Dynamic confirmation, compensation and routing for combined transportation of passengers and parcels

Y. Wang, M.Xu*

Department of Industrial and Systems Engineering, The Hong Kong Polytechnic University, Hung Hom, Hong Kong
yilun.wang@connect.polyu.hk, min.m.xu@polyu.edu.hk
* 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 26, 2024

Keywords: integrated passenger-and-parcel transportation; dynamic stochastic vehicle routing; dynamic compensation; value function approximation.

1 INTRODUCTION

In recent years, the mobility-on-demand services have expanded into a market worth over $500 billion in 2021, with projections exceeding $1600 billion by 2031 (Mohnish & Sonia, 2022). Typically, these services are provided by separate fleets, either on-demand ride services for individuals (e.g., e-hailing) or on-demand delivery services for parcels (e.g., same-day delivery). However, the isolated operation overlooks the potential of consolidating the two types of requests, which often come out in the same region. To capitalize on this opportunity, the concept of combined urban transportation (CUT) has been proposed recently. The idea is to utilize a single fleet to cater to both types of requests (Cheng *et al.*, 2023). While CUT has the potential to optimize fleet utilization, its operational success is not guaranteed considering the potential detours required for parcel delivery during passenger trips. Research on integrated passenger-parcel transportation systems has received much attention in recent years (Cheng *et al.*, 2023). Examples include Li *et al.* (2014), Li *et al.* (2016), Schlenther *et al.* (2020), etc. Those studies for CUT often ignored the potential detour or restricted the detour by enforcing deterministic constraints. Passengers' acceptance of detour, which could be uncertain in real-time operations, has not been considered. Moreover, the solution algorithms proposed by most of the existing studies for the dynamic CUT problem were purely reactive, without adequately considering the stochastic information of future requests and the impact of current decisions on the future profitability.

To fill the research gaps, this study investigates a new dynamic and stochastic confirmation, compensation, and routing problem, referred to as DCR problem. We consider a single fleet of vehicles used to serve passenger ride requests and parcel delivery requests, which arrive dynamically. Each passenger/parcel should be picked up at a specified location and dropped off at a destination before a deadline. To mitigate the negative impact of detours caused by parcel delivery on passenger ride experience and satisfaction, we consider providing compensation for affected passengers who are willing to accept detours during the trips. Specifically, passengers would be asked for the acceptance of an detour option associated with a specific amount of compensation and detour time limit. Passengers acceptance is unknown until their feedback is received. If the detour option is accepted, the service provider will make use of the detour time to deliver parcels and offer the compensation amount to the concerned passenger; otherwise, other alternative solutions will be sought. In addition to compensation decision, confirmation and routing decisions will be jointly decided during the operation. The confirmation decision determines whether to serve an upcoming request, whereas the routing decision confirms the sequence of locations to be visited by the fleet. The objective of the DCR problem is to maximize the total profit of the service provider during the service period.

2 METHODOLOGY

2.1 MDP formulation

The DCR problem is a dynamic and stochastic problem, which can be formulated as a Markov decision process (MDP) with five components: state variables, decision variables, exogenous information, transition function, and objective function. Let $\mathcal{K}$ be the set of decision epochs. Each epoch $k \in \mathcal{K}$ corresponds to the time point t_k to make a decision triggered by the arrival of a new request o_k. The state variable $\mathcal{S}_k$ includes the information for decision making at epoch k, i.e., $\mathcal{S}_k = (t_k, o_k, \Theta_k)$, where Θ_k is the current routing plan of the fleet. Each element $\theta \in \Theta_k$ represents the path of an individual vehicle. The decision variables at epoch k are determined based on state variable $\mathcal{S}_k$ and the feedback from passengers of ride requests at epoch k, which involves confirmation, compensation, and routing decisions. Let $x_k := (\zeta_k, \tau_k, \eta_k)$ denote the decision variables at epoch k, which include confirmation decision variable ζ_k, routing decision variables τ_k, and compensation decision variables η_k. The confirmation decision variable ζ_k is binary, which equals 1 if request o_k, is served and 0 otherwise. Given the current routing plan Θ_k, let Θ_k^τ denote the a set of potential routing paths for serving the requests in set $\mathcal{O}_k := \mathcal{O}_k^c \cup \{o_k\}$, where $\mathcal{O}_k^c$ is the set of requests confirmed to be served, but have not been completely served by epoch k. Then $\tau_k :=\{\tau_{\theta k} \in \{0,1\} | \theta \in \Theta_k^\tau\}$ is defined as the corresponding path-based routing decision variables, with $\tau_{\theta k} = 1$ if path $\theta \in \Theta_k^\tau$ is chosen, and $\tau_{\theta k} = 0$ otherwise. The exogenous information $\mathcal{W}_{k+1}$ is the arrival of the next new request or the termination of the operation period. The transition function $\mathcal{S}_{k+1} = \mathcal{F}(\mathcal{S}_k, x_k, \mathcal{W}_{k+1})$ transfers the current state $\mathcal{S}_k$ into $\mathcal{S}_{k+1}$ by decision x_k and exogenous information $\mathcal{W}_{k+1}$. The reward resulted from the decision in epoch k is denoted by $\mathcal{R}(\mathcal{S}_k, x_k)$. Let π be the policy that maps a state $\mathcal{S}_k$ to decision $x_k = \mathcal{X}^\pi(\mathcal{S}_k)$. We can formulate the objective function as finding the policy $\pi \in \Pi$ that maximizes the expected total reward conditional on the initial state as follows:

$$\max_{\pi \in \Pi} \; E\left\{ \sum_{k=0}^{K} \mathcal{R}(\mathcal{S}_k, \mathcal{X}^\pi(\mathcal{S}_k)) \middle| \mathcal{S}_0 \right\}. \tag{1}$$

2.2 Anticipatory policy

To efficiently solve the MDP, we propose an anticipatory policy ARCC based on approximate dynamic programming. Specifically, a routing heuristic is first proposed to generate potential feasible paths for serving the new request. We will then develop an anticipatory model (AM) based on these new paths and propose a compensation strategy. In addition, a new VAF with a slide memory is designed to learn the reward-to-go value through simulation. The proposed ARCC is expected to generate the optimal solutions rapidly in dynamic scenarios.

Routing heuristic: We design an insertion-based routing heuristic to rapidly identify potential routing plans for serving the new request. The algorithm will insert request o_k into any feasible locations along each path $\theta \in \Theta_k$ to generate the set of potential paths for serving request o_k. All the newly generated paths are grouped in set Θ_k^π. If there is no feasible paths for serving the new request o_k, the request will be declined directly and the current routing plan remains unchanged.

Anticipatory model: Based on the newly generated paths in set Θ_k^π and the existing paths in set Θ_k, we formulate an AM to determine the approximate solution at each epoch k. The objective of the AM is to maximize the total of the current reward $\mathcal{R}(\mathcal{S}_k, x_k)$ and the approximated future reward-to-go value $\hat{V}(\mathcal{S}_k, x_k)$, which will be estimated by a VFA method. We prove that the AM can be solved directly in certain cases without determining the compensation by searching among all feasible path solutions and identify the one with the largest objective function value $f^R(x_k)$.

Compensation strategy: In the other cases that AM can not be solved directly, we will determine the optimal compensation for each newly-generated path by maximizing the expected anticipatory profit considering the probability of a passenger accepting the detour option. We prove that the upper bound of the optimal solution to this maximization problem $\hat{f}$ can be calculated. Then the compensation amount is set to be $\lambda\hat{f}$, where $\lambda \in [0, 1]$ is a pre-specified parameter.

Value function approximation: The ARCC requires an estimate of expected profit. To this end, we propose a machine learning-empowered VFA to learn the value function $\hat{V}(\cdot)$ by simulating the MDP offline. Four representative attributes are used to aggregate the sequential and temporal information in each routing plan θ: the arrival time on the heading location, the remaining available time after visiting all locations, the average available time to serve future requests along θ, and the minimal available time to serve future requests along θ. Different from conventional VFA based on lookup table, the proposed VFA employs a regression tree model to learn $\hat{V}(\cdot)$, which avoids exponential scale creation of subsets as the dimension increases. In addition, the prediction of regression tree may not be reliable due to the poor observations in the early stage. We thus employ a slide memory which reserves a certain number of recent observations for training.

3 NUMERICAL EXPERIMENTS

Table 1 – *Results for Improvement in Profits*

Request number	Passenger request ratio	Fleet size	ARCC	ARCC-FB	ARCC-B	ARC-FC	ARC-M
300	0.3	3	**0.66**	0.46	0.37	0.52	0.43
300	0.3	5	**0.46**	0.31	0.25	0.30	0.37
300	0.3	7	**0.19**	0.10	0.18	0.17	0.12
300	0.5	3	**0.77**	0.70	0.42	0.63	0.63
300	0.5	5	**0.46**	0.43	0.34	0.36	0.28
300	0.5	7	**0.13**	0.10	0.04	0.06	0.04
400	0.3	3	**0.68**	0.49	0.42	0.57	0.41
400	0.3	5	**0.59**	0.34	0.33	0.37	0.39
400	0.3	7	**0.45**	0.22	0.37	0.24	0.31
400	0.5	3	**0.88**	0.79	0.44	0.71	0.75
400	0.5	5	**0.66**	0.40	0.51	0.47	0.48
400	0.5	7	**0.36**	0.23	0.19	0.32	0.19
500	0.3	3	**0.77**	0.44	0.40	0.45	0.54
500	0.3	5	**0.61**	0.17	0.47	0.31	0.27
500	0.3	7	**0.56**	0.37	0.40	0.39	0.36
500	0.5	3	**0.88**	0.82	0.51	0.80	0.59
500	0.5	5	**0.81**	0.76	0.49	0.71	0.55
500	0.5	7	**0.59**	0.37	0.58	0.55	0.40

We extensively evaluate the performance of the proposed policy in a wide range of randomly generally instances against five benchmark policies: (i) a myopic policy (MYP-M) for CUT with parcel delivery only when no passengers on board, which serves as the baseline for all other policies; (ii) an anticipatory policy (ACR-M) for the same CUT service with MYP-M; (iii) an anticipatory policy (ARCC-B) which learns the VFA via a lookup table with arrival time and remaining available time as attributes; (iv) an anticipatory policy (ARCC-FB) that learns the VFA in the same way as ARCC-B but with an additional attribute, average available time;

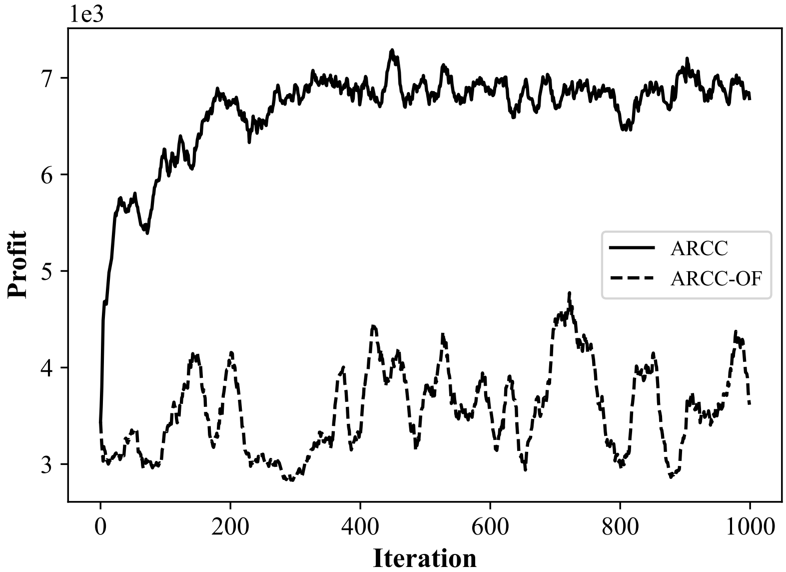

Figure 1 – *Learning process of ARCC and ARCC-OF*

and (v) an anticipatory policy (ARC-FC) with a fixed compensation amount per unit detour time. Table 1 reports the profit improvement of each policy over that of policy MYP-M for each instance. It shows that the proposed ARCC achieves the best performance in 34 out of all 36 instances, which gains 49.4% improvement in profits on average. In addition, the ARC-M achieves 31.2% more profits on average compared to that of MYP-M. The results demonstrate the efficacy of anticipatory decision. Moreover, the ARC-FC achieves 35.4% improvements, which outperforms ARCC-B and ARCC-FB with 30.0% and 33.9% improvements, respectively. Even with an anticipatory compensation strategy, ARCC-B and ARCC-FB gain poor solution quality compared to ARC-FC with fixed compensation strategy. The results show the significance of design of VFA to DCR problem. We also evaluate the performance of the proposed VFA compared with a benchmark VFA without a slide memory, referred to as ARCC-OF. Figure 1 illustrates the learning process of ARCC and ARCC-OF along the iteration of simulation. It shows that the proposed VFA significantly outperforms the benchmark.

References

Cheng, Rong, Jiang, Yu, & Nielsen, Otto Anker. 2023. Integrated people-and-goods transportation systems: from a literature review to a general framework for future research. *Transport Reviews*, 1–24.

Li, Baoxiang, Krushinsky, Dmitry, Reijers, Hajo A, & Van Woensel, Tom. 2014. The share-a-ride problem: People and parcels sharing taxis. *European Journal of Operational Research*, **238**(1), 31–40.

Li, Baoxiang, Krushinsky, Dmitry, Van Woensel, Tom, & Reijers, Hajo A. 2016. The share-a-ride problem with stochastic travel times and stochastic delivery locations. *Transportation Research Part C: Emerging Technologies*, **67**, 95–108.

Mohnish, K, & Sonia, M. 2022 (Oct). *Mobility on demand market size, analysis, share, trends - 2031.*

Schlenther, Tilmann, Martins-Turner, Kai, Bischoff, Joschka Felix, & Nagel, Kai. 2020. Potential of private autonomous vehicles for parcel delivery. *Transportation Research Record*, **2674**(11), 520–531.

Acknowledgement

The work described in this paper was fully supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (Project No. PolyU 25222822).