2026 IEEE/SICE International Symposium on System Integration (SII)
January 11-14, 2026. Cancún, México

# Intermodal Journey Planning to Transportation Hubs in a Microscopic Environment: A Multi-Objective Multi-Agent Reinforcement Learning Optimization Framework

Dominik Wittenberg[1,⋆], Nick Schade[1] and Jürgen Pannek[1]

***Abstract*— Intermodal journey planning remains a challenge in intelligent transportation systems, particularly when accounting for heterogeneous passenger preferences and the integration into smart cities. Traditional planning approaches often fail to capture dynamic traffic conditions and the passenger-centric view required for future transportation systems. This study proposes a Multi-Objective Multi-Agent Reinforcement Learning (MOMARL) framework for individual intermodal journey planning across multiple modes. Two microscopic traffic models were developed in Simulation of Urban Mobility (SUMO), creating simulation environments in which passengers plan their journeys to arrive on time at transportation hubs. One simpler model for the verification of the framework and a calibrated model reflecting the dynamics of a real city. The transportation networks were modeled as multilayered graphs. Since each passenger has different preferences and access to transport modes, their individual cost-minimal paths are formulated as a multi-objective optimization (MOO) problem. From this, the scalarized reward signals used in the MOMARL framework are derived. Simulation results show that the proposed approach enables agents to generate feasible intermodal routes in a microscopic traffic environment, demonstrating the use of MOMARL for passenger-centric coordination in multimodal transport systems. Application to the calibrated model of Ingolstadt posed challenges regarding simulation complexity, highlighting the need to expand research in methods that allow the systematic reduction of model fidelity and granularity while retaining realistic dynamics.**

## I. Introduction

Due to the imminent climate crisis and the ratification of agreements such as the Paris Climate Accord, manufacturers and public authorities are under the obligation to recognize their ecological responsibilities and prioritize zero-emission vehicles and sustainable transport systems [1]. Consequently, the transportation sector in the European Union (EU) is undergoing a significant transformation toward a more sustainable, passenger-centric, efficient and seamless transportation system [2]. While discussions often focus on the transition to zero emission vehicles, the considerable portion of urban air pollution and the blockage of urban space resulting from the continuous growth of individual transportation and the use of private vehicles has frequently been overlooked [3], [4]. In recent years, there has been a notable shift in focus in various countries, with cities aiming to transition to smart cities, announcing support for the integration of alternative and more sustainable forms of transportation [5]. Further reinforced by the European Commission's "Horizon Europe" strategy, which advocates a transition from a modal/ driver-centric to a multimodal/mobility-user-oriented and demand-driven approach [6].

However, planning intermodal journeys, especially for large numbers of travelers in complex multimodal systems, such as smart cities, with different stakeholders, remains a challenging problem [7]. A multitude of factors contribute to the planning process and influence people's travel attitudes towards intermodality. These factors include the limited capacities of individual transport modes, the increasing flexibility offered by new transportation systems and Mobility-as-a-Service (MaaS) concepts, as well as the personal satisfaction with each transportation mode and the overall trip [8], [7].

In this paper, we integrate the aforementioned factors and present a multi-agent, multi-objective reinforcement learning optimization framework, which will allow the optimization and integration of intermodal journey planning, while taking into account individual attitudes for each traveler.

To this purpose, Section II presents the state of the art in intermodal route planning, the use of reinforcement learning in this context, and multi-objective optimization. Section III outlines the proposed methodology. In Section IV, the simulation setup is explained and the methodology is applied to a traffic scenario. The results are then presented and discussed in Section V. Finally, Section VI summarizes the findings and provides an outlook on future work.

## II. State of the Art

### A. Intermodal Route Planning Methods

Numerous well-established optimization algorithms exist in the field of multimodal transport, including mathematical programming, genetic algorithms, particle swarm algorithms, and ant colony algorithms [9]. Traditional mathematical programming approaches such as Dijkstra's and A* algorithms have been adapted for multimodal contexts. For instance, Pajor [10] integrated time-dependent Dijkstra with label-constrained shortest path techniques [11]. Linear and mixed-integer programming have also been used to model journey planning with user preferences and mode compatibility [12]. Despite their speed, these methods often struggle with scalability and non-linear dynamics [13].

Funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy - EXC 2163/1 - Sustainable and Energy Efficient Aviation Project-ID 390881007.

[1]Dominik Wittenberg, Nick Schade and Jürgen Pannek are with the Institute for Intermodal Transport and Logistics Systems, TU Braunschweig, Hermann-Blenk-Straße 42, 38108 Brunswick, Germany
⋆Corresponding Author
dominik.wittenberg@tu-braunschweig.de

To address these issues, heuristic and metaheuristic algorithms offer better adaptability to complex, dynamic networks. Genetic algorithms, as a subset of metaheuristics improve route solutions over multiple generations, selecting, crossing, and mutating them to optimize the desired objectives [14], as seen in the variable-length chromosome approach of Yu and Lu [15]. Particle swarm optimization, inspired by social behavior, balances multiple objectives like cost and emissions effectively [16], [17]. Ant colony optimization, a population-based metaheuristic, mimics pheromone trails to discover optimal paths, especially in dynamic conditions [18], [19].
However, these algorithms require careful parameter adjustment to avoid falling into local optima and may perform poorly if the scale of the multi-objective optimization problem is too large [9].

### B. Reinforcement Learning

Reinforcement Learning (RL) has emerged as a promising approach for journey planning, given its ability to adapt to dynamic environments. Specifically, an RL agent can adjust its route selections in response to observed traffic patterns, delays, or fluctuations in transportation availability [9]. This process of learning optimal actions in a given environment is facilitated by the agent's reception of feedback through rewards and penalties.
The utilization of RL in the domain of freight transport has been extensively documented [20], [21], although recent studies explored its application in passenger transportation. Chu and Guo [22] used Deep RL to personalize route planning, thus improving both passenger satisfaction and operator profit. Feng et al. [23] combined ride-sourcing with public transit using RL, improving both revenue and service utilization in urban settings.
However, a single agent RL approach may be insufficient in environments where multiple agents (e.g., multiple travelers or transportation services) interact. In such cases, optimizing the behavior of a single agent may not result in an optimal solution for the entire system. Multi-Agent Reinforcement Learning (MARL) addresses this limitation by considering multiple agents operating within a shared environment. These agents act autonomously, yet their decisions impact one another. In the context of intermodal route planning, MARL is especially useful to manage coordination between different transportation modes and to balance the different needs of multiple travelers. Although often applied to freight transportation [9], MARL has also been tested in event-based passenger routing by Codeca and Cahill [7] to generate multimodal travel plans for large-scale events.
However, a significant limitation of many studies is that they focus on a single optimization target, such as travel time, distance, or emissions. To address this issue, multi-objective optimization is a necessary approach.

### C. Multi-Objective Optimization

Multi-objective optimization (MOO) is the process of finding optimal solutions when multiple, often conflicting, objectives must be considered simultaneously. Instead of a single best answer, MOO identifies tradeoffs between goals, helping to explore balanced solutions between conflicting criteria [24]. In the context of intermodal route planning, common objectives include minimizing travel time, cost, and environmental impact, while maximizing traveler comfort and convenience.
Early research primarily focused on minimizing transportation costs and time, often employing deterministic methods that combined multiple objectives into a single weighted function, yielding limited optimal solutions. For instance, Sun et al. [25] introduced a bi-objective mixed-integer linear programming model to optimize multimodal transportation routing, balancing transportation costs and time, and utilized Pareto optimality to derive a set of optimal solutions.
To handle the complexity of large-scale networks, metaheuristics like Non-dominated Sorting Genetic Algorithm II (NSGA-II) have gained popularity. A comprehensive review by Chau and Gkiotsalitis [26] examined the use of metaheuristics to optimize multimodal transportation, highlighting their efficacy in addressing the high-dimensional challenges inherent in such networks. Furthermore, a study demonstrated the application of NSGA-II in a multimodal transportation network, effectively finding multiple optimal routes that balance objectives such as cost, time and environmental impact [27]. Further works have also integrated sustainability considerations into intermodal route planning. Vale and Ribeiro [28] proposed a routing model that simultaneously optimized travel time and $CO_2$ emissions, contributing to more sustainable transportation solutions. Additionally, the integration of a decision rule-based sorting approach with NSGA-III was proposed to address customer-centered intermodal freight routing problems, emphasizing the importance of customer satisfaction alongside traditional objectives [29].

### D. Multi-Objective Reinforcement Learning

The combination of MOO and RL results in Multi-Objective Reinforcement Learning (MORL), allowing the optimization of multiple conflicting objectives within a sequential decision-making process [30]. These methods are typically categorized into single-policy approaches, which aim to learn one optimal policy based on a scalarized objective, and multi-policy approaches, which seek to approximate the Pareto front [31]. Extending this concept to multiple agents leads to Multi-Objective Multi-Agent Reinforcement Learning (MOMARL), a field that remains relatively underexplored. A brief overview and the cooperative method MOMIX are presented in [32].

### E. Research Gap

As shown, the previously discussed aspects have been extensively addressed in the existing literature. However, their concurrent integration remains limited. Only a few works, such as [7], [9], attempt to consider these aspects simultaneously in artificial environments.
Moreover, a transition towards passenger-centric intermodal planning models that incorporate personnel preferences and

behavioral diversity remains rare. This is particularly evident in the context of a centralized approach that encompasses a significant number of journeys and travelers. This paper aims to address this issue by presenting an intermodal journey planning framework that uses MOO and MARL as a MOMARL and applies this to a microscopic traffic simulation.

## III. Methodology

The methodology consists of four parts: First, we state a mathematical definition of the overall problem setting. Second, the multi-layer network model of the multimodal system is introduced. Third, a MOO problem for the individual passenger and a baseline are defined. Finally, the MOO problem is transformed and an algorithm for MOMARL is derived.

### A. Problem Setting

Given a transport hub with scheduled transportation services $\mathcal{H}$ and the respective catchment area $\mathcal{C}_{\mathcal{H}}$. During a specified time $\mathcal{T} \in \mathbb{R}^{+}$, transport units $u \in \mathcal{U}$, like vehicles, depart from the hub $\mathcal{H}$. Consequently, there exist $b_u \in \mathcal{B}$ boarding times, where all passengers of a unit $P_u \subseteq \mathcal{P}$ want to board it. Each individual passenger $p$ of a specific unit $p_{u,i} \in P_u$ needs and wants to arrive at the hub $\mathcal{H}$ during an individual time slot $T_i \subseteq \mathcal{T}, T_i = [t_0, t_e]$ to be on time. To do so, every passenger needs an individual journey plan to the hub, based on the modal opportunities $\mathcal{M}$ of the catchment area $\mathcal{C}_{\mathcal{H}}$ and the attitudes of the individual traveling $\mathcal{A}_i$. Implicitly assumed, all starting points of the passengers $\mathbf{x}_{p_i}(0)$ are part of the catchment area $\mathcal{C}_{\mathcal{H}}$, $\mathbf{x}_{p_i}(0) \in \mathcal{C}_{\mathcal{H}}$ and the catchment area has multiple modes of transportation $m \in \mathcal{M}$.

### B. Multi-Layer Network Approach

To systematically describe the transportation net, the catchment area may be modeled as network $\mathcal{N} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ is a set of vertices and $\mathcal{E}$ is a set of edges connecting the vertices. The multimodal setting will be represented by a multi-layer approach, where each mode $m \in \mathcal{M}$ operates on its own graph layer $\mathcal{N}^m = (\mathcal{V}^m, \mathcal{E}^m)$. The resulting vertical connecting vertices represent locations such as stations, parking lots, or intermodal hubs where passengers can switch between modes. The corresponding subset $\mathcal{V}^{mm'}$ defines these intermodal transfer points between mode $m$ and mode $m'$. As a result of this, the overall network can be reduced to a transport-focused network, as visualized in Figure 1. In there, each edge $e \in \mathcal{E}$ represents a travel connection between two locations within a mode $e^m$ or a transition to another mode $e^{mm'}$. Traversing within a mode from the current vertex $i$ to the next vertex $j$ through an edge is represented by $e^m_{ij}$.

### C. Multi-Objective Optimization

To describe how passengers act in the multi-layer network, the state $\mathbf{x}_{p_i}(t) \in \mathcal{X}$ at timestep $t$ consists of:

$$\mathbf{x}_{p_i}^{\top} = (v, m, t^{\star}(p_i) - t) \tag{1}$$

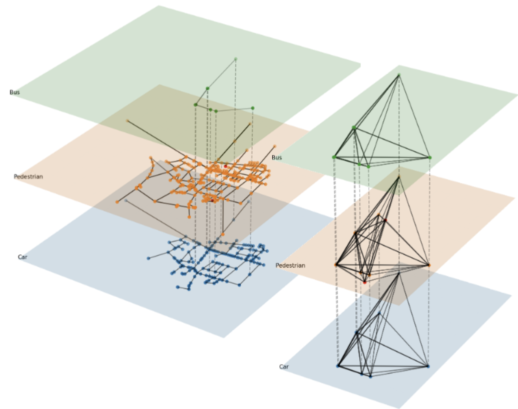


Fig. 1. Exemplary representation of the multilayer multimodal transportation graph, with car, bus and pedestrians as available modes.

where $v$ denotes the current vertex, $m$ the current mode, and $t^{\star} - t$ the remaining time until the planned arrival time $t^{\star}$ at the transport hub. To represent individual attitudes and conflicting objectives, we formalize an individual cost minimal path for each passenger $J_{p_i}$, where the weights $w_{p_i}$ reflect the specific attitudes of the passenger regarding the relevance of a certain type of cost $\mathrm{C}_k \in C$,

$$J_{p_i}(w_{p_i}, \tau_{p_i}) = min \sum_{k=1}^{n_{\mathrm{C}}} w_{p_i,k} \cdot \mathrm{C}_k(\tau_{p_i}) \tag{2}$$

$$\sum_{k=1}^{n_{\mathrm{C}}} w_k = 1, w_k \geq 0. \tag{3}$$

with $\tau_{p_i}$ the path the passenger traversed and $n_{\mathrm{C}} = |C|$ the number of different types of costs.
To provide a baseline, the total generic cost of a path $C(\tau_p)$, assuming a uniform weight distribution, is determined through the costs associated with the edges and vertices traversed as a sum of different types of costs:

$$C(\tau_p) = \sum_{k=1}^{n_{\mathrm{C}}} \mathrm{C}_k(\tau_p) \tag{4}$$

$$\mathrm{C}_k(\tau_p) = \sum_{e \in \mathcal{E}(\tau_p)} \mathrm{C}_k(e) + \sum_{v \in \mathcal{V}(\tau_p)} \mathrm{C}_k(v) \tag{5}$$

To determine the cost minimal paths for all passengers, we minimize over the sum of the individual cost minimal paths across all individuals

$$J_{\mathcal{P}}(w, \tau) = min \sum_{p_i=1}^{\mathcal{P}} J_{p_i}(w_{p_i}, \tau_{p_i}). \tag{6}$$

### D. Multi-Objective Reinforcement Learning

To apply MORL, each passenger $p_i$ is represented as an agent $A_{\{1,2,\ldots,\mathcal{P}\}}$. The available actions $a$ consist of:

$$a = \left\{ \begin{array}{c} \text{change to a reachable vertex} \\ \text{wait} \end{array} \right\} \tag{7}$$

By this, the action of changing to a reachable vertex defines the new mode and the next goal position of an agent.

Each agent aims to find a policy $\pi_p$ that maximizes the expected discounted cumulative scalar reward, where $\gamma \in (0,1]$ is the discount factor and $T$ is the time horizon:

$$J(\pi_A) = \mathbb{E}\left[\sum_{t=0}^{T} \gamma^t \cdot J(w,\tau)\right] \tag{8}$$

It is important to note that the costs $J(w,\tau)$ must be implemented in a manner that facilitates maximization rather than minimization, as indicated in the multi-objective optimization.

## IV. Case Study

To assess the applicability and feasibility of the methodology, the MARL algorithm is implemented using the open-source distributed execution framework RAY, together with the reinforcement learning library RLlib [33]. The actor-critic algorithm Proximal Policy Optimization (PPO) [34] is utilized using the implementation provided by RLlib.

The generic multi-objective formulation from the methodology is adapted for the specific use case, where passengers have to arrive within a certain time window at a transportation hub and board a common mode. Additional objectives, such as emissions or cost, may be incorporated in future work but are omitted here to reduce environment complexity and ease verification. To minimize travel time and achieve optimal arrival of passengers at a defined time at the hub, the travel time and the deviations of late and early arrival are defined. In addition to the travel time cost $\mathrm{C}_1(\tau_p)$, we define the terminal costs for each passenger path with regard to late $\mathrm{C}_2(\tau_p)$ and early $\mathrm{C}_3(\tau_p)$ arrival with their respective weights. Both reflect the absolute difference between the planned arrival time and the actual arrival time. The minimal cost for a passenger becomes:

$$J_{p_i}(w_{p_i},\tau_{p_i}) = \min\Big\{w_{p_i,1}\cdot \mathrm{C}_1(\tau_{p_i}) + w_{p_i,2}\cdot \mathrm{C}_2(\tau_{p_i}) + w_{p_i,3}\cdot \mathrm{C}_3(\tau_{p_i})\Big\}. \tag{9}$$

To ensure passengers arrive within a defined time window around the scheduled arrival time, we impose the constraints:

$$|\Delta t_{\text{arr}}| \le l_{\text{acc}} \tag{10}$$

and

$$\mathrm{C}(\tau_{p_i},\tau^*) = |\arg\min_t\{\tau_{p_i}(t) - \tau^* = 0\} - t^*| \tag{11}$$

where $l_{\text{acc}}$ defines the maximum allowable deviation and $|\Delta t_{\text{arr}}| = t_{\text{arr}} - t^\star$ denotes the deviation between the actual arrival time $t_{\text{arr}}$, the planned arrival time $t^\star$ and the planned destination $x^*$. To ensure that early and late costs are mutually exclusive, we impose

$$\mathrm{C}_3 \cdot \mathrm{C}_2 = 0. \tag{12}$$

As a result, the scalarized reward for the MARL algorithm at time $t$ becomes:

$$r^{\text{scalar}} = r_1 + r_2 + r_3 \tag{13}$$

with

$$r_1 = -w_1 \cdot \left(\alpha\sqrt{|\Delta t_{\text{arr}}|} + t_{\text{travel, best}}\right), \tag{14}$$

$$r_2 = \begin{cases} 0, & |\Delta t_{\text{arr}}| < l_{\text{acc}} \\ -w_2 \cdot (|\Delta t_{\text{arr}}| - l_{\text{acc}}), & \Delta t_{\text{arr}} > l_{\text{acc}} \end{cases} \tag{15}$$

and

$$r_3 = \begin{cases} 0, & |\Delta t_{\text{arr}}| < l_{\text{acc}} \\ -w_3 \cdot (|\Delta t_{\text{arr}}| - l_{\text{acc}}), & \Delta t_{\text{arr}} < -l_{\text{acc}}, \end{cases} \tag{16}$$

where $\alpha$ represents the scaling factor to balance arrival deviation $t_{\text{arr}}$ and shortest travel time $t_{\text{travel, best}}$.

To encourage agents to depart as late as possible, an additional stepwise reward is formulated which provides a reward to the agents for waiting before departure. Each agent uses an individual policy and value function, modeled as fully connected neural networks with hidden layers of 256, 512, 512, and 256 neurons. A linear activation function is applied at the output layer. Policies output discrete action logits from a masked observation vector encoding the agent's current location, transport mode, and remaining time until planned arrival. The action space is dynamically masked to exclude infeasible options based on the current state and network constraints.

The PPO algorithm uses a discount factor of $1.0$ and a clipping parameter of $0.2$. Value function clipping is disabled by setting the threshold to $100,000$. A scheduled learning rate starts at $10^{-4}$, decreasing to $1.0\times10^{-5}$ after $30,000$ iterations, and to $10^{-6}$ after $50,000$ iterations. Training proceeds for 120 iterations with a sample batch size of $2,000$. Each training iteration collects at least 100 environment steps.

The simulation environment is implemented in SUMO, accessed via TraCI for real-time interaction. It supports time-dependent, intermodal routing across pedestrian, bus, and car networks, and applies dynamic feasibility constraints through a labeled node graph. Agents operate under partial observability with local observations and use stored observation histories.

Fifty agents are initialized with a random pedestrian-accessible origin edge, a fixed common destination edge, random car possession ($30\%$ probability), and a set planned arrival time with an upper and lower bound. Cars are assigned random parking edges and positions.

Two simulation environments are used: a simplified setup based on the village of Schapen, Germany, consisting of a specific bus schedule, parking spots at every edge, and synthetically generated traffic demand, and a calibrated microscopic model of Ingolstadt, Germany [35], offering a high resolution urban road infrastructure with realistic traffic signals, public transportation schedules, and demand profiles. For the environment, data from June 19, 2023, starting at 08:00 AM are used.

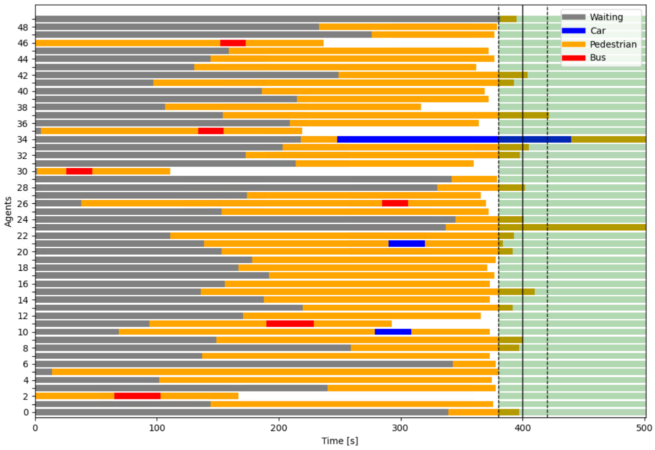


Fig. 2. Intermodal journeys of the agents over time, including individual mode choice and alignment with the scheduled arrival window.

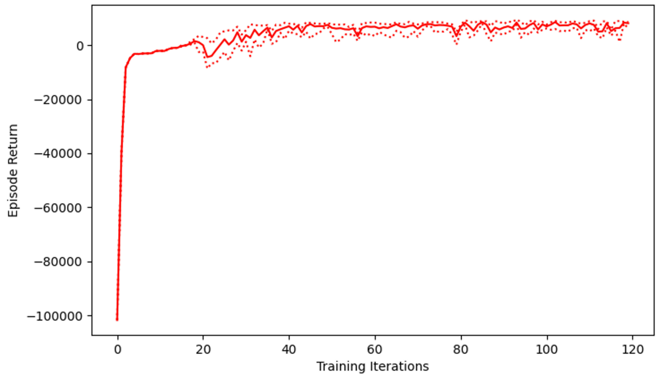


Fig. 3. Evolution of the total episode return during training. The solid red line shows the mean episode return per iteration, while the red dotted lines represent the spread of individual agent returns.

## V. Results and Discussion

The performance of the MOMARL framework in planning intermodal passenger journeys, with the objective of minimizing travel time and arriving within a specified time window at an intermodal hub, was evaluated in both simplified and calibrated microscopic traffic simulations. The simplified simulation evaluation focused on the modal split, the travel times of the agents, and the performance of the MARL algorithm. For the calibrated simulation, the focus was on assessing the feasibility of applying the framework in a realistic environment. Figure 2 presents the travel behavior of fifty agents during the simulation period. Each horizontal bar represents the journey of an individual agent, segmented into pedestrian (orange), bus (red), car (blue), and waiting phases (gray). Two agents failed to complete their journeys within the time window, which may be attributed to suboptimal decisions in rarely encountered situations and the heuristic nature of the RL algorithm. The majority of agents arrived before the lower bound of the time window, as the reward function incentivizes early arrival more strongly than late arrival. Combined with the stochastic nature of the environment, this leads agents to arrive earlier rather than risk being too late. The use of cars indicates that car segments were only chosen when available and necessary due to infeasible pedestrian and bus connections and the proximity of the starting position of the car to the agent's origin and destination. Due to the limited size of the environment shown in Figure 1 and the random distribution of agents, remaining pedestrian for the journey tends to be the best option for most starting positions. Figure 3 shows the evolution of the total episode return during training. The return increased sharply within the first iterations. After this initial phase, the learning curve continued to increase slightly and then plateaued, exhibiting only minor oscillations without significant regressions. The rapid increase in return during early training indicates that the agents quickly learned effective policies for planning. The subsequent stabilization of the return values suggests convergence toward a locally optimal policy under the given task and environment conditions. These results demonstrate that the MARL training process successfully improved agent behavior to minimize travel time and meet scheduled arrival windows. Despite successful initialization and agent spawning in the calibrated simulation, learning progress remained limited. Episode returns increased slightly and decreasing entropy values suggested a shift from exploration to early exploitation. However, the persistently high value function loss and near-zero explained variance indicated that the critic failed to learn meaningful predictions. The average episode length decreased, but this was mainly due to early terminations rather than improved planning behavior. Additional challenges originated from the simulation environment. The runtime of SUMO per episode averaged 11.4 minutes, significantly slowing down the training iterations. Its limitation to single-core execution further constrained learning efficiency. Encountered scenario specific issues included missing pedestrian edge links, absent parking area definitions, and the manual initialization of more than 7,000 vehicles at simulation start to reflect real-world traffic at 08:00 AM, to mitigate a warm up time for SUMO, all of which added additional bottlenecks and instability. Overall, the extent of environmental uncertainty and scenario induced variability makes it difficult to reliably separate genuine learning progress from artifacts caused by simulation dynamics.

## VI. Conclusions

In this research, we considered the problem of intermodal journey planning in a multimodal transportation system using MARL in two microscopic traffic simulations, a simplified one of the village of Schapen for verification purposes, and a calibrated one of the city of Ingolstadt to determine the viability to integrate this framework into a realistic environment. Each environment was represented as a multilayered graph, with car, bus, and pedestrian modes on separate layers. The dynamics of the environment were modeled using SUMO. The simulation results showed that the agents were able to compute feasible intermodal paths that minimize travel time and deviation from the planned arrival time. Allowing us to verify the ability of the MOMARL single-policy algorithm in the Schapen environment to perform intermodal journey planning, where each passenger had specific preferences and attributes and therefore different optimal journeys.

The application of this framework to the Ingolstadt environment was possible, but resulted in slow convergence and inconsistent agent behavior due to high simulation complexity. In the next steps, additional passenger preferences, such as comfort, emissions, and price, will be included in the cost function. Concurrently, the benefits of implementing a multi-policy algorithm will be evaluated. To apply the framework to realistic environments, further work will be required to optimize the MARL algorithm and systematically reduce environment complexity, ensuring that intermodal route planning continues to reflect calibrated dynamics while minimizing computational demand through model abstraction. A viable approach may be to pretrain the MARL algorithm using curriculum learning on a static environment graph, where edge weights represent average travel times derived from SUMO simulations, before fine-tuning in the calibrated environment. Furthermore, the Schapen environment will be extended to include an airport catchment area with additional transportation modes, allowing the evaluation of intermodal travel behavior in a multimodal transportation system where the airport serves as an intermodal hub. This will enable the evaluation and optimization of passenger-centric journey planning, where each passenger follows an individually optimal route under varying preferences and constraints, supporting the EU vision of a seamless door-to-door passenger journey.

## REFERENCES


[1] J. Reul, T. Grube, and D. Stolten, “Urban transportation at an inflection point: An analysis of potential influencing factors,” *Transportation Research Part D: Transport and Environment*, vol. 92, 2021.

[2] European Commission, “Data driven approach for a Seamless Efficient European Travelling in 2050 (DATASET2050),” 2014.

[3] R. Arnott and E. Inci, “An integrated model of downtown parking and traffic congestion,” *Journal of Urban Economics*, vol. 60, 2006.

[4] M. Barth and K. Boriboonsomsin, “Real-World Carbon Dioxide Impacts of Traffic Congestion,” *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2058, pp. 163–171, 2008.

[5] D. Müller-Eie and I. Kosmidis, “Sustainable mobility in smart cities: a document study of mobility initiatives of mid-sized nordic smart cities,” *European Transport Research Review*, vol. 15, p. 36, 2023.

[6] European Commission, *Horizon Europe strategic plan 2025-2027*. LU: Publications Office, 2024.

[7] L. Codeca and V. Cahill, “Using Deep Reinforcement Learning to Coordinate Multi-Modal Journey Planning with Limited Transportation Capacity,” *SUMO Conference Proceedings*, vol. 2, pp. 13–32, 2022.

[8] T. Meyer-Hollatz, M. Kaiser, R. Keller, and M. Schober, “Understanding information needs for seamless intermodal transportation: Evidence from Germany,” *Transportation Research Part D: Transport and Environment*, vol. 130, 2024.

[9] T. Zhang, J. Cheng, and Y. Zou, “Multimodal transportation routing optimization based on multi-objective Q-learning under time uncertainty,” *Complex & Intelligent Systems*, vol. 10, 2024.

[10] T. Pajor, “Multi-Modal Route Planning,” Diplomarbeit, Universität Karlsruhe, Karlsruhe, 2009.

[11] C. Barrett, R. Jacob, and M. Marathe, “Formal-Language-Constrained Path Problems,” *SIAM Journal on Computing*, vol. 30, 2000.

[12] G. Aifadopoulou, A. Ziliaskopoulos, and E. Chrisohoou, “Multi-objective Optimum Path Algorithm for Passenger Pretrip Planning in Multimodal Transportation Networks,” *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2032, pp. 26–34, 2007.

[13] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to algorithms*. MIT Press, 2022.

[14] S. Katoch, S. S. Chauhan, and V. Kumar, “A review on genetic algorithm: past, present, and future,” *Multimedia Tools and Applications*, vol. 80, 2021.

[15] H. Yu and F. Lu, “A multi-modal route planning approach with an improved genetic algorithm,” *Adv. Geo-Spatial Inform. Sci*, vol. 38, 2012.

[16] A. G. Gad, “Particle Swarm Optimization Algorithm and Its Applications: A Systematic Review,” *Archives of Computational Methods in Engineering*, vol. 29, 2022.

[17] X. Zhang, H. Liu, and L. Tu, “A modified particle swarm optimization for multimodal multi-objective optimization,” *Engineering Applications of Artificial Intelligence*, vol. 95, 2020.

[18] P. He, G. Jiang, S.-K. Lam, and Y. Sun, “ML-MMAS: Self-learning ant colony optimization for multi-criteria journey planning,” *Information Sciences*, vol. 609, 2022.

[19] C. Blum, “Ant colony optimization: Introduction and recent trends,” *Physics of Life Reviews*, vol. 2, pp. 353–373, 2005.

[20] H. Aghazadeh and X. Wang, “Reinforcement Learning for Intermodal Transportation Planning with Time Windows and Limited Cargo Capacity,” in *Proceedings of the 16th ACM SIGSPATIAL International Workshop on Computational Transportation Science*, 2023, pp. 28–31.

[21] A. Farahani, L. Genga, and R. Dijkman, “Online Multimodal Transportation Planning using Deep Reinforcement Learning,” in *IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 2021.

[22] K.-F. Chu and W. Guo, “Deep reinforcement learning of passenger behavior in multimodal journey planning with proportional fairness,” *Neural Computing and Applications*, vol. 35, 2023.

[23] S. Feng, P. Duan, J. Ke, and H. Yang, “Coordinating ride-sourcing and public transport services with a reinforcement learning approach,” *Transportation Research Part C: Emerging Technologies*, vol. 138, 2022.

[24] M. J. Kochenderfer and T. A. Wheeler, *Algorithms for Optimization*. The MIT Press, 2019.

[25] Y. Sun and M. Lang, “Bi-objective optimization for multi-modal transportation routing planning problem based on Pareto optimality,” *Journal of Industrial Engineering and Management*, vol. 8, 2015.

[26] M. L. Y. Chau and K. Gkiotsalitis, “A systematic literature review on the use of metaheuristics for the optimisation of multimodal transportation,” *Evol. Intell.*, vol. 18, 2025.

[27] H. Faroqi, “Multiobjective route finding in a multimode transportation network by NSGA-II,” *J. Eng. Appl. Sci.*, vol. 71, 2024.

[28] C. Vale and I. M. Ribeiro, “Intermodal routing model for sustainable transport through multi-objective optimization,” in *Intelligent Transport Systems, From Research and Development to the Market Uptake*, J. C. Ferreira, A. L. Martins, and V. Monteiro, Eds. Springer International Publishing, 2019, pp. 144–154.

[29] C. Shao, H. Wang, and M. Yu, “Multi-objective optimization of customer-centered intermodal freight routing problem based on the combination of DRSA and NSGA-III,” *Sustainability*, vol. 14, 2022.

[30] C. Liu, X. Xu, and D. Hu, “Multiobjective reinforcement learning: A comprehensive overview,” *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 45, pp. 385–398, 2015.

[31] P. Vamplew, R. Dazeley, A. Berry, R. Issabekov, and E. Dekker, “Empirical evaluation methods for multiobjective reinforcement learning algorithms,” *Mach. Learn.*, vol. 84, p. 51–80, 2011.

[32] T. Hu, B. Luo, C. Yang, and T. Huang, “Mo-mix: Multi-objective multi-agent cooperative decision-making with deep reinforcement learning,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 45, 2023.

[33] P. Moritz, R. Nishihara, S. Wang, A. Tumanov, R. Liaw, E. Liang, M. Elibol, Z. Yang, W. Paul, M. I. Jordan, and I. Stoica, “Ray: a distributed framework for emerging ai applications,” in *Proceedings of the 13th USENIX Conference on Operating Systems Design and Implementation*. USENIX Association, 2018, p. 561–577.

[34] S. V. Albrecht, F. Christianos, and L. Schäfer, *Multi-Agent Reinforcement Learning: Foundations and Modern Approaches*. MIT Press, 2024.

[35] M. Harth, M. Langer, and K. Bogenberger, “Automated calibration of traffic demand and traffic lights in sumo using real-world observations,” *SUMO Conference Proceedings*, vol. 2, p. 133–148, 2022.