

Dual-Phase Playtime-guided Recommendation: Interest Intensity Exploration and Multimodal Random Walks

Jingmao Zhang*
 Harbin Institute of
 Technology
 Shenzhen, China
 220810426@stu.hit.edu.cn

Zhitong Zhao*
 Harbin Institute of
 Technology
 Shenzhen, China
 24s151185@stu.hit.edu.cn

Yunqi Lin
 Harbin Institute of
 Technology
 Shenzhen, China
 210110519@stu.hit.edu.cn

Jianghong Ma†
 Harbin Institute of
 Technology
 Shenzhen, China
 mjianghong@hit.edu.cn

Tianjun Wei†
 Nanyang Technological
 University
 Singapore
 tjwei2-c@my.cityu.edu.hk

Haijun Zhang
 Harbin Institute of
 Technology
 Shenzhen, China
 hjzhang@hit.edu.cn

Xiaofeng Zhang
 Harbin Institute of
 Technology
 Shenzhen, China
 zhangxiaofeng@hit.edu.cn

Abstract

The explosive growth of the video game industry has created an urgent need for recommendation systems that can scale with expanding catalogs and maintain user engagement. While prior work has explored accuracy and diversity in recommendations, existing models underutilize playtime, a rich behavioral signal unique to gaming platforms, and overlook the potential of multimodal information to enhance diversity. In this paper, we propose *DP²Rec*, a novel Dual-Phase Playtime-guided Recommendation model designed to jointly optimize accuracy and diversity. First, we introduce a playtime-guided interest intensity exploration module that separates strong and weak preferences via dual-beta modeling, enabling fine-grained user profiling and more accurate recommendations. Second, we present a playtime-guided multimodal random walks module that simulates player exploration using transitions guided by both playtime-derived interest similarity and multimodal semantic similarity. This mechanism preserves core preferences while promoting cross-category discovery through latent semantic associations and adaptive category balancing. Extensive experiments on a real-world game dataset show that *DP²Rec* outperforms existing methods in both recommendation accuracy and diversity. The dataset and source code are released at <https://github.com/zqxwcevrtyui/DP2Rec>

CCS Concepts

- Information systems → Recommender systems; Information extraction; Personalization;

*Both authors contributed equally to this research.

†Corresponding author

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

MM '25, Dublin, Ireland

© 2025 Copyright held by the owner/author(s). Publication rights licensed to ACM.
 ACM ISBN 979-8-4007-2035-2/2025/10
<https://doi.org/10.1145/3746027.3755597>

Keywords

Game Recommendation, Playtime, Accuracy, Diversity

ACM Reference Format:

Jingmao Zhang, Zhitong Zhao, Yunqi Lin, Jianghong Ma, Tianjun Wei, Haijun Zhang, and Xiaofeng Zhang. 2025. Dual-Phase Playtime-guided Recommendation: Interest Intensity Exploration and Multimodal Random Walks. In *Proceedings of the 33rd ACM International Conference on Multimedia (MM '25)*, October 27–31, 2025, Dublin, Ireland. ACM, New York, NY, USA, 10 pages. <https://doi.org/10.1145/3746027.3755597>

1 Introduction

The video game industry is growing rapidly, with 18,949 new games released on Steam in 2024—a 32.42% increase from 2023. Steam also hit a record 39.3 million concurrent users in December 2024, reflecting the rising demand driven by rich visual and textual content. This surge underscores the need for domain-specific recommendation systems [5, 7, 12, 22, 33, 44] to help users navigate expanding game catalogs and enhance engagement. This accelerated market expansion underscores the critical need for advanced, domain-specific recommendation systems in the gaming domain [5, 7, 12, 22, 33, 44]. Such systems are instrumental in helping users navigate the growing volume of releases by recommending preferred games, thereby enhancing user experience and platform retention.

While accuracy, driven by users' historical behavior [8, 29, 35, 43, 50], remains the primary focus of recommender systems, recent work has broadened the evaluation scope to include diversity, which is essential for reducing redundancy and promoting discovery [10, 42, 48, 51, 53]. Striking a balance between the two ensures sustained engagement without sacrificing personalization [14, 17, 18, 32]. While these dimensions have been extensively studied in general recommendation settings, domain-specific adaptations—such as game recommendation—have seen incremental advances. For instance, SCGRec [45] boosts accuracy via social-contextualized graph neural networks; DRGame [28] augments diversity through clustered node selection; and CPGRec [23] imposes category-aware constraints to harmonize both objectives.

Crucially, *playtime*—a behavioral signal unique to gaming platforms like Steam—provides an untapped granularity for preference modeling. Unlike discrete interactions (e.g., clicks or ratings) in

traditional domains, playtime offers a temporal proxy for latent interest, yet current methods fail to exploit its full potential. CPGRRec disregards playtime entirely, relying only on categorical graph construction, while SCGRec and DRGame inadequately model interaction strength, failing to resolve fine-grained preference distinctions encoded in playtime patterns. Specifically, two key challenges persist in game recommendation research:

- **Playtime-aware Dual-Interest Modeling.** Current game recommenders underutilize playtime data, often treating interactions as binary preferences. While some methods incorporate playtime, they miss the interest intensity spectrum it encodes - prolonged engagement indicates strong interest while brief interactions suggest weak interest. Our analysis (Section 3.1) shows two distinct playtime-driven interest types, necessitating separation of core preferences from casual ones. Though traditional domains [4, 13, 27, 39] model dynamic interests, they lack playtime integration for gaming contexts, highlighting the need for fine-grained interest differentiation.
- **Cross-Category Diversity Through Multimodal Similarity.** While multimodal learning is prevalent in graph [26, 40] and recommenders, existing methods (both general and game-specific) primarily use modalities (text, images) for accuracy, overlooking diversity potential. Current approaches limit content similarity to historical preferences, causing homogenized recommendations. We show multimodal representations can reveal latent cross-category associations, enabling diverse yet relevant recommendations. Our analysis (Section 3.2) demonstrates that high multimodal similarity often occurs across distinct categories, proving modalities can transcend categorical constraints.

To address the above challenges, we propose a novel model: Dual-Phase Playtime-guided Recommendation: Interest Intensity Exploration and Multimodal Random Walks (DP^2Rec)—designed to improve both the accuracy and diversity of game recommendations. **First**, we introduce a *playtime-guided interest intensity exploration* module to enhance accuracy, which identifies strong and weak preferences in players' gameplay duration via dual-beta distribution separation based on Expectation-Maximization (EM) algorithm [15], thereby enabling fine-grained modeling of user interest levels. **Second**, to enhance diversity, we present a *playtime-guided multimodal random walks* module. This module uses a novel random walks mechanism to simulate the user exploration process where node transitions are jointly determined by semantic alignment (via multimodal embeddings) and behavioral alignment (via playtime-based interest). The dual-alignment preserves users' core interest while enabling cross-category discovery through multimodal space traversal and adaptive category balance coefficients. Collectively, these two modules lead to significant improvements in both recommendation accuracy and diversity.

The main contributions can be summarized as follows:

- To our knowledge, this work presents the **first** framework in game recommendation systems that leverages playtime dynamics to explicitly optimize both accuracy and diversity.
- Through in-depth data analysis, we uncover two key insights: (1) playtime reflects distinct levels of user interest intensity, supported by theoretical validation, revealing fine-grained preference structures; and (2) multimodal representations of games

uncover latent semantic links across categories, enabling effective cross-category exploration.

- Building on the above-mentioned insights, we design two complementary modules: an interest intensity exploration module to boost accuracy, and a multimodal random walk module to enhance diversity. Theoretical analysis further substantiates the effectiveness of the proposed diversity-enhancing mechanism.
- Extensive experiments on real-world game data demonstrate that our method outperforms state-of-the-art recommendation techniques in both accuracy and diversity.

2 Related Work

2.1 Game Recommendation

Early game recommendation research focused on Collaborative Filtering (CF) and Content-Based Filtering (CBF). GAMBIT [2] applies CF by modeling users and games separately, while BharathiPriya *et al.* [5] combine CF and CBF, incorporating playtime for implicit ranking. With the rise of deep learning, models like DeepNN and DeepFM [12] have outperformed traditional methods on the Steam dataset. Graph Neural Networks (GNNs) have also gained traction for their strength in modeling complex relations. For instance, SCGRec [44] enhances accuracy using game context and social links, while CPGRec [23] improves propagation by enforcing strict game connections.

Discussion. Our approach differs from existing methods in two key ways: (1) As far as we know, we are the **first** to model player behavior by extracting dual-interest signals from playtime; and (2) as far as we know, we are also the **first** to introduce a random walks-based item selection strategy for game recommendation.

2.2 Diversified Recommendation

Since Ziegler *et al.* [55] introduced diversified recommendation, many follow-up approaches have been developed to enhance diversity. Re-ranking methods like Maximum Marginal Relevance (MMR) [1, 6, 25, 34] and Determinantal Point Process (DPP) [11, 19, 21] enhance diversity by adjusting output rankings. More recently, GNN-based methods [38] leverage high-order graph structures for diversity. DGCN [52] incorporates category-aware sampling, while DGRec [46] introduces submodular neighbor selection. DRGame [28] explores cluster-based diversification, and DDGraph [47] dynamically updates the user-item graph for diverse recommendations. Emerging LLM-based approaches, such as DL-CRec [9], offer controllable frameworks to balance category diversity using language models.

Discussion. To the best of our knowledge, we are the **first** to leverage multimodal similarity specifically to enhance recommendation diversity. Our playtime-guided multimodal random walks module integrates multimodal information and adopts a hierarchical exploration strategy. By modeling the dynamic shift of user interests through path-level cumulative similarity, it preserves personalization while effectively expanding cross-category discovery.

3 Data Analysis

We begin with data analysis of the Steam gaming platform dataset¹.

¹<https://store.steampowered.com>

3.1 User Interest Intensity Modeling Based on Dual-Beta Distribution

3.1.1 Rationale for Dual-Beta Modeling. To characterize users' interest intensity in games, we categorize player interactions into strong and weak interest interactions, using Beta distribution for precise modeling. First, players' gameplay time is normalized to the interval $[0, 1]$ using the percentile method:

$$\hat{t}_i^u = \frac{\text{rank}(t_i^u)}{\text{total_players}(i)} \quad (1)$$

where t_i^u represents user u 's gameplay time for game i , $\text{rank}(t_i^u)$ denotes the ascending rank position of this time among all users who played game i (longer time corresponds to larger value), and $\text{total_players}(i)$ represents the total number of users who played game i . This normalization method converts time to relative percentiles, perfectly matching the domain of Beta distribution and forming an ideal modeling foundation.

The Beta distribution is determined by shape parameters α and β , with probability density function:

$$f(x; \alpha, \beta) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha, \beta)}, \quad x \in (0, 1) \quad (2)$$

where $B(\alpha, \beta)$ is the Beta function². Notably, **strong interest** often concentrates in high percentiles (close to 1), which can be represented by Beta distributions where $\alpha > \beta$, such as Beta(5, 2). With this parameter configuration, the density function peaks near 1, precisely reflecting users' behavior of investing long gameplay time in preferred games. Conversely, **weak interest** can be represented by Beta distributions where $\alpha < \beta$, such as Beta(2, 5), with distributions skewed toward lower percentiles, matching users' shallow interaction characteristics.

Based on above discussions, we establish separate Beta distributions for each user's strong and weak interests, collectively called the dual-beta model. The probability density function of the dual-beta distribution is:

$$f(x) = \pi f_1(x; \alpha_1, \beta_1) + (1 - \pi) f_2(x; \alpha_2, \beta_2)$$

where π represents the weight of the strong interest distribution, and f_1 and f_2 are the Beta distribution density functions for strong and weak interests, respectively.

3.1.2 Validation of Dual-Beta Modeling via KS Test. We select users with no fewer than 10 game interactions from the original pool of 60,742 players, resulting in a total of 53,025 valid users as our study sample. These users have relatively complete behavioral records, better reflecting their intrinsic interest structures. For each user, we use Maximum Likelihood Estimation (MLE) to separately fit the Beta distribution parameters for strong and weak interests.

Table 1: KS Fitting Test Statistics

Successful Fits	Average KS Statistic	Average p -value
53003 / 53025 (99.96%)	0.0696	0.9638

²The percentile-based normalized playtime $\hat{t}_i^u \in [0, 1]$ may include boundary values 0 or 1. To ensure compatibility with the open interval $(0, 1)$ required by the Beta distribution and to avoid numerical instability, we apply a small offset ϵ (e.g., 10^{-6}) to shift extreme values into $(0, 1)$.

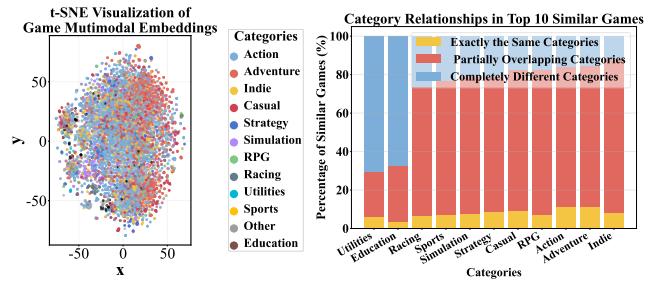
To evaluate the fitting quality, we apply the Kolmogorov-Smirnov (KS) test [30]. A fitting is considered successful if the p -value exceeds the significance level of 0.05. The statistical results are shown in Table 1. The experimental results indicate that among the 53,025 users, **53,003 users (99.96%)** had their interest distributions accurately modeled by the dual-beta distribution. Notably, the **average p -value reached 0.9638, far exceeding the 0.05 threshold**, which strongly supports the effectiveness and accuracy of using dual-beta distributions for modeling user interests.

Table 2: Beta Distribution Parameter Statistics

Interest Type	Average α	Average β
Strong Interest	9.58	2.26
Weak Interest	4.68	8.37

Further analysis of the fitted parameters yields the statistical characteristics for strong and weak interest distributions as shown in Table 2. From the parameter perspective, **strong interest** distributions generally exhibit a **left-skewed** shape with $\alpha > \beta$ (averaging Beta(9.58, 2.26)), which aligns with theoretical expectations and indicates players tend to engage deeply and consistently with games of strong interest. In contrast, **weak interest** distributions exhibit a **right-skewed** form with $\alpha < \beta$ (averaging Beta(4.68, 8.37)), reflecting shallow or occasional interactions. This also captures a subset of users with more exploratory behavioral patterns.

3.2 Cross-Category Structure Analysis Based on Multimodal Game Representations



(a) t-SNE visualization of game multimodal embeddings **(b) Category relationships in Top 10 modally similar games**

Figure 1: Multimodal distribution and structural relations across game categories.

To assess whether multimodal information facilitates cross-category exploration, we analyze the multimodal embeddings of games. We first use a large language model (LLM) to generate textual descriptions of game cover images and fuse them with the original game introductions. The fused text is then encoded using BERT [16] to obtain a semantic embedding for each game. In this study, we adopt Qwen-2.5-VL-72B-Instruct [3] as the LLM for both image captioning and text fusion, due to its strong contextual understanding and natural language generation capabilities, making it well-suited for aligning visual and textual modalities.

As shown in Figure 1a, the t-SNE visualization of multimodal embeddings reveals that games from different categories do not form clearly separated clusters in the embedding space, but rather exhibit significant overlap. Figure 1b further illustrates the category distribution of the top-10 games retrieved based on multimodal similarity. Specifically, for each category, we retrieve the top-10 similar games for every game in that category, then compute the average distribution of category relationships among those similar games. For example, in the *Action* category, only 11.30% of the similar games belong to exactly the same category, 72.82% partially overlap, and 15.88% fall into completely different categories. In contrast, in the *Indie* category, 79.89% of the similar games exhibit partial overlap. These results suggest that **multimodal similarity can support cross-category discovery while preserving semantic relevance**.

4 Method

4.1 Overview

As depicted in Figure 2, the proposed *DP²Rec* model consists of two main modules: *Playtime-guided Interest Intensity Exploration (IIE)* and *Playtime-guided Multimodal Random Walks (MRW)*. The former is designed to enhance accuracy, whereas the latter promotes diversity. Lastly, a balance module is incorporated to fuse the result embeddings from the aforementioned modules.

4.2 Playtime-guided Interest Intensity Exploration

As detailed in Section 3.1, we model strong and weak interests using a dual-beta distribution tailored to the playtime signal. To disentangle these distinct interest types, we employ the EM algorithm, enabling a principled separation of underlying interest patterns.

4.2.1 Dual-Interest Recognition via EM. The EM algorithm offers an efficient framework for parameter estimation in probabilistic models with latent variables from mixture distributions. By alternating between E and M steps, it uncovers hidden patterns from unlabeled data. In our case, EM is used to classify playtime distributions into strong and weak interests.

To model the distributional patterns of user playtime, we adopt a dual-beta mixture model within the EM algorithm:

$$p(\hat{t}_i^u) = \pi \cdot \text{Beta}(\hat{t}_i^u; \alpha_S, \beta_S) + (1 - \pi) \cdot \text{Beta}(\hat{t}_i^u; \alpha_W, \beta_W) \quad (3)$$

where $p(\hat{t}_i^u)$ denotes the probability density function of player u 's normalized playtime for game i , and π is the mixing weight for the strong interest component. The Beta parameters (α_S, β_S) and (α_W, β_W) correspond to strong and weak interests.

We independently perform the EM algorithm on each player's normalized playtime set $\{\hat{t}_i^u\}_{i \in \mathcal{I}_u}$, where \mathcal{I}_u is the set of games interacted by player u . Initialization begins by sorting playtime in descending order and randomly selecting the top 40% as the initial strong interest set, from which initial parameter estimates are derived. Interactions with zero playtime are treated as weak interest and excluded from EM updates. During the E-step, for each non-zero interaction (u, i) , the posterior probability of representing

strong interest is calculated as:

$$\gamma_{u,i}^{(S)} = \frac{\pi \cdot \text{Beta}(\hat{t}_i^u; \alpha_S, \beta_S)}{\pi \cdot \text{Beta}(\hat{t}_i^u; \alpha_S, \beta_S) + (1 - \pi) \cdot \text{Beta}(\hat{t}_i^u; \alpha_W, \beta_W)} \quad (4)$$

where the Beta probability density function is defined in Equation 2. In the M-step, we treat the posterior probabilities from the E-step as soft labels and perform weighted MLE to update the distribution parameters:

$$\alpha_S = \frac{\sum_{i \in \mathcal{I}_u} \gamma_{u,i}^{(S)} \ln \hat{t}_i^u}{-\sum_{i \in \mathcal{I}_u} \gamma_{u,i}^{(S)} \ln \hat{t}_i^u + \sum_{i \in \mathcal{I}_u} \gamma_{u,i}^{(S)} \ln (1 - \hat{t}_i^u)} \quad (5)$$

$$\beta_S = \alpha_S \cdot \frac{\sum_{i \in \mathcal{I}_u} \gamma_{u,i}^{(S)} \ln (1 - \hat{t}_i^u)}{\sum_{i \in \mathcal{I}_u} \gamma_{u,i}^{(S)} \ln \hat{t}_i^u} \quad (6)$$

The mixing weight π is updated as the average of the posterior probabilities:

$$\pi = \frac{1}{|\mathcal{I}_u|} \sum_{i \in \mathcal{I}_u} \gamma_{u,i}^{(S)} \quad (7)$$

By iteratively performing E and M steps until convergence, we obtain the final posterior probability $\gamma_{u,i}^{(S)}$ for each interaction (u, i) being classified as strong interest.

4.2.2 Dual-Interest Graph Learning. Based on the posterior probabilities, interactions with $\gamma_{u,i}^{(S)} > 0.5$ are classified as strong interest; others are considered weak. We construct two complementary graph views: the full interest graph $\mathcal{G}_F = (\mathcal{V}, \mathcal{E}_F)$, which includes all interactions, and the strong interest graph $\mathcal{G}_S = (\mathcal{V}, \mathcal{E}_S)$, which retains only edges satisfying $\gamma_{u,i}^{(S)} > 0.5$. Notably, (1) \mathcal{G}_S filters out weak interactions to highlight **users' core interests**, while (2) the \mathcal{G}_F retains both dominant and exploratory signals, offering a **comprehensive view of user behavior**.

To enhance strong-interest signals while effectively leveraging the supplementary information in weak-interest interactions, we apply contrastive learning between the embeddings derived from both graph views, formulated by the following loss function:

$$\begin{aligned} \mathcal{L}_{ssl, user} &= \sum_{u \in \mathcal{U}} -\log \frac{\exp(\phi(\mathbf{e}_u^F, \mathbf{e}_u^S) / \tau)}{\sum_{v \in \mathcal{U}} \exp(\phi(\mathbf{e}_u^F, \mathbf{e}_v^S) / \tau)} \\ \mathcal{L}_{ssl, item} &= \sum_{i \in \mathcal{I}} -\log \frac{\exp(\phi(\mathbf{e}_i^F, \mathbf{e}_i^S) / \tau)}{\sum_{j \in \mathcal{I}} \exp(\phi(\mathbf{e}_i^F, \mathbf{e}_j^S) / \tau)} \end{aligned} \quad (8)$$

where \mathbf{e}_u^F and \mathbf{e}_i^F are the embeddings learned from the full graph \mathcal{G}_F , and \mathbf{e}_u^S and \mathbf{e}_i^S are the corresponding embeddings from the strong interest graph \mathcal{G}_S . $\phi(\cdot, \cdot)$ denotes the cosine similarity, and τ is the temperature hyperparameter.

After optimization, we obtain enhanced representations \mathbf{e}_u^{IIE} and \mathbf{e}_i^{IIE} that fuse dual-interest signals. These embeddings strengthen users' core interests while retaining potentially valuable signals from weak interactions, providing richer and more personalized representations for downstream recommendation tasks.

The overall contrastive loss to balance the contributions of user-side and item-side contrastive learning is:

$$\mathcal{L}_{ssl} = \mathcal{L}_{ssl, user} + \lambda \cdot \mathcal{L}_{ssl, item} \quad (9)$$

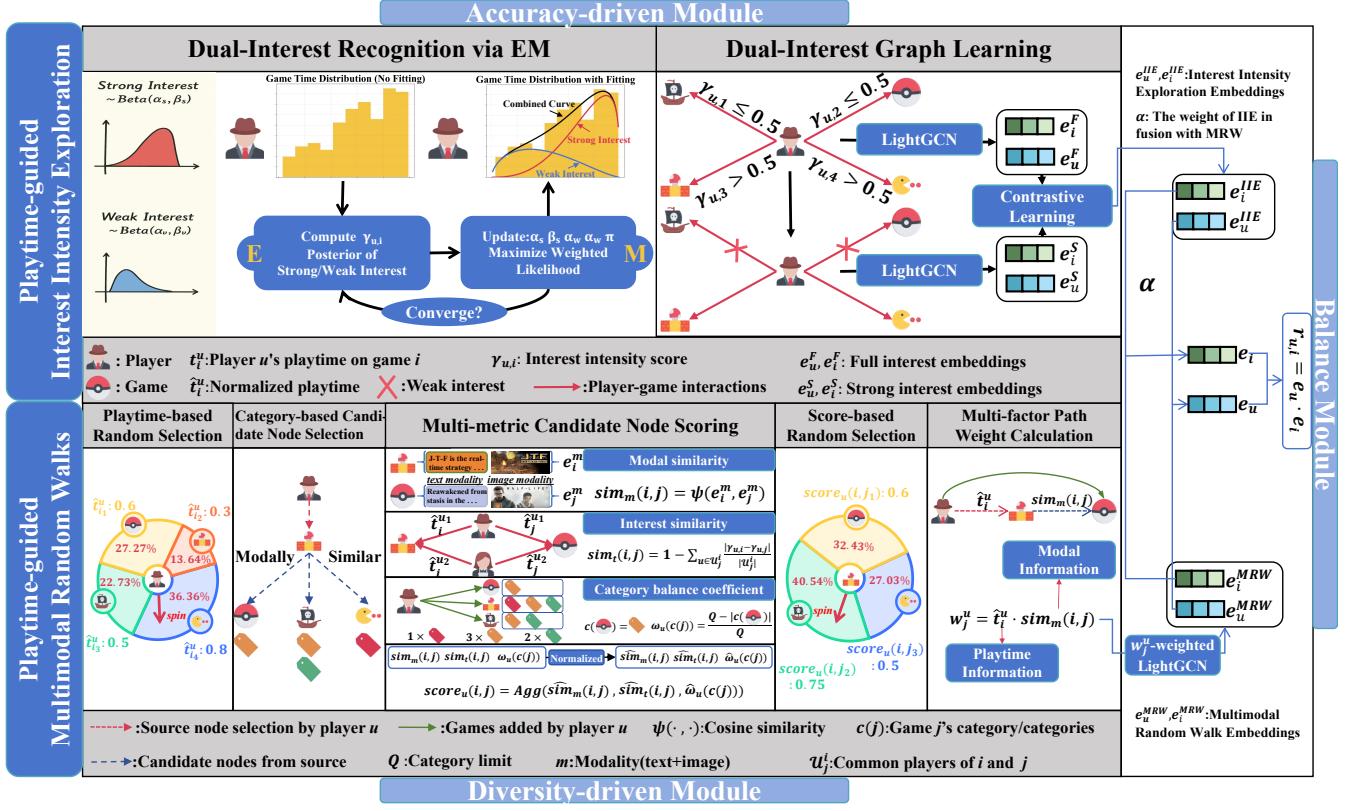


Figure 2: An illustration of the proposed DP^2Rec model. The upper module applies *Playtime-guided IIE*, where strong and weak interests are disentangled and encoded through two submodules. The lower module features a *Playtime-guided MRW*, where multimodal and interest-based similarities guide user exploration through five submodules, preserving core preferences while enhancing diversity via cross-category discovery based on multimodal similarity and adaptive category balance coefficients. The balance module on the right fuses both modules using weighted integration to improve both accuracy and diversity.

where λ is a hyperparameter controlling the relative importance of the two loss terms.

4.3 Playtime-guided Multimodal Random Walks

To model users' interest migration, we adopt a random walks strategy guided by multimodal similarity and playtime-informed patterns, ensuring alignment with core preferences. This approach enhances category coverage by leveraging multimodal embeddings that capture semantic associations beyond explicit category boundaries (see Section 3.2) and by introducing a category balance coefficient that encourages cross-category exploration.

4.3.1 Playtime-based Random Selection (Initial node Selection). For each player u , we perform multiple rounds of random walks. At the beginning of each exploration round, we first select an initial exploration node from the user's strong interest game set I_u^S . To reflect the user's interest intensity across different games, we use the normalized playtime \hat{t}_i^u on game i as the sampling weight and apply a multinomial sampling strategy. The probability of selecting game node i is defined as follows:

$$P_u(i) = \frac{\hat{t}_i^u}{\sum_{j \in I_u^S} \hat{t}_j^u} \quad (10)$$

During multi-round exploration, we record the set of games \mathcal{A}_u^i expanded from each initial node i for user u . If node i is selected again, a game from \mathcal{A}_u^i is randomly chosen as a new starting point to maintain interest continuity and expand the exploration space, thereby improving category coverage and recommendation diversity. This progressive strategy avoids redundant paths and better utilizes recommendation opportunities. All added games form $\mathcal{A}_u = \bigcup_{i \in I_u^S} \mathcal{A}_u^i$, which is used to prevent repeated recommendations.

4.3.2 Category-based Candidate Node Selection. Starting from the initial game node i , we conduct hierarchical exploration. For the current node i , we first find the most modally similar game in each game category, and these most similar games from different categories form the candidate neighbor set $\Omega_u(i)$ of game i .

$$\Omega_u(i) = \bigcup_{c \in C} \{j \mid j = \arg \max_{j \in I_c, j \neq i} \phi(\mathbf{e}_i^m, \mathbf{e}_j^m)\} \quad (11)$$

where C is the set of all categories, \mathcal{I}_c represents the set of games in category c , \mathbf{e}_i^m represents the modal embedding vector of game i , which is extracted based on Section 3.2.

To prevent repeated recommendations, we perform preliminary filtering on the candidate set $\Omega_u(i)$ to obtain a new filtered candidate set $\Omega'_u(i)$:

$$\Omega'_u(i) = \{j \in \Omega_u(i) \mid j \notin \mathcal{I}_u \wedge j \notin \mathcal{A}_u \wedge |c^*(j)| < Q\} \quad (12)$$

where $|c^*(j)|$ denotes the number of times category $c^*(j)$ appears in the already recommended set \mathcal{A}_u , $c^*(j)$ represents the category that game j stands for in the candidate generation process, and Q represents the maximum number of game nodes that can be added per category for each user.

4.3.3 Multi-metric Candidate Node Scoring. For each game j in the candidate set $\Omega'_u(i)$, we consider three key metrics to select game nodes to establish new connections with player u :

a) *Multimodal Similarity.* Based on the cosine similarity of game modal content embedding vectors:

$$\text{sim}_m(i, j) = \phi(\mathbf{e}_i^m, \mathbf{e}_j^m) \quad (13)$$

Multimodal similarity captures content-level semantic connections between games by integrating textual descriptions and visual features. It enables the recommendation of content-related games that may differ in collaborative patterns, enhancing both relevance and the potential for novel discovery.

b) *Interest Similarity.* Based on the similarity of players' interest intensity scores on two games, we define the interest similarity as:

$$\text{sim}_t(i, j) = 1 - \frac{1}{|\mathcal{U}_{i,j}|} \sum_{u \in \mathcal{U}_{i,j}} |\gamma_{u,i} - \gamma_{u,j}| \quad (14)$$

where $\gamma_{u,i}$ denotes the interest intensity score of user u for game i (i.e., $\gamma_{u,i}^{(S)}$), and $\mathcal{U}_{i,j}$ denotes the set of users who have played both games i and j . Interest similarity captures implicit associations between games by measuring consistency in user engagement patterns. Complementing content-based multimodal analysis, it reveals behaviorally similar games across categories, enhancing recommendation diversity while preserving alignment with user preferences.

c) *Category Balance Coefficient.* It controls the balance of recommendation quantities across categories, enhancing diversity selection:

$$\omega_u(j) = \frac{Q - |c^*(j)|}{Q} \quad (15)$$

The category balance coefficient adjusts selection probabilities during random walks to prevent over-concentration in a few categories, promoting broader category coverage.

These three metrics are normalized in the candidate game set to obtain a comprehensive score:

$$\text{score}_u(i, j) = \text{Agg}(\hat{\text{sim}}_m(i, j), \hat{\text{sim}}_t(i, j), \hat{\omega}_u(c^*(j))) \quad (16)$$

where $\hat{\text{sim}}_m(i, j)$, $\hat{\text{sim}}_t(i, j)$, and $\hat{\omega}_u(c^*(j))$ are the normalized multimodal similarity, interest similarity, and category balance coefficient, respectively. $\text{Agg}(\cdot)$ can be any differentiable function. We use equal-weight averaging to avoid tuning and ensure fairness.

This fused scoring strategy integrates semantic relevance, behavioral patterns, and category diversity to balance recommendation accuracy and exploration.

4.3.4 Score-based Random Selection (Node Walking). Based on the comprehensive game scores, we apply multinomial sampling to the filtered candidate set $\Omega'_u(i)$ to select nodes to establish connections with player u , with the probability of selecting the next node being:

$$P_i^u(j) = \frac{\text{score}_u(i, j)}{\sum_{k \in \Omega'_u(i)} \text{score}_u(i, k)} \quad (17)$$

This probabilistic strategy ensures relevance while promoting diversity among eligible recommendations.

4.3.5 Multi-factor Path Weight Calculation. To quantify the influence of explored nodes, we define a path-based weighting scheme. Given a path $i_1 \rightarrow i_2 \rightarrow \dots \rightarrow i_l$, the edge weight from user u to node i_l is computed as:

$$w_{i_l}^u = \begin{cases} \hat{t}_{i_1}^u & \text{if } l = 1 \\ \hat{t}_{i_1}^u \cdot \text{sim}_m(i_1, i_2) & \text{if } l = 2 \\ \hat{t}_{i_1}^u \cdot \prod_{k=1}^{l-1} \text{sim}_m(i_k, i_{k+1}) & \text{if } l \geq 3 \end{cases} \quad (18)$$

This formulation combines the user's initial interest (playtime) with accumulated multimodal similarity along the path, balancing personalization and exploratory diversity during random walks. We adopt weighted LightGCN [20] to compute the final multimodal random walks embeddings. Incorporating our path weights, node embeddings are updated as:

$$\mathbf{e}_u^{(l)} = \sum_{i \in \mathcal{N}_u} \frac{w_i^u}{\sqrt{|\mathcal{N}_u||\mathcal{N}_i|}} \cdot \mathbf{e}_i^{(l-1)}, \mathbf{e}_i^{(l)} = \sum_{u \in \mathcal{N}_i} \frac{w_i^u}{\sqrt{|\mathcal{N}_i||\mathcal{N}_u|}} \cdot \mathbf{e}_u^{(l-1)} \quad (19)$$

where \mathcal{N}_u and \mathcal{N}_i are the neighbor sets of user u and game i respectively, and w_i^u is the path-based edge weight. This weighted aggregation emphasizes nodes with stronger semantic similarity and user interest. After L propagation layers, the final user and item embeddings are defined as $\mathbf{e}_u^{MRW} = \mathbf{e}_u^{(L)}$, $\mathbf{e}_i^{MRW} = \mathbf{e}_i^{(L)}$.

4.3.6 Theoretical Analysis: Diversity Optimization via Controlled Martingale Process. We model the MRW as a controlled stochastic process. Specifically, a user's exploration path on the game graph is treated as a discrete-time random walk, where diversity at time step t is defined by the number of distinct categories covered so far.

Let $P = \{i_1, i_2, \dots, i_t\}$ denote the set of game nodes visited by the user up to step t , and let each node i belong to a set of categories $c(i) \subseteq C$. We define the diversity metric at time t as:

$$C_t(P) := \left| \bigcup_{j \in P} c(j) \right| \quad (20)$$

Let \mathcal{F}_t represent the filtration containing all historical information up to time t . We model the sequence $\{C_t(P)\}_{t=0}^T$ as a controlled stochastic process.

To encourage exploration into underrepresented categories, the transition probability $P_i^u(j)$ incorporates a category balance coefficient. The marginal contribution of a candidate node j is quantified by its expected diversity gain, defined as:

$$\Delta C(j) := \mathbb{E}[C_t(P \cup \{j\}) - C_t(P) \mid \mathcal{F}_t] \quad (21)$$

Here, $\mathbb{E}[\cdot \mid \mathcal{F}_t]$ denotes the conditional expectation given the current path history \mathcal{F}_t .

Given the definition of $C_t(P)$, it follows that:

$$\Delta C(j) = |c(j) \setminus c(P)| \quad (22)$$

where $c(P) = \bigcup_{i \in P} c(i)$ denotes the set of categories already covered by the path P , and \setminus denotes the set difference operator. $\Delta C(j)$ represents the number of new categories introduced by node j .

Substituting this into the expectation, we obtain:

$$\mathbb{E}[C_{t+1}(P) - C_t(P) \mid \mathcal{F}_t] = \sum_{j \in \Omega'_u(i)} P_i^u(j) \cdot \Delta C(j) \quad (23)$$

This confirms that $\{C_t(P)\}$ forms a controlled submartingale [41], indicating that the expected diversity is non-decreasing over time. Now let T denote the stopping time when the user's exploration terminates. By the optional stopping theorem, the expected diversity at time T satisfies:

$$\mathbb{E}[C_T(P)] = \mathbb{E}[C_0(P)] + \mathbb{E}\left[\sum_{t=0}^{T-1} \sum_{j \in \Omega'_u(i)} P_i^u(j) \cdot \Delta C(j)\right] \quad (24)$$

This provides an exact characterization of expected diversity gains, showing that MRW's category-aware mechanism systematically guides users to novel semantic regions. The node sequence P forms a subgraph revealing behaviors like cross-category transitions, highlighting that diversity is structured, controllable graph traversal.

4.4 Balance Module

To integrate the above two modules, we introduce a balance module that dynamically controls their relative influence. This design enables flexible trade-offs between recommendation accuracy and diversity. We define the fused user and item embeddings as:

$$\mathbf{e}_u = \alpha \cdot \mathbf{e}_u^{IIE} + \mathbf{e}_u^{MRW}, \mathbf{e}_i = \alpha \cdot \mathbf{e}_i^{IIE} + \mathbf{e}_i^{MRW} \quad (25)$$

where $\alpha > 0$ is a tunable parameter: a larger α emphasizes behavior-driven personalization (IIE), enhancing accuracy; a smaller α favors semantic-guided exploration (MRW), boosting diversity.

We introduce a balance-aware loss function inspired by CPGRec that adaptively reweights negative samples to balance accuracy and diversity:

$$\mathcal{L}_{\text{balance}} = - \sum_{(u,i,j)} \log \sigma(s_{u,i} - \tilde{s}_{u,j}) \quad (26)$$

where $s_{u,i} = \mathbf{e}_u \cdot \mathbf{e}_i$ is the predicted score between user u and positive item i , and $\tilde{s}_{u,j}$ is the reweighted score of the negative item j :

$$\tilde{s}_{u,j} = s_{u,j} \cdot \left(\frac{1}{1 + e^{-s_{u,j} \cdot \zeta}} \cdot K \right) \quad (27)$$

where ζ controls the reweighting intensity, and K is a scaling factor. This formulation suppresses easy negatives and highlights informative ones, enhancing the model's training efficiency.

5 Experiments

5.1 Experimental Setup

5.1.1 Dataset. To evaluate DP^2Rec , we use the Steam platform—the most common data source in game recommendation research due to its large user base, rich interaction data, and public APIs. We collected the Steam dataset following DRGame's method and crawled

additional multimodal content (text and images). The filtered dataset contains 60,742 players, 7,726 games, 4,145,359 interactions, and 20 genres.

5.1.2 Evaluation Metrics. To evaluate recommendation quality, we measure accuracy using NDCG@K, Recall@K, Hit Ratio@K, and Precision@K. For diversity, we assess Coverage@K, which tracks the number of distinct game categories recommended. K is set to $\{5, 10, 20\}$.

5.1.3 Baselines. For **accuracy**, we include non-modal SCGRec [45] and RGCF [37], and accuracy-oriented CPGRec [23]. We also consider multimodal models—BM3 [54], SLMRec [36], MGNC [49], and SMORE [31], which leverage visual and textual game content. For **diversity**, we compare DP^2Rec with non-modal diversity-enhancing methods: EDUA [24], DGCN [52], DGRec [46], and diversity-focused CPGRec [23], as existing multimodal models do not explicitly target diversity. To assess **accuracy-diversity trade-off**, we highlight comparisons with balance-driven CPGRec [23], which optimizes both objectives.

5.2 Overall Performance

5.2.1 Performance Evaluation. We evaluated our proposed model, DP^2Rec , against several mainstream baseline models. The evaluation results are presented in Table 3, where the best and second-best results for each metric are highlighted in **bold** and underlined, respectively. Each result represents the average over five runs. Based on the results, we summarize the following key conclusions:

- **Advantages of DP^2Rec .** DP^2Rec outperforms diversity-based baselines, surpassing DGRec in **Coverage** by 13–21%, while maintaining competitive **NDCG** and **Hit Ratio** with the SOTA accuracy-focused method in game recommendation (RGCF). This shows it balances diversity and accuracy effectively.
- **Superiority on Dual Objectives.** The accuracy-focused DP^2Rec surpasses the best non-modal (CPGRec) and multimodal (SMORE) baselines by 17% and 11% on NDCG@5, respectively. Meanwhile, its diversity-focused variant exceeds the leading diversity model (DGRec) in Coverage by 30–50%, showing strong performance across both objectives.
- **Inherent Trade-off Between Accuracy and Diversity.** While SMORE leads among current accuracy-oriented models (excluding the accuracy-optimized DP^2Rec), it falls short on Coverage compared to diversity-focused models. This highlights the inherent trade-off—pursuing accuracy often comes at the cost of diversity, which can reduce recommendation variety.

5.2.2 Ablation Study. We perform ablation experiments to evaluate the contribution of four core submodules: (1) Dual-Interest Recognition via EM, (2) Dual-Interest Graph Learning, (3) Category-based Candidate Node Selection, and (4) Category Balance Coefficient. The corresponding variants are denoted as *w/o DIR*, *w/o DIGL*, *w/o CCNS*, and *w/o CBC*. Results in Table 4 offer the following insights:

1) **w/o DIR.** Without the EM-based dual-beta modeling, strong and weak interests remain entangled, limiting the model's capacity to capture true two-level user preferences and reducing accuracy. Additionally, the inability to leverage weak interests for cross-category exploration results in diminished diversity.

Table 3: Performance comparison of different recommendation methods.

Method	NDCG			Recall			Hit Ratio			Precision			Coverage			
	@5	@10	@20	@5	@10	@20	@5	@10	@20	@5	@10	@20	@5	@10	@20	
Accuracy-based Methods	SCGRec	0.2150	0.2283	0.2592	0.1341	0.2139	0.3208	0.6039	0.7663	0.8785	0.1867	0.1619	0.1331	3.2348	4.7899	6.5225
	RGCF	0.2628	0.2525	0.2607	0.1308	0.1901	0.2651	0.6423	0.7414	0.8179	0.2272	0.1822	0.1377	4.8120	6.7300	8.5280
	BM3	0.2344	0.2291	0.2415	0.1296	0.1912	0.2662	0.6304	0.7561	0.8439	0.1933	0.1513	0.1144	5.1940	6.8601	8.8693
	SLMRec	0.2621	0.2638	0.2841	0.1557	0.2328	0.3279	0.6780	0.8076	0.8954	0.2205	0.1785	0.1359	5.4632	6.8632	8.3717
	MGCN	0.2615	0.2595	0.2777	0.1437	0.2163	0.3090	0.6739	0.7939	0.8720	0.2252	0.1822	0.1407	4.9905	6.6059	8.2074
	SMORE	0.2741	0.2720	0.2903	0.1558	0.2302	0.3242	0.6970	0.8112	0.8871	0.2320	0.1860	0.1427	5.0582	6.5760	8.1871
Accuracy-focused Methods	CPGRec	0.2594	0.2591	0.2791	0.1444	0.2213	0.3186	0.6675	0.7933	0.8783	0.2227	0.1812	0.1398	5.1467	6.6551	8.3357
	Accuracy-focused D^2Rec	0.3032	0.2978	0.3169	0.1704	0.2504	0.3503	0.7280	0.8342	0.9008	0.2570	0.2050	0.1555	5.2076	6.6935	8.2847
Diversity-based Methods	EDUA	0.1421	0.1455	0.1571	0.0908	0.1390	0.1897	0.4500	0.5700	0.6600	0.1200	0.0935	0.0685	5.6850	7.4700	9.1200
	DGCN	0.0963	0.0900	0.0799	0.0601	0.1031	0.1689	0.3668	0.5327	0.6925	0.0923	0.0850	0.0732	5.3660	7.4756	9.2498
	DGRec	0.1134	0.1075	0.0931	0.0959	0.1529	0.2165	0.4988	0.6660	0.7860	0.1298	0.1128	0.0896	5.8438	7.9892	9.9746
	Diversity-focused CPGRec	0.2257	0.2203	0.2343	0.1208	0.1851	0.2671	0.6076	0.7376	0.8341	0.1941	0.1544	0.1163	6.0860	7.6519	9.3547
Balance-based Methods	Diversity-focused D^2Rec	0.2346	0.2322	0.2518	0.1273	0.1942	0.2862	0.6298	0.7626	0.8586	0.1953	0.1577	0.1241	7.9133	11.2227	14.7972
	CPGRec (trade-off model)	0.2347	0.2333	0.2532	0.1272	0.1952	0.2893	0.6326	0.7557	0.8512	0.2027	0.1666	0.1308	5.6811	7.0266	8.4586
	D^2Rec (trade-off model)	0.2518	0.2497	0.2688	0.1403	0.2100	0.3038	0.6587	0.7846	0.8722	0.2121	0.1712	0.1330	6.6499	9.3627	12.0228

Table 4: Ablation study. We show D^2Rec 's performance when removing each of the submodules.

Method	NDCG			Recall			Hit Ratio			Precision			Coverage		
	@5	@10	@20	@5	@10	@20	@5	@10	@20	@5	@10	@20	@5	@10	@20
Accuracy-focused D^2Rec	0.3032	0.2978	0.3169	0.1704	0.2504	0.3503	0.7280	0.8342	0.9008	0.2570	0.2050	0.1555	5.2076	6.6935	8.2847
	0.2877	0.2844	0.3030	0.1609	0.2395	0.3381	0.7069	0.8190	0.8918	0.2446	0.1962	0.1496	5.0780	6.5365	8.1875
	0.2743	0.2694	0.2848	0.1454	0.2180	0.3083	0.6785	0.7841	0.8561	0.2370	0.1914	0.1456	4.8388	6.2882	7.9043
Diversity-focused D^2Rec	0.2346	0.2322	0.2518	0.1273	0.1942	0.2862	0.6298	0.7626	0.8586	0.1953	0.1577	0.1241	7.9133	11.2227	14.7972
	0.2528	0.2506	0.2678	0.1403	0.2120	0.3011	0.6606	0.7848	0.8705	0.2122	0.1703	0.1306	5.6348	6.9416	8.5220
	0.2493	0.2472	0.2629	0.1388	0.2095	0.2952	0.6572	0.7817	0.8653	0.2101	0.1686	0.1282	5.9722	7.4482	11.2468

2) **w/o DIGL.** Removing contrastive learning causes the model to focus solely on strong interests, reducing both accuracy and diversity. This over-reliance overlooks weak interests, which not only contribute to finer-grained preference modeling but also promote exploration across categories.

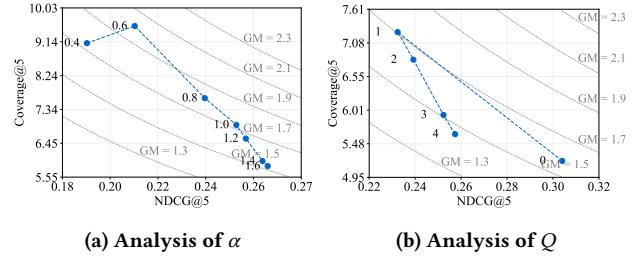
3) **w/o CCNS.** This submodule enables cross-category exploration via semantic similarity. Its removal increases accuracy slightly, by focusing on familiar regions, but significantly reduces diversity, confirming its importance in discovering semantically relevant games outside core categories.

4) **w/o CBC.** Dropping this submodule leads to more recommendations from dominant categories, improving accuracy but sharply reducing diversity. This confirms its role in ensuring balanced exposure across categories and preventing over-concentration.

5.2.3 Parameter Sensitivity.

Sensitivity Analysis on α . As shown in Figure 3(a), increasing α from 0.4 to 1.6 steadily boosts NDCG@5 (from 0.19 to 0.26), reflecting improved capture of users' core preferences. However, this gain in accuracy comes with a drop in diversity, as Coverage@5 declines from 9.10 to 5.84—indicating a narrowing of recommendation scope. To quantify this trade-off, we compute the **geometric mean (GM)** of NDCG@5 and Coverage@5 ($GM = \sqrt{NDCG@5 \times Coverage@5}$). The contour map reveals that the optimal balance occurs at $\alpha = 0.6$, where GM peaks at **2.00**.

Sensitivity Analysis on Q . As shown in Figure 3(b), increasing Q from 0 to 4 causes NDCG@5 to initially drop from 0.30 to 0.24 before gradually recovering to 0.26. Meanwhile, Coverage@5 peaks

**Figure 3: Analysis of key hyperparameters α and Q .**

at $Q = 1$ (7.25) and then slightly declines. This suggests that moderate category control enhances structural diversity, while excessive flexibility narrows recommendations. To evaluate the overall performance, we compute the GM of NDCG@5 and Coverage@5. The contour plot shows the highest GM at $Q = 1$, reaching **1.71**, indicating it offers the best balance between accuracy and diversity.

6 Conclusion

In this work, we address two core challenges in game recommendation: the underutilization of playtime as a behavioral signal and the limited capacity to promote cross-category diversity based on multimodal similarity. To enhance accuracy, we introduce a *IIE* module that distinguishes strong and weak interests based on playtime distributions. To promote diversity, we design a *MRW* that incorporates both behavioral and semantic signals to guide user exploration across game categories.

Acknowledgments

This work was partially supported by the National Natural Science Foundation of China under Project No. 62202122 and No. 62073272; by the Guangdong Basic and Applied Basic Research Foundation under Grant No. 2024A1515011949; by the Shenzhen Science and Technology Program under Grant No. GXWD20231130110308001, JCYJ20240813104843058, and JCYJ20240813104837050; by the Shenzhen Education Science “14th Five-Year Plan” 2023 Annual Project on Artificial Intelligence Special Project under Grant No. rgzn23001; by the Guangdong Province Higher Education Research and Reform Project under Grant No. YueJiaoGaoHan(2024) No. 9; and by the Shenzhen Key Laboratory of Internet Information Collaboration under Grant No. ZDSY20120613125016389.

References

- [1] Mustafa Abdool, Malay Haldar, Prashant Ramanathan, Tyler Sax, Lanbo Zhang, Aamir Manaswala, Lynn Yang, Bradley Turnbull, Qing Zhang, and Thomas Legrand. 2020. Managing diversity in airbnb search. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. 2952–2960.
- [2] Syed Muhammad Anwar, Talha Shahzad, Zunaira Sattar, Rahma Khan, and Muhammad Majid. 2017. A game recommender system using collaborative filtering (GAMBIT). In *2017 14th International Bhurban Conference on Applied Sciences and Technology (IBCAST)*. IEEE, 328–332.
- [3] Shuai Bai, Keqin Chen, Xuejing Liu, Jialin Wang, Wenbin Ge, Sibo Song, Kai Dang, Peng Wang, Shijie Wang, Jun Tang, et al. 2025. Qwen2. 5-vl technical report. *arXiv preprint arXiv:2502.13923* (2025).
- [4] Yong Bai, Rui Xiang, Kaiyuan Li, Yongxiang Tang, Yanhua Cheng, Xialong Liu, Peng Jiang, and Kun Gai. 2025. CHIME: A Compressive Framework for Holistic Interest Modeling. *arXiv preprint arXiv:2504.06780* (2025).
- [5] C BharathiPriya, Akash Sreenivasu, and Sampath Kumar. 2021. Online Video Game Recommendation System Using Content And Collaborative Filtering Techniques. In *2021 International Conference on Advancements in Electrical, Electronics, Communication, Computing and Automation (ICAEECA)*. IEEE, 1–7.
- [6] Jaime Carbonell and Jade Goldstein. 1998. The use of MMR, diversity-based reranking for reordering documents and producing summaries. In *Proceedings of the 21st annual international ACM SIGIR conference on Research and development in information retrieval*. 335–336.
- [7] Loïc Caroux. 2023. Presence in video games: A systematic review and meta-analysis of the effects of game design choices. *Applied Ergonomics* 107 (2023), 103936.
- [8] Chaochao Chen, Yizhao Zhang, Yuyuan Li, Jun Wang, Lianyong Qi, Xiaolong Xu, Xiaolin Zheng, and Jianwei Yin. 2024. Post-training attribute unlearning in recommender systems. *ACM Transactions on Information Systems* 43, 1 (2024), 1–28.
- [9] Jiaju Chen, Chongming Gao, Shuai Yuan, Shuchang Liu, Qingpeng Cai, and Peng Jiang. 2024. DLCRec: A Novel Approach for Managing Diversity in LLM-Based Recommender Systems. *arXiv preprint arXiv:2408.12470* (2024).
- [10] Jiaju Chen, Chongming Gao, Shuai Yuan, Shuchang Liu, Qingpeng Cai, and Peng Jiang. 2025. DLCRec: A Novel Approach for Managing Diversity in LLM-Based Recommender Systems. In *Proceedings of the Eighteenth ACM International Conference on Web Search and Data Mining*. 857–865.
- [11] Laming Chen, Guoxin Zhang, and Eric Zhou. 2018. Fast greedy map inference for determinantal point process to improve recommendation diversity. *Advances in Neural Information Processing Systems* 31 (2018).
- [12] Germán Cheque, José Guzmán, and Denis Parra. 2019. Recommender systems for online video game platforms: The case of steam. In *Companion Proceedings of The 2019 World Wide Web Conference*. 763–771.
- [13] Junsu Cho, Dongmin Hyun, Dong won Lim, Hyeon jae Cheon, Hyoung-iel Park, and Hwanju Yu. 2023. Dynamic multi-behavior sequence modeling for next item recommendation. In *Proceedings of the AAAI conference on artificial intelligence*, Vol. 37. 4199–4207.
- [14] Erica Coppolillo, Giuseppe Manco, and Aristides Gionis. 2024. Relevance meets diversity: A user-centric framework for knowledge exploration through recommendations. In *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*. 490–501.
- [15] Arthur P Dempster, Nan M Laird, and Donald B Rubin. 1977. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the royal statistical society: series B (methodological)* 39, 1 (1977), 1–22.
- [16] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. Bert: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 conference of the North American chapter of the association for computational linguistics: human language technologies, volume 1 (long and short papers)*. 4171–4186.
- [17] Tomislav Duricic, Dominik Kowald, Emanuel Lacic, and Elisabeth Lex. 2023. Beyond-accuracy: a review on diversity, serendipity, and fairness in recommender systems based on graph neural networks. *Frontiers in big data* 6 (2023), 1251072.
- [18] Farzad Eskandanian and Bamshad Mobasher. 2020. Using stable matching to optimize the balance between accuracy and diversity in recommendation. In *Proceedings of the 28th ACM Conference on User Modeling, Adaptation and Personalization*. 71–79.
- [19] Lu Gan, Diana Nurbakova, Léa Laporte, and Sylvie Calabretto. 2020. Enhancing recommendation diversity using determinantal point processes on knowledge graphs. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*. 2001–2004.
- [20] Xiangnan He, Kuan Deng, Xiang Wang, Yan Li, Yongdong Zhang, and Meng Wang. 2020. LightGCN: Simplifying and powering graph convolution network for recommendation. In *Proceedings of the 43rd International ACM SIGIR conference on research and development in Information Retrieval*. 639–648.
- [21] Yanhua Huang, Weikun Wang, Lei Zhang, and Ruiwen Xu. 2021. Sliding spectrum decomposition for diversified recommendation. In *Proceedings of the 27th ACM SIGKDD conference on knowledge discovery & data mining*. 3041–3049.
- [22] Fasiha Ikram and Humera Farooq. 2022. Multimedia Recommendation System for Video Game Based on High-Level Visual Semantic Features. *Scientific Programming* 2022, 1 (2022), 6084363.
- [23] Xiping Li, Jianghong Ma, Kangzhe Liu, Shanshan Feng, Haijun Zhang, and Yutong Wang. 2024. Category-based and Popularity-guided Video Game Recommendation: A Balance-oriented Framework. In *Proceedings of the ACM on Web Conference 2024*. 3734–3744.
- [24] Lile Liang, Tieyan Qian, Qing Li, and Hongzhi Yin. 2021. Enhancing domain-level and user-level adaptivity in diversified recommendation. In *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 747–756.
- [25] Zihan Lin, Hui Wang, Jingshu Mao, Wayne Xin Zhao, Cheng Wang, Peng Jiang, and Ji-Rong Wen. 2022. Feature-aware diversified re-ranking with disentangled representations for relevant recommendation. In *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*. 3327–3335.
- [26] Chengliang Liu, Jie Wen, Zhihao Wu, Xiaoling Luo, Chao Huang, and Yong Xu. 2023. Information recovery-driven deep incomplete multiview clustering network. *IEEE Transactions on Neural Networks and Learning Systems* 35, 11 (2023), 15442–15452.
- [27] Haobing Liu, Jianyu Ding, Yanmin Zhu, Feilong Tang, Jiadi Yu, Ruobing Jiang, and Zhongwen Guo. 2023. Modeling multi-aspect preferences and intents for multi-behavioral sequential recommendation. *Knowledge-Based Systems* 280 (2023), 111013.
- [28] Kangzhe Liu, Jianghong Ma, Shanshan Feng, Haijun Zhang, and Zhao Zhang. 2024. DRGame: Diversified Recommendation for Multi-category Video Games with Balanced Implicit Preferences. In *International Conference on Database Systems for Advanced Applications*. Springer, 254–263.
- [29] Mingyang Lv, Xiangfeng Liu, and Yuanbo Xu. 2025. Dynamic Multi-Interest Graph Neural Network for Session-Based Recommendation. (2025).
- [30] Frank J Massey Jr. 1951. The Kolmogorov-Smirnov test for goodness of fit. *Journal of the American statistical Association* 46, 253 (1951), 68–78.
- [31] Rongqing Kenneth Ong and Andy WH Khong. 2025. Spectrum-based Modality Representation Fusion Graph Convolutional Network for Multimodal Recommendation. In *Proceedings of the Eighteenth ACM International Conference on Web Search and Data Mining*. 773–781.
- [32] Kenny Peng, Manish Raghavan, Emma Pierson, Jon Kleinberg, and Nikhil Garg. 2024. Reconciling the accuracy-diversity trade-off in recommendations. In *Proceedings of the ACM Web Conference 2024*. 1318–1329.
- [33] Javier Pérez-Marcos, Lucía Martín-Gómez, Diego M Jiménez-Bravo, Vivian F López, and María N Moreno-García. 2020. Hybrid system for video game recommendation based on implicit ratings and social networks. *Journal of Ambient Intelligence and Humanized Computing* 11 (2020), 4525–4535.
- [34] Ladislav Peska and Patrik Dokoupil. 2022. Towards results-level proportionality for multi-objective recommender systems. In *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 1963–1968.
- [35] Youchen Sun, Zhu Sun, Yingpeng Du, Jie Zhang, and Yew Soon Ong. 2024. Self-Supervised Denoising through Independent Cascade Graph Augmentation for Robust Social Recommendation. In *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*. 2806–2817.
- [36] Zhulin Tao, Xiaohao Liu, Yewei Xia, Xiang Wang, Lifang Yang, Xianglin Huang, and Tat-Seng Chua. 2022. Self-supervised learning for multimedia recommendation. *IEEE Transactions on Multimedia* 25 (2022), 5107–5116.
- [37] Changxin Tian, Yuxiang Xie, Yaliang Li, Nan Yang, and Wayne Xin Zhao. 2022. Learning to denoise unreliable interactions for graph collaborative filtering. In *Proceedings of the 45th international ACM SIGIR conference on research and development in information retrieval*. 122–132.

- [38] Yijun Tian, Chuxu Zhang, Zhichun Guo, Chao Huang, Ronald Metoyer, and Nitesh V Chawla. 2022. RecipeRec: A heterogeneous graph learning model for recipe recommendation. *arXiv preprint arXiv:2205.14005* (2022).
- [39] Wei Wei, Lianghao Xia, and Chao Huang. 2023. Multi-relational contrastive learning for recommendation. In *Proceedings of the 17th ACM conference on recommender systems*. 338–349.
- [40] Jie Wen, Gehui Xu, Zhenyan Tang, Wei Wang, Lunke Fei, and Yong Xu. 2023. Graph regularized and feature aware matrix factorization for robust incomplete multi-view clustering. *IEEE Transactions on Circuits and Systems for Video Technology* 34, 5 (2023), 3728–3741.
- [41] David Williams. 1991. *Probability with martingales*. Cambridge university press.
- [42] Shuai Xiao and Zaifan Jiang. 2024. DivNet: Diversity-Aware Self-Correcting Sequential Recommendation Networks. In *Proceedings of the 33rd ACM International Conference on Information and Knowledge Management*. 5000–5006.
- [43] Yuanbo Xu, Fuzhen Zhuang, En Wang, Chaozhuo Li, and Jie Wu. 2024. Learning without Missing-At-Random Prior Propensity-A Generative Approach for Recommender Systems. *IEEE Transactions on Knowledge and Data Engineering* (2024).
- [44] Liangwei Yang, Zhiwei Liu, Yu Wang, Chen Wang, Ziwei Fan, and Philip S Yu. 2022. Large-scale personalized video game recommendation via social-aware contextualized graph neural network. In *Proceedings of the ACM Web Conference 2022*. 3376–3386.
- [45] Liangwei Yang, Zhiwei Liu, Yu Wang, Chen Wang, Ziwei Fan, and Philip S. Yu. 2022. Large-scale Personalized Video Game Recommendation via Social-aware Contextualized Graph Neural Network. In *WWW '22: The ACM Web Conference 2022, Virtual Event, Lyon, France, April 25 - 29, 2022*. ACM, 3376–3386. doi:10.1145/3485447.3512273
- [46] Liangwei Yang, Shengjie Wang, Yunzhe Tao, Jiankai Sun, Xiaolong Liu, Philip S Yu, and Taiqing Wang. 2023. Dgrec: Graph neural network for recommendation with diversified embedding generation. In *Proceedings of the sixteenth ACM international conference on web search and data mining*. 661–669.
- [47] Rui Ye, Yuqing Hou, Te Lei, Yunxing Zhang, Qing Zhang, Jiale Guo, Huaiwen Wu, and Hengliang Luo. 2021. Dynamic graph construction for improving diversity of recommendation. In *Proceedings of the 15th ACM Conference on Recommender Systems*. 651–655.
- [48] Qing Yin, Hui Fang, Zhu Sun, and Yew-Soon Ong. 2024. A Simple Yet Effective Approach for Diversified Session-Based Recommendation. *arXiv preprint arXiv:2404.00261* (2024).
- [49] Penghang Yu, Zhiyi Tan, Guanming Lu, and Bing-Kun Bao. 2023. Multi-view graph convolutional network for multimedia recommendation. In *Proceedings of the 31st ACM international conference on multimedia*. 6576–6585.
- [50] Juja Zhao, Wang Wenjie, Yiyuan Xu, Teng Sun, Fuli Feng, and Tat-Seng Chua. 2024. Denoising diffusion recommender model. In *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 1370–1379.
- [51] Yuying Zhao, Minghua Xu, Huiyuan Chen, Yuzhong Chen, Yiwei Cai, Rashidul Islam, Yu Wang, and Tyler Derr. 2024. Can one embedding fit all? a multi-interest learning paradigm towards improving user interest diversity fairness. In *Proceedings of the ACM Web Conference 2024*. 1237–1248.
- [52] Yu Zheng, Chen Gao, Liang Chen, Depeng Jin, and Yong Li. 2021. DGCN: Diversified recommendation with graph convolutional networks. In *Proceedings of the Web Conference 2021*. 401–412.
- [53] Yongsen Zheng, Guohua Wang, Yang Liu, and Liang Lin. 2024. Diversity Matters: User-Centric Multi-Interest Learning for Conversational Movie Recommendation. In *Proceedings of the 32nd ACM International Conference on Multimedia*. 9515–9524.
- [54] Xin Zhou, Hongyu Zhou, Yong Liu, Zhiwei Zeng, Chunyan Miao, Pengwei Wang, Yuan You, and Feijun Jiang. 2023. Bootstrap latent representations for multimodal recommendation. In *Proceedings of the ACM web conference 2023*. 845–854.
- [55] Cai-Nicolas Ziegler, Sean M McNee, Joseph A Konstan, and Georg Lausen. 2005. Improving recommendation lists through topic diversification. In *Proceedings of the 14th international conference on World Wide Web*. 22–32.