Professorship of Operations and Supply Chain Management
TUM School of Management
Technical University of Munich

Deep reinforcement learning for last-mile crowdsourced delivery

To cope with increased e-commerce sales volumes and customer expectations with regards to same-day or same-hour delivery, a number of businesses rely on crowdsourced couriers to perform last-mile logistics. Crowdsourced couriers are independent contractors that work when and where they want. The logistics service provider (LSP), e.g., an on-demand service company like Postmates (US) or Gorillas (Germany), can therefore not rely safely on their availability. Consequently, they hire a small pool of in-house couriers (equivalent to permanent employees) to support operations.

Crowdsourced drivers prefer platforms that offer competitive compensation. Therefore, LSPs need to, in addition to operating cost-efficiently, ensure high utilization rates of their crowdsourced couriers. Due to the stochasticity of the problem in the demand and the supply dimension and the complexity of the objective function, finding a routing and matching policy of delivery requests to crowdsourced and in-house couriers is hard. To address this issue, we employ deep reinforcement learning (DRL) to approximate the values of matching and routing decisions, from which we can derive near-optimal matching and routing policies.

Aims and scope of the thesis
It is subject of this thesis to model a dynamic pick-up and delivery problem in a many-to-many network in which the LSP operates a mixed fleet of crowdsourced and in-house couriers. Crowdsourced couriers' availability as well as demand are uncertain. Costs consist of variable costs from in-house and crowdsourced courier routing and penalties on not served requests. One objective is the minimization of operating costs for the LSP. A second objective is the maximization of crowdsourced couriers' compensation. We can model this as a MARKOV Decision Process (MDP), where a state describes, for example, the location of each courier and whether they currently serve a request or not. The LSP's decision space can, for example, consist of decisions on which request to match to which courier or which route to take. To summarize, the work consists of following research tasks:

- Review literature on crowdsourced delivery systems and state of the art DRL approaches
- Formally define the problem as MDP – the objective being minimization of operational costs and maximization of crowdsourced courier compensation
- Implement an efficient algorithm able to solve real-sized instances based on state-of-the-art deep reinforcement learning approaches (e.g., Deep Q-Networks, actor-critic) to find optimal matching and routing policies
- Study performance based on convergence analysis
- Design and conduct experiments
- Compare benefits of such an approach with literature

Requirements
This thesis targets students of the TUM-BWL (with a major in Supply Chain Management), Informatics, Engineering or similar study programs. Knowledge of mathematical programming, optimization, and a general-purpose programming language (e.g. C++, Java, Python) is required. Prior participation in the chair's lecture *Introduction to Deep Reinforcement Learning* is recommended. The thesis should be written in English.

Related Research
- Yildiz B, Savelsbergh M (2019). Provably high-quality solutions for the meal delivery routing problem. Transportation Science, 53(5), 1372-1388.
- Arslan A, Zuidwijk R (2016). Crowdsourced delivery—a dynamic pickup and delivery problem with ad hoc drivers. Transportation Science, 53(1), 472–480.
- Sutton S, Barto A G (2015). Reinforcement Learning: An Introduction. Second edition, MIT Press.

Begin: as soon as possible

Advisor: Julius Luy

Application: See `https://www.osm.wi.tum.de/education/masters-thesis/`