Developing a Personalised End-to-End Optimisation Algorithm for Smart Parking Systems

J. Qiu[a*], Y. Feng[a], S. Dale[b], M. Quddus[a] and W. Ochieng[a]

[a] Department of Civil and Environmental Engineering, Imperial College London, London, United Kingdom

[b] Nottingham City Council, Nottingham, United Kingdom

jingshuo.qiu19@imperial.ac.uk, y.feng19@imperial.ac.uk, simon.dale@nottinghamcity.gov.uk, m.quddus@imperial.ac.uk, w.ochieng@imperial.ac.uk

[*] Corresponding author

Extended abstract submitted for presentation at the Conference in Emerging Technologies in Transportation Systems (TRC-30)
September 02-03, 2024, Crete, Greece

April 16, 2024

Keywords: Reinforcement Learning, Personalised Parking Allocation, Multi-Agent Simulation, Intermodal Optimisation

1 INTRODUCTION

The rapid economic growth and technological advancement have increased reliance on cars, evident in the significant rise in car ownership around the world. However, despite the critical role of vehicles in modern life, parking-related challenges persist, leading to negative externalities such as time, fuel consumption and environmental impact (INRIX, 2017). Emerging solutions such as Smart Parking Systems (SPS) offer advanced technologies to improve parking efficiency, providing services like real-time parking availability information, online reservation, and mobile payment (Rizvi *et al.*, 2019). The efficiency of SPS has been proved in reducing fuel consumption, driving costs and parking search time, thereby potentially alleviating traffic congestion and emissions (Berenger Vianna *et al.*, 2004).

However, current parking solutions often have limited search capabilities, typically offering parking spaces with direct walking access to destinations, thereby restricting parking options in close proximity to intended destinations (Postma, 2022). Secondly, existing parking allocation algorithms lack the flexibility to accommodate personalised parking preferences for individuals. Instead, they assume uniform preferences among drivers, which may not align with real-world practices where individual parking preferences vary significantly (Nakazato *et al.*, 2022). To address these challenges, this study proposes a personalised end-to-end parking allocation algorithm. Firstly, the algorithm uses a Multi-Agent Reinforcement Learning (MARL) paradigm to simulate the end-to-end parking process with multiple traffic modes involved, which enables drivers to explore parking options beyond the constraints of walking access to their destinations. Secondly, Grey Relational Analysis (GRA) is applied to personalise parking profiles for individuals, especially under the assumption that the relative importance of selected preferences is unknown (Deng, 1982). Lastly, a comprehensive analysis is conducted to evaluate the effectiveness of GAR in personalisation compared to the Technique for Order Preference by Similarity to Ideal Solution (TOPSIS) algorithm within the MARL paradigm.

2 METHODOLOGY

2.1 Reinforcement Learning Approach

A MARL environment is usually characterised by Markov Games (MGs), also known as stochastic games. A MG is generally defined by a tuple $(N, S, A^i(s), P, R^i, \gamma | i \in N, s \in S, a \in A^i(s))$ where $N = 1, 2, \ldots, n$ denotes the number of agents; S is the state space observed from the environment. It includes number of agents n^t on the network at time t and the occupied number o_t^j of parking zones j and the confidence level c_t^j of securing an available parking space where $c_t^j = \frac{1-(n_t^j)}{N^j} \times 100$ and N^j is the capacity of the parking zone j. $A^i(s)$ is the action space of the i-th agent and $A := A^1 \times A^2 \times \cdots \times A^n$ is the joint action space for all agents. In this paper, $A^i(s)$ is defined as $A^i(s) = [a_1,\ a_2,\ \ldots,\ a_j, a_w,\ a_{pt}]$, where a_j is the action for driving to car park j, a_w denotes the walking action and a_{pt} is the action for taking public transport. If agents are driving from the origin to the car park, $A^i(s) = [a_1,\ a_2,\ \ldots,\ a_j]$; otherwise, $A^i(s) = [a_w,\ a_{pt}]$. $P : S \times A \to \delta(S)$ is the transition probability to the next state $s^{'} \in S$ given a current state $s \in S$ and a joint state $a \in A$; $R^i : S \times A \times S \to R$ is the reward function representing the instantaneous reward received by the i-th agent when transitioning from a state-action pair (s, a) to $s^{'}$, denoted by $R_t^i : \sum_{k=t}^{T-1} \gamma^{k-1} r_{k+1}^i$. $\gamma \in [0, 1]$ is the discount factor.

In a Multi-Objective MARL problem, rewards are represented as vectors rather than scalar values: $R(s, a) = [R^1(s, a), R^2(s, a), \ldots, R^i(s, a), \ldots, R^n(s, a)]$. Each $R^i(s, a)$ component represents a multi-objective reward for agent i: $R^i(s, a) = \sum_{k=1} w_k^i f_k^i(x)$, where $f_k^i(x)$ denotes the reward calculated for different parking preferences, weighted by w_k^i to determine relative importance. The rewards are transformed into monetary values within a linear programming model. Specifically, time reward is the difference between observed and free-flow travel times, multiplied by the Value of Time. Cost reward includes ticket fees, hourly parking pricing, and estimated parking time. Fuel consumption reward is determined by the fuel price multiplied by total consumption while carbon footprint reward is the total carbon emission multiplied by the value of carbon. To evaluate the learning performance of different models under MARL, Deep Q-Network (DQN) and Advantage Actor Critic (A2C) algorithms are employed in Section 3.

2.2 Parking Preference Personalisation

This paper employs GRA to personalise and determine the relative importance of parking preferences, assuming only the chosen parking preference is known a priori. GRA, initially proposed by Deng (1982), deals with real-world situations only containing partial information, also known as grey systems. In a grey system, decision-making becomes challenging when multi-attributes are considered. To represent this problem using the GRA framework, the parking preferences of agent i can be denoted by relative factor sequences, as expressed below:

$$X_q^i = (x_q^i(1), x_q^i(2), \ldots, x_q^i(p), \ldots, x_q^i(n)) \tag{1}$$

where $q = 0, 1, 2, \ldots, m$ represents the number of preferences or objectives with $q = 0$ denoting the total distance, $p = 1, 2, \ldots, n$ represents the number of parking zones and the element $x_q^i(p)$ represents the value computed for the parking zone p under the q-th preference.

The grey relational coefficient $\gamma(x_0^i(p), x_p^i(q))$ of X_q^i and X_0^i at point p can be defined as (Deng, 1982):

$$\begin{aligned}
\gamma(x_0^i(p), x_q^i(p)) &= \frac{\underset{q}{min}\underset{p}{min} \left| x_0^i(p) - x_q^i(p) \right| + \xi \underset{q}{max}\underset{p}{max} \left| x_0^i(p) - x_q^i(p) \right|}{\left| x_0^i(p) - x_q^i(p) \right| + \xi \underset{q}{max}\underset{p}{max} \left| x_0^i(p) - x_q^i(p) \right|} \\
\gamma(X_0^i, X_q^i) &= \frac{1}{n} \sum_{p=1}^{n} \gamma(x_0^i(p), x_q^i(p))
\end{aligned} \tag{2}$$

where ξ is the distinguishing coefficient, typically set to 0.5 and $\gamma(X_0^i, X_q^i)$ is the grey relational degree that can be normalised to serve as the weights in the reward function.

3 RESULTS

3.1 Experiment Setting

The study focuses on the Meadows district of Nottingham, England, known for its diverse mixed-use spaces. In the study area, a total of 562 parking spaces are distributed across 7 parking zones, including both on-street and off-street areas. To effectively manage computational intensity, the total number of parking spaces was scaled down to 70 parking spaces for 20 agents involved in the environment. The traffic flow data was extracted from smart cameras provided by the Vivacity Lab while the parking data was analysed based on the transactions of parking metres sourced from RingGo. Both traffic flow data and parking data for this study area were pre-processed and integrated into the simulation environment using Simulation of Urban MObility (SUMO). Two analyses are conducted: one comparing the learning performance of DQN and A2C algorithms, and another comparing GRA with TOPSIS for personalising parking preferences. Five distinct scenarios, based on the number of chosen preferences, are defined for comparison. Scenario 1 evenly distributes agents selecting one to four preferences, while Scenarios 2, 3 and 4 feature agents predominantly choosing one, two, three, or four preferences, respectively. In addition, three metrics are defined to evaluate the practical performance of the proposed algorithms, including average total travel time, average total travel distance, and average walking distance.

3.2 Findings

To analyse the convergence of learning algorithms, the median results of the last 100 episodes are summarised in Table 1. The total reward is not simply the summation of all individual rewards, as the weights assigned to each preference for each agent vary. In this table, the trend of individual rewards does not exhibit a consistent increase or decrease across scenarios, reflecting the complexity of the scenario configurations based on the dominance of preference numbers and the origin-destination pairs of each agent. However, it can be concluded that A2C-GRA demonstrates superiority over DQN-GRA in the aspects of total reward, as evidenced by the increased rewards of A2C-GRA. On average, the total reward of A2C-GRA is increased by 12.58% relative to that of DQN-GRA. Therefore, in the comparison of preference personalisation algorithms, GRA and TOPSIS were trained using the A2C paradigm for optimal results.

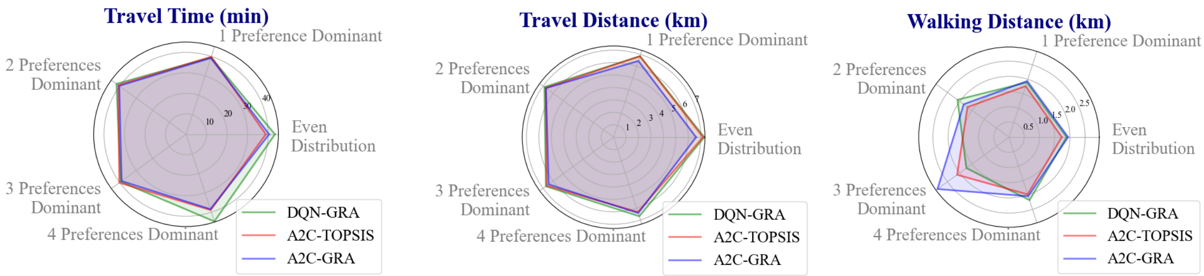
In addition, the A2C-GRA algorithm demonstrates significant differences in learning performance compared to the A2C-TOPSIS algorithm, with an average improvement of 27.24% in total rewards. Furthermore, the individual rewards for preferences of the A2C-TOPSIS algorithm also fail to surpass those of the A2C-GRA algorithm, except for some rewards for carbon footprint. Figure 1 summarises the metrics evaluation of the DQN-GRA, A2C-GRA, and A2C-TOPSIS algorithms. The superiority of the A2C-GRA algorithm is observed in minimising travel time and travel distance compared with DQN-GRA, and A2C-TOPSIS. Conversely, A2C-TOPSIS outperforms others in minimising walking distance across all scenarios, except in scenarios dominated by three preferences. Although the A2C-GRA algorithm does not minimise the walking distance to the same content as other algorithms, it can be inferred that the driving distance it generates is smaller, indicating reduced fuel consumption and carbon footprint.

4 DISCUSSION

The analysis compared the performance of A2C-GRA and DQN-GRA, finding A2C-GRA to be notably more efficient, achieving a total reward 12.58% higher on average and demonstrating improvements in travel time, travel distance, and walking distance by up to 14.05% in certain

Table 1 – *Learning Performance of DQN-GRA, A2C-GRA, and A2C-TOPSIS*

Scenarios	Algorithms	Reward	Time	Cost	Fuel	Carbon
Scenario 1	DQN-GRA	-1.200	-10.146	-2.177	-0.433	-0.282
(Even Distribution)	A2C-GRA	-1.047	-8.302	-2.251	-0.422	-0.172
	A2C-TOPSIS	-1.364	-7.488	-2.490	-0.428	-0.180
Scenario 2	DQN-GRA	-0.872	-9.173	-3.124	-0.644	-0.284
(1 Preference Dominant)	A2C-GRA	-0.868	-7.591	-2.325	-0.490	-0.222
	A2C-TOPSIS	-1.091	-7.751	-2.205	-0.502	-0.141
Scenario 3	DQN-GRA	-1.338	-8.497	-2.039	-0.527	-0.188
(2 Preferences Dominant)	A2C-GRA	-1.143	-8.432	-1.758	-0.432	-0.171
	A2C-TOPSIS	-1.420	-11.405	-2.608	-0.685	-0.214
Scenario 4	DQN-GRA	-1.218	-10.189	-2.225	-0.675	-0.269
(3 Preferences Dominant)	A2C-GRA	-1.195	-8.462	-1.935	-0.434	-0.161
	A2C-TOPSIS	-1.894	-8.480	-2.260	-0.344	-0.126
Scenario 5	DQN-GRA	-1.766	-8.931	-1.543	-0.637	-0.233
(4 Preferences Dominant)	A2C-GRA	-1.179	-8.951	-1.891	-0.423	-0.133
	A2C-TOPSIS	-1.845	-10.604	-2.346	-0.582	-0.254



(a) *Travel Time (min)* (b) *Travel Distance (km)* (c) *Walking Distance (km)*

Figure 1 – *Metrics Evaluation for DQN-GRA, A2C-GRA, and A2C-TOPSIS*

scenarios. Although both GRA and TOPSIS algorithms were trained using A2C, A2C-GRA consistently outperformed A2C-TOPSIS across all scenarios, with a notable 27.24% increase on average. The evaluation of metrics showed that A2C-GRA excels in minimising travel time and travel distance compared with A2C-TOPSIS, while A2C-TOPSIS outperforms A2C-GRA in walking distance. The study suggests potential future research involving real-world preference data to further assess GRA and TOPSIS effectiveness and expand the applicability and accuracy by considering additional preferences.

References

Berenger Vianna, Marcello Marinho, Portugal, Licínio da Silva, & Balassiano, Ronaldo. 2004. Intelligent transportation systems and parking management: implementation potential in a Brazilian city. *Cities*, **21**(2), 137–148. Number: 2.

Deng, Ju-Long. 1982. Control problems of grey systems. *Systems & Control Letters*, **1**(5), 288–294.

INRIX. 2017. *Searching for Parking Costs the UK £23.3 Billion a Year.*

Nakazato, Takuya, Fujimaki, Yuto, & Namerikawa, Toru. 2022. Parking Lot Allocation Using Rematching and Dynamic Parking Fee Design. *IEEE Transactions on Control of Network Systems*, 1–1. Conference Name: IEEE Transactions on Control of Network Systems rate: 4.

Postma, Alexander. 2022. *RingGo | The UK's No 1 Parking App.*

Rizvi, Syed R., Zehra, Susan, & Olariu, Stephan. 2019. ASPIRE: An Agent-Oriented Smart Parking Recommendation System for Smart Cities. *IEEE Intelligent Transportation Systems Magazine*, **11**(4), 48–61. Number: 4 Conference Name: IEEE Intelligent Transportation Systems Magazine.