APEM journal

Advances in Production Engineering & Management
Volume 19 | Number 2 | June 2024 | pp 268–280
https://doi.org/10.14743/apem2024.2.506

ISSN 1854-6250
Journal home: apem-journal.org
Original scientific paper

Simulation analysis of dual-end queuing ride-hailing system considering driver-side queue management

Tang, M.C.[a,b,*], Cao, J.[a], Gong, D.Q.[c], Xue, G.[c], Khoa, B.T.[b]

[a]Xuzhou University of Technology, P.R. China
[b]Industrial University of Ho Chih Minh City, Vietnam
[c]International Center for Informatics Research, Beijing Jiaotong University, P.R. China

ABSTRACT

Ride-hailing services have transformed urban transportation through convenience yet introduced new complexities around efficiency and traffic management. This study investigates the dual-queuing problem in ride-hailing from the driver perspective using a multi-agent simulation approach. The focus is dissecting the dynamics between driver search times and passenger wait times, which critically influence operational efficiency especially during peak demand. Exploring these interactions aims to uncover insights that could improve service efficiency and customer satisfaction. Addressing such ride-hailing challenges is vital not just for individual providers but also for advancing sustainable mobility across rapidly growing metropolitan regions. Enhanced efficiency connects to broader urban development narratives around livability, accessibility, and responsible mobility ecosystems.

ARTICLE INFO

Keywords:
Dual-end queuing;
Multi-agent simulation;
System operational efficiency;
Cumulative passenger count

**Corresponding author:*
Tang12290@gmail.com
(Tang, M.C.)

Article history:
Received 21 December 2023
Revised 2 June 2024
Accepted 12 June 2024

1. Introduction

As China swiftly emerges as a global leader in on-demand mobility, ride-hailing services have profoundly transformed urban transportation paradigms. Leveraging mobile technologies, these innovative platforms efficiently match passengers and drivers, catering to escalating mobility demands within metropolitan expanses [1, 2]. Having altered commuting habits, ride-sharing now constitutes an integral mobility component for contemporary urban living [3, 4].

As an emergent mode of transport, ride-hailing presents service characteristics unprecedented in traditional travel methods. It has altered commuting habits and has had a substantial impact on the urban transportation landscape. With the rapid development of the ride-hailing market and fierce competition, enhancing service efficiency and customer satisfaction has become pivotal to the industry's growth [5]. A major challenge faced by the sector, particularly during peak times and within congested city environments, is the effective management of the matching process between drivers and passengers to reduce waiting times and improve operational efficiency. A notable operational issue within the ride-hailing industry is the often inefficient process of drivers seeking passengers, which not only diminishes drivers' work efficiency but also extends passengers' waiting periods [6]. This issue becomes especially pronounced during urban peak hours. The efficiency of the match between drivers and passengers directly affects the entire service system's efficiency, making the effective balance of driver search time and passenger wait time a pressing problem to resolve [7].

The necessity of this study lies in exploring and addressing the dual-queue problem on the driver's end within the ride-hailing industry, which is crucial for enhancing the service efficiency and customer satisfaction of the entire industry. This research aims to delve into the processes of driver search and passenger wait through in-depth analysis and simulation to identify strategies that elevate service efficiency, thereby providing effective operational and decision-making support for ride-hailing platforms. Moreover, by studying the driver-side queuing model, this research also aims to offer insights beneficial to urban traffic management and service optimization [8].

This study faces multiple challenges. First is the technical challenge of constructing a simulation model that accurately mirrors real traffic and service conditions, requiring precise modeling of complex urban traffic environments and dynamic passenger demands [9]. Secondly, processing and analyzing the vast amount of data generated by the simulation model to extract meaningful insights and strategies is non-trivial. Additionally, designing efficient algorithms to optimize driver search routes and reduce passenger waiting times is key to our research [10].

To achieve these objectives, this study employs a multi-agent simulation model. This method, by simulating the interactive behavior of drivers and passengers within a virtual environment, allows us to replicate and analyze complex urban traffic environments and ride-hailing service processes under controlled conditions. The simulation model is capable of emulating various traffic conditions and passenger demand patterns, providing an effective tool for analyzing driver search times and passenger waiting times. The crux of this approach is its ability to reproduce the complex scenarios of real life and provide real-time feedback on strategies.

The primary contribution of this research is the development of an innovative simulation model that effectively simulates the dual-queue system from the driver's perspective, analyzing its impact on the overall efficiency of ride-hailing services. Through this model, we can propose specific strategies to reduce driver search times and passenger wait times, thereby enhancing the operational efficiency of ride-hailing platforms. Additionally, this research offers valuable insights into ride-hailing services, particularly on how to effectively manage the match between drivers and passengers during peak periods. These contributions are significant not only for the practical operations of the ride-hailing industry but also for providing theoretical support for future urban traffic management and service optimization. Through comprehensive and in-depth research, this article aims to contribute significantly to the academic field of ride-hailing services and provide practical and feasible recommendations for business operations.

Tabel 1 Reminder of this paper

Section	Description	Key points
Abstract	Overview of the study focusing on ride-hailing services and dual-end queuing problems.	Highlights the significance of driver search times and passenger waiting times in ride-hailing efficiency.
Introduction	Contextualizes the rise of ride-hailing services in urban transportation, especially in China.	Emphasizes the impact of ride-hailing on urban mobility and the importance of managing driver-passenger matching.
Related works	Review of existing literature on ride-hailing system efficiency and multi-agent simulation studies.	Discusses previous findings on operational efficiency, user behavior, and dispatch strategies in ride-hailing services.
Methodology	Details the approach for analyzing the dual-end queuing problem using a multi-agent simulation model.	Describes research hypothesis, ride-hailing vehicle arrivals analysis, dual-end queuing model, and parameter settings.
Simulation results	Presents the findings from the simulation, focusing on passenger accumulation and efficiency metrics.	Analyzes the correlation between passenger numbers, driver search times, and system efficiency.
Discussion	Theoretical and practical implications of the research, along with limitations and future directions.	Reflects on the study's contribution to urban transportation and ride-hailing service management.
Conclusions	Summarizes the study's findings and their implications for optimizing ride-hailing services.	Emphasizes the importance of managing driver search and passenger wait times for service efficiency.

Table 1 provides a structured summary this paper. It outlines the key sections of the research paper, including the Abstract, Introduction, Related Works, Methodology, Simulation Results, Discussion, Conclusions, and References. Each section is concisely described, emphasizing the study's focus on the complexities of ride-hailing services, particularly the efficiency of driver-passenger matching and the impact of driver search and passenger wait times on overall service efficiency. The table encapsulates the research's theoretical and practical implications, methodological approaches, and key findings, offering a coherent overview of the paper's content and contributions to the field of urban transportation and ride-hailing service management.

2. Related works

2.1 Ride-hailing system operational efficiency

Within the contemporary transportation ecosystem, the operational efficiency, service quality, and user satisfaction of ride-hailing services constitute key research areas. Multidimensional studies in this domain have indicated that Feng *et al*. [4] engaged in a comprehensive systemic analysis to explore avenues for enhancing the operational efficiency and service quality of ride-hailing services. Additionally, Xu *et al*. [5] centered their research on the supply curve of ride-hailing systems, revealing how market conditions affect the balance of supply and demand for these services.

Further advancing the field, another study by Xu *et al*. [6] focused on optimizing vehicle dispatch within ride-hailing systems, demonstrating the application of technological means to increase service efficiency. Complementarily, Li *et al*. [7] conducted an in-depth exploration of user behavior within ride-hailing systems, particularly the decision-making processes of boundedly rational users. On another front, Lu *et al*. [8] sought to balance efficiency and fairness in ride-hailing services, especially in the context of carpooling design during specific times such as late-night hours. Meanwhile, Schlenther *et al*. [9] emphasized the issue of spatial equity in the provision of ride-hailing services.

Collectively, these studies offer a comprehensive understanding of the operations of ride-hailing systems, covering aspects such as service efficiency, user behavior, dispatch strategies, and equity. Not only do they provide new theoretical perspectives for the academic sphere, but they also offer practical guidance and strategy recommendations for operational practices.

Building on the foundation of existing research, our study delves further into the dual-queue problem at the driver's end within ride-hailing systems. We pay special attention to the relationship between driver search behavior and passenger wait times and how this relationship impacts the efficiency of the entire service system. Through the development of an integrated simulation model, we plan to simulate and analyze driver and passenger behaviors under varying conditions, seeking novel pathways to enhance service efficiency. Simultaneously, we will explore how to maintain system fairness and sustainability while ensuring efficient service, providing new theoretical insights into the ride-hailing service domain and offering feasible solutions to meet market demands and urban transportation challenges.

2.2 Multi-agent simulation studies

In the pursuit of improving shared mobility services, numerous studies leveraging multi-agent simulation models have provided significant insights. Inturri *et al*. [10] utilized a multi-agent simulation approach to plan and design new shared mobility services. Their research team, by simulating diverse traffic and shared mobility scenarios, discussed how to effectively implement shared mobility services in urban environments. This work underscored the potential applications of multi-agent systems in understanding and optimizing shared mobility services. Li *et al*. [11] employed mean field multi-agent reinforcement learning for the efficient allocation of ride-hailing orders, showcasing how advanced machine learning algorithms can be used to optimize ride-hailing order distribution, thus enhancing overall service efficiency and reducing passenger waiting times. Ke *et al*. [12] used a multi-agent deep reinforcement learning framework to study delay strategies within ride-sourcing systems, focusing on how to effectively dispatch ride-sourcing vehicles under high-demand conditions to balance supply and demand and improve system efficiency. Galland *et al.* [13] simulated individual mobility behavior in carpooling using a multi-agent simulation, providing insights on how to optimize carpooling services to reduce traffic congestion and environmental impact. Mei *et al.* [14] explored multi-modal travel policies to improve park-and-ride efficiency through multi-agent simulation, demonstrating how policy interventions can optimize the connection between public transportation and private car usage, enhancing overall travel efficiency.

These studies demonstrate the tremendous potential of multi-agent simulations in understanding and optimizing shared mobility services. By simulating complex traffic environments and user behaviors, they provide valuable insights for the planning and design of shared mobility services. Particularly with the use of advanced machine learning techniques to optimize service distribution and scheduling, these studies illustrate how data-driven approaches can enhance service efficiency and user satisfaction. Additionally, these studies consider environmental sustainability and policy impacts while improving travel efficiency, highlighting the importance of integrated approaches in solving transportation issues.

Building upon these foundations, our research further explores the application of multi-agent simulation in optimizing ride-hailing services. We focus on how simulation models can better understand and optimize the interactions between drivers and passengers and how these interactions affect service efficiency and user experience. Additionally, our research will investigate how to reduce environmental impacts and consider policy factors while ensuring efficient service. In this way, our study aims to provide new theoretical insights and practical strategies for the development of shared mobility services.

3. Methodology

3.1 Research hypothesis

In the domain of urban transportation, the dynamics of taxi services can be conceptualized as a queuing system where both the taxis and the passengers are involved in a bidirectional waiting mechanism—waiting to board and to pick up, respectively. Within this system, the metric of efficiency in boarding is intrinsically linked to the length of the queue and the associated waiting time. Theoretically, the waiting time is directly proportional to the queue length; an increase in the number of waiting passengers or taxis invariably leads to longer waiting periods.

Optimizing the efficiency of such a taxi-boarding system necessitates a strategic determination of the quantity of boarding points. An optimal number of boarding points can theoretically minimize the queue length on both ends of the system, thereby reducing the waiting time and enhancing the overall efficiency of the service. For instance, a singular boarding point would result in a bottleneck scenario where all taxis and passengers converge, leading to extended queues and increased waiting times, thereby diminishing the system's efficiency. Conversely, the introduction of multiple boarding points could distribute the demand, subsequently shortening the queue lengths and reducing waiting times, culminating in an elevated efficiency of the boarding process.

Therefore, viewing the process of taxi boarding as a queuing system sheds light on the pivotal role of queuing lengths and waiting times as critical indicators of efficiency. It emphasizes the necessity of a judicious allocation of boarding points to minimize queuing lengths, thereby optimizing the efficacy of the taxi-boarding system. This theoretical framework can guide the structuring of urban taxi services to achieve maximum operational efficiency.

3.2 Analysis of ride-hailing vehicle arrivals

Assuming that the vehicle flow is relatively low when ride-hailing cars reach their destination, the Poisson distribution is applied to model the arrival pattern of these vehicles. The number of ride-hailing taxis at the destination can thus be represented by the following equations:

$$X \sim P(\lambda) \tag{1}$$

$$P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda} \quad (k = 0,1,2,\cdots) \tag{2}$$

To further refine our model, it is essential to estimate the parameter λ within the Poisson distribution, which represents the mean rate of arrivals. Given that the overall distribution is known and randomness is adequately accounted for, it is prudent to employ the method of maximum likelihood estimation (MLE) over moment estimation or other point estimation techniques. MLE is particularly advantageous in this context due to its efficacy in estimating parameters for well-defined probability distributions like the Poisson distribution.

The specific steps to solve for the parameter λ involve setting up the likelihood function based on the Poisson probability mass function, taking its natural logarithm, and then finding the value of λ that maximizes this log-likelihood function. This optimization process typically involves taking the derivative of the log-likelihood function with respect to λ, setting it equal to zero, and solving for λ.

$$\dot{\lambda} = \frac{1}{n}\sum_{i=1}^{n} x_i = \bar{x} \tag{3}$$

Ultimately, the estimate of λ is obtained, as shown in Eq. 3. This estimate provides the expected number of ride-hailing cars arriving at the destination within a given timeframe, allowing for more effective management and allocation of resources to match service capacity with customer demand.

3.3 Dual-end queuing model

In the realm of ride-hailing services, a bi-directional queuing model encapsulates the interactions between drivers and passengers, each seeking "service" in terms of securing a ride or a fare, respectively. This model can be typified as a multi-server queuing system, where the average queue length is a critical measure of operational efficiency.

The formulas for calculating average queue length is:

$$L_q = \frac{p_0 \rho^s \rho_s}{s!(1-\rho_s)^2} \tag{4}$$

The probability of no customers in the system is

$$p_0 = \left[\sum_{n=0}^{s-1} \frac{\rho^n}{n!} + \frac{\rho^s}{s!(1-\rho_s)}\right]^{-1}, \tag{5}$$

and the probability of the service station being busy

$$\rho = \frac{\lambda}{\mu}, \tag{6}$$

$$\rho_s = \frac{\rho}{s} = \frac{\lambda}{s\mu}, \tag{7}$$

The average number of customers in the system is:

$$L = L_q + \rho, \tag{8}$$

L can be derived from the average queue length. This, in turn, informs the average time customers spend in the system

$$W = \frac{L_q}{\lambda} + \frac{1}{\mu}, \tag{9}$$

and their average waiting time:

$$W_q = \frac{L_q}{\lambda}. \tag{10}$$

Additionally, the probability of k vehicles waiting in the system is:

$$P_k = \begin{cases} \frac{\rho^{\mathrm{k}}}{k!} p_0, \ k < s \\ \frac{\rho^{\mathrm{k}}}{s! s^{k-s}} p_0, \ \mathrm{k} \geq s \end{cases}, \tag{11}$$

which serves as a significant indicator of service capacity and customer demand alignment.

These metrics – average waiting time, average number of customers, and average time spent in the system – collectively gauge the performance and quality of the service system. They reflect the system's operational tempo and efficiency, providing insights into resource allocation and customer behavior patterns. Importantly, the queuing probability when k vehicles are present aids in evaluating the system's capability to meet customer service requirements and manage busy periods.

In essence, these metrics are more than mathematical expressions; they are vital tools for assessing system performance and service quality. By understanding and applying these indicators, business decision-makers can enhance service processes, optimize resource distribution, and ultimately, elevate customer satisfaction.

In summary, the dual-queue model in ride-hailing systems, represented by a series of equations, offers a robust framework for analyzing and improving service efficiency. By refining these mathematical models and focusing on the core indicators, researchers and practitioners alike can better navigate the complexities of ride-hailing services, ensuring that operational decisions are informed, strategic, and customer-centric.

3.4 Parameter settings

To conduct the simulation for a ride-hailing service system, a set of parameters is input into a designated Input folder. These parameters are critical in shaping the simulation's framework and include:

N_{pass} (100000): This denotes the total number of passengers within the simulation, representing the demand side of the service.

N_{cars} (50): The total number of cars available, indicating the supply side of the ride-hailing system.

N_m (10000): The total simulation time, which reflects the operational time frame for the service analysis.

N_x (200) and N_y (200): These parameters define the number of grid points in the x and y directions, respectively, forming a grid that simulates the geographical area over which the service operates.

L_{block} (5): This specifies the number of points in a block, providing a measure for the simulation's spatial resolution.

$Scan_d$ (5): The scope of taxi service, which could potentially be used to define the radius within which a taxi searches for passengers.

max customer (10): The maximum number of customers that a taxi can serve before it's considered full and unable to accept new passengers.

$t_{customer}$ (100): The time step interval at which new passengers are generated in the simulation, dictating the influx of service requests.

The output data derived from these inputs are organized into three folders: pass_sum, search_sum, and wait_time. The pass_sum folder contains a log of timestamps and the corresponding cumulative number of passengers waiting for or currently on a ride. The search_sum keeps a record of all the times taken by taxis to find passengers, reflecting the efficiency of the service in matching supply with demand. Lastly, the wait_time folder tracks the duration that passengers wait before boarding, a direct measure of service responsiveness. In the pass_sum dataset, the first column lists the timestamps starting from zero, while the fourth column details the accumulated number of passengers at each timestamp. The search times and waiting times are illustrated in the third column of their respective datasets.

These parameters and the resulting data provide an intricate picture of the ride-hailing system's dynamics, allowing for an assessment of its efficiency, capacity to meet demand, and overall service quality.

Table 2 Simulation process of this paper

Stage	Description	Key considerations
Collection of input data	Gathering essential parameters such as passenger demand and driver availability.	Accurate data collection is crucial for a realistic simulation.
Simulation modeling	Simulating interactions between drivers and passengers based on the collected data.	This phase replicates real-world scenarios to understand the dynamics of the ride-hailing service.
Analysis of output data	Examining key metrics like system efficiency, driver search times, and passenger waiting times.	The analysis provides insights into the operational effectiveness and areas needing optimization.

Table 2 captures the simulation process of a dual-end queuing ride-hailing system through a detailed flow. This flowchart methodically outlines each step involved in the simulation, providing a clear and sequential depiction of the entire process. The simulation begins with the collection of input data, which includes crucial parameters such as passenger demand and driver availability. This initial stage sets the groundwork for the simulation by establishing the fundamental variables that will influence the system's dynamics. Following data collection, the process advances to the simulation modeling phase. Here, the interactions between drivers and passengers are simulated, considering the various factors collected in the initial stage. This modeling is pivotal as it replicates the real-world scenarios of a ride-hailing service, thereby enabling a comprehensive analysis of the system. The final stage of the simulation process involves analyzing the output data. Key metrics such as system efficiency, driver search times, and passenger waiting times are examined. This analysis provides valuable insights into the operational effectiveness of the ride-hailing system and highlights areas that require optimization.

Overall, Figure 1 effectively illustrates the structured approach of the simulation process, from the gathering of initial data to the final analysis of the system's performance. This step-by-step representation is crucial for understanding the methodology behind the simulation and the resulting conclusions drawn about the ride-hailing system's efficiency.

3.5 Evaluation of the dual-end queuing model simulation

In the evaluation of dual-end queuing models for ride-hailing simulations, two critical performance metrics are employed: the average waiting time for passengers and the average search time for drivers. The average waiting time, is calculated as the ratio of the total waiting time experienced by all customers to the total number of passengers served within the system, as shown in Eq. 12:

$$T_{\text{wait}} = \frac{T_{\text{wait\_total}}}{N_{\text{total}}} \tag{12}$$

Similarly, the average search time for drivers, is determined by the ratio of the total search time for all ride-hailing vehicles to the number of vehicles in the system, indicated in Eq. 13:

$$T_{\text{search}} = \frac{T_{\text{search\_totd}}}{N_{\text{\_cars}}} \tag{13}$$

The operational excellence of a ride-hailing system is inversely proportional to both the average search time and the average waiting time—the lower these values, the more optimized the system is considered to be. A comprehensive score, denoted as G is utilized to assess the overall efficiency of the ride-hailing operation, which is expressed in Eq. 14:

$$G = -\theta_1 T_{\text{wait}} - \theta_2 \tag{14}$$

Here, G represents the total score of the system, while θ_1 and θ_2 symbolize the respective weights assigned to the average waiting time and the average search time. The maximization of G is equivalent to finding the optimal weights θ_1 and θ_1, such that the conditions for the equality in the arithmetic and geometric means inequality, represented by Eqs. 15 and 16:

$$G_n = \sqrt[n]{\prod_{i=1}^{n} x_i} = \sqrt[n]{x_1 x_2 \, \mathrm{L} x_n} \tag{15}$$

$$A_n = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{x_1 + x_2 + \mathrm{L} + x_n}{n} \tag{16}$$

are satisfied. When the equality holds true, we have:

$$\frac{\theta_1}{\theta_2} = \frac{T_{\text{search}}}{T_{\text{wait}}} \tag{17}$$

Ultimately, as Eq. 18 suggests:

$$\frac{\theta_1}{\theta_2} = \frac{T_{\text{search\_totd}}}{T_{\text{wait\_totd}}} \cdot \frac{N_{\text{total}}}{N_{\_}\text{cars}} \tag{18}$$

when the input data for the number of passengers and other system inputs are fixed. Hence, the ratio of weights for the average waiting time and the average search time in the efficiency score solely depends above parameters.

By analyzing these performance metrics, it is possible to strategically manage ride-hailing operations to improve overall service quality. The implications of this model extend beyond mere operational statistics; they encompass strategic decision-making regarding resource allocation, service policy adjustments, and enhancements to customer experience. The model provides a quantitative framework to dissect and ameliorate the complex dynamics of ride-hailing services, ultimately guiding improvements in system design and management for enhanced service delivery.

4. Simulation results

In the pass_sum dataset of the ride-hailing simulation, the data represents a correlation between specific timestamps and the cumulative number of passengers at each of these moments, thus defining a relationship between the timestamp and N_{total} – the total accumulated passenger count. The ratio of θ_1 to θ_2 is positively correlated with N_{total}. This means that as time progresses, if N_{total} increases, so does the ratio of θ_1 to θ_2; if N_{total} remains constant, the ratio stays the same; and if N_{total} decreases, the ratio diminishes.

A scatter plot (Fig. 1) is generated using 100 data sets from pass_sum, with the horizontal axis representing timestamps (chosen at every 500th moment from a total of 50,000) and the vertical axis showing the cumulative number of passengers at those times. This scatter plot illustrates the general distribution of accumulated passenger numbers across 50,000 timestamps. From Figure 1, it's observed that peak passenger accumulation occurs near timestamps 10,000, 28,000, 40,000, and 44,000, while the lowest passenger numbers are near 0, 35,000, and 50,000.

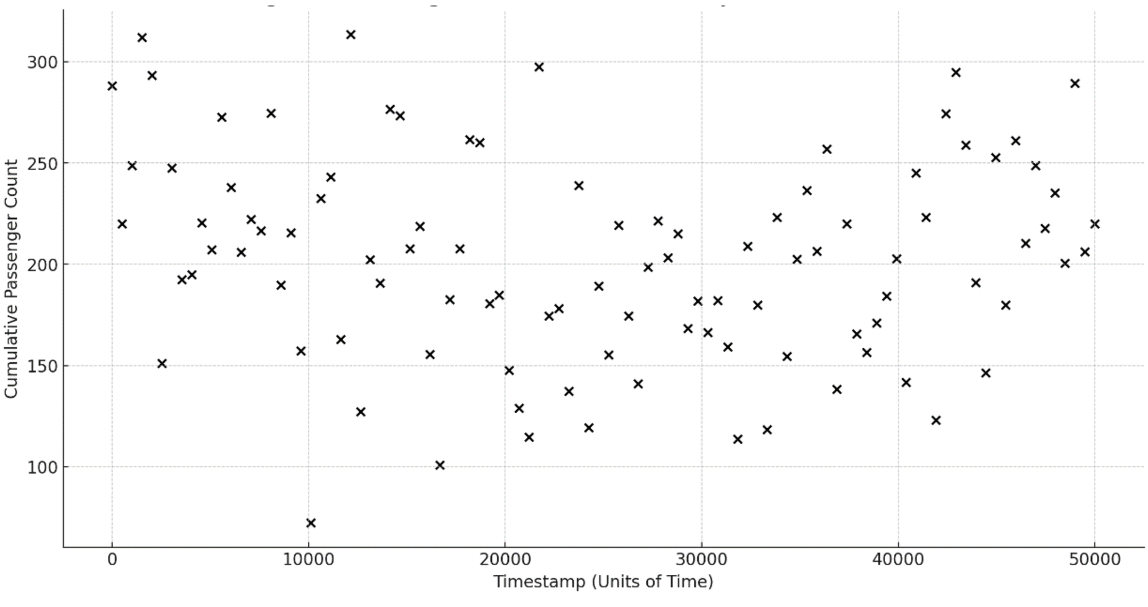

Fig. 1 Passengers accumulated in the system

The analysis of this data and the resulting plot indicates that overall, N_{total} tends to increase over time, though there are moments of decline and troughs. The positive correlation between the ratio of θ_1 and θ_2 and N_{total} suggests that the change in the ratio can approximate the trend of accumulated passenger numbers in the system. Hence, at the point where N_{total} is at its maximum, the ratio of θ_1 to θ_2 also peaks, indicating that compared to the search time, the passenger waiting time has the most significant impact on system efficiency at those moments, and vice versa. Particularly around timestamp 40,000, where the ratio is highest, the passenger waiting time most significantly impacts system efficiency; at timestamp 0, where the ratio is lowest, the search time has the greatest impact on efficiency.

In practical terms, this simulation provides insights into operational dynamics of ride-hailing services. It highlights the importance of managing passenger wait times and driver search times to optimize system efficiency. In real-world scenarios, when a system begins operation, an absence of generated passengers equates to no customer waiting time, suggesting that such extreme cases should be excluded from analysis. As the system accumulates more passengers, search times for drivers tend to increase. In contrast, when there are enough drivers waiting, passenger waiting times do not significantly increase. This reflects real-life situations and suggests that reducing driver search times might be more crucial than reducing passenger wait times when passenger numbers are high; the opposite might be true when there are fewer passengers.

The implications of these findings are significant for the practical operation of ride-hailing services. They emphasize the need for dynamic management strategies that adjust to varying demand levels throughout the day. For example, during peak hours with high passenger numbers, ride-hailing platforms might focus more on deploying drivers efficiently to reduce their search times, thereby ensuring that passengers are picked up promptly and the overall system efficiency is maintained. Conversely, during off-peak hours with fewer passengers, the focus could shift to ensuring that passengers have shorter waiting times, perhaps by predicting demand patterns and pre-positioning vehicles in strategic locations.

Moreover, the simulation results can guide policy-making in urban transport management, where balancing the need for efficient ride-hailing services with broader concerns like traffic congestion and environmental impact is crucial. For instance, during times of maximum passenger accumulation, city planners and ride-hailing platforms could collaborate to implement measures that ease traffic congestion, such as opening additional pick-up points or providing incentives for shared rides.

In summary, the simulation results offer a comprehensive understanding of how passenger accumulation trends affect the efficiency of ride-hailing systems. By analyzing these trends, ride-hailing services can optimize their operations, improving customer satisfaction and contributing to more efficient urban transport systems. This research not only provides a model for analyzing and enhancing ride-hailing services but also serves as a template for understanding similar demand-responsive transport systems, potentially leading to innovations in urban mobility solutions.

Table 3 encapsulates the simulation results, highlighting the relationship between different timestamps, the total number of passengers at these points, and the ratio, a metric indicating the system efficiency. It shows that at peak times (e.g., 10,000, 28,000, 40,000, and 44,000 timestamps), the efficiency is most affected by passenger waiting times, indicating the need for efficient passenger handling during high-demand periods. Conversely, during low passenger accumulation (e.g., timestamps 0, 35,000, and 50,000), driver search times have a more significant impact on efficiency, suggesting a focus on optimizing driver deployment during these periods. This table serves as a concise summary of the simulation's findings, guiding operational strategies for ride-hailing services.

Fig. 2 presents a simulated bivariate line graph, which vividly captures the relationship between the average search time for drivers and the average waiting time for passengers in a ride-hailing scenario. As time progresses, both lines exhibit a slight inverse trend, suggesting a degree of interaction between the two. The fluctuations in average search time appear to correspond to changes in passenger waiting time, particularly near the points where the lines intersect, which may indicate a direct link between the efficiency of matching passengers to vehicles and passenger satisfaction at certain times. These insights offer strategic implications for operational tactics, especially in balancing resource allocation during periods of high demand and low activity.

Table 3 The simulation results

Timestamp (Units)	Cumulative Passenger Count	Ratio	Implications for System Efficiency
0	Low/Minimum	Low	Greater impact of driver search time on efficiency
10,000	High/Peak	High	Passenger waiting time significantly impacts efficiency
28,000	High/Peak	High	Passenger waiting time significantly impacts efficiency
35,000	Low	Lower	Driver search time becomes more crucial for efficiency
40,000	High/Peak	Highest	Passenger waiting time most significantly impacts efficiency
44,000	High/Peak	High	Passenger waiting time significantly impacts efficiency
50,000	Low/Minimum	Low	Greater impact of driver search time on efficiency

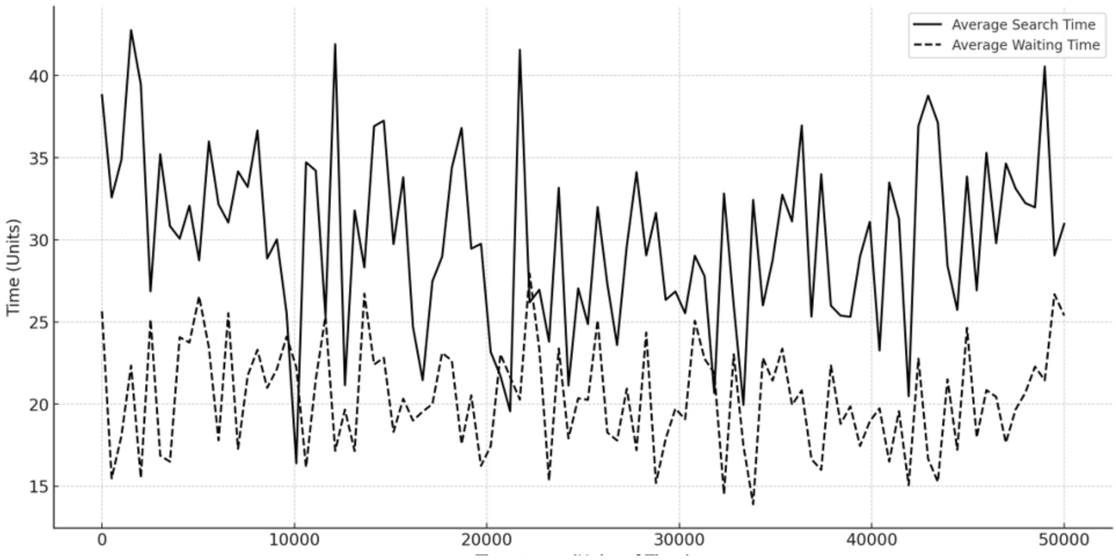

Fig. 2 Simulation of average search and waiting times in ride-hailing

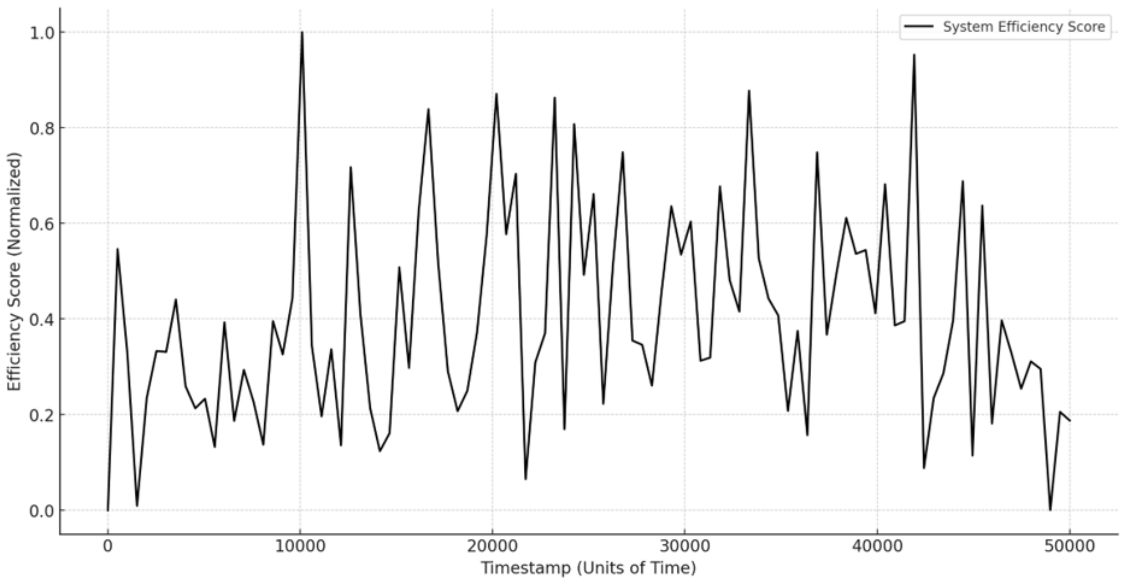

Fig. 3 Simulation of system efficiency score in ride-hailing

Fig. 3 illustrates the normalized System Efficiency Score in a ride-hailing service simulation, traced over time. The efficiency score, normalized between 0 and 1 for clarity, exhibits notable variability, with peaks suggesting instances of heightened efficiency and troughs indicating dips in performance. This oscillation reflects the dynamic nature of ride-hailing operations, where efficiency is impacted by various factors, such as driver availability and passenger demand. The graph serves as a crucial analytical tool, indicating that strategic adjustments are essential to sustain high-efficiency levels and to address the periods of lower performance that could affect customer satisfaction and operational success.

5. Discussion

5.1 Theoretical implications

This research puts forth valuable theoretical insights for the realms of urban transportation and ride-hailing service optimization. At its core, the work illustrates the potency of deploying a dual-end queuing approach to closely examine and enhance ride-hailing efficacy. This framework grants more nuanced comprehension of the complex interplay between passenger wait times and driver vacant times, alongside the collective impact on system-wide efficiency.

The revelations spotlight the intricate nature of responsive transport networks, underscoring the necessity of integrating real-time tracking and predictive analytics into flexible operational strategies. By unpacking the dynamics between various performance variables, the study enriches prevailing theoretical models surrounding ride-share platforms. The findings affirmatively highlight the need for adaptive, data-driven management practices.

Additionally, the research contributes to the evolving discourse on urban mobility ecosystems, suggesting efficient ride-hailing mechanisms can constitute a pivotal mobility component within modern cities. The work emphasizes the value of bridging theoretical constructs with practical implementations, paving inroads for innovative solutions that synthesize consumer and societal transportation needs.

In summary, the study yields multi-faceted theoretical insights around structuring sophisticated ride-hailing models to empower decision-makers in enhancing system-level and customer outcomes. The research harbors valuable potential to inform existing knowledge as well as future platforms at the intersection of transportation technology and urban advancement.

5.2 Practical implications

The practical implications of this research on ride-hailing services are wide-ranging and impactful, proffering valuable insights for operators, regulators, and urban designers alike.

Foremost, the study spotlights the potency of dynamic resource allocation in ride-hailing stewardship. By comprehending passenger volume patterns and driver availability fluxes, providers can optimize fleet positioning to slash wait times during peak-demand. This adaptive approach enhances overall customer experiences. Additionally, the findings underline the merits of demand-responsive tactics to address volatility. Platforms can harness predictive data analytics to forecast spikes and pre-deploy supply to strategic zones accordingly. Such proactive maneuvers help balance stability and efficiency.

Moreover, the work endorses greater cooperation between ride-hailing entities and metropolitan authorities to ease congestion. By collaborating to share insights on ridership and performance, these stakeholders may co-create innovative solutions such as designated access points or carpool incentives. These symbiotic measures stand to simultaneously improve services, customer satisfaction, and environmental sustainability.

In summary, this research puts forth an insightful optimization framework primed to shape decision-making within the rapidly changing urban mobility arena. The illuminated dynamics and partnerships can empower operators to unlock enhanced strategies, authorities to encourage responsible innovation, and cities to progress towards accessibility and sustainability goals [15, 16].

5.3 Limitations and future directions

While this research proffers valuable insights into ride-hailing ecosystems, the study simultaneously highlights avenues for prospective exploration given certain inherent limitations.

Primarily, the simulation model hinges on defined assumptions that cannot encapsulate the full dynamism of real-world systems. Future works could address this constraint by integrating more diverse, live data covering traffic flows and passenger behavioral variances to enhance experiential representativeness. Additionally, supplemental research factors including regulatory policies, economic climates, and sustainability impacts warrant investigation to construct more holistic perspectives.

As intelligent technologies progress, opportunities abound to cultivate sophisticated algorithms that sharpen predictive capacities and operational optimization. Comparative studies across diverse urban geographies would also elucidate how localized elements like layouts and cultural attitudes influence ride-hailing efficacy. Such international juxtapositions could unveil tailored strategies for assorted metropolitan contexts.

In summary, while constituting an informative foundation, this research indicates ample avenues to surmount current limitations as the intelligent transportation domain continues advancing. Exploring human dynamics and technological possibilities can lead to increasingly realistic and hyper-responsive urban ride systems [17, 18].

6. Conclusion

This research utilized an innovative multi-agent simulation to provide new perspectives on the dual-queuing conundrum in ride-hailing services. By meticulously tracking driver search times versus passenger wait times, the study revealed crucial insights around managing operational efficiency. As the simulation indicates, these time-based factors have an asymmetrical, proportional impact on overall system performance. When passenger demand surges during peak periods, the influence of driver search time becomes more pronounced.

Consequently, a strategic priority emerges for ride-hailing firms seeking to optimize efficiency, especially under high congestion scenarios. The focus should shift to policies and mechanisms that reduce driver vacant times between rides. Targeting search time not only directly improves system-level performance but also alleviates passenger wait time, thereby enhancing service quality and satisfaction.

As urban transportation networks evolve, such evidence-based findings can guide the sustainable growth of ride-sharing. The multi-agent simulation approach could be extended to determine optimization thresholds and evaluate the impact of interventions like driver incentives or demand-responsive pricing. Accounting for human dynamics is vital towards mobility ecosystems that balance convenience, profitability, and responsibility.

References

[1] Sun, J., Liu, S.F., Zhang, X.H., Gong, D.Q. (2022). Simulation-based modelling of the impact of ridesharing on urban system, *International Journal of Simulation Modelling*, Vol. 21, No. 1, 148-159, doi: 10.2507/IJSIMM21-1-CO2.

[2] Huang, Q.L., Wang, W.J., Liang, X.J., Xu, L., Niu, X.Y., Yang, X.Y. (2022). Last-mile delivery optimization considering the demand of market distribution methods: A case studies using adaptive large neighborhood search algorithm, *Advances in Production Engineering & Management*, Vol. 17, No. 3, 350-366, doi: 10.14743/apem2022.3.441.

[3] Han, X., Zhao, P.X., Kong, D.X. (2022). A bi-objective optimization of airport ferry vehicle scheduling based on heuristic algorithm: A real data case study, *Advances in Production Engineering & Managemen*t, Vol. 17, No. 2, 183-192, doi: 10.14743/apem2022.2.429.

[4] Feng, G., Kong, G., Wang, Z. (2021). We are on the way: Analysis of on-demand ride-hailing systems, *Manufacturing & Service Operations Management*, Vol. 23, No. 5, 1237-1256, doi: 10.1287/msom.2020.0880.

[5] Xu, Z., Yin, Y., Ye, J. (2020). On the supply curve of ride-hailing systems, *Transportation Research Part B: Methodological*, Vol. 132, 29-43, doi: 10.1016/j.trb.2019.02.011.

[6] Xu, Y., Wang, W., Xiong, G., Liu, X., Wu, W., Liu, K. (2022). Network-flow-based efficient vehicle dispatch for city-scale ride-hailing systems, *IEEE Transactions on Intelligent Transportation Systems*, Vol. 23, No. 6, 5526-5538, doi: 10.1109/TITS.2021.3054893.

[7] Li, Y., Liu, Y. (2021). Optimizing flexible one-to-two matching in ride-hailing systems with boundedly rational users, *Transportation Research Part E: Logistics and Transportation Review*, Vol. 150, Article No. 102329, doi: 10.1016/j.tre.2021.102329.

[8] Lu, C., Wu, J., Wu, C., Qin, Y., Li, Q., Ma, N. (2021). Efficiency or fairness? Carpooling design for online ride-hailing platform in transport hubs at midnight, In: *Proceedings of the 29th International Conference on Advances in Geographic Information Systems,* Beijing, China, 244-255, doi: 10.1145/3474717.3483953.

[9] Schlenther, T., Leich, G., Maciejewski, M., Nagel, K. (2023). Addressing spatial service provision equity for pooled ride-hailing services through rebalancing, *IET Intelligent Transport Systems*, Vol. 17, No. 3, 547-556, doi: 10.1049/itr2.12279.

[10] Inturri, G., Le Pira, M., Giuffrida, N., Ignaccolo, M., Pluchino, A., Rapisarda, A., D'Angelo, R. (2019). Multi-agent simulation for planning and designing new shared mobility services, *Research in Transportation Economics*, Vol. 73, 34-44, doi: 10.1016/j.retrec.2018.11.009.

[11] Li, M., Qin, Z., Jiao, Y., Yang, Y., Wang, J., Wang, C., Wu, G., Ye, J. (2019). Efficient ridesharing order dispatching with mean field multi-agent reinforcement learning, In: *Proceedings of WWW '19: The Web Conference*, San Francisco, California, USA, 983-994, doi: 10.1145/3308558.3313433.

[12] Ke, J., Xiao, F., Yang, H., Ye, J. (2022). Learning to delay in ride-sourcing systems: A multi-agent deep reinforcement learning framework, *IEEE Transactions on Knowledge and Data Engineering*, Vol. 34, No. 5, 2280-2292, doi: 10.1109/TKDE.2020.3006084.

[13] Galland, S., Knapen, L., Yasar, A.-U.-H., Gaud, N., Janssens, D., Lamotte, O., Koukam, A., Wets, G. (2014). Multi-agent simulation of individual mobility behavior in carpooling, *Transportation Research Part C: Emerging Technologies*, Vol. 45, 83-98, doi: 10.1016/j.trc.2013.12.012.

[14] Mei, Z., Wei, D., Ding, W., Wang, D., Ma, D. (2023). Multi-agent simulation for multi-mode travel policy to improve park and ride efficiency, *Computers & Industrial Engineering*, Vol. 185, Article No. 109660, doi: 10.1016/j.cie.2023.109660.

[15] Dragan, D., Šinko, S., Keshavarzsaleh, A., Rosi, M. (2022). Road freight transport forecasting: A fuzzy Monte-Carlo simulation-based model selection approach, *Tehnički Vjesnik – Technical Gazette,* Vol. 29, No. 1, 81-91, doi: 10.17559/TV-20210110140112.

[16] Gao, X., Kong, C., Wang, H., Dong, B., Ma, Z., Ren, D. (2022). Calculation of stability limit displacement of surrounding rock of deep-buried soft rock tunnel construction based on fuzzy logic matching algorithm, *Tehnički Vjesnik – Technical Gazette,* Vol. 29, No. 2, 441-448, doi: 10.17559/TV-20210607100722.

[17] Kim, Y.-J., Jung, K.-H. (2022). Social data anlysis on the perception of emergency simulation education of nursing college students using the q method, *Journal of Logistics, Informatics and Service Science*, Vol. 9, No. 1, 291-306, doi: 10.33168/LISS.2022.0118.

[18] Wei, G., Sarman, A.M., Li, M., Shen, L. (2023). A comprehensive approach for thermal comfort analysis in green intelligent buildings using BIM technology, *Journal of System and Management Sciences*, Vol. 13, No. 2, 515-528, doi: 10.33168/JSMS.2023.0235.