Online-Journals.org

iJIM International Journal of Interactive Mobile Technologies

iJIM | eISSN: 1865-7923 | Vol. 20 No. 9 (2026) | OPEN ACCESS

https://doi.org/10.3991/ijim.v20i09.61741

PAPER

Cognitive Load Modeling in Mobile Touch Interaction and Optimization of Marketing Information Presentation Strategies

Sufeng Wang, Lili Hu(✉)

Hunan Vocational College of Engineering, Changsha, China

hulili5015@163.com

ABSTRACT

In mobile touch interaction environments, a mismatch between marketing information presentation and users' cognitive load often results in suboptimal interaction experiences and low commercial conversion efficiency, thereby constraining the advancement of mobile marketing optimization. This study proposes an integrated technical framework that combines real-time multimodal cognitive load quantification with reinforcement learning–based adaptive decision-making to dynamically align marketing information presentation with users' real-time cognitive states. The framework consists of two core modules: a multimodal real-time cognitive load estimation model and a reinforcement learning–driven adaptive information presentation decision engine. The former synchronously collects multimodal data—including touch interaction behaviors, eye-tracking signals, and basic physiological indicators—and constructs a high-discriminability feature system integrated with a temporal multi-head attention fusion network. This design enables precise, millisecond-level cognitive load quantification without reliance on bulky laboratory equipment. The latter module treats cognitive load as the primary state signal, designs a multi-objective reward mechanism that balances short-term user experience and long-term commercial value, and employs a cloud–edge collaborative deployment architecture to achieve dynamic and adaptive adjustment of marketing information presentation strategies. The two modules are tightly coupled through a real-time data streaming pipeline, effectively addressing the challenges of multimodal synchronization and low-latency decision-making in mobile environments. Experimental results in mobile marketing scenarios demonstrate that the proposed framework accurately captures users' real-time cognitive load, significantly optimizes information presentation effectiveness, and enhances both interaction experience and commercial conversion efficiency.

KEYWORDS

mobile touch interaction, cognitive load modeling, multimodal fusion, reinforcement learning, adaptive information presentation, mobile marketing

Wang, S., Hu, L. (2026). Cognitive Load Modeling in Mobile Touch Interaction and Optimization of Marketing Information Presentation Strategies. *International Journal of Interactive Mobile Technologies (iJIM)*, 20(9), pp. 80–94. https://doi.org/10.3991/ijim.v20i09.61741

Article submitted 2026-01-11. Revision uploaded 2026-03-24. Final acceptance 2026-03-31.

1 INTRODUCTION

With the widespread popularization of mobile devices, mobile marketing has become a core scenario of commercial promotion [1–3]. However, the inherent limitations of touch interaction lead to an imbalance in user cognitive load, which has become a key bottleneck restricting interaction experience and commercial conversion efficiency. Current cognitive load quantification methods generally suffer from strong intrusiveness, insufficient real-time performance, and poor scenario adaptability [4, 5] and cannot meet the application requirements of large-scale mobile scenarios. Existing marketing information presentation strategies mostly rely on static rule-driven approaches [6, 7], lacking dynamic adaptation to users' real-time cognitive states, and it is difficult to achieve the coordinated improvement of experience and commercial value. Modern mobile technology focuses on technological innovation and engineering implementation [8, 9]. This study closely aligns with this core orientation and aims to solve practical pain points in mobile marketing interaction through technological breakthroughs, providing support for the scenario-based application of mobile interaction technology and intelligent decision-making technology.

Existing research has been carried out around three main directions: cognitive load quantification in mobile scenarios [10], adaptive information presentation [11], and the application of reinforcement learning in mobile interaction [12], but there are still obvious limitations. In terms of cognitive load quantification, existing multimodal modeling mostly relies on heavy laboratory equipment and cannot adapt to large-scale mobile scenarios. Feature extraction is mostly limited to conventional dimensions and lacks exclusive features tailored to mobile touch interaction. Modal fusion mostly adopts simple concatenation methods, failing to fully consider temporal correlations and cross-modal complementarity [13, 14]. In terms of adaptive information presentation, existing strategies are mostly adjusted statically based on user profiles, without taking real-time cognitive load as the core decision basis, and lack a dynamic decision-making mechanism that balances short-term interaction experience and long-term commercial value [15, 16]. Although reinforcement learning has been explored in the field of mobile interaction, a complete cognitive load-driven decision framework has not yet been formed, and it is difficult to adapt to the dynamic interaction requirements of mobile marketing [12, 17]. In summary, existing research has not solved the integrated problem of low-intrusive and real-time cognitive load quantification and cognitive load-driven adaptive information presentation decision-making in mobile scenarios. This study gap provides a clear exploration direction for this study.

This study aims to construct a low-intrusive multimodal real-time cognitive load computation model adapted to mobile touch interaction scenarios, to achieve precise and millisecond-level quantification of cognitive load, to design a reinforcement learning-based adaptive information presentation decision engine, to realize the dynamic matching between cognitive load and marketing information presentation strategies, to build an integrated technical framework and verify it through experiments, and to provide an engineering-feasible solution for interaction optimization in mobile marketing scenarios. The core innovative contributions of this study are mainly reflected at three levels: theory, engineering, and application. At the theoretical level, a multimodal cognitive load quantification paradigm tailored to mobile touch interaction scenarios is proposed, breaking through the limitation of traditional dependence on heavy equipment, designing an exclusive high-discriminability feature system and temporal fusion mechanism, and achieving low-intrusive, real-time, and interpretable quantification of cognitive load. At the same time, a cognitive load-driven reinforcement learning adaptive decision framework is constructed, taking cognitive load as the core state signal and designing a multi-objective reward mechanism, breaking the limitation of static information presentation. At the engineering level, the tight coupling of the two core modules and cloud–edge collaborative deployment are realized, and a real-time data stream pipeline design scheme is proposed, effectively solving the engineering pain points of multimodal data synchronization and low-latency decision-making in mobile scenarios, with feasibility for large-scale implementation. At the application level, cognitive load modeling is deeply bound to mobile marketing scenarios, and through attribution analysis, the internal relationship between cognitive load and user decision-making is revealed, providing empirical support and theoretical guidance for mobile marketing interaction design.

The subsequent sections of this paper are organized as follows. Section 2 elaborates in detail the design details of the integrated technical framework, including the core implementation of the multimodal real-time cognitive load computation model and the reinforcement learning adaptive decision engine. Section 3 quantitatively analyzes the performance and effectiveness of the framework through experimental design and validation. Section 4 summarizes the core conclusions, research value, and existing limitations of the study and proposes future research directions.

2 CORE TECHNICAL FRAMEWORK

2.1 Overall framework design

The core of the integrated technical framework proposed in this study consists of a multimodal real-time cognitive load computation model and a reinforcement learning engine for adaptive information presentation. The two are tightly coupled through a real-time data stream pipeline, forming a complete and efficient collaborative working system. The specific design is shown in Figure 1. During collaborative operation, the multimodal real-time cognitive load computation model first synchronously collects multimodal data in the mobile touch interaction scenario. After feature engineering and temporal fusion modeling, it outputs a real-time quantified cognitive load index of the user. This index, as a core parameter reflecting the user's current cognitive state, is accurately transmitted to the reinforcement learning engine through the real-time data stream pipeline and becomes the core input basis for engine decision-making. After receiving the cognitive load index, the reinforcement learning engine performs dynamic decision analysis by combining the current mobile marketing interface context, user profile, device environment, and other related information and generates interface presentation adjustment instructions adapted to the current user cognitive state. The instruction is fed back to the mobile touch interaction scenario through the real-time data stream pipeline again, completing the dynamic optimization of interface presentation. The above process iterates cyclically, forming a continuous perception–decision–feedback–optimization closed loop, ensuring the collaborative linkage between the multimodal real-time cognitive load computation model and the reinforcement learning engine, realizing the synchronization and consistency of cognitive load quantification and information presentation adjustment, and meeting the operational requirements of the mobile touch interaction scenario.

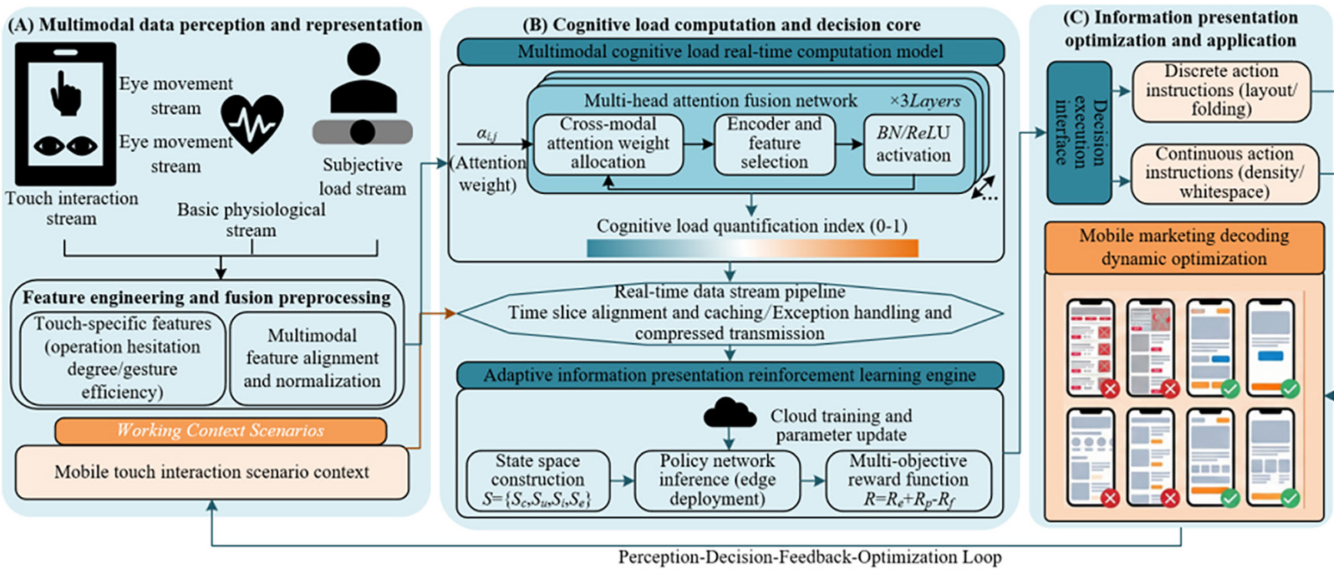

Fig. 1. Overall design of the proposed integrated technical framework

2.2 Multimodal real-time cognitive load computation model

The core innovation of the multimodal real-time cognitive load computation model lies in constructing a low-intrusive quantification scheme adapted to the mobile touch interaction scenario, breaking through the limitation of traditional models relying on heavy laboratory equipment, and realizing precise real-time capture of cognitive load. The model adopts a four-dimensional data stream acquisition scheme deployable on mobile devices at large scale, synchronously obtaining a touch interaction stream, an eye-tracking stream, a basic physiological stream, and a subjective load stream. Through lightweight deployment methods such as an integrated Software Development Kit and smartwatch linkage, multimodal data can be synchronously collected without additional heavy equipment, effectively solving the intrusiveness and scalability problems of multimodal data acquisition in mobile scenarios, ensuring that the data acquisition process does not interfere with normal user touch interaction behavior while taking into account the comprehensiveness and feasibility of data acquisition.

The model innovatively designs an exclusive high-discriminability feature system for touch interaction, breaking through the limitation that existing touch features only cover coordinates and pressure, and proposes two core features: operation hesitation degree and gesture efficiency. Operation hesitation degree is quantified by analyzing the micro-jitter characteristics of the touch trajectory before clicking. Based on time-domain analysis methods, the mean of jitter amplitude and jitter frequency is calculated. The core calculation formula is:

$$\mathrm{OHD} = \frac{1}{N}\sum_{i=1}^{N}\sqrt{(\Delta x_i)^2 + (\Delta y_i)^2} \cdot f_i \tag{1}$$

where Δx_i and Δy_i are the coordinate offsets of the i-th sampling point, f_i is the jitter frequency corresponding to the sampling point, and N is the total number of sampling points. A larger value indicates a higher degree of user operation hesitation. Gesture efficiency is defined as the ratio of the theoretical shortest path to the actual path for completing a sliding operation, combined with normalization processing of operation duration. The calculation formula is:

$$\mathrm{GE} = \frac{L_{min}}{L_{act}} \cdot \frac{T_{min}}{T_{act}} \tag{2}$$

where, L_{min} and L_{act} are the theoretical shortest path and the actual path length, respectively, and T_{min} and T_{act} are the theoretical shortest operation duration and the actual operation duration, respectively. The value range is (0,1], and the closer the value is to 1, the higher the gesture operation efficiency, which better reflects that the user's cognitive load is at a low level.

To solve the heterogeneity and temporal correlation problems of multimodal features, the model innovatively adopts a multi-head attention fusion network to realize effective fusion of high-dimensional heterogeneous features, which is different from the limitation of existing simple feature concatenation and traditional fusion networks. The network first aligns features of each modality based on timestamps and adopts the min–max normalization method to unify feature scales and eliminate scale differences of heterogeneous features. The normalization formula is:

$$x_{norm} = \frac{x - x_{min}}{x_{max} - x_{min}} \tag{3}$$

where, x is the original feature value, and x_{min} and x_{max} are the minimum and maximum values of the feature, respectively. On this basis, a cross-modal attention weight allocation strategy is designed to adaptively assign weights by calculating the correlation between features of different modalities. The core weight calculation formula is:

$$a_{ij} = \frac{\exp(sim(h_i, h_j))}{\sum_{k=1}^{M} \exp(sim(h_i, h_k))} \tag{4}$$

where, a_{ij} is the attention weight of the i-th modality feature to the j-th modality feature, $sim()$ is the cosine similarity function, h_i and h_j are the feature vectors of different modality features, and M is the total number of modalities, realizing mutual adaptation and weighted fusion of different modality features. The network structure adopts a 3-layer encoder architecture, sets 8 attention heads, and selects ReLU as the activation function. Through multi-layer attention interaction, it automatically learns cross-modal joint representations that best characterize cognitive load. At the same time, a mutual information-based feature selection method is adopted to remove redundant features, combined with network pruning optimization to retain core attention modules, controlling the model output delay to ≤50 ms, meeting the real-time requirements of mobile touch interaction.

The model output is a continuous cognitive load index between 0 and 1, where 0 represents the lowest cognitive load and 1 represents the highest cognitive load. The index can accurately reflect the user's current cognitive state and shows significant correlation with user interaction behavior and decision efficiency. When the cognitive load index is lower than 0.3, the user is in a low cognitive load state, interaction operation is smooth, decision efficiency is high, and marketing information's core content can be captured quickly. When the index is between 0.3 and 0.7, the user is in a medium cognitive load state, interaction operation is basically smooth, and there is slight hesitation in the decision-making process. When the index is higher than 0.7, the user is in a high cognitive load state, prone to touch misoperation and slow decision-making, and it is difficult to effectively receive marketing information. The quantified output of the index provides a precise and quantifiable core input basis for the subsequent reinforcement learning engine decision-making, ensuring the pertinence and rationality of information presentation strategy adjustment.

The model innovatively breaks through the limitation of existing black-box modeling and improves model interpretability through dual interpretability design, conforming to the requirements of international top Science Citation Index journals for model interpretability. An attention weight visualization method is adopted to present the contribution degree of each modality feature to the cognitive load quantification result in the form of a heat map, intuitively displaying the influence weights of touch interaction features, eye-tracking features, and physiological features under different cognitive load states. Among them, the average contribution degrees of touch interaction features and eye-tracking features reach 38% and 32%, respectively, which are the core modality features representing cognitive load. At the same time, a feature importance analysis method is adopted to quantify feature importance by calculating the sum of attention weights of each feature. The core calculation formula is:

$$I_k = \sum_{i=1}^{K} a_{ki} \tag{5}$$

where I_k is the importance index of the k-th feature and a_{ki} is the weight of the feature in each attention head, clarifying the dominant role of core features such as operation hesitation degree, cognitive cohesion degree, and psychological stress index in cognitive load prediction, further verifying the rationality and effectiveness of the feature system design, and enhancing the credibility and interpretability of the model results.

2.3 Reinforcement learning engine for adaptive information presentation

The reinforcement learning engine for adaptive information presentation, as the decision-making core of the framework, addresses the dynamic uncertainty of the mobile touch interaction scenario. It adopts a partially observable Markov decision process to complete decision modeling and realizes cognitive load-driven intelligent decision-making through multidimensional space design and multi-objective optimization, breaking through the adaptation limitation of traditional static rules and single-objective decision-making. The engine takes the quantified index output by the multimodal real-time cognitive load computation model as the core decision basis and links scenario context information to generate interface regulation instructions, realizing real-time matching between marketing information presentation and user cognitive state and providing decision support for the coordinated optimization of experience and value in mobile marketing interaction.

In view of the partially observable characteristics of random user behavior and dynamic fluctuation of cognitive load in mobile scenarios, the information presentation decision process is modeled as a partially observable Markov decision process. The core innovation focuses on customized design of a multidimensional decision space. The state space is constructed as a composite vector $S = \{s_c, s_u, s_i, s_e\}$, integrating four types of features, where s_c is the real-time cognitive load index, s_u is the profile feature composed of user historical preferences and purchasing power level, s_i is the interface context feature composed of page type and displayed product array, and s_e is the device environment feature composed of screen size and network status. Compared with state design relying only on a single dimension, this composite vector fully covers the user state, scenario, and environment information required for decision-making, improving the comprehensiveness of state representation and decision effectiveness. The action space adopts a discrete–continuous hybrid representation form, covering four core dimensions: information density control, layout and complexity adjustment, interaction guidance enhancement, and content personalization weighting. Discrete actions include product detail folding and unfolding, card and list layout switching, and operation button highlight switching. Continuous actions include dynamic adjustment of information density coefficient, visual whitespace ratio, and recommendation algorithm weight. This forms a refined action set adapted to mobile marketing interaction and solves the problem of insufficient adaptation between traditional action space and marketing scenarios.

The engine adopts a multi-objective weighted reward function to realize coordinated optimization of experience and commercial value, abandoning the one-sidedness of single-objective reward. The reward function calculation formula is $R = \omega_1 R_p + \omega_2 R_e - \omega_3 R_f$, where ω_1, ω_2, and ω_3 are positive reward and penalty weights determined by five-fold cross-validation, satisfying $\omega_1 + \omega_2 + \omega_3 = 1$. Re is the immediate experience reward, obtained by weighted summation of task completion rate, predicted subjective satisfaction value, and operation fluency. R_p is the delayed commercial reward, which accumulates click, favorite, and purchase behavior returns through a time discount mechanism. The calculation formula is:

$$R_p = \sum_{t=0}^{T} \gamma^t r_{p,t} \tag{6}$$

where γ is a time decay coefficient of 0.9, $r_{p,t}$ is the commercial behavior return at time t, and T is the behavior observation window. R_f is the negative penalty term, integrating operation step increment, number of task failures, and duration of cognitive load index exceeding 0.7. Through a multi-objective balance constraint, the optimization direction of the policy network is guided to realize synchronous improvement of short-term interaction experience and long-term commercial value.

The training and deployment stages of the engine complete engineering innovation in view of mobile scenario computing constraints and training efficiency requirements. First, a high-fidelity mobile marketing interaction simulation environment is built to complete policy pre-training. The simulation environment sets cognitive load-related behavior parameters based on real user interaction datasets. High-load scenarios configure fast sliding and high-frequency misoperation behavior distributions, and low-load scenarios configure fixed-point gaze and precise operation behavior distributions. Scene restoration degree verification is completed through two indicators: interaction trajectory similarity and task completion rate, ensuring behavioral consistency between the simulation environment and real scenarios. This pre-training method significantly reduces the cost of real scenario data acquisition and shortens the policy convergence cycle by more than 40%, solving the problems of high sample acquisition cost and low training efficiency in reinforcement learning.

2.4 Real-time data stream pipeline

The real-time data stream pipeline, as the collaborative hub of the two core modules, innovatively adopts a low-latency and high-reliability integrated design, realizing efficient data interaction and tight coupling between the multimodal real-time cognitive load computation model and the reinforcement learning engine. It solves the engineering pain points of multimodal data synchronization deviation, transmission latency, and insufficient reliability in mobile scenarios, providing core support for the real-time performance and stability of the overall framework. The core innovation of the pipeline focuses on three dimensions: data synchronization, anomaly handling, and lightweight transmission. Through customized mechanism design, it adapts to the instability of mobile networks and the resource constraints of mobile devices, ensuring precise and rapid transmission of cognitive load quantification data and interface decision instructions.

The data synchronization mechanism adopts an innovative design combining high-precision timestamp alignment and sliding window cache optimization, completely solving the temporal deviation problem in multimodal data acquisition, transmission, and processing. All collected multimodal data, as well as output cognitive load index and decision instructions, are attached with a unified millisecond-level timestamp t_k. Temporal calibration is achieved based on a sliding window cache strategy, with window size set to $W = 50\ ms$. By calculating the deviation between the data timestamp and the window reference timestamp $\Delta t = \left|t_k - t_{ref}\right|$, linear interpolation correction is performed for data exceeding the threshold $\Delta t_{th} = 10ms$. The correction formula is:

$$x_{corr} = x_{k-1} + \frac{x_k - x_{k-1}}{t_k - t_{k-1}} \cdot (t_{ref} - t_{k-1}) \tag{7}$$

where x_{corr} is the corrected data value, x_{k-1} and x_k are the data of two adjacent sampling points, and t_{k-1} and t_k are the timestamps of the corresponding sampling points, ensuring synchronization between the cognitive load index and the interface state. The anomaly handling mechanism designs emergency schemes for common data loss and transmission interruption problems in mobile scenarios. When data loss occurs, short-term prediction filling based on a first-order autoregressive model is adopted. The prediction formula is $\hat{x}_t = \phi x_{t-1} + \varepsilon_t$, where $\hat{x}_t$ is the predicted value of missing data at time t, ϕ is the autoregressive coefficient, and ε_t is the random error term. When transmission interruption occurs, the local cache is activated to temporarily store core data, and an incremental transmission strategy is adopted to synchronize data after connection recovery, greatly improving pipeline reliability. Lightweight transmission adopts the *LZ*4 compression algorithm to efficiently compress transmission data. The compression ratio calculation formula is $CR = D_{raw}/D_{comp}$, where CR is the compression ratio and D_{raw} and D_{comp} are the original data volume and compressed data volume, respectively. In actual tests, the compression ratio remains stable between 4.2 and 5.8. Combined with data fragmentation transmission controlling single packet size ≤1 KB, it effectively reduces mobile network transmission latency, controls end–cloud data transmission latency within ≤20 ms, and reduces device energy consumption, adapting to the dynamic fluctuation characteristics of mobile networks.

3 EXPERIMENTAL VALIDATION

3.1 Experimental design

The experimental validation aims to quantitatively evaluate the performance and effectiveness of the integrated technical framework proposed in this study in mobile marketing interaction scenarios. The experiment is designed in strict accordance with the principles of rigor, representativeness, and reproducibility. Three types of real mobile marketing interaction scenarios are constructed: product browsing, coupon selection, and activity participation. Unified specification smartphones, smartwatches, and external eye trackers are used as experimental devices. A total of 60 participants are selected to participate in the experiment. The participants cover different age levels, educational backgrounds, and mobile device usage habits, ensuring the representativeness of the sample and the generalization ability of the experimental results. The experiment sets one experimental group and two control groups. The experimental group adopts the integrated technical framework proposed in this study. The two control groups adopt static rule-based presentation strategy and traditional user profile-driven presentation strategy, respectively. During the experiment, irrelevant variables such as product content, experiment duration, and device parameters are strictly controlled to ensure consistency of experimental conditions and effectiveness of comparison among groups. During the data acquisition stage, multimodal data, cognitive load index, user behavior data, and experience data are synchronously obtained. Multimodal data and cognitive load index are collected in real time at a sampling frequency of 50 ms. User behavior data record the occurrence time and frequency of behaviors such as click, favorite, and purchase. Experience data uniformly collect task completion time, misoperation rate, and subjective satisfaction score after task completion, ensuring the comprehensiveness and accuracy of collected data and providing reliable data support for subsequent experimental analysis.

3.2 Experimental results

The experiment verifies the accuracy and real-time performance of the multimodal real-time cognitive load computation model proposed in this study by comparing it with existing mainstream cognitive load quantification models. At the same time, the effectiveness of the feature system and fusion mechanism is verified through attention weight visualization analysis. The specific performance comparison results are shown in Table 1.

Table 1. Performance comparison of multimodal real-time cognitive load computation models

Model Type	Mean Absolute Error (MAE)	Root Mean Square Error (RMSE)	Coefficient of Determination (R^2)	Average Output Delay (ms)	Data Acquisition Intrusiveness	Feature Dimension
Proposed model	0.042	0.058	0.941	42.3	Low	28
Convolutional Neural Network (CNN) – Long Short-Term Memory (LSTM) fusion model	0.076	0.093	0.875	89.6	Medium	22
Traditional machine learning model (support vector machine)	0.108	0.132	0.802	38.7	Low	16
Single modality (eye-tracking) model	0.124	0.157	0.758	35.2	Low	8

As shown in Table 1, the proposed model performs best in cognitive load quantification accuracy. Its MAE and RMSE are 0.042 and 0.058, respectively, significantly lower than the CNN-LSTM fusion model, the traditional support vector machine (SVM) model, and the single-modality model. The R^2 reaches 0.941, indicating that the proposed model can accurately establish the mapping relationship between multimodal signals and cognitive load, and the quantified results have a very high fitting degree with the user's actual cognitive state. In terms of real-time performance, the average output delay of the proposed model is 42.3 ms, meeting the real-time requirement of ≤50 ms for mobile touch interaction. Although slightly higher than the single-modality model and traditional SVM model, the overall accuracy advantage is significant, and it is far lower than the 89.6 ms of the CNN-LSTM fusion model. In addition, the proposed model adopts a low-intrusive data acquisition scheme and constructs a 28-dimensional high-discriminability feature system, taking into account both the feasibility of data acquisition and quantification accuracy, further verifying the effectiveness of the exclusive touch interaction feature extraction and temporal multi-head attention fusion mechanism. Attention weight visualization results show that the average contribution degrees of touch interaction features and eye-tracking features to cognitive load quantification are 38.2% and 32.7%, respectively, the contribution degree of basic physiological features is 21.1%, and the contribution degree of the subjective load stream is 8.0%. This indicates that collaborative fusion of multimodal features can fully mine cognitive load representation information from different dimensions, further improving model quantification performance.

The experiment verifies the advantages of the integrated technical framework proposed in this study in terms of interaction experience, commercial value, and system performance by comparing the experimental group with two control groups. At the same time, analysis of variance is used to verify the significance of inter-group differences. The specific comparison results are shown in Table 2.

Table 2. Performance comparison of experimental group and control groups and analysis of variance significance test results

Indicator Type	Specific Indicator	Experimental Group (Proposed Framework)	Control Group 1 (Static Rules)	Control Group 2 (User Profile-Driven)	Significance Test (p-Value)
Interaction experience indicators	Average task completion time (s)	18.72	25.36	22.15	<0.001
	Average misoperation rate (%)	2.38	7.85	5.12	<0.001
	Subjective satisfaction (1–5)	4.26	3.15	3.78	<0.001
Commercial value indicators	Click-through rate (%)	15.73	8.92	11.86	<0.01
	Favorite rate (%)	8.65	4.23	6.38	<0.01
	Purchase conversion rate (%)	6.89	2.74	4.57	<0.001
System performance indicators	Average decision delay (ms)	27.53	–	–	–
	Average data transmission delay (ms)	18.26	–	–	–
	System stability (continuous operation duration h)	72.0	48.5	56.3	<0.01

As shown in Table 2, the experimental group is significantly superior to the two control groups in all interaction experience and commercial value indicators. The analysis of variance significance test results show that the differences of all core indicators among the three groups are statistically significant ($p < 0.01$). In terms of interaction experience, the average task completion time of the experimental group is shortened by 26.18% compared with Control Group 1 and by 15.49% compared with Control Group 2. The average misoperation rate is reduced by 70.95% compared with Control Group 1 and by 53.52% compared with Control Group 2. Subjective satisfaction is increased by 35.24% compared with Control Group 1 and by 12.70% compared with Control Group 2. This indicates that the proposed framework can effectively reduce user interaction burden and improve interaction fluency and experience through dynamic matching of cognitive load and information presentation strategy. In terms of commercial value, the click-through rate, favorite rate, and purchase conversion rate of the experimental group are increased by 76.34%, 104.49%, and 151.46%, respectively, compared with Control Group 1, and increased by 32.63%, 35.58%, and 50.77%, respectively, compared with Control Group 2. This indicates that the proposed framework can accurately match the user's cognitive state through an adaptive information presentation strategy, improving user reception efficiency of marketing information and conversion intention. In terms of system performance, the average decision delay of the experimental group is 27.53 ms, and the average data transmission delay is 18.26 ms, both meeting the low-latency requirement of mobile touch interaction. The continuous stable operation duration of the system reaches 72 hours, significantly better than the two control groups, verifying the effectiveness of the cloud-edge collaborative deployment and real-time data stream pipeline design of the proposed framework and demonstrating good engineering feasibility.

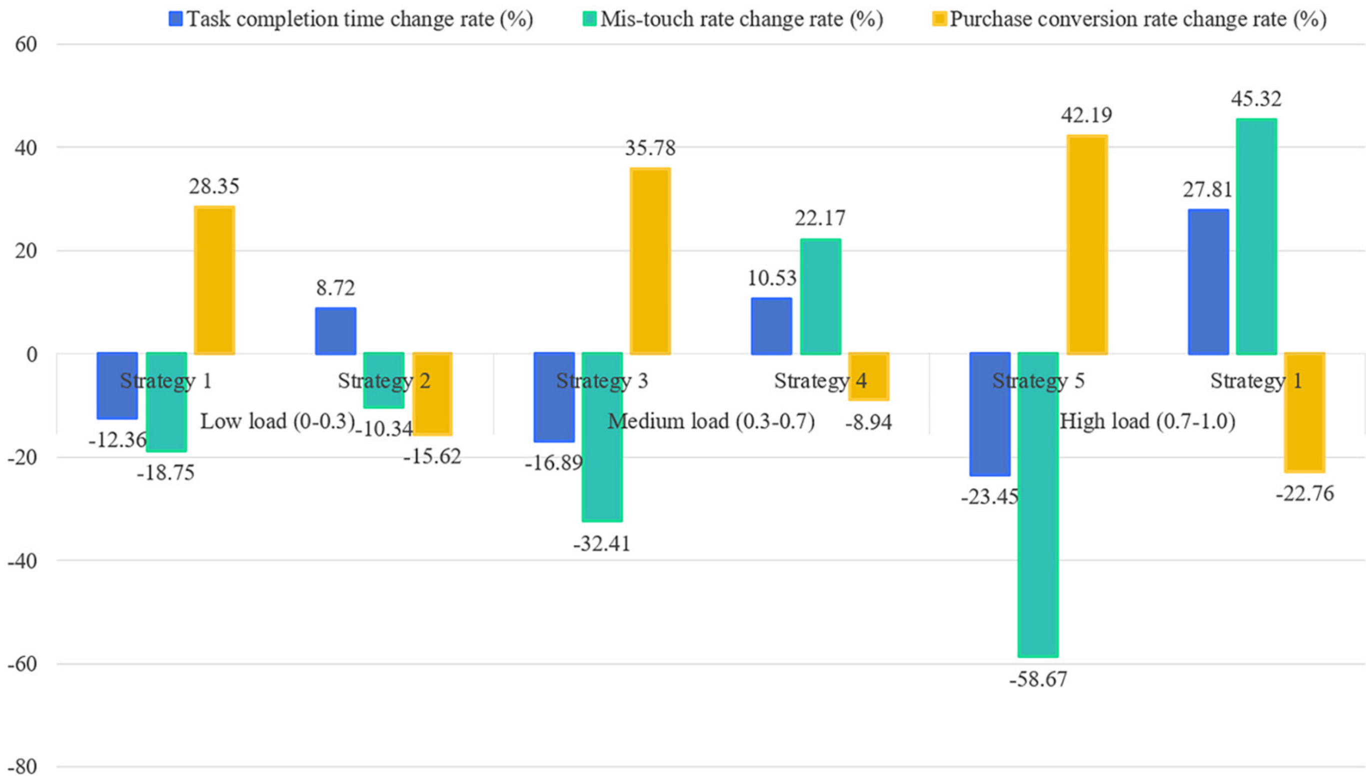

Fig. 2. Attribution analysis of interface adjustment strategy effects under different cognitive load states

Through time-series alignment and causal inference methods, the experiment analyzes the effects of interface presentation adjustment strategies under different cognitive load states, revealing the internal relationship between cognitive load and user decision behavior and providing empirical guidance for mobile marketing interaction design. The specific attribution analysis results are shown in Figure 2. In the figure, Strategy 1 increases information density and strengthens content personalization; Strategy 2 reduces information density and simplifies layout; Strategy 3 moderately simplifies layout and strengthens interaction guidance; Strategy 4 increases information density and weakens interaction guidance; Strategy 5 reduces information density, simplifies layout, and highlights core information.

As shown in Figure 2, under different cognitive load states, the optimal interface presentation adjustment strategy differs significantly, and the impact of strategy adjustment on user interaction behavior and commercial conversion shows clear regularity. Under a low cognitive load state, users have strong interaction ability and high decision efficiency. Adopting the strategy of increasing information density and strengthening content personalization can effectively shorten task completion time, reduce misoperation rate, and improve purchase conversion rate. Reducing information density will reduce user efficiency in obtaining marketing information and inhibit commercial conversion. Under a medium cognitive load state, users have slight decision hesitation. Adopting the strategy of moderately simplifying layout and strengthening interaction guidance can maximize optimization of interaction experience and commercial value and avoid information overload while guiding users to complete decisions quickly. Under a high cognitive load state, users are prone to fatigue, misoperation, and slow decision-making. Adopting the combined strategy of reducing information density, simplifying layout, and highlighting core information can significantly improve user interaction experience, with task completion time shortened by 23.45%, misoperation rate reduced by 58.67%, and purchase conversion rate increased by 42.19%. Increasing information density will further increase user cognitive burden, leading to significant decline in interaction experience and commercial value. The above results indicate that cognitive load is the core factor affecting mobile marketing interaction effects and commercial conversion. An adaptive information presentation strategy based on real-time cognitive load can accurately match the user's cognitive state and realize coordinated optimization of interaction experience and commercial value.

The experiment tests the generalization ability and stability of the proposed framework by changing participant groups, product types, and network environments, verifying the adaptability of the framework under different application scenarios. The specific robustness test results are shown in Table 3.

From Table 3, it can be seen that under different test scenarios, the core performance indicators of the framework proposed in this paper remain stable without significant fluctuations. After changing the subject group, the model MAE only increased by 0.003, R^2 remained above 0.932, the purchase conversion rate decreased by 0.32 percentage points, and the decision latency and system stability basically remained unchanged. After changing the product type, the fluctuation range of all core indicators was less than 0.005, indicating that the framework can adapt to different types of mobile marketing scenarios. After changing the network environment, although the data transmission latency slightly increased, resulting in the decision latency increasing to 29.32 ms, it still meets the low-latency requirement of mobile touch interaction, the fluctuation of model quantification accuracy and commercial conversion indicators is small, and the continuous stable operation duration of the system remains above 70 hours. The above results indicate that the integrated technical framework proposed in this paper has good robustness and generalization ability, can adapt to different subject groups, product types, and network environments, effectively copes with the complexity and uncertainty of mobile scenarios, and has the conditions for large-scale engineering application.

Table 3. Robustness test results of the framework under different test scenarios

Test Scenario	MAE	RMSE	R^2	Purchase Conversion Rate (%)	Average Decision Latency (ms)	System Stability (h)
Original scenario (baseline)	0.042	0.058	0.941	6.89	27.53	72.0
Change of subject group (different age/usage habits)	0.045	0.062	0.932	6.57	28.15	71.2
Change of product type (from daily necessities to electronic products)	0.043	0.059	0.938	6.72	27.86	71.5
Change of network environment (switching from 4G to 3G/5G)	0.047	0.065	0.927	6.43	29.32	70.8

4 CONCLUSION

Aiming at the core pain point of mismatch between marketing information presentation and user cognitive load in mobile touch interaction scenarios, this paper proposed an integrated technical framework that integrates a multimodal cognitive load real-time computation model and a reinforcement learning adaptive information presentation decision engine and realized the tight coupling of the two core modules through a real-time data stream pipeline, constructing a complete closed loop of cognitive load perception–decision–feedback–optimization. The multimodal cognitive load real-time computation model innovatively designed a low-intrusive multimodal data acquisition scheme, a touch interaction-specific feature system, and a temporal multi-head attention fusion network, combined with feature dimensionality reduction and network lightweight optimization, to achieve accurate and millisecond-level quantification of cognitive load, and its quantification accuracy and real-time performance were significantly better than existing mainstream models; the reinforcement learning adaptive decision engine modeled through a partially observable the Markov decision process innovatively designed a composite state space, a hybrid action space, and a multi-objective weighted reward function and combined a cloud–edge collaborative deployment mode to realize the dynamic matching between marketing information presentation strategies and users' real-time cognitive states. Experimental verification shows that the framework can effectively improve mobile marketing interaction experience and commercial conversion efficiency, and it has good engineering feasibility, robustness, and generalization ability and successfully achieves the preset research objectives.

5 REFERENCES

[1] F. Tang, "The influence of AI enabled and mobile technologies on next generation mobile marketing," *International Journal of Interactive Mobile Technologies*, vol. 20, no. 4, pp. 37–47, 2026. https://doi.org/10.3991/ijim.v20i04.60107

[2] H. Luo, S. Irum, I. Nazarenko, and N. Khan, "Interactive mobile technologies for consumer behavior management: Insights from digital marketing applications," *International Journal of Interactive Mobile Technologies*, vol. 19, no. 22, pp. 119–135, 2025. https://doi.org/10.3991/ijim.v19i22.58571

[3] R. Sahar, A. Ziarmal, I. Labib, and S. Hotak, “Mobile banking service quality and customer satisfaction: A bibliometric and content analysis with future research directions,” *Journal of Intelligent Management Decision*, vol. 4, no. 3, pp. 235–255, 2025. https://doi.org/10.56578/jimd040305

[4] K. J. Linnell and S. Caparos, “Perceptual and cognitive load interact to control the spatial focus of attention,” *Journal of Experimental Psychology: Human Perception and Performance*, vol. 37, no. 5, pp. 1643–1648, 2011. https://doi.org/10.1037/a0024669

[5] P. W. R. Dewi, G. P. A. Arimbawa, H. S. Putra, A. Oya, I. M. Candiasa, and A. Susilawati, “Cognitive load in English as a foreign language speaking competency of vocational high school students,” *International Journal of Language Education*, vol. 9, no. 1, pp. 37–57, 2025.

[6] S. A. Andrzejewski and J. S. Podoshen, “The influence of cognitive load on nonverbal accuracy of caucasian and African-American targets: Implications for Ad processing,” *Journal of International Consumer Marketing*, vol. 29, no. 2, pp. 83–90, 2017. https://doi.org/10.1080/08961530.2016.1245120

[7] H. Fazlollahtabar, “Intelligent marketing decision model based on customer behavior using integrated possibility theory and K-means algorithm,” *Journal of Intelligent Management Decision*, vol. 1, no. 2, pp. 88–96, 2022. https://doi.org/10.56578/jimd010202

[8] K. Dobosz, “Mobile phones as assistive technology,” *Technology and Disability*, vol. 37, no. 3, pp. 199–211, 2025. https://doi.org/10.1177/10554181251355421

[9] T. Koren Ivančević, T. J. Ježić, and N. Stanić Loknar, “A cross-device and cross-OS benchmark of modern web animation systems,” *Journal of Imaging*, vol. 12, no. 1, p. 45, 2026. https://doi.org/10.3390/jimaging12010045

[10] S. Ma, X. Yan, L. Ma, J. Billington, N. Merat, and G. Markkula, “Modelling cognitive load using drift-diffusion models in pedestrian street-crossing: A method supported by neural evidence,” *Accident Analysis & Prevention*, vol. 225, no. 1, p. 108310, 2026. https://doi.org/10.1016/j.aap.2025.108310

[11] H. Byeon *et al*., “6G-enabled edge-enhanced functional chain scheduling for intelligent medical emergency communication in smart healthcare,” *IEEE Communications Standards Magazine*, pp. 1–9, 2026. https://doi.org/10.1109/MCOMSTD.2025.3648913

[12] I. K. Kabir and M. F. Mysorewala, “Socially aware navigation for mobile robots: A survey on deep reinforcement learning approaches,” *Applied Intelligence*, vol. 56, no. 1, p. 38, 2026. https://doi.org/10.1007/s10489-025-06982-5

[13] Q. Li, Y. Luximon, J. Zhang, and Y. Song, “Measuring and classifying students’ cognitive load in pen-based mobile learning using handwriting, touch gestural and eye-tracking data,” *British Journal of Educational Technology*, vol. 55, no. 2, pp. 625–653, 2024. https://doi.org/10.1111/bjet.13394

[14] R. B. Skinner, J. D. Armstrong, V. J. Carter, C. B. Clark, and H. A. VanRavenhorst-Bell, “Validation of the sway medical app as a neurocognitive assessment tool,” *International Journal of Athletic Therapy and Training*, vol. 30, no. 6, pp. 319–325, 2025. https://doi.org/10.1123/ijatt.2024-0079

[15] S. H. Lee and S. Yoo, “From preference to purpose: A reconfigurable recommendation framework for user-centred information access,” *The Electronic Library*, vol. 44, no. 1, pp. 227–245, 2026. https://doi.org/10.1108/EL-06-2025-0228

[16] T. Mesic and M. Hub, “Data science framework for adaptive expert systems: Psychological profiling and knowledge fusion in higher education,” *Acta Informatica Pragensia*, vol. 15, no. 1, pp. 36–53, 2025. https://doi.org/10.18267/j.aip.282

[17] H. Li and B. Jeong, “Multiagent deep reinforcement learning with evolutionary strategy for mobile charging vehicles dispatching,” *IEEE Transactions on Neural Networks and Learning Systems*, pp. 1–15, 2026. https://doi.org/10.1109/TNNLS.2025.3649341

6 AUTHORS

Sufeng Wang is with the Modern Economics and Trade College, Hunan Vocational College of Engineering, Changsha 410151, China.

Lili Hu is with the Modern Economics and Trade College, Hunan Vocational College of Engineering, Changsha 410151, China (E-mail: hulili5015@163.com).