Inferring Targets from Calibrated Hesitations via Mutual Information Maximization in Multi-Behavior Recommendation

Cheng Li
Anhui Normal University
Wuhu, China
jxxylc@ahnu.edu.cn

Yong Xu*
Anhui Normal University
Wuhu, China
yxull@ahnu.edu.cn

Suhua Tang*
The University of Electro-Communications
Tokyo, Japan
shtang@uec.ac.jp

Xin He*
Anhui Normal University
Wuhu, China
xin.he@ahnu.edu.cn

Jianfeng Sun
Anhui Normal University
Wuhu, China
sunjf@ahnu.edu.cn

Jinde Cao
Southeast University
Nanjing, China
jdcao@seu.edu.cn

Abstract

Multi-behavior recommendation enriches user preference modeling by incorporating diverse auxiliary interactions. However, most existing methods simply treat interactions without target behaviors as absolute negative feedback. This strategy ignores an important intermediate state known as user hesitation, where users exhibit strong intent but fail to complete the final conversion due to various reasons. Consequently, models cannot distinguish true disinterest from intended but hesitant behavior, which introduces substantial noise into preference modeling. To address this issue, we propose a novel framework named **C**alibrated **H**esitation **A**nalysis for Multi-Behavior **R**ecommendation via **M**utual Information Maximization (CHARM). Specifically, we aggregate auxiliary behaviors that lead to successful conversions into latent intent representations and train an inference network by maximizing the mutual information between these intents and observed target behaviors. We then apply this network to auxiliary behaviors without conversion, under the assumption that the conversion had occurred, in order to infer latent conversion probabilities and identify high-intent hesitation candidates. Furthermore, to distinguish genuine hesitation from interaction termination caused by competing item choices, we design a competitor substitution penalty strategy to refine hesitation confidence scores. Finally, the calibrated hesitation set is incorporated into the recommendation process to improve ranking quality. Extensive experiments on three real-world datasets demonstrate that CHARM consistently outperforms existing state-of-the-art methods. The source code is available at https://github.com/city59/CHARM.

CCS Concepts

• **Information systems → Recommender systems**.

Keywords

Multi-Behavior Recommendation, Graph Neural Networks, Mutual Information, Collaborative Filtering

ACM Reference Format:
Cheng Li, Yong Xu, Suhua Tang, Xin He, Jianfeng Sun, and Jinde Cao. 2026. Inferring Targets from Calibrated Hesitations via Mutual Information Maximization in Multi-Behavior Recommendation. In *Proceedings of the 49th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '26), July 20–24, 2026, Melbourne, VIC, Australia.* ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3805712.3809574

1 Introduction

Recommender systems [18, 66] have become indispensable tools for alleviating information overload [1]. They serve as the core engines of personalized content delivery in electronic commerce [30] and online platforms [45]. Within existing paradigms, collaborative filtering (CF) [20, 21, 46], which captures user preferences by mining latent associations in user-item interaction matrices [41, 63], has long dominated the field. However, traditional CF methods primarily rely on single target behaviors[1] such as purchases for model training [54]. The inherent scarcity of such supervisory signals makes it difficult for these models to learn high-quality user representations [27, 58], particularly in data sparsity and cold-start scenarios. In practice, as interaction ecosystems continue to evolve, user behavior patterns have become increasingly diverse, encompassing various forms such as browsing, favoriting, and adding items to carts [62]. These heterogeneous interactions not only provide complementary multi-perspective information for preference modeling, but also involve complex dynamics of user interest evolution and behavior conversion [35, 56]. As such, they offer valuable opportunities to overcome the limitations of single-behavior modeling and to construct more expressive and robust recommendation [34].

Multi-behavior recommendation [60, 67] enhances user preference modeling through high-frequency auxiliary behaviors and learned dependencies between auxiliary and target behaviors [23]. These auxiliary signals provide complementary perspectives on

*Corresponding authors.

SIGIR '26, Melbourne, VIC, Australia

ACM ISBN 979-8-4007-2599-9/2026/07
https://doi.org/10.1145/3805712.3809574

[1] In this paper, "target behaviors" and "target actions" are used interchangeably. "Successful conversion" refers to the occurrence of a target behavior.

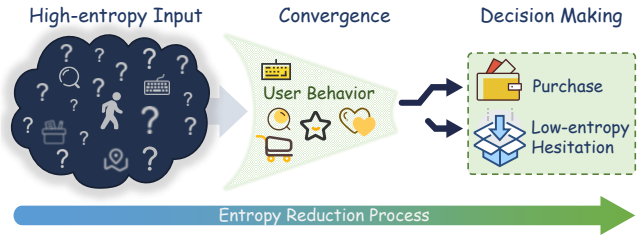

Figure 1: User hesitation from an information-theoretic perspective of entropy reduction.

user preferences and help alleviate data sparsity while improving representation quality [31]. Despite the effectiveness of existing methods, most approaches [9, 24] implicitly treat all non-converted auxiliary interactions as homogeneous negative feedback or weakly-supervised signals at the target behavior prediction stage. Although auxiliary behaviors are exploited at the feature or multi-task level [28], their non-conversion states are rarely further distinguished or explicitly modeled. This simplification may overlook heterogeneous preference signals embedded in different auxiliary behavior failure patterns, thereby neglecting a critical psychological state in the user behavior process, namely user hesitation [70]. In practice, users may frequently interact with a particular item, indicating strong purchase intent, yet ultimately fail to complete the transaction due to factors such as price sensitivity or decision paralysis. Treating these hesitation cases as purely negative samples introduces false-negative noise, confuses the model, and suppresses the ranking of potentially convertible items [43].

Although highly appealing from an intuitive perspective, explicitly modeling user hesitation faces two major theoretical and practical challenges. First, due to the unobservability of hesitation signals, it is difficult to rigorously define hesitation from a theoretical standpoint. As an implicit intermediate psychological state, hesitation lacks explicit label supervision. This requires models to infer unobserved states in an unsupervised manner by accurately quantifying the discrepancy between users accumulated interaction intent and the absence of final conversion behavior [2]. Second, substantial causal ambiguity [16, 49] exists behind non-converted behaviors, making it highly challenging to distinguish genuine hesitation signals from noise. Frequent interactions with an item that do not culminate in a purchase do not necessarily indicate hesitation, as users may instead discover and purchase a superior competing item during the decision process [64]. In such cases, the non-conversion of the original item effectively reflects an implicit rejection rather than hesitation. Without effectively disentangling this substitution effect, models are prone to misclassifying items abandoned due to satisfied demand as high-potential hesitation candidates, thereby introducing severe false-positive noise. This not only degrades recommendation accuracy but also harms user experience by repeatedly exposing users to redundant items [6].

To address these issues, we propose the **C**alibrated **H**esitation **A**nalysis for Multi-Behavior **R**ecommendation via **M**utual Information Maximization (CHARM). As illustrated in Fig. 1, we view user decision-making as progressively reducing the uncertainty of the eventual target action. Formally, let Z denote a latent intent inferred from auxiliary behaviors and let E denote the target-domain interaction. Let $H(\cdot)$ be the Shannon entropy, $H(\cdot \mid \cdot)$ be the conditional entropy, and $I(\cdot;\cdot)$ denote the mutual information between two random variables. By the information-theoretic identity $I(Z;E) = H(E) - H(E \mid Z)$ [7, 26], a larger mutual information implies a lower conditional entropy of the target action given the inferred intent (i.e., higher intent certainty w.r.t. the target). Based on this principle, CHARM identifies hesitation as auxiliary behaviors that induce highly certain intent about a target action but do not culminate in conversion. Specifically, we first learn a target-aware intent representation from auxiliary behaviors that lead to successful conversions, and train a dependency estimator by maximizing a variational lower bound of $I(Z;E)$ between the inferred intents and observed target behaviors. Next, we perform virtual interaction modeling to evaluate non-converted auxiliary behaviors by constructing their hypothetical target-action representations and scoring their dependency with the inferred intent. The resulting score is mapped to a continuous hesitation confidence and used to weight candidate hesitation signals. Finally, to reduce spurious signals caused by substitution, we introduce a competitor substitution penalty that downweights demands likely satisfied by competing items, yielding purified hesitation cues that are incorporated into the recommendation objective. The main contributions of this work are summarized as follows:

- We revisit the user decision-making process in recommendation from an information-theoretic entropy reduction perspective and highlight the importance of explicitly modeling user hesitation for mitigating false-negative noise.
- We propose the CHARM framework, which novelly integrates mutual information maximization with virtual interaction modeling to effectively discover and quantify implicit hesitation intent in an unsupervised manner. In addition, we design a competitor-aware mechanism to eliminate spurious hesitation signals caused by competing items.
- Extensive experiments on three real-world datasets demonstrate that our approach significantly outperforms existing state-of-the-art baseline models.

2 Related Work

2.1 Multi-Behavior Recommendation

Multi-behavior recommendation enriches sparse supervisory signals by integrating high-frequency auxiliary behaviors and have become a key paradigm for addressing data sparsity [53]. Early studies primarily relied on matrix factorization. For example, CMF [44] alleviates data sparsity by jointly factorizing interaction matrices of target and auxiliary behaviors and sharing latent representations across behaviors. In recent years, due to their strong capability in capturing high-order connectivity, graph neural networks (GNN) have gradually become dominant in this area. For instance, NMTR [15] pioneers a neural multi-task learning framework that constructs a cascaded prediction chain to explicitly model hierarchical dependencies among behaviors. Building on this idea, MBGCN [23] further extends behavior dependencies to graph structures and refines user representations on heterogeneous graphs through behavior-aware graph convolutional propagation. Beyond

architectural innovations, CML [52] explores the learning paradigm by aligning representations across different behavior views via contrastive learning, which significantly enhances model robustness and generalization. However, most of the aforementioned methods rely on a uniform negative sampling assumption and homogenize all samples without target interactions as absolute negative feedback. This coarse-grained treatment overlooks a critical intermediate state known as user hesitation, causing samples with latent intent but without final conversion to be mislabeled as negative instances.

2.2 Causal Inference in Recommendation

Recommender systems rely on observational data for model training. However, such data are inevitably subject to selection bias and confounding bias. To infer users true preferences from biased observations, causal inference has been introduced as a core solution in this domain [16]. Early studies adopted inverse propensity scoring [4] and doubly robust techniques [14] to eliminate systematic bias induced by popularity or exposure mechanisms through sample reweighting. For example, MF-IPS [42] formulates recommendation as a treatment effect estimation problem and constructs an unbiased estimator by weighting observed interactions with propensity scores. More recently, counterfactual reasoning has attracted increasing attention and is often used to generate augmented samples to alleviate data sparsity. For instance, CASR [50] synthesizes counterfactual sequence data by modifying interaction items in user sequences, thereby improving the performance of sequential recommendation models under sparse settings. MACR [51] adopts a model-agnostic causal inference framework that decouples user interest from item popularity through multi-task learning and removes the direct causal effect of popularity during inference. However, most existing causal debiasing methods focus on correcting exposure bias, namely distinguishing between non-exposure and non-interaction, while rarely investigating the deeper causal mechanisms underlying exposed but non-interacted behaviors.

3 Preliminaries

Let $\mathcal{U} = \{u_1, u_2, \ldots, u_M\}$ and $\mathcal{I} = \{i_1, i_2, \ldots, i_N\}$ denote the sets of M users and N items, respectively. The set of behaviors is defined as $\mathcal{B} = \{1, 2, \ldots, B\}$, which includes one target behavior b^* (e.g., purchase) and $B-1$ auxiliary behaviors $\mathcal{B}_{\text{aux}}$ (e.g., view, cart, favorite), such that $\mathcal{B} = \mathcal{B}_{\text{aux}} \cup \{b^*\}$. For each behavior type $b \in \mathcal{B}$, we define a user-item interaction graph as $\mathcal{G}_b = (\mathcal{V}, \mathcal{E}_b)$, where $\mathcal{V} = \mathcal{U} \cup \mathcal{I}$ is the set of nodes, and $\mathcal{E}_b$ denotes the set of observed interaction edges. Based on these interactions, we categorize user-item pairs into three distinct subsets: (1) **Positive samples** $\mathcal{E}^+ = \{(u,i) \mid (u,i) \in \mathcal{E}_{b^*}\}$, where the user has performed the target behavior; (2) **Candidate hesitation space** $\mathcal{E}^? = \{(u,i) \mid (\exists b \in \mathcal{B}_{\text{aux}}, (u,i) \in \mathcal{E}_b) \wedge ((u,i) \notin \mathcal{E}_{b^*})\}$, which consists of pairs where the user has exhibited interest through auxiliary behaviors but has not converted; and (3) **Negative samples** $\mathcal{E}^- = (\mathcal{U} \times \mathcal{I}) \setminus \left(\bigcup_{b \in \mathcal{B}} \mathcal{E}_b\right)$, representing pairs with no observed interactions. Unlike previous works that indiscriminately treat $\mathcal{E}^?$ as static negative samples, we posit that this set contains heterogeneous signals, encompassing both genuine disinterest (e.g., the user viewed but disliked the item) and implicit hesitation (e.g., the user intended to purchase but suspended the decision). Consequently, our framework aims to disentangle these sub-populations within $\mathcal{E}^?$ to mitigate false-negative noise. The objective of the multi-behavior recommendation task is as follows: **Input:** The multi-behavior interaction graphs $\mathcal{G} = \bigcup_{b=1}^{B} \mathcal{G}_b$. **Output:** A predictive model which estimates the probability that a user $u \in \mathcal{U}$ will interact with an item $i \in \mathcal{I}$ under the target behavior type b^*.

4 Methodology

As shown in Fig. 2, CHARM follows three stages to handle unobservable intent, virtual interaction modeling, and disambiguation. First, we use a LightGCN [19] encoder on the heterogeneous behavior graph to learn target-aware representations, and aggregate conversion-related auxiliary behaviors into latent intent vectors as intent proxies. Second, we train a mutual information estimator on converted samples to quantify intent-action alignment, producing a dependency score indicative of successful conversion. Third, we apply this scorer to non-converted auxiliary interactions by constructing target-action representations and evaluating their dependency, where high scores under missing conversion suggest hesitation. Finally, we add a competitor substitution penalty to reduce substitution-induced false positives by downweighting cases where a competing purchase plausibly explains the non-conversion.

4.1 Multi-Behavior Encoding

We adopt LightGCN as the backbone encoder and apply behavior-specific propagation on each behavior subgraph. We do so to preserve heterogeneous preference signals, since merging all behaviors into a single graph would conflate distinct interaction semantics (e.g., browsing reflects interest exploration whereas cart additions/purchases indicate stronger intent), leading to semantic smoothing and degraded representations as observed in prior work [12, 15, 55]. For each behavior view b, we obtain behavior-specific embeddings by layer-wise averaging:

$$\begin{aligned} \mathbf{E}^{b,(l)} &= \hat{\mathbf{A}}_b \, \mathbf{E}^{b,(l-1)}, \quad l = 1, \ldots, L, \\ \mathbf{e}_u^b &= \frac{1}{L+1} \sum_{l=0}^{L} \mathbf{e}_u^{b,(l)}, \quad \mathbf{e}_i^b = \frac{1}{L+1} \sum_{l=0}^{L} \mathbf{e}_i^{b,(l)}, \end{aligned} \tag{1}$$

where $\mathbf{E}^{b,(l)} = [\mathbf{e}_u^{b,(l)}; \mathbf{e}_i^{b,(l)}]$ stacks user/item embeddings and $\hat{\mathbf{A}}_b$ is the normalized adjacency of $\mathcal{G}_b$.

4.2 Mutual Information Learning

4.2.1 Latent Intent Aggregation. Hesitation reflects a mismatch between a user's latent decision intent and the realized target action [28]. Since intent is unobservable, we infer it from auxiliary behaviors. Crucially, such intent is not a static user trait but a goal-conditioned latent state: the same user may exhibit multiple coexisting intents across categories and price ranges, and only a subset of historical signals is relevant when assessing a specific target item. If we aggregate auxiliary behaviors with uniform pooling, incompatible intents (e.g., browsing laptops and shoes) are conflated, yielding a diluted intent representation that weakens the identifiability of hesitation and undermines subsequent dependency estimation between intent and target action.

To address this, we adopt a target-conditioned contextual activation mechanism: the target item i serves as a contextual stimulus

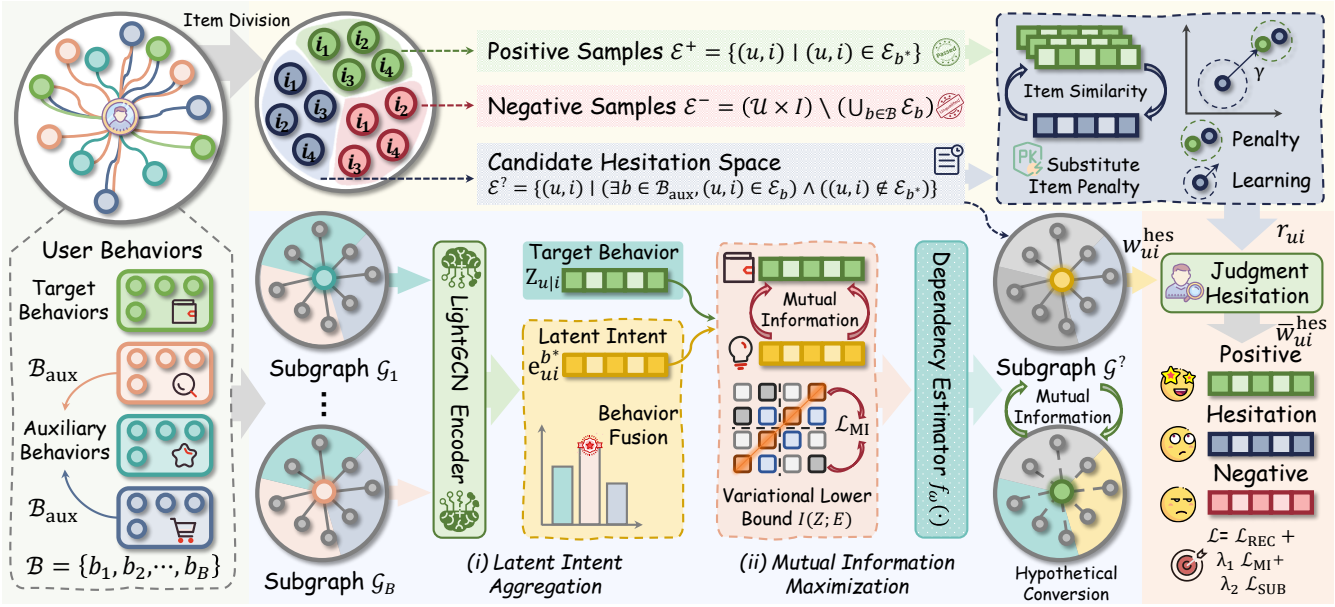

Figure 2: Overall framework of CHARM. CHARM learns user and item representations from multi-behavior interaction graphs via GNN, and jointly models latent intention and hesitation signals using mutual information estimation and a competitor substitution penalty strategy for recommendation optimization.

that selects and amplifies the auxiliary behavior views most relevant to the current decision. Formally, given user u and target item i, we use the target-view item embedding $\mathbf{e}_i^{b^*}$ as a query to activate the user's auxiliary-view representations $\{\mathbf{e}_u^b\}_{b \in \mathcal{B}_{\text{aux}}}$. The attention weight α_{ui}^b is computed as:

$$\alpha_{ui}^b = \text{Softmax}_{b \in \mathcal{B}_{\text{aux}}} \left(\frac{(\mathbf{W}_K \mathbf{e}_u^b)\,(\mathbf{W}_Q \mathbf{e}_i^{b^*})^\top}{\sqrt{d}} \right), \tag{2}$$

where $\mathbf{W}_K, \mathbf{W}_Q \in \mathbb{R}^{d \times d}$ are trainable projections. We deliberately use the query from the target graph rather than auxiliary graphs, because the intent representation is constructed to be predictive of the target action: conditioning on $\mathbf{e}_i^{b^*}$ yields an item-specific intent $\mathbf{z}_{u|i}$, instead of a global preference vector. This target conditioning is essential for our later mutual information criterion, as it prevents irrelevant auxiliary signals from artificially reducing the estimated dependency between intent and target behavior. We then aggregate auxiliary signals into the latent intent vector:

$$\mathbf{z}_{u|i} = \mathbf{W}_P^\top \sum_{b \in \mathcal{B}_{\text{aux}}} \alpha_{ui}^b (\mathbf{W}_V^\top \mathbf{e}_u^b), \tag{3}$$

where $\mathbf{W}_V$ projects heterogeneous auxiliary-view embeddings into a shared intent space before aggregation, and $\mathbf{W}_P$ further adapts the aggregated intent to the target domain for subsequent comparison. Correspondingly, we define the target interaction representation under the target view as:

$$\mathbf{e}_{ui}^{b^*} = \mathbf{e}_u^{b^*} \odot \mathbf{e}_i^{b^*}, \tag{4}$$

where $\odot$ denotes element-wise product. This interaction vector encodes the realized target action and serves as the anchor for mutual information estimation.

4.2.2 Mutual Information Maximization. Mutual information offers a criterion for hesitation detection by casting it as uncertainty reduction. The residual uncertainty of the action given intent is $H(E \mid Z)$, and $I(Z;E) = H(E) - H(E \mid Z)$ measures how much intent reduces this uncertainty; thus, larger $I(Z;E)$ implies lower $H(E \mid Z)$ and a more certain intent-action decision state. This also motivates using mutual information instead of frequency-based heuristics (e.g., counting auxiliary interactions) [13, 37]. While additional auxiliary interactions may provide more evidence, their count is agnostic to whether this evidence is consistent with a specific target conversion and is often confounded by user activity level, since interaction volume does not reveal whether signals are aligned with the target action, while $I(Z;E)$ directly quantifies such dependency and is less sensitive to absolute activity. Accordingly, we learn a dependency model on converted samples: high $I(\mathbf{z}_{u|i}; \mathbf{e}_{ui}^{b^*})$ on $\mathcal{E}^+$ indicates that inferred intent strongly determines the realized action. If a non-converted interaction yields similarly high inferred dependency, it indicates strong alignment despite missing conversion, which characterizes hesitation rather than disinterest. Formally, $I(Z;E)$ can be defined as the Kullback-Leibler (KL) divergence [38] between the joint distribution $\mathbb{P}_{ZE}$ and the product of marginals $\mathbb{P}_Z \otimes \mathbb{P}_E$:

$$\mathcal{I}(Z;E) = D_{\text{KL}}(\mathbb{P}_{ZE} \,\|\, \mathbb{P}_Z \otimes \mathbb{P}_E), \tag{5}$$

where $\|$ denotes divergence between two distributions and $\otimes$ denotes the product of marginals. Maximizing $\mathcal{I}(\mathbf{z}_{u|i}; \mathbf{e}_{ui}^{b^*})$ on $\mathcal{E}^+$ encourages the aggregated intent to be predictive of the realized target action, yielding a transferable dependency estimator for hesitation inference. Since exact mutual information computation is intractable in high-dimensional spaces [5], we adopt a neural variational estimator that maximizes a tractable lower bound derived

from a variational characterization of KL divergence:

$$D_{\mathrm{KL}}(\mathbb{P} \,\|\, \mathbb{Q}) = \sup_{T \in \mathcal{T}} \left\{ \mathbb{E}_{\mathbf{x}\sim\mathbb{P}}[T(\mathbf{x})] - \log\left(\mathbb{E}_{\mathbf{x}\sim\mathbb{Q}}[\exp(T(\mathbf{x}))]\right) \right\}. \quad (6)$$

In our setting, $\mathbf{x} = (\mathbf{z}_{u|i}, \mathbf{e}_{ui}^{b^*})$ and we parameterize T with a dependency estimator $f_\omega(\cdot)$:

$$f_\omega(\mathbf{z}_{u|i}, \mathbf{e}_{ui}^{b^*}) = \mathrm{MLP}\left(\mathbf{z}_{u|i} \odot \mathbf{e}_{ui}^{b^*}\right), \quad (7)$$

where $\mathrm{MLP}(\cdot)$ denotes a Multi-Layer Perceptron. The optimal f_ω under this objective approximates the log density ratio between the joint distribution and the product of marginals, $\log \frac{p(\mathbf{z}_{u|i}, \mathbf{e}_{ui}^{b^*})}{p(\mathbf{z}_{u|i})p(\mathbf{e}_{ui}^{b^*})}$. Therefore, $f_\omega(\mathbf{z}_{u|i}, \mathbf{e}_{ui}^{b^*})$ serves as a calibrated dependency score rather than an arbitrary classifier.

To optimize the lower bound using empirical samples, we construct two types of pairs from converted interactions. Given a sampled set $\mathcal{S} \subset \mathcal{E}^+$, aligned pairs $(\mathbf{z}_{u|i}, \mathbf{e}_{ui}^{b^*})$ correspond to samples from the joint $\mathbb{P}_{ZE}$. We obtain mismatched pairs $(\mathbf{z}_{u|i}, \tilde{\mathbf{e}}_{u\pi(i)}^{b^*})$ by randomly permuting target-action representations within $\mathcal{S}$, which correspond to samples from the product of marginals $\mathbb{P}_Z \otimes \mathbb{P}_E$. We then minimize the following mutual information loss:

$$\mathcal{L}_{\mathrm{MI}} = \log\left(\mathbb{E}_{(u,i)\in\mathcal{S}}\left[\exp\left(f_\omega(\mathbf{z}_{u|i}, \tilde{\mathbf{e}}_{u\pi(i)}^{b^*})\right)\right]\right) - \mathbb{E}_{(u,i)\in\mathcal{S}}\left[f_\omega(\mathbf{z}_{u|i}, \mathbf{e}_{ui}^{b^*})\right], \quad (8)$$

where $\pi(\cdot)$ is a random permutation operator used to break the dependence within the batch.

4.3 Virtual Interaction Modeling

Since f_ω captures dependency patterns of successful conversions, applying it to non-converted interactions provides a plausibility score: a high value indicates strong intent-action alignment despite missing conversion, which is characteristic of hesitation. We thus focus on candidate pairs in $\mathcal{E}^?$. For each $(u, i) \in \mathcal{E}^?$, we compute $\mathbf{z}_{u|i}$ from auxiliary behaviors and construct a virtual target-interaction representation, because the static target-view embedding cannot reflect the structural effect of an assumed target behavior. To do this efficiently, we adopt a virtual structural perturbation that intervenes on the user representation without rebuilding the graph. As LightGCN performs degree-normalized linear neighbor aggregation, inserting an edge (u, i) can be interpreted as adding a new normalized neighbor contribution from item i that shifts the user embedding. We approximate the post-interaction state by keeping the encoder fixed and retaining the dominant local change from this new one-hop neighbor, while only updating degree-related normalization terms. Specifically, we estimate the post-interaction embedding $\tilde{\mathbf{e}}_u^{b^*}$ via linear feature injection:

$$\tilde{\mathbf{e}}_u^{b^*} \approx \frac{\sqrt{(|\mathcal{N}_u^{b^*}|)}}{\sqrt{(|\mathcal{N}_u^{b^*}|+1)}}\mathbf{e}_u^{b^*} + \frac{1}{\sqrt{(|\mathcal{N}_u^{b^*}|+1)(|\mathcal{N}_i^{b^*}|+1)}}\mathbf{e}_i^{b^*}, \quad (9)$$

where $\mathbf{e}_u^{b^*}$ and $\mathbf{e}_i^{b^*}$ are embeddings from the frozen target-view encoder and $\mathcal{N}_u^{b^*}$ and $\mathcal{N}_i^{b^*}$ denote the set of neighbors. This approximation is motivated by the observation that message-passing GNN is often dominated by local aggregation, while long-range information transmission becomes increasingly difficult as depth increases due to repeated normalized propagation, leading to phenomena such as over-smoothing and over-squashing [3, 29, 36, 40, 47]. Although the above injection ignores higher-order effects after multi-layer propagation, such local structural perturbations are commonly adopted in sensitivity analyses and contrastive objectives as efficient surrogates of structural changes [59, 69]. We obtain the virtual target interaction representation $\tilde{\mathbf{e}}_{ui}^{b^*}$ following Eq. (4). This $\tilde{\mathbf{e}}_{ui}^{b^*}$ serves as an efficient surrogate of the virtual action realization, capturing the primary structural consequence of the hypothetical purchase. We then feed the inferred intent $\mathbf{z}_{u|i}$ together with $\tilde{\mathbf{e}}_{ui}^{b^*}$ into the dependency estimator to compute the hesitation score h_{ui}:

$$h_{ui} = f_\omega\left(\mathbf{z}_{u|i}, \tilde{\mathbf{e}}_{ui}^{b^*}\right), \quad (10)$$

where h_{ui} measures the dependency strength between the user's desire and the hypothetical outcome. We assign a probabilistic confidence weight to each candidate pair to capture the degree of hesitation. This weight allows downstream modules to adjust their influence based on model confidence. Specifically, we map the score h_{ui} to the range $(0, 1)$ using a temperature-scaled sigmoid function:

$$w_{ui}^{\mathrm{hes}} = \sigma\left(\frac{h_{ui}}{\tau}\right) = \frac{1}{1+\exp(-h_{ui}/\tau)}, \quad (11)$$

where $\sigma(\cdot)$ is the sigmoid function and τ controls the sharpness of the mapping. The resulting weight w_{ui}^{hes} reflects the model's confidence that (u, i) corresponds to hesitation and will be incorporated into the subsequent substitute item penalty module to modulate the penalty strength.

4.4 Competitor Substitution Penalty

In some cases, non-conversion reflects indecision, where the user's intent is aligned with the target item but the final action is delayed. In other cases, it reflects substitution, where the user's underlying demand has been satisfied by purchasing an alternative item i', making the non-converted candidate i no longer a pending decision. Importantly, purchasing an alternative does not deterministically imply rejecting i (e.g., users may buy multiple similar items such as clothes). Therefore, we treat substitution as a soft, latent cause of non-conversion whose role is to suppress substitution-induced false positives in hesitation labels, rather than enforcing a hard rule. For each user u, we denote the set of purchased items under the target behavior b^* as $\mathcal{I}_u^+ = \{\, i' \in \mathcal{I} \mid (u, i') \in \mathcal{E}^+ \,\}$. We model substitution through a demand-satisfaction view [32, 33] from consumer choice: items can substitute each other if they satisfy a similar need under the target consumption context. Since substitutability is context-dependent and may not coincide with surface-level similarity (e.g., different vacation destinations can substitute each other for a "travel" need), we compute substitute affinity conditioned on the target-induced intent. Specifically, for a candidate item i and a purchased item $i' \in \mathcal{I}_u^+$, we define a substitution evidence score:

$$s_{u,ii'} = \sigma\Big(\cos\left(\mathbf{W}_S[\mathbf{e}_i^{b^*} \oplus \mathbf{z}_{u|i}],\ \mathbf{W}_S[\mathbf{e}_{i'}^{b^*} \oplus \mathbf{z}_{u|i}]\right)\Big), \quad (12)$$

where $\sigma(\cdot)$ is the sigmoid function, $\cos(\cdot,\cdot)$ denotes cosine similarity, $\oplus$ denotes element-wise vector addition, and $\mathbf{W}_S \in \mathbb{R}^{d\times d}$ projects representations into a shared space for comparing substitutes. Intuitively, $s_{u,ii'}$ estimates how likely i' can satisfy the same demand that drives u's hesitation on i under the target context. We

then aggregate the evidence over purchased items by modeling substitution as a latent event that occurs when at least one purchased alternative sufficiently satisfies the demand:

$$r_{ui} = 1 - \prod_{i' \in \mathcal{I}_u^+} \left(1 - \frac{1}{\sqrt{|\mathcal{I}_u^+|}} s_{u,ii'}\right), \quad \bar{w}_{ui}^{\text{hes}} = w_{ui}^{\text{hes}} \cdot (1 - r_{ui}), \tag{13}$$

where r_{ui} can be interpreted as the probability that the missing conversion of i is replaced by substitution. We introduce the factor $1/\sqrt{|\mathcal{I}_u^+|}$ to normalize substitution evidence, preventing many weak similarities from cumulatively inflating r_{ui} and maintaining discriminability across users with different volumes. This aggregation preserves the dominant-evidence property that motivates max pooling, while remaining probabilistic and robust to interaction volume. Accordingly, we downweight hesitation confidence when substitution is more plausible, so that candidates whose demand has likely been resolved contribute less as hesitation positives.

Based on the revised hesitation weight, we define the competitor substitution penalty as a weighted margin ranking loss:

$$\mathcal{L}_{\text{SUB}} = \sum_{(u,i) \in \mathcal{E}^?} \sum_{i' \in \mathcal{I}_u^+} \bar{w}_{ui}^{\text{hes}} \cdot s_{u,ii'} \cdot \max(0, \hat{y}_{ui} - \hat{y}_{ui'} + \gamma),$$
$$\gamma = \text{softplus}\left(\log \frac{s_{u,ii'} + \epsilon}{1 - s_{u,ii'} + \epsilon}\right), \hat{y}_{ui} = \left(\mathbf{e}_u^{b^*} \oplus \mathbf{z}_{u|i}\right)^\top \left(\mathbf{e}_i^{b^*} \oplus \mathbf{z}_{u|i}\right), \tag{14}$$

where $\text{softplus}(x) = \log(1 + \exp(x))$ is a smooth approximation of the ReLU activation function, ϵ is a small constant for numerical stability, and $\hat{y}_{ui}$ and $\hat{y}_{ui'}$ denote the preference scores. The loss adopts an evidence-adaptive margin γ computed from substitution evidence, so stronger evidence requires a larger separation between the purchased competitor i' and the candidate item i. When $s_{u,ii'}$ is small, both the pair weight and γ remain small, making the penalty negligible and avoiding comparisons between unrelated items. When $s_{u,ii'}$ is large, γ increases and the constraint $\hat{y}_{ui'} \geq \hat{y}_{ui} + \gamma$ is emphasized, pushing i' above i and thereby reducing substitution-induced false positives.

4.5 Joint Optimization

Finally, we integrate hesitation into the recommendation objective. Standard BPR [39] treats all items without observed target behavior as negatives and enforces them to rank below observed target interactions, but this unobserved set can include high-intent items whose conversion is missing due to hesitation. Such mislabeling distorts preference learning and blurs the boundary between uncertainty and irrelevance. Unlike prior soft-positive or multi-behavior methods [11, 22] that often handle non-conversions uniformly or rely on interaction frequency, CHARM assigns each candidate a calibrated hesitation confidence and promotes only plausible hesitation cases as intermediate positives. Using the calibrated weight $\bar{w}_{uh}^{\text{hes}}$, we treat a high-confidence hesitation item as lying between observed positives and true negatives and optimize a two-step pairwise ranking objective. Concretely, for each user u we sample quadruples (u, i, h, j) from $O = \{(u, i, h, j) \mid (u, i) \in \mathcal{E}^+, (u, h) \in \mathcal{E}^?, (u, j) \in \mathcal{E}^-\}$, and define the recommendation loss as:

$$\mathcal{L}_{\text{REC}} = -\sum_{(u,i,h,j) \in O} \left(\ln \sigma(\hat{y}_{ui} - \hat{y}_{uh}) + \bar{w}_{uh}^{\text{hes}} \cdot \ln \sigma(\hat{y}_{uh} - \hat{y}_{uj})\right). \tag{15}$$

This objective enforces the ordinal constraint $\hat{y}_{ui} > \hat{y}_{uh} > \hat{y}_{uj}$ while allowing the model to adaptively control how strongly a candidate h should be promoted above true negatives. When $\bar{w}_{uh}^{\text{hes}}$ is small, the model avoids over-enforcing noisy or substitution-affected candidates as intermediate positives; when $\bar{w}_{uh}^{\text{hes}}$ is large, the constraint between h and j is strengthened, reducing the risk of mistakenly suppressing high-intent yet unconverted items. We apply $\bar{w}_{uh}^{\text{hes}}$ only to the intermediate-to-negative term because the positive-to-intermediate preference is anchored by observed target interactions and should be enforced consistently, whereas the main uncertainty lies in whether h should truly behave as an intermediate positive rather than a noisy or substitution-affected item. The overall objective combines the recommendation loss, the mutual information objective, and the competitor substitution penalty:

$$\mathcal{L} = \mathcal{L}_{\text{REC}} + \lambda_1 \mathcal{L}_{\text{MI}} + \lambda_2 \mathcal{L}_{\text{SUB}} + \lambda_3 \|\Theta\|_2^2, \tag{16}$$

where Θ denotes trainable parameters, λ_1 and λ_2 control the contributions of mutual information learning and competitor substitution penalty, and $\lambda_3 \|\Theta\|_2^2$ is an L_2 regularization term.

4.6 Further Analysis

4.6.1 Time Complexity. CHARM is dominated by multi-behavior LightGCN propagation and dependence scoring, where LightGCN over B behavior graphs with L layers costs $O\left(\sum_{b=1}^{B} L\,|\mathcal{E}_b|\,d\right)$; for each pair (u, i), intent aggregation costs $O(|\mathcal{B}_{\text{aux}}|\,d)$, and mutual information learning with a dependence scoring network incurs $O(|S|\,d)$ per mini-batch S; at inference, virtual construction and scoring over $\mathcal{E}^?$ candidates costs $O\left(|\mathcal{E}^?|\,(d + f_\omega(d))\right)$, plus an optional substitution term $O(|\mathcal{E}^?|\,K\,d)$ when comparing against K sampled substitutes, and the time complexity is stable within an acceptable range.

4.6.2 Space Complexity. CHARM stores behavior-specific user/item embeddings for B views, requiring $O((|U| + |I|)\,B\,d)$ memory, and sparse adjacency for each behavior subgraph, requiring $O(\sum_b |\mathcal{E}_b|)$; the parameters of the dependency estimator contribute $O(f_\omega(d))$, typically negligible compared with embeddings.

5 Experiments

5.1 Experimental Setup

5.1.1 Dataset Description. We evaluate CHARM on three multi-behavior recommendation benchmarks, namely **Tmall**[2], **Taobao**[3], and **Beibei**[4]. Following the common setting in prior work [34, 61], we treat the strongest behavior (i.e., *purchase*) as the target behavior b^* and use the remaining behaviors as auxiliary signals. The dataset statistics after preprocessing are reported in Table 1.

[2] https://tianchi.aliyun.com/dataset/53
[3] https://tianchi.aliyun.com/dataset/649
[4] https://github.com/akaxlh/MB-GMN/tree/main/Datasets

Table 1: Statistics of datasets.

Stats.	#Users	#Items	#Interactions	#Behavior Type
Tmall	41,738	11,953	2.3×10^6	{View, Collect, Cart, Purchase}
Taobao	15,449	11,953	1.2×10^6	{View, Cart, Purchase}
Beibei	21,716	7,977	3.3×10^6	{View, Cart, Purchase}

5.1.2 Parameter Settings. We implement CHARM in PyTorch. All model parameters are initialized with Xavier [17] initialization. Unless otherwise specified, the embedding size is set to $d = 64$ and the batch size is 2048. We use Adam [25] for optimization with a fixed learning rate of 1e-5. The overall objective is optimized with loss weights $\lambda_1 = 0.6$ and $\lambda_2 = 0.2$, and we set the L_2 regularization coefficient to $\lambda_3 = 0.001$. We conduct grid search for two key hyperparameters of CHARM on the validation set. Specifically, the LightGCN propagation depth L is searched in $\{1, 2, 3, 4\}$, and the temperature τ in the hesitation mapping is searched in $\{0.1, 0.3, 0.5, 0.7\}$. All experiments are repeated five times with different random seeds, and we report the average results.

5.1.3 Evaluation Metrics. We evaluate top-K ranking performance using Hit Ratio (HR@K) and Normalized Discounted Cumulative Gain (NDCG@K), with $K = 10$. Following standard implicit feedback evaluation, for each user we rank the ground-truth test item against a set of randomly sampled unobserved items as negatives, compute HR@10 and NDCG@10 on the resulting candidate set.

5.1.4 Baseline Models. We compare CHARM with representative baselines from four categories: **(1) Single-behavior models** only use the target behavior and treat unobserved target interactions as negatives, including MF-BPR [39] and LightGCN [19]. **(2) Multi-behavior models without multi-task learning** incorporate auxiliary behaviors but do not explicitly formulate multi-task objectives across behaviors, including NMTR [15], MBGCN [23], and KHGT [55]. **(3) Soft-positive and causal multi-behavior baselines** either assign soft positives for implicit feedback or multi-behavior learning, or incorporate causal reasoning to mitigate confounding. Specifically, the causal baselines include PDA [68], DecRs [48], CausalD [65], and CVID [10], while the soft-positive baselines include CDR [8], MBA [57], and SoftRec [11]. **(4) Multi-behavior models with multi-task learning** explicitly couple behaviors via multi-task objectives or cross-behavior transfer mechanisms, including CML [52], RCL [53], PKEF [34], and COPF [61].

5.2 Performance Comparison

Table 2 reports the overall comparison between CHARM and all baselines. First, CHARM achieves the best results on all three benchmarks for both HR@10 and NDCG@10, and consistently improves over the strongest baseline on each dataset by 7.67% on Tmall, 8.68% on Taobao, and 7.70% on Beibei, suggesting strong robustness across different behavior distributions and sparsity levels. Target-only methods are uniformly weaker than multi-behavior approaches, which confirms that auxiliary behaviors provide crucial evidence for intent inference when purchases are sparse; for example, CHARM substantially outperforms LightGCN on all datasets, indicating that target-view interactions alone cannot recover the evolving intent signals contained in auxiliary feedback. Second, soft-positive and causal multi-behavior baselines are strong competitors because they soften supervision or reduce bias, but their signals are typically generic and do not explicitly determine whether a specific non-conversion reflects hesitation or is explained by other mechanisms, so false positives can still be introduced. CHARM strengthens this line by grounding the intermediate signal in target-conditioned intent and intent-action alignment, and further suppressing substitution-induced false positives, which yields cleaner training supervision and stronger top-K ranking performance than softening or debiasing alone. Third, strong multi-behavior encoders, including both graph-based and multi-task variants, benefit from heterogeneous supervision and perform competitively (e.g., KHGT and COPF on Taobao), yet CHARM still improves upon them, implying that representation fusion across behaviors is not sufficient when the semantics of missing conversions are ambiguous. In realistic logs, auxiliary interactions often mix exploratory and decisive intents, and non-conversion can be caused by hesitation or substitution rather than irrelevance, which leads to noisy gradients if treated as uniformly negative; CHARM mitigates this by calibrating auxiliary evidence into a target-conditioned intermediate preference signal and using it selectively in optimization.

Table 2: The overall performance comparison. The best results are highlighted in bold and the second-best results are underlined. ⋆ represents significance level p-value < 0.05 of comparing CHARM with the best baseline.

Model	Tmall		Taobao		Beibei	
	HR	NDCG	HR	NDCG	HR	NDCG
MF-BPR	0.0228	0.0156	0.0074	0.0030	0.0186	0.0054
LightGCN	0.0398	0.0206	0.0406	0.0244	0.0394	0.0212
NMTR	0.0451	0.0228	0.0505	0.0263	0.0345	0.0270
MBGCN	0.0550	0.0281	0.0508	0.0297	0.0471	0.0254
KHGT	0.0599	0.0308	0.1180	0.0662	0.0534	0.0259
PDA	0.1089	0.0517	0.1009	0.0500	0.1067	0.0518
DecRs	0.1231	0.0649	0.1108	0.0532	0.1178	0.0593
CausalD	0.1433	0.0794	0.1327	0.0787	0.1399	0.0817
CVID	0.1643	0.0840	0.1512	0.0791	0.1577	<u>0.0924</u>
CDR	0.0959	0.0535	0.0868	0.0492	0.0911	0.0513
MBA	0.1261	0.0780	0.1179	0.0720	0.1235	0.0724
SoftRec	0.1588	0.0899	0.1487	<u>0.0878</u>	0.1516	0.0914
CML	0.0532	0.0284	0.0682	0.0632	0.0429	0.0303
RCL	0.0835	0.0441	0.1249	0.0737	0.0559	0.0325
PKEF	0.1281	0.0724	0.1390	0.0780	0.1129	0.0576
COPF	<u>0.1656</u>	<u>0.0913</u>	<u>0.1649</u>	0.0842	<u>0.1588</u>	0.0898
CHARM	**0.1803⋆**	**0.0972⋆**	**0.1766⋆**	**0.0968⋆**	**0.1750⋆**	**0.0972⋆**
%Improv.	**8.88%**	**6.46%**	**7.10%**	**10.25%**	**10.20%**	**5.19%**

5.3 Ablation Study

5.3.1 Impact of the Key Components. Table 3 reports ablations of CHARM to isolate the contribution of each component. The first block removes one major module from the full framework, including the *Backbone* (reducing to LightGCN), *w/o Attn* that replaces intent aggregation with uniform pooling, *w/o MI* that replaces the learned dependency estimator with cosine similarity, *w/o VIM* that disables

Table 3: Performances of different CHARM variants.

Model	Tmall		Taobao		Beibei	
	HR	NDCG	HR	NDCG	HR	NDCG
Backbone	0.0398	0.0206	0.0406	0.0244	0.0394	0.0212
w/o Attn	0.1664	0.0905	0.1654	0.0904	0.1672	0.0920
w/o MI	0.1681	0.0914	0.1666	0.0907	0.1644	0.0905
w/o VIM	0.1720	0.0937	0.1701	0.0939	0.1698	0.0947
w/o Reweight	0.1755	0.0916	0.1727	0.0915	0.1719	0.0933
w/o SUB	0.1738	0.0921	0.1719	0.0923	0.1693	0.0934
w/o Ordinal	0.1652	0.0899	0.1648	0.0895	0.1667	0.0918
Uniform	0.1652	0.0875	0.1619	0.0863	0.1603	0.0885
HardSet	0.1766	0.0941	0.1737	0.0948	0.1721	0.0958
Heu-Freq	0.1707	0.0913	0.1651	0.0895	0.1664	0.0916
Sub-no-w	0.1771	0.0955	0.1752	0.0954	0.1737	0.0964
CHARM	**0.1803**	**0.0972**	**0.1766**	**0.0968**	**0.1750**	**0.0969**

virtual target-interaction construction and uses the static target interaction representation, *w/o Reweight* that removes substitution-aware reweighting while keeping the substitution loss, *w/o SUB* that drops the substitution loss while still using the reweighted hesitation confidence, and *w/o Ordinal* that reverts the ordinal objective to standard BPR. The second block examines signal-level choices, including *Uniform* (equal weights for all hesitation candidates), *HardSet* (hard thresholding), *Heu-Freq* (frequency-based hesitation scores), and *Sub-no-w* (training the substitution loss without modulating it by calibrated hesitation confidence).

Across datasets, removing any major component consistently degrades performance, indicating that CHARM gains from a coupled pipeline rather than a single isolated design. Among the intent and calibration modules, both *w/o Attn* and *w/o MI* show clear declines, suggesting that target-conditioned intent construction and a learned, calibrated intent-action dependency estimator is crucial for separating plausible missing conversions from weak exploratory interest. Compared with *w/o VIM*, the full model yields consistent improvements, supporting that explicitly modeling the structural effect of an assumed purchase produces a more faithful action representation than a static target interaction vector and therefore enables more reliable hesitation scoring. In addition, substitution modeling and ordinal training act as necessary complements: *w/o Reweight* and *w/o SUB* underperform CHARM, implying that suppressing substitution-induced false positives is important, while *w/o Ordinal* shows that treating hesitation candidates as an intermediate preference level provides a stronger optimization signal than collapsing training back to standard ranking.

The signal-level variants further suggest that the gains do not come from naive reweighting or heuristic soft positives. *Uniform* is consistently the weakest among strategy variants, indicating that indiscriminately promoting all candidates introduces substantial noise. *Heu-Freq* also lags behind CHARM, showing that interaction counts are an unreliable proxy for hesitation even though they provide a form of supervision. *HardSet* improves over several weakened variants, which implies that filtering extremely noisy candidates can help, but it remains inferior because hard selection discards the calibration offered by confidence weights. Finally, *Sub-no-w* is competitive yet still below CHARM, suggesting that substitution constraints are most effective when coupled with calibrated hesitation confidence so that the penalty concentrates on genuinely ambiguous non-conversions rather than being applied uniformly.

Table 4: Detailed ablation study on structural perturbation and substitution modeling.

Variant	Tmall		Taobao		Beibei	
	HR	NDCG	HR	NDCG	HR	NDCG
Virtual Interaction Modeling						
VIM-NaiveAdd	0.1739	0.0932	0.1711	0.0936	0.1710	0.0945
VIM-ScaleOnly	0.1714	0.0927	0.1708	0.0932	0.1691	0.0947
VIM-2HopComp	0.1809	0.0986	0.1771	0.0977	0.1758	0.0974
VIM-Recompute	0.1811	0.0982	0.1774	0.0979	0.1764	0.0977
Substitution Modeling Strategy						
(i) Evidence						
Sub-ItemOnly	0.1769	0.0942	0.1732	0.0936	0.1728	0.0955
Sub-NoProj	0.1780	0.0964	0.1743	0.0956	0.1738	0.0962
(ii) Aggregation						
Sub-MeanPool	0.1775	0.0957	0.1740	0.0944	0.1732	0.0953
Sub-MaxPool	0.1791	0.0966	0.1754	0.0965	0.1743	0.0961
(iii) Weighting						
Sub-FixedMargin	0.1759	0.0940	0.1728	0.0937	0.1712	0.0952
Sub-NoSWeight	0.1746	0.0939	0.1713	0.0929	0.1705	0.0943
Sub-NoWWeight	0.1765	0.0957	0.1732	0.0940	0.1721	0.0957
CHARM	0.1803	0.0972	0.1766	0.0968	0.1750	0.0969

5.3.2 Impact of Structural Perturbation. Table 4 presents an ablation study on two key components of CHARM, namely structural perturbation and substitution modeling. For structural perturbation, we compare four virtual variants: *VIM-NaiveAdd*, which adds the item embedding to the user embedding without graph normalization, *VIM-ScaleOnly*, which only applies degree-based scaling, *VIM-2HopComp*, which introduces a second-order compensation term for two-hop effects, and *VIM-Recompute*, which inserts (u, i) and reruns message passing to obtain the virtual embedding. *VIM-NaiveAdd* and *VIM-ScaleOnly* perform substantially worse than CHARM, indicating that simple additive or scaling interventions do not capture the structural semantics of a hypothetical purchase. *VIM-Recompute* achieves the best performance, confirming that explicitly modeling structural change is beneficial, but it is computationally impractical because it requires rerunning the encoder for each candidate. *VIM-2HopComp* brings only marginal gains while noticeably increasing runtime due to denser aggregation and the extra compensation term. Overall, CHARM's first-order residual injection achieves a better accuracy efficiency trade-off, retaining comparable performance to higher-fidelity variants with essentially the same training cost as the base setting.

5.3.3 Impact of Substitution Modeling . Substitution modeling is ablated along three design dimensions: evidence, aggregation, and penalty weighting. For evidence, *Sub-ItemOnly* computes cosine similarity using only item embeddings and performs worse than CHARM, implying that item-only similarity is insufficient to capture context-dependent substitutability; *Sub-NoProj* removes the projection layer W_S and improves over *Sub-ItemOnly*, which suggests that learning a suitable space for measuring substitution evidence is beneficial. For aggregation, *Sub-MaxPool* outperforms *Sub-MeanPool*, indicating that emphasizing the strongest substitute

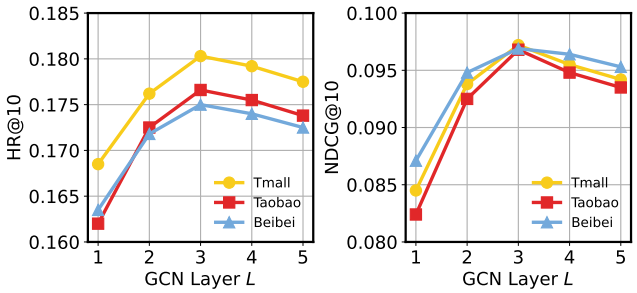

Figure 3: Impact of LightGCN propagation layers L.

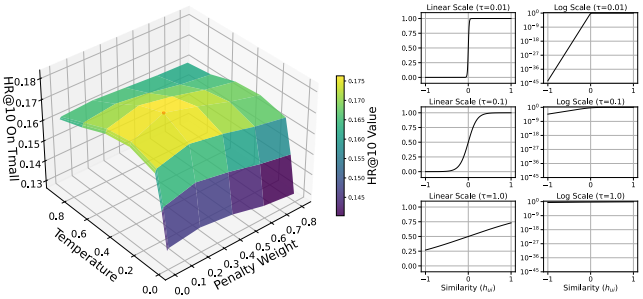

Figure 4: Impact of temperature parameter τ.

signal is more effective than averaging across all purchased items, although both remain below CHARM, which yields a more reliable aggregation of evidence. For penalty weighting, *Sub-FixedMargin* and *Sub-NoSWeight* degrade performance, showing that the penalty should adapt to the strength of substitution evidence rather than using a fixed margin or ignoring $s_{u,ii'}$; *Sub-NoWWeight* also underperforms, confirming that calibrated hesitation confidence $\bar{w}^{\text{hes}}$ is important for focusing the penalty on truly ambiguous non-conversions instead of uniformly penalizing all candidates.

5.4 Hyperparameter Analysis

5.4.1 Impact of LightGCN Layers. Fig. 3 shows CHARM's performance under propagation depths $L \in \{1, \ldots, 5\}$. Performance improves substantially from $L = 1$ to $L = 3$, indicating that moderate high-order neighborhood aggregation helps capture intent signals beyond direct interactions, which is important under sparse target behaviors. The best results at $L = 3$ suggest a favorable balance between leveraging collaborative structure and avoiding noise. When $L > 3$, performance declines, because deeper propagation increasingly mixes heterogeneous and less relevant signals and aggravates over-smoothing, making node representations less discriminative.

5.4.2 Impact of the Temperature Parameter. The temperature τ controls the sharpness of the hesitation $\sigma(h_{ui}/\tau)$, and thus the effective boundary between high-intent hesitation and low-confidence noise. We sweep $\tau \in \{0.1, 0.3, 0.5, 0.7, 0.9\}$ and observe a clear ridge on Tmall (Fig. 4, Left), with the best performance at $\tau = 0.3$ and a sharp collapse at very small temperatures. This behavior is explained by the gradient profile of the sigmoid (Fig. 4, Right): when

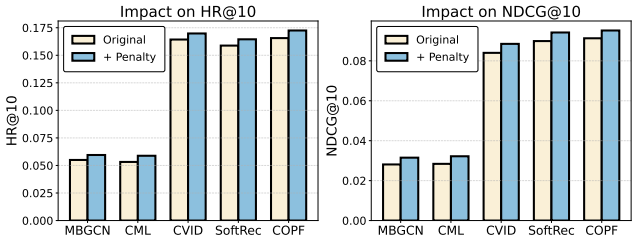

Figure 5: Compatibility analysis on Tmall. The "+ Penalty" denotes the baseline model equipped with our competitor substitution penalty strategy.

Figure 6: Case study of user u_{3518} interpreting hesitation confidence and substitution score.

$\tau \to 0$, the mapping approaches a hard step function and induces severe gradient vanishing (e.g., non-activated samples yield near-zero gradients, preventing effective back-propagation and collapsing training), whereas when τ grows toward 1, the mapping becomes overly flat/linear, producing nearly uniform weights that weaken discrimination between true hesitation candidates and noise. The setting $\tau = 0.3$ achieves the best trade-off. It preserves enough nonlinearity for discriminative weighting while maintaining stable gradient flow, so we fix $\tau = 0.3$ in all subsequent experiments.

5.5 Universality and Explainability

5.5.1 Universality of Penalty Strategy. To verify the universality of our strategy, we conduct a compatibility study by incorporating the competitor substitution penalty into five representative multi-behavior baselines (MBGCN, CML, CVID, SoftRec, and COPF) and reporting results on Tmall in Fig. 5. We observe consistent improvements for all backbones after adding the penalty, including strong methods such as SoftRec and COPF. This suggests that even high-capacity multi-behavior models can misinterpret substitution-driven non-conversions as a lack of interest. A reason is that standard objectives treat missing purchases as uniformly negative evidence, so items that are actually high-intent but replaced by competitors are pushed down during optimization. Our penalty provides an explicit corrective signal by down-weighting such ambiguous negatives and encouraging the model to preserve the ranking of truly high-intent candidates. Importantly, the gains hold across qualitatively different learning paradigms, indicating that the substitution phenomenon arises from data generation and user choice behavior rather than from a particular encoder design.

5.5.2 Case Study. Fig. 6 shows a user-level case illustrating how CHARM detects hesitation from multi-behavior logs and distinguishes it from substitution. By maximizing mutual information,

CHARM assigns high hesitation confidence to several non-converted candidates, suggesting that they are plausible missing conversions rather than uniform negatives. The right panel further demonstrates how CHARM refines such candidates by accounting for substitution. Although item i_{1462} initially receives a high hesitation confidence (0.84) from intent-action alignment, CHARM identifies that the user has purchased a highly plausible competing alternative that can satisfy the same demand. Consequently, the substitution module substantially downweights its contribution, reducing the calibrated hesitation weight of i_{1462} from 0.84 to 0.21. This adjustment prevents CHARM from incorrectly promoting a candidate whose missing conversion is better explained by demand satisfaction through a competitor purchase, while still retaining high-intent candidates that are not supported by strong substitution evidence.

6 Conclusion

In this work, we study user hesitation in multi-behavior recommendation, where high-intent items may remain unconverted and are easily misinterpreted as negatives. We propose CHARM, a unified framework that estimates hesitation as the strength of intent-action alignment via virtual structural perturbation, and further calibrates the inferred signals by accounting for competitor substitution. By turning non-conversions into calibrated intermediate preference signals rather than uniformly negative feedback, CHARM improves ranking quality and reduces redundancy caused by substitution-induced false positives. Extensive experiments on benchmark datasets demonstrate consistent gains over strong multi-behavior baselines, and ablations validate the importance of each stage in the inference pipeline. Future work includes improving causal modeling with richer interventions and incorporating finer-grained choice contexts (e.g., time, budget, and category constraints) to better distinguish substitution from multi-purchase behavior.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grants No. 62573122 and 62576098) and the Special Project of Anhui Cultural Development Institute (Grant No. 2026ZKZX32).

References

[1] Gediminas Adomavicius and Alexander Tuzhilin. 2005. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE Transactions on Knowledge and Data Engineering* 17, 6 (2005), 734–749.

[2] Aman Agarwal, Kenta Takatsu, Ivan Zaitsev, and Thorsten Joachims. 2019. A general framework for counterfactual learning-to-rank. In *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval*. 5–14.

[3] Uri Alon and Eran Yahav. 2020. On the bottleneck of graph neural networks and its practical implications. *arXiv preprint arXiv:2006.05205* (2020).

[4] Peter C Austin and Elizabeth A Stuart. 2015. Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies. *Statistics in medicine* 34, 28 (2015), 3661–3679.

[5] Mohamed Ishmael Belghazi, Aristide Baratin, Sai Rajeshwar, Sherjil Ozair, Yoshua Bengio, Aaron Courville, and Devon Hjelm. 2018. Mutual information neural estimation. In *International conference on machine learning*. PMLR, 531–540.

[6] Dirk G. F. M. Bollen, Bart P. Knijnenburg, Martijn C. Willemsen, and Mark P. Graus. 2010. Understanding choice overload in recommender systems. In *Proceedings of the Fourth ACM Conference on Recommender Systems (RecSys '10)*. 63–70. doi:10.1145/1864708.1864724

[7] Leon Brillouin. 2013. *Science and information theory*. Courier Corporation.

[8] Hong Chen, Yudong Chen, Xin Wang, Ruobing Xie, Rui Wang, Feng Xia, and Wenwu Zhu. 2021. Curriculum disentangled recommendation with noisy multi-feedback. *Advances in Neural Information Processing Systems* 34 (2021), 26924–26936.

[9] Xiaoqing Chen, Zhitao Li, Weike Pan, and Zhong Ming. 2023. A survey on multi-behavior sequential recommendation. *arXiv preprint arXiv:2308.15701* (2023).

[10] Yuzhe Chen, Jie Cao, Youquan Wang, Jia Wu, Huanhuan Chen, and Guandong Xu. 2025. Causal variational inference for deconfounded multi-behavior recommendation. *ACM Transactions on Information Systems* 43, 6 (2025), 1–26.

[11] Mingyue Cheng, Fajie Yuan, Qi Liu, Shenyang Ge, Zhi Li, Runlong Yu, Defu Lian, Senchao Yuan, and Enhong Chen. 2021. Learning recommender systems with implicit feedback via soft target enhancement. In *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 575–584.

[12] Zhiyong Cheng, Sai Han, Fan Liu, Lei Zhu, Zan Gao, and Yuxin Peng. 2023. Multi-behavior recommendation with cascading graph convolution networks. In *Proceedings of the ACM Web Conference 2023*. 1181–1189.

[13] Paweł Czyż, Frederic Grabowski, Julia Vogt, Niko Beerenwinkel, and Alexander Marx. 2023. Beyond normal: On the evaluation of mutual information estimators. *Advances in neural information processing systems* 36 (2023), 16957–16990.

[14] Michele Jonsson Funk, Daniel Westreich, Chris Wiesen, Til Stürmer, M Alan Brookhart, and Marie Davidian. 2011. Doubly robust estimation of causal effects. *American journal of epidemiology* 173, 7 (2011), 761–767.

[15] Chen Gao, Xiangnan He, Dahua Gan, Xiangning Chen, Fuli Feng, Yong Li, Tat-Seng Chua, and Depeng Jin. 2019. Neural multi-task recommendation from multi-behavior data. In *2019 IEEE 35th international conference on data engineering (ICDE)*. IEEE, 1554–1557.

[16] Chen Gao, Yu Zheng, Wenjie Wang, Fuli Feng, Xiangnan He, and Yong Li. 2024. Causal inference in recommender systems: A survey and future directions. *ACM Transactions on Information Systems* 42, 4 (2024), 1–32.

[17] Xavier Glorot and Yoshua Bengio. 2010. Understanding the difficulty of training deep feedforward neural networks. In *Proceedings of the thirteenth international conference on artificial intelligence and statistics*. JMLR Workshop and Conference Proceedings, 249–256.

[18] Qingyu Guo, Fuzhen Zhuang, Chuan Qin, Hengshu Zhu, Xing Xie, Hui Xiong, and Qing He. 2020. A survey on knowledge graph-based recommender systems. *IEEE Transactions on Knowledge and Data Engineering* 34, 8 (2020), 3549–3568.

[19] 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.

[20] Xiangnan He, Lizi Liao, Hanwang Zhang, Liqiang Nie, Xia Hu, and Tat-Seng Chua. 2017. Neural collaborative filtering. In *Proceedings of the 26th international conference on world wide web*. 173–182.

[21] Jonathan L Herlocker, Joseph A Konstan, and John Riedl. 2000. Explaining collaborative filtering recommendations. In *Proceedings of the 2000 ACM conference on Computer supported cooperative work*. 241–250.

[22] Yifan Hu, Yehuda Koren, and Chris Volinsky. 2008. Collaborative filtering for implicit feedback datasets. In *2008 Eighth IEEE international conference on data mining*. Ieee, 263–272.

[23] Bowen Jin, Chen Gao, Xiangnan He, Depeng Jin, and Yong Li. 2020. Multi-behavior recommendation with graph convolutional networks. In *Proceedings of the 43rd international ACM SIGIR conference on research and development in information retrieval*. 659–668.

[24] Kyungho Kim, Sunwoo Kim, Geon Lee, Jinhong Jung, and Kijung Shin. 2025. Multi-Behavior Recommender Systems: A Survey. *arXiv preprint arXiv:2503.06963* (2025).

[25] Diederik P Kingma and Jimmy Ba. 2014. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980* (2014).

[26] Alexander Kraskov, Harald Stögbauer, and Peter Grassberger. 2004. Estimating mutual information. *Physical Review E—Statistical, Nonlinear, and Soft Matter Physics* 69, 6 (2004), 066138.

[27] Seunghan Lee, Geonwoo Ko, Hyun-Je Song, and Jinhong Jung. 2024. MuLe: Multi-Grained Graph Learning for Multi-Behavior Recommendation. In *Proceedings of the 33rd ACM International Conference on Information and Knowledge Management*. 1163–1173.

[28] Cheng Li, Yong Xu, Suhua Tang, Wenqiang Lin, Xin He, and Jinde Cao. 2025. User Hesitation and Negative Transfer in Multi-Behavior Recommendation. *arXiv e-prints* (2025), arXiv–2511.

[29] Qimai Li, Zhichao Han, and Xiao-Ming Wu. 2018. Deeper insights into graph convolutional networks for semi-supervised learning. In *Proceedings of the AAAI conference on artificial intelligence*, Vol. 32.

[30] Tzu-Heng Lin, Chen Gao, and Yong Li. 2019. Cross: Cross-platform recommendation for social e-commerce. In *Proceedings of the 42nd International ACM SIGIR conference on research and development in information retrieval*. 515–524.

[31] Weiwen Liu, Wei Guo, Yong Liu, Ruiming Tang, and Hao Wang. 2023. User behavior modeling with deep learning for recommendation: Recent advances. In *Proceedings of the 17th ACM Conference on Recommender Systems*. 1286–1287.

[32] Andreu Mas-Colell, Michael Dennis Whinston, Jerry R Green, et al. 1995. *Microeconomic theory*. Vol. 1. Oxford university press New York.

[33] Daniel McFadden. 1972. Conditional logit analysis of qualitative choice behavior. (1972).

[34] Chang Meng, Chenhao Zhai, Yu Yang, Hengyu Zhang, and Xiu Li. 2023. Parallel knowledge enhancement based framework for multi-behavior recommendation. In *Proceedings of the 32nd ACM international conference on information and knowledge management*. 1797–1806.

[35] Chang Meng, Hengyu Zhang, Wei Guo, Huifeng Guo, Haotian Liu, Yingxue Zhang, Hongkun Zheng, Ruiming Tang, Xiu Li, and Rui Zhang. 2023. Hierarchical projection enhanced multi-behavior recommendation. In *Proceedings of the 29th ACM SIGKDD conference on knowledge discovery and data mining*. 4649–4660.

[36] Kenta Oono and Taiji Suzuki. 2019. Graph neural networks exponentially lose expressive power for node classification. *arXiv preprint arXiv:1905.10947* (2019).

[37] Ben Poole, Sherjil Ozair, Aaron Van Den Oord, Alex Alemi, and George Tucker. 2019. On variational bounds of mutual information. In *International conference on machine learning*. PMLR, 5171–5180.

[38] Fiana Raiber and Oren Kurland. 2017. Kullback-leibler divergence revisited. In *Proceedings of the ACM SIGIR international conference on theory of information retrieval*. 117–124.

[39] Steffen Rendle, Christoph Freudenthaler, Zeno Gantner, and Lars Schmidt-Thieme. 2012. BPR: Bayesian personalized ranking from implicit feedback. *arXiv preprint arXiv:1205.2618* (2012).

[40] T Konstantin Rusch, Michael M Bronstein, and Siddhartha Mishra. 2023. A survey on oversmoothing in graph neural networks. *arXiv preprint arXiv:2303.10993* (2023).

[41] Badrul Sarwar, George Karypis, Joseph Konstan, and John Riedl. 2001. Item-based collaborative filtering recommendation algorithms. In *Proceedings of the 10th international conference on World Wide Web*. 285–295.

[42] Tobias Schnabel, Adith Swaminathan, Ashudeep Singh, Navin Chandak, and Thorsten Joachims. 2016. Recommendations as treatments: Debiasing learning and evaluation. In *international conference on machine learning*. PMLR, 1670–1679.

[43] Kexin Shi, Yun Zhang, Bingyi Jing, and Wenjia Wang. 2022. Enhancing Recommender Systems: A Strategy to Mitigate False Negative Impact. *arXiv preprint arXiv:2211.13912* (2022).

[44] Ajit P Singh and Geoffrey J Gordon. 2008. Relational learning via collective matrix factorization. In *Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*. 650–658.

[45] Xiaodan Song, Belle L Tseng, Ching-Yung Lin, and Ming-Ting Sun. 2006. Personalized recommendation driven by information flow. In *Proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval*. 509–516.

[46] Xiaoyuan Su and Taghi M Khoshgoftaar. 2009. A survey of collaborative filtering techniques. *Advances in artificial intelligence* 2009, 1 (2009), 421425.

[47] Jake Topping, Francesco Di Giovanni, Benjamin Paul Chamberlain, Xiaowen Dong, and Michael M Bronstein. 2021. Understanding over-squashing and bottlenecks on graphs via curvature. *arXiv preprint arXiv:2111.14522* (2021).

[48] Wenjie Wang, Fuli Feng, Xiangnan He, Xiang Wang, and Tat-Seng Chua. 2021. Deconfounded recommendation for alleviating bias amplification. In *Proceedings of the 27th ACM SIGKDD conference on knowledge discovery & data mining*. 1717–1725.

[49] Yixin Wang, Dawen Liang, Laurent Charlin, and David M Blei. 2020. Causal inference for recommender systems. In *Proceedings of the 14th ACM conference on recommender systems*. 426–431.

[50] Zhenlei Wang, Jingsen Zhang, Hongteng Xu, Xu Chen, Yongfeng Zhang, Wayne Xin Zhao, and Ji-Rong Wen. 2021. Counterfactual data-augmented sequential recommendation. In *Proceedings of the 44th international ACM SIGIR conference on research and development in information retrieval*. 347–356.

[51] Tianxin Wei, Fuli Feng, Jiawei Chen, Ziwei Wu, Jinfeng Yi, and Xiangnan He. 2021. Model-agnostic counterfactual reasoning for eliminating popularity bias in recommender system. In *Proceedings of the 27th ACM SIGKDD conference on knowledge discovery & data mining*. 1791–1800.

[52] Wei Wei, Chao Huang, Lianghao Xia, Yong Xu, Jiashu Zhao, and Dawei Yin. 2022. Contrastive meta learning with behavior multiplicity for recommendation. In *Proceedings of the 15th ACM international conference on web search and data mining*. 1120–1128.

[53] 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.

[54] Lianghao Xia, Chao Huang, Yong Xu, Peng Dai, Bo Zhang, and Liefeng Bo. 2020. Multiplex behavioral relation learning for recommendation via memory augmented transformer network. In *Proceedings of the 43rd international ACM SIGIR conference on research and development in information retrieval*. 2397–2406.

[55] Lianghao Xia, Chao Huang, Yong Xu, Peng Dai, Xiyue Zhang, Hongsheng Yang, Jian Pei, and Liefeng Bo. 2021. Knowledge-enhanced hierarchical graph transformer network for multi-behavior recommendation. In *Proceedings of the AAAI conference on artificial intelligence*, Vol. 35. 4486–4493.

[56] Lianghao Xia, Chao Huang, Yong Xu, and Jian Pei. 2022. Multi-behavior sequential recommendation with temporal graph transformer. *IEEE Transactions on Knowledge and Data Engineering* 35, 6 (2022), 6099–6112.

[57] Xin Xin, Xiangyuan Liu, Hanbing Wang, Pengjie Ren, Zhumin Chen, Jiahuan Lei, Xinlei Shi, Hengliang Luo, Joemon M Jose, Maarten de Rijke, et al. 2023. Improving implicit feedback-based recommendation through multi-behavior alignment. In *Proceedings of the 46th international ACM SIGIR conference on research and development in information retrieval*. 932–941.

[58] Jingcao Xu, Chaokun Wang, Cheng Wu, Yang Song, Kai Zheng, Xiaowei Wang, Changping Wang, Guorui Zhou, and Kun Gai. 2023. Multi-behavior self-supervised learning for recommendation. In *Proceedings of the 46th international ACM SIGIR conference on research and development in information retrieval*. 496–505.

[59] Yuning You, Tianlong Chen, Yongduo Sui, Ting Chen, Zhangyang Wang, and Yang Shen. 2020. Graph contrastive learning with augmentations. *Advances in neural information processing systems* 33 (2020), 5812–5823.

[60] Enming Yuan, Wei Guo, Zhicheng He, Huifeng Guo, Chengkai Liu, and Ruiming Tang. 2022. Multi-behavior sequential transformer recommender. In *Proceedings of the 45th international ACM SIGIR conference on research and development in information retrieval*. 1642–1652.

[61] Chenhao Zhai, Chang Meng, Yu Yang, Kexin Zhang, Xuhao Zhao, and Xiu Li. 2025. Combinatorial Optimization Perspective based Framework for Multi-behavior Recommendation. In *Proceedings of the 31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining*. 1891–1902.

[62] Chi Zhang, Rui Chen, Xiangyu Zhao, Qilong Han, and Li Li. 2023. Denoising and prompt-tuning for multi-behavior recommendation. In *Proceedings of the ACM web conference 2023*. 1355–1363.

[63] Fuzheng Zhang, Nicholas Jing Yuan, Defu Lian, Xing Xie, and Wei-Ying Ma. 2016. Collaborative knowledge base embedding for recommender systems. In *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*. 353–362.

[64] Mingyue Zhang and Jesse Bockstedt. 2020. Complements and substitutes in online product recommendations: The differential effects on consumers' willingness to pay. *Information & Management* 57, 6 (2020), 103341.

[65] Shengyu Zhang, Ziqi Jiang, Jiangchao Yao, Fuli Feng, Kun Kuang, Zhou Zhao, Shuo Li, Hongxia Yang, Tat-seng Chua, and Fei Wu. 2023. Causal distillation for alleviating performance heterogeneity in recommender systems. *IEEE Transactions on Knowledge and Data Engineering* 36, 2 (2023), 459–474.

[66] Shuai Zhang, Lina Yao, Aixin Sun, and Yi Tay. 2019. Deep learning based recommender system: A survey and new perspectives. *ACM computing surveys (CSUR)* 52, 1 (2019), 1–38.

[67] Weifeng Zhang, Jingwen Mao, Yi Cao, and Congfu Xu. 2020. Multiplex graph neural networks for multi-behavior recommendation. In *Proceedings of the 29th ACM international conference on information & knowledge management*. 2313–2316.

[68] Yang Zhang, Fuli Feng, Xiangnan He, Tianxin Wei, Chonggang Song, Guohui Ling, and Yongdong Zhang. 2021. Causal intervention for leveraging popularity bias in recommendation. In *Proceedings of the 44th international ACM SIGIR conference on research and development in information retrieval*. 11–20.

[69] Yanqiao Zhu, Yichen Xu, Feng Yu, Qiang Liu, Shu Wu, and Liang Wang. 2020. Deep graph contrastive representation learning. *arXiv preprint arXiv:2006.04131* (2020).

[70] Kuan Zou, Aixin Sun, Xuemeng Jiang, Yitong Ji, Hao Zhang, Jing Wang, and Ruijie Guo. 2024. Hesitation and Tolerance in Recommender Systems. *arXiv preprint arXiv:2412.09950* (2024).