concrete instance characteristics, available time budget, and online indicators such as convergence behavior, solution diversity, and stagnation detection.

Intensification is achieved through a high-performance local search engine consisting of a variety of move operators inspired by several state-of-the-art VRP algorithms (Arnold et al., 2019; Vidal, 2022; Accorsi and Vigo, 2021, 2024; Máximo et al., 2024). The neighborhood structure and size can be extended dynamically during the search, in the spirit of variable neighborhood search (VNS). In particular, neighborhood sizes are adapted dynamically according to the current search phase to mitigate stagnation. This enables broader exploration in diversification phases and more focused neighborhoods during intensification. The local search implementation relies on efficient delta evaluation (amortized constant time for many supported moves), which enables parallel evaluation of multiple candidate moves to ensure high computational efficiency.

To escape from local optima, TQrouting employs several complementary diversification mechanisms, which are dynamically selected depending on the problem type, instance size, and search phase:

- **Hybrid Genetic Search (HGS)** follows key principles from Vidal (2022), including diversity management to maintain population variety and dynamic penalty management to adapt penalties for constraint violations during the search. This allows exploration of infeasible or partially feasible regions when beneficial, thereby balancing exploration and exploitation. Each individual goes through a local-search-driven education phase; in our implementation, multiple individuals can be educated in parallel.

- **Large Neighborhood Search (Ruin-and-Recreate)** is inspired by shaking/perturbation procedures proposed by state-of-the-art large-scale routing heuristics (Arnold et al., 2019; Accorsi and Vigo, 2021, 2024; Máximo et al., 2024). Conceptually, this procedure follows a large neighborhood search principle: a subset of customers is removed and then reinserted using insertion heuristics. The choice of destroy/repair operators and the affected region are adjusted adaptively depending on the current search phase. The degree of destruction is also adapted based on diversification needs, using solution-distance indicators to adapt the perturbation strength and overall diversification pressure (Máximo et al., 2024).

- **Adaptive Threshold Acceptance** is inspired by adaptive threshold acceptance mechanisms used in iterated local search frameworks Máximo et al. (2024). Instead of accepting only improving moves, TQrouting allows controlled acceptance of slightly inferior solutions under a dynamic threshold, thereby diversifying the search trajectory and helping escape deep local minima while still promoting convergence. Beyond the monotonically tightening threshold as proposed by Máximo et al. (2024), TQrouting provides several adaptive acceptance-control strategies that adjust the acceptance parameter online to react to stagnation and modulate exploration versus convergence throughout the search.

The optimization engine is implemented in modern C++ with a strong focus on computational efficiency and parallelization. This design enables TQrouting to exploit multi-core processors and deliver high-quality solutions quickly across a wide range of VRP instances.

References

Santini, A.; Schneider, M.; Vidal, T.; Vigo, D. Decomposition Strategies for Vehicle Routing Heuristics. *INFORMS Journal on Computing* **2023**, *35*, 543–559.

Arnold, F.; Gendreau, M.; Sörensen, K. Efficiently Solving Very Large-Scale Routing Problems. *Computers & Operations Research* **2019**, *107*, 32–42.

Vidal, T. Hybrid Genetic Search for the CVRP: Open-Source Implementation and SWAP* Neighborhood. *Computers & Operations Research* **2022**, *140*, 105643.

Accorsi, L.; Vigo, D. A Fast and Scalable Heuristic for the Solution of Large-Scale Capacitated Vehicle Routing Problems. *Transportation Science* **2021**, *55*, 832–856.

Accorsi, L.; Vigo, D. Routing one million customers in a handful of minutes. *Computers & Operations Research* **2024**, *164*, 106562.

Máximo, V. R.; Cordeau, J.-F.; Nascimento, M. C. V. AILS-II: An Adaptive Iterated Local Search Heuristic for the Large-Scale Capacitated Vehicle Routing Problem. *INFORMS Journal on Computing* **2024**, *36*, 974–986.