Online-Journals.org

iJIM International Journal of Interactive Mobile Technologies

iJIM | eISSN: 1865-7923 | Vol. 20 No. 14 (2026) | OPEN ACCESS

https://doi.org/10.3991/ijim.v20i14.62555

PAPER

# Modelling Behavior for the Optimization of Mobile Interactions and Supply–Demand Alignment for Sharing Economy Platforms

**Zhaohe Ma** (✉)

Henan Logistics Vocational College, Zhengzhou, China

hedy_mazhaohe@163.com

## ABSTRACT

Shared mobility platforms rely on mobile terminals to support large-scale service scheduling. However, the limited computational resources of mobile devices and the reliance of existing approaches on trajectory data alone have resulted in insufficient demand perception accuracy and elevated supply–demand matching latency, primarily because dynamic user interaction behaviors have been overlooked. To address these limitations, a progressive and integrated framework encompassing mobile sensing, behavior modeling, and edge-assisted matching was developed to enable fine-grained user behavior modeling and real-time supply–demand optimization in shared mobility environments. Mobile interaction entropy and conditional interaction entropy were introduced to quantify the degree of operational disorder and the temporal regularity of user behaviors, respectively. A lightweight dual-stream attention network was designed to fuse trajectory-based spatial features with interaction-behavior features, thereby facilitating accurate short-term travel demand prediction. A weighted loss function was incorporated to strengthen the fitting performance of difficult-to-predict scenarios. Within an edge computing architecture, dynamic contextual matching criteria were established through online metric learning with positive-definite constraints, while locality-sensitive hashing was employed to accelerate candidate retrieval and achieve millisecond-level supply–demand matching. Furthermore, a closed-loop edge–cloud collaborative updating mechanism was developed to balance local environmental adaptability with global model generalization.

## KEYWORDS

mobile terminal sensing, mobile interaction entropy, multimodal behavior modeling, dual-stream attention network, edge computing, online metric learning, locality-sensitive hashing, supply–demand matching in the sharing economy

Ma, Z. (2026). Modelling Behavior for the Optimization of Mobile Interactions and Supply–Demand Alignment for Sharing Economy Platforms. *International Journal of Interactive Mobile Technologies (iJIM)*, 20(14), pp. 117–131. https://doi.org/10.3991/ijim.v20i14.62555

Article submitted 2026-04-14. Revision uploaded 2026-06-01. Final acceptance 2026-06-06.

# 1 INTRODUCTION

The widespread adoption of mobile intelligent terminals has accelerated the rapid expansion of sharing-economy services [1–3], including shared mobility services [4, 5] and on-demand delivery platforms [6]. Platform operations have increasingly been centered on user interaction, state perception, and service dispatching through mobile devices. Consequently, the massive scale and highly dynamic nature of supply–demand interactions have imposed more stringent requirements on the timeliness and accuracy of platform scheduling. At present, shared-service dispatching systems remain largely dependent on cloud-centric centralized computing architectures. However, mobile terminals are inherently constrained by limited computational resources, energy sensitivity, and restricted communication bandwidth. Under such conditions, processing paradigms based on long-distance data transmission and centralized cloud inference [7, 8] have been found to be insufficient for satisfying the low-latency scheduling requirements associated with highly concurrent service environments. Furthermore, supply–demand decision-making mechanisms implemented by existing platforms are primarily driven by the mining of user travel demand from global positioning system trajectory data alone [9, 10]. Dynamic interaction behaviors generated on mobile devices, including clicking, page refreshing, and vehicle-type switching, are generally overlooked. As a result, latent demand characteristics, such as waiting anxiety and decision uncertainty, cannot be effectively represented, thereby increasing the likelihood of demand estimation errors and matching failures.

Substantial progress has been achieved in recent years in the areas of mobile behavior analysis, travel demand prediction, and supply–demand matching. Nevertheless, significant limitations remain in the practical deployment of these approaches within mobile shared-economy environments. From the perspective of user behavior modeling, existing mobile behavior analysis methods have primarily focused on general human–computer interaction scenarios [11]. Dedicated quantitative frameworks capable of characterizing the transient and irregular operational patterns observed in shared mobility services have not yet been fully established. Consequently, temporal dependency relationships embedded within user behaviors cannot be adequately captured. In addition, spatial trajectory information and dynamic interaction behaviors are commonly modeled independently, resulting in the loss of intrinsic correlations between these complementary information sources and limiting the accurate identification of uncertain travel demand [12]. From the perspective of demand prediction, mainstream multimodal deep learning models are often characterized by complex network architectures and substantial inference overhead, rendering them unsuitable for deployment on resource-constrained mobile terminals. Moreover, refined fusion mechanisms specifically designed for temporal interaction behavior features remain insufficiently explored. As a consequence, model adaptability to rapidly changing operational environments remains limited [13–15].

From the perspective of supply–demand matching, conventional matching algorithms generally rely on fixed distance metrics [16, 17]. Such static criteria cannot adapt effectively to dynamic spatiotemporal variations in traffic conditions, service capacity distribution, and user demand patterns. Furthermore, centralized matching architectures are frequently associated with excessive communication and inference delays, making it difficult to satisfy the millisecond-level real-time scheduling requirements of large-scale platforms. From the perspective of edge-side data processing, most existing studies have not optimized data transmission mechanisms for mobile devices [18, 19]. Raw sensing data are typically uploaded directly to remote servers, resulting in substantial redundant transmission overhead.

Consequently, terminal energy consumption and communication resource utilization are significantly increased, which is inconsistent with the fundamental design principles of lightweight mobile computing systems.

To address the aforementioned limitations, a series of methodological innovations is developed around mobile behavior quantification, feature fusion, dynamic matching, retrieval acceleration, and system-level optimization. Mobile interaction entropy and conditional interaction entropy are introduced as quantitative metrics to enable accurate characterization of interaction irregularity and temporal behavioral patterns in shared mobility environments. A lightweight dual-stream attention network is designed to facilitate the adaptive fusion of multimodal features derived from spatial trajectories and interaction behaviors, thereby improving demand prediction accuracy while maintaining strict control over model complexity and inference overhead. A dynamic contextual matching criterion is established through edge-based online metric learning, and positive-definite matrix constraints are incorporated to overcome the limitations of conventional fixed-distance metrics in dynamically evolving travel scenarios. Furthermore, a hashing-based retrieval strategy compatible with dynamic metrics is introduced to substantially reduce matching complexity, thereby enabling millisecond-level real-time matching at the network edge. An end–edge–cloud collaborative closed-loop optimization architecture is constructed to simultaneously preserve local contextual adaptability and global model generalization capability, ensuring stable long-term system evolution and continuous performance enhancement.

## 2 METHODOLOGY

### 2.1 Overall system architecture

To address the challenges associated with resource-constrained mobile terminals, highly dynamic user demand, and real-time supply–demand matching in sharing economy environments, an integrated hierarchical framework is developed that progressively incorporates mobile terminal sensing, behavior modeling, and edge-enabled matching. In addition, a collaborative end–edge–cloud closed-loop feedback mechanism is introduced to support continuous system adaptation and optimization. The proposed architecture is designed through a layered optimization strategy that explicitly addresses the constraints encountered throughout the entire mobile service lifecycle. At the mobile terminal layer, lightweight acquisition, temporal synchronization, and compressed preprocessing of multisource behavioral and sensing data are performed to minimize terminal energy consumption and communication overhead at the data source. At the behavior modeling layer, dynamic interaction irregularities exhibited by users are quantitatively characterized. By integrating dual-modal information derived from spatial trajectories and interaction behaviors, accurate modeling and intensity prediction of time-varying travel demand are achieved. At the edge matching layer, an adaptive contextual matching criterion is established through online metric learning, and low-latency intelligent dispatching is realized through the integration of efficient retrieval algorithms. Furthermore, a closed-loop optimization mechanism combining local incremental updating at edge nodes with global parameter calibration in the cloud is incorporated. Through this design, dynamic adaptation to regional operational contexts is preserved while maintaining strong global model generalization capability. Consequently, the integrated objectives of lightweight and low-overhead operation on mobile terminals, highly responsive scheduling at edge nodes, and high-accuracy supply–demand matching across the entire service network are simultaneously achieved. The overall technical architecture is illustrated in Figure 1.

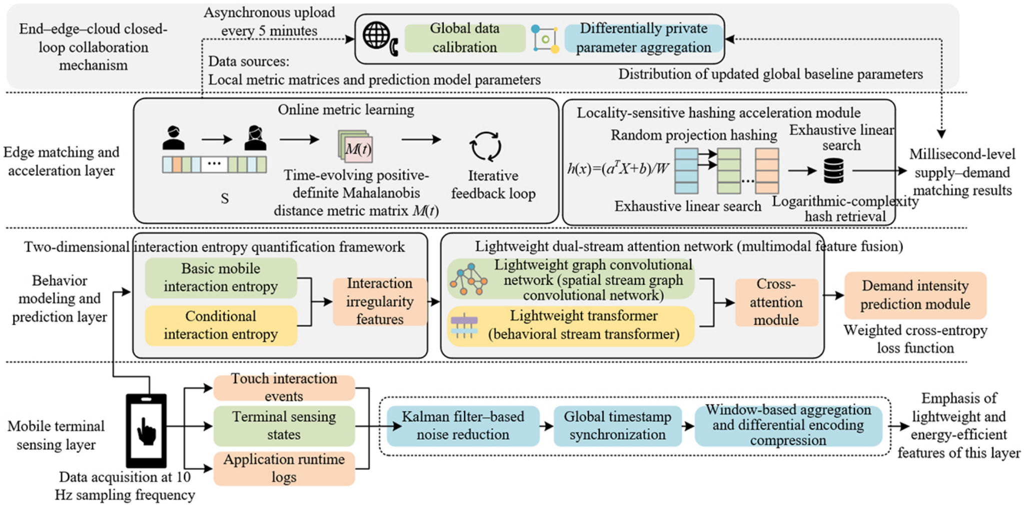

**Fig. 1.** Progressive integrated architecture for mobile terminal behavior sensing and supply–demand matching

## 2.2 Layer I: Lightweight multisource data acquisition and preprocessing on mobile terminals

This layer serves as the sensing foundation of the overall framework, where synchronized acquisition, temporal calibration, and lightweight preprocessing of multisource heterogeneous data are performed. The primary objective is to ensure that data collection and transmission remain compatible with the energy and bandwidth constraints of mobile terminals. To achieve this objective, embedded lightweight data acquisition modules are deployed within both user-side and service-provider-side mobile applications. Three categories of information are synchronously collected at a fixed sampling frequency of 10 Hz: touch interaction events, terminal sensing states, and application runtime logs. Through this design, user operational behaviors, spatial mobility states, and service interaction processes are comprehensively captured. To address temporal misalignment and random noise commonly present in raw sensing data, noise reduction is performed using a Kalman filtering algorithm, while temporal synchronization across heterogeneous data sources is achieved through a global timestamp alignment mechanism. Subsequently, standardized dynamic user behavior sequences are constructed using a sliding time-window strategy. The resulting sequence is defined as:

$$S_u(t) = \{(e_1, t_1, s_1), (e_2, t_2, s_2), \ldots\} \quad (1)$$

where, $e_k$ denotes the event type associated with the $k$-th terminal interaction, $t_k$ represents the corresponding occurrence timestamp, and $s_k$ denotes the sensing-state vector composed of geographic coordinates, movement velocity, and device orientation. Collectively, these variables provide a comprehensive representation of users' temporally evolving travel behaviors. To avoid the communication redundancy and excessive energy consumption associated with direct transmission of raw sensing data, window-based aggregation and differential encoding compression are performed locally on mobile terminals. Redundant information is eliminated by aggregating repetitive temporal features within each window and encoding only the differences between consecutive observations. As a result, the volume of uplink data transmission is substantially reduced. While preserving the integrity of behavioral sequence representations, the computational burden, device energy consumption, and communication overhead imposed on mobile terminals are effectively minimized. Consequently, high-quality and low-overhead data inputs are provided for subsequent behavior quantification modeling and real-time matching at edge nodes.

### 2.3 Layer II: Behavior modeling and demand prediction based on mobile interaction entropy and a dual-stream attention network

In shared mobility environments, decision uncertainty and waiting anxiety frequently induce irregular interaction behaviors. Such latent demand characteristics cannot be effectively captured through conventional trajectory data alone. To quantitatively characterize the degree of interaction irregularity and the dynamic evolution of user behaviors, a two-dimensional interaction entropy quantification framework is established based on fixed-length sliding time windows, thereby enabling a fine-grained mathematical representation of user behavior. Within the time interval $[t–\Delta t, t]$, the occurrence frequencies of different interaction events are first calculated. The degree of behavioral disorder within the current window is then measured through mobile interaction entropy, which is defined as:

$$H_u(t) = -\sum_{i=1}^{M} p_i(t)\log p_i(t) \tag{2}$$

where $M$ denotes the total number of interaction event categories supported by the platform, and $p_i(t)$ represents the probability of occurrence of the $i$-th interaction event within the current time window. Although mobile interaction entropy effectively characterizes the static distribution of interaction behaviors, temporal dependency relationships among successive events cannot be captured. To address this limitation, conditional interaction entropy is further introduced to quantify the dynamic dependencies and behavioral transition patterns embedded within interaction sequences. Conditional interaction entropy is formulated as:

$$H_u^{cond}(t) = -\sum_{j=1}^{L}\sum_{i=1}^{M} q_{j,i}(t)\log \frac{q_{j,i}(t)}{\sum_i q_{j,i}(t)} \tag{3}$$

where, $q_{j,i}(t)$ denotes the joint probability that an interaction event of type $j$ is immediately followed by an interaction event of type $i$ within the current time window, and $L$ is equal to the total number of event types. The two entropy measures are subsequently concatenated and fused to construct a user-specific behavioral feature vector capable of simultaneously representing interaction irregularity and temporal behavioral evolution.

User travel behavior is characterized by two heterogeneous modalities, namely spatial mobility trajectories and terminal interaction behaviors. Dynamic travel demand cannot be comprehensively represented through single-modality feature modeling alone. To address this limitation, a lightweight dual-stream attention network is constructed, in which feature extraction is performed independently for the spatial trajectory stream and the interaction behavior stream, while adaptive cross-modal fusion is achieved through a cross-attention mechanism. Owing to its compact architecture, the proposed model can be effectively deployed under the computational constraints of resource-limited mobile terminals. For the spatial stream, the urban road network is partitioned into uniform grid cells of 200 m × 200 m. User location coordinates recorded at different time steps are mapped onto the corresponding grid cells, resulting in a spatiotemporal feature tensor defined as $X_u^{spatial}$ÎR$^{T \times N_g \times 2}$. A lightweight graph convolutional network is employed to capture spatial correlations within the road network and the spatiotemporal mobility patterns of users. For the interaction behavior stream, discrete interaction events are transformed into unified embedding representations, generating a temporal feature sequence $X_u^{act}$ÎR$^{T \times d_e}$, where the embedding dimension is set to ($d_e = 64$). A lightweight Transformer architecture consisting of only two encoder layers and four-head self-attention is adopted. Through this design, temporal behavioral dynamics can be effectively captured while maintaining a minimal parameter footprint and computational overhead. To exploit the complementary information contained in the two modalities, a cross-attention module is introduced. Spatial features are treated as query vectors, whereas interaction behavior features are utilized as key and value vectors. Through adaptive learning of cross-modal correlation weights, a global feature representation $f_u(t)$, which simultaneously incorporates spatiotemporal information and user-specific behavioral characteristics, is finally output.

To enable quantitative prediction of short-term travel demand, a travel demand intensity prediction module is constructed based on the fused multimodal feature representation. High-level feature mapping is performed through a two-layer fully connected network, and the predicted travel demand intensity at the subsequent time step is obtained through sigmoid-based normalization. The prediction function is defined as:

$$\hat{r}_u(t+\delta) = \sigma(W_2 \cdot ReLU(W_1 \cdot f_u(t) + b_1) + b_2) \tag{4}$$

where, $W_1$ and $W_2$ denote the weight matrices of the two fully connected layers, $b_1$ and $b_2$ represent the corresponding bias vectors, and $\sigma$ denotes the sigmoid activation function. Conventional static loss functions exhibit limited capability in modeling highly dynamic scenarios characterized by substantial behavioral irregularity and rapidly evolving travel demand patterns. Consequently, a weighted binary cross-entropy loss function is introduced based on conditional interaction entropy. Through adaptive weighting, higher training priority is assigned to samples associated with elevated behavioral uncertainty. The loss function is formulated as:

$$L_{demand} = -\sum_u w_u [r_u \log \hat{r}_u + (1 - r_u)\log(1 - \hat{r}_u)] \tag{5}$$

where, $r_u$ denotes the ground truth demand label of the user. The weighting factor is computed as $w_u = 1 + \alpha H_u^{cond}(t)^2$, where $\alpha$ is a hyperparameter controlling the weighting gain. Through this mechanism, training weights are adaptively adjusted according to the real-time behavioral irregularity of individual users. This improves predictive accuracy and generalization robustness under complex and dynamically evolving travel scenarios.

### 2.4 Layer III: Edge-side supply–demand matching optimization based on online metric learning and locality-sensitive hashing acceleration

In shared mobility environments, geographic locations, traffic conditions, user demand, and service capacity continuously evolve over time. Consequently, fixed distance metrics are unable to accurately characterize matching suitability under dynamically changing spatiotemporal conditions. To address this challenge, an extended feature representation is constructed by integrating multidimensional information, including user and driver locations, travel demand intensity, order acceptance propensity, and traffic condition coefficients. Based on this representation, a context-aware matching criterion is established through a time-varying Mahalanobis metric. The corresponding distance function is defined as:

$$d(u, v) = \sqrt{(p_u - p_v)^{\top} M(t)(p_u - p_v)} \tag{6}$$

where, $p_u$ and $p_v$ denote the extended feature vectors of the user and driver, respectively, and $M(t)$ represents a symmetric positive-definite metric matrix of the same dimensionality. The parameters are continuously updated over time. The matrix can adaptively adjust the weights of individual feature dimensions according to real-time scenarios, enabling the measurement results to accurately reflect the matching cost under the current environment and thereby achieve effective adaptation to dynamic travel scenarios.

To enable autonomous optimization of the metric matrix, online metric learning is performed at edge nodes using historical matching records. Parameter convergence is guided through triplet-based constraints. Drivers involved in successful matches are treated as positive samples, whereas candidate drivers who were not selected are treated as negative samples. Subsequently, a corresponding contrastive loss function is designed as follows:

$$L_{metric} = \sum_{(u,v_+,v_-)\in T} \max(0, d^2(u,v_+) - d^2(u,v_-) + \gamma) \tag{7}$$

Where $T$ denotes the triplet set obtained through batch sampling, and $\gamma$ is the margin the parameter used to control the separation between positive and negative samples. Incremental parameter optimization is performed through stochastic gradient descent, and the metric matrix is updated according to:

$$M(t+1) = M(t) - \eta \nabla_M L_{metric} \tag{8}$$

where, $\eta$ denotes the learning rate. To preserve the positive-definite property of the matrix, Cholesky decomposition is applied to $M(t)$ such that $M = LL^{\top}$, and the lower triangular matrix $L$ is subsequently updated within the manifold space. To further limit computational and storage overhead at edge nodes, only matching records generated within the most recent ten-minute time window are retained for model updates.

When the size of the driver pool becomes large, exhaustive search over all candidate drivers introduces substantial computational latency and is unable to satisfy the real-time requirements of intelligent dispatching. To address this limitation, a dedicated locality-sensitive hashing strategy is designed for the dynamic Mahalanobis metric. A family of random-projection hash functions is employed to partition feature representations into hash buckets. The hash function is defined as:

$$H = \left\{ h(x) = \left\lfloor \frac{a^{\top} x + b}{w} \right\rfloor \right\} \tag{9}$$

where, $a$ is the random vector that follows a Gaussian distribution with covariance matrix $M^{-1}$, $b$ is the offset uniformly sampled from the interval [0, $w$], and $w$ denotes the hash-bucket width parameter. Multiple hash tables are deployed to perform feature mapping. During the retrieval stage, distance computations are performed only among driver candidates located within the same hash buckets. As a result, the computational complexity of candidate retrieval is reduced from linear complexity to approximately logarithmic complexity. After candidate filtering, the final matching decision is determined according to a predefined distance threshold. When executed on edge nodes, the entire workflow maintains an end-to-end latency of less than 5 ms, thereby effectively ensuring the real-time performance of platform scheduling.

## 3 EXPERIMENTAL EVALUATION AND RESULTS ANALYSIS

### 3.1 Experimental setup and dataset description

The experimental platform was implemented using an end–edge–cloud hierarchical deployment architecture comprising mobile terminals, edge-computing nodes, and cloud servers. Mobile-side testing was conducted on representative Android and iOS devices. The computational capabilities of these devices were consistent with those of commercially available consumer-grade mobile terminals, while their operating power consumption remained within the typical range observed in practical mobile deployment environments. Edge nodes were deployed on computing units associated with 5G base stations and were equipped with low-latency processing and local caching capabilities. Cloud servers were configured with high-performance processors and were responsible for global data aggregation and baseline model calibration. The software environment was established on the Linux operating system. Model training and inference were implemented using the PyTorch deep learning framework, while computationally intensive algorithmic components were optimized and compiled in C++ to improve execution efficiency at edge nodes.

Experimental data were collected from real-world mobile-terminal records obtained from a large-scale shared mobility platform. Data acquisition was conducted continuously over a 90-day period and covered a diverse range of travel environments, including central business districts, residential areas, and transportation hubs. The dataset contained operational records from more than 20,000 users and over 3,000 active drivers. Comprehensive information was included, comprising touch interaction events, multisource sensor time-series data, application runtime logs, and ground-truth order records. The dataset was partitioned into training, validation, and testing subsets using a ratio of 7:1:2. The training set was utilized for iterative model optimization, the validation set was employed for hyperparameter tuning, and the testing set was reserved for unbiased assessment of model generalization capability. Through this experimental design, the reliability and practical relevance of the evaluation results were ensured.

### 3.2 Validation of mobile interaction entropy

To evaluate the representational capability of mobile interaction entropy and conditional interaction entropy for dynamic user behavior characterization, as well as their contribution to travel demand prediction, a series of comparative experiments was conducted using different feature combinations. The contributions of the two entropy-based indicators were quantitatively assessed through feature ablation studies.

As illustrated in Figure 2, the lowest predictive performance was observed when only trajectory features were utilized. Under this configuration, the prediction accuracy reached only 82.36%, while the mean absolute error remained as high as 0.187, indicating that spatial location information alone is insufficient to fully characterize complex travel demand patterns. After the incorporation of basic interaction statistical features, the prediction accuracy was improved by 3.43 percentage points, and the mean absolute error was reduced to 0.142. These results suggest that terminal interaction behaviors provide valuable complementary information beyond trajectories. However, because only shallow statistical descriptors were employed, the underlying behavioral patterns could not be adequately captured. A substantial improvement in predictive performance was achieved after the introduction of mobile interaction entropy and conditional interaction entropy features. Compared with the trajectory-only configuration, prediction accuracy was increased by 9.09 percentage points, while the mean absolute error was reduced by more than 60%. Furthermore, relative to the configuration combining trajectories and basic interaction statistical features, an additional improvement of 5.66 percentage points in accuracy was achieved, accompanied by a further reduction in prediction error. The observed performance gains provide clear evidence that mobile interaction entropy and conditional interaction entropy effectively capture user behavioral irregularity and temporal behavioral dependencies embedded within interaction sequences. Consequently, latent demand characteristics that cannot be identified through conventional statistical descriptors are successfully represented. These entropy-based indicators constitute critical features for improving demand prediction accuracy in dynamic scenarios.

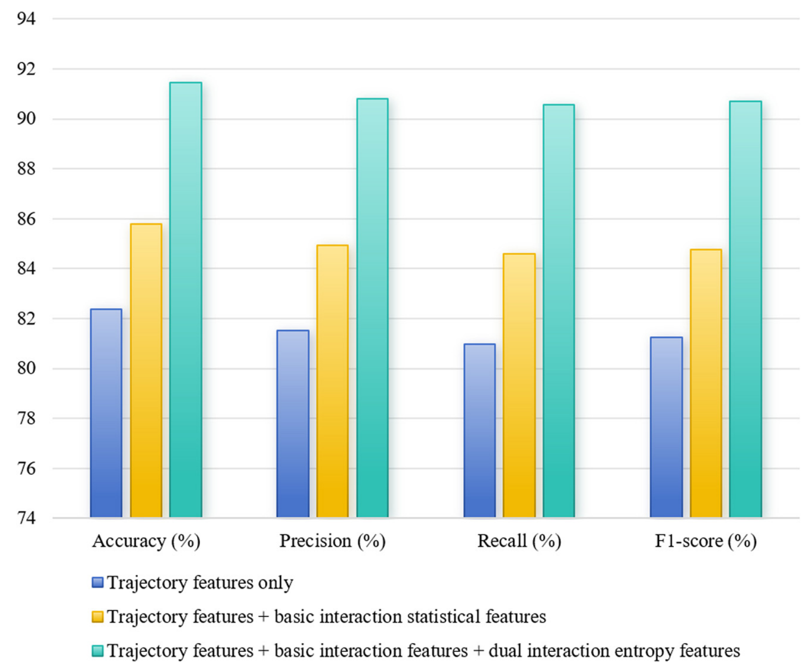


**Fig. 2.** Comparison of demand prediction performance under different feature combinations

### 3.3 Performance comparison of the lightweight dual-stream attention network

To evaluate the comprehensive advantages of the proposed lightweight dual-stream attention network in terms of prediction accuracy, model size, and mobile-side inference efficiency, three representative prediction models were selected as benchmark methods and compared under identical experimental conditions.

As shown in Table 1, the single-stream convolutional neural network model and the standard Transformer model were constrained by their limited ability. Consequently, the highest prediction accuracy achieved by these baseline models was only 88.63%. In addition, model sizes exceeding 2000 KB and inference latencies greater than 12 ms were observed in both cases, indicating limited suitability for deployment on resource-constrained mobile terminals. Although the non-lightweight dual-stream fusion model benefited from multimodal feature fusion and achieved an accuracy of 90.12%, substantial computational redundancy remained. As a result, the model size reached 2945 KB, and the single-sample inference latency increased to 15.48 ms. The proposed lightweight dual-stream attention network achieved the best overall performance among all evaluated approaches. A prediction accuracy of 91.45% was obtained, while substantial reductions in both model size and inference latency were simultaneously achieved. The model size was reduced to only 896 KB, corresponding to a reduction of 69.1%–74.9% relative to the competing approaches. Furthermore, the single-sample inference latency was decreased to 3.25 ms, representing less than one-fifth of that required by the standard Transformer model. These results demonstrate that the streamlined network architecture and cross-modal attention mechanism effectively preserve the benefits of dual-modal feature fusion while substantially reducing computational complexity. Consequently, an improved balance between prediction accuracy and mobile-terminal deployment constraints is achieved.

**Table 1.** Comprehensive performance comparison of different prediction models

| Model | Accuracy (%) | F1-score (%) | Model Size (KB) | Single-Sample Inference Latency (ms) |
|---|---|---|---|---|
| Single-stream convolutional neural network model | 84.21 | 83.96 | 2186 | 12.36 |
| Standard Transformer model | 88.63 | 88.41 | 3572 | 18.72 |
| Non-lightweight dual-stream fusion model | 90.12 | 89.95 | 2945 | 15.48 |
| Proposed lightweight dual-stream attention network | 91.45 | 90.69 | 896 | 3.25 |

### 3.4 Comparative evaluation of dynamic online metric learning for supply–demand matching

To evaluate the effectiveness of the proposed dynamic contextual metric matrix and online metric learning strategy in supply–demand matching optimization, the proposed approach was compared with three representative matching methods. Algorithmic performance was assessed from the perspective of practical dispatching effectiveness.

**Table 2.** Performance comparison of different matching algorithms

| Matching Algorithm | Matching Success Rate (%) | Average Rriver Empty-Travel Distance (m) | Average User Waiting Time (s) | Total Global Matching Cost |
|---|---|---|---|---|
| Euclidean distance matching | 83.27 | 896.35 | 28.64 | 1.000 |
| Fixed Mahalanobis distance matching | 86.53 | 742.68 | 22.31 | 0.826 |
| Static metric learning matching | 88.19 | 685.42 | 19.57 | 0.753 |
| Proposed dynamic online metric learning matching | 93.74 | 512.83 | 11.42 | 0.518 |

The quantitative results presented in Table 2 clearly demonstrate the advantages of the dynamic metric strategy. When Euclidean distance was employed, only spatial proximity was considered during the matching process. Consequently, the lowest matching success rate was observed, while both average driver empty-travel distance and average user waiting time reached the highest values among all evaluated methods. The corresponding total global matching cost was normalized to a baseline value of 1.000. Performance improvements were achieved after the incorporation of multidimensional feature representations through fixed Mahalanobis distance matching and static metric learning. Matching success rates were increased by 3.26 and 4.92 percentage points, respectively. However, because the underlying metric parameters remained fixed after training, adaptation to continuously evolving conditions could not be effectively achieved. The proposed dynamic online metric learning approach achieved the best overall performance across all evaluation metrics. A matching success rate of 93.74% was obtained, corresponding to an improvement of 10.47 percentage points relative to Euclidean distance matching. In addition, the average driver empty-travel distance was reduced to 512.83 m, representing a reduction of more than 42%. The average user waiting time was shortened from 28.64 s to 11.42 s, while the total global matching cost decreased substantially to 0.518. These results indicate that the continuously updated metric matrix is capable of dynamically adjusting feature weights in response to changes in traffic conditions and supply–demand distributions. As a consequence, spatiotemporal variations in shared mobility environments can be more effectively accommodated, leading to substantial improvements in dispatching efficiency and overall service quality of the platform.

## 3.5 Ablation study of core modules

To systematically evaluate the necessity and effectiveness of the proposed methodological components, an ablation study was conducted by progressively removing individual core modules from the complete framework. Through this experimental design, the contribution of each module to overall system performance could be quantitatively assessed.

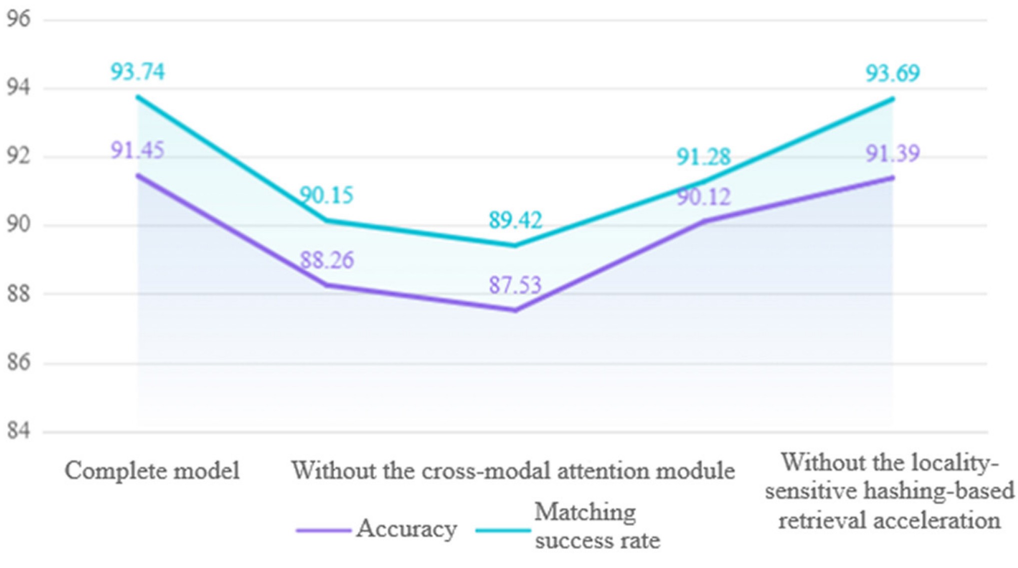


**Fig. 3.** Ablation results of core modules in terms of prediction accuracy and matching success rate

The quantitative results presented in Figure 3 and Table 3 provide clear evidence of the functional role and contribution of each core module. When conditional interaction entropy was removed, prediction accuracy decreased by 3.19 percentage points, while the matching success rate declined by 3.59 percentage points. Simultaneously, the total global matching cost increased from 0.518 to 0.643. These results indicate that conditional interaction entropy constitutes a critical component for capturing temporal behavioral dependencies and improving the accuracy of travel demand estimation. The largest performance degradation was observed after removal of the cross-modal attention module. Under this configuration, prediction accuracy decreased to 87.53%, while the matching success rate dropped to 89.42%. These findings demonstrate that adaptive fusion of spatial trajectory features and interaction behavior features represents the principal mechanism through which effective multimodal behavior modeling is achieved.

When the Cholesky-based positive-definite constraint was disabled, the stability of metric matrix optimization deteriorated slightly, resulting in a marginal decline across performance metrics. However, the overall impact remained limited. In contrast, removal of the locality-sensitive hashing-based retrieval acceleration module produced almost no change in prediction accuracy or matching success rate. However, the end-to-end latency increased dramatically from 4.96 ms to 28.75 ms. The findings indicate that locality-sensitive hashing-based retrieval acceleration has a negligible influence on modeling accuracy but plays a decisive role in ensuring low-latency operation of the system. Taken together, the ablation results demonstrate that the proposed modules perform complementary yet distinct functions within the overall architecture. Through their coordinated operation, a unified framework characterized by high prediction accuracy, low latency, and strong robustness is achieved, thereby confirming the necessity of all core design components.

**Table 3.** Ablation study results of core modules in terms of accuracy, matching performance, and system efficiency

| Experimental Configuration | Accuracy (%) | Matching Success Rate (%) | End-to-End Latency (ms) | Total Global Matching Cost |
|---|---|---|---|---|
| Complete model | 91.45 | 93.74 | 4.96 | 0.518 |
| Without conditional interaction entropy | 88.26 | 90.15 | 4.92 | 0.643 |
| Without the cross-modal attention module | 87.53 | 89.42 | 4.89 | 0.671 |
| Without the Cholesky-based positive-definite constraint | 90.12 | 91.28 | 5.01 | 0.585 |
| Without locality-sensitive hashing-based retrieval acceleration | 91.39 | 93.69 | 28.75 | 0.521 |

## 4 CONCLUSION

To address the challenges posed by highly dynamic user demand, resource-constrained mobile terminals, and the limited adaptability of conventional supply–demand matching mechanisms in sharing economy environments, an integrated optimization framework incorporating mobile terminal sensing, behavior modeling, and edge-enabled supply–demand matching was developed. Furthermore, an end–edge–cloud collaborative closed-loop feedback mechanism was introduced to enable continuous system optimization. A two-dimensional interaction entropy quantification framework based on mobile interaction entropy and conditional interaction entropy was established to characterize user behavioral irregularity and temporal behavioral dependencies. Through this design, the limitations of conventional travel demand identification approaches that rely exclusively on trajectory information were effectively overcome. A lightweight dual-stream attention network was subsequently constructed to facilitate efficient dual-modal feature fusion between spatial trajectories and terminal interaction behaviors. As a result, travel demand prediction accuracy was substantially improved while maintaining strict control over computational complexity and inference overhead on resource-constrained mobile terminals. A dynamic contextual matching criterion was established through online metric learning at edge nodes. Real-time adaptive updating of the metric matrix was achieved through positive-definite constraints. In addition, a locality-sensitive hashing-based algorithm adapted to the dynamic metric was incorporated to substantially reduce retrieval latency, thereby enabling millisecond-level real-time dispatching. Experimental evaluations conducted across multiple real-world scenarios demonstrated that the proposed framework significantly improved travel demand prediction accuracy and supply–demand matching success rates. Simultaneously, reductions in driver empty-travel distance, user waiting time, mobile-terminal energy consumption, and communication overhead were achieved. Moreover, superior long-term operational stability was observed when compared with conventional dispatching strategies.

The proposed framework advances the theoretical foundations of dynamic user behavior quantification and intelligent supply–demand matching in mobile shared-service environments. The limitations associated with traditional static feature modeling and fixed-metric matching approaches are effectively addressed. The resulting lightweight, low-latency, and highly adaptive end–edge–cloud collaborative scheduling paradigm provides a practical foundation for intelligent dispatching optimization in shared mobility platforms. In addition, the proposed framework can be readily extended to a broad range of mobile sharing economy services, including on-demand urban delivery and local service platforms.

## 5 REFERENCES


[1] C. B. Chandrakala, P. Somarajan, S. Jadhav, and A. Kapoor, "Empowering safety-conscious women travelers: Examining the benefits of electronic word of mouth and mobile travel assistant," *International Journal of Interactive Mobile Technologies*, vol. 18, no. 5, pp. 112–134, 2024. https://doi.org/10.3991/ijim.v18i05.43573

[2] H. Gözlükaya and G. D. O. Ertekin, "A systematic review and comparative policy analysis of sharing cities and their urban governance," *Challenges in Sustainability*, vol. 13, no. 4, pp. 587–608, 2025. https://doi.org/10.56578/cis130409

[3] S. Al-Masaeed *et al.*, "Factors affecting consumers' intention to use mobile ride hailing services in developing countries," *International Journal of Interactive Mobile Technologies (iJIM)*, vol. 16, no. 11, pp. 207–223, 2022. https://doi.org/10.3991/ijim.v16i11.30579

[4] R. Sumitkumar and A. S. Al-Sumaiti, “Shared autonomous electric vehicle: Towards social economy of energy and mobility from power-transportation nexus perspective,” *Renewable and Sustainable Energy Reviews*, vol. 197, p. 114381, 2024. https://doi.org/10.1016/j.rser.2024.114381

[5] J. N. Rosa, A. R. Pereira, F. Santoro, J. Gordijn, and M. Mira da Silva, “Modeling Waymo’s shared autonomous vehicle service in Phoenix using e3value,” *International Journal of Transport Development and Integration*, vol. 7, no. 3, pp. 153–165, 2023. https://doi.org/10.18280/ijtdi.070301

[6] A. K. Sahu, M. Z. Khan, and P. Gupta, “Instant food on your table: The role of logistics service quality dimensions in the adoption of instant online food delivery services,” *Transportation Research Part E: Logistics and Transportation Review*, vol. 200, p. 104205, 2025. https://doi.org/10.1016/j.tre.2025.104205

[7] S. Velmurugan, M. Prakash, S. Neelakandan, and A. Radhakrishnan, “Provably secure data selective sharing scheme with cloud-based decentralized trust management systems,” *Journal of Cloud Computing*, vol. 13, no. 1, p. 86, 2024. https://doi.org/10.1186/s13677-024-00634-8

[8] L. Yang, W. Zou, J. Wang, and Z. Tang, “Edgeshare: A blockchain-based edge data-sharing framework for industrial Internet of Things,” *Neurocomputing*, vol. 485, pp. 219–232, 2022. https://doi.org/10.1016/j.neucom.2021.01.147

[9] Y. Ma, X. Guan, J. Cao, and H. Wu, “A multi-stage fusion network for transportation mode identification with varied scale representation of GPS trajectories,” *Transportation Research Part C: Emerging Technologies*, vol. 150, p. 104088, 2023. https://doi.org/10.1016/j.trc.2023.104088

[10] M. Leodolter, C. Plant, and N. Brändle, “Rotation invariant GPS trajectory mining,” *GeoInformatica*, vol. 28, no. 1, pp. 89–115, 2023. https://doi.org/10.1007/s10707-023-00495-4

[11] J. Zhou, M. Du, and A. Chen, “Multimodal urban transportation network capacity model considering intermodal transportation,” *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2676, no. 9, pp. 357–373, 2022. https://doi.org/10.1177/03611981221086931

[12] H. Abkarian, H. S. Mahmassani, and M. Hyland, “Modeling the mixed-service fleet problem of shared-use autonomous mobility systems for on-demand Ridesourcing and carsharing with reservations,” *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2676, no. 8, pp. 363–375, 2022. https://doi.org/10.1177/03611981221083617

[13] F. Jin, Y. Cheng, X. Li, and Y. J. Hu, “Connecting the last mile: The impact of dockless bike-sharing on public transportation,” *Production and Operations Management*, vol. 34, no. 12, pp. 3904–3919, 2024. https://doi.org/10.1177/10591478231224953

[14] S. R. Leat, K. E. Ravi, A. Nordberg, and R. Voth Schrag, “Exploring the feasibility of shared mobility services for reducing transportation disadvantage among survivors of intimate partner violence,” *Journal of Transport & Health*, vol. 27, p. 101517, 2022. https://doi.org/10.1016/j.jth.2022.101517

[15] R. Mucci and G. D. Erhardt, “Understanding ride-hailing sharing and matching in Chicago using travel time, cost, and choice models,” *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2678, no. 2, pp. 293–306, 2023. https://doi.org/10.1177/03611981231173636

[16] Z. El Khaled, W. Ajib, and H. Mcheick, “Log distance path loss model: Application and improvement for sub 5 GHz rural fixed wireless networks,” *IEEE Access*, vol. 10, pp. 52020–52029, 2022. https://doi.org/10.1109/access.2022.3166895

[17] Z. Liu, H. Deng, X. Zhu, and L. Li, "Distance constrained sparse representation approach for scene based nonuniformity correction in infrared imaging," *IEEE Access*, vol. 12, pp. 141116–141129, 2024. https://doi.org/10.1109/access.2024.3462820

[18] S. Y. Park, H. Choi, Y. Boztuğ, and H. Moon, "Forecasting transition of personal travel behavior in a sharing economy: Evidence from consumer preferences of travel modes," *Journal of Forecasting*, vol. 44, no. 4, pp. 1563–1577, 2025. https://doi.org/10.1002/for.3255

[19] C. He, Z. Zhang, Y. Cai, H. Wang, L. Chen, and F. Huang, "Perceptual data importance and freshness aware transmission in millimeter wave vehicular networks," *Vehicular Communications*, vol. 54, p. 100939, 2025. https://doi.org/10.1016/j.vehcom.2025.100939

## 6 AUTHOR

**Zhaohe Ma** is with the School of Finance and Accounting, Henan Logistics Vocational College, Zhengzhou 453514, China (E-mail: hedy_mazhaohe@163.com).