Fast Approximate Solutions using Reinforcement Learning for Dynamic Capacitated Vehicle Routing with Time Windows

Nazneen N Sultana, Vinita Baniwal, Ansuma Basumatary, Piyush Mittal, Supratim Ghosh, Harshad Khadilkar

nn.sultana@tcs.com

April 2021

Vehicle Routing Problem

tcs research

- The basic version involves computation of the optimal route for a single or multiple identical vehicles, given the nodes to be visited.
- Two basic variants of the problem are:
  - Static routing
  - Dynamic routing
- The variants that we have included in this paper are:
  - Vehicle capacity in terms of load and range.
  - Arbitrary number of customers and locations.
  - Arbitrary number of vehicles.
  - Customer service time windows.
  - Dynamic arrival of demand at arbitrary locations.

RL in Vehicle Routing Problem

- Learning based techniques provide good solutions to problems with real world constraints.
- RL provide significantly faster solutions than exact methods and meta-heuristics.

Problem Formulation

- Objective of the problem is to find the total distance J that minimises,

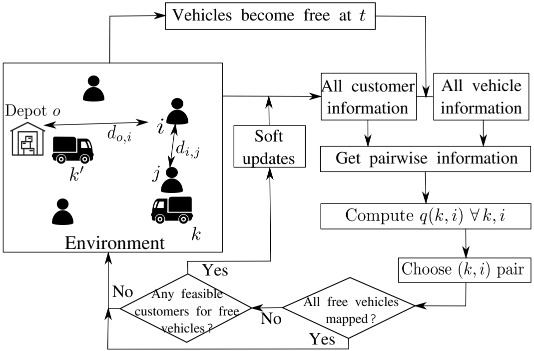
$$J = \min_{a_{i,j,k},\, f_{i,k},\, l_{i,k}} \left(\sum_{i,j,k} d_{i,j}\, a_{i,j,k} + \sum_{i,k} d_{o,i}\, f_{i,k} + \sum_{i,k} d_{o,i}\, l_{i,k} \right), \tag{1}$$

Proposed Approach

Formulate the problem as a Markov Decision Process (MDP) where each vehicle independently evaluates the value of serving each available customer at each time step after which a centralised allocation heuristic maps each vehicles to customers.

Vehicle-Customer mapping

tcs research



- Workflow of decision-making at each time step t.
- The decision loop terminates either when all free vehicles are allocated a customer, or when all feasible customers are exhausted.

Dataset Description

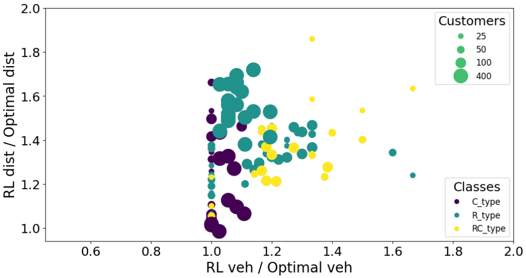
- Train Data
  - ★ 20 randomly generated dataset ($|\mathcal{C}| = 20$, $|\mathcal{V}| = 4$).
  - ★ Customer locations (x_i, y_i) are uniformly random distribution.
  - ★ The demand of each customer is drawn from an exponential distribution with $\beta = 0.1$.
  - ★ The time window width $(T_{i,max} - T_{i,min})$ is Gaussian with $(\mu = 35, \sigma = 5)$.
- Test Data
  - ★ Solomon benchmark dataset.
  - ★ Gehring and Homberger's extended CVRP-TW benchmark.
  - ★ Dynamic arrival of customers: we use a simulated dataset that is a modified version of Solomon CVRP-TW.
  - ★ An instance with X% dynamicity implies that X% of customers are not revealed at the start, and appear during execution.

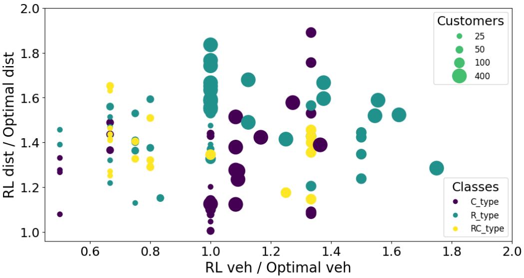
Baseline Algorithms

- Mixed integer linear program formulation (MILP)
  - Due to excessive computation times, we only report the results of the MILP formulation for the 25-customer case.
- Meta-heuristic algorithm based on genetic algorithms (GA).
- Best-known results from literature.[1]

[1] https://www.sintef.no/projectweb/top/vrptw/

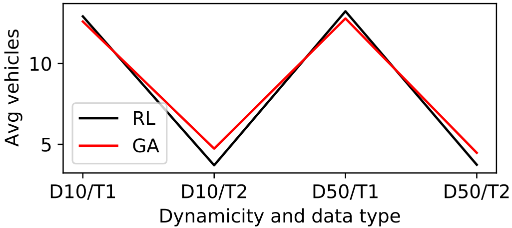
Results on static test data (CVRP-TW)

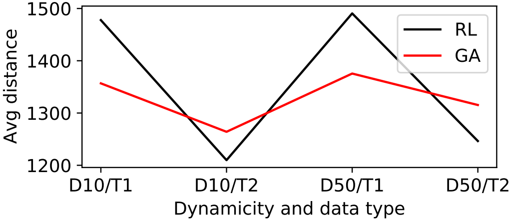




- Scatter plots of relative distance and vehicle counts for RL in comparison with best-known solutions, for type-1 (left) and type-2 (right) data.
- RL performs close to the best-known solutions in terms of number of vehicles, and is approximately 35% above optimal in terms of distance travelled.

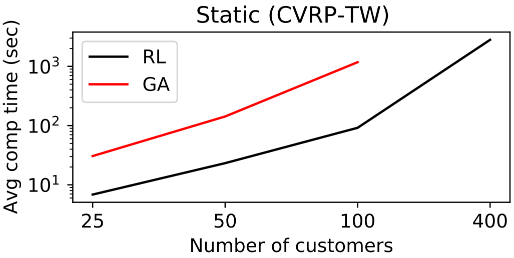
Results on dynamic test data (CVRP-TWDR)

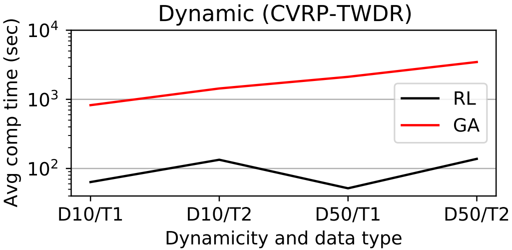




- RL outperforms GA in Type-2 data both in terms of vehicle counts and distance travelled.

Comparison of computation times





- RL has a significant advantage over GA in terms of computation times for both static and dynamic version.
- RL computation times are nearly constant for both levels of dynamicity and types for dynamic case
- RL computation time for dynamic is approximately equal to static whereas GA computation time increases with dynamicity.

Conclusion

tcs research

- The proposed approach was inherently suited for both dynamic and static vehicle routing with capacity constraints.
- Capable of producing competitive results with completely out-of-distribution test data demonstrate the scalability and flexibility of the proposed approach.
- We further wish to explore the options of parallelisability, and flexibility to handle even more constraints and stochasticity.

Thank You..