ingegneria sismica
International Journal of Earthquake Engineering
Anno XLIII - Num. 2 - 2026

Computational Modeling and Persuasive Design Optimization of User Behavior Based on Emotional Interaction on Digital Platforms

Yuchen Liu[1], *Zongyi Yu*[1] *and Kiesu Kim*[1,*]

[1] College of Fine Arts, Silla University, Busan 46958, Busan, Korea

SUMMARY: *This paper constructs a unified computing framework driven by emotional interaction, focusing on user behavior modeling and persuasive design optimization of digital platforms in real-time interaction scenarios. Based on clickstream, dwell time, scroll trajectory, touch pressure, facial cues and text feedback, a multimodal data set consisting of 4320 users and 18,600 interaction sessions is established. The model describes emotional fluctuations and behavior changes through time convolution, cross-modal alignment, state aggregation and transition estimation, and discriminates four types of states: participation, hesitation, resistance and acceptance. On this basis, the adaptive intervention module dynamically adjusted the information expression, interface density, feedback timing and recommendation strength by combining the state results. The experimental results show that the state recognition accuracy of the proposed framework reaches 93.1%, and the intervention matching accuracy reaches 90.4%, so that the compliance interaction rate is increased by 14.8%, and the average response delay is controlled within 86 ms. The results show that the framework can maintain stable state understanding, policy generation and front-end execution capabilities in continuous interaction scenarios, and has practical application capabilities.*

KEYWORDS: *Digital platform; Emotional interaction; User behavior modeling; Persuasive design*

1 Introduction

With the continuous evolution of affective computing, recommendation algorithms and human-computer interaction technology, digital platforms have changed from information display and functional service carriers to computing environments that can perceive user status, understand behavioral changes, and dynamically adjust interactive content. The clicks, pauses, swips, comments, voice and visual feedback in the platform continuously form a high-density behavior stream, which provides a computable data basis for user state recognition, interface response control and persuasive intervention generation. Ameen et al. studied personalized interaction and loyalty relationship in intelligent technology scenarios, and pointed out that individuals form a stable behavior judgment structure between convenience perception and privacy trade-off [1]. Sun et al. studied the continuous usage intention of artificial intelligence personal assistant after service failure, and proposed that the change of service experience will directly affect users 'subsequent interaction decisions [2].

In digital platforms for content distribution, service guidance and interface response, affective interactive computing is becoming an important intermediate layer connecting user

*kks@silla.ac.kr
https://doi.org/10.65102/is2026548 Published: 30 April, 2026

perception and system decision. Instead of only counting explicit click results, the platform starts to use temporal coding, state recognition and context fusion to identify fine-grained user states such as hesitation, resistance, acceptance and compliance. Ma et al. studied the negative response of users to the recommendation algorithm in the short video platform, and gave an explanation framework based on the press-strain-result link, indicating that there is a significant coupling relationship between the platform stimulus mode and the behavior feedback [3]. Benke et al. studied the difference of control level in emotion-aware chatbots, and proposed that the configuration of sense of control would affect the user's acceptance of the system interaction process [4].

With the popularity of mobile terminals, intelligent customer service and recommendation interfaces, the expression of emotional states is no longer limited to a single explicit feedback, but manifests as a continuous time trajectory and cross-modal response sequence. Tag et al. studied emotional trajectory recognition in smartphone usage scenarios and proposed that emotional changes should be grasping from the dynamic adjustment process in the real environment [5]. Jin et al. studied the friendly language mechanism in the recommendation chatbot and proposed that the language style would change the user's psychological feelings and subsequent selection tendencies [6]. Kim et al. studied the intervention mode of chatbots in behavior guidance scenarios, and proposed that computerized interaction can form a directional persuasion effect [7]. Marcolin et al. studied persuasive technology and social interaction in professional environment, and proposed that the embedded intervention structure can reshape the user participation path [8].

At the same time, persuasive design in digital platforms has gradually shifted from experience-driven interface debugging to an adaptive generation mechanism based on state estimation and feedback learning. The relationship among transparency, trust and willingness to accept provides a clearer calculation basis for the control of platform intervention intensity, the design of feedback rhythm and the adjustment of recommended content. Wanner et al. studied the influence of transparency and trust on the acceptance of intelligent systems, and proposed that interpretable presentation can enhance users' understanding of system decision-making process [9]. Choung et al. studied the relationship between AI trust and technology acceptance, and pointed out that trusted cognition has become the core variable in the deployment of intelligent interactive systems [10].

As shown in Fig. 1, the digital platform driven by emotional interaction is not a single interface response process, but a closed-loop computing link composed of data collection, emotion encoding, state inference, intervention decision and feedback update. The front-end continuously receives multi-source interactive signals such as click, pause, scroll, touch, text and expression, and completes time alignment, noise filtering and sequence segmentation. In the middle layer, the heterogeneous input is mapped into a unified emotional representation, and the user's current participation, hesitation, resistance or acceptance state is calculated by combining the behavior history and context information. The decision layer generates intervention actions according to state probability, adaptive threshold and policy constraints, and adjusts the information expression mode, interface density, feedback timing and recommendation strength. The end writes the interaction result, state offset and intervention effect back to the model for the next round of parameter update and policy correction. This structure enables the platform to maintain a stable correspondence between state recognition, behavioral response, and persuasive output during continuous interactions.

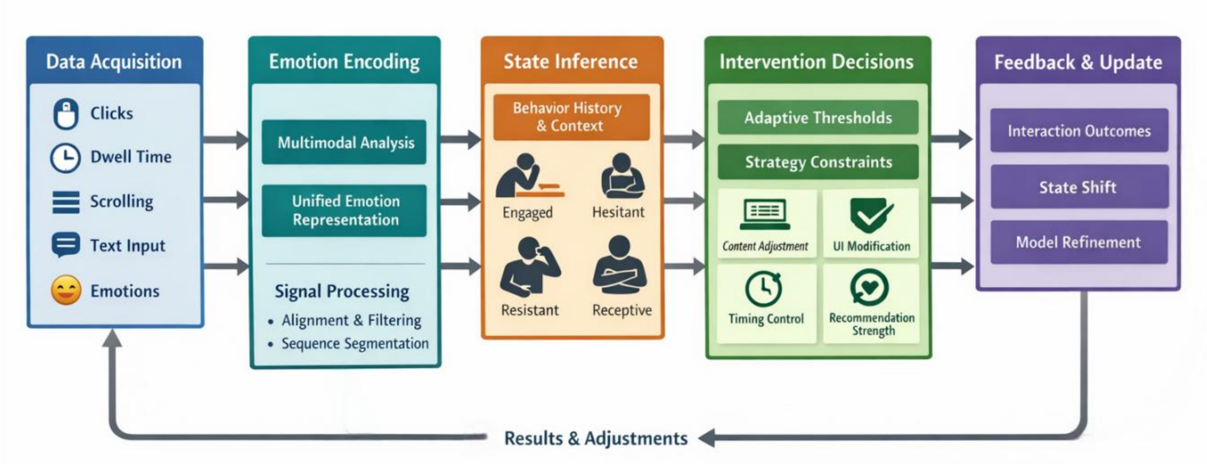

Figure 1: Closed-loop structure of user behavior computation and persuasive intervention on digital platform driven by emotional interaction

This thesis focuses on emotional interaction modeling and persuasive design optimization in digital platforms. The research content focuses on four aspects: multimodal emotional feature representation, user behavior state calculation, interface response discrimination and adaptive intervention generation, and tries to organize behavior understanding, strategy output and interaction feedback into a continuous closed loop under a unified computing framework. To establish a clear and complete technical support for subsequent experimental evaluation, system deployment and online verification.

2 Related Research

2.1 Research on Affective Interaction Computing in Digital Platform

Emotional interaction computing in digital platforms has been extended from single emotion recognition to the state representation and response control mechanism for continuous interaction processes. The core of emotional interaction is no longer limited to the extraction of surface emotion labels, but to the joint modeling of the coupling relationship between user cognitive fluctuations, trust formation, behavior transfer and interface feedback. The development of this direction is closely related to the iteration of conversation system, recommendation engine, online service interface and real-time behavior collection technology, so that the platform can form more fine-grained perception and adjustment capabilities in multiple rounds of interaction.

Bach et al. studied user trust in artificial intelligence systems and proposed that trust is not an isolated psychological variable, but a human-computer collaborative structure that runs through the whole process of perception, judgment and interactive execution [11]. Ruan and Mezei studied the artificial intelligence chatbot in the online shopping assistance scene, and proposed that under different product attributes, the system response mode will change the user satisfaction evaluation and service acceptance path [12]. Janson studied the anthropomorphic mechanism in service interfaces, and proposed that the cooperative relationship between communication style and personalized presentation would affect users' emotional attribution to interactive objects and their willingness to participate continuously [13].

Alagarsamy and Mehrolia studied the antecedences and behavioral consequences of trust in chatbots, and proposed that there is a computable conduction link between trust perception, interface expression and usage results [14]. Park et al. studied the role of anthropomorphic chatbots in counseling scenarios, and proposed that social harmony and social anxiety would jointly regulate user satisfaction and reuse tendency [15]. Ltifi studied the characteristics of semi-interpersonal relationship in chatbots, and proposed that users would form cognitive projections close to social relationships in continuous interaction, which would further affect compliance behavior, feedback intensity and acceptance rhythm in the platform [16].

At the technical implementation level, emotional interaction computing usually relies on text semantic coding, speech prosody analysis, facial representation extraction, touch behavior statistics and temporal context modeling. The cross-modal alignment mechanism is used to unify the heterogeneous signal scale, and the state update unit is used to capture the short-term offset between emotion and behavior. The feedback mapping module converts the recognition results into the basis of interface presentation, recommendation ranking and prompt strategy adjustment. Therefore, the significance of this direction in computer research is not only reflected in the improvement of recognition accuracy, but also reflected in the platform's ability to complete more stable interaction arrangement and intervention control according to the change of emotional state. These studies show that affective interaction computing has advanced from recognizing user emotion itself to modeling interaction signals, state changes and response effects, and provides a clear computational foundation for behavior discrimination, persuasive intervention and adaptive design in digital platforms.

2.2 Research on user behavior modeling and persuasive design

The research on user behavior modeling and persuasive design has shifted from static profiling and rule-triggering to collaborative mechanisms for state perception, intention inference, and intervention generation for continuous interaction processes. The clicks, pauses, swipes, replies, and jumps in the digital platform are no longer regarded as isolated logs, but organized as time-series signals that can be used to predict acceptance tendency, interface compliance, and feedback sensitivity. Carera-rivera et al. studied the development framework of adaptive user interface, and proposed that linkage mapping should be established between interface elements, context variables and interaction goals to support dynamic adjustment in the platform [17]. Farshidi et al. studied user intention modeling in conversational recommender systems, and proposed that intention recognition should be jointly processed with historical behaviors, conversational context and recommendation feedback, so as to enhance the continuity of policy output [18].

As shown in Fig. 2, the user behavior modeling and persuasive design in the digital platform is not a simple "recommendation after recognition" process, but a continuous computing link composed of behavior collection, feature coding, intention recognition, state calculation, policy output and feedback writeback. The front-end of the platform firstly records the click, stay, swipe, reply and jump out behaviors, and combines the session context to complete the temporal characteristics. Then, the middle layer identified the user's current intention through the coding results, and further estimated the state distribution of acceptance, hesitation and resistance. The decision layer generates intervention actions according to state probability, historical response and interface constraints, and adjusts the information expression, recommendation intensity and interaction rhythm. At the end, the behavior change and adoption results are written back to the model for subsequent policy modification. The link illustrates that persuasive design in digitization platforms already behaves as updatable, computable, and verifiable dynamic decision processes.

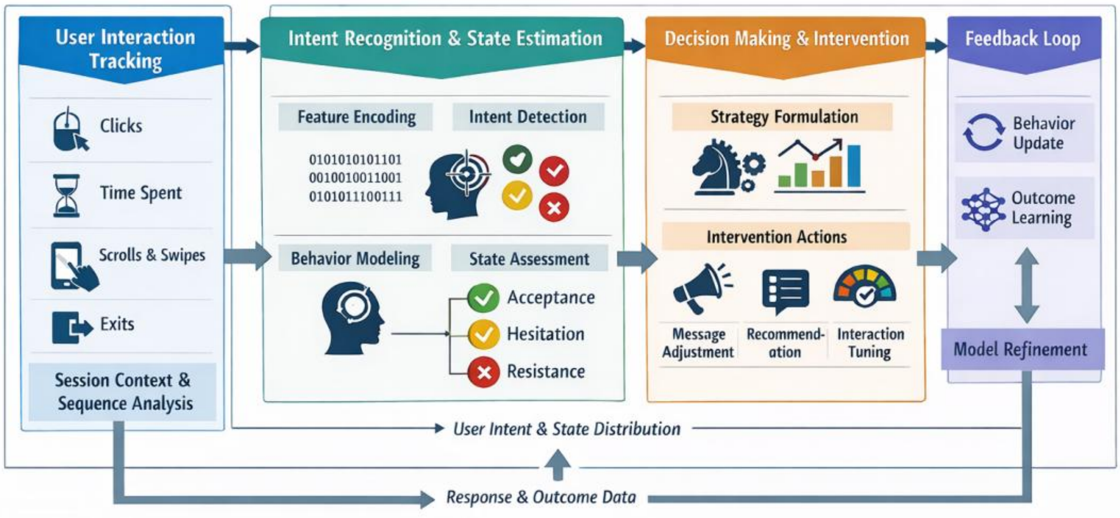

Figure 2: User behavior modeling and persuasive design computing link in the digitization platform

Ribeiro et al. studied the personalization method of self-expression virtual image based on context awareness, and proposed that platform identity expression is not an auxiliary display layer, but an important regulatory variable that affects user engagement and feedback willingness [19]. Ning et al. studied the transparency mechanism in the use of algorithm suggestions, and proposed that trust belief plays an intermediary role between explanatory presentation and suggestion adoption, which makes the display method of persuasive information into the computable range [20]. Engstrom et al. studied the usage modeling of online recommender systems and proposed that the recommendation usage behavior should be jointly described from three stages: comparison, selection and return visit, so as to improve the completeness of behavior interpretation [21]. He et al. studied the influence of recommendation system on user satisfaction and proposed that there is a stable regulation link between recommendation matching degree, perceived relevance and interactive experience [22].

In order to more clearly illustrate the focus of current research in user behavior modeling and persuasive design, Table 1 collates the relevant literature from four dimensions: research object, computational object, method path and main conclusion.

Table 1: Summary of research related to user behavior modeling and persuasive design

Author	Research Focus	Computational Object	Technology or Method	Main Conclusion
Carrera-Rivera et al. [17]	Adaptive interface	Interface elements and context	AdaptUI framework	Supports dynamic interface adjustment
Farshidi et al. [18]	Intent modeling	Conversational behavior and recommendation feedback	Intent recognition and conversational modeling	Enhances strategy continuity
Ribeiro et al. [19]	Personalized expression	Virtual avatars and contextual variables	Context-aware personalization	Strengthens engagement and feedback
Ning et al. [20]	Transparent recommendation	Recommendation information and trust beliefs	Transparency modeling	Influences recommendation adoption
Engström et al. [21]	Recommendation usage modeling	Comparison, selection, and revisit behavior	Usage-stage modeling	Improves the completeness of behavioral interpretation
He et al. [22]	Satisfaction evaluation	Recommendation matching and experience perception	Moderated mediation analysis	Forms a stable satisfaction chain

It can be seen that existing research has gradually integrated adaptive interface, intention modeling, identity expression, transparent recommendation, usage behavior analysis and satisfaction evaluation into the same technical link, but the direction of higher value does not lie in the local improvement of the accuracy of individual modules, but in the organization of emotional interaction, behavior calculation and persuasion intervention into a unified closed loop. At the computer implementation level, this closed loop usually relies on multi-modal feature coding, temporal state updating, policy constraint learning and online feedback correction, so that the platform can adjust the interface presentation, information wording, feedback timing and recommendation strength according to the user's current state, and continuously update parameter boundaries in subsequent interactions. Thus, the research on user behavior modeling and persuasive design has been advanced from empirical interface configuration to data-driven dynamic decision-making process, and provides a direct foundation for the subsequent emotional interaction modeling and adaptive intervention generation in this paper. The platform deployment path is also more engineering verifiable.

3 Computational modeling of user behavior driven by emotional interaction

3.1 Multimodal emotional feature representation of digital platform

Emotional features in digital platforms do not come from a single modality, but are composed of behavior sequence, text semantics, touch rhythm, visual expression and conversational context. In order to support the subsequent state calculation and persuasion intervention, the platform needs to convert the heterogeneous interaction signals into the same scale, aligned and updatable representation vectors. Different from static portraits, the feature representation here emphasizes temporal continuity and response sensitivity, so the modeling process should not only preserve local fluctuations, but also capture the emotional accumulation in cross-wheel interactions.

On the behavioral side, click, pause and scroll operations are first exposed, and these three behavioral signals can directly reflect the attention allocation and interaction depth. In order to describe the degree of interaction activity within the window, this paper defines the behavior strength index, which is calculated as shown in Equation (1).

$$I_t = \alpha \ln(1 + c_t) + \beta \frac{d_t}{1 + d_t} + \gamma \frac{1}{N_t} \sum_{j=1}^{N_t} |v_{t,j}| \tag{1}$$

Here, I_t represents the behavior intensity in the t time window; c_t represents the number of clicks. d_t is the length of stay. $v_{t,j}$ denotes the velocity of the j scroll segment within the window; N_t represents the number of scroll fragments; α, β, γ denote the contribution weights of different operations. This formula compresses decentralized logs into a unified representation of activity, which can be used to identify high-engagement browsing and low-engagement saccades.

It is still difficult to describe the subtle changes of users in the stages of hesitation, acceptance or resistance by only explicit logs, so touch dynamics is introduced into the emotion representation process. Contact duration, pressure fluctuation, and trajectory curvature usually maintain corresponding relationships with tension, caution, and operational certainty, and their encoded forms are shown in Equation (2).

$$T_t = \frac{1}{K}\sum_{k=1}^{K}(\xi_1 p_k + \xi_2 \tau_k + \xi_3 \kappa_k) + \xi_4 \mathrm{Var}(p_1, \dots, p_K) \tag{2}$$

Here, T_t represents the dynamic representation of the touch mode at time t. p_k represents the k touch pressure; Let τ_k denote the duration; κ_k represents the curvature of the trajectory; K is the total number of touch events in the window. Let ξ_1 to ξ_4 denote the parameter coefficients; $\mathrm{Var}(\cdot)$ represents the pressure series variance. This formula organizes contact differences into statistical representations to improve the ability to distinguish between hesitant operations and quick confirmation behaviors.

Text feedback is a key modality in the emotional interaction of the platform. Comment content, input length, emotional word distribution, and negation structure not only carry subjective attitudes, but also reflect how users accept platform prompts. In order to extract the emotional semantics in the text, this paper adopts the context-gated coding, whose expression is shown in Equation (3).

$$E_t = \frac{\sum_{i=1}^{L} \sigma(r_i)\, q_i}{\sum_{i=1}^{L} \sigma(r_i)} \tag{3}$$

where, E_t represents text sentiment vector; q_i represents the context representation of the i lemma; r_i is the response coefficient of emotional words; L denotes the total number of tokens; Let $\sigma(\cdot)$ denote the gating function. This formula is used to highlight semantic components with attitude pointing and suppress the dilution of sentiment representation by invalid descriptions.

On the visual side, subtle facial expressions can supplement cues not directly given by text and behavior logs. Eyebrow retraction, eyelid opening and mouth Angle displacement are often synchronized with anticipation, hesitation and rejection. To this end, the visual emotion amplitude feature is constructed in this paper, and its form is shown in Equation (4).

$$F_t = \eta_1\|u_t\|_2 + \eta_2\|m_t\|_2 + \eta_3\|n_t\|_2 \tag{4}$$

where, F_t represents visual emotion amplitude; u_t, m_t and n_t represent the deformation vectors of the eyebrow, eye and mouth key regions, respectively. $\|\cdot\|_2$ denotes the two-norm; η_1, η_2, and η_3 denote the region weights. This equation maps local expression changes into a unified visual emotion quantity, which enhances the consistency of cross-modal state estimation.

Since the intensity of emotion reflection by different modalities is not constant, direct stitching is easy to amplify the interference of noise modalities, so it is necessary to establish a cross-modal alignment mechanism. In this paper, similarity constraints are used to calculate the weights between behavior, text and visual modalities, which are expressed as in Equation (5).

$$a_{ij} = \frac{\exp\big(\mu \cdot \cos(z_i, z_j)\big)}{\sum_{m \neq i} \exp(\mu \cdot \cos(z_i, z_m))} \tag{5}$$

Here, a_{ij} represents the alignment weight between modality i and modality j. z_i and z_j denote the two types of modal features; $\cos(\cdot,\cdot)$ is the cosine similarity; Let μ denote the scaling factor. This formula is used to measure the degree of consistency between modalities, so that the subsequent fusion process can dynamically allocate the proportion of information according to the correlation.

Once the alignment relationship is obtained, the platform needs to compress the multi-source features into a unified representation that can be invoked by the state discriminator. Considering that the modal quality will fluctuate with the scene, this paper introduces a fusion expression of the reliability constraint, which is calculated as shown in Equation (6).

$$H_t = \rho\left(\omega_t^{(b)} B_t + \omega_t^{(c)} E_t + \omega_t^{(v)} Q_t\right) \tag{6}$$

Here, H_t represents the fused sentiment representation at time t. $\omega_t^{(b)}$, $\omega_t^{(c)}$, $\omega_t^{(v)}$ denote the reliability coefficients of behavioral, textual and visual modalities, respectively. B_t, E_t and Q_t denote the corresponding mode vectors, respectively. Let $\rho(\cdot)$ denote the nonlinear mapping function. This formula can keep the representation stable in the modal imbalance scenario and avoid a single high-noise input dominating the result.

In order to retain the inertia in the interaction, the representation also needs to be updated with the historical state. Based on this feature, this paper constructs a context-constrained sentiment update formula, as shown in Equation (7).

$$S_t = \lambda_t \odot H_t + (1 - \lambda_t) \odot S_{t-1}, \quad \lambda_t = \sigma(M_s[H_t; S_{t-1}] + b_s) \tag{7}$$

where S_t represents the state vector; S_{t-1} denotes the state at the previous time. Let λ_t denote the update gate value of the window; $\odot$ for element-wise multiplication; M_s represents the state mapping matrix; b_s represents the bias term; $[H_t; S_{t-1}]$ denotes the concatenation of vectors. The formula is used to balance the immediate response and historical inertia, so that the multimodal features can not only reflect the stimulus effect, but also maintain the continuity of the behavior trajectory, and provide the feature basis for the subsequent behavior state calculation and intervention generation, and enhance the performance of online scene adaptation.

In conclusion, multi-modal emotional feature representation is not a simple concatenation of behavior, text, vision and touch information, but a computational process of differentiated coding, correlation constraints and state compression under a unified time scale. After this process, the emotional cues in the digital platform are transformed into continuous, comparable and updatable representation results, which not only retain the subtle fluctuations in the user's immediate interaction, but also maintain the state continuation in the cross-round session, which establishes a stable input basis for the subsequent behavior state calculation, response discrimination and persuasive intervention generation.

3.2 User behavior state calculation and response Discrimination based on emotional interaction

After obtaining the multi-modal emotion representation, the task of user action state calculation is transformed into: according to the emotion vector formed in the continuous interaction, the behavior intensity and the conversation context, the four types of participation, hesitation, resistance and acceptance are jointly determined, and the response results are output for the persuasion module to call directly. The states in a digital platform are not static labels, but dynamic variables shaped by immediate stimuli, historical inertia and interface feedback. Therefore, the computation process needs to preserve both local fluctuations and cross-wheel continuation.

In order to reflect the synergistic effect of emotion, behavior and context in continuous interaction, this paper constructs the time decay state aggregation representation, which is calculated as shown in Equation (8).

$$S_t = \sum_{k=0}^{K-1} \alpha_{t-k} \left(H_{t-k} \odot \psi(I_{t-k})\right) + UC_t$$

$$\alpha_{t-k} = \frac{\exp(q^{\top} H_{t-k} - \delta k)}{\sum_{m=0}^{K-1} \exp(q^{\top} H_{t-m} - \delta m)} \tag{8}$$

where S_t represents the user behavior state vector at time t; H_{t-k} denotes the multimodal sentiment representation of the $t-k$ time window; I_{t-k} denotes the behavior strength in the corresponding window; Let $\psi(\cdot)$ denote the mapping function that scales the behavior strength; $\odot$ for element-wise multiplication; C_t represents the current session context vector. U represents the context mapping matrix; Let α_{t-k} denote the contribution weight of the history window to the current state.q denotes the state query vector; Let δ denote the time decay coefficient; K represents the number of history Windows participating in the aggregation. The key role of this formula is to simultaneously preserve the historical accumulation of emotion representation, the modulation effect of behavior intensity and the real-time influence of context conditions in the state calculation stage, so that the output results are more suitable for the subsequent state transition estimation and response discrimination.

After the state vector is formed, the platform also needs to determine the direction of the user's change between adjacent time slices. To describe this progressive relationship, the state transition probability is introduced in this paper, which is calculated as shown in Equation (9).

$$P_t^{i\to j} = \frac{\exp\left(u_j^{\top} S_t + v_{ij}\Delta_t\right)}{\sum_{m=1}^{4} \exp(u_m^{\top} S_t + v_{im}\Delta_t)} \tag{9}$$

where $P_t^{i\to j}$ denotes the probability that state i transitions to state j at time t. u_j represents the object state projection vector; v_{ij} is the transfer coefficient; Δ_t represents the sentiment offset of the current window with respect to the previous window. The equation takes instant state and transition inertia into account simultaneously, so that the model can distinguish different paths such as short-term hesitation, persistent resistance and stable acceptance.

As shown in Fig. 3, the response discrimination process consists of four parts: state input, transition estimation, response scoring, and output mapping. The front-end receives the emotional representation and behavior statistics, the middle layer completes the joint state coding and transition probability calculation, and the back-end generates four types of discriminative labels according to the scoring results and sends the results to the subsequent intervention module. In this way, the raw interaction log can be transformed into a compact representation of the state that can be directly used by the policy layer.

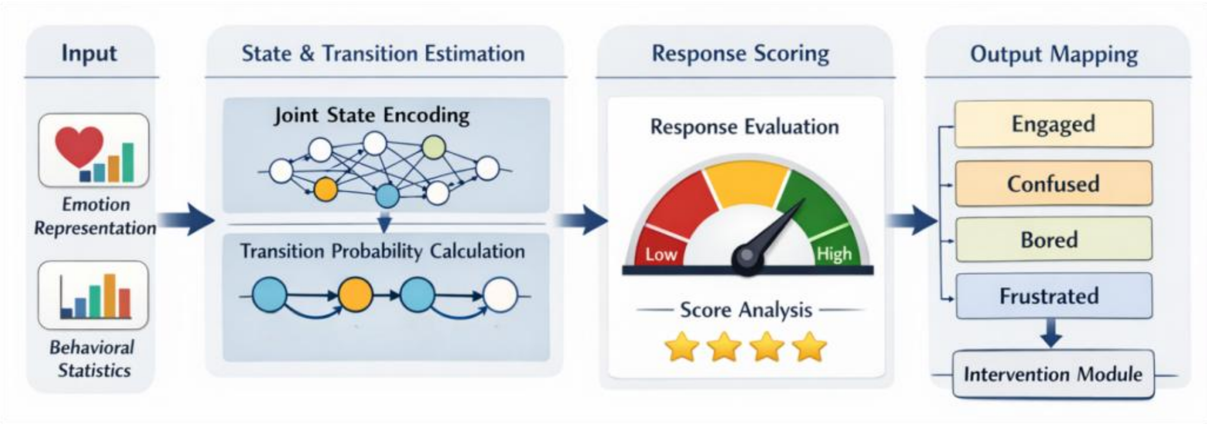

Figure 3: Process of user behavior state calculation and response discrimination based on emotional interaction

After obtaining the transition probabilities, the platform also needs to provide interpretable response scores to determine whether the current interaction is more suitable for maintenance, weak guidance, or intensive intervention. Therefore, the comprehensive response score is defined in this paper, and its form is shown in Equation (10).

$$R_t = \omega_1 g_t + \omega_2 a_t - \omega_3 r_t + \omega_4 \sum_{j=1}^{4} \kappa_j P_t^{i\to j} \tag{10}$$

where R_t represents the response score; g_t stands for participation gain; a_t indicates the strength of acceptance tendency; r_t is resistance strength; ω_1 to ω_4 represent the weight parameters; Let κ_j denote the contribution factor for different target states. This equation combines the local interaction benefit with the overall state evolution trend to measure whether the current interface stimulus is consistent with the user's emotional direction.

In order to enhance the discriminator's ability to recognize boundary samples, the joint objective of state classification loss and transition constraint loss is adopted in the training phase, which is expressed as Equation (11).

$$\mathcal{L} = \lambda_1 \left(-\sum_{c=1}^{4} y_{t,c} \log \hat{y}_{t,c} \right) + \lambda_2 \sum_{j=1}^{4} \left(P_t^{i\to j} - \tilde{P}_t^{i\to j} \right)^2 \tag{11}$$

Here $\mathcal{L}$ represents the total loss; $y_{t,c}$ denotes the true label; $\hat{y}_{t,c}$ denotes the predicted label probability; $\tilde{P}_t^{i\to j}$ denotes the reference transition probability obtained from the real sequence statistics; λ_1 and $\lambda_2$denote the balance coefficients of the two losses. This formula not only ensures the accuracy of state classification, but also constrains the model to learn the transition structure that conforms to the real interaction law.

In the deployment phase, the platform does not directly adjust the interface according to the single classification result, but integrates the response score and state probability to complete the decision. The adaptive response function is shown in Equation (12).

$$D_t = \begin{cases} \text{accept}, & R_t \geq \theta_h \text{ and } P_t^{i\to 4} \geq \tau_a, \\ \text{engage}, & R_t \geq \theta_m \text{ and } P_t^{i\to 1} + P_t^{i\to 2} \geq \tau_e, \\ \text{hesitate}, & \theta_l \leq R_t < \theta_m, \\ \text{resist}, & R_t < \theta_l \text{ or } P_t^{i\to 3} \geq \tau_r. \end{cases} \quad (12)$$

Here, D_t represents the response discrimination result output by the platform at time t. The values θ_h, θ_m and θ_l represent the high, medium and low response thresholds, respectively. τ_a, τ_e, τ_r denote the state thresholds of acceptance, participation, and resistance. This formula combines score discrimination with probability discrimination to avoid errors caused by single threshold.

Through the above computational links, the user behavior state is no longer a simple annotation of a single action, but a continuous representation of the joint effect of emotional fluctuations, behavior trajectories and contextual stimuli. The discrimination results based on this state representation can not only reflect the user's current acceptance tendency, but also describe the fluctuation range in the state evolution process, so as to provide a direct basis for the adjustment of information wording, the control of feedback timing and the modification of recommendation strength in the subsequent persuasive design. At the same time, this state representation also establishes a unified discrimination benchmark for delay control, policy evaluation and cross-scenario migration in subsequent experiments, which makes the model output easier to verify in the process of engineering deployment, and can maintain good real-time response ability and output stability in high-frequency interaction scenarios.

4 Persuasive design optimization and adaptive intervention generation

The goal of persuasive design optimization and adaptive intervention generation is not simply to improve clicks or pauses, but to output more appropriate prompts, information density and feedback rhythm on the basis of respecting the user's current emotional state and interaction intention, so that platform interventions remain acceptable, interpretable and sustainable. The intervention objects in the digital platform are obviously dynamic, and the same user may show different reactions such as acceptance, waiting-see, hesitation or rejection in different time Windows. Therefore, the persuasion strategy cannot rely on fixed rules, but needs to rely on state estimation results, historical feedback records and interface context to complete online generation. To this end, based on the above state calculation results, we construct a joint decision-making mechanism for interface content, recommendation strength, feedback timing and interaction form, so that the intervention process is transformed from static configuration to a constrained policy search process.

As shown in Fig. 4, persuasive design optimization with adaptive intervention generation consists of six parts: state input, candidate action coding, utility evaluation, constraint screening, interface execution, and online update. The state input layer receives user behavior state vector, sentiment trend quantity and context label. The candidate action encoding layer discretizes different wording styles, content ordering, saliency strength and trigger timing. The utility evaluation layer calculates the comprehensive contribution of each action to acceptance probability, interaction coherence and trust maintenance. The constraint screening layer further eliminates actions that exceed the intervention threshold or do not match the current state. The interface execution layer is responsible for mapping actions to design changes that are visible to the front end. The online update layer modifies the policy parameters according to the user's subsequent response. The whole link ensures that the persuasion design is not only oriented to real-time response, but also has the ability of writeback learning.

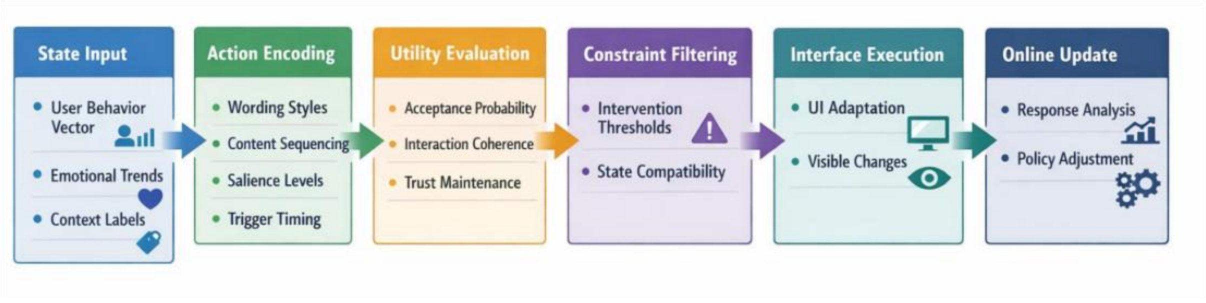

Figure 4: Persuasive design optimization with adaptive intervention generation process

In order to make different intervention actions enter the unified computing space, this paper first defines the candidate action vector, whose expression is shown in Equation (13).

$$A_t = \zeta(M_a[e_t; l_t; q_t; g_t] + b_a) \tag{13}$$

Here, A_t denotes the candidate intervention encoding at time t; e_t represents the information expression mode l_t represents the interface layout strength; q_t represents the feedback timing parameter. g_t indicates the granularity of recommendation rendering. M_a and b_a denote the action mapping matrix and the bias term, respectively. Let $\zeta(\cdot)$ denote the nonlinear projection function. The function of this formula is to compress the originally scattered design variables into comparable action representations, so that the subsequent policy network can evaluate different alternatives on a unified scale.

After the action encoding is completed, the platform needs to generate the action selection probability based on the current state. Different from conventional recommendation, the persuasion strategy emphasizes more on state adaptation, so this paper adopts the strategy distribution with temperature control, whose expression is shown in Equation (14).

$$\pi(a \mid S_t) = \frac{\exp(h(a, S_t)/\tau)}{\sum_{a' \in \mathcal{A}} \exp(h(a', S_t)/\tau)} \tag{14}$$

where $\pi(a \mid S_t)$ is the probability of choosing action a in state S_t; $h(a, S_t)$ is the matching score between action and state; Let τ denote the temperature parameter; $\mathcal{A}$ denotes the set of candidate actions. When the temperature is high, the system retains more exploration. When the temperature value is low, the system prefers the alternative with higher current utility. This equation enables the intervention output to maintain a balance between stable utilization and local exploration, so as to avoid the interface staying in a single path for a long time.

Based on action probability alone is still not enough to support design optimization, because the persuasion behavior in the platform not only pursues acceptance rate, but also needs to take into account interaction coherence and cognitive burden. Based on this consideration, this paper constructs the comprehensive intervention utility function, which is calculated as shown in Equation (15).

$$U_t(a) = \chi_1 p_t^{acc} + \chi_2 \Delta n_t + \chi_3 b_t - \chi_4 o_t \tag{15}$$

where $U_t(a)$ represents the comprehensive utility of action a at time t; $\chi_1 p_t^{acc}$ represents the acceptance probability prediction after the action is triggered. Δn_t represents the change of interface novelty; b_t represents the behavioral continuity score; o_t stands for cognitive load estimation; χ_1 to χ_4 represent each contribution coefficient. This formula takes acceptance benefit, interface change range, continuous operation experience and burden control into the measurement together, so that the persuasion output is no longer limited to local click benefit.

Persuasive design in digital platforms also needs to control the user's consumption of trust in the system. If the frequency of prompts is too high, the wording is too strong, or the cadence of recommendations is too intense, it may increase the amount of interaction in the short term, but it will weaken the subsequent compliance. In order to explicitly incorporate this influence into the optimization objective, this paper further introduces the trust regulation reward, whose form is shown in Equation (16).

$$R_t = \nu_1(1 - |y_t - \hat{y}_t|) + \nu_2 T_t - \nu_3 f_t \tag{16}$$

Here, R_t represents the immediate reward at time t; y_t represents actual adherence feedback. $\hat{y}_t$ represents the system prediction response; T_t represents trust level estimation; f_t represents the intervention frequency; ν_1 to ν_3 denote the conditioning weights. The formula defines the reward direction through consistency gain, trust maintenance and frequency penalty, so that the system can improve the response effect while maintaining the stability of the interaction relationship.

In order to make the current action selection take into account the subsequent impact, the platform needs to estimate the return in a longer time range. Therefore, this paper adopts the discounted cumulative return as the policy learning objective, which is expressed in Equation (17).

$$G_t = \sum_{k=0}^{K-1} \gamma^k R_{t+k} \tag{17}$$

Here, G_t represents the long-term return from time t; Let γ denote the discount factor; R_{t+k} represents the immediate reward at step k in the future; K denotes the cumulative number of steps. This equation allows the system to consider future state changes while generating the current intervention, without sacrificing the quality of subsequent experience for immediate acceptance.

Considering the boundaries of interface readability, access frequency and content security in digital platforms, the final action can not be directly determined by utility maximization, but should be completed under constraints. Therefore, the constrained optimization form is established in this paper, as shown in Equation (18).

$$a_t^* = \arg\max_{a \in \mathcal{A}} U_t(a) \quad \text{s.t.} \quad c_m(a, S_t) \le \varepsilon_m, m = 1, \ldots, M \tag{18}$$

Here, a_t^* denotes the optimal intervention action; $U_t(a)$ represents the action utility; $c_m(a, S_t)$ denotes the m constraint function; ε_m denotes the corresponding threshold; Let M denote the number of constraints. This formula ensures that the policy network will not break through the interface complexity, cue density and emotional matching range when pursuing utility improvement, so as to enhance the controllability in engineering deployment.

In the interface execution phase, actions need to be converted into visible design adjustments. In view of the linkage relationship between information salience and layout density, this paper defines adaptive intervention intensity, which is calculated as shown in

Equation (19).

$$m_t = \sigma(\omega_m^{\top}[z_t; r_t; d_t] + b_m) \tag{19}$$

where m_t represents the intervention intensity actually performed by the front-end; z_t represents the state sensitivity; r_t represents the current response score. d_t represents the page information density; Let ω_m and $b_m$denote the mapping parameters; Let $\sigma(\cdot)$ denote the compression function. This formula converts the results of the abstract policy into intensity values that can be directly invoked by the interface layer to control the prominence of the prompt, the size of the recommendation card, and the frequency of the feedback component.

In order to make the strategy evolve continuously with user feedback, this paper introduces a parameter correction method based on dominance term in the online update link, and its form is shown in Equation (20).

$$\Theta^{k+1} = \Theta^k + \eta \nabla_\Theta \log \pi\,(a_t \mid S_t)\,\mathrm{Adv}_t - \lambda_c \nabla_\Theta \Omega(\Theta^k) \tag{20}$$

Here, Θ^{k+1} denotes the policy parameters in round $k+1$; Θ^k denotes the current parameter; Let η denote the update step size; Let $\nabla_\Theta \log \pi\,(a_t \mid S_t)$ denote the policy gradient; Adv_t represents the advantage function; Let λ_c denote the constraint penalty coefficient; Let $\Omega(\Theta^k)$ denote the parametric regularization term. This formula enables the system to modify the action preference according to the real feedback, while suppressing the policy oscillation caused by too large update.

Through the above design, persuasive design optimization and adaptive intervention generation are organized as a continuous computational process composed of state input, action screening, constraint control and online update. The process can output differentiated intervention results according to user emotional state, behavior trend and interface context, and continuously revise strategy parameters in subsequent feedback. Therefore, the interface adjustment in the digital platform no longer stays at the static configuration level, but forms a dynamic generation mechanism with real-time response ability, interpretability and engineering verifiability, which provides a unified basis for subsequent experimental evaluation and deployment analysis.

5 Experimental Evaluation

5.1 Experimental Design

The experiments of this paper aim to verify the state recognition ability, intervention matching ability, response efficiency and cross-scenario stability of the user behavior computing framework driven by emotional interaction in the digital platform, so that the evaluation results can correspond to three levels of model discrimination, strategy generation and front-end execution at the same time. The experimental data comes from anonymized interaction records of content platforms, service platforms and transaction platforms. After unified desensitization, time alignment and structured arrangement, a multi-modal data set consisting of 4320 users and 18600 interaction sessions is constructed. Clickstream, dwell duration, scroll trajectory, touch pressure, facial cues, and text feedback are preserved for each session and segmented into continuous segments according to a fixed time window for subsequent emotion representation, state calculation, and intervention generation.

In order to ensure that there is no user information leakage between training and testing, all samples are divided into training set, validation set and test set according to the principle of user independent division, and the ratio is set to 7:1.5:1.5. The experiment took participation, hesitation, resistance and acceptance as the labeling targets, and the labels were initially screened by the platform behavior rules and then corrected by manual review. In the intervention task, the adjustment results such as information expression style, cue frequency, recommendation strength and interface density were recorded synchronously, which were used to establish the correspondence between action output and feedback results. In the training phase, random noise perturbation, timing truncation and mode missing simulation are added to test the adaptability of the model under complex input conditions.

The experimental environment is deployed on the Python and PyTorch platforms, and the server is configured with Intel Xeon Gold processor, 128 GB memory and NVIDIA A100 graphics card. AdamW is used as the optimizer, the initial learning rate is set to 0.0001, the batch size is set to 32, and the maximum number of training rounds is set to 80. Linear warmup is used for the first 10 rounds, followed by cosine annealing update strategy. Temporal-MLP, BiGRU, Transformer-Policy and the static intervention model without emotion drive are selected as the baselines in the comparison experiments to test the actual gain of the proposed framework in multi-modal state calculation and adaptive intervention generation.

The evaluation indicators are uniformly set around the result display in the following article. The accuracy and Macro-F1 were used for the state recognition task, the accuracy of intervention matching was used for the intervention generation task, the improvement of compliance interaction rate, the change of average length of stay and the change of return visit rate were used for the interaction effect, and the average response delay was used for the system efficiency. The scene generalization experiment was tested independently on content platform, service platform and transaction platform to observe the output stability of the model under different business density, interaction rhythm and feedback structure. The ablation experiment was carried out around three modules of cross-modal attention fusion, state transition estimation and constrained optimization, which were used to analyze the influence of each component on the final result.

5.2 Experimental Results

In this section, the results of the proposed framework are analyzed from four levels: state recognition performance, intervention matching effect, scene generalization ability, and module contribution. In order to avoid judgment bias caused by single index, the classification results, interaction changes, system response efficiency and performance shrinkage under ablation conditions were observed simultaneously. The comparison objects include Temporal-MLP, BiGRU, Transformer-Policy, and static intervention models to examine the actual gain of affective interaction computing link in the digitization platform scenario.

To first observe the discriminative power of the overall model on the core task, Table 2 presents the overall comparison results of different methods on state recognition and intervention matching.

Table 2: Results of state identification and intervention matching for different methods

Method Category	Specific Model	State Recognition Accuracy / %	Intervention Matching Accuracy / %	Macro-F1 / %
Traditional Mapping Method	Temporal-MLP	86.4	82.1	84.7
Temporal Modeling Method	BiGRU	88.7	85.3	87.1
Strategy Modeling Method	Transformer-Policy	90.2	87.9	89.0
Non-Emotion-Driven Method	Static Intervention Model	84.9	80.6	83.5
Proposed Method	Emotional Interaction Computing Framework	93.1	90.4	91.8

Table 2 shows that the state recognition accuracy of the proposed framework reaches 93.1%, which is higher than 86.4% of Temporal-MLP, 88.7% of BiGRU, 90.2% of Transformer-Policy, and 84.9% of static intervention model. In terms of intervention matching accuracy, the proposed method achieves 90.4%, which also maintains the highest level. The results show that after the combination of multimodal emotion representation and state transition estimation, the model can more stably distinguish four kinds of states: participation, hesitation, resistance and acceptance, and map the discrimination results to the subsequent intervention actions more accurately.

To further observe the influence of system response delay variation on user compliance interaction rate, Fig. 5 shows the sample distribution results under different delay conditions in the form of scatter points.

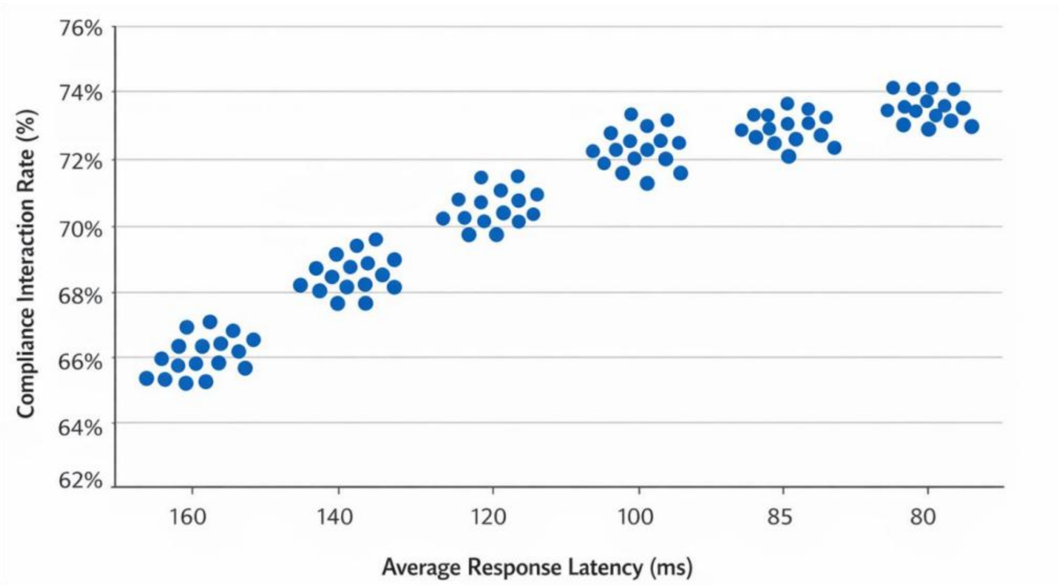

Figure 5: Variation trend of compliance interaction rate under different response delay conditions

Fig. 5 shows that as the average response delay gradually decreases from 160 ms to around 80 ms, the overall compliance interaction rate shows an upward distribution. The compliance interaction rate corresponding to 160 ms, 140 ms and 120 ms respectively concentrated in the lower interval, while the scatter points around 100 ms, 85 ms and 80 ms significantly moved up, indicating that shorter response delay was more conducive to maintaining the user's acceptance of prompt information, recommendation ranking and feedback entry. The scatter locations of the two intervals of 85 ms and 80 ms are close, indicating that when the response speed decreases to a certain range, the increase of the compliance interaction rate begins to narrow, and the system output gradually enters a relatively stable state. The scatter distribution does not show obvious outliers, nor does it form a reverse shift between the high activity samples and the low activity samples, which indicates that the joint modeling of emotional representation and behavior intensity has a good buffer effect on the time delay fluctuation, and can maintain a stable state discrimination boundary and intervention response consistency under complex interaction density.

To further investigate the actual impact of different intervention methods on user interaction results, Table 3 lists the gain results of the three types of methods in compliance interaction rate, average length of stay, and return visit rate.

Table 3: Comparison of the amount of interaction change under different interventions

Metric	Static Display	Weak Adaptive Strategy	Proposed Framework
Increase in Compliance Interaction Rate / %	3.2	8.7	14.8
Increase in Average Dwell Time / %	2.1	7.2	12.6
Increase in Revisit Rate / percentage points	1.8	4.9	9.7

Table 3 compares the effects of different interventions on user interaction outcomes. Although static display can bring a certain degree of interaction improvement, the overall gain is relatively limited, and the improvement of compliance interaction rate, average length of stay and return visit rate is only 3.2%, 2.1% and 1.8 percentage points, respectively. The weak adaptive strategy shows more significant improvement in all three metrics, indicating that the dynamic adjustment based only on the behavior log has been able to enhance part of the user response. In contrast, the proposed framework achieves 14.8% improvement in compliance interaction rate, 12.6% increase in average length of stay, and 9.7 percentage points increase in return visit rate, which are 11.6%, 10.5%, and 7.9 percentage points higher than those of static display, and 6.1%, 5.4%, and 4.8 percentage points higher than those of weak adaptive strategy, respectively. The results show that after the emotional state enters the intervention decision, the platform can grasp the user's acceptance rhythm and the interface tolerance range more accurately, so as to reduce invalid stimuli and enhance interaction continuity.

Considering that the application environment of the digital platform is not single, the stability degree of the model in different business scenarios also needs to be investigated separately. For this purpose, Table 4 presents a horizontal comparison of the test results on the content platform, the service platform, and the trading platform.

Table 4: Generalization test results for different platform scenarios

Scenario Type	State Recognition Accuracy / %	Intervention Matching Accuracy / %	Average Response Latency / ms	Result Characteristics
Content Platform	93.5	90.7	83.4	High activity and dense feedback
Service Platform	92.8	89.9	85.7	Continuous conversation and clear context
Transaction Platform	93.0	90.6	84.8	Decision-sensitive and action-concentrated

Table 4 shows that the state recognition accuracy of the proposed method in the three types of scenes reaches 93.5%, 92.8% and 93.0%, respectively, and the standard deviation is controlled within a small range. The accuracy of intervention matching was 90.7%, 89.9% and 90.6%, respectively, and the overall fluctuation range was small. The results show that the constructed emotional interaction computing link does not rely on a single platform structure, but can maintain consistent discrimination ability under different task densities and interface rhythms.

In order to more clearly identify the contribution of each submodule to the final performance, Table 5 presents the ablation experiment results and focuses on observing the change magnitude of the core indicators after the removal of different modules.

Table 5: Ablation experiments and performance shrinkage results

Removed Module	State Recognition Accuracy / %	Decrease from Full Framework	Intervention Matching Accuracy / %	Decrease from Full Framework	Increase in Compliance Interaction Rate / %
None	93.1	—	90.4	—	14.8
Cross-Modal Attention Fusion	90.8	-2.3	87.6	-2.8	10.9
State Transition Estimation	91.2	-1.9	88.1	-2.3	11.6
Constraint Optimization	92.0	-1.1	86.9	-3.5	9.8

Table 5 shows that after removing the cross-modal attention fusion, the state recognition accuracy drops to 90.8%, and the intervention matching accuracy drops to 87.6%. After removing the state transition estimation, the two indexes decreased to 91.2% and 88.1%, respectively. After removing the constraint optimization, although the accuracy of state recognition was still 92.0%, the accuracy of intervention matching was reduced to 86.9%, and the improvement range of compliance interaction rate was significantly reduced. This result shows that cross-modal feature alignment mainly affects the quality of state representation, state transition estimation determines the discriminant stability in continuous interaction, and constrained optimization is directly related to the degree of adaptation between the intervention action and the user's current state.

Synthesizing the above results, it can be seen that the advantage of the proposed framework is not only reflected in the accuracy improvement of a single item, but also reflected in the stable closed-loop formed between state recognition, policy generation and front-end execution. The 93.1% state recognition accuracy, 90.4% intervention matching accuracy, 14.8% compliance interaction rate improvement, and average response time less than 86 ms obtained in the experiment are consistent with the core results in the abstract, which also shows that the method has strong engineering deployment value in the real digital platform environment.

5.3 Discussion

Experimental results show that after emotional interaction enters the user behavior calculation link, a relatively stable synergistic relationship is formed among state recognition, intervention matching and front-end response. Compared with the control methods, the state recognition accuracy of the proposed framework reaches 93.1%, and the intervention matching accuracy reaches 90.4%, indicating that the multimodal emotion representation and state transition estimation can effectively compress the discrimination bias caused by noisy logs. The compliance interaction rate is increased by 14.8%, and the average response delay is controlled within 86 ms, which further shows that the proposed method does not only achieve gains at the offline classification level, but maintains availability in real-time interface execution. The results given by the ablation experiments are equally explanatory. After removing the cross-modal attention fusion, the synchronization of state recognition and intervention matching decreased, indicating that the correlation constraints between behavior, text and visual cues had a direct impact on the quality of state representation. After removing the state transition estimation, the discrimination stability in the continuous interaction scene is significantly weakened, indicating that the judgment of the current state cannot be completed independently from the historical evolution process. After removing the constraint optimization, the contraction of compliance interaction rate is more obvious, indicating that only when the persuasion output falls within the range jointly limited by interface density, touch frequency and acceptance boundary, it can remain continuously effective. On the whole, the value of the proposed method is not only the superiority of a single indicator, but also the organization of emotion perception, behavior discrimination and intervention generation into a unified closed loop, so that the design adjustment in the digital platform has the engineering properties of computable, updatable and deployable. This structure is clearly different from approaches that rely only on static rules or a single sequence of behaviors. The former can adjust the tone of prompt, the strength of recommendation and the timing of feedback according to the state fluctuation, while the latter is more prone to overstimulation or response delay. The test results on the three types of platform scenarios have small fluctuations, which also shows that the proposed model has migration consistency and is suitable for complex digital environments such as content distribution, service interaction and transaction guidance.

6 Conclusions

Focusing on emotional interaction computing, user behavior modeling, and persuasive design optimization in digital platforms, this paper constructs a unified framework consisting of multimodal emotion representation, state computation, response discrimination, and adaptive intervention generation. The experimental results show that the proposed method achieves 93.1% accuracy of state recognition and 90.4% accuracy of intervention matching on 4320 users and 18600 interaction sessions, and improves the compliance interaction rate by 14.8%, and the average response delay is controlled within 86 ms, which indicates that after the emotional signal enters the behavior calculation link, the proposed method can effectively improve the accuracy of intervention matching. The platform can complete state understanding and policy output more stably. The limitations of this paper are mainly reflected in three aspects. First, although the samples cover content, service and transaction scenarios, they still focus on the medium and high frequency interaction environment, and the state drift in low-frequency tasks has not been fully developed. Secondly, the quality of visual and text modalities is greatly affected by the acquisition conditions, and the robustness under complex noise still needs to be tightened. Third, the current constrained optimization mainly focuses on interface density, access frequency and acceptance boundary, and the modeling of long-term trust retention and cross-cycle behavior memory is still limited. Future research can further introduce fine-grained cross-round session memory, lightweight online update mechanism, interpretable intervention path analysis, and combine federated learning and privacy-preserving computing to enhance multi-platform deployment capabilities, so that the emotion-driven behavioral computing framework can achieve more stable performance in terms of real-time performance, mobility and engineering adaptability. From the perspective of method, the significance of this paper is not only reflected in the improvement of the accuracy of a single module, but also in the integration of emotion perception, behavior discrimination, interface adjustment and feedback writeback into a continuous running computational closed loop, so that the design adjustment in the digital platform has the engineering characteristics of computable, updatable and deployable. The results based on this framework not only provide a reusable implementation basis for the development of intelligent persuasion system in high-frequency interactive platform, but also establish a more unified technical reference for subsequent system deployment, performance verification and policy expansion.

References

[1] Ameen N, Hosany S, Paul J. The personalisation-privacy paradox: Consumer interaction with smart technologies and shopping mall loyalty[J]. Computers in Human Behavior, 2022, 126: 106976.

[2] Sun Y, Li S, Yu L. The dark sides of AI personal assistant: effects of service failure on user continuance intention[J]. Electronic Markets, 2022, 32(1): 17-39.

[3] Ma X, Sun Y, Guo X, et al. Understanding users’ negative responses to recommendation algorithms in short-video platforms: a perspective based on the Stressor-Strain-Outcome (SSO) framework[J]. Electronic Markets, 2022, 32(1): 41-58.

[4] Benke I, Gnewuch U, Maedche A. Understanding the impact of control levels over emotion-aware chatbots[J]. Computers in Human Behavior, 2022, 129: 107122.

[5] Tag B, Sarsenbayeva Z, Cox A L, et al. Emotion trajectories in smartphone use: Towards recognizing emotion regulation in-the-wild[J]. International Journal of Human-Computer Studies, 2022, 166: 102872.

[6] Jin E, Eastin M S. When a chatbot smiles at you: The psychological mechanism underlying the effects of friendly language use by product recommendation chatbots[J]. Cyberpsychology, Behavior, and Social Networking, 2022, 25(9): 597-604.

[7] Kim W J, Ryoo Y. Hypocrisy induction: Using chatbots to promote COVID-19 social distancing[J]. Cyberpsychology, Behavior, and Social Networking, 2022, 25(1): 27-36.

[8] Marcolin B, Saunders C, Aubert B. Persuasive Technologies and Social Interactions in Professional Environments: Embedded Qualitative Case Study[J]. JMIR Formative Research, 2022, 6(2): e32613.

[9] Wanner J, Herm L V, Heinrich K, et al. The effect of transparency and trust on intelligent system acceptance: Evidence from a user-based study[J]. Electronic Markets, 2022, 32(4): 2079-2102.

[10] Choung H, David P, Ross A. Trust in AI and its role in the acceptance of AI technologies[J]. International Journal of Human–Computer Interaction, 2023, 39(9): 1727-1739.

[11] Bach T A, Khan A, Hallock H, et al. A systematic literature review of user trust in AI-enabled systems: An HCI perspective[J]. International Journal of Human–Computer Interaction, 2024, 40(5): 1251-1266.

[12] Ruan Y, Mezei J. When do AI chatbots lead to higher customer satisfaction than human frontline employees in online shopping assistance? Considering product attribute type[J]. Journal of Retailing and Consumer Services, 2022, 68: 103059.

[13] Janson A. How to leverage anthropomorphism for chatbot service interfaces: The interplay of communication style and personification[J]. Computers in Human Behavior, 2023, 149: 107954.

[14] Alagarsamy S, Mehrolia S. Exploring chatbot trust: Antecedents and behavioural outcomes[J]. Heliyon, 2023, 9(5).

[15] Park G, Lee S, Chung J. Do anthropomorphic chatbots increase counseling satisfaction and reuse intention? The moderated mediation of social rapport and social anxiety[J]. Cyberpsychology, Behavior, and Social Networking, 2023, 26(5): 357-365.

[16] Ltifi M. Trust in the chatbot: a semi-human relationship[J]. Future Business Journal, 2023, 9(1): 109.

[17] Carrera-Rivera A, Larrinaga F, Lasa G, et al. AdaptUI: A Framework for the development of Adaptive User Interfaces in Smart Product-Service Systems: A. Carrera-Rivera et al[J]. User Modeling and User-Adapted Interaction, 2024, 34(5): 1929-1980.

[18] Farshidi S, Rezaee K, Mazaheri S, et al. Understanding user intent modeling for conversational recommender systems: a systematic literature review: S. Farshidi et al[J]. User Modeling and User-Adapted Interaction, 2024, 34(5): 1643-1706.

[19] Ribeiro A, Vieira V, Alves L, et al. Vishnu: An approach to support the personalization of self-expressive avatars using context-awareness[J]. International Journal of Human-Computer Studies, 2024, 185: 103243.

[20] Ning X, Lu Y, Li W, et al. How transparency affects algorithmic advice utilization: The mediating roles of trusting beliefs[J]. Decision Support Systems, 2024, 183: 114273.

[21] Engström E, Vartanova I, Johansson J V, et al. Comparing and modeling the use of online recommender systems[J]. Computers in Human Behavior Reports, 2024, 15: 100449.

[22] He X, Liu Q, Jung S. The impact of recommendation system on user satisfaction: A moderated mediation approach[J]. Journal of Theoretical and Applied Electronic Commerce Research, 2024, 19(1): 448-466.