



Deutsche Post Lehrstuhl
für Optimierung von
Distributionsnetzwerken

RWTHAACHEN
UNIVERSITY

Master Thesis

Reinforcement Learning a Stochastic Dynamic Vehicle Routing Problem with Incomplete Information

Murwan Siddig, Ph.D.

Kackertstrasse 7
52072 Aachen
GERMANY
Telefon: +49 241 80-96187

siddig@dpo.rwth-aachen.de

01.02.2022

Our chair

Our chair is devoted to the development and implementation of decision support systems for solving and analyzing planning problems in logistics and production, especially transportation, network design, location planning, warehouse management and workforce scheduling problems. The methodological focus lies on mathematical modeling, exact and heuristics optimization, and machine learning techniques. Our intensive collaboration with Deutsche Post DHL and other industry partners provides a strong application focus for many of our projects.

Your task

Vehicle Routing Problems (VRPs) concern the decision of determining a set of routes for a fleet of vehicles that travel from (possibly more than) one central depot to provide service to (or collect items from) a set of customers, often in an urban environment. VRPs have a colorful range of applications in logistics and mobility services, such as transporting goods or passengers, or conducting services at customer homes. In many cases of VRPs, the decision maker (DM) is uncertain about the problem data. In such situations, the DM learns about such information sequentially, over multiple periods, in a dynamic manner. For example, additional customer requests might arrive during the routing mission, the travel time might increase due to road congestions, and the size of the available fleet might change due to accidents and malfunction of vehicles. In such situations, the dynamic nature of how the stochastic information unfolds gives rise to a class of VRPs known as Stochastic Dynamic Vehicle Routing Problem (SDVRP).

The goal of this thesis is to study an SDVRP in which a single vehicle is available to serve customer requests in a given time horizon defined by the vehicle driver's shift. In this problem, some customer requests are known in advance, but others are received during the day. When a new request occurs, the DM needs to decide whether the new request is accepted or rejected. If the new request is accepted, the DM has to decide on how to update the route to include the location of the newly accepted request. Otherwise, the rejected requests are lost forever. The goal of the DM is to maximize the number of fulfilled requests.

A common approach to proceed with solving an SDVRP is to (1) assume that the underlying stochastic process can be analyzed and transferred into a predictive model (e.g., nominal probability distribution), (2) use the predictive model to anticipate future scenarios (e.g., the likelihood of a customer request occurring at a given location), and (3) incorporate the predicted scenarios into the framework of optimizing the routing decision. Nevertheless, constructing the predictive model requires sufficient past observations (e.g., previous customer requests), which might not always be available. This can be circumvented by leveraging a reinforcement learning framework that allows the DM to use an adaptive prediction model. In this case, initially, the DM is unsure how the predictive model should reflect the uncertainty. However, as time progresses, the DM observes new realizations of the stochastic process (e.g., new demand requests) and updates their belief about how the predictive model should anticipate the uncertainty. For instance, this could be a situation in which the DM does not know in advance the demand distribution at different locations. However, as more demand requests arrive, the DM revises the predictive model to reflect the different levels

of demand requests at different locations.

Your profile

- Reliable, independent and motivated way of working
- Conscientious and structured approach to work
- Good knowledge of English and programming skills

If you are interested, please feel free to write an email with your CV and your transcript of records to Murwan Siddig (siddig@dpo.rwth-aachen.de).