A Multi-Interest Evolution Story: Applying Psychology in Query-based Recommendation for Inferring Customer Intention

Yuqi Qin*
Beijing University of Posts and Telecommunications
Beijing, China
yuqiqin@bupt.edu.cn

Pengfei Wang*†
Beijing University of Posts and Telecommunications
Beijing, China
wangpengfei@bupt.edu.cn

Biyu Ma
Xidian University
Xi'an, China
mabiyujy@126.com

Zhe Zhang
Zhejiang University
Hangzhou, China
1014623855@qq.com

ABSTRACT

The query-based recommendation now is becoming a basic research topic in the e-commerce scenario. Generally, given a query that a user typed, it aims to provide a set of items that the user may be interested in. In this task, the customer intention (*i.e.,* browsing or purchase) is an important factor to configure the corresponding recommendation strategy for better shopping experiences (i.e, providing diverse items when the user prefers to browse or recommending specific items when detecting the user is willing to purchase). Though necessary, this is usually overlooked in previous works. In addition, the diversity and evolution of user interests also bring challenges to inferring user intentions correctly.

In this paper, we propose a predecessor task to infer two important customer intentions, which are purchasing and browsing respectively, and we introduce a novel **P**sychological **I**ntention **P**rediction **M**odel (***PIPM*** for short) to address this issue. Inspired by cognitive psychology, we first devise a multi-interest extraction module to adaptively extract interests from the user-item interaction sequence. After this, we design an interest evolution layer to model the evolution of the mined multiple interests. Finally, we aggregate all evolved multiple interests to infer users' intentions in his/her next visit. Extensive experiments are conducted on a large-scale Taobao industrial dataset. The results demonstrate that *PIPM* gains a significant improvement on AUC and GAUC than state-of-the-art baselines. Notably, PIPM has been deployed on the Taobao e-commerce platform and obtained over 10% improvement on PCTR.

*Beijing Key Laboratory of Intelligent Telecommunications Software and Multimedia.
†Corresponding author.

CIKM '22, October 17–21, 2022, Atlanta, GA, USA.

ACM ISBN 978-1-4503-9236-5/22/10...$15.00
https://doi.org/10.1145/3511808.3557221

CCS CONCEPTS

• **Information systems** → *Decision support systems*; *Retrieval models and ranking*.

KEYWORDS

e-commerce, recommendation system, customer intention inference, multi-interest modeling, psychology

ACM Reference Format:
Yuqi Qin, Pengfei Wang, Biyu Ma, and Zhe Zhang. 2022. A Multi-Interest Evolution Story: Applying Psychology in Query-based Recommendation for Inferring Customer Intention . In *Proceedings of the 31st ACM International Conference on Information and Knowledge Management (CIKM '22), October 17–21, 2022, Atlanta, GA, USA.* ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3511808.3557221

1 INTRODUCTION

The query-based recommendation now is becoming a crucial and fundamental service in e-commerce platforms. Different from traditional item recommendations, it aims to infer a shopping list that a user may be interested by considering both the query the user typed and the items he/she has interacted. In this specific scenario, inferring the customer intention (*i.e.,* browsing or purchase) is crucial for configuring the corresponding recommendation strategy for better shopping experiences (i.e, providing diverse items when the user prefers to browse, or recommending specific items when detecting the user is willing to purchase) [8]. Considering its effectiveness, it is necessary to infer user intentions and further utilize these intentions to guide the following downstream recommendation task for better performance.

Though ideal, it is non-trivial to perform customer intention inference due to the following two aspects: 1) **The diversity of user interests**: users' interests are random and diverse, making it difficult to infer user willingness in the next visit. As Fig. 1 shows, the user will interact with different types of items in a short period, which revealed his/her several interests in T-shirts, sleepwear, and dresses. Though the platform can also utilize categories to represent users' interests, the rough management strategy does not fit users' complex personalized cognitive classification criteria (*i.e.,* both the sleep dress and dress belong to the skirt category, but their applications and requirements behind them are different). 2) **The evolving of user interests**: User interests are evolving during the

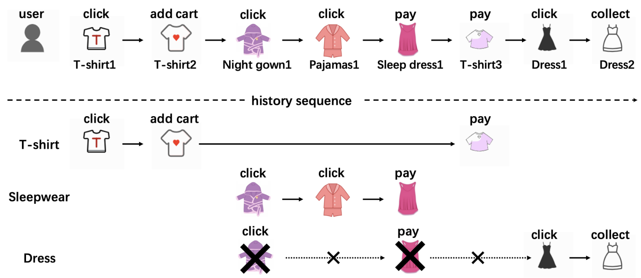

Figure 1: A motivating example to illustrate the diversity and evolution of multiple interests during the user-item interaction process.

process of user-item interaction. To explain this, let's review Fig. 1 again. We see that the user clicked on the first T-shirt and then added the second T-shirt to the cart. Intuitively, after this process, the user's interest to buy a T-shirt was constantly increased so that the user bought a T-shirt after some time. However, after this, the user's need for T-shirts is satisfied, resulting in a sharp decline in his/her T-shirts interest.

Following this line, many efforts are also devoted to modeling users' multiple interests for better recommendation [3, 14]. However, seldom of them have considered the impact of interests' evolution on users' next choices. In addition, user intentions are also ignored, resulting in a restriction on the model's performance. How to model the evolution of multiple interests to infer user intentions to guide the following downstream recommendation task, is becoming an interesting and challenging task.

Recently, the cognitive psychology is attracting many concerns [2, 5], with many psychology models are designed and applied. Among these models, the perceptual model [5] is a widely used one. Based on the participation of stored knowledge experience, the perceptual generation model can transform original stimulating factors into higher-level cognitive information according to a bottom-up and up-bottom strategy. It exactly sheds light on the challenged mentioned above, bringing the dawn of modelling the evolution of multiple interests for better intention inference.

In light of the challenges mentioned above, in this paper, we propose an end-to-end Psychological Intention Prediction Model (*PIPM* for short) to automatically model the evolution process of multiple interests for intention inference. The proposed *PIPM* mainly consists of a multi-interest extraction module (MECM) and a multi-interest evolution module (MEVM). Specifically, inspired by cognitive psychology, we first devise the MECM to adaptively extract multiple cognitive interests from his/her interaction sequence, hoping to model the process of human gradually building cognitive classification criteria during consumption as much as possible. With the mined multiple interests, we further split the whole interaction sequence into multiple disjoint sub-sequences just as the three interests mentioned in Fig. 1. For each chain, we conduct interest evolution based on its constituent items through the designed MEVM, which has ability to enhance the influence of users' different behaviors during interest evolution by a novel interest evolution layer (IEL). Ultimately, we aggregate useful information mostly relevant with the current query from the evolved multiple interests for prediction. Furthermore, a relatively simple auxiliary task is proposed to enhance the generalization and accuracy of the model and is optimized jointly in an annealing-based way to obtain a balance between the two tasks.

To summary, the contributions of this paper are summarized as follows:

- We formulate a predecessor task before query-based recommendation to infer customer intention (*i.e.,* browsing or purchase) for providing more flexible downstream recommendation strategies.
- Inspired by cognitive psychology, we design a novel multi-interest extraction model to adaptively extract multiple interests from history, hoping to eliminate the gap between users' real cognitive interests and the mined.
- For simulating the process of browsing and comparing related item in a real shopping scenario, we devise a novel interest evolution layer (IEL) to enhance the influence of users' different behaviors (*i.e., add_to_cart*) on interest evolution.
- Extensive offline experiments on the large-scale Taobao industrial dataset show that our *PIPM* outperforms state-of-the-art baselines in terms of key metrics AUC and GAUC. We also conduct online experiments on Taobao e-commerce platform and obtained over 10% improvement on PCTR.

2 RELATED WORK

In this section, we provide a brief overview of the related work from two perspectives, including user interest modeling and sequential modeling respectively.

2.1 User Interest Modeling

Collaborative filtering methods [1, 15, 19] model user preferences by finding similar users and items for recommendation, which have been proven successful in real-world. Matrix factorization [11, 22] methods map both users and items to a joint latent factor space to measure the user's preference for the item. DeepFM [7] combines low-order and high-order feature interactions to improve the power of representation. For further modeling user interests, DIN [31] uses the attention mechanism to activate related user behaviors, and DIEN [30] further models the evolution of relative interest by introducing attentional update into GRU.

However, these methods utilize a fixed-length vector to represent the user's multiple interests, limiting the model's performance due to the diversity and randomness of users' behaviors. Taking it into consideration, researchers recently focused on multi-interest modeling methods. MIND [14] clusters past behaviors based on the capsule routing mechanism to extract multiple interests. ComiRec [3] proposes a self-attention based extraction method to implement interest extraction. KA-MemNN [33] evaluates users' interest tendencies through category addressing with the representation of each category generated by an attention network according to related items and the target user, which needs to explicitly group items in the history sequence by category. HLN [6] proposes a hierarchical leaping network to capture the users' multiple preferences iteratively. SINE [23] proposes a sparse-interest module to explicitly model multiple interests for prediction by designing a large concept pool. As far as we know, most of those existing multi-interest methods ignore the evolution of interests.

2.2 Sequential Modeling

Sequential modeling is a crucial problem in the e-commerce scenario owing to its ability to capture the sequential patterns and model users' dynamic interests based on users' historical logs. FPMC [18] designed a personalized Markov chain to provide recommendations. HRM [25] extends the FPMC model and models complex interactions between users and items by a two-layer structure. Although effective, these MC-based methods only analyze the impacts of single-step sequential behaviors.

With the prosperity of deep models, for further modeling multi-step sequential behaviors, RNN-based [9, 12, 13, 28] methods are introduced to capture users' interests and sequential features, which assign the same weight to each item in a historical sequence. In order to model user preference more accurately, attention-based methods [16, 21, 24, 26, 27, 30, 31] are introduced to allow a model to focus on the more informative parts of the target with different weights. SASRec [10] introduced a novel self-attention-based sequential approach to model the entire user sequence. BERT4Rec [20] used a bidirectional self-attention network to model user sequential behaviors. Recently, there are also some pre-training methods [17, 32] to derive the intrinsic data correlations.

3 PRELIMINARY

Notations. Without losing generality, we let $\mathcal{U} = \{u_1, u_2, \cdots, u_{|\mathcal{U}|}\}$ denote all users, $\mathcal{I} = \{i_1, i_2, \cdots, i_{|\mathcal{I}|}\}$ denote all items and $\mathcal{Q} = \{q_1, q_2, \cdots, q_{|\mathcal{Q}|}\}$ denote all searched queries, where $|\mathcal{U}|$, $|\mathcal{I}|$ and $|\mathcal{Q}|$ represent the total number of unique users, items and queries respectively. For each user $u \in \mathcal{U}$, we use $\mathcal{S}^u = \{i^u_1, i^u_2, \cdots, i^u_{|\mathcal{S}|}\}$ to represent the interaction sequence of items ordered by time, where $i^u_j \in \mathcal{I}$ represents the item that u has interacted with at j-th time step and $|\mathcal{S}|$ denotes the length of the sequence. Corresponding to the historical item sequence, $\mathcal{A}^u = \{a^u_1, a^u_2, \cdots, a^u_{|\mathcal{S}|}\}$ denotes a sequence of interaction behaviors, where $a_j \in \{a_1, a_2, a_3, a_4\}$ represents the behavior u interacted on the j-th item with a_1 meaning $click$, a_2 meaning pay, a_3 meaning add_to_cart and a_4 meaning $add\_to\_collect$. In addition, we utilize q to represent the current query that u typed.

Task Definition. Based on this notation, our task of customer intention inference aims to infer whether the user u wants to purchase or browse under the current query q:

$$F(u, q, \mathcal{S}^u, \mathcal{A}^u) \rightarrow y \tag{1}$$

where y denotes the purchase probability in the next visit.

Perceptual Model. We first briefly introduce a perceptual model in cognitive psychology, and the framework of the perceptual model is shown in Fig. 2.

In cognitive psychology, perception [5] is to organize sensory information into meaningful objects, that is, to understand the meaning of current stimuli with the participation of stored knowledge experience. Taking vision as an example, the information from

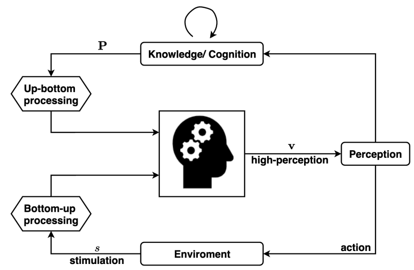

Figure 2: The perceptual model in cognitive psychology, which divides the formation process of perception into bottom-up processing, up-bottom processing and their interaction.

the sense organs provides us with individual attributes such as certain colors, boundaries, line segments, and so on. After processing by the mind, we recognize that "this is an apple" and "that is a bag". That is to say, sensation is the primary stage of perceptual cognition, and various sensations are the representations of nerve impulses generated by stimulation acting on the perceiver; perceptual cognition is higher than sensation, although it is based on sensation, it is not limited by realistic stimulation. And it also involves a variety of psychological components such as memory and thinking.

Therefore, cognitive psychology divides the formation process of perception into bottom-up processing, up-bottom processing and their interaction. which can be simplified as follows: (1) Without the participation of advanced cognitive processes such as complex thinking and reasoning, the perceiver first combines the small sensory information sensed by the sensory organs from the environment to form primary perception. This process is called for bottom-up processing $\mathbf{o} = f_{b\text{-}u}(s)$, where the stimulation s is from the environment. (2) Influenced by experiential knowledge, expectations and motivation, the perceiver carries out information selection, integration, and construction of high-level perceptual representations based on primary perception. This process is known as up-bottom processing $\mathbf{v} = f_{u\text{-}b}(\mathbf{P}, \mathbf{o})$, where $\mathbf{P}$ represents the existing experience knowledge. And it is consistent with the theory of constructive perception. (3) Furthermore, constructivism holds that knowledge is dynamic, that is, the perceiver constantly revises and improves the existing knowledge through the interaction with the experience world $\mathbf{P}_t = f_g(\mathbf{P}_{t-1}, \mathbf{v}, s)$. (4) Ultimately, different perceptual outcomes will affect the perceiver's subsequent behavior, thereby affecting the environment.

4 OUR APPROACH

In this section, we first introduce the problem formulation of the predecessor task and then describe the proposed *PIPM* model in detail. For simplicity, we drop subscript u in the notations for concise presentation.

In order to enhance the representation of each item, we introduce a lot of additional feature information. For item, we utilize item title, item price, item category, item brand, item transaction amount, item ctr, item pv (page view) and so on; for query, we utilize query title, query category, query transaction amount, query uv (unique visitor), query pv (page view) and so on.. And following [4, 8], we

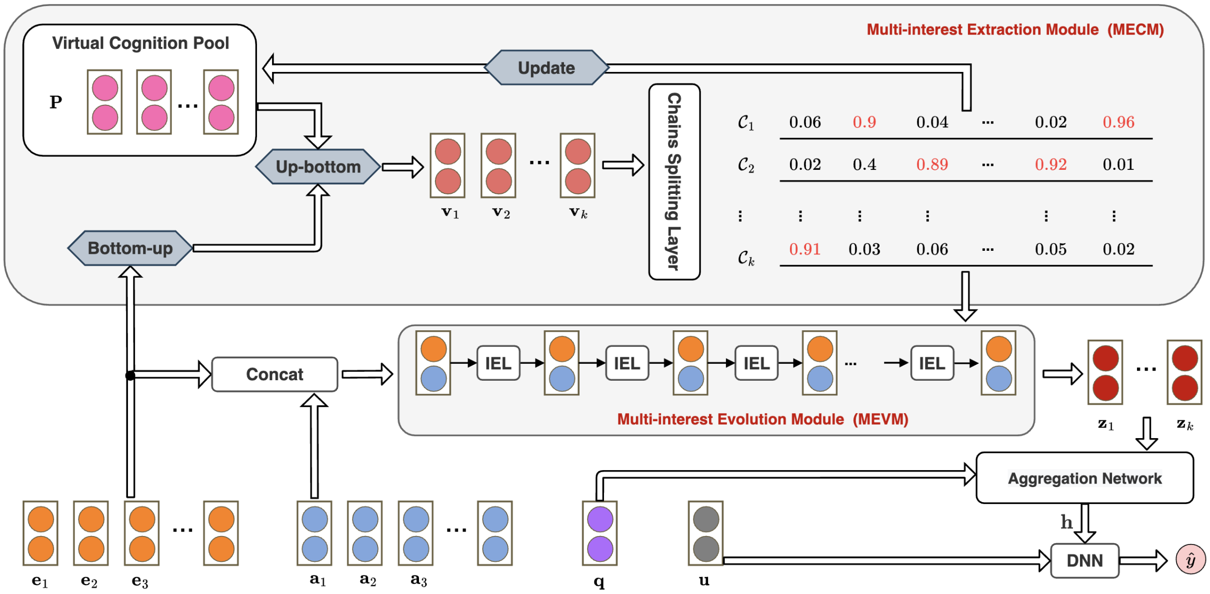

Figure 3: The network structure of the proposed *PIPM*, which is composed of a Multi-interest Extraction Module (MECM) that aims to adaptively extract K cognitive interests $[\mathbf{v}_1, \ldots, \mathbf{v}_k]$ from user-item interaction history, a Multi-interest Evolution Module (MEVM) devised for interest evolution. Finally, we aggregate all evolved multiple interests $\mathbf{z}_{i\in\{1,\cdots,k\}}$ by a simple attention-based aggregation network for customer intention inference.

utilize similar methods that the term embedding mechanism and discretization method to process these features, and then the final item representation $\mathbf{e}_j$ is derived by concatenating the representation of all these features related to the item i_j. Similar to item, we can obtain the enhanced representation of the current query that the user typed, denoted as $\mathbf{q}$.

4.1 Multi-interest Extraction Module (MECM)

Since the diversity and complexity of user interests, we use the perceptual model in cognitive psychology to frame the multi-interest extraction process, hoping to model the impact of human's complex cognitive classification criteria formed in consumption. In the perceptual model, on the basis of not involving any advanced thinking, raw environmental stimulus sensed by the sensory organs is firstly combined by the perceiver to primary perception through bottom-up processing. Similarly, we design an attention-based module named bottom-up as a primary feature extractor to extract k primary interests $\mathbf{O} = [\mathbf{o}_1, \mathbf{o}_2, \cdots, \mathbf{o}_k]$ from user history.

Bottom-Up Module. Taking the historical item representation sequence $\mathbf{S} = \{\mathbf{e}_1, \cdots, \mathbf{e}_{|\mathcal{S}|}\}$ as an environmental stimulus, we firstly calculate the relevance between each item $\mathbf{e}_i \in \mathbf{S}$ and the j-th primary interest as follows:

$$\beta_{i,j} = \frac{exp(MLP([\mathbf{e}_i, \vec{0}]) \cdot \mathbf{w}_{1,j}{}^{\top})}{\sum_{\mathbf{e}_i \in \mathbf{S}} exp(MLP([\mathbf{e}_i, \vec{0}]) \cdot \mathbf{w}_{1,j}{}^{\top})} \tag{2}$$

where $\mathbf{w}_{1,j}$ is the learnable feature extractor embedding for the j-th primary interest. $\vec{0}$ is a zero-vector and it will be replaced in our proposed auxiliary task, which is described in detail in Session 4.3. According to such a design, we can derive each primary interest $\mathbf{o}_j \in \mathbf{O}$ by weighted clustering of the interacted items, which can be seen as a combination of sensory information sensed:

$$\mathbf{o}_j = \sum_{\mathbf{e}_i \in \mathbf{S}} \beta_{i,j} \cdot \mathbf{e}_i \tag{3}$$

After that, influenced by experiential knowledge and expectation, the perceiver carries out information selection, integration, and construction of high-level perceptual information on the basis of primary perception through up-bottom processing. Similarly, we design an up-bottom module to imitate the advanced cognitive processes. Specifically, we construct a virtual cognition pool $\mathbf{P} = [\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_{|P|}]$ to represent individual prior knowledge.

Up-Bottom Module. Given the individual prior knowledge $\mathbf{P}$ and the clustered primary interests $\mathbf{O}$, we formulate the cognitive interests $\mathbf{V} = [\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k]$ as follows:

$$\begin{aligned} \gamma_{i,j} &= \frac{exp(\sigma(\mathbf{o}_i \cdot \mathbf{w}_{2,j}{}^{\top}))}{\sum_{j=1}^{|P|} exp(\sigma(\mathbf{o}_i \cdot \mathbf{w}_{2,j}{}^{\top}))} \\ \mathbf{v}_i &= \sum_{j=1}^{|P|} \gamma_{i,j} \cdot \mathbf{p}_j \end{aligned} \tag{4}$$

where σ is the tanh activation function to perform nonlinear transfer, $\gamma_{i,j}$ represents the relevance between the primary interest $\mathbf{o}_i$ and the j-th virtual cognition $\mathbf{p}_j$, $\mathbf{w}_{2,j}$ is the learnable embedding. Notably, both k and $|P|$ are hyper-parameters.

Then, in the third stage of the perceptual model, the perceiver constantly revises and improves the existing knowledge through interacting with the experience world. Therefore, In order to better learn the virtual cognition pool, an update module is designed to

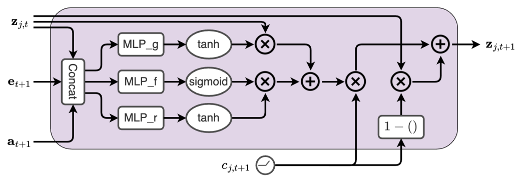

Figure 4: A novel Interest Evolution Layer (IEL), which can model the impact of users' different behaviors during the interest evolution process.

imitate the process of humans continuously revising their prior knowledge from practice.

Update Module. Generally, people always get revelation from the interaction results of current cognition and the real world to update their knowledge. To simply formalize this process, we first split the raw historical item sequence $\mathcal{S}$ into different interest chains $\{C_1, C_2, \cdots, C_k\}$ according to the multiple cognitive interests mined, aiming to represent the user's interaction result with the environment under the current cognition. For each interest chain $C_j = [c_{j,1}, c_{j,2}, \cdots, c_{j,|\mathcal{S}|}]$, the calculation process is as follows:

$$
\begin{aligned}
s(\mathbf{e}_i, \mathbf{v}_j) &= \frac{LayerNorm(\mathbf{e}_i) \cdot LayerNorm(\mathbf{v}_j)}{\sqrt{d}} \\
c_{j,i} &= \frac{exp(s(\mathbf{e}_i, \mathbf{v}_j)/\tau_0)}{\sum_{j=1}^{k} exp(s(\mathbf{e}_i, \mathbf{v}_j)/\tau_0)}
\end{aligned}
\tag{5}
$$

where $s(\mathbf{e}_i, \mathbf{v}_j)$ measures the similarity between the i-th item and the j-th cognitive interest. Notably, $\tau_0 > 0$ is a temperature parameter to tune, and we simulate the hard-coding by setting $\tau_0 \rightarrow 0$ because the hard-coding allows only the items most relevant to the cognitive interest $\mathbf{v}_j$ retained. According to such a design, our model can focus on the valuable item information related with $\mathbf{p}_l$ to update $\mathbf{p}_l$ in the following updating.

After obtained k different interest chains $\{C_1, \cdots, C_k\}$, we update the virtual cognition pool as:

$$
\begin{aligned}
\tilde{\mathbf{p}}_l &= \sum_{j=1}^{k} \gamma_{j,l} (\sum_{\mathbf{e}_i \in \mathbf{S}} c_{j,i} \cdot \mathbf{e}_i) \\
\mathbf{p}_l &= \theta_l * \mathbf{p}_l + (1 - \theta_l) \cdot \tilde{\mathbf{p}}_l
\end{aligned}
\tag{6}
$$

where $\tilde{\mathbf{p}}_l$ is the valuable information learned from the history $\mathbf{S}$ concerning the virtual cognition $\mathbf{p}_l$, and $\theta_l = MLP(\tilde{\mathbf{p}}_l, \mathbf{p}_l)$ is a gate to control the update of $\mathbf{p}_l$ with the sigmoid activation.

4.2 Multi-interest Evolution Module (MEVM)

Up to now, we have mined multiple interests people perceive and their corresponding interest chains by imitating the perceptual model in cognitive psychology. However, since users' interests are evolving, it is necessary to model the transformation of interest over time and behaviors just as Fig. 1. Therefore, we design a Multi-interest Evolution Module (MEVM) to model the evolution of each cognitive interest via a novel Interest Evolution Layer (IEL), aiming to imitate people's process of browsing and comparing related items to obtain the final evolved interests $\mathbf{Z} = [\mathbf{z}_1, \mathbf{z}_2, \cdots, \mathbf{z}_k]$. The IEL is described in Fig. 4 in detail.

For the interest chain C_j, we set the initial interest as $\mathbf{z}_{j,0} = MLP_u(\mathbf{u})$, the candidate interest state $\tilde{\mathbf{z}}_{j,t+1}$ in preparation for the subsequent state updating is calculated firstly by taking the last state $\mathbf{z}_{j,t}$, the current item $\mathbf{e}_{t+1}$ and the corresponding behavior $\mathbf{a}_{t+1} \in \mathbf{A}$:

$$
\begin{aligned}
\mathbf{g}_{t,j} &= \sigma_g(MLP_g([\mathbf{z}_{j,t}, \mathbf{a}_{t+1}, \mathbf{e}_{t+1}])) \\
\mathbf{f}_{t,j} &= \sigma_f(MLP_f([\mathbf{z}_{j,t}, \mathbf{a}_{t+1}, \mathbf{e}_{t+1}])) \\
\mathbf{r}_{t,j} &= \sigma_r(MLP_r([\mathbf{z}_{j,t}, \mathbf{a}_{t+1}, \mathbf{e}_{t+1}])) \\
\tilde{\mathbf{z}}_{j,t+1} &= \mathbf{g}_{t,j} \cdot \mathbf{z}_{j,t} + \mathbf{f}_{t,j} \cdot \mathbf{r}_{t,j}
\end{aligned}
\tag{7}
$$

where $\mathbf{g}_{t,j}$,$\mathbf{f}_{t,j}$ and $\mathbf{r}_{t,j}$ are the vectors of retaining gate, update gate and resetting gate. σ_g,σ_f and σ_r are the tanh, sigmoid and tanh nonlinear activation functions respectively. Notably, we utilize the retaining gate $\mathbf{g}_{t,k}$ with the tanh activation to control how much state information from the last step should be retained, which differs from GRU using the sigmoid activation. As we all know, the range of sigmoid and tanh functions are $(0, 1)$ and $(-1, 1)$ respectively. When we use the sigmoid activation, there is always a positive correlation in different degrees between the last interest state and the current interest state, which is reasonable in general. However, in real life, a negative correlation may exist during the process of browsing and comparing related items w.r.t a specific interest. Just as the T-shirt in Fig. 1, we can see that the user's purchase willingness concerning T-shirts increased with he/she clicking the first T-shirt and then adding the second T-shirt into the cart. However, when the user bought a T-shirt after some time, his/her demand for T-shirts was met, which may lead the user no longer to buy related items in the short term. Therefore, we argue that it is more reasonable to use the tanh instead of sigmoid in our scenario.

Based on the candidate interest state $\tilde{\mathbf{z}}_{j,t+1}$ and the interest chain C_j, we update interest state of the current time step as:

$$
\mathbf{z}_{j,t+1} = c_{j,t+1} \cdot \tilde{\mathbf{z}}_{j,t+1} + (1 - c_{j,t+1}) \cdot \mathbf{z}_{j,t} \tag{8}
$$

Here, we particularly use the hard-attention weight calculated in Eq. 5 to update the interest state. That is to say, when $c_{j,t+1} = 1$, the i_{t+1} can be seen as the item related with the cognitive interest $\mathbf{v}_j$, and $\mathbf{z}_{j,t+1} = \tilde{\mathbf{z}}_{j,t+1}$. On the contrary, when $c_{j,t+1} = 0$, we regard the item i_{t+1} as the irrelevant and set $\mathbf{z}_{j,t+1} = \mathbf{z}_{j,t}$. With the process of interest evolution under $\mathbf{v}_j$, the last interest state is set to the final interest representation of the cognitive interest, that is, $\mathbf{z}_j = \mathbf{z}_{j,|\mathcal{S}|}$.

4.3 Prediction and Learning Strategy

After this, through a simple aggregation network, we aggregate all evolved cognitive interests $\{\mathbf{z}_1, \mathbf{z}_2, \cdots, \mathbf{z}_k\}$ based on the relevance between each cognitive interest and the current query $\mathbf{q}$ for prediction:

$$
\begin{aligned}
\delta_j &= \frac{exp(MLP(\mathbf{q}, \mathbf{z}_j))}{\sum_{j=1}^{k} exp(MLP(\mathbf{q}, \mathbf{z}_j))} \\
\mathbf{h} &= \sum_{j=1}^{k} \delta_j \cdot \mathbf{z}_j
\end{aligned}
\tag{9}
$$

where $\mathbf{h}$ is the extracted information mostly relevant with the current query. Finally, through a 2-layer DNN network, we calculate the final score $\hat{y}$ as:

$$
\hat{y} = \sigma(DNN([\mathbf{u}, \mathbf{h}, \mathbf{q}])) \tag{10}
$$

where σ is the sigmoid nonlinear activation function.

We utilize negative log-likelihood function as the loss function:

$$L_1 = -\frac{1}{|\mathcal{D}|}\sum_{\mathcal{D}} ylog\sigma(\hat{y}) + (1-y)log(1-\sigma(\hat{y})) \tag{11}$$

where $\mathcal{D}$ is the training set of size $|\mathcal{D}|$. $y \in \{0,1\}$ represents whether the user wants to purchase or browse under the current query. $\hat{y}$ is the output of *PIPM*, which is the predicted probability that a user wants to make a purchase currently.

Furthermore, in order to enhance the generalization of *PIPM*, we design a relatively simple auxiliary task that first inputs the point-to-point concatenation of the historical item representation sequence S and its corresponding behavior sequence $\mathbf{A} = \{\mathbf{a}_1, \cdots, \mathbf{a}_T\}$ into the bottom-up module in MECM like $\beta_{i,j} = \frac{exp(MLP([\mathbf{e}_i,\mathbf{a}_i])\cdot \mathbf{w}_j^{1\top})}{\sum_{\mathbf{e}_i \in \mathbf{S}} exp(MLP([\mathbf{e}_i,\mathbf{a}_i])\cdot \mathbf{w}_j^{1\top})}$ to get multiple interest representations $\{\overline{\mathbf{z}}_1, \cdots, \overline{\mathbf{z}}_k\}$ by Eq. 3, and then input them directly into the aggregation network to predict a score $\tilde{y}$. The loss function of the auxiliary task is:

$$L_2 = -\frac{1}{|\mathcal{D}|}\sum_{\mathcal{D}} ylog\sigma(\tilde{y}) + (1-y)log(1-\sigma(\tilde{y})) \tag{12}$$

Then, the final loss function is as follows:

$$L = \alpha * L_1 + (1-\alpha) * L_2 \tag{13}$$

where α is a hyper-parameter that aims to obtain the trade-off between the main and the auxiliary task.

5 EXPERIMENT

In this section, we conduct extensive experiments on a large-scale industrial dataset to evaluate the efficacy of *PIPM* against the up-to-date state-of-the-art alternatives. All results are consistently significant at 0.05 level.

5.1 Experimental Setup

Dataset. We construct a real-world large-scale industrial dataset from the log data of Taobao's search scenario from May 9, 2021, to May 29, 2021. The dataset is page-level. Specifically, each page-turning request made by the user in the scenario will generate a sample data, which contains $< u, q, t_{cur}, \mathcal{S}^u, \mathcal{A}^u, label >$ representing that the user u who has the historical item sequence $\mathcal{S}^u$ and the corresponding behavior sequence $\mathcal{A}^u$ makes a page-turning request at time t_{cur} under the typed query q. And the label indicates whether the user u has purchase intention at time t_{cur}. In the experiment, if the user makes a purchase under any category related to the current query q from the current moment t_{cur} to 24:00 of the next day, we believe that the user has a purchase intention under q and set label to 1; otherwise, label is 0. Notably, we take the most recent 30 interactions from the user's history log of the past 14 days prior to the timestamp t_{cur} to construct $\mathcal{S}^u$ and $\mathcal{A}^u$.

Based on the above description of the construction method of each sample data, in our experiment, we construct the large-scale dataset by utilizing the sampled sample data for 6 days that is from May 23, 2021, to May 28, 2021, consisting of about 500 million users, 100 million items and 100 million queries. To get more robust results, we conduct experiments on training sets of different sizes (4 days and 5 days). Therefore, we have two types of datasets with different scales marked as 5_*Day* and 4_*Day*. Specifically, for the dataset 5_*Day*, we randomly sampled 10 billion samples from the data in May 23, 2021 - May 27, 2021 to construct a training set, that is, 2 billion per day. Similarly, we obtain a training set of $4 - Day$ consisting of 8 billion samples based on the data from May 24, 2021 to May 27, 2021. The validation set is composed of 2 billion samples randomly sampled on May 28, 2021. As we utilized tens of billions of data to guarantee the robustness and effectiveness of our proposed model, considering such a huge dataset, we then utilized the unified platform AOP used in Taobao.

Baselines. We compare the proposed *PIPM* against up-to-date multi-interest modeling methods and sequential modeling models.

- *MIND* [14]: MIND clusters past behaviors based on the capsule routing mechanism to extract multiple interests. Due to the difference in tasks, in the experiment, we directly input the multiple interests extracted by MIND into the aggregation network for predicting a score.
- *ComiRec-SA* [3] : It is a self-attention based multi-interest extraction method. In the experiment, we take a approach similar to MIND to calculate a score for prediction.
- *DIN*: [31]. DIN utilizes a simple attention-based activation unit to activate related user behaviors.
- *DIN*+: [31]. For the activation unit of DIN, multi-head attention with the feedforward network is used to replace it.
- *DIEN* [30]: Considering interest evolving phenomenon, DIEN designs an interest evolving layer named AUGRU to model the evolution of relative interest related to the target by introducing attentional update gate.
- *MIAN* [29]: By utilizing the fine-grained user-specific, context as well as historical interaction information, MIAN designs a global and three local interaction modules to extract the relationship among all kinds of fine-grained features.

For our proposed methods:

- PIPM: *PIPM* is a psychological intention prediction model, which utilizes Eq. 11 to optimize the model.
- PIPM $^+$: Designing a relatively simple auxiliary task with an annealing-based joint optimization method to enhance the generalization of the model, we use Eq. 13 and Eq. 14 to optimize *PIPM*.
- PIPM $^-$: For further validating the validity of the proposed annealing-based joint optimization method in balancing the main task and auxiliary task, we simply set α as 0.5 in Eq. 13.

Evaluation Metrics. In our task, we employ the widely used AUC and GAUC as evaluation metrics and regard the top-level category of the current query existing in the dataset as a grouping basis for GAUC. The large AUC and GAUC values mean better performance.

Experimental Details *PIPM* is trained with Adagrad optimizer in a warmup way that the learning rate increased from 0.001 to 0.01 in the first hundred thousand steps. The dimension d is set to 128, the initial temperature τ_0 in Eq. 5 is set to 0.001, $|P|$ is set to 1000 and α_0 in Eq. 14 is tuned in the range of [0.1,0.3,0.5,0.7]. The batch size is set to 512. For a fair comparison, in other methods, the model dimension and the batch size are also set to 128 and 512 respectively. We train *PIPM* using a distributed TensorFlow with 20 parameter servers and 800 workers.

Table 1: Performance comparison between the baselines and our model with % omitted. The best results in each row are highlighted in boldface and the best results in several baselines are highlighted in * . ▲% refers to the absolute performance gain against the best baseline, which is consistently significant at 0.05 level.

		Multi-intenion Modeling		Sequential Modeling				Our Proposed			
	Metrics	MIND	ComiRec-SA	MIAN	DIN	DIN$^+$	DIEN	*PIPM* $^-$	*PIPM*	*PIPM* $^+$	▲%
5_*Day*	AUC	78.51	78.93	78.35	78.62	78.71	78.99*	79.43	79.78	**79.85**	0.86
	GAUC	75.12	75.63	75.07	75.31	75.49	75.69*	76.27	76.63	**76.70**	1.01
4_*Day*	AUC	78.33	78.57	78.13	78.34	78.55	78.75*	79.34	79.71	**79.79**	1.01
	GAUC	74.89	75.23	74.80	75.03	75.30	75.46*	76.18	76.55	**76.63**	1.17

Table 2: Performance comparison of three different variants against *PIPM*. All numbers in the table are percent numbers with % omitted.

	Metrics	*PIPM* $^{-ec}_{v1}$	*PIPM* $^{-ec}_{v2}$	*PIPM* $^{-evo}$	*PIPM*
5_*Day*	AUC	79.67	79.70	79.33	79.78
	GAUC	79.50	76.55	76.14	76.63
4_*Day*	AUC	79.64	79.66	79.27	79.71
	GAUC	79.47	76.50	76.06	76.55

Table 3: Performance comparison of different interest evolution networks against proposed IEL. All numbers in the table are percent numbers with % omitted.

	Metrics	*PIPM* $^{-evo}$	*PIPM* ^{soft}	*PIPM* ^{aigru}	*PIPM* ^{augru}	*PIPM*
5_*Day*	AUC	79.33	79.63	79.53	79.64	79.78
	GAUC	76.14	76.49	76.37	76.48	76.63
4_*Day*	AUC	79.27	79.58	79.45	79.62	79.71
	GAUC	76.06	76.42	76.27	76.44	76.55

In particular, due to the simplicity of the auxiliary task compared with the main task, we design an annealing-based joint optimization method to obtain a balance between the main task and the auxiliary task for performance improvement.

$$\alpha = min\big(1.0, \alpha_0 + (1 - \alpha_0) * iter_{cur}/iter_0\big) \tag{14}$$

where α_0 is the initial value, $iter_0$ is the iteration number of annealing and $iter_{cur}$ is the current iteration number.

5.2 Performance Comparison

Table 1 presents a summary of the experimental results of different methods. Here, we have the following observations:

(1) For multi-interest modeling methods, ComiRec-SA performs better than MIND, which verifies the strong ability to capture user interests by the self-attention mechanism and is consistent with the result of [3]. However, Both ComiRec-SA and MIND perform worse than *PIPM* due to ignoring the evolution of user interests.

(2) For sequential modeling methods, MIAN performs worst in our scenario, the reason might be that it is capable for models to learn the fine-grained latent relationship among all kinds of features in the condition of large-scale datasets and rich side information. It is expected that DIN$^+$ performs better than DIN because the muti-head attention mechanism helps the network capture richer information. DIEN enhances the evolution of relative interest by designing an attentional update gate, outperforms DIN$^+$, DIN and MIAN, which verifies the necessity of interest evolution. Though effective, due to regardless of multi-interest modeling in DIEN, our proposed *PIPM* achieves the best performance.

(3) For our proposed methods, we observe that *PIPM* $^+$>*PIPM* >*PIPM* $^-$. The reason for *PIPM* >*PIPM* $^-$ might be that adding the loss of the auxiliary task to the main task in a simple way would decrease the performance of the main task. Due to an annealing-based joint optimization method designed to balance the two tasks, *PIPM* $^+$ performs better than *PIPM*.

(4) Interestingly, the performance gap of 5_*Day* and 4_*Day* over our proposed model is smaller than other baselines, such as from 79.79% to 79.85%. The phenomenon demonstrates that *PIPM* can capture the underlying pattern behind users' purchasing behavior in real life, reflecting that it is an effective direction to introduce psychology and cognitive psychology into user interests modeling.

5.3 Further Analysis

Ablation Study In this paper, we design a MECM module to imitate the perceptual model to extract multiple cognitive interests and devise a MEVM module to simulate people's process of browsing and comparing related items. Specifically, we design a bottom-up module, an up-bottom module and an update module respectively to model the impact of people's prior knowledge. Therefore, for fine-grained verification of the impact of each design choice on model performance, we design different variants of *PIPM* as follows. The results of *PIPM* and its variants models are shown in Table. 2.

In Eq. 6, we utilize $\tilde{\mathbf{p}}_j$ as the valuable information that the user learns from the interaction result of the current cognition and the environment, and update the virtual cognition pool based on this for better performance. Hence, we first remove the update model in MECM to investigate the effect of this updating. We named this model ignoring as *PIPM* $^{-ec}_{v2}$. We can see that comparing with *PIPM* $^{-ec}_{v2}$, *PIPM* obtains a better performance. Moreover, we further remove the up-bottom module in MECM on the basis of *PIPM* $^{-ec}_{v2}$. In this sense, the multiple interests mined are just normal category interests without any participation of advanced cognitive processes. We named this variant ignoring both up-bottom module and update module as *PIPM* $^{-ec}_{v1}$. Comparing with *PIPM* $^{-ec}_{v2}$ and *PIPM*, *PIPM* $^{-ec}_{v1}$ performs worst, which is expected. It demonstrates the correctness of introducing the perceptual model in multi-interest extraction.

Moreover, we also examine the effect of the multi-interest evolution by replacing the MEVM module with sum-pooling methods (namely *PIPM* $^{-evo}$). We can see that *PIPM* $^{-evo}$<*PIPM* $^{-ec}_{v1}$<*PIPM*, which indicates the necessity of modeling the evolution of user interests. This can also be verified by the result that *PIPM* $^{-ec}_{v1}$>ComiRec-SA. It is worthwhile to highlight that *PIPM* $^{-evo}$ also performs better than MIND and ComiRec-SA, which coincides with the previous finding that it is correct to extract multiple cognitive interests rather than normal category interests. In summary, this set of experimental comparisons suggests that each design choice in *PIPM* is rational to enhance customer intention inference.

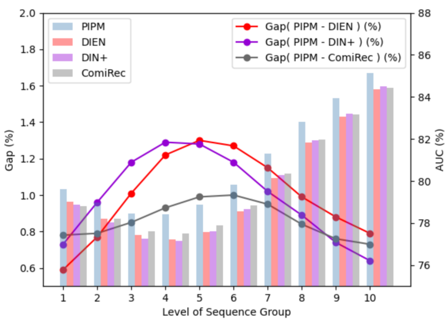

Figure 5: Performance comparison of the proposed model *PIPM*, DIN$^+$, DIEN and ComiRec-SA on different levels of sequence. The higher the level, the greater the number of unique categories contained. Bars represent the performance of different methods on AUC. Broken lines represent the AUC gap between PIPM and other baselines.

Analysis on IEL Recall that we design a novel interest evolution layer (IEL) to simulate the process of browsing and comparing related items in a real shopping scenario. Specifically, we adopt a hard-coding strategy to compute the attention weights (Eq. 5 and Eq. 8) to make the most relevant items retained for better evolution. And in order to enhance the impact of users' different behaviors (*i.e., add_to_cart*) on interest evolution, we replace the traditional sigmoid activation with the tanh activation in the retaining gate of GRUs, which is used to control how much state information from the last step should be retained (Eq. 7). Here, we first examine the effectiveness of the hard-attention weight by removing the value τ_0 in Eq. 5. The new variant of *PIPM* is named *PIPM* ^{soft}. Furthermore, we also replace the IEL with the following two attention-based GRUs that are referred to in [30], namely AIGRU and AUGRU. Specifically, AIGRU uses attention score to affect the input of interest evolution and AUGRU combines the attention mechanism and GRU seamlessly by designing an attentional update gate. Notably, all of these attention-based GRUs utilize $c_{k,t}$ as the attention score. We name these variants as *PIPM* ^{aigru} and *PIPM* ^{augru}, respectively.

The results is shown in Table 3. *PIPM* $^{-evo}$ here is the same model with *PIPM* $^{-evo}$ in Table 2. We can see that *PIPM* $^{-evo}$<*PIPM* ^{soft}<*PIPM* on both two datasets. It indicates the necessity of dropping irrelevant information to enhance interest evolution. When applying soft attention to items, *PIPM* ^{soft} can weaken the importance of irrelevant items and performs better than *PIPM* $^{-evo}$. Also, in contrast to *PIPM* ^{soft}, *PIPM* uses hard-coding to directly drop these noise items irrelevant to the specific cognitive interest and shows a better performance than *PIPM* ^{soft}. Moreover, we can see that *PIPM* ^{aigru} performs worse than *PIPM* ^{augru} and *PIPM* because zero input can change the hidden state of GRU in the condition of $c_{k,t}=0$. And *PIPM* ^{augru} gain comparable performances by using the attention score $c_{k,t}$ to control the update of hidden state in a different way. However, owing to utilizing the sigmoid activation to control the information retention of the last step, it is always a positive correlation between the last interest state and the current interest state, which is not consistent with the fact in our scenario, which is explained in detail in Session 4.2.

Analysis on Different Levels of Sequence To further investigate the performance of *PIPM* on multi-interest modeling, we split all samples into ten levels based on the number of the unique top-level categories in the historical sequence. The higher the level, the greater the number of unique categories contained. As is shown in Fig. 5, bars and broken lines represent the performance of different methods on AUC and the AUC gap between *PIPM* and other baselines over sequences of different levels, respectively.

We observe that all broken lines show a trend of rising and then falling, which is reasonable. The main idea of our proposed *PIPM* is the evolution of multi-interest. With the increase of the sequence level from level 1, the advantages of multi-interest modeling are becoming apparent, leading to a trend of rising first. When a sufficient number of unique categories exists, such as level-10, it makes no sense to model the evolution of each interest. Therefore, the trend falls subsequently. Furthermore, we observe that DIEN performs better than DIN$^+$ in level [1,5] because of the modeling of relative interest evolution. Though effective, DIEN performs worse than *PIPM* owing to the difference between evolving networks that AUGRU and IEL, a more detailed comparison on which will be made in Table 3. Moreover, ComiRec obtains better performance than DIEN and DIN$^+$ in level [3,8] because of the modeling of multiple intentions. For the phenomenon that ComiRec is worse than DIEN and DIN$^+$ in level [9,10], the reason might be that it is difficult for ComiRec to capture each interest in the condition of too many unique categories.

Analysis on Different Interest Numbers At last, we study the impact of different interest numbers k in the range of 1 to 8. The performance patterns on 5_*Day* dataset are plotted in Fig. 7. We see that *PIPM* achieves optimal performance with $k=6$ and achieves worst performance with $k=1$, which is expected. It also coincides with the previous opinion: modeling users' multiple interests is a benefit for a better recommendation.

5.4 Case Study

In order to better understand why *PIPM* is useful, we further perform analysis with a case study on 5_*Day*. Specifically, as Fig. 6 shows, given the user's historical log and the current query, we use *PIPM* and DIEN to infer whether the user has purchase intention. We can see that *PIPM* can infer all purchase intentions correctly. For example, when determining whether the user wants to buy under 'irregular short sleeve', *PIPM* can better capture the phenomenon that the user has bought an irregular short sleeve recently and will

Figure 6: Case study on customer intention inference task. Red boxes represent the *pay* behavior and blue boxes represent the *add_to_cart* behavior. √ represents that the user has purchase intention; otherwise, it indicates browsing intention.

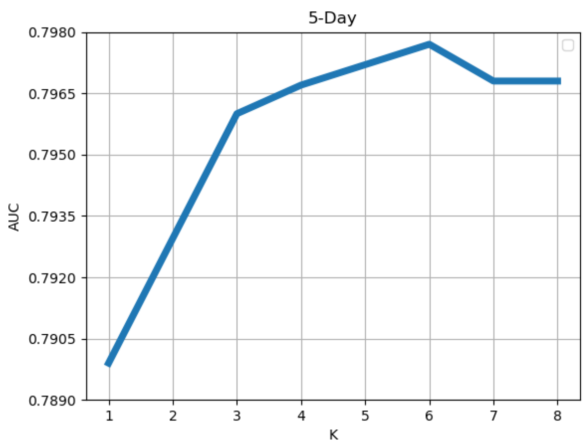

Figure 7: Performance comparison of different interest numbers k in terms of AUC on 5_*Day* dataset. The interest numbers is increased from 1 to 8.

no longer want to buy again in the short time because the IEL pays more attention to the impact of users' behaviors compared with AUGRU. And when determining whether the user wants to buy under 'night table', although the user has clicked some items related to the night table, it is difficult for DIEN to capture the user's interest in the category because the shoes are the closest items to the current query. However, *PIPM* can capture the user's interest on the night table accurately owing to the interest evolution over the interest chain composed of items related to the night table.

6 ONLINE EXPERIMENT

From 20220315 to 20220401, we conducted a bucket testing online in Taobao mobile App. Since there is no direct application scenario, for fairness, we directly replaced the Mind model deployed in the online matching stage of queries recommendation as our proposed *PIPM* to verify the ability of *PIPM* in multi-interest modeling. And we use the metrics PCTR, UCTR, PVR, and UV to evaluate the online performance.

The results are shown in Table 4. We can see that, *PIPM* outperforms Mind in all metrics, and gains the improvement of 10.13%, 8.79%, 5.59% and 1.2% on PCTR, UCTR, PVR and UV respectively. UV indicates how many unique visitors, PVR represents how many queries were exposed, UCTR indicates how many unique visitors click the recommended query and PCTR indicates how many exposed queries were clicked. All of these improvements verify that multi-interest evolution modeling can better capture users' multiple interests.

Table 4: Performance comparison of MIND and *PIPM* from Online A/B testing, which is in terms of PCTR, UCTR, PVR and UV.

Model	PCTR Gain	UCTR Gain	PVR Gain	UV Gain
Mind	0%	0%	0%	0%
PIPM	+10.13%	+8.79%	+5.59%	+1.2%

7 CONCLUSION

In this paper, we propose a predecessor new task named customer intention inference before query-based recommendation by considering the actual demand of e-commerce platforms and design a Psychological Intention Prediction Model (PIPM) to model multi-interest evolution effectively. Specially, we introduce the perceptual model in cognitive psychology for the first time to learn the cognitive interests for performance improvement. After this, a novel interest evolution layer named IEL is further designed to simulate the process of browsing and comparing related items in a real shopping scenario. Extensive experiments confirm the effectiveness of our proposed model.

To our knowledge, it is the first time considering the impact of users' complex cognition on user interests. Currently, our focus only considers the perceptual model to enhance user understanding, and there is much work to be done. In the future, we will further analyze users' purchasing behavior from the perspective of psychology, predict their possible behavior trajectory, and then promote the transformation of users' behavior path.

ACKNOWLEDGMENTS

This research work was supported by fundamental Research for the National Natural Science Foundation of China (No.61802029). We would like to thank the anonymous reviewers for their valuable comments.

REFERENCES

[1] Christos George Bampis, Cristian Rusu, Hazem M. Hajj, and Alan C. Bovik. 2017. Robust matrix factorization for collaborative filtering in recommender systems. In *51st Asilomar Conference on Signals, Systems, and Computers, ACSSC 2017, Pacific Grove, CA, USA, October 29 - November 1, 2017*, Michael B. Matthews (Ed.). IEEE, 415–419. https://doi.org/10.1109/ACSSC.2017.8335371

[2] Amin Beheshti, Shahpar Yakhchi, Salman Mousaeirad, Seyed Mohssen Ghafari, Srinivasa Reddy Goluguri, and Mohammad Amin Edrisi. 2020. Towards Cognitive Recommender Systems. *Algorithms* 13, 8 (2020). https://doi.org/10.3390/a13080176

[3] Yukuo Cen, Jianwei Zhang, Xu Zou, Chang Zhou, Hongxia Yang, and Jie Tang. 2020. Controllable Multi-Interest Framework for Recommendation. In *KDD '20: The 26th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, Virtual Event, CA, USA, August 23-27, 2020*, Rajesh Gupta, Yan Liu, Jiliang Tang, and B. Aditya Prakash (Eds.). ACM, 2942–2951. https://doi.org/10.1145/3394486.3403344

[4] Shaohua Fan, Junxiong Zhu, Xiaotian Han, Chuan Shi, Linmei Hu, Biyu Ma, and Yongliang Li. 2019. Metapath-guided Heterogeneous Graph Neural Network for Intent Recommendation. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, KDD 2019, Anchorage, AK, USA, August 4-8, 2019*. ACM, 2478–2486. https://doi.org/10.1145/3292500.3330673

[5] E.B. Goldstein. 2014. *Cognitive Psychology: Connecting Mind, Research and Everyday Experience.* Cengage Learning. https://books.google.com/books?id=4LI8AwAAQBAJ

[6] Cheng Guo, Mengfei Zhang, Jinyun Fang, Jiaqi Jin, and Mao Pan. 2020. Session-based Recommendation with Hierarchical Leaping Networks. In *Proceedings of the 43rd International ACM SIGIR conference on research and development in Information Retrieval, SIGIR 2020, Virtual Event, China, July 25-30, 2020*, Jimmy Huang, Yi Chang, Xueqi Cheng, Jaap Kamps, Vanessa Murdock, Ji-Rong Wen, and Yiqun Liu (Eds.). ACM, 1705–1708. https://doi.org/10.1145/3397271.3401217

[7] Huifeng Guo, Ruiming Tang, Yunming Ye, Zhenguo Li, and Xiuqiang He. 2017. DeepFM: A Factorization-Machine based Neural Network for CTR Prediction. In *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence, IJCAI 2017, Melbourne, Australia, August 19-25, 2017*, Carles Sierra (Ed.). ijcai.org, 1725–1731. https://doi.org/10.24963/ijcai.2017/239

[8] Long Guo, Lifeng Hua, Rongfei Jia, Binqiang Zhao, Xiaobo Wang, and Bin Cui. 2019. Buying or Browsing?: Predicting Real-time Purchasing Intent using Attention-based Deep Network with Multiple Behavior. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, KDD 2019, Anchorage, AK, USA, August 4-8, 2019*. ACM, 1984–1992. https://doi.org/10.1145/3292500.3330670

[9] Balázs Hidasi, Alexandros Karatzoglou, Linas Baltrunas, and Domonkos Tikk. 2016. Session-based Recommendations with Recurrent Neural Networks. In *4th International Conference on Learning Representations, ICLR 2016, San Juan, Puerto Rico, May 2-4, 2016, Conference Track Proceedings*.

[10] Wang-Cheng Kang and Julian J. McAuley. 2018. Self-Attentive Sequential Recommendation. In *IEEE International Conference on Data Mining, ICDM 2018, Singapore, November 17-20, 2018*. 197–206.

[11] Yehuda Koren, Robert M. Bell, and Chris Volinsky. 2009. Matrix Factorization Techniques for Recommender Systems. *Computer* 42, 8 (2009), 30–37. https://doi.org/10.1109/MC.2009.263

[12] Duc-Trong Le, Hady Wirawan Lauw, and Yuan Fang. 2017. Basket-Sensitive Personalized Item Recommendation. In *IJCAI 19-25, 2017*. 2060–2066.

[13] Duc-Trong Le, Hady W. Lauw, and Yuan Fang. 2019. Correlation-Sensitive Next-Basket Recommendation. In *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, IJCAI 2019, Macao, China, August 10-16, 2019*. 2808–2814.

[14] Chao Li, Zhiyuan Liu, Mengmeng Wu, Yuchi Xu, Huan Zhao, Pipei Huang, Guoliang Kang, Qiwei Chen, Wei Li, and Dik Lun Lee. 2019. Multi-Interest Network with Dynamic Routing for Recommendation at Tmall. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management, CIKM 2019, Beijing, China, November 3-7, 2019*. ACM, 2615–2623. https://doi.org/10.1145/3357384.3357814

[15] Hao Li, Kenli Li, Ji-yao An, and Keqin Li. 2018. MSGD: A Novel Matrix Factorization Approach for Large-Scale Collaborative Filtering Recommender Systems on GPUs. *IEEE Trans. Parallel Distributed Syst.* 29, 7 (2018), 1530–1544. https://doi.org/10.1109/TPDS.2017.2718515

[16] Jiacheng Li, Yujie Wang, and Julian J. McAuley. 2020. Time Interval Aware Self-Attention for Sequential Recommendation. In *WSDM '20: The Thirteenth ACM International Conference on Web Search and Data Mining, Houston, TX, USA, February 3-7, 2020*, James Caverlee, Xia (Ben) Hu, Mounia Lalmas, and Wei Wang (Eds.). ACM, 322–330. https://doi.org/10.1145/3336191.3371786

[17] Zhiwei Liu, Ziwei Fan, Yu Wang, and Philip S. Yu. 2021. Augmenting Sequential Recommendation with Pseudo-Prior Items via Reversely Pre-training Transformer. In *SIGIR '21: The 44th International ACM SIGIR Conference on Research and Development in Information Retrieval, Virtual Event, Canada, July 11-15, 2021*. ACM, 1608–1612. https://doi.org/10.1145/3404835.3463036

[18] Steffen Rendle, Christoph Freudenthaler, and Lars Schmidt-Thieme. 2010. Factorizing personalized Markov chains for next-basket recommendation. In *Proceedings of the 19th international conference on World wide web - WWW '10*. ACM Press, New York, New York, USA, 811–820. https://doi.org/10.1145/1772690.1772773

[19] Badrul Munir Sarwar, George Karypis, Joseph A. Konstan, and John Riedl. 2001. Item-based collaborative filtering recommendation algorithms. In *Proceedings of the Tenth International World Wide Web Conference, WWW 10, Hong Kong, China, May 1-5, 2001*, Vincent Y. Shen, Nobuo Saito, Michael R. Lyu, and Mary Ellen Zurko (Eds.). ACM, 285–295. https://doi.org/10.1145/371920.372071

[20] Fei Sun, Jun Liu, Jian Wu, Changhua Pei, Xiao Lin, Wenwu Ou, and Peng Jiang. 2019. BERT4Rec: Sequential Recommendation with Bidirectional Encoder Representations from Transformer. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management, CIKM 2019, Beijing, China, November 3-7, 2019*, Wenwu Zhu, Dacheng Tao, Xueqi Cheng, Peng Cui, Elke A. Rundensteiner, David Carmel, Qi He, and Jeffrey Xu Yu (Eds.). ACM, 1441–1450. https://doi.org/10.1145/3357384.3357895

[21] Leilei Sun, Yansong Bai, Bowen Du, Chuanren Liu, Hui Xiong, and Weifeng Lv. 2020. Dual Sequential Network for Temporal Sets Prediction. In *Proceedings of the 43rd International ACM SIGIR conference on research and development in Information Retrieval, SIGIR 2020, Virtual Event, China, July 25-30, 2020*, Jimmy Huang, Yi Chang, Xueqi Cheng, Jaap Kamps, Vanessa Murdock, Ji-Rong Wen, and Yiqun Liu (Eds.). ACM, 1439–1448. https://doi.org/10.1145/3397271.3401124

[22] Panagiotis Symeonidis and Andreas Zioupos. 2016. *Matrix and Tensor Factorization Techniques for Recommender Systems.* Springer. https://doi.org/10.1007/978-3-319-41357-0

[23] Qiaoyu Tan, Jianwei Zhang, Jiangchao Yao, Ninghao Liu, Jingren Zhou, Hongxia Yang, and Xia Hu. 2021. Sparse-Interest Network for Sequential Recommendation. In *WSDM '21, The Fourteenth ACM International Conference on Web Search and Data Mining, Virtual Event, Israel, March 8-12, 2021*, Liane Lewin-Eytan, David Carmel, Elad Yom-Tov, Eugene Agichtein, and Evgeniy Gabrilovich (Eds.). ACM, 598–606. https://doi.org/10.1145/3437963.3441811

[24] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. 2017. Attention is All you Need. In *Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017, December 4-9, 2017, Long Beach, CA, USA*, Isabelle Guyon, Ulrike von Luxburg, Samy Bengio, Hanna M. Wallach, Rob Fergus, S. V. N. Vishwanathan, and Roman Garnett (Eds.). 5998–6008. https://proceedings.neurips.cc/paper/2017/hash/3f5ee243547dee91fbd053c1c4a845aa-Abstract.html

[25] Pengfei Wang, Jiafeng Guo, Yanyan Lan, Jun Xu, Shengxian Wan, and Xueqi Cheng. 2015. Learning Hierarchical Representation Model for NextBasket Recommendation. In *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval, Santiago, Chile, August 9-13, 2015*. 403–412.

[26] Shoujin Wang, Liang Hu, Longbing Cao, Xiaoshui Huang, Defu Lian, and Wei Liu. 2018. Attention-Based Transactional Context Embedding for Next-Item Recommendation. In *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence, (AAAI-18), the 30th innovative Applications of Artificial Intelligence (IAAI-18), and the 8th AAAI Symposium on Educational Advances in Artificial Intelligence (EAAI-18), New Orleans, Louisiana, USA, February 2-7, 2018*. 2532–2539.

[27] Haochao Ying, Fuzhen Zhuang, Fuzheng Zhang, Yanchi Liu, Guandong Xu, Xing Xie, Hui Xiong, and Jian Wu. 2018. Sequential Recommender System based on Hierarchical Attention Networks. In *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence, IJCAI 2018, July 13-19, 2018, Stockholm, Sweden*, Jérôme Lang (Ed.). ijcai.org, 3926–3932. https://doi.org/10.24963/ijcai.2018/546

[28] Feng Yu, Qiang Liu, Shu Wu, Liang Wang, and Tieniu Tan. 2016. A Dynamic Recurrent Model for Next Basket Recommendation. In *Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval, SIGIR 2016, Pisa, Italy, July 17-21, 2016*. 729–732.

[29] Kai Zhang, Hao Qian, Qing Cui, Qi Liu, Longfei Li, Jun Zhou, Jianhui Ma, and Enhong Chen. 2021. Multi-Interactive Attention Network for Fine-grained Feature Learning in CTR Prediction. In *WSDM '21, The Fourteenth ACM International Conference on Web Search and Data Mining, Virtual Event, Israel, March 8-12, 2021*. ACM, 984–992. https://doi.org/10.1145/3437963.3441761

[30] Guorui Zhou, Na Mou, Ying Fan, Qi Pi, Weijie Bian, Chang Zhou, Xiaoqiang Zhu, and Kun Gai. 2019. Deep Interest Evolution Network for Click-Through Rate Prediction. In *The Thirty-Third AAAI Conference on Artificial Intelligence, AAAI 2019, The Thirty-First Innovative Applications of Artificial Intelligence Conference, IAAI 2019, The Ninth AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2019, Honolulu, Hawaii, USA, January 27 - February 1, 2019*. AAAI Press, 5941–5948. https://doi.org/10.1609/aaai.v33i01.33015941

[31] Guorui Zhou, Xiaoqiang Zhu, Chengru Song, Ying Fan, Han Zhu, Xiao Ma, Yanghui Yan, Junqi Jin, Han Li, and Kun Gai. 2018. Deep Interest Network for Click-Through Rate Prediction. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, KDD 2018, London, UK, August 19-23, 2018*, Yike Guo and Faisal Farooq (Eds.). ACM, 1059–1068. https://doi.org/10.1145/3219819.3219823

[32] Kun Zhou, Hui Wang, Wayne Xin Zhao, Yutao Zhu, Sirui Wang, Fuzheng Zhang, Zhongyuan Wang, and Ji-Rong Wen. 2020. S3-Rec: Self-Supervised Learning for Sequential Recommendation with Mutual Information Maximization. In *CIKM '20: The 29th ACM International Conference on Information and Knowledge Management, Virtual Event, Ireland, October 19-23, 2020*. ACM, 1893–1902. https://doi.org/10.1145/3340531.3411954

[33] Nengjun Zhu, Jian Cao, Yanchi Liu, Yang Yang, Haochao Ying, and Hui Xiong. 2020. Sequential Modeling of Hierarchical User Intention and Preference for Next-item Recommendation. In *WSDM '20: The Thirteenth ACM International Conference on Web Search and Data Mining, Houston, TX, USA, February 3-7, 2020*, James Caverlee, Xia (Ben) Hu, Mounia Lalmas, and Wei Wang (Eds.). ACM, 807–815. https://doi.org/10.1145/3336191.3371840