Optimizing Employee Transportation : A Novel Approach Combining Reinforcement Learning and Genetic Algorithms

Mohamed Tifroute[1], Anouar Lahmdani[2]

[1] LISTI,ENSA and E2D team, ESTG - Ibn Zohr University, PO Box 1136, Agadir, Morocco.
m.tifroute@uiz.ac.ma

[2] Mathematical and computer engineering laboratory - Faculty of Applied Science, Ait Melloul, Ibn Zohr University ; Morocco.
a.lahmdani@uiz.ac.ma

Mots-clés : *Employee Transportation Problem (ETP) , Genetic Algorithm (GA) , Reinforcement Learning (RL).*

Résumé : *This research focuses on creating the best possible decision-making strategies for providing employees with a transportation service that is both fast, cost-effective, and punctual . What makes this study unique is its comprehensive approach, which combines the determination of bus stop locations, bus assignment, and the formation of bus routes within the context of a transportation service tailored for a large corporation. The primary objective of this article is to reduce the overall expenses related to employee travel, including the total travel costs and the collective rental expenses incurred by a fleet of vehicles of different sizes. This particular problem has been identified as NP-hard in existing literature [1, 2], prompting us to develop a new algorithm. Our approach combines reinforcement learning and a genetic algorithm to address this issue.*

1 Introduction

Our proposed algorithm takes advantage of data processed by an intelligent transportation system framework. This integration allows the algorithm to modify genetic algorithm parameters using reinforcement learning in real-time. This guidance increases the efficiency of the genetic algorithm, resulting in near-optimal schedules in a shorter period of time. To test our model, we conducted experiments with geospatial data including road networks, bus trip trajectories, and employee addresses. The results show that the genetic algorithm outperforms the existing schedule in terms of reducing both travel distance and time.The genetic algorithm empowered by reinforcement learning demonstrates superior performance in optimizing objective functions, and it requires fewer generations than other state-of-the-art evolutionary algorithms. Comparing the reinforcement learning-enabled genetic algorithm with the schedule generated by the initial state process, we observe significant savings in travel distance for buses and employees.

1.1 Proposed approach

The solution strategy for addressing this problem involves tackling two distinct sub-problems :

Bus Stop Selection : This phase focuses on identifying the nearest roadside locations for employees based on their respective addresses.

Bus scheduling : Within this aspect, the scheduling method divides the pickup locations into sectors, ensuring that the pickup points within the same sector are in close proximity. This approach allows you to group these points together for visits by the same bus. Consequently, this method generates initial solutions for the evolution process. These initial solutions are refined by altering the phasing, where "phasing" refers to the initial state of scanning akin to a radar mode, measured clockwise from the north line. This alteration aids in deriving improved and more viable initial solutions.

1.1.1 Reinforcement learning-enabled genetic algorithm

The genetic algorithm (GA) uses specific terminology to operate. Each gene represents a roadside point, P_j, where employees are allocated. A single gene stores information about the number of employees (H_j) at that roadside point, the source (src_j) and destination ($dest_j$) nodes for the direction of the road segment, and the bus ID (B_i) assigned to it. To generate child entities, the GA uses a 3-point crossover and 3-point mutation operator, resulting in the production of 14 entities for $2k+1$ (where $k! > 1$ produces 5). This is regulated by mutation probability (p_m) and crossover probability (p_c) parameters. These child entities are tested for feasibility to ensure they meet the problem's constraints. To improve the GA's efficiency, we have adopted an adaptive learning-based controller called RL-enabled GA. This controller dynamically determines optimal probability parameter values during operation. Unlike self-adaptive methods, which may experience premature convergence issues [3], these parameters evolve alongside the solutions. The adaptive controller is based on reinforcement learning (RL). Here, an agent driven by goals learns the best probabilities for operators through interaction with the population, making decisions based on environmental feedback. SARSA (state-action-reward-state-action) is used as an iterative, on-policy method, allowing the agent to continuously refine its learning process, transitioning between state-action pairs while receiving rewards, aiming to explore and optimize the best policy.

The value of the current policy guides the on-policy methods in their iterative improvement process, allowing them to learn and refine the agent's existing policy.

To enable Reinforcement Learning (RL) to regulate the probabilities of Genetic Algorithm (GA) operators, the parameter ranges are utilized to define a state space. Both parameter ranges undergo discretization into $'d'$ intervals, resulting in $'d2'$ potential combinations, each representing distinct GA states (p_m, p_c). Based on the current state, the specific ranges of values for both operator probabilities are determined. Random values are then assigned to these probabilities, drawn from these predefined ranges using a uniform distribution. Additionally, actions are defined to transition between states, involving adjustments to increase, decrease, or maintain the current parameter range.

For each state and its associated five actions, an estimated state-action Q-table ($Q(S, A)$) is constructed. Actions are chosen from this Q-table using an epsilon-greedy policy, where either a random action is selected from the Q-table with a certain probability, or the action with the highest Q-value is chosen.

Moreover, an eligibility trace, coupled with SARSA, acts as a temporary memory, recording event occurrences and the impact of actions on subsequent states. This trace illustrates the influence of actions from specific states on the accrued reward.

2 Conclusions and perspectives

This study focused on the Employee Transportation Problem (ETP) to improve the efficiency, reliability, and safety of employee transportation. The goal is to optimize both the overall travel costs and the rental costs of a diverse fleet of capacity-equipped vehicles.

Références

[1] Leksakul, Komgrit, Uttapol Smutkupt, Raweeroj Jintawiwat, and Suriya Phongmoo. 2017. Heuristic Approach for Solving Employee Bus Routes in a Large-Scale Industrial Factory. Advanced Engineering Informatics 32 : 176-187.

[2] Brandão, José. 2009. A Deterministic Tabu Search Algorithm for the Fleet Size and Mix Vehicle Routing Problem. European Journal of Operational Research 195 (3) : 716–728.

[3] Aleti, A., and Moser, I. (2016). A systematic literature review of adaptive parameter control methods for evolutionary algorithms. ACM Computing Surveys, 49(3), 1-35.