Understanding the lock-in effect of on-demand ride-hailing services

Yu Zhang [1,2], Ren-Yong Guo *[1], Zhenliang Ma [2]

[1] School of Economics and Management, Beihang University, Beijing 100191, China

[2] Department of Civil and Architectural Engineering, KTH Royal Institute of Technology, Stockholm 10044, Sweden

SHORT SUMMARY

On-demand ride-hailing platforms have recently been criticized for being less efficient than traditional street-hailing services in certain scenarios. To characterize this phenomenon, we propose "the lock-in effect" in ride-hailing services, in which initial matches are locked in despite better matches becoming possible later. To better understand the lock-in effect and mitigate its negative impacts, we propose a matching-flexible reactive ride-hailing framework with realistic operation settings. Case studies using real-world data are conducted to explore the existence conditions and impacts of the lock-in effect. The key findings are: (i) The lock-in effect in ride-hailing services tends to occur in low-supply scenarios. Its inefficiencies primarily affect either service rate or total service distance (but not simultaneously). (ii) The proposed mitigation strategy enhances ride-hailing efficiency across all supply scenarios (low-supply and high-supply) and demand patterns (dense, medium, and sparse demand). The findings will support the efficient operation and management of on-demand mobility services.

Keywords: Shared mobility; Ride-hailing; Street-hailing; Matching strategies; Lock-in effect; On-demand mobility services.

1. INTRODUCTION

The high efficiency and convenience of on-demand ride-hailing platforms, such as Uber, Didi, and Bolt, have fundamentally transformed contemporary travel behavior (Agatz et al., 2012; Ta et al., 2017; Simonetto et al., 2019; Alonso-Mora et al., 2017; Guo et al., 2021, 2022; Ma et al., 2022). However, the rapid growth of these platforms has significantly disrupted the traditional taxi industry (Zhong et al., 2022). While ride-hailing intuitively seems more efficient than street-hailing, recent studies suggest that it may be inefficient in certain scenarios and may even underperform compared to street-hailing (Feng et al., 2021; Castillo et al., 2017). Castillo et al. (2017) introduce the "wild goose chase" (WGC) to describe situations where an overburdened dispatch system sends drivers to pick up distant riders, wasting drivers' time and reducing earnings. As a result, driver availability decreases, both directly due to extended en-route distances and indirectly as drivers leave the platform because of lower earnings, further compounding the problem. Feng et al. (2021) corroborate these findings, highlighting the inefficiency caused by prolonged en-route distances.

Although Feng et al. (2021) and Castillo et al. (2017) have identified some inefficiencies in ride-hailing services, their continuous models address the problem from a long-term, static, and macroscopic perspective, focusing on macro-level analysis and policy evaluations. Consequently, they ignore some realistic and microscopic aspects. For example, their models do not account for the following factors affecting system performance:

- Different utility indicators for trips: These models typically consider en-route distance only (i.e., the distance that a driver travels to pick up a rider) to evaluate trip utility. In realistic scenarios, other indicators, such as travel and service distance, may also influence the trip evaluation.

- The matching approach: For simplicity, these continuous models adopt a first-in-first-out (FIFO) matching approach in general. However, in practice, ride-hailing services use a rolling-horizon matching approach.
- Demand distribution: These continuous models frequently assume idealized demand distributions. In practice, demand distributions are complex and crucial in determining system performance.

Therefore, the factors considered by the aforementioned studies to influence the performance of ride-hailing services are not comprehensive. To better understand the factors contributing to poor performance in ride-hailing services, we propose a more general definition. In the matching process of ride-hailing services, due to the dynamic and uncertain nature of participant information, the matching scheduler is constrained to make sequential and locally optimal decisions. Once matches between drivers and riders are made, these participants are occupied until they complete their trips. Even if more suitable partners for these occupied participants appear, they cannot be matched with these occupied participants. In this paper, we characterize this phenomenon as the lock-in effect in ride-hailing services, i.e., better matches cannot be formed due to the inflexibility of ride-hailing matching. The lock-in effect creates a performance gap between dynamic matching (where participant information evolves in real-time) and oracle matching (where all participant information is known in advance). In real-world scenarios, it is impractical to expect all riders to submit travel requests in advance. Consequently, the inefficiency caused by the lock-in effect is almost inevitable.

The presence of the lock-in effect limits the potential benefits of ride-hailing services. In this paper, we explore the lock-in effect of on-demand ride-hailing services. Specifically, we address two key questions: (*1*) Under what circumstances the lock-in effect occurs in on-demand ride-hailing services. (*2*) What matching strategies can mitigate the lock-in effect to improve on-demand ride-hailing performance.

To achieve this objective, we propose a matching-flexible reactive ride-hailing framework with realistic operation settings to investigate the lock-in effect in ride-hailing services. A dual-flexible matching strategy is proposed to improve the matching flexibility of drivers and riders. In the basic ride-hailing model, only idle drivers within the same timeframe are available to meet the rider requests. The dual-flexible matching strategy expands this scope by allowing en-route drivers and incoming available drivers to be eligible for new matches. This approach increases matching flexibility and enhances the system performance. We verify the proposed framework and strategy by comparing their performance with the matching radius limitation strategy in the literature using real-world data from three cities.

The main contributions of this paper are summarized as follows:

(*1*) A matching-flexible reactive ride-hailing framework with realistic operation settings is proposed.

(*2*) The concept of lock-in effect is introduced and its existence conditions and impacts in on-demand ride-hailing services are investigated.

(*3*) The dual-flexible matching strategy is proposed to improve the service efficiency and its performance is verified by comparing it with an existing model using real-world data from three cities.

2. METHODOLOGY

In this paper, we propose a matching-flexible reactive ride-hailing framework with realistic operation settings to address the real-time on-demand ride-hailing problem. The framework is matching-flexible in accommodating a dual-flexible matching strategy to improve the matching flexibility and mitigate the lock-in effect in ride-hailing services. The dual-flexible matching strategy expands the alternative set of available partners for each participant to form trips. Our framework is reactive, focusing solely on serving real-time on-demand requests without considering reservation requests or predictions of future demand. Compared with the continuous models, our framework does not rely on idealized demand distributions, it is able to utilize real-world data to investigate and analyze the complexities of real-time on-demand ride-hailing scenarios. Moreover, the rolling-horizon matching, the routing process, and the demand-supply interactions incorporated in our framework more closely mirror the operational characteristics of real-world ride-hailing platforms. Thus, the operation settings of our framework are more realistic.

The real-time on-demand ride-hailing problem can be described as follows: A fleet of drivers, each starting from an initial location, serves dynamically emerging ride requests over time. As new requests appear, the ride-hailing platform must rapidly decide whether they can be fulfilled with minimal latency. This problem aims to optimize system performance over the entire operational period, measured by service rate and total service distance metrics.

The rolling horizon framework is employed to manage the dynamic and sequential decision-making process of this problem. Fig.1 illustrates the structure of the proposed ride-hailing framework. The process is as follows: Time is divided into sequential timeframes through timeframe segmentation. In this paper, we use fixed time intervals to generate timeframes. In each timeframe, available drivers and newly emerging riders are gathered and an optimization problem is solved to address the matching in this timeframe. The decision-making timeframe scrolls horizontally over time, allowing the matching decision process to be executed continuously. The information for future timeframes remains uncertain during the matching process within a given timeframe.

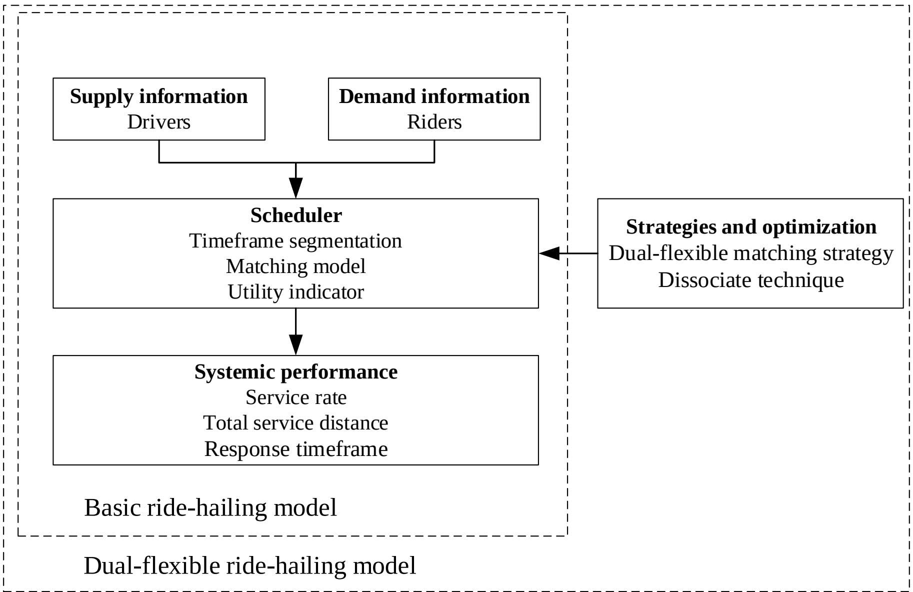


Fig.1. The structure of the ride-hailing framework.

To optimize overall performance, the framework first employs a utility indicator to estimate the benefit of each trip and then maximizes the total utility of all trips in each timeframe. Our framework considers four utility indicators to comprehensively evaluate trips: Sensible hybrid distance (SHD), negative travel distance (NTD), negative en route distance (NED), and service distance (SD). These utility indicators of the trip composed by driver d and rider r are governed by

$$u_{d,r}^1 = \xi_{d,r} - \delta_{d,r}, \tag{1}$$

$$u_{d,r}^2 = -\xi_{d,r} - \delta_{d,r}, \tag{2}$$

$$u_{d,r}^3 = -\delta_{d,r}, \tag{3}$$

$$u_{d,r}^4 = \xi_{d,r}, \tag{4}$$

respectively. These indicators are all benefit indicators, meaning that a higher value indicates a greater positive impact on the system performance. $\xi_{d,r}$ and $\delta_{d,r}$ are the service distance and en route distance of the trip, respectively.

The aforementioned process constitutes the basic ride-hailing model (RM-BA). To enhance matching flexibility and system performance, this paper proposes a dual-flexible matching strategy. This strategy incorporates two key flexibilities to improve matching efficiency. The first flexibility, termed the pre-release flexibility, allows drivers approaching the completion of their trips to participate in the next matching round. The second flexibility, termed the re-assignment flexibility, enables riders who have been matched but not yet picked up to be rematched with newly emerging and potentially more suitable drivers. By expanding the pool of candidate trips, the dual-flexible matching strategy increases matching flexibility and significantly enhances system performance. The framework implementing this strategy is referred to as the dual-flexible ride-hailing model (RM-DF).

3. RESULTS AND DISCUSSION

In this section, we conduct experiments to investigate the lock-in effect in ride-hailing services and evaluate the effectiveness of the dual-flexible matching strategy in mitigating its negative impacts. We use the Chengdu DIDI Dataset for the comprehensive model analysis and datasets from Haikou and Manhattan for framework generalizability analysis.

The existence conditions and impacts of the lock-in effect in ride-hailing services

We compare the performance of the street-hailing simulation model (SHS) with that of the basic ride-hailing models (RM-BAs) with four different utility indicators. Fig. 2 shows the service rate performance of SHS and RM-BAs with different indicators. In general, the SRs of RM-BAs with u_2 and u_3 consistently surpass that of SHS. For RM-BAs with u_1 and u_4, their SRs are initially lower at smaller fleet scales but eventually exceed that of SHS as the fleet scale increases. The results presented in Fig. 2 reveal the existence of the lock-in effect in ride-hailing services on service rates. This effect occurs under conditions of insufficient travel supply when u_1 or u_4 is adopted as the utility indicator.

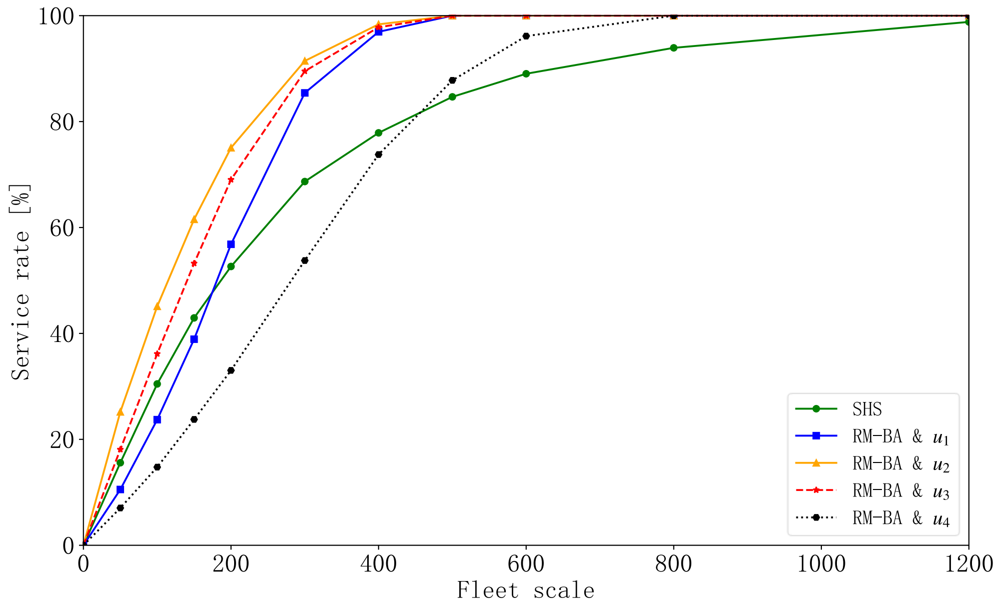


Fig. 2. Service rate performance of the street-hailing model and basic ride-hailing models with different utility indicators.

Fig. 3 shows the total travel distance performance of SHS and RM-BAs with different indicators. The TSDs of RM-BAs with u_1 and u_3 consistently exceed that of SHS. For RM-BAs with u_2 and u_4, their TSDs are initially lower at smaller fleet scales but surpass that of SHS as the fleet scale increases. The results in Fig. 3 also reveal the existence of the lock-in effect in ride-hailing services on total service distance. This effect occurs under conditions of insufficient travel supply when u_2 or u_4 is used as the utility indicator.

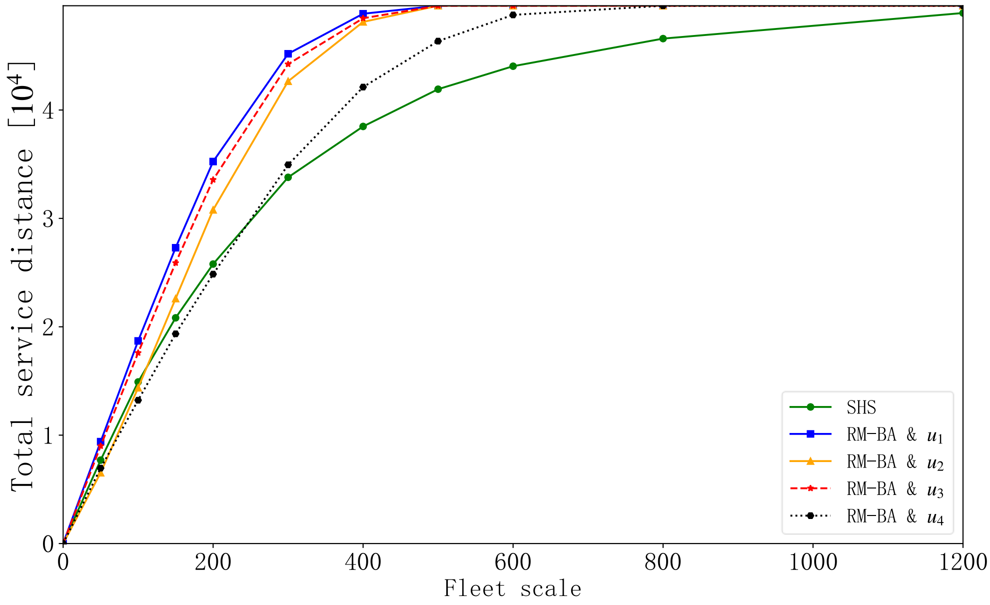


Fig. 3. Total service distance performance of the street-hailing model and basic ride-hailing models with different utility indicators.

When considering both SR and TSD performance, the choice of utility indicator distinctly influences the performance of RM-BA distinctly. These ride-hailing models are not entirely inefficient but rather Pareto-efficient.

The lock-in effect generally exists in ride-hailing services, particularly under conditions of insufficient travel supply. However, the choice of utility indicator in a ride-hailing model is critical in determining its system performance. When an appropriate utility indicator is adopted, ride-hailing services can achieve a system performance that surpasses the performance of street-hailing services in terms of either service rate or total service distance, or both.

The effectiveness and advantages of the dual-flexible strategy

In the subsection, we first verify the effectiveness of the dual-flexible strategy by comparing its performance with that of the ride-hailing model with the capped matching mechanism (RM-CMM), which is a well-established radius limitation strategy documented in the existing literature. The matching radius limitation strategy is widely employed in the existing literature to mitigate the lock-in effect and is considered an effective measure. The specific implementation of this strategy discussed in this subsection is based on the work of Feng et al. (2021). Their approach limits the en-route distance of each trip to a threshold T_d.

Figs. 4 and 5 show the comparative results on service rate performance and total service distance performance, respectively. When the fleet scale is below 200, the SR of RM-BA is lower than that of SHS. In this fleet scale range, RM-CMMs improve both SR and TSD. However, when the fleet scale exceeds 200, RM-CMMs lead to reductions in both SR and TSD. The impact of the capped matching mechanism becomes more pronounced as T_d is less, amplifying either improvements or reductions in performance. As T_d increases, the curves of RM-CMM approach to those of RM-BA, which is reasonable since RM-BA is equivalent to RM-CMM when $T_d = \infty$. In contrast, the RM-DF enhances both SR and TSD across all fleet scales and yields the most significant improvements compared with the other strategies.

Feng's research reveals several negative impacts of the lock-in effect in ride-hailing services and proposes the CMM which has a certain effect in practical scenarios. The experiments verify that the proposed dual-flexible strategy improves the system performance more significantly across the full range of fleet scales.

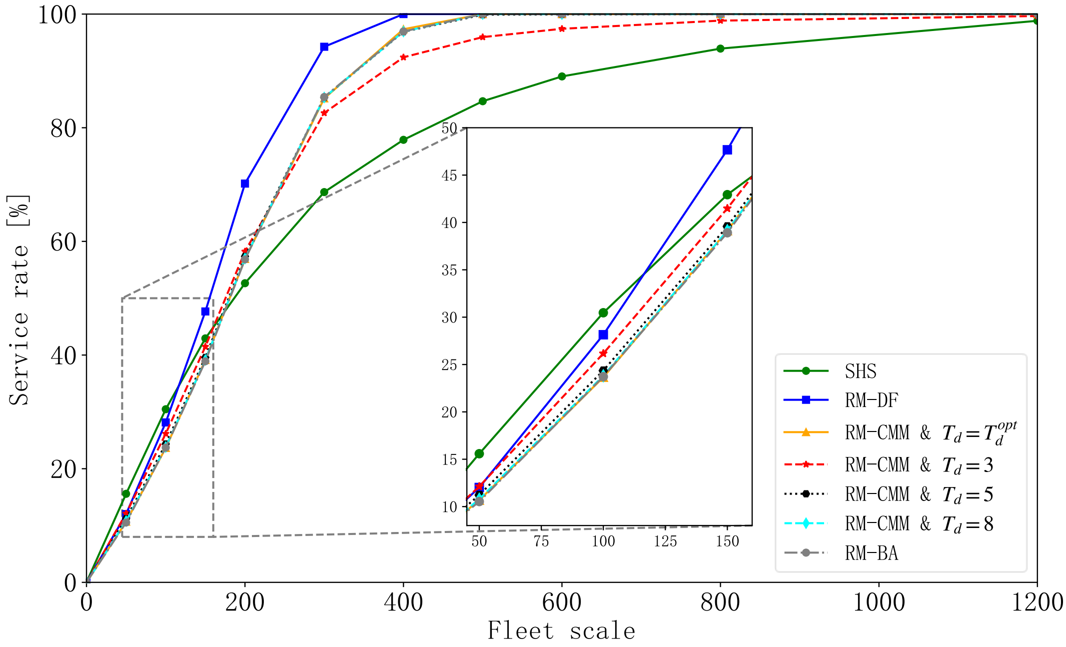


Fig. 4. SR performance of SHS, RM-DF, and RM-CMMs with different T_d.

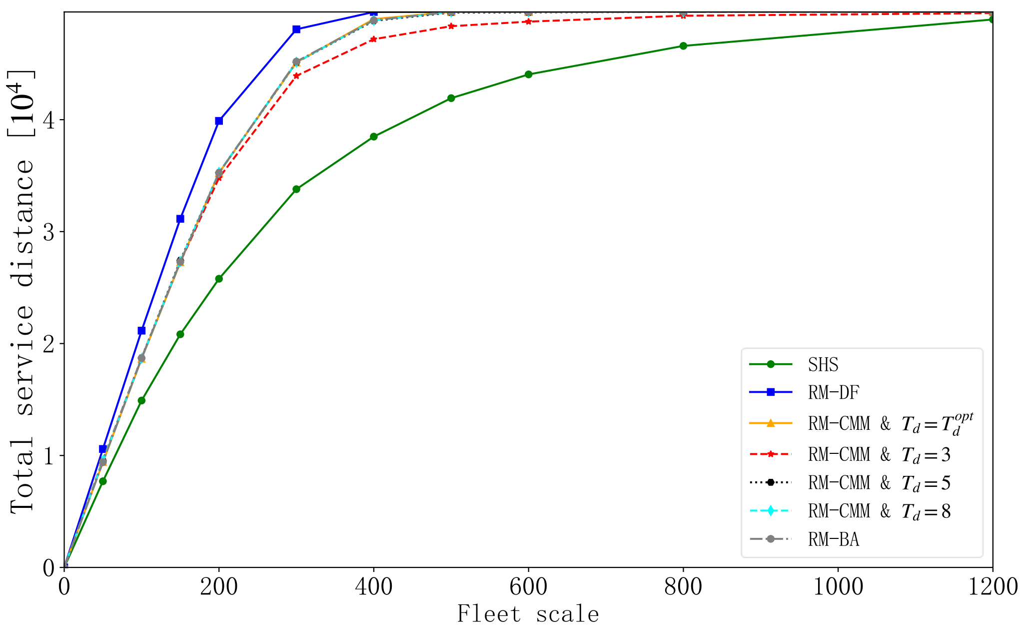
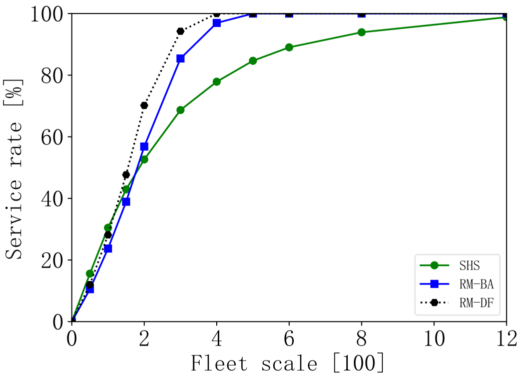
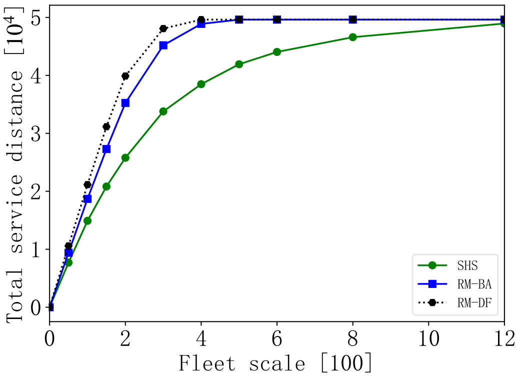


Fig. 5. TSD performance of SHS, RM-DF, and RM-CMMs with different T_d.

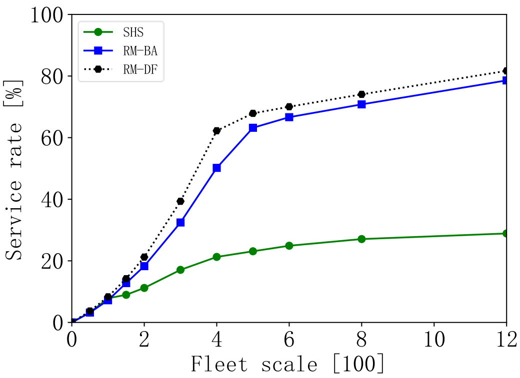
We then illustrate the advantages and generality of the dual-flexible strategy by applying it to dataset from three cities which exhibit different demand patterns: Chengdu, Haikou, and Manhattan. Haikou's demand level is comparable to that of Chengdu; however, it covers an area approximately 46 times larger, making Haikou a sparser city in terms of demand density. In contrast, Manhattan's study area is 4.7 times that of Chengdu, with its demand levels are 7.2 times that of Chengdu, resulting in a denser city regarding demand density. We implement SHS, RM-BA, and RM-DF on the three datasets to estimate the generality and stability of the proposed framework. The results are shown in Fig. 6.



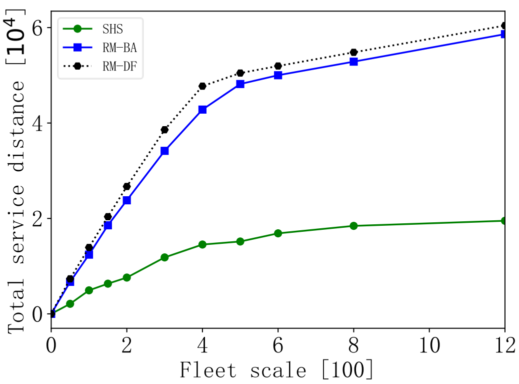
(a)



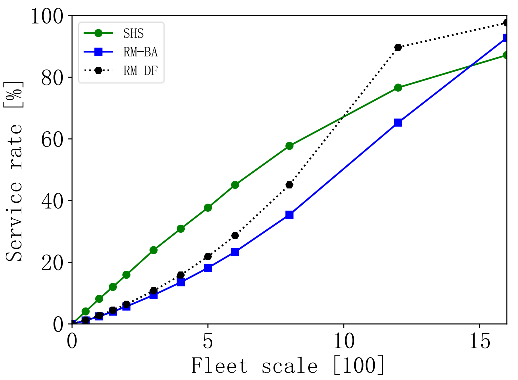
(b)



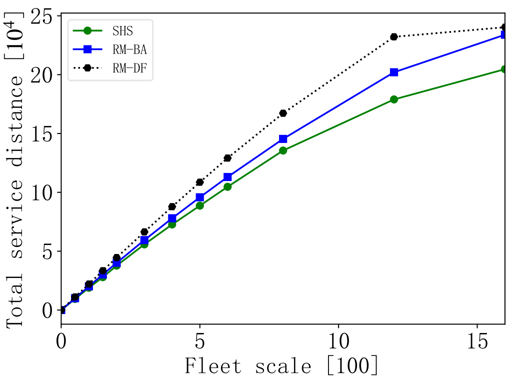
(c)



(d)



(e)



(f)

Fig. 6. Performance of the street-hailing and ride-hailing models on varying cities. (a) SR in Chengdu. (b) TSD in Chengdu. (c) SR in Haikou. (d) TSD in Haikou. (e) SR in Manhattan. (f) TSD in Manhattan.

The results of Chengdu have been discussed in previous experiments and we now focus our attention on the results for Haikou and Manhattan.

- Haikou: It is obvious that SHS performs significantly poorly in this sparser city. The larger area reduces the probability for drivers to encounter riders, resulting in low SR and TSD for SHS. Although there is also a performance reduction in RM-BA and RM-DF, this reduction remains within acceptable limits.
- Manhattan: The higher demand density increases the fleet scale crossover thresholds between the ride-hailing models and the street-hailing model. Although SHS outperforms RM-BA and RM-DF in terms of SR across the initial range of fleet scales in Manhattan, the fleet scales required to provide travel services with 100% service rates for RM-BA and RM-DF are still less than that of SHS. In addition, across the full range of fleet scales, the SR performance of RM-BA is lower than that of SHS by 24.00%. However, under the implementation of the dual-flexible strategy, this inefficiency in RM-DF is reduced to only 5.78%, and the improvement rate of TSD performance of RM-DF increases to 24.04%, compared with this rate which is only 10.85% for RM-BA.

These results indicate that the dual-flexible strategy improves the performance of RM-BA in all three cities; the advantages of RM-DF are general and stable, regardless of whether the city has normal, sparse, or dense demand. Furthermore, in the dense-demand city where the SR performance of RM-BA is likely to be worse than that of SHS, the advantages of RM-DF are more pronounced.

4. CONCLUSIONS

The paper introduces the lock-in effect in on-demand ride-hailing services, in which initial matches between drivers and riders are locked in prematurely despite better matches becoming possible later in the dynamic ride-hailing process. We propose a matching-flexible reactive ride-hailing framework with realistic operation settings to investigate the lock-in effect. The model allows a dual-flexible matching strategy including pre-release flexibility and re-assignment flexibility, enabling en route drivers and incoming available drivers to be dynamically reassigned to new riders.

We analyze case studies using real-world data to investigate the lock-in effect, evaluate its impacts, and assess the effectiveness of the proposed mitigation strategy. First, the lock-in effect tends to occur in low-supply scenarios; its inefficiencies primarily affect either the service rate or the total service distance, depending on the choice of utility indicator. Second, the proposed mitigation strategy enhances ride-hailing efficiency across all supply scenarios (low-supply and high-supply) and demand patterns (dense, medium, and sparse demand). The findings will support the efficient operation and management of on-demand mobility services. Traditional street-hailing services remain more competitive in cities with dense demand. The service rate of the basic ride-hailing model in Manhattan has a reduction of 24.00% compared with that of the street-hailing model. The dual-flexible matching strategy mitigates this reduction to 5.78% and also improves the total service distance performance to 24.04% compared with 10.85% with no such strategy. In addition, it also consistently enhances performance across all other demand patterns.

REFERENCES

Agatz, N., Erera, A., Savelsbergh, M., & Wang, X. (2012). Optimization for dynamic ride-sharing: A review. European Journal of Operational Research, 223(2), 295-303.

Alonso-Mora, J., Samaranayake, S., Wallar, A., Frazzoli, E., & Rus, D. (2017). On-demand high-capacity ride-sharing via dynamic trip-vehicle assignment. Proceedings of the National Academy of Sciences, 114(3), 462-467.

Castillo, J. C., Knoepfle, D., & Weyl, G. (2017, June). Surge pricing solves the wild goose chase. In Proceedings of the 2017 ACM Conference on Economics and Computation (pp. 241-242).

Feng, G., Kong, G., & Wang, Z. (2021). We are on the way: Analysis of on-demand ride-hailing systems. Manufacturing & Service Operations Management, 23(5), 1237-1256.

Guo, Y., Zhang, Y., & Boulaksil, Y. (2021). Real-time ride-sharing framework with dynamic timeframe and anticipation-based migration. European Journal of Operational Research, 288(3), 810-828.

Guo, Y., Zhang, Y., Boulaksil, Y., & Tian, N. (2022). Multi-dimensional spatiotemporal demand forecasting and service vehicle dispatching for online car-hailing platforms. International Journal of Production Research, 60(6), 1832-1853.

Ma Z, Koutsopoulos H N. (2022). Near-on-demand mobility. The benefits of user flexibility for ride-pooling services[J]. Transportation Research Part C: Emerging Technologies, 135: 103530.

Simonetto, A., Monteil, J., & Gambella, C. (2019). Real-time city-scale ridesharing via linear assignment problems. Transportation Research Part C: Emerging Technologies, 101, 208-232.

Ta, N., Li, G., Zhao, T., Feng, J., Ma, H., & Gong, Z. (2017). An efficient ride-sharing framework for maximizing shared route. IEEE Transactions on Knowledge and Data Engineering, 30(2), 219-233.

Zhong, Y., Yang, T., Cao, B., & Cheng, T. C. E. (2022). On-demand ride-hailing platforms in competition with the taxi industry: Pricing strategies and government supervision. International Journal of Production Economics, 243, 108301.